Comparison of destructive and non-destructive DNA extraction methods for the metabarcoding of arthropod bulk samples

Ameli Kirse*,[1], Sarah J. Bourlat[1], Kathrin Langen[1], Björn Zapke[1], Vera MA Zizka[1]

[1] LIB/Zoological Research Museum Alexander Koenig (ZFMK), Centre for Biodiversity Monitoring, Adenauerallee 127, 53113 Bonn, Germany

*Correspondence

Ameli Kirse

A.Kirse@leibniz-lib.de

Abstract

DNA metabarcoding is routinely used for biodiversity assessment, especially targeting highly diverse groups for which limited taxonomic expertise is available. Various protocols are currently in use, although standardization is key to its application in large-scale monitoring. DNA metabarcoding of arthropod bulk samples can be either conducted destructively from sample tissue, or non-destructively from sample fixative or lysis buffer. Non-destructive methods are highly desirable for the preservation of sample integrity but have yet to be experimentally evaluated in detail. Here, we compare diversity estimates from 14 size sorted Malaise trap samples processed consecutively with three non-destructive approaches (one using fixative ethanol and two using lysis buffers) and one destructive approach (using homogenized tissue). Extraction from commercial lysis buffer yielded comparable

species richness and high overlap in species composition to the ground tissue extracts. A significantly divergent community was detected from preservative ethanol-based DNA extraction. No consistent trend in species richness was found with increasing incubation time in lysis buffer. These results indicate that non-destructive DNA extraction from incubation in lysis buffer could provide a comparable alternative to destructive approaches with the added advantage of preserving the specimens for post-metabarcoding taxonomic work.

Introduction

Biodiversity is declining rapidly worldwide, but the full extent of these losses is yet unknown. A recent study showed that individual numbers of European breeding birds decreased by up to 19% since 1980 (Burns et al., 2021). More dramatic data have been published for terrestrial arthropods showing a decline in flying insect biomass of more than 75% over the last 27 years in Germany (Hallmann et al., 2017). Even if biomass already provides a useful indicator of insect declines, detailed understanding of changes in community composition and their potential causes and consequences need to be based on high-resolution diversity assessments. For insects, this has long been limited to a small number of charismatic groups, and hampered by the shortage in taxonomic expertise also referred to as taxonomic impediment (Wheeler, Raven, & Wilson, 2004). Given the fact that Malaise trap samples can contain up to several thousand species (Geiger et al., 2016), this lack of taxonomic expertise has prevented large scale diversity studies targeting flying insect diversity for a long time. Ongoing developments in DNA-based approaches and high-throughput sequencing techniques (HTS) have enabled diversity assessments of complex bulk samples, requiring comparably less time and effort. For example, in metabarcoding, DNA from specimen mixtures is bulk extracted, a specific marker gene is PCR amplified, sequenced using high-throughput sequencing techniques and subsequently assigned to

a species by matching against a reference database (Compson, McClenaghan, Singer, Fahner, & Hajibabaei, 2020).

However, choices made at every step of the metabarcoding protocol can directly affect the resulting species list, e.g. primer choice (Elbrecht et al., 2019; Elbrecht & Leese, 2017; Hajibabaei, Porter, Wright, & Rudar, 2019; Marquina, Andersson, & Ronquist, 2019), extraction method (Deiner, Walser, Mächler, & Altermatt, 2015; Dopheide, Xie, Buckley, Drummond, & Newcomb, 2019; Kirse, Bourlat, Langen, & Fonseca, 2021b; Majaneva, Diserud, Eagle, Hajibabaei, & Ekrem, 2018), sequencing depth (Alberdi, Aizpurua, Gilbert, & Bohmann, 2018) but also bioinformatic analysis of sequencing data (Brandon-Mong et al., 2015; Brandt et al., 2021; Kopylova et al., 2016), emphasizing that the trade-offs associated with different protocols should be considered (Bruce et al., 2021; Kirse et al., 2021b). While laboratory protocol optimization has focused on post-extraction steps such as primer choice (Elbrecht et al., 2019), less attention has been paid to the choice of DNA extraction method. DNA extraction from homogenized sample tissue, a method which is currently widely used (Elbrecht et al., 2021), results in sample destruction and thereby prevents all subsequent morphological investigation of the sample (Buchner, Haase, & Leese, 2021; Elbrecht et al., 2019; Kocher et al., 2017; Mata et al., 2020). Consequently, post-metabarcoding morphological analysis, species discovery or vouchering of the original sample are no longer possible. Unexpected diversity patterns or rare and potentially undescribed species which could help complete reference databases cannot be re-examined (Carew, Coleman, & Hoffmann, 2018). Non-destructive extraction methods are therefore highly desirable to enable the long-term preservation of samples and retrospective analysis. Various non-destructive DNA barcoding and metabarcoding protocols have been tested on single-specimens or bulk samples, these either include a chemical lysis step prior to extraction (Giebner et al., 2020; Ji et al., 2020; Kirse, Bourlat, Langen, & Fonseca, 2021a; Svenningsen et al., 2021; Vesterinen et al., 2016) or the isolation of DNA directly from the preservative ethanol (Marquina, Esparza-Salas, Roslin, & Ronquist, 2019; Zenker, Specht, & Fonseca, 2020; Zizka, Leese, Peinert, & Geiger, 2018). While Nielsen et al. (2019) usefully

demonstrated the efficiency of non-destructive DNA extraction from mock communities, studies on complex bulk samples such as Malaise traps resulted in different assessments. In particular, assessments from highly diverse terrestrial insect samples seemed incomplete or resulted in different diversity patterns when eDNA extracts from preservative ethanol were directly compared to tissue extracts, which highlights the need for further experimental testing of non-destructive DNA metabarcoding approaches (Marquina, Esparza-Salas, et al., 2019; Zenker et al., 2020). Despite these methodological inconsistencies, DNA metabarcoding is already widely used in various large-scale studies as the tool of choice for diet-, biodiversity assessment and monitoring (Bonato, Peretti, Sandionigi, & Bortolin, 2021; Cabodevilla et al., 2021; Ingala et al., 2021; Lozano Mojica & Caballero, 2021; Svenningsen et al., 2021). The processing of large sample numbers requires fast, reliable, and cost-efficient protocols. In particular, bulk insect samples from Malaise traps are often dried and homogenized in a bead mill or mixer mill. This is a time-consuming procedure that creates a fine powder which needs to be handled carefully to avoid cross contamination. Non-destructive approaches would therefore significantly speed up the metabarcoding process and reduce contamination risk by eliminating unnecessary handling steps.

In this study, we compare various extraction methods consecutively from the same 14 Malaise trap bulk samples (see Fig. 1), obtained from two habitat types. One destructive method (1. extraction from homogenized tissue) is compared to three non-destructive approaches (2. extraction from commercial ATL lysis buffer (Qiagen, Hilden Germany), 3. extraction from home-made lysis buffer (Vesterinen et al. 2016, modified from Aljanabi & Martinez 1997) and 4. extraction from sample preservative ethanol (fixative)). In addition, we investigate the effect of different incubation times (2h, 4h, 8h, 12h) on DNA yield and species detection from commercial ATL lysis buffer. Alpha diversity (number of arthropod species obtained) and overlap in species composition are used to assess the efficacy of different methods. In addition, workload, handling time and costs are considered in the analysis and discussion of trade-offs between alternative methods.

Material and Methods

Sample collection

12 Malaise trap samples (1L collection bottles filled with 96% denatured ethanol (1% MEK, type 641)) were collected on an agricultural field near Hennef, North Rhine Westphalia (Germany, test area campus Wiesengut, 50.7869°N. 7.2756°E, referred to as WG samples, WG 1-12) using automated and non-automatated Malaise traps (Wägele et al., 2022). Malaise traps were set up between the 7th of August and 10th of November 2020 and collection bottles were changed every second week. For the present study, six samples collected on 10th of September and six samples collected on 24th of September were processed. In addition, two Malaise trap samples were collected at the Bislicher Insel nature reserve, North Rhine Westphalia (Germany, 51.6529°N. 6.5231°E, referred to as BI samples, BI 1-2). Sampling was conducted on 28th of October 2020. After collection, samples were stored at -20 °C in a freezing chamber until further processing.

Extraction from sample preservative ethanol

Two weeks after collection, the total volume of approximately 800 ml preservative ethanol from samples WG1 - WG12 were filtered through 0.45 micrometer nitrocellulose filters (Nalgene™ Sterile Analytical Filter Units, 0.45 µm, 150 ml, Thermo Scientific) connected to a mini-membrane-gas-vacuum pump (VWR, Type PM20405-86.18). Insect samples were covered with new 96% denatured ethanol and stored at -20°C until further processing. Filters were torn into small pieces with fine tweezers and dried until the ethanol was completely evaporated. Subsequently, DNA was extracted from the filters with the DNeasy 96 Blood and Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. Lab bench and equipment were sterilized with ethanol, DNA-AWAY (Thermo Scientific

Molecular BioProduct) and UV radiation for 5 minutes before use. Ethanol filtering was carried out in a clean room free of DNA amplicons and DNA extraction was carried out under a PCR workstation sterilized with UV surface irradiation (Labcaire Systems Ltd.). Extraction success and DNA quality was checked on a 1 % agarose gel.

Extraction from commercial buffer (ATL, Qiagen, Hilden Germany) and column-based extraction

Samples WG 1-6 and BI 1 were divided into two size fractions by wet sieving through a 4 mm x 4 mm mesh sized stainless steel sieve with a wire diameter of 0.5 mm (Elbrecht et al., 2021). From here on, the size fractions are referred to as S (small < 4 mm) and L (large > 4 mm). Each size fraction was extracted and further processed separately. Specimens were left to air dry for five minutes before proceeding with the protocol.

In the next step, 45 ml of ATL buffer (Qiagen, Hilden Germany) mixed with 400 µg Proteinase K per ml buffer were added to both size fractions of 6 WG (WG 1-6) and one BI (BI 1) sample and placed in a shaking incubator (INCU Line ILS6, shaking incubator, VWR) at 56°C and 200 rpm. 10 ml of lysate were taken from each sample and placed into a new 50 ml Falcon tube (Corning Life Sciences) after two, four, eight and twelve hours respectively and immediately stored at -20°C. At each time point, a negative control was also taken (10 ml purified water). After the removal of lysis buffer, specimens were carefully washed with 96% denatured ethanol before drying in an incubator for three days at 50°C to prepare for tissue-based DNA extraction. Lysis buffer removed at each time point (2 h, 4 h, 8 h, 12 h for the 7 samples WG 1-6, BI 1) was processed in two centrifugation steps. As shown in several studies, a simple differential centrifugation protocol can be used to isolate mitochondria (Djafarzadeh & Jakob, 2017; Macher, Zizka, Weigand, & Leese, 2018). Cell debris were pelletized in a first centrifugation step at 12000 g for 2 minutes. The supernatant was transferred to a new 50 ml falcon tube and centrifuged at 12000 g for 60 min to achieve mitochondrial enrichment. The supernatant was

discarded and pelletized mitochondria were resuspended in 180 µl ATL buffer with 20 µl Proteinase K and further processed using the Qiagen DNeasy Blood and Tissue kit following the manufacturer's protocol. Extraction success, DNA quality and negative controls were checked on a 1 % agarose gel.

Extraction with the home-made buffer (HM) and salt precipitation

Samples (WG 7-12 and BI 2) were divided into two size fractions by wet sieving through a 4 mm x 4 mm mesh sized stainless steel sieve with a wire diameter of 0.5 mm (Elbrecht et al., 2021). From here on the size fractions are referred to as S (small < 4 mm) and L (large > 4 mm). Each size fraction was extracted and further processed separately. In the following step, 45 ml of a home-made lysis buffer (from Vesterinen 2016, modified from Aljanabi & Martinez (1997): 0.4 M NaCl, 10 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0 and 2% SDS were added to both size fractions of 6 WG and 1 BI sample. Subsequently, 400 µg Proteinase K per ml buffer were added to the sample. Samples were placed in a shaking incubator at 56°C. 10 ml of solution was taken from each sample and transferred into a new 50 ml Falcon tube (Corning Life Sciences) after two, four, eight and twelve hours respectively, and immediately stored at -20 °C. At each time point, a negative control was also taken (10 ml purified water). In the following step, a 6 M NaCl saturated salt solution was added to the lysate to a final concentration of 4 mmol and the mixture was vortexed for 30 s, to precipitate the proteins and cell membranes. Tubes were centrifuged at 4700 rpm for 30 s, the supernatant was transferred to a new Falcon tube and an equal volume of isopropanol was added to precipitate the DNA in solution. After mixing by inversion, tubes were placed at -20°C for one hour and subsequently centrifuged at 4700 rpm for 60 min to pellet the DNA. The supernatant was discarded, and the resulting DNA pellet was washed with 20 ml ice cold 70% ethanol, then centrifuged at 4700 rpm for 15 min. The remaining ethanol was discarded, and the pellet was left to dry at room temperature overnight. The next day,

166 the pellet was resuspended in 1 ml of sterile H_2O and stored at -20 °C until further processing.
167 Extraction success, DNA quality and negative controls were checked on a 1 % agarose gel.

168

169 Extraction from the sample tissue

170 Both size fractions of the 14 samples (WG 1-12, BI 1-2) were dried in a shaking incubator at 50 °C for
171 up to 3 days until ethanol was completely evaporated. The dried samples were homogenized for 3 min
172 in a Turax mixer mill (Tube Mill 100 Control) at 25,000 rpm. Approximately 25 mg of homogenized
173 tissue powder were transferred per sample to a 1.5 ml Eppendorf tube, where 180 µl ATL buffer and
174 400 mg proteinase K were added. DNA was extracted for each size fraction separately using the DNeasy
175 96 Blood and Tissue Kit (Qiagen, Hilden, Germany) including 6 negative controls following the
176 manufacturer's instructions. Extraction success and DNA quality was checked on a 1 % agarose gel.

177

178 Library preparation

179 Mitochondrial cytochrome oxidase 1 (COI) metabarcoding was carried out using a two-step PCR
180 protocol. The first PCR (termed amplicon PCR) is carried out using amplicon-specific primers with
181 Illumina adapter overhangs and the second (termed index PCR) allows the incorporation of Illumina
182 index adapters (Bourlat, Haenel, Finnman, & Leray, 2016). The amplicon PCR was set up as follows:
183 12.5 µl of PCR Multiplex Plus Mastermix (Qiagen, Hilden, Germany), 1 µl of template DNA, 0.2 µM of
184 the fwhF2 forward primer (GGDACWGGWTGAACWGTWTAYCCHCC (Vamos, Elbrecht, & Leese, 2017))
185 and 0.2 µM of the Fol_degen_rev reverse primer (TANACYTCNGGRTGNCCRAARAAYCA (Yu et al., 2012))
186 with 10.5 µl ddH_2O for a 25 µl final reaction volume. PCR was run on an Applied Biosystems 2720
187 thermocycler with the following program: initial denaturation at 95 °C for 5 min; 25 cycles of 30 s at
188 95 °C, 30 s at 50 °C and 50 s 72 °C; final extension of 5 min at 72 °C. The resulting PCR product was used

for the index PCR, set up as follows: 1 µl DNA template from PCR 1, 0.2 µM of each tagging primer (Nextera, Illumina, San Diego, USA), 12.5 µl PCR Multiplex Plus Mastermix (Qiagen, Hilden, Germany) and 10.5 µl H_2O. The PCR program was as above, running for 15 cycles instead of 25. PCRs were set up on a 96 well plate including eight extraction negative controls. PCR success was evaluated on a 1 % agarose gel before normalization to a final concentration of 25 ng per sample using a SequalPrep normalization plate (Thermo Fisher Scientific, MA, USA) following the manufacturer's instructions. 10 µl of each sample were pooled and two left sided size selections were carried out on the sample pool using magnetic beads at a ratio 0.76x to remove primer dimers (SPRIselect, Beckman Coulter). Library concentration was measured with a Quantus™ Fluorometer with the QuantiFluor® dsDNA System (Promega) and on a Fragment Analyzer (Agilent Technologies, Santa Clara, CA, USA). The pooled library was sent for sequencing on one lane of Illumina HiSeq covering 2x 250 bp at Macrogen Europe, Netherlands.

Bioinformatic Analysis

Demultiplexing was carried out by the sequencing company. Before loading demultiplexed samples into QIIME2 version 2020.11 (Bolyen et al., 2019), primer pairs were removed using Cutadapt 3.5. (Martin, 2011) with the following settings: maximum error rate (-e): 0.1, minimum overlap (-O): 20, minimum sequence length (-m): 150. Only sequences with both forward and reverse primers were retained for further analysis. In Qiime2, sequences were truncated to 175 bp for the forward reads and 170 bp for the reverse reads, respectively. Dada2 was used for merging paired-end reads, quality filtering and denoising resulting in a total of 24393 ASVs (Callahan et al. 2016). An initial blasting of ASV representative sequences against each other was conducted with blastn (version 2.9.0). The following parameter settings were chosen: 'query coverage high-scoring sequence pair percent' (-qcov_hsp_perc) was set to 80 and minimum percent identity (-perc_identity) was set to 84. The

resulting ASV match list was uploaded into R (version 3.5) and the R-package 'lulu' (version 0.1.0) was used to perform post clustering curation using standard settings (Frøslev et al., 2017). As an additional filtering step, ASVs found in the negative controls were extracted from the samples, resulting in a total of 6728 ASVs. Taxonomic assignment was carried out against the BOLD database (https://www.boldsystems.org) using Boldigger (Buchner & Leese, 2020). Only ASVs identified to species level were included for further analysis. If several ASVs were assigned to the same species indicating intraspecific diversity, ASVs and assigned reads were merged on species level (see supplementary Table 1). In addition, for the samples separated into two size fractions of < 4mm and > 4mm in the lab, the datasets were merged in silico.

UpsetR plots were made using the R packages UpSetR (version 1.4.0) (Conway, Lex, & Gehlenborg, 2017), ComplexUpset (version 1.4.1) (Krassowski, 2021) and ggplot2 (version 3.3.3) (Wickham, 2016) for visualization of the number of shared arthropod species between extraction methods as well as between incubation times. Resulting plots were further modified using Microsoft PowerPoint (Microsoft Corporation). To analyze dissimilarities between assessed communities depending on extraction methods, Permutational Multivariate Analysis of Variance (PERMANOVA) using Jaccard distance matrices for incidence data of detected arthropod species (blastID ≥ 98%) were performed using dplyr (version 0.8.3) (Wickham, Francois, Henry, & Müller, 2015) betapart (version 1.5.1) (Baselga & Orme, 2012) and vegan (version 2.5-7) (Dixon, 2003).

Boxplots showing differences in mean number of assessed species depending on extraction method and insect order were prepared using the R package ggplot2 (version 3.3.3.) and the package rstatix (version 0.7.0) (Kassambara, 2021). A t-test was conducted to check for significant differences in number of assessed species between extraction methods.

Differences in assessed species communities depending on incubation time were visualized with a PCOA plot using the R package vegan (version 2.5-7). To statistically analyze dissimilarities between

assessed communities depending on incubation time Permutational Multivariate Analysis of Variance (PERMANOVA) using Jaccard distance matrices for incidence data of detected arthropod species (blastID ≥ 98%) were performed using dplyr (version 0.8.3) betapart (version 1.5.1) and vegan (version 2.5-7). Subsequently, Kruskal-Wallis and t-tests were conducted to test for significant differences in number of assessed species depending on incubation time using the R package Rstatix (version 0.7.0).

Samples were rarefied to even sequence depth (same number of sequences per Malaise trap sample with each extraction method) using the phyloseq command rarefy_even_depth (McMurdie & Holmes, 2013)

Results

On average, 178,000 reads were recovered per sample. We observed that sequencing depth was slightly higher for the tissue and HM extraction with 207,000 and 217,000 reads respectively in comparison to the extraction from ethanol (153,000 reads) and ATL (134,000 reads). While only a single extraction per sample was conducted from the ethanol, for the extraction from tissue and lysis buffer, samples were size sorted into two fractions that were processed and sequenced separately. The reads from both size fractions were then merged in silico, resulting in twice the sequencing depth for the tissue, HM and ATL samples in comparison to the ethanol samples. Subsequent rarefaction to the same sequencing depth resulted in no change in the observed results presented here (see supplementary Table S2, S3, S4, S5).

Extraction from HM lysis buffer was more successful in terms of DNA yield than extraction from commercial ATL buffer (Supplementary Table S1). However, PCR amplification of HM buffer samples was less successful. Even if a band was visible on the DNA extraction gel, 14 of the 56 samples could not be amplified by PCR (Supplementary Table S1). Analysis of 260/280 nm ratio did not show significant differences in purity of DNA extracts between the two lysis buffers (Supplementary Table

S1). Because fewer PCR amplifications were obtained with the HM buffer, our comparison and statistical analysis focus on the latter three DNA extraction approaches (ATL, ethanol and tissue).

Species number, composition and overlap between different extraction methods

The number of species detected within the highly diverse insect orders Coleoptera, Hemiptera, Hymenoptera, and Lepidoptera did not vary significantly between extraction methods. Dipterans and other arthropods not included within the above-mentioned insect orders (Arachnida, Chilopoda, Collembola and Malacostraca as well as the insect orders Trichoptera, Neuroptera, Ephemeroptera, Orthoptera, Raphidioptera, Psocodea, Dermaptera, Thysanoptera, Mecoptera, Megaloptera, Blattodea and Zygentoma) showed a significantly higher diversity when samples were extracted using ATL buffer in comparison to extraction directly from the preservative ethanol. However, no significant differences in numbers of assessed species were detected between the ATL and tissue-based extraction for any of the target groups (Fig.2).

From 811 arthropod species identified in total, 224 (27.6%) were shared between all three extraction methods (Fig. 3). An additional 149 (18.4%), 99 (12.2%) and 61 (7.5%) species were exclusively found in extraction from ethanol, tissue and ATL respectively. The number of species shared between tissue and ATL buffer-based extraction was with 176 species (21.7%) higher than the number of shared species between the ethanol and ATL based extraction (88 species / 10.9%) and the number of shared species between extraction from tissue and ethanol (14 species / 1.7%).

Coleoptera

The total number of coleopterans detected with ATL buffer extraction was with 66 species higher than detection with the other two extraction methods (ethanol: 61 species; tissue: 52 species). The number of unique species was highest in the ethanol (21 species), followed by the tissue (16 species) and the

ATL buffer-based extraction (9 species) (Fig. 4A). In addition to 17 species detected with all three methods, 22 species were overlapping between ATL and ethanol-based extraction.

Diptera

Overall, the highest number of dipterans was detected when DNA was extracted from the ATL buffer (278 species) followed by extraction from the tissue (225 species) and the lowest species number was retrieved through extraction from the ethanol (209 species) (Fig. 4B). In total 139 dipteran species were detected with all three extraction methods (Fig. 4B). An additional 72 species were overlapping between the lysis buffer and tissue-based extraction. A high number of species were exclusively found in only one of the extraction methods, accounting for 31 species in the ethanol extracts, 34 species in the ATL lysis buffer extracts and 8 species when DNA was extracted from the homogenized tissue.

Hemiptera

The number of hemipterans identified when DNA was extracted from the ATL buffer (33 species) and the tissue-based approach (31 species) was higher in comparison to detection from the ethanol-based approach (21 species) (Fig. 4C). Additionally, a high number of species was shared between extraction from ATL and tissue (19). However, 13 species were exclusively detected when DNA was extracted from the ethanol, while 9 and 6 species were only found with the ATL and tissue methods respectively.

Hymenoptera

The highest number of hymenopteran species was found when DNA was extracted from the ethanol (98 species) while with the ATL and tissue-based extraction 94 and 84 species were detected respectively (Fig. 4D). From the 98 species detected with the ethanol extraction, 57 were not found with the two other extraction methods. However, 24 and 22 species were uniquely found with the tissue and ATL-based extraction respectively. An overlap of 35 species was found between extraction from tissue and ATL buffer while 19 species were detected with all three extraction methods (Fig. 4D).

Overlap between the ethanol and the other two approaches was comparatively low (ATL-ethanol: 18 species; ethanol-tissue: 4 species) (Fig. 4D).

Lepidoptera

In total, the highest number of species was found when DNA was extracted from the ATL buffer (52 species). 43 and 38 species were detected with the ethanol and tissue-based extraction method respectively. 22 lepidopterans were found with all three extraction methods (Fig. 4E). Additionally, 15 species intersected with the ATL- and tissue-based extraction while 8 species were exclusively found with the extraction from ATL buffer. 14 species were exclusively detected with the ethanol-based extraction, while seven additional species were shared between ethanol and ATL buffer.

Others

The category 'Others' includes species assigned to the arthropod classes Arachnida, Chilopoda, Collembola and Malacostraca as well as the insect orders Trichoptera, Neuroptera, Ephemeroptera, Orthoptera, Raphidioptera, Psocodea, Dermaptera, Thysanoptera, Mecoptera, Megaloptera, Blattodea and Zygentoma, which are infrequently collected with malaise trapping devices and constitute low overall species numbers in the traps. The most species were picked up with the DNA extraction from ATL buffer (64 species), while 47 and 43 species were detected with the tissue- and ethanol-based approach respectively. 24 species were detected with all three extraction methods and an overlap of further 17 taxa could be detected between ATL and tissue. 17, 13 and 6 species were uniquely found with ATL, ethanol and tissue-based extraction (Fig. 4F).

Incubation Time

No significant effect of incubation time on total number of identified species was observed (Kruskal-Wallis $p = 0.78$) (Fig. 5). While for three samples a slightly positive relationship of species detection

was found with increasing incubation time (WG 1 – WG 3), species detection varied with incubation time for the other samples (WG 4 – WG 6, BI 1) (Fig. 6, see Supplementary Table 1 for a list of sample names and treatments). As mentioned earlier, out of the seven samples treated with the home-made (HM) buffer DNA amplification was possible only for three. All three samples showed only slight variations in number of detected species with incubation time.

As expected, arthropod community composition did not show major differences depending on incubation time (Fig. 5, Fig. 6).

On order level, no consistent trend was found in species richness fluctuation with incubation time for four out of five investigated orders (Fig. 6, Fig. 7). The numbers of detected species of the orders Diptera, Hemiptera and Lepidoptera varied only slightly across incubations times (Fig. 7 A-C; E). Although all orders had in common that a high was reached after 12 hours of incubation, this observation was not significant. For the coleopterans, it was found that after 4 hours of incubation a significantly lower number of species was detected. The number of detected hymenopteran species was positively correlated with incubation time (Fig. 7D).

Discussion

In comparing several DNA extraction approaches sequentially on the same bulk samples, we demonstrate that non-destructive extraction from lysis buffer produces comparable results in terms of species number and composition than destructive extraction from sample tissue (Fig. 1, 2). Here it needs to be considered that extraction from ATL lysis buffer included four different time points (2h, 4h, 8h and 12h) resulting in four different extractions, compared to only one extraction in the tissue-based approach resulting in significant differences in sequencing depth. However, incubation time showed no impact on total species number nor on number of shared species with the tissue-based extraction, indicating that one replicate is sufficient to produce comparable results to the destructive

353 approach (Fig. 6). Additionally, observed patterns did not change after rarefying datasets to the same
354 read number.

355

356 Sample integrity after incubation in lysis buffer

357 The integrity of samples after the metabarcoding process is crucial for the subsequent morphological
358 identification of specimens for methodological refinements, for the completion of reference databases
359 and the long-term storage of samples (Carew et al., 2018; Martoni, Valenzuela, & Blacket, 2019). While
360 sample integrity is fully given when using ethanol as the source for DNA extraction, incubation in lysis
361 buffer directly influences the quality of preserved specimens. Lysed samples were checked by highly
362 skilled taxonomists specialized in the insect orders Hymenoptera, Lepidoptera and Diptera. While after
363 two and four hours of incubation no decrease in specimen integrity was observed, internal structures
364 in the abdomen of very small and soft specimens (e.g. Nematocera) started to dissolve after 8 hours
365 of incubation. However, the majority of specimens were still identifiable on species level after 12 hours
366 of incubation (Supplementary Fig. 1).

367

368 Trade-offs of different methods and potential for large-scale biodiversity assessments

369 In addition to preserved sample integrity, the low workload, low processing time and low risk of cross-
370 contamination associated with non-destructive approaches presents a high potential for large-scale
371 biodiversity assessments or biodiversity monitoring (Nielsen, Gilbert, Pape, & Bohmann, 2019).
372 However, choice of lysis buffer and extraction method should be carefully considered. Here, we
373 observed that the two lysis buffers and extraction methods compared (ATL and column based
374 extraction versus HM and salt based extraction) performed qualitatively very differently. While the
375 DNA extraction from commercial ATL buffer solution and blood & tissue DNA extraction kit (Qiagen,

Hilden Germany) was successful, processing of samples treated with a home-made lysis buffer and a salt precipitation protocol (HM) was more challenging. Here, several extracts required special treatment e.g. dilution and/or an increase in number of PCR cycles to retrieve sufficiently high amplicon concentration for library preparation and sequencing. However, for several HM samples amplification was not successful and only three out of seven samples could be included with all four incubation times, pointing to an issue with inhibition or DNA purity. Similar issues have been observed with this protocol in previous experiments and could be resolved with further DNA purification of the samples (DNeasy PowerClean Pro Cleanup Kit (Qiagen, Hilden Germany), increasing handling time and cost.

Overall costs for the extraction of one sample with the tested lysis buffers were as follows: while a single extraction with ATL buffer (50 ml per sample) costs around 57 $, 63 $ had to be budgeted for a single extraction with the home-made buffer. The large volume of Proteinase K added (400 µg Proteinase K per ml buffer) to counteract the activity of DNase during lysis is the costliest reagent in the protocol (21.47 $ per ml). While complete omission of Proteinase K can lead to a decrease in DNA yield due to DNA degradation (Tsuji, Yamanaka, & Minamoto, 2017), a reduced volume has already been successfully tested (Kirse et al., 2021a). Additionally, the extraction from ATL buffer has the added advantage of using a ready-made solution, giving more reliable results.

Discussion of observed differences between methods

Aside extraction from ATL lysis buffer, the extraction directly from the ethanol revealed similar species numbers compared with the tissue-based approach and required a shorter processing time. However, the species composition differed greatly compared to ATL buffer and tissue-based assessments (Fig. 3). This is congruent with a previous study investigating Malaise traps (Marquina, Esparza-Salas, et al., 2019) and can be explained by the different sclerotization levels of species present in the samples, but

400 also by the different DNA sources the approaches are based on (bulk DNA versus eDNA). Body structure
401 but also size are the main parameters defining the amount of DNA released into the fixative ethanol
402 leading to the underrepresentation of hard and highly sclerotized taxa in final sequencing results
403 (Erdozain et al., 2019; Marquina, Esparza-Salas, et al., 2019; Zizka et al., 2018). Conversely, DNA
404 extracted from bulk tissue constitutes a highly concentrated source of DNA, termed community DNA
405 (cDNA) (Deiner et al., 2017). The grinding of specimens of different biomass means that specimens of
406 low biomass and DNA traces (e.g stomach contents) are unlikely to be PCR amplified. On the other
407 hand, DNA extraction from sample fixative ethanol and lysis buffer targets the DNA released from the
408 specimens into the processing liquid. This method is therefore considered as intermediate between
409 community DNA (cDNA) and environmental DNA (eDNA) metabarcoding. This method facilitates the
410 detection of DNA traces from e.g. stomach contents or endoparasites especially in ethanol-based
411 extraction due to the frequently observed regurgitation of arthropods when fixed live in ethanol
412 (Linard et al., 2018; Tiede, Scherber, Mutschler, McMahon, & Gratton, 2017; Zizka et al., 2018). Less
413 abundant species and species with lower biomass will also be preferentially detected with non-
414 destructive methods, as biomass biases are less prominent than with a tissue-based approach. As a
415 result, it is likely that metabarcoding from preservative ethanol or lysis buffer not only reflects the
416 insect community caught but also gives a glimpse of associated species e.g. stomach contents or
417 parasites. Our results support this assumption as several taxa which are rarely detected with Malaise
418 trapping were mainly found with non-destructive methods (e.g. Fungi, Chordata, Nematoda).

419 Remarkably, the highest number of hymenopterans was detected through DNA extraction from the
420 ethanol. Various representatives of this highly diverse insect order perform important ecological
421 functions in ecosystems and are of special interest in ecological studies (Anderson et al., 2011; Gallai,
422 Salles, Settele, & Vaissière, 2009). If a very detailed assessment of diversity pattern is needed, the
423 combination of different DNA extraction methods can further increase resolution.

Effect of sub-sampling, replication, and incubation time

However, the different number of extraction replicates might further enhance the observed differences in the species lists between the ethanol extraction and the ATL and tissue samples respectively. Studies have already shown that replication can increase species detection rates in metabarcoding studies (Beentjes, Speksnijder, Schilthuizen, Hoogeveen, & van der Hoorn, 2019; Buchner et al., 2021; Ficetola et al., 2015; Zizka et al., 2022). Here we show that although combining different time replicates increased species detection rates with the ATL buffer-based extraction, a single extraction yields comparable species numbers to the tissue-based approach. Except for the Hymenoptera, incubation time during lysis had no significant influence on total number of detected species nor on assessed community composition. While for Lepidoptera, Diptera und Hemiptera results reveal more than 50% of all species detected at all incubation times, this was only observed for approximately 30% of coleopterans and hymenopterans. Hymenopterans are one of the most difficult groups to target with metabarcoding which could be explained by less conserved primer binding sites inducing primer mismatches when universal primers are used (Brandon-Mong et al., 2015). Additionally, several representatives of low biomass are often hard to detect contributing only low amounts to the DNA mixture for sequencing.

Here we found that after 4 hours of incubation DNA concentration significantly increases until a high was reached at 12 hours incubation time. It is likely that relative proportions of DNA originating from the larger specimens increases with incubation time causing a change in lysis DNA soup composition, resulting in shifts in DNA proportions, and probably resulting in overrepresentation of the large specimens. However, our approach of size sorting the samples into two size fractions (>4mm, <4mm) prior to incubation in lysis buffer seems to have remedied some of the biomass biases and contributed to comparable species lists to the ground tissue extracts.

448 It cannot be excluded that the observed random fluctuations in number of assessed species but also
449 the species composition during incubation time are artefacts of the subsampling strategy used. At each
450 time point, 10 ml of lysis buffer were removed from the sample. Although samples were thoroughly
451 mixed throughout the experiment, it cannot be guaranteed that DNA from each specimen was
452 contained in each subsample. As this effect would be enhanced with lower subsample volumes
453 (Batovska, Piper, Valenzuela, Cunningham, & Blacket, 2021; Giebner et al., 2020; Martins et al., 2019),
454 we would recommend using the total volume of lysate for the DNA extraction.

455 Overall, further studies are needed to evaluate the effect of using a higher number of subsamples of
456 ground tissue extract and the observed overlap with lysis buffer extracts. In any case, a higher number
457 of extraction replicates will increase costs. While with the ATL buffer a single extraction (50 ml
458 extraction buffer, one size fraction) costs around 57 $, extraction from 25 mg tissue (a complete ground
459 single Malaise trap bulk sample often weighs more than 10 g) will cost between 8 and 18 $ depending
460 on homogenization approach used.

461

462 **Conclusion**

463 We demonstrated here that using non-destructive lysis buffer based DNA extraction methods on size-
464 sorted arthropod malaise trap samples delivers comparable results in terms of species richness and
465 composition as destructive, tissue-based approaches. In agreement with previous studies, DNA
466 extraction from ethanol results in high numbers of species but yields different community
467 compositions than lysis buffer and tissue-based extractions. Considering sample integrity, time-
468 efficiency and a low contamination risk, non-destructive extraction from lysis buffer is a promising
469 alternative to sample tissue homogenization, especially in large scale projects where sample handling
470 needs to be kept to a minimum. Using a combination of extraction methods can strongly increase
471 assessed species richness. This is consistent with previous studies which have already shown that

combining several methods for the extraction of DNA from the same sample type can increase species detection rates.

Acknowledgements

This project was carried out in the context of AMMOD (https://ammod.de) and funded by the German Federal Ministry of Education and Research (BMBF). We thank the Krefeld Entomological Society for help with the Malaise traps and sampling permits at the Bislicher Insel nature reserve. We thank Marianne Espeland, Björn Rulik, Ralph Peters and Ximo Mengual for examination of the specimens post lysis.

References

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

Aljanabi, S. M., & Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, *25*(22), 4692–4693.

Anderson, A., McCormack, S., Helden, A., Sheridan, H., Kinsella, A., & Purvis, G. (2011). The potential of parasitoid Hymenoptera as bioindicators of arthropod diversity in agricultural grasslands. *Journal of Applied Ecology*, *48*(2), 382–390. doi: 10.1111/j.1365-2664.2010.01937.x

Baselga, A., & Orme, C. D. L. (2012). betapart: An R package for the study of beta diversity. *Methods in Ecology and Evolution*, *3*(5), 808–812.

Batovska, J., Piper, A. M., Valenzuela, I., Cunningham, J. P., & Blacket, M. J. (2021). Developing a non-destructive metabarcoding protocol for detection of pest insects in bulk trap catches. *Scientific Reports*, *11*(1), 1–14.

Beentjes, K. K., Speksnijder, A. G., Schilthuizen, M., Hoogeveen, M., & van der Hoorn, B. B. (2019). The effects of spatial and temporal replicate sampling on eDNA metabarcoding. *PeerJ*, *7*, e7335.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., … Asnicar, F. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857.

Bonato, L., Peretti, E., Sandionigi, A., & Bortolin, F. (2021). The diet of major predators of forest soils: A first analysis on syntopic species of Chilopoda through DNA metabarcoding. *Soil Biology and Biochemistry*, *158*, 108264.

Bourlat, S. J., Haenel, Q., Finnman, J., & Leray, M. (2016). Preparation of amplicon libraries for metabarcoding of marine eukaryotes using Illumina MiSeq: The dual-PCR method. In *Marine Genomics* (pp. 197–207). Springer.

Brandon-Mong, G.-J., Gan, H.-M., Sing, K.-W., Lee, P.-S., Lim, P.-E., & Wilson, J.-J. (2015). DNA metabarcoding of insects and allies: An evaluation of primers and pipelines. *Bulletin of Entomological Research*, *105*(6), 717–727.

Brandt, M. I., Trouche, B., Quintric, L., Günther, B., Wincker, P., Poulain, J., & Arnaud-Haond, S. (2021). Bioinformatic pipelines combining denoising and clustering tools allow for more comprehensive prokaryotic and eukaryotic metabarcoding. *Molecular Ecology Resources*.

Bruce, K., Blackman, R. C., Bourlat, S. J., Hellström, M., Bakker, J., Bista, I., … Clark, K. (2021). A practical guide to DNA-based methods for biodiversity assessment. *Advanced Books*. doi: https://doi.org/10.3897/ab.e68634

Buchner, D., Haase, P., & Leese, F. (2021). Wet grinding of invertebrate bulk samples–a scalable and cost-efficient protocol for metabarcoding and metagenomics. *Metabarcoding and Metagenomics*, *5*, e67533.

519 Buchner, D., & Leese, F. (2020). BOLDigger–a Python package to identify and organise sequences
520 with the Barcode of Life Data systems. *Metabarcoding and Metagenomics*, *4*, e53535.
521 Burns, F., Eaton, M. A., Burfield, I. J., Klvaňová, A., Šilarová, E., Staneva, A., & Gregory, R. D. (2021).
522 Abundance decline in the avifauna of the European Union reveals cross-continental
523 similarities in biodiversity change. *Ecology and Evolution*, *n/a*(n/a). doi: 10.1002/ece3.8282
524 Cabodevilla, X., Mougeot, F., Bota, G., Mañosa, S., Cuscó, F., Martínez-García, J., … Madeira, M. J.
525 (2021). Metabarcoding insights into the diet and trophic diversity of six declining farmland
526 birds. *Scientific Reports*, *11*(1), 1–13.
527 Carew, M. E., Coleman, R. A., & Hoffmann, A. A. (2018). Can non-destructive DNA extraction of bulk
528 invertebrate samples be used for metabarcoding? *PeerJ*, *6*, e4980. doi: 10.7717/peerj.4980
529 Compson, Z. G., McClenaghan, B., Singer, G. A., Fahner, N. A., & Hajibabaei, M. (2020).
530 Metabarcoding from microbes to mammals: Comprehensive bioassessment on a global scale.
531 *Frontiers in Ecology and Evolution*, *8*, 379.
532 Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: an R package for the visualization of
533 intersecting sets and their properties. *Bioinformatics*, *33*(18), 2938–2940.
534 Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez,
535 L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant
536 communities. *Molecular Ecology*, *26*(21), 5872–5895. doi: 10.1111/mec.14350
537 Deiner, K., Walser, J.-C., Mächler, E., & Altermatt, F. (2015). Choice of capture and extraction
538 methods affect detection of freshwater biodiversity from environmental DNA. *Biological*
539 *Conservation*, *183*, 53–63.
540 Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation*
541 *Science*, *14*(6), 927–930.

Djafarzadeh, S., & Jakob, S. M. (2017). Isolation of intact mitochondria from skeletal muscle by differential centrifugation for high-resolution respirometry measurements. *Journal of Visualized Experiments: JoVE*, (121).

Dopheide, A., Xie, D., Buckley, T. R., Drummond, A. J., & Newcomb, R. D. (2019). Impacts of DNA extraction and PCR on DNA metabarcoding estimates of soil biodiversity. *Methods in Ecology and Evolution*, *10*(1), 120–133.

Elbrecht, V., Bourlat, S. J., Hörren, T., Lindner, A., Mordente, A., Noll, N. W., … Zizka, V. M. A. (2021). Pooling size sorted Malaise trap fractions to maximize taxon recovery with metabarcoding. *PeerJ*, *9*, e12177. doi: 10.7717/peerj.12177

Elbrecht, V., Braukmann, T. W., Ivanova, N. V., Prosser, S. W., Hajibabaei, M., Wright, M., … Steinke, D. (2019). Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ Preprints*, *7*, e27801v1.

Elbrecht, V., & Leese, F. (2017). Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Science*, *5*, 11.

Erdozain, M., Thompson, D. G., Porter, T. M., Kidd, K. A., Kreutzweiser, D. P., Sibley, P. K., … Hajibabaei, M. (2019). Metabarcoding of storage ethanol vs. Conventional morphometric identification in relation to the use of stream macroinvertebrates as ecological indicators in forest management. *Ecological Indicators*, *101*, 173–184. doi: 10.1016/j.ecolind.2019.01.014

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., … Pompanon, F. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*, *15*(3), 543–556.

Frøslev, T. G., Kjøller, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., & Hansen, A. J. (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

566 Gallai, N., Salles, J.-M., Settele, J., & Vaissière, B. E. (2009). Economic valuation of the vulnerability of
567 world agriculture confronted with pollinator decline. *Ecological Economics*, *68*(3), 810–821.
568 doi: 10.1016/j.ecolecon.2008.06.014
569 Geiger, M. F., Moriniere, J., Hausmann, A., Haszprunar, G., Wägele, W., Hebert, P. D., & Rulik, B.
570 (2016). Testing the Global Malaise Trap Program–How well does the current barcode
571 reference library identify flying insects in Germany? *Biodiversity Data Journal*, *4*(4), e10671.
572 Giebner, H., Langen, K., Bourlat, S. J., Kukowka, S., Mayer, C., Astrin, J. J., … Fonseca, V. G. (2020).
573 Comparing diversity levels in environmental samples: DNA sequence capture and
574 metabarcoding approaches using 18S and COI genes. *Molecular Ecology Resources*, *20*(5),
575 1333–1345. doi: 10.1111/1755-0998.13201
576 Hajibabaei, M., Porter, T. M., Wright, M., & Rudar, J. (2019). COI metabarcoding primer choice affects
577 richness and recovery of indicator taxa in freshwater systems. *PloS One*, *14*(9), e0220953.
578 Hallmann, C. A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., … Hörren, T. (2017). More
579 than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PloS*
580 *One*, *12*(10), e0185809.
581 Ingala, M. R., Simmons, N. B., Wultsch, C., Krampis, K., Provost, K. L., & Perkins, S. L. (2021).
582 Molecular diet analysis of neotropical bats based on fecal DNA metabarcoding. *Ecology and*
583 *Evolution*, *11*(12), 7474–7491.
584 Ji, Y., Huotari, T., Roslin, T., Schmidt, N. M., Wang, J., Yu, D. W., & Ovaskainen, O. (2020). SPIKEPIPE: A
585 metagenomic pipeline for the accurate quantification of eukaryotic species occurrences and
586 intraspecific abundance change using DNA barcodes or mitogenomes. *Molecular Ecology*
587 *Resources*, *20*(1), 256–267. doi: 10.1111/1755-0998.13057
588 Kassambara, A. (2021). Rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package version
589 0.7.0. (Version 0.7.0). Retrieved from https://CRAN.R-project.org/package=rstatix

590 Kirse, A., Bourlat, S. J., Langen, K., & Fonseca, V. G. (2021a). Metabarcoding Malaise traps and soil
591 eDNA reveals seasonal and local arthropod diversity shifts. *Scientific Reports*, *11*(1), 10498.
592 doi: 10.1038/s41598-021-89950-6
593 Kirse, A., Bourlat, S. J., Langen, K., & Fonseca, V. G. (2021b). Unearthing the Potential of Soil eDNA
594 Metabarcoding—Towards Best Practice Advice for Invertebrate Biodiversity Assessment.
595 *Frontiers in Ecology and Evolution*, *9*, 337.
596 Kocher, A., Gantier, J.-C., Gaborit, P., Zinger, L., Holota, H., Valiere, S., … Murienne, J. (2017). Vector
597 soup: High-throughput identification of Neotropical phlebotomine sand flies using
598 metabarcoding. *Molecular Ecology Resources*, *17*(2), 172–182. doi: 10.1111/1755-
599 0998.12556
600 Kopylova, E., Navas-Molina, J. A., Mercier, C., Xu, Z. Z., Mahé, F., He, Y., … Knight, R. (2016). Open-
601 Source Sequence Clustering Methods Improve the State Of the Art. *MSystems*, *1*(1), e00003-
602 15. doi: 10.1128/mSystems.00003-15
603 Krassowski, M. (2021). ComplexUpset (Version v1.3.3). Zenodo. doi: 10.5281/ZENODO.3700590
604 Linard, B., Crampton-Platt, A., Moriniere, J., Timmermans, M. J. T. N., Andújar, C., Arribas, P., …
605 Vogler, A. P. (2018). The contribution of mitochondrial metagenomics to large-scale data
606 mining and phylogenetic analysis of Coleoptera. *Molecular Phylogenetics and Evolution*, *128*,
607 1–11. doi: 10.1016/j.ympev.2018.07.008
608 Lozano Mojica, J. D., & Caballero, S. (2021). Applications of eDNA metabarcoding for vertebrate
609 diversity studies in northern Colombian water bodies. *Frontiers in Ecology and Evolution*,
610 522.
611 Macher, J., Zizka, V. M. A., Weigand, A. M., & Leese, F. (2018). A simple centrifugation protocol for
612 metagenomic studies increases mitochondrial DNA yield by two orders of magnitude.
613 *Methods in Ecology and Evolution*, *9*(4), 1070–1074.

614 Majaneva, M., Diserud, O. H., Eagle, S. H., Hajibabaei, M., & Ekrem, T. (2018). Choice of DNA
615 extraction method affects DNA metabarcoding of unsorted invertebrate bulk samples.
616 *Metabarcoding and Metagenomics*, *2*, e26664.
617 Marquina, D., Andersson, A. F., & Ronquist, F. (2019). New mitochondrial primers for metabarcoding
618 of insects, designed and evaluated using in silico methods. *Molecular Ecology Resources*,
619 *19*(1), 90–104.
620 Marquina, D., Esparza-Salas, R., Roslin, T., & Ronquist, F. (2019). Establishing arthropod community
621 composition using metabarcoding: Surprising inconsistencies between soil samples and
622 preservative ethanol and homogenate from Malaise trap catches. *Molecular Ecology*
623 *Resources*, *19*(6), 1516–1530. doi: 10.1111/1755-0998.13071
624 Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads.
625 *EMBnet.Journal*, *17*(1), 10–12. doi: 10.14806/ej.17.1.200
626 Martins, F. M. S., Galhardo, M., Filipe, A. F., Teixeira, A., Pinheiro, P., Paupério, J., … Beja, P. (2019).
627 Have the cake and eat it: Optimizing nondestructive DNA metabarcoding of
628 macroinvertebrate samples for freshwater biomonitoring. *Molecular Ecology Resources*,
629 *19*(4), 863–876. doi: 10.1111/1755-0998.13012
630 Martoni, F., Valenzuela, I., & Blacket, M. J. (2019). Non-destructive DNA extractions from fly larvae
631 (Diptera: Muscidae) enable molecular identification of species and enhance morphological
632 features. *Austral Entomology*, *58*(4), 848–856. doi: 10.1111/aen.12419
633 Mata, V. A., Ferreira, S., Campos, R. M., da Silva, L. P., Veríssimo, J., Corley, M. F. V., & Beja, P. (2020).
634 Efficient assessment of nocturnal flying insect communities by combining automatic light
635 traps and DNA metabarcoding. *Environmental DNA*, *n/a*(n/a). doi: 10.1002/edn3.125
636 McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis
637 and graphics of microbiome census data. *PloS One*, *8*(4), e61217.

638 Nielsen, M., Gilbert, M. T. P., Pape, T., & Bohmann, K. (2019). A simplified DNA extraction protocol
639 for unsorted bulk arthropod samples that maintains exoskeletal integrity. *Environmental*
640 *DNA*, *1*(2), 144–154. doi: 10.1002/edn3.16
641 Svenningsen, C. S., Frøslev, T. G., Bladt, J., Pedersen, L. B., Larsen, J. C., Ejrnæs, R., … Dunn, R. R.
642 (2021). Detecting flying insects using car nets and DNA metabarcoding. *Biology Letters*, *17*(3),
643 20200833.
644 Tiede, J., Scherber, C., Mutschler, J., McMahon, K. D., & Gratton, C. (2017). Gut microbiomes of
645 mobile predators vary with landscape context and species identity. *Ecology and Evolution*,
646 *7*(20), 8545–8557. doi: 10.1002/ece3.3390
647 Tsuji, S., Yamanaka, H., & Minamoto, T. (2017). Effects of water pH and proteinase K treatment on
648 the yield of environmental DNA from water samples. *Limnology*, *18*(1), 1–7. doi:
649 10.1007/s10201-016-0483-x
650 Vamos, E. E., Elbrecht, V., & Leese, F. (2017). *Short COI markers for freshwater macroinvertebrate*
651 *metabarcoding* (No. 2167–9843). PeerJ Preprints.
652 Vesterinen, E. J., Ruokolainen, L., Wahlberg, N., Peña, C., Roslin, T., Laine, V. N., … Lilley, T. M. (2016).
653 What you need is what you eat? Prey selection by the bat Myotis daubentonii. *Molecular*
654 *Ecology*, *25*(7), 1581–1594.
655 Wägele, J. W., Bodesheim, P., Bourlat, S. J., Denzler, J., Diepenbroek, M., Fonseca, V., … Glöckner, F.
656 O. (2022). Towards a multisensor station for automated biodiversity monitoring. *Basic and*
657 *Applied Ecology*.
658 Wheeler, Q. D., Raven, P. H., & Wilson, E. O. (2004). Taxonomy: Impediment or Expedient? *Science*,
659 *303*(5656), 285–285. doi: 10.1126/science.303.5656.285
660 Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.
661 Wickham, H., Francois, R., Henry, L., & Müller, K. (2015). dplyr: A grammar of data manipulation. *R*
662 *Package Version 0.8.3*, *3*.

Yu, D. W., Ji, Y., Emerson, B. C., 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

Zenker, M. M., Specht, A., & Fonseca, V. G. (2020). Assessing insect biodiversity with automatic light traps in Brazil: Pearls and pitfalls of metabarcoding samples in preservative ethanol. *Ecology and Evolution*, *10*(5), 2352–2366.

Zizka, V. M., Geiger, M. F., Hörren, T., Kirse, A., Noll, N. W., Schäffler, L., … Sorg, M. (2022). Recommendations for tissue homogenisation and extraction in DNA metabarcoding of Malaise trap samples. *BioRxiv*, 2022.01.25.477667. doi: 10.1101/2022.01.25.477667

Zizka, V. M., Leese, F., Peinert, B., & Geiger, M. F. (2018). DNA metabarcoding from sample fixative as a quick and voucher-preserving biodiversity assessment method. *Genome*, *62*(3), 122–136.

Data accessibility statement

Raw sequence data for this project has been submitted to NCBI's SRA archive under accession number PRJNA817517

Benefit-Sharing Statement

Benefits from this research accrue from the sharing of our data and results on public databases as described above.

Author contributions

A.K. and V.Z. conceived and planned the experiments. A.K., V.Z., B.Z. and K.L. carried out the experiments. A.K., V.Z. and S.B. contributed to the interpretation of the results. A.K. took the lead in

writing the manuscript with contribution from VZ and SB. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

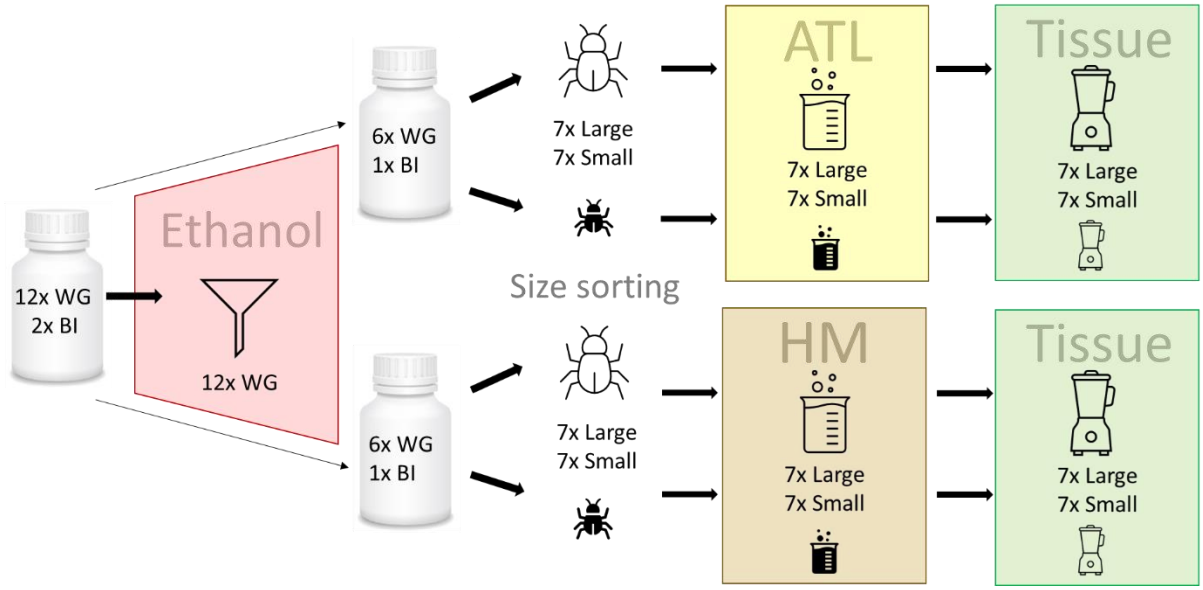

Figure 1: 14 Malaise trap samples were processed consecutively with three non-destructive and one destructive approach from homogenized tissue. In the non-destructive approaches, DNA was extracted using fixative ethanol and two lysis-based approaches (home-made buffer (HM), ATL buffer (ATL)).

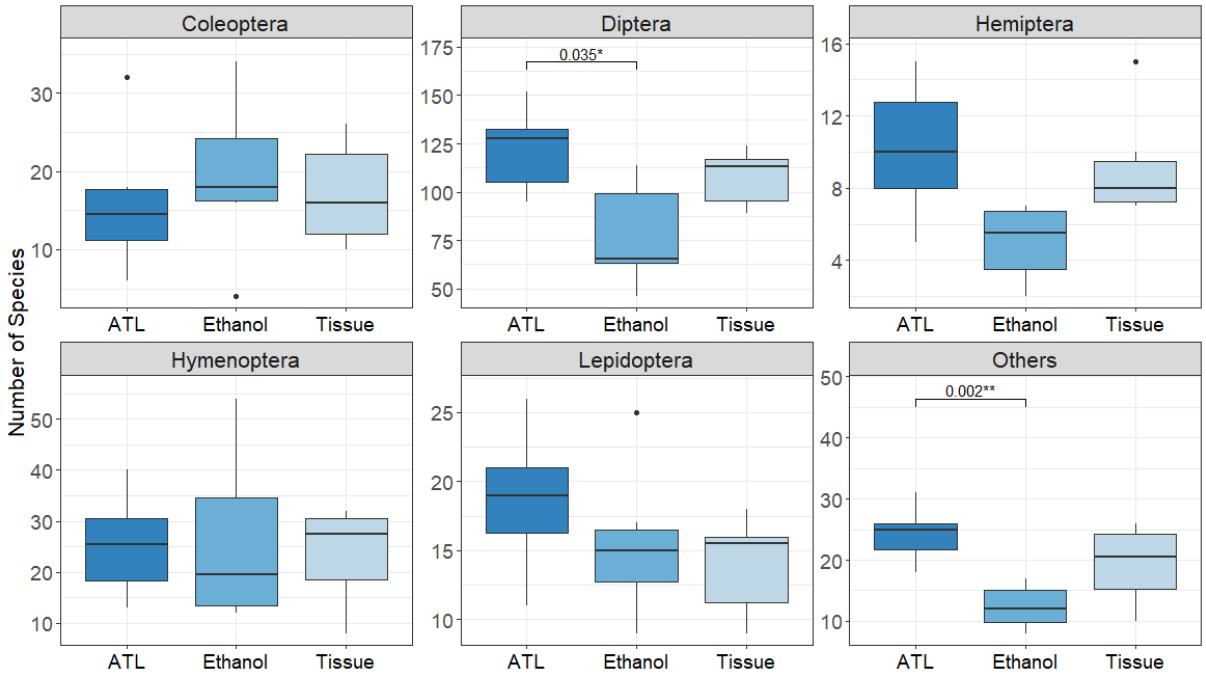

Figure 2: Mean number of species per taxonomic group depending on extraction method. Analysis includes the samples WG 1-6 and BI 1 with both size classes per sample combined (S + L).

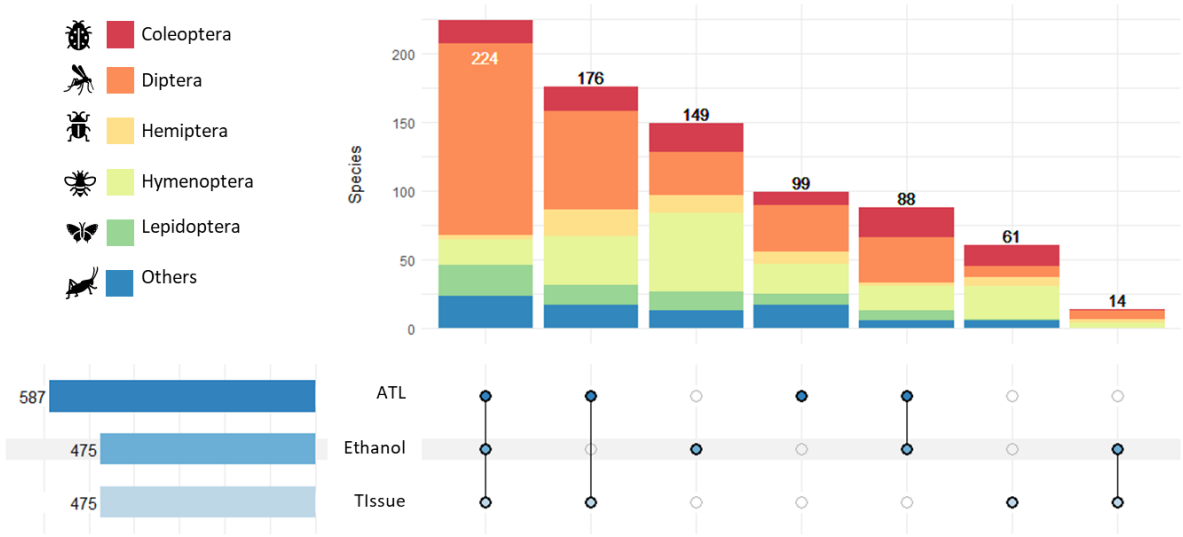

Figure 3: Total number of detected arthropod species depending on extraction method and taxon. Horizontal bars on the left indicate total number of detected arthropod species per extraction method (Blast ID >98%). Vertical bars indicate number of shared and unique species within and between extraction methods. Individual combinations of extraction methods are indicated with points of similar shading. Analysis includes the samples WG 1-6 with both size classes per sample combined (S + L). In addition, lysis buffer analysis includes all four incubation times combined (2h, 4h, 8h, 12h).

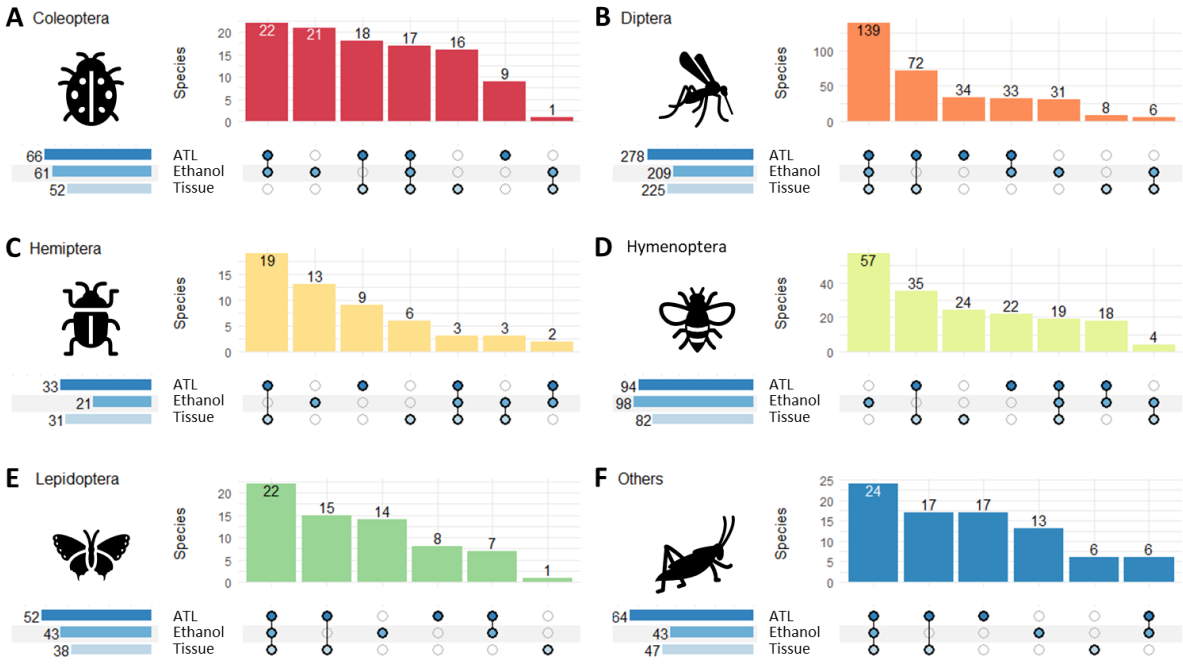

Figure 4: Total number of detected arthropod species depending on extraction method and order. Bars on the left indicate total number of detected species per extraction method (Blast ID >98%). Vertical bars indicate number of shared and unique species with and between extraction methods. Individual combinations of extraction methods are indicated with points of similar shading. Analysis include the samples WG 1-6 and both size classes (S + L) per sample and all time points (2h, 4h, 8h, 12h) combined. A) Coleoptera, B) Diptera, C) Hemiptera, D) Hymenoptera, E) Lepidoptera, F) Others.

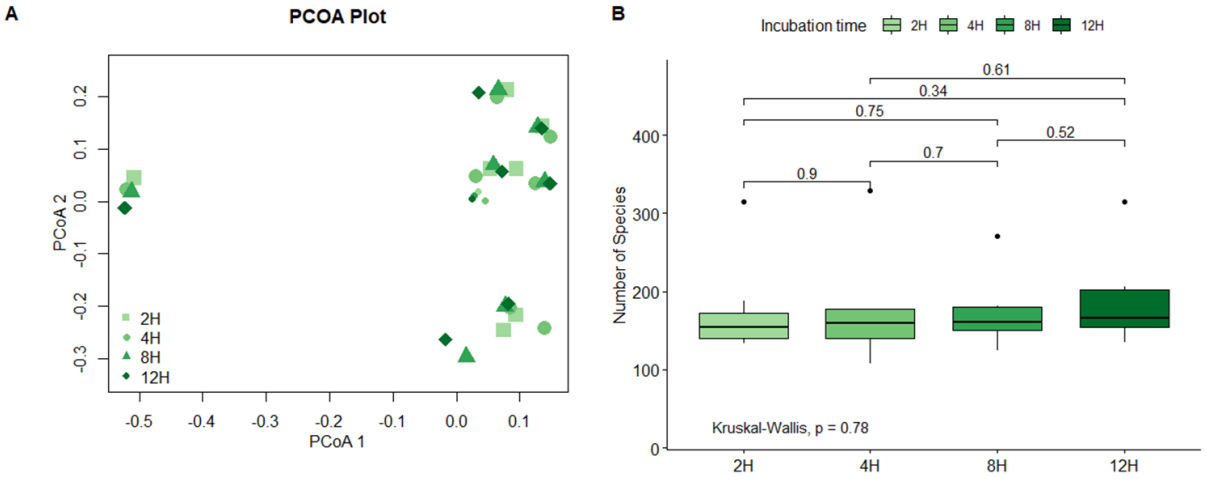

Figure 5: PCOA plot indicates no significant differences in assessed species communities (both size fractions merged) (PERMANOVA: F2 = 0.02, p = 1) depending on incubation time (A). Number of species did not show major differences depending on incubation time (B).

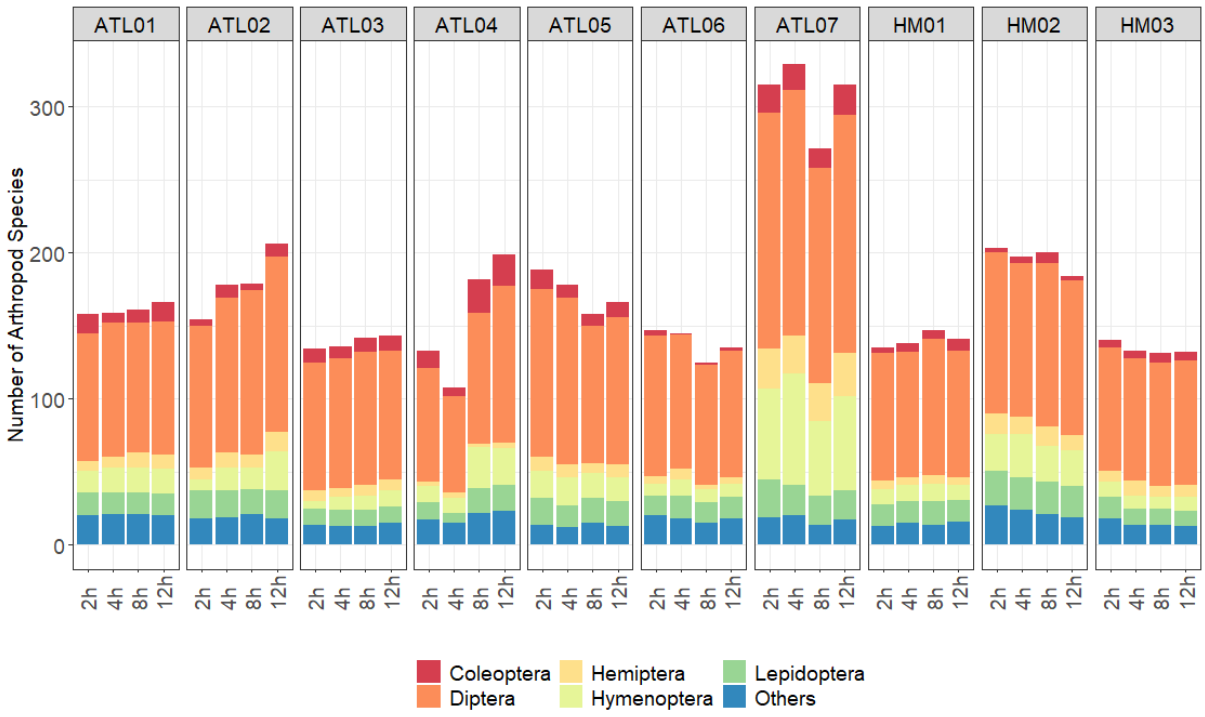

Figure 6: Number of identified species (blastID >= 98%) per sample depending on incubation time using two different extraction buffers (ATL & HM). Analysis includes both size classes per sample (S + L).

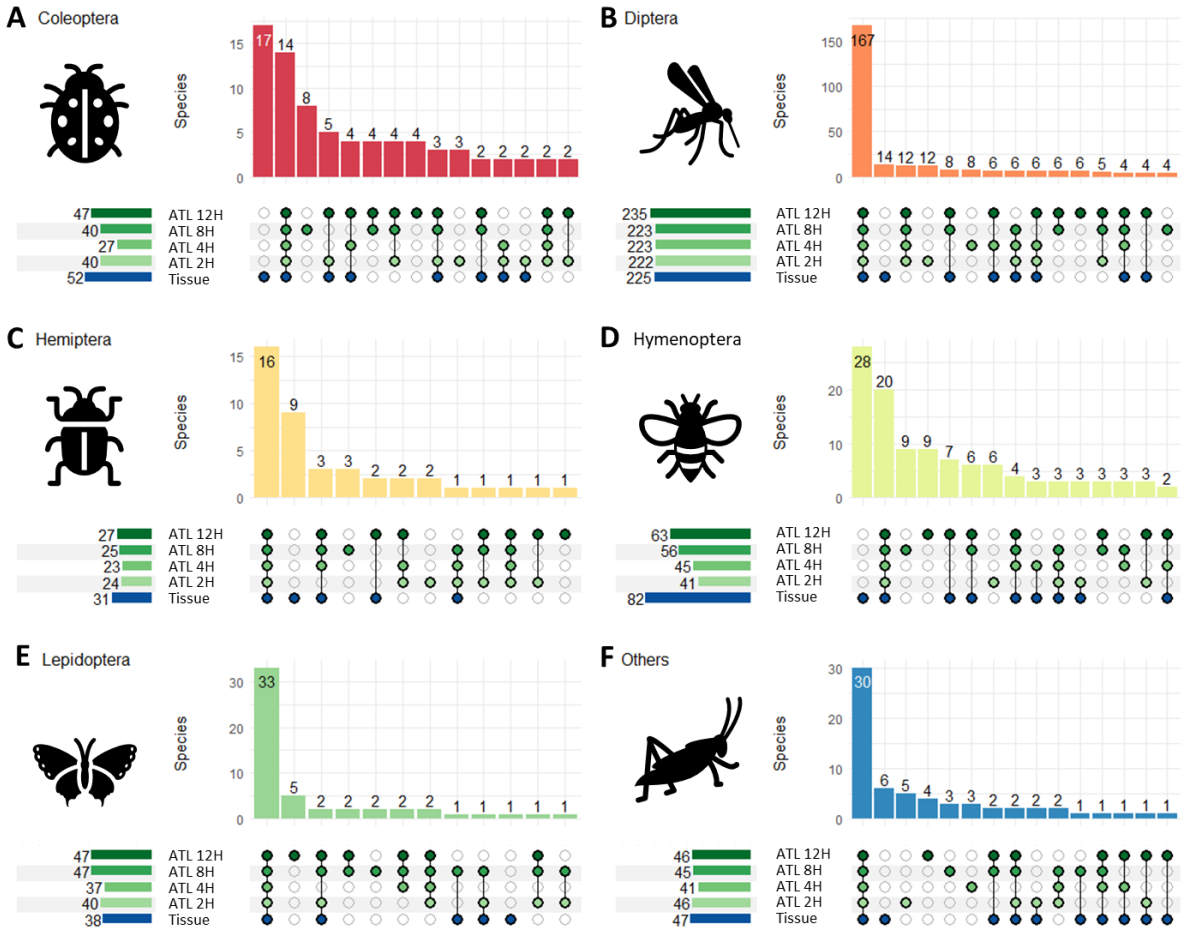

Figure 7: Total number of identified species (blastID >= 98%) with the ATL lysis buffer at different incubation times. Analysis includes both size classes per sample (S + L).