Check for updates

Received: 24 September 2019 | Revised: 3 June 2020 | Accepted: 9 June 2020

DOI: 10.1111/1755-0998.13212

RESOURCE ARTICLE

MOLECULAR ECOLOGY RESOURCES WILEY

DNA metabarcoding for biodiversity monitoring in a national park: Screening for invasive and pest species

Laura A. Hardulak[1,2] | Jérôme Morinière[1] | Axel Hausmann[1] | Lars Hendrich[1] | Stefan Schmidt[1] | Dieter Doczkal[1] | Jörg Müller[3,4] | Paul D. N. Hebert[5] | Gerhard Haszprunar[1]

[1]SNSB-Zoologische Staatssammlung München, Munich, Germany

[2]Ludwig-Maximilians-Universität München, Munich, Germany

[3]National Park Bavarian Forest, Grafenau, Germany

[4]Field Station Fabrikschleichach, Department of Animal Ecology and Tropical Biology, University of Würzburg, Biocenter, Rauhenebrach, Germany

[5]Centre for Biodiversity Genomics, University of Guelph, Guelph, ON, Canada

Correspondence
Laura A. Hardulak, SNSB-Zoologische Staatssammlung München, Münchhausenstraße 21, 81247 Munich, Germany.
Email: hardulak@snsb.de

Present address
Jérôme Morinière, AIM – Advanced Identification Methods GmbH, München, Germany

Funding information
Bundesministerium für Bildung und Forschung, Grant/Award Number: GBOL: BMBF FKZ 01LI1101 and 01LI1501; Bayerisches Staatsministerium für Wissenschaft, Forschung und Kunst

Abstract

DNA metabarcoding was utilized for a large-scale, multiyear assessment of biodiversity in Malaise trap collections from the Bavarian Forest National Park (Germany, Bavaria). Principal component analysis of read count-based biodiversities revealed clustering in concordance with whether collection sites were located inside or outside of the National Park. Jaccard distance matrices of the presences of barcode index numbers (BINs) at collection sites in the two survey years (2016 and 2018) were significantly correlated. Overall similar patterns in the presence of total arthropod BINs, as well as BINs belonging to four major arthropod orders across the study area, were observed in both survey years, and are also comparable with results of a previous study based on DNA barcoding of Sanger-sequenced specimens. A custom reference sequence library was assembled from publicly available data to screen for pest or invasive arthropods among the specimens or from the preservative ethanol. A single 98.6% match to the invasive bark beetle *Ips duplicatus* was detected in an ethanol sample. This species has not previously been detected in the National Park.

KEYWORDS

biodiversity, DNA barcoding, invasive species, metabarcoding, monitoring, pest species

1 | INTRODUCTION

The worldwide decline in biodiversity currently presents an urgent challenge facing humanity, and slowing down or halting this decline is an objective of broad international political agreement (Thomsen & Willerslev, 2015). A major barrier to achieving this objective is the lack of knowledge of biodiversity states and patterns on a global scale (Geijzendorffer et al., 2016; Lindenmayer et al., 2012). Hundreds or possibly thousands of species become extinct each year (Ceballos & Ehrlich, 2018; Chivian & Bernstein, 2008), and conservation of biodiversity depends upon ongoing monitoring efforts which can elucidate patterns of distribution and abundances of species

[Correction added on 20-October-2020, after first online publication: Ludwig-Maximilians-Universität München was added for Laura A. Hardulak.]

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2020 The Authors. *Molecular Ecology* Resources published by John Wiley & Sons Ltd

and populations (Geijzendorffer et al., 2016; Honrado, Pereira, & Guisan, 2016; Schmeller et al., 2015; Thomsen & Willerslev, 2015). A well-designed monitoring effort should provide an early warning of changes in the ecosystem which could otherwise become problems that are difficult or impossible to remediate (Bohmann et al., 2014; Lindenmayer et al., 2012). One such change is the introduction of animal and plant species to non-native geographical areas. With the globalization of trade, reduced travel time and immense passenger travel, species invasions have recently intensified (Keller, Geist, Jeschke, & Kühn, 2011; Sala et al., 2000), and are now one of the major recognized causes of biodiversity loss (Bellard, Cassey, & Blackburn, 2016; Ehrenfeld, 2010).

Accurate, rapid identifications of invasive species are needed to better manage the risks associated with alien species. An estimated 1% of all neozoans and neophytes become invasive with serious economic impacts (Meyerson & Reaser, 2002; Williamson, 1996). Some taxa which are innocuous or only minor pests in their native regions have unforeseen consequences after arriving in new areas lacking microbial control, competition or predators. For example, of the six most serious forestry pests introduced in North America, only the European gypsy moth had pest status in its indigenous range (Cock, 2003). In New Zealand, the introduced painted apple moth, *Orgyia anartoides* (Walker, 1855), from Australia was predicted to cause €33–205 million in damage if it was not eradicated (Armstrong & Ball, 2005).

Traditional biodiversity monitoring has relied on visual observation and identification of species and counting of individuals. These efforts may be hampered by a lack of available taxonomic expertise for morphological identifications, as well as nonstandard sampling techniques (Beng et al., 2016; Corlett, 2017; Ji et al., 2013; Thomsen & Willerslev, 2015). Towards the aim of fulfilling an urgent need for accurate large-scale biodiversity monitoring, molecular methods have been applied in recent years, particularly since the advent of DNA barcoding (Hebert, Ratnasingham, & de Waard, 2003). DNA barcoding (Hebert et al., 2003), the characterization of sequence variation in a standard DNA fragment, is a broadly applicable and objective method, which increases the speed and taxonomic resolution of specimen identification as well as reducing costs. In this way, DNA barcoding and, more recently, metabarcoding (Hajibabaei, Shokralla, Zhou, Singer, & Baird, 2011)—a process by which genetic material is extracted from mixed or bulk samples, amplified, sequenced by high-throughput sequencing (HTS) and analysed holistically—assist in augmenting biodiversity monitoring efforts (Ji et al., 2013). In its first few years, metabarcoding was shown to recover significant portions of existing biodiversity (Aylagas, Borja, & Rodríguez-Ezpeleta, 2014; Yu et al., 2012) and to reveal unknown patterns of biodiversity (Leray & Knowlton, 2015), and it has been successfully applied to large-scale biodiversity assessments (e.g. Elbrecht, Peinert, & Leese, 2017; Epp et al., 2012; Ji et al., 2013; Morinière et al., 2016; Shokralla, Spall, Gibson, & Hajibabaei, 2012; Taberlet, Coissac, Pompanon, Brochmann, & Willerslev, 2012; Yu et al., 2012). DNA barcoding and metabarcoding also permit species-level identifications when only eggs, larvae or parts of specimens are available for analysis. These may be intercepted at borders (e.g. wooden pallets at airports, ports, railway stations) as they are transported by vectors or accidentally by humans, such as in the ballast waters of ships, or with animals and plants in the food trade (Borrell, Miralles, Do Huu, Mohammed-Geba, & Garcia-Vazquez, 2017). For these reasons, HTS has been considered the ideal method for early warning of invasive species (Comtet, Sandionigi, Viard, & Casiraghi, 2015).

In terrestrial ecosystems, macroinvertebrates are often stored directly in ethanol following their collection. DNA can subsequently be harvested either directly from the specimens or from the preservative. Maceration of the specimens followed by subsequent extraction of DNA from a subsample of the homogenate is commonly practised (Yu et al., 2012), and it is probably both the simplest and the most effective way of securing a representative DNA extract from a bulk sample for subsequent metabarcoding (Elbrecht et al., 2017). However, there is a growing need to integrate sequence-based with morphological research (Silva-Santos, Ramirez, Galetti, & Freitas, 2018), and requirements to keep specimens intact for subsequent morphological control sometimes exist. Therefore, the efficiency and effectiveness of various nondestructive methods of sample preparation and DNA extraction of mixed samples for metabarcoding is a subject of ongoing research.

Additionally, an issue impacting the ability of metabarcoding to recover sequences representing the total biodiversity of a holistically homogenized sample is the bias in primer competition due to unequal specimen size (Elbrecht & Leese, 2015; Elbrecht et al., 2017; Leray & Knowlton, 2015). Larger specimens have more biomass and thus more DNA to contribute to lysed tissue pools. Therefore, larger individuals become overrepresented in sequencing results, and smaller ones underrepresented, increasing the risk of failure to detect taxa with small body sizes. Nondestructive ethanol-based DNA extraction methods have been recommended for their potential to provide solutions to sampling and vouchering challenges of metabarcoding (Hajibabaei, Spall, Shokralla, & van Konynenburg, 2012); and specifically, an ethanol filtration method has been shown to exhibit weak or even no correlation between specimen biomass and read numbers (Zizka, Leese, Peinert, & Geiger, 2019), thus potentially remediating the size-bias problem. As an objective of the present study is qualitative biodiversity analysis of mixed samples of invertebrates, we decided to supplement the standard homogenized tissue DNA extraction method with ethanol-based methods in 2018, in order to improve taxon recovery rates. The aims of the present study are to (a) perform biodiversity analysis comparing collection sites in and around the Bavarian Forest National Park (Nationalpark Bayerischer Wald, NPBW) and in two study years; and (b) construct a custom database of potential pest and invasive arthropod species in Germany based on public data sets and literature, and use it to screen our samples for these taxa.

The results reported in this study derive from two major DNA barcoding campaigns: "Barcoding Fauna Bavarica" (BFB, www.faunabavarica.de, Haszprunar, 2009) and the "German Barcode of Life" project (GBOL, www.bolgermany.de, Geiger, Astrin, et al., 2016), which aim to establish a DNA barcode reference library for all German species. Since their initiation in 2009, DNA barcodes for more than 23,000 metazoan species in Germany have been

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

assembled. Through the analysis of more than 250,000 specimens, the SNSB – Bavarian State Collection of Zoology (ZSM, see www.barcoding-zsm.de) has made a major contribution to parameterization of the global DNA barcode library maintained in the Barcode of Life Data System (BOLD, www.boldsystems.org, Ratnasingham & Hebert, 2007). Currently, the DNA barcode library created by researchers at the ZSM represents the second-most comprehensive library of any nation, with good coverage for Coleoptera, Diptera, Heteroptera, Hymenoptera, Lepidoptera, Neuroptera, Orthoptera, Araneae and Opiliones, and Myriapoda (see Table 1).

2 | MATERIALS AND METHODS

2.1 | Sample collection

Nine Malaise traps were deployed around the perimeter of the Bavarian Forest National Park from May to September in 2016 and in 2018 (Figure 1; Table 2). Traps were emptied twice a month, producing 10 samples for each trap year (collection periods designated 1 May to 2 September), for a total of 90 samples annually. In 2016, the original preservative ethanol was changed prior to transportation to the laboratory. Samples were stored in 80% ethanol at room temperature until laboratory analysis. 2016 samples were processed in the laboratory in November 2016. The first 54 samples of 2018 were processed in the laboratory in August 2018, and the latter 36 were processed in November 2018; the original preservative ethanol was processed in December 2018.

2.2 | DNA extraction

2.2.1 | Destructive methods

Preservative ethanol was removed, and specimens were transferred to 500-ml PET bottles, dried at 70°C for at least 3 hr and then left at room temperature overnight if necessary, to evaporate off

TABLE 1 Major arthropod orders and respective species and specimen numbers represented by DNA barcode sequences from the ZSM

Order	Number of barcoded individuals	Number of species	Reference
Coleoptera	15,948	3,514	Hendrich et al. (2015)
	819	78	Raupach, Hannig, Morinière, and Hendrich (2016)
	690	47	Raupach, Hannig, Morinière, and Hendrich (2018)
	13,516	2,846	Rulik et al. (2017)
Diptera	45,040	2,453	Morinière et al. (2019)
Ephemeroptera, Plecoptera and Trichoptera	2,613	363	Morinière et al. (2017)
Heteroptera	1,742	457	Raupach et al. (2014)
	712	67	Havemann et al. (2018)
Hymenoptera	4,118	561	Schmidt, Schmid-Egger, Morinière, Haszprunar, and Hebert (2015)
	4,362	1,037	Schmidt et al. (2017)
	3,695	661	Schmid-Egger et al. (2019)
Lepidoptera	1,395	331	Hausmann, Haszprunar, and Hebert (2011)
	3,467	957	Hausmann, Haszprunar, Segerer, et al. (2011)
	2,130	219	Hausmann et al. (2013)
Neuroptera	237	83	Morinière et al. (2014)
Orthoptera	748	122	Hawlitschek et al. (2017)
Araneae and Opiliones	3,537	598	Astrin et al. (2016)
Myriapoda	320	122	Spelda, Reip, Oliveira Biener, and Melzer (2011)

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 1 Overview of the Malaise trap sample sites in the Bavarian Forest National Park (left). Example image of a Malaise trap setup in the National Park (right) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Locations of the nine Malaise traps deployed in this study in 2016 and 2018

Plot	Location	Latitude (deg.)	Longitude (deg.)	Altitude (m a.s.l.)	In the NPBW?
Igg35	Iggensbach	48.73	13.10	379	N
Jos21	Assmann	48.52	13.72	364	N
Sal25	Saldenburg	48.80	13.35	505	N
T1_2	Plattenhausen_1	48.92	13.40	740	Y
T1_34	Plattenhausen_1	48.94	13.42	819	Y
T1_52	Plattenhausen_1	48.95	13.44	945	Y
T1_63	Plattenhausen_1	48.96	13.45	1,287	Y
T3_50	Scheuereck_3	49.10	13.32	1,182	Y
T4_64	Lackenberg_4	49.10	13.28	1,137	Y

[a]Abbreviation: NPBW, nationalpark bayerischer wald.

the residual ethanol. In 2016, dried specimens were ground with a sterilized pestle to homogenize the tissue. Samples from 2018 were homogenized in 500-ml PET bottles with 5–10 sterile steel balls using a FastPrep 96 (MP Biomedicals). Because the specimens were not quantified (e.g. by weighing or counting them) prior to homogenization, a 9:1 mixture of insect lysis buffer and Proteinase K was added in sufficient amounts to cover the ground specimens. Lysis was performed overnight at 56°C. Lysates were then allowed to cool to 20°C and 200-µl aliquots were used for DNA extraction using a Qiagen DNEasy Blood & Tissue Kit (Qiagen) following the manufacturer's instructions.

2.2.2 | Nondestructive methods

DNA extraction from preservative ethanol

For extraction of DNA from the preservative ethanol, we followed protocols employed by Hajibabaei et al. (2012). This evaporative ethanol technique was performed on five samples (1 May to 1 July) from each of the nine traps in 2018. A 50-ml aliquot of preservative ethanol was taken from each bottle. From this, two 1-ml aliquots were placed into Eppendorf tubes and allowed to dry overnight at 56°C. Fifty microlitres of molecular water was added the next morning, and the tubes were vortexed. Afterwards, DNA extraction was performed on the entire 50-µl sample using the DNeasy Blood and Tissue kit.

For another five samples (trap T3-50B 2018; 2 July, 1 August, 2 August, 1 September, 2 September II) a 50-ml aliquot of ethanol was used for filtration of DNA and tissue residuals using analytical test filter funnels (0.45 µm, Fisher Scientific) equipped with a water jet pump. After ethanol was filtered, the filter funnels were lysed overnight at 56°C. DNA extraction was performed using the DNeasy Blood and Tissue kit following the manufacturer's instructions and eluted into 50 µl of molecular-grade water.

Semilysis of bulk samples

Five bulk samples of 2018 (Sal-25, 2 July; T1-02, 2 July; T1-52, 2 July; T1-34, 2 July; and T3-50, 1 July) were used for semilysis and

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

subsequent DNA extraction. PET bottles (500 ml) were filled with sufficient amounts of lysis mixture (9:1 insect lysis buffer/Proteinase K) and incubated overnight at 56°C. For DNA extraction, 1 ml of the lysate was used following the above-mentioned methods using the DNeasy Blood and Tissue kit. The remaining bulk sample was then dried, and the residual insect lysis buffer was discarded. Samples were then homogenized as described in the Section 2.2.1 above.

2.3 | Amplification of the CO1 barcode fragment

From each sample, 5 µl of extracted genomic DNA was used, along with 20 µl of the following mixture: 1.5 µl Mango TAQ (Bioline), 5 µl forward and 5 µl reverse HTS-adapted minibarcode primers of Leray et al. (2013), 6.25 µl $MgCl_2$, 10 µl dNTPs, 25 µl Mango Buffer and 62.5 µl molecular-grade water. DNA extractions from preservative ethanol were amplified using a MyTaq Plant-PCR Kit (Bioline). PCR conditions were as follows: 2 min at 96°C; three cycles of 15 s at 96°C, 30 s at 48°C and 90 s at 65°C; 30 cycles of 15 s at 96°C, 30 s at 55°C and 90 s at 65°C; 10 min at 72°C (see Morinière et al., 2016). Amplification success and fragment lengths (~350 bp) were observed using gel electrophoresis on a 1% agarose gel.

2.4 | Purification and next generation sequencing

Amplified DNA was cleaned up by centrifugation of each sample with a 1:10 mixture of 3 m sodium acetate and ice cold 100% ethanol and resuspended in 50 µl molecular-grade water before proceeding. Illumina Nextera XT (Illumina Inc.) indices were ligated to the samples by PCR, and ligation success was confirmed by gel electrophoresis (as described in Morinière et al., 2019). DNA concentrations were measured using a Qubit fluorometer (Life Technologies), and samples were combined into 40-µl pools containing equimolar concentrations of 100 ng each. Pools were loaded into a 1.5% agarose gel, run at 90 V for 45 min, and bands of target amplicons were excised with sterilized razor blades, and purified with a GeneJet Gel Extraction kit (Life Technologies), following the manufacturer's instructions. A final elution volume of 20 µl was used. Sequencing runs were performed on an Illumina MiSeq using V2 chemistry (2 × 250 bp, 500 cycles, 20 million paired-end reads maximum).

2.5 | Pre-processing and clustering of sequence data

All FASTQ files generated were combined although they were sequenced on separate runs throughout the study period. Sequence processing was performed with the VSEARCH version 2.4.3 suite (Rognes, Flouri, Nichols, Quince, & Mahé, 2016) and CUTADAPT version 1.14 (Martin, 2011). Because some runs did not yield reverse reads of sufficiently high quality to enable paired-end merging, only forward reads were utilized. Forward primers were removed with CUTADAPT. Quality filtering was with the fastq_filter program of VSEARCH, fastq_maxee 2, with a minimum length of 100 bp. Sequences were dereplicated with derep_fulllength, first at the sample level, and then concatenated into one FASTA file, which was then dereplicated. Chimeric sequences were removed from the FASTA file using uchime_denovo. Remaining sequences were clustered into operational taxonomic units (OTUs) at 97% identity with cluster_size, and an OTU table was created with usearch_global. To reduce probable false positives, a cleaning step was employed which excluded read counts in the OTU table that represented less than 0.01% of the total read count for their respective sample (see Elbrecht & Steinke, 2019).

2.6 | Construction of reference databases and sequence identification

2.6.1 | BIN-based reference library

All arthropod sequences on BOLD were downloaded (FASTA format, including private and public data) to create a general reference database containing hierarchical taxonomic information and barcode index numbers (BINs). To create this database, downloaded FASTA files were concatenated and imported into GENEIOUS (version 10 Biomatters) (Kearse et al., 2012). To aid the monitoring of species of interest, a broad list of potentially relevant arthropod species was compiled from the following literature sources: Index of Economically Important Lepidoptera (Zhang, 1994), and *Die Forstschädlinge Europas* ("The Forest Pests of Europe") (Pschorn-Walcher & Schwenke, 1982). Of the Index of Economically Important Lepidoptera, 2,684 species names were found on BOLD. Of the Forest Pests of Europe, 294 species names were found on BOLD. About two-thirds (1,962/2,978) of these species had BINs. OTUs were BLASTed (MEGABLAST, default parameters) against the downloaded database. The result was joined to the OTU table in LIBREOFFICE, where the spreadsheet of pest names and BINs was used to cross-check with the BLAST results. All of these BINs and species names available on BOLD were added to a publicly available data set named "Dataset – DS-BWPST Database of Pest Species of Insects in Germany" (data set https://doi.org/10.5883/DS-BWPST).

2.6.2 | Pest and invasive species custom reference libraries

Reference sequences for species from the following sources were compiled into a list of 1,017 names: Nature protection warning list of the German Federal Office for Nature Conservation in Bonn ("Erstellung einer Warnliste in Deutschland noch nicht vorkommender invasiver Tiere und Pflanzen") (Rabitsch, Gollasch, Isermann, Starfinger, & Nehring, 2013), terrestrial arthropods only; "Die invasiven gebietsfremden Arten der Unionsliste der Verordnung (EU) Nr.1143/2014 -Erste Fortschreibung

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2017" (Nehring and Skowronek); The International Union for Conservation of Nature's Red List of Threatened Species (IUCN, 2019), accessed online, https://www.iucnredlist.org, filter criteria of phylum = Arthropoda, land regions = Europe, Geographical scale = global, Red List Category = Critically Endangered, Endangered, Extinct in the wild, Lower risk/Conservation dependent, near threatened, or vulnerable; the European Plant Protection Global Database (https://gd.eppo.int/country/DE), filter criteria of "Germany"; as well as the following 28 widely known invasive species (with one synonym), if not already listed: *Periplaneta americana* (Linnaeus, 1758), *Harmonia axyridis* (Pallas, 1773), *Stictocephala bisonia* (Kopp and Yonke, 1977), *Anoplophora chinensis* (Forster, 1771), *Corythucha ciliata* (Say, 1832), *Rhagoletis completa* (Cresson, 1929), *Sceliphron curvatum* (Smith, 1870), *Leptinotarsa decemlineata* (Say 1824), *Reticulitermes flavipes* (Kollar, 1837), *Anoplophora glabripennis* (Motschulsky, 1853), *Hulecoeteomyia japonica* (Theobald, 1901), *Aedes japonicus* (Theobald, 1901), *Aedes koreicus* (Edwards, 1917), *Dryocosmus kuriphilus* (Yasumatsu, 1951), *Aproceros leucopoda* (Takeuchi, 1939), *Cacyreus marshalli* (Butler, 1898), *Dreyfusia nordmannianae* (Eckstein, 1890), *Frankliniella occidentalis* (Pergande, 1895), *Leptoglossus occidentalis* (Heidemann, 1910), *Cameraria ohridella* (Deschka and Dimic, 1986), *Cydalima perspectalis* (Walker, 1859), *Monomorium pharaonis* (Linnaeus, 1758), *Hypoponera punctatissima* (Roger, 1859), *Phyllonorycter robiniella* (Clemens, 1859), *Drosophila suzukii* (Matsumura, 1931), *Trialeurodes vaporariorum* (Westwood, 1856), *Diabrotica virgifera* (J.L. LeConte, 1868), *Viteus vitifoliae* (Fitch, 1855) and *Ectobius vittiventris* (Costa, 1847).

Sequences were downloaded using the R (R Core Team, 2019) package *BOLD* (Chamberlain, 2018). Of the 1,004 total species names, 361 were found in BOLD. These were exported as a tab-separated file and processed into FASTA format with Linux command lines. The remaining species were searched for on NCBI GenBank (advanced search, criteria including ["COI" OR "CO1" OR "COXI" OR "COX1"]). Forty-one of the species names were found and downloaded as FASTA files. To combine the sequences from both sources into a single database and BLAST, we used BOLD_NCBI_Merger (Macher, Macher, & Leese, 2017). The highest scoring pair of the top hit (NCBI BLAST+, outfmt 6) for each OTU was imported into LIBREOFFICE, joined with the OTU table, and filtered. A taxonomic neighbour-joining tree was constructed using the BOLD website. All arthropod species and corresponding BINs on the list that were available on BOLD were added to a publicly available data set named "Dataset – DS-BFNWARN Bundesamt für Naturschutz Warnliste, Arthropoden" (data set https://doi.org/10.5883/DS-BFNWARN).

2.7 | Biodiversity analysis

As DNA metabarcoding is not quantitative (Krehenwinkel et al., 2017; Piñol, Senar, & Symondson, 2019) we utilized presence–absence data of BINs recovered at ≥97% identity over geographical areas represented by Malaise trap locations to calculate many of the biodiversity metrics. The OTU table indicates which BINs (or higher corresponding taxa) were detected in each collection event. To calculate detection frequencies, all counts in the table greater than zero were set to one. In this way, row sums across the table indicate the number of samples from which a particular taxon was recovered, while column sums indicate the total numbers of taxa recovered from a sample. Presence–absence data for the homogenized samples for all traps from 2016 and 2018 were also analysed together with a data set from the Global Malaise Trap Program (GMTP) downloaded from BOLD, project "GMTPE Germany Malaise 2012" (see Geiger, Moriniere, et al., 2016). Frequencies of BIN detection throughout the growing seasons could then be compared for each of the three years. Bar and line charts were created with GGPLOT2 (Wickham, 2016) or base R.

The presence of BINs in the 2016 and 2018 samples was used to calculate Jaccard distances and dissimilarity matrices for traps inside and outside the National Park, with the R packages VEGAN (Dixon, 2003) and BETAPART (Baselga & Orme, 2012). A Mantel test was performed to compare the study years in terms of their dissimilarities among trap sites, utilizing the R packages GEOSPHERE (Hijmans, Williams, Vennes, & Hijmans, 2017) and ADE4 (Dray & Dufour, 2007). Analysis of similarities (ANOSIM) tests to compare BIN compositions of trap sites inside and outside of the park were performed with the anosim function of VEGAN: *Community Ecology Package* (Oksanen et al., 2010). Additionally, principal component analyses for the 2016 and 2018 taxonomic composition data for each trap site were performed based on seven-level taxonomic identifications of OTUs and their read counts, with the R package AMPVIS2 (Andersen, Kirkegaard, Karst, & Albertsen, 2018), amp_ordinate function, Hellinger transform.

3 | RESULTS

3.1 | Biodiversity analysis (BOLD BIN-based database)

A total of 19,727 OTUs were produced by the pipeline. Of these, 12,513 matched at ≥73% identity to the database downloaded from BOLD. After filtering for alignment lengths of ≥100 bp, *E*-value of 10e-6 and ≥97% identity to the reference sequences, 5,782 matches remained. The majority of matches belonged to Arthropoda, with the majority of those belonging to Diptera (3,169), Hymenoptera (1,173), Lepidoptera (527) and Coleoptera (411). Table 3 lists total BIN detections broken down by order in 2016 and 2018, and the proportion of BINs which were recovered in both years (percentage overlap). Total read numbers produced per sample are given in Table S1, and rarefaction curves for BINs detected are in Figure S1.

Of the BOLD BIN-based database records to which OTUs matched at ≥97%, roughly half (2,918) had species-level taxonomic classifications in BOLD. The rest of the records to which OTUs

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Comparison of total BIN detections within Malaise trap surveys in 2016 and 2018. The overlap indicates the number of identical BINs detected in both survey years

Class	Order	2016	2018	Overlap (%)
Arachnida	Araneae	67	42	42
	Mesostigmata	2	3	25
	Opiliones	2	4	50
	Sarcoptiformes	2	2	100
Collembola	Entomobryomorpha	6	6	71
	Symphypleona	4	2	50
Insecta	Blattodea	2	3	67
	Coleoptera	268	234	40
	Dermaptera	3	3	100
	Diptera	2,119	1,900	61
	Ephemeroptera	2	2	0
	Hemiptera	94	92	46
	Hymenoptera	731	709	45
	Lepidoptera	328	351	44
	Mecoptera	3	3	100
	Neuroptera	19	17	44
	Odonata	0	14	0
	Orthoptera	13	17	50
	Plecoptera	16	10	53
	Psocodea	9	9	100
	Raphidioptera	4	3	75
	Thysanoptera	1	1	0
	Trichoptera	24	19	59
Malacostraca	Isopoda	0	3	0
Gastropoda	Stylommatophora	1	3	0

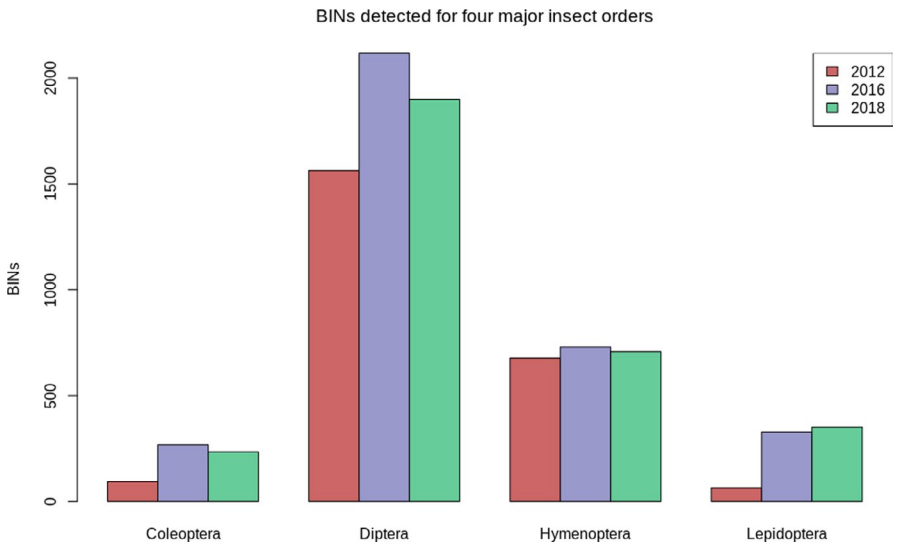

FIGURE 2 Detected BINs belonging to the orders Coleoptera, Diptera, Hymenoptera, and Lepidoptera within study years 2016 and 2018 [Colour figure can be viewed at wileyonlinelibrary.com]

matched were classified to lower levels. This is a consequence of the BIN system assigning BINs to sequence clusters algorithmically, whereas taxonomic classifications must be assigned by taxonomists to voucher specimens from which barcode sequences are obtained, a process which requires more time. At the time of writing, an effort is underway to provide taxonomic classifications for all records in BOLD with BINs, with particular emphasis on Diptera and Hymenoptera.

In 2016, 3,430 total BIN matches were detected from all tissue-based (homogenized) samples, and 2,957 in 2018 (counts include BINs belonging to classes Arachnida, Chilopoda, Clitellata, Collembola, Diplopoda, Gastropoda, Insecta and Malacostraca).

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

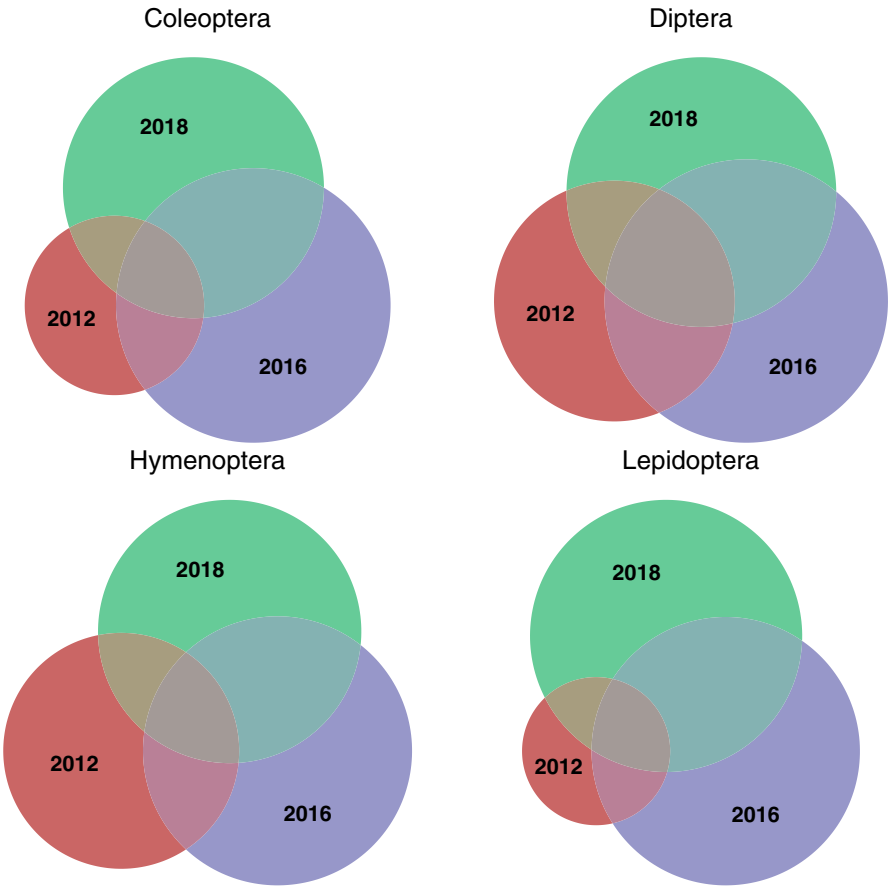

FIGURE 3 Venn diagrams depict the overlaps in BINs belonging to Coleoptera, Diptera, Hymenoptera, and Lepidoptera which were detected in the 2012 GMTP dataset and the two study years [Colour figure can be viewed at wileyonlinelibrary.com]

Figure 2 compares BIN detections within four major insect orders for the two study years and for the 2012 GMTP data set. Figure 3 depicts proportions of shared BINs between the three years for the same orders. BIN recoveries tended to peak in June or early July of each year (Figure 4). Counts of shared BINs between 2016 and 2018 for the four orders are shown in Figure 4 as black lines; for comparison, coloured dotted lines represent counts of individual BINs (presence–absence data for each collection period) for each year. Coloured solid lines take into account how many times BINs were detected in each collection period (total BIN detections). Diptera was the largest order by BIN count. In this order, 2,119 BINs were detected in 2016 (homogenized tissue), 1,900 in 2018 (homogenized tissue) and 2,021 in 2018 (all extraction methods in total).

Based on presence and absence of BINs, a Mantel test revealed a significant correlation between matrices of the mean Jaccard distances by trap sites in 2016 with those of 2018 (r = .4995, p = .005). Based on read abundances, biodiversity analyses of taxa in each trap for 2016 and 2018 are shown as principal component analyses in Figure 5. Malaise traps Igg35, Jos21 and Sal25, which are outside of the National Park, can be observed here clustering the furthest along PC2 in 2016 and PC1 in 2018, compared to all other traps, which are within the park. Additionally, ANOSIM tests showed significant differences between BIN detections in traps inside versus outside of the park in both years (2016 r = .2, p = 2e-04; 2018 r = .239, p = 1e-04).

3.2 | Economically important terrestrial arthropods and other species of interest

A total of 83 species names and 118 BINs from the list compiled from *Economically Important Lepidoptera* and *Forest Pests of Europe* matched in the BOLD database BLAST results for all samples (≥97% sequence similarity, *E*-value ≤ 10e-6, highest scoring pairs). We chose two cases of detected species of interest from this list: the noctuid *Lymantria dispar* (Linnaeus, 1758), a common forest pest, and the tortricid *Epinotia tedella* (Clerck, 1759), the presence of which relates to that of a potential regulatory parasitoid species of ichneumonid wasps (*Lissonota dubia* Holmgren, 1856). Total numbers of collection events in which these species of interest were detected in each year are shown in Figures 6 and 7. *Lymatria dispar* is an invasive lepidopteran listed in the *Index of Economically Important Lepidoptera* (Zhang, 1994). Eurasian in origin, it was introduced to the USA in the 19th century. We detected its sequences at 100% match to the database in Malaise trap Jos21 in May and the second collection of July 2016; in 2018 it was found in the same trap but more frequently: in every collection through August, and also in trap T1-34 in the first collection of June (Figure 6). Interestingly, we also observed similar patterns of presence/absence for *E. tedella* and its parasite, *Lissonota dubia* (Figure 7).

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

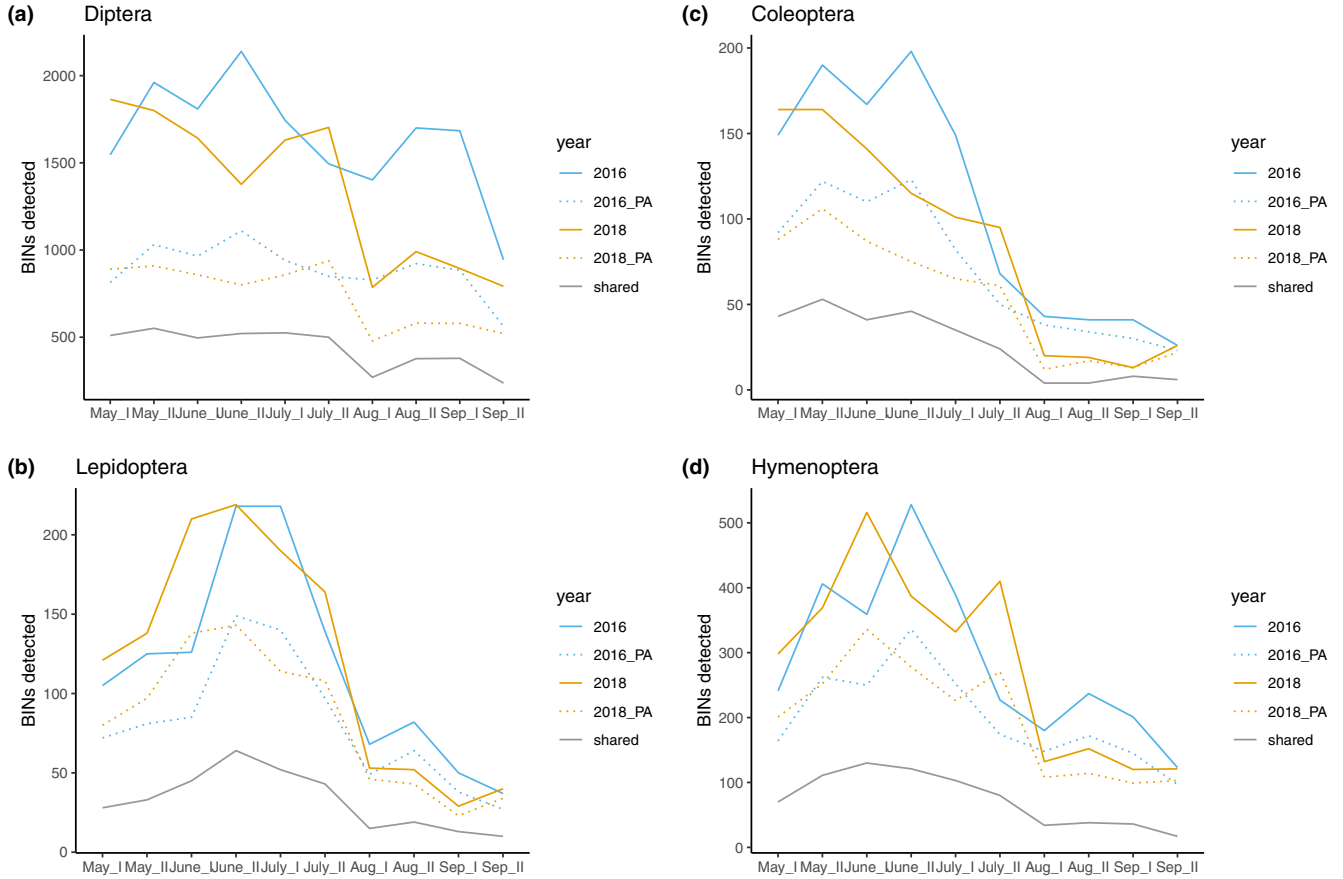

FIGURE 4 Breakdown of BINs detected in the two survey years by the four Orders (Diptera [a], Coleoptera [b], Lepidoptera [c], and Hymenoptera [d]). Colored solid lines take into account how many times BINs were detected in each collection period. "PA" denotes presence-absence BIN counts. Black lines indicate counts of BINs shared between both years [Colour figure can be viewed at wileyonlinelibrary.com]

3.3 | Species of interest custom database

Two species from our species of interest database matched to the samples' OTUs by BLAST at ≥97%: the lasiocampid moth *Dendrolimus superans* (Butler, 1877) and the bark beetle *Ips duplicatus* (Sahlberg, 1836) (Table 4), both from the warning list of the German Federal Office for Nature Conservation (Rabitsch et al., 2013). *D. superans* (BOLD: AAB6845) matched at 99.55% identity in Malaise trap sample T1-52 (inside the National Park), collection 1 September 2016.

Dendrolimus superans is endemic to Siberia and is a pest of over 20 species of coniferous plants. It has not yet been observed in Germany (Rabitsch et al., 2013). It also shares the BIN BOLD: AAB6845 with *Dendrolimus pini* (Linnaeus, 1758), which is known throughout most of Europe, including Germany. This result illustrates that, because a small custom database was used for this task, consisting of *only* species of interest, hits must be investigated further when the possibility exists that a specimen actually belongs to a closely related species not in this database.

Figure 8 presents a section of a neighbour-joining tree from barcode sequences on BOLD showing representatives of these species clustering together, also with the OTU sequence in question ("Unknown Specimen"). As observed by the BLAST against the general BOLD database, *D. pini* was also detected at a similar identity (99.5%) in the same trap in the BLAST against the BOLD BIN-based database. Therefore, it is probable that the latter was the species which was collected. Further integrative taxonomic study is needed to examine whether *superans* may better be downgraded to subspecies rank or synonymy of *pini*.

Ips duplicatus (BOLD: ACD5566) matched at 98.64% identity to the database in Malaise trap T3-50 (inside the National Park), collection 2 July 2018, filtered ethanol sample (Table 4). *I. duplicatus* is endemic to northern Europe, where it is a pest of pine trees (*Pinus* spp.), whereas it is unknown if it additionally poses a threat to biodiversity. The species was unknown in Germany at the time of publication of the warning list, but has recently been spreading southward, through central, eastern and southern Europe (Fiala & Holuša, 2019). Although another congeneric species, *Ips typographus* (Linnaeus, 1758) (BOLD: ACT0826), a keystone pest species in the Bavarian Forest National Park (Müller, Bußler, Goßner, Rettelbach, & Duelli, 2008), was also detected in the same trap at 100% identity, these two species' barcode sequences cluster less closely together, and they do not share a BIN. The present result is therefore likely to be a case of correct molecular identification of *I. duplicatus*, and to represent the first detection of this invasive saproxylic beetle in the National Park.

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

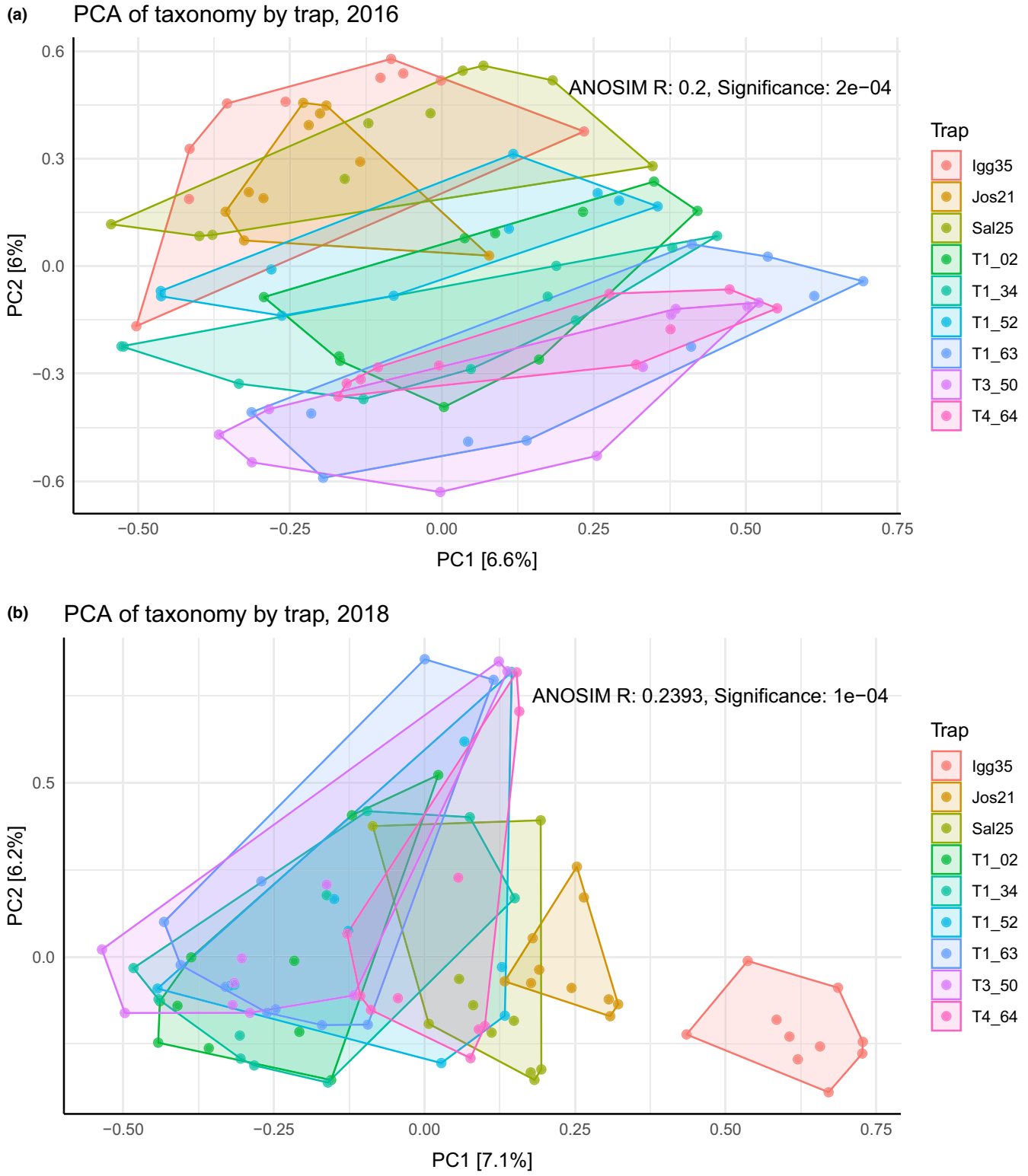

FIGURE 5 Principal component analyses of read abundances and 7-level taxonomic assignments of OTUs, for survey years 2016 (a) and 2018 (b) [Colour figure can be viewed at wileyonlinelibrary.com]

4 | DISCUSSION

In the present study, we have been able to accomplish large-scale biomonitoring of the largest national park in Europe, using DNA metabarcoding. By way of presence–absence and read count-based biodiversity analyses, we observed trends in frequencies of observations of taxa throughout two years, utilizing bulk samples from Malaise traps at sites inside and outside of the park, de novo OTU generation and existing reference libraries. Analysing the data from homogenized samples from 2016 and 2018 together with data from a GMTP

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

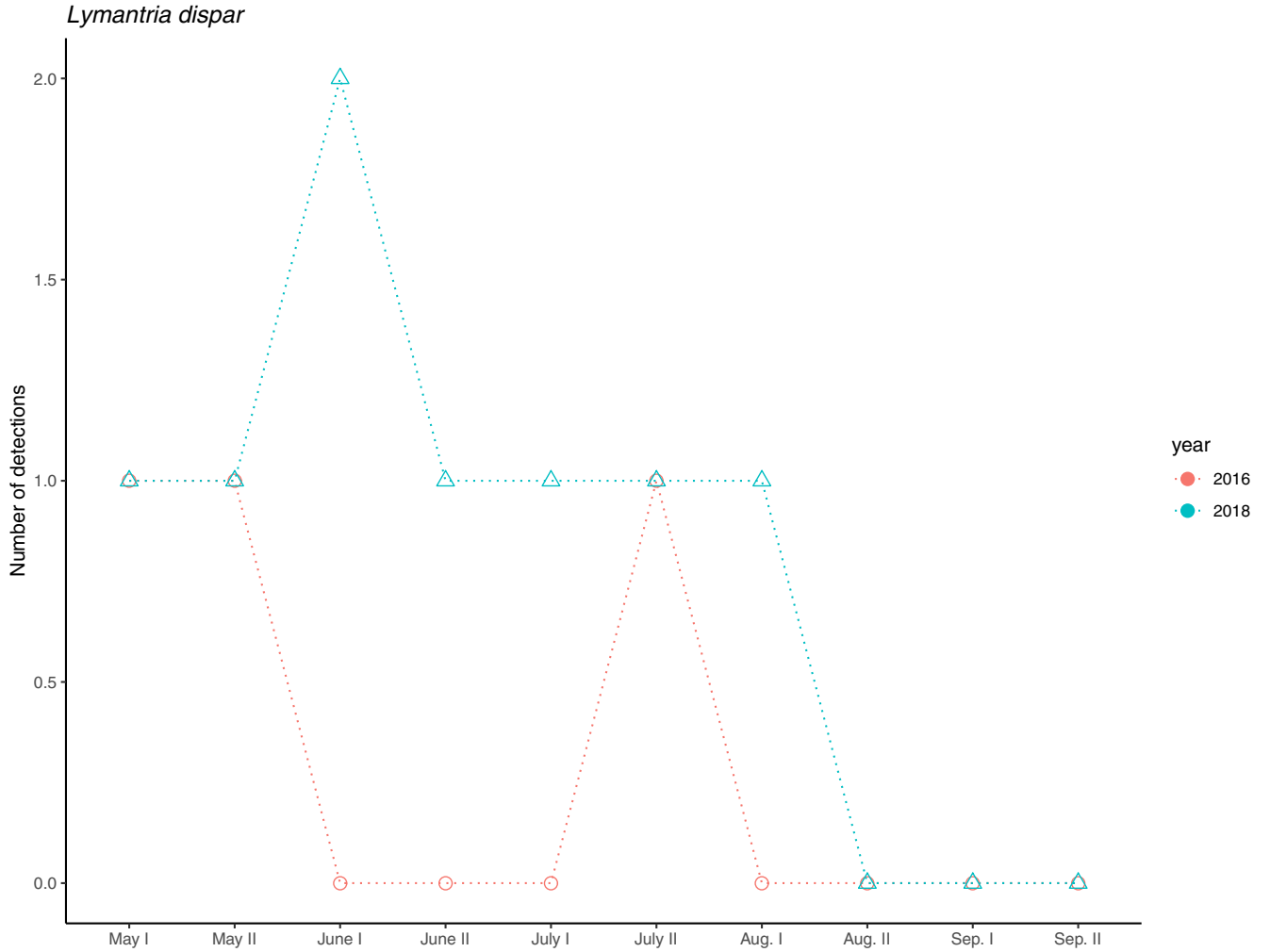

FIGURE 6 Patterns of detections of *Lymantria dispar* in 2016 and 2018 [Colour figure can be viewed at wileyonlinelibrary.com]

voucher-based DNA barcoding survey in the Bavarian Forest National Park during 2012 (see Geiger, Moriniere, et al., 2016), we have examined patterns in biodiversity over time. Comparison with the DNA barcoding reference library offers an interesting opportunity to compare local ecosystems with digitized voucher animals over a longer period.

For survey years 2016 and 2018, as well as from the GMTP data, yearly trends in BIN detection overall, as well as on a per-site basis, followed a similar pattern, peaking in June or July, and gradually declining again throughout the remainder of the growing season. Although the samples in the GMTP were screened by morphotype species and DNA barcoded individually, BIN detection for major insect orders was similar to that of both years of the present study (Figure 2). In particular among the dipteran families Cecidomyiidae and Chironomidae, and in the hymenopteran families Braconidae and Ichneumonidae, a BOLD BLAST of our metabarcoding sequences yielded many matches to sequences which had been uploaded to BOLD from voucher specimens collected as part of the GMTP at the very same sites within the Bavarian Forest National Park utilized in the present study. This observation provides support for the exactness and efficacy of metabarcoding for the re-detection of local species.

Detection frequencies of species of interest could also be examined. Same-time detection of host and parasite species was observed, in *Epinotia tedella* and *Lissonota dubia*, in both study years (Figures 6 and 7). These results provide support for the use of metabarcoding as a reliable method for informing phenologies of individual species. It is noteworthy, too, that detection patterns of *Lymantria dispar*, a known pest, potentially suggest an increase in its abundance throughout the National Park. Efforts to track the spread of pest and invasive arthropods should be continued, and metabarcoding represents a viable time- and cost-efficient method of their early detection. We think that implementation of biodiversity data from various sources—such as bulk data on BOLD—will be valuable for ongoing monitoring efforts. Spatial biodiversity analysis revealed a strong correlation of similarity indices of collection sites between the two study years based on presence–absence data of BINs. Furthermore, principal component analysis revealed clustering patterns of abundance-based biodiversities by collection site in each year; and ANOSIM tests showed significant differences in BIN detection between groups of traps located inside and outside of the park (Figure 5). These results provide evidence in support of multi-year repeatability of the methods.

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

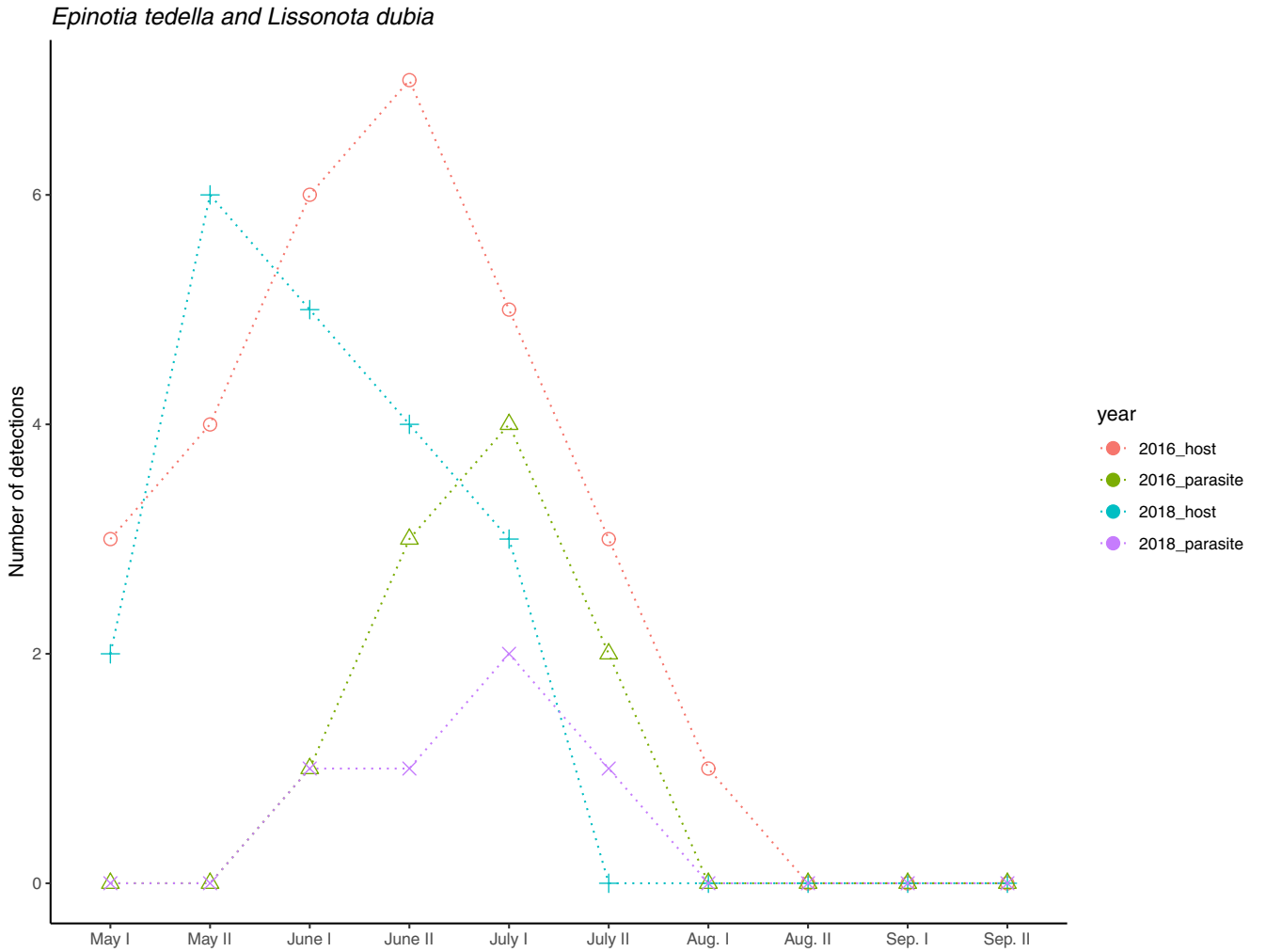

FIGURE 7 Patterns of detections of *Epinotia tedella* (host) and *Lissonota dubia* (parasite) in both survey years [Colour figure can be viewed at wileyonlinelibrary.com]

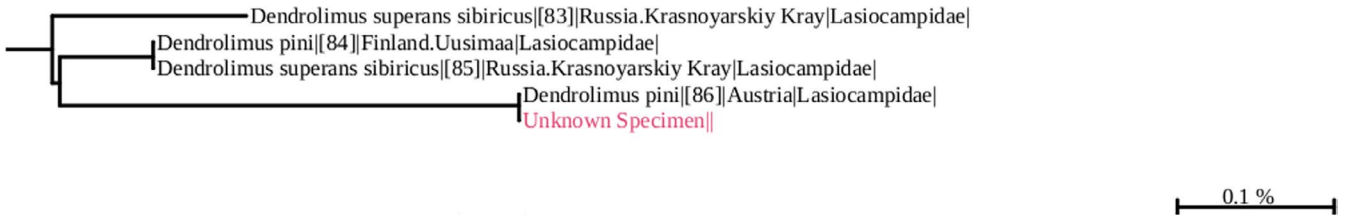

FIGURE 8 Neighbor Joining tree from BOLD shows barcode sequences of *Dendrolimus superans* and *Dendrolimus pini* specimens clustering together [Colour figure can be viewed at wileyonlinelibrary.com]

A comprehensive, well-curated reference sequence library is necessary to realize the full potential of metabarcoding. Barcode databases, most notably NCBI GenBank (Benson et al., 2017) and BOLD (Ratnasingham & Hebert, 2007), now contain millions of reference sequences, especially for the 5′ segment of the mitochondrial cytochrome *c* oxidase I (*COI*) gene (see Porter & Hajibabaei, 2018), designated as the barcode region (Hebert et al., 2003). As OTUs from metabarcoding reads are generally employed for comparison by algorithms such as BLAST, reference sequences should ideally represent intraspecific variation in all taxa. As downloading or comparing against all sources is generally impractical due to their size, the standard approach in metabarcoding is to download only taxa of interest and format them into a local database for comparison by, for example, BLAST. Studies have shown, however, that combining multiple databases provides increased taxonomic coverage and reliability of results (Macher et al., 2017).

In the present study, we have utilized species lists from the literature and publicly available gene banks to create a custom reference database for taxa of potential interest as pests or invasive species, using multiple sources of reference sequences. The

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 Selection from the OTU and BLAST result table showing invasive terrestrial arthropods from the German national institute for nature conservation warning list (2013) detected in samples

Hit description	Percentage identity	Alignment length	X2016_T1_34B_2LMai	X2016_T1_34B_2LSep	X2016_T1_52B_1LSep	X2016_T1_52B_2LAug	X2016_T1_52B_2LSep	X2018_T3_50B_1LAug	X2018_T3_50_2LAug_filter	X2018_T3_50_2LJuli_filter
Dendrolimus_superans BOLD:AAB6845	99.548	221	0	0	6	0	0	0	0	0
Ips_duplicatus BOLD:ACD5566	98.636	220	0	0	0	0	0	0	0	11

database, however, could have been even more comprehensive if *COI* reference sequences for more of the species of interest were publicly available, underlining the ongoing need for comprehensive reference libraries for DNA metabarcoding. In conjunction with the application of multiple methods of DNA extraction, this database enabled us to find a match to a warning-list species in our samples (Table 4). Of two potential matches above 97% identity to database sequences, one was a participant in BIN-sharing, clustering together with an endemic species. Therefore, *Ips duplicatus* was the only molecular identification from the warning list. The unambiguous molecular identification of the heavily invasive pest *I. duplicatus* represents a new record of this pest in the Bavarian Forest National Park. Bark beetles of the genus *Ips* are of interest to biologists for the roles they play in the decomposition of pine and spruce trees in forest ecosystems. Although this species was detected in only one sample with one extraction method (filtered ethanol) with low read numbers (11), it nevertheless remained in the OTU table after applying our cleaning steps; and although the possibility of a false positive (e.g. from contamination) cannot be definitively excluded, it may have been a result of traces of this species in the environment, especially in light of its invasive patterns observed recently (Fiala & Holuša, 2019). One possibility is regurgitated gut contents from a predator species in the trap (see Zizka et al., 2019). Detection of this pest may suggest a need for follow-up monitoring with particular attention to this species. If this result is indeed an early detection of a pest species at its invasive front, it may assist in the implementation of timely measures to reduce the risk of damage to the ecosystem. Additionally, the fact that this species was detected exclusively by ethanol filtration provides further support for our recommendation of the use of multiple methods of DNA extraction in conjunction for metabarcoding efforts, whenever possible.

With the rapidly growing demand for large-scale biodiversity data, metabarcoding has gained popularity as the method of choice for any major biomonitoring initiative. Our study shows that the method qualifies as a cost- and time-efficient alternative to traditional approaches. However, despite its apparent advantages, more research is needed to overcome its current limitations in both the laboratory and informatic areas. We encourage further studies towards this aim, to investigate patterns of biodiversity across all varieties and scales of ecosystems and environments, in order to increase the ability of scientists to effectively manage resources and conserve the biodiversity upon which life on Earth depends.

ACKNOWLEDGEMENTS

We express our sincere gratitude to Olaf Schubert for administrating the collection of the specimens from the Malaise traps in the field. Without his dedicated work this study would not have been possible. We would also like to thank Dr Marina Querejeta Coma for her diligent work on the laboratory portion of this study, as well as Niklas Franzen and Jonas Isemann for their assistance in support of the laboratory work. We are grateful to the team at the Centre for Biodiversity Genomics in Guelph (Ontario, Canada) for their support and help, and particularly to Sujeevan Ratnasingham for developing

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the BOLD database (BOLD; www.boldsystems.org) infrastructure and the BIN management tools.

AUTHOR CONTRIBUTIONS

Obtained funding: G.H., A.H., J.Mü., P.D.N.H. Designed the research: G.H., J.Mo., L.A.H., J.Mü. Analysed the data: L.A.H., J.Mü. Wrote the paper: L.A.H., J.Mo. Contributed additions/corrections to the manuscript: P.D.N.H., L.H., G.H., A.H, S.S., D.D.

ETHICAL APPROVAL

Specimens were collected by Malaise traps, which were deployed in 2016 and 2018 in the National Park Bavarian Forest. Fieldwork permits were issued by the responsible state environmental offices of Bavaria (*Bayerisches Staatsministerium für Umwelt und Gesundheit, Munich, Germany, project: "Barcoding Fauna Bavarica"; confirmed by the regional governments "Bezirksregierungen"*).

DATA ACCESSIBILITY

Table generated from OTU table and BLAST results: Dryad https://doi.org/10.5061/dryad.xd2547dcb (Hardulak et al., 2019). BOLD data set "DS-BWPST Database of Pest Species of Insects in Germany" https://doi.org/10.5883/DS-BWPST. BOLD data set "DS-BFNWARN Bundesamt für Naturschutz Warnliste, Arthropoden" https://doi.org/10.5883/DS-BFNWARN.

ORCID

Laura A. Hardulak https://orcid.org/0000-0002-4821-7343
Jérôme Morinière https://orcid.org/0000-0001-9167-6409
Stefan Schmidt https://orcid.org/0000-0001-5751-8706
Jörg Müller https://orcid.org/0000-0002-1409-1586
Paul D. N. Hebert https://orcid.org/0000-0002-3081-6700

REFERENCES

Andersen, K. S. S., Kirkegaard, R. H., Karst, S. M., & Albertsen, M. (2018). ampvis2: an R package to analyse and visualise 16S rRNA amplicon data. *BioRxiv*, 299537.

Armstrong, K. F., & Ball, S. L. (2005). DNA barcodes for biosecurity: Invasive species identification. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1462), 1813–1823. https://doi.org/10.1098/rstb.2005.1713

Astrin, J. J., Höfer, H., Spelda, J., Holstein, J., Bayer, S., Hendrich, L., … Muster, C. (2016). Towards a DNA barcode reference database for spiders and harvestmen of Germany. *PLoS One*, *11*(9), e0162624 (24 pp plus supplements).

Aylagas, E., Borja, Á., & Rodríguez-Ezpeleta, N. (2014). Environmental status assessment using DNA metabarcoding: Towards a genetics based marine biotic index (gAMBI). *PLoS One*, *9*(3), e90529. https://doi.org/10.1371/journal.pone.0090529

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. https://doi.org/10.1111/j.2041-210x.2012.00224.x

Bellard, C., Cassey, P., & Blackburn, T. M. (2016). Alien species as a driver of recent extinctions. *Biology Letters*, *12*(2), 20150623. https://doi.org/10.1098/rsbl.2015.0623

Beng, K. C., Tomlinson, K. W., Shen, X. H., Surget-Groba, Y., Hughes, A. C., Corlett, R. T., & Slik, J. F. (2016). The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. *Scientific Reports*, *6*(1), 1–13. https://doi.org/10.1038/srep24965

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2017). GenBank. *Nucleic Acids Research*, *45*(D1), D37–D42. https://doi.org/10.1093/nar/gkw1070

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., … De Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in ecology & evolution*, *29*(6), 358–367.

Borrell, Y. J., Miralles, L., Do Huu, H., Mohammed-Geba, K., & Garcia-Vazquez, E. (2017). DNA in a bottle—Rapid metabarcoding survey for early alerts of invasive species in ports. *PLoS One*, *12*(9), e0183347. https://doi.org/10.1371/journal.pone.0183347

Ceballos, G., & Ehrlich, P. R. (2018). The misunderstood sixth mass extinction. *Science*, *360*(6393), 1080–1081.

Chamberlain, S. (2018). *Bold: Interface to Bold systems API*. R package version 0.8.6. Retrieved from https://CRAN.R-project.org/package=bold

Chivian, E., & Bernstein, A. (Eds.) (2008). *Sustaining life: How human health depends on biodiversity*. Oxford, UK: Oxford University Press.

Cock, M. J. W. (2003). *Biosecurity and forests: An introduction – With particular emphasis on forest pests*. FAO Forest Health and Biosecurity Working Paper FBS/2E, 2003.

Comtet, T., Sandionigi, A., Viard, F., & Casiraghi, M. (2015). DNA (meta) barcoding of biological invasions: A powerful tool to elucidate invasion processes and help managing aliens. *Biological Invasions*, *17*(3), 905–922. https://doi.org/10.1007/s10530-015-0854-y

Corlett, R. T. (2017). A bigger toolbox: Biotechnology in biodiversity conservation. *Trends in Biotechnology*, *35*(1), 55–65. https://doi.org/10.1016/j.tibtech.2016.06.009

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, *14*(6), 927–930. https://doi.org/10.1658/1100-9233(2003)014[0927:vaporf]2.0.co;2

Dray, S., & Dufour, A. B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*(4), 1–20.

Ehrenfeld, J. G. (2010). Ecosystem consequences of biological invasions. *Annual review of ecology, evolution, and systematics*, *41*, 59–80.

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). https://doi.org/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. https://doi.org/10.1002/ece3.3192

Elbrecht, V., & Steinke, D. (2019). Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. *Freshwater Biology*, *64*(2), 380–387.

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., … Brochmann, C. (2012). New environmental metabarcodes for analysing soil DNA: Potential for studying past and present ecosystems. *Molecular Ecology*, *21*(8), 1821–1833. https://doi.org/10.1111/j.1365-294X.2012.05537.x

Fiala, T., & Holuša, J. (2019). Occurrence of the invasive bark beetle *Phloeosinus aubei* on common Juniper trees in the Czech Republic. *Forests*, *10*(1), 12. https://doi.org/10.3390/f10010012

Geiger, M. F., Astrin, J. J., Borsch, T., Burkhardt, U., Grobe, P., Hand, R., … Wägele, W. (2016). How to tackle the molecular species inventory for an industrialized nation—Lessons from the first phase of the German Barcode of Life initiative GBOL (2012–2015). *Genome*, *59*(9), 661–670. https://doi.org/10.1139/gen-2015-0185

Geiger, M. F., Moriniere, J., Hausmann, A., Haszprunar, G., Wägele, W., Hebert, P. D. N., & Rulik, B. (2016). Testing the Global Malaise Trap Program – How well does the current barcode reference library

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

identify flying insects in Germany? *Biodiversity Data Journal*, *4*, e10671 (22 pp plus supplements).

Geijzendorffer, I. R., Regan, E. C., Pereira, H. M., Brotons, L., Brummitt, N., Gavish, Y., … Walters, M. (2016). Bridging the gap between biodiversity data and policy reporting needs: An essential biodiversity variables perspective. *Journal of Applied Ecology*, *53*(5), 1341–1350. https://doi.org/10.1111/1365-2664.12417

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. A., & Baird, D. J. (2011). Environmental barcoding: A next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS One*, *6*(4), e17497. https://doi.org/10.1371/journal.pone.0017497

Hajibabaei, M., Spall, J. L., 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*, 28. https://doi.org/10.1186/1472-6785-12-28

Hardulak, L. A., Morinière, J., Hausmann, A., Hendrich, L., Schmidt, S., Doczkal, D., … Haszprunar, G. (2019). DNA metabarcoding for biodiversity monitoring in a national park: screening for invasive and pest species. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.xd2547dcb

Haszprunar, G. (2009). Barcoding Fauna Bavarica–eine Chance für die Entomologie. *Nachrichtenblatt Der Bayerischen Entomologen*, *58*(1/2), 45–47.

Hausmann, A., Godfray, H. C. J., Huemer, P., Mutanen, M., Rougerie, R., van Nieukerken, E. J., … Hebert, P. D. N. (2013). Genetic patterns in European geometrid moths revealed by the Barcode Index Number (BIN) system. *PLoS One*, *8*(12), e84518 (11 pp).

Hausmann, A., Haszprunar, G., & Hebert, P. D. N. (2011). DNA barcoding the geometrid fauna of Bavaria (Lepidoptera): Successes, surprises, and questions. *PLoS One*, *6*(2), e17134 (9 pp).

Hausmann, A., Haszprunar, G., Segerer, A. H., Speidel, W., Behounek, G., & Hebert, P. D. N. (2011). Now DNA-barcoded: The butterflies and larger moths of Germany. *Spixiana*, *34*(1), 47–58.

Havemann, N., Gossner, M. M., Hendrich, L., Morinière, J., Niedringhaus, R., Schäfer, P., & Raupach, M. J. (2018). From water striders to water bugs: the molecular diversity of aquatic Heteroptera (Gerromorpha, Nepomorpha) of Germany based on DNA barcodes. *PeerJ*, *6*, e4577 (30 pp).

Hawlitschek, O., Morinière, J., Lehmann, G. U. C., Lehmann, A. W., Kropf, M., Dunz, A., … Haszprunar, G. (2017). DNA barcoding of crickets, katydids and grasshoppers (Orthoptera) from Central Europe with focus on Austria. *Germany and Switzerland*. *Molecular Ecology Resources*, *17*(5), 1037–1053. https://doi.org/10.1111/1755-0998.12638

Hebert, P. D. N., Ratnasingham, S., & de Waard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(Suppl_1), S96–S99.

Hendrich, L., Morinière, J., Haszprunar, G., Hebert, P. D., Hausmann, A., Köhler, F., & Balke, M. (2015). A comprehensive DNA barcode database for Central European beetles with a focus on Germany: Adding more than 3500 identified species to BOLD. *Molecular Ecology Resources*, *15*(4), 795–818. https://doi.org/10.1111/1755-0998.12354

Hijmans, R. J., Williams, E., Vennes, C., & Hijmans, M. R. J. (2017). *Package 'geosphere'*. Spherical Trigonometry. R package version 1.5-7, 2017.

Honrado, J. P., Pereira, H. M., & Guisan, A. (2016). Fostering integration between biodiversity monitoring and modelling. *Journal of Applied Ecology*, *53*(5), 1299–1304. https://doi.org/10.1111/1365-2664.12777

IUCN. (2019). *The IUCN red list of threatened species*. Version 2019–1. Retrieved from http://www.iucnredlist.org

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245–1257. https://doi.org/10.1111/ele.12162

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647–1649.

Keller, R. P., Geist, J., Jeschke, J. M., & Kühn, I. (2011). Invasive species in Europe: Ecology, status, and policy. *Environmental Sciences Europe*, *23*(1), 23 (17 pp).

Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports*, *7*(1), 17668. https://doi.org/10.1038/s41598-017-17333-x

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), 2076–2081. https://doi.org/10.1073/pnas.1424997112

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., … Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, *10*, 34. https://doi.org/10.1186/1742-9994-10-34

Lindenmayer, D. B., Gibbons, P., Bourke, M., Burgman, M., Dickman, C. R., Ferrier, S., … Zerger, A. (2012). Improving biodiversity monitoring. *Austral Ecology*, *37*(3), 285–294. https://doi.org/10.1111/j.1442-9993.2011.02314.x

Macher, J.-N., Macher, T.-H., & Leese, F. (2017). Combining NCBI and BOLD databases for OTU assignment in metabarcoding and metagenomic datasets: The BOLD_NCBI _Merger. *Metabarcoding and Metagenomics*, *1*, e22262. https://doi.org/10.3897/mbmg.1.22262

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

Meyerson, L. A., & Reaser, J. K. (2002). A unified definition of biosecurity. *Science*, *295*(5552), 44. https://doi.org/10.1126/science.295.5552.44a

Morinière, J., Balke, M., Doczkal, D., Geiger, M. F., Hardulak, L. A., Haszprunar, G., … Hebert, P. D. N. (2019). A DNA barcode library for 5,200 German flies and midges (Insecta: Diptera) and its implications for metabarcoding-based biomonitoring. *Molecular Ecology Resources*, *9*(4), 900–928. https://doi.org/10.1111/1755-0998.13022

Morinière, J., Cancian de Araujo, B., Lam, A. W., Hausmann, A., Balke, M., Schmidt, S., … Haszprunar, G. (2016). Species identification in Malaise Trap samples by DNA barcoding based on NGS technologies and a scoring matrix. *PLoS One*, *11*(5), e0155497. https://doi.org/10.1371/journal.pone.0155497

Morinière, J., Hendrich, L., Balke, M., Beermann, A. J., König, T., Hess, M., … Haszprunar, G. (2017). A DNA barcode library for Germany's mayflies, stoneflies and caddisflies (Ephemeroptera, Plecoptera and Trichoptera). *Molecular Ecology Resources*, *17*(6), 1293–1307. https://doi.org/10.1111/1755-0998.12683

Morinière, J., Hendrich, L., Hausmann, A., Hebert, P., Haszprunar, G., & Gruppe, A. (2014). Barcoding Fauna Bavarica: 78% of the Neuropterida fauna barcoded! *PLoS One*, *9*(10), e109719 (8 pp plus supplements).

Müller, J., Bußler, H., Goßner, M., Rettelbach, T., & Duelli, P. (2008). The European spruce bark beetle *Ips typographus* in a national park: From pest to keystone species. *Biodiversity and Conservation*, *17*(12), 2979–3001. https://doi.org/10.1007/s10531-008-9409-1

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O'hara, R. B., Simpson, G. L., … Wagner, H. (2010). *Vegan: community ecology package*. R package version 1.17-4. Retrieved from http://CRAN.R-project.org/package=vegan

Piñol, J., Senar, M. A., & Symondson, W. O. C. (2019). The choice of universal primers and the characteristics of the species mixture

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

determine when DNA metabarcoding can be quantitative. *Molecular Ecology*, *28*(2), 407–419.

Porter, T. M., & Hajibabaei, M. (2018). Scaling up: A guide to high-throughput genomic approaches for biodiversity analysis. *Molecular Ecology*, *27*(2), 313–338. https://doi.org/10.1111/mec.14478

Pschorn-Walcher, H., & Schwenke, W. (1982). *Die Forstschädlinge Europas. 3 Band – Schmetterlinge*. Hamburg, Germany: Parey.

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rabitsch, W., Gollasch, S., Isermann, M., Starfinger, U., & Nehring, S. (2013). "Erstellung einer Warnliste in Deutschland noch nicht vorkommender invasiver Tiere und Pflanzen" Bundesamt für Naturschutz. *BfN-Skripten*, *331*, 126–131.

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes*, *7*(3), 355–364.

Raupach, M. J., Hannig, K., Morinière, J., & Hendrich, L. (2016). A DNA barcode library for ground beetles (Insecta, Coleoptera, Carabidae) of Germany: The genus *Bembidion* Latreille, 1802 and allied taxa. *ZooKeys*, *592*, 121–141. https://doi.org/10.3897/zookeys.592.8316

Raupach, M. J., Hannig, K., Morinière, J., & Hendrich, L. (2018). A DNA barcode library for ground beetles of Germany: The genus Amara Bonelli, 1810 (Insecta, Coleoptera, Carabidae). *ZooKeys*, *759*, 57–80. https://doi.org/10.3897/zookeys.759.24129

Raupach, M. J., Hendrich, L., Küchler, S. M., Deister, F., Morinière, J., & Gossner, M. M. (2014). Building-up of a DNA barcode library for true bugs (Insecta: Hemiptera: Heteroptera) of Germany reveals taxonomic uncertainties and surprises. *PLoS One*, *9*(9), e106940 (13 pp).

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

Rulik, B., Eberle, J., von der Mark, L., Thormann, J., Jung, M., Köhler, F., … Fritzlar, F. (2017). Using taxonomic consistency with semiautomated data preprocessing for high quality DNA barcodes. *Methods in Ecology and Evolution*, *8*(12), 1878–1887.

Sala, O. E., Chapin, F. S., Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., … Leemans, R. (2000). Global biodiversity scenarios for the year 2100. *Science*, *287*(5459), 1770–1774.

Schmeller, D. S., Julliard, R., Bellingham, P. J., Böhm, M., Brummitt, N., Chiarucci, A., … Belnap, J. (2015). Towards a global terrestrial species monitoring program. *Journal for Nature Conservation*, *25*, 51–57. https://doi.org/10.1016/j.jnc.2015.03.003

Schmid-Egger, C., Straka, J., Ljubomirov, T., Blagoev, G. A., Morinière, J., & Schmidt, S. (2019). DNA barcodes identify 99 per cent of apoid wasp species (Hymenoptera: Ampulicidae, Crabronidae, Sphecidae) from the Western Palearctic. *Molecular Ecology Resources*, *19*(2), 476–484. https://doi.org/10.1111/1755-0998.12963

Schmidt, S., Schmid-Egger, C., Morinière, J., Haszprunar, G., & Hebert, P. D. N. (2015). DNA barcoding largely supports 250 years of classical taxonomy: Identifications for Central European bees (Hymenoptera, Apoidea partim). *Molecular Ecology Resources*, *15*(4), 985–1000.

Schmidt, S., Taeger, A., Morinière, J., Liston, A., Blank, S. M., Kramp, K., … Stahlhut, J. (2017). Identification of sawflies and horntails (Hymenoptera, 'Symphyta') through DNA barcodes: Successes and caveats. *Molecular Ecology Resources*, *17*(4), 670–685. https://doi.org/10.1111/1755-0998.12614

Shokralla, S., Spall, J. L., Gibson, J. F., & Hajibabaei, M. (2012). Next-generation sequencing technologies for environmental DNA research. *Molecular Ecology*, *21*(8), 1794–1805. https://doi.org/10.1111/j.1365-294X.2012.05538.x

Silva-Santos, R., Ramirez, J. L., Galetti, P. M. Jr, & Freitas, P. D. (2018). Molecular evidences of a hidden complex scenario in *Leporinus* cf. *friderici*. *Frontiers in Genetics*, *9*, 47. https://doi.org/10.3389/fgene.2018.00047

Spelda, J., Reip, H. S., Oliveira Biener, U., & Melzer, R. R. (2011). Barcoding Fauna Bavarica: Myriapoda – a contribution to DNA sequence-based identifications of centipedes and millipedes (Chilopoda, Diplopoda). *ZooKeys*, *115*, 123–139. https://doi.org/10.3897/zookeys.156.2176

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, *183*, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019

Wickham, H. (2016). *ggplot2*. Cham, Switzerland: Springer.

Williamson, M. (1996). *Biological invasions*. London, UK: Chapman & Hall.

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/British Ecological Society*, *3*(4), 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x

Zhang, B. C. H. (1994). *Index of economically important Lepidoptera*. Wallingford, UK: CAB International.

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

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Hardulak LA, Morinière J, Hausmann A, et al. DNA metabarcoding for biodiversity monitoring in a national park: Screening for invasive and pest species. *Mol Ecol Resour*. 2020;20:1542–1557. https://doi.org/10.1111/1755-0998.13212

17550998, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13212 by Cochrane Germany, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License