Illuminating arthropod diversity in a tropical forest: Assessing biodiversity by automatic light trapping and DNA metabarcoding

by Souto- Vilarós, D., Basset, Y., Blažek, P., Laird-Hopkins, B., Segar, S.T., Navarro-Valencia, E., Zamora, A.C., Campusano, Y., Čtvrtečka, R., Savage, A.F., Perez, F., Lopez, Y., Bobadilla, R., Ramírez Silva, J.A. and Lamarre, G.P.A.

Souto- Vilarós, D., Basset, Y., Blažek, P., Laird-Hopkins, B., Segar, S.T., Navarro-Valencia, E., Zamora, A.C., Campusano, Y., Čtvrtečka, R., Savage, A.F., Perez, F., Lopez, Y., Bobadilla, R., Ramírez Silva, J.A. and Lamarre, G.P.A. (2024) 'Illuminating arthropod diversity in a tropical forest: Assessing biodiversity by automatic light trapping and DNA metabarcoding'. *Environmental DNA*, 6(2), Article number 540.

Received: 13 June 2023 | Revised: 26 February 2024 | Accepted: 15 March 2024

DOI: 10.1002/edn3.540

Environmental DNA
Open Access
Dedicated to the study and use of environmental DNA for basic and applied sciences

WILEY

ORIGINAL ARTICLE

Illuminating arthropod diversity in a tropical forest: Assessing biodiversity by automatic light trapping and DNA metabarcoding

Daniel Souto-Vilarós[1,2] | Yves Basset[1,3,4,5] | Petr Blažek[1,4] | Benita Laird-Hopkins[1,4] | Simon T. Segar[6] | Eduardo Navarro-Valencia[3,7] | Ana Cecilia Zamora[3] | Yahir Campusano[3] | Richard Čtvrtečka[1] | Amanda F. Savage[3,8] | Filonila Perez[3] | Yacksecari Lopez[3] | Ricardo Bobadilla[3] | José Alejandro Ramírez Silva[3] | Greg P. A. Lamarre[1,3]

[1]Institute of Entomology, Biology Centre, Czech Academy of Sciences, Ceske Budejovice, Czechia

[2]Department of Biology, Utah State University, Logan, Utah, USA

[3]Smithsonian Tropical Research Institute, Balboa, Ancon, Panama

[4]Faculty of Science, University of South Bohemia, Ceske Budejovice, Czechia

[5]Maestria de Entomologia, Universidad de Panamá, Panama City, Panama

[6]Department of Agriculture and Environment, Harper Adams University, Newport, UK

[7]School of Biology, University of Panama, Ciudad Universitaria, Estafeta Universitaria, Panama City, Panama

[8]Department of Biological Sciences, Boise State University, Boise, Idaho, USA

Correspondence
Daniel Souto-Vilarós and Greg P. A. Lamarre, Institute of Entomology, Biology Centre, Czech Academy of Sciences, Branisovska 31, CZ – 37005 Ceske Budejovice, Czechia.
Email: daniel.souto.v@gmail.com and greglamarre973@gmail.com

Funding information
Grantová Agentura České Republiky, Grant/Award Number: 19-15645Y and 20-31295S

Abstract

Although studies of insect decline have recently dominated headlines worldwide, their interpretation requires caution since for most species, we lack long-term population baselines. In the tropics, where most insect species thrive, our knowledge is even more limited and so reliable insect assessments must originate from well-established long-term monitoring efforts. Combining the extensive monitoring data from the Arthropod Program of the Smithsonian Tropical Research Institute (STRI) on Barro Colorado Island (BCI), Panama, we compare whether known arthropod diversity can be detected through metabarcoding of bulk insect samples obtained through automatic light-trapping. Our study detected 4402 species based on Barcode Index Numbers (BIN) and detected fine-scale differences between wet and dry seasons and sampling localities. We further refined our analysis to indicate which families and genera explained seasonal turnover. Using samples collected in parallel, but sorted manually as part of the ongoing arthropod monitoring program, we compared these methods. Out of 538 BINs recovered through manual sorting, there was a 70% overlap with the metabarcoding data; however, it represented 30% of the total BINs detected through metabarcoding. Expecting higher detection through metabarcoding, we also compare the results with the 14 years of sampling in BCI to better understand how well the monitoring program has captured the diversity of focal groups. Our results revealed a ~50% overlap between both methods and similar total catch. Barcode Index Numbers manually detected but not recovered by metabarcoding highlight some of the limitations of molecular detection methods such as primer bias. Contrastingly, BINs detected with metabarcoding, but not recovered by the traditional monitoring scheme, highlight the importance of local and regional barcode reference libraries.

KEYWORDS
arthropod monitoring, Barcode Index Number, Barro Colorado Island, biodiversity, light-trapping, metabarcoding

1 | INTRODUCTION

With an estimated ~6 million species worldwide, terrestrial arthropods represent the majority of eukaryote diversity on earth (Hamilton et al., 2010; Stork et al., 2015). Given their provision of essential ecosystem services, and their intimate association with host plants, soil formation, trophic interactions, and functional diversity, arthropod populations are inextricably linked to ecosystem functioning and stability (Weisser & Siemann, 2008). With thousands of arthropod species awaiting description, including complexes of often cryptic species, the taxonomic impediment remains a significant challenge (Engel et al., 2021). The estimated cost of describing all species is in excess of US$260 billion and would take centuries to complete (Carbayo & Marques, 2011). As Wagner et al. (2021) suggest (with some understatement), nature is under siege as the planet enters its sixth mass extinction event. Understandably, apocalyptic reports on the decline of insect abundance have gained worldwide attention (Lister & Garcia, 2018; Sánchez-Bayo & Wyckhuys, 2019). Nevertheless, insect decline needs to be interpreted with caution, mainly because for most arthropod species, we do not have long-term population baselines (Didham et al., 2020; Wagner et al., 2021). Additionally, the few long-term monitoring programs that exist target a number of indicator species or focal groups, given the need for specialized taxonomic expertise for most insect groups (Hallmann et al., 2020, 2021; Ji et al., 2013; Yu et al., 2012). In the tropics, where the majority of arthropod species and biomass occurs (Basset et al., 2012), the knowledge gap is even wider and particularly alarming since tropical arthropods may face greater threats from habitat loss and the impacts of climate change (Deutsch et al., 2008; Harvey et al., 2023). Coincidentally, the extremely high diversity of species in tropical regions is coupled with a lack of specialized taxonomists (Engel et al., 2021; Paknia et al., 2015) and regular funding (Basset & Lamarre, 2019; Donkersley et al., 2022), placing further constraints on comprehensive monitoring schemes in the region.

Born from the accelerated need for monitoring species communities, molecular advances have allowed for an integrative approach for identifying molecular operational taxonomic units (OTUs) based on DNA barcodes (Ratnasingham & Hebert, 2013). These barcodes are standard gene regions that enable species discrimination based on sequence variation; in the field of microbiology, a 3% sequence variation of the 16S rDNA region delineates distinct bacterial linages while fungal studies use a 2% divergence of the ITS spacer region to distinguish between species (Kauserud et al., 2008; Stackebrandt & Goebl, 1994). For animals, there are more than two million available *cytochrome c oxidase I* (COI) sequences which rarely exceed a 2% variation within species making this a useful barcode for delimiting species with incomplete taxonomy (Hebert et al., 2003). These barcodes alongside collaborative repositories allow us to generate highly curated reference databases, which include not only the DNA sequence but also sample metadata such as geographic, morphological, and taxonomic information with the final goal of building a barcode library for all eukaryotic life (Porter & Hajibabaei, 2020; Ratnasingham & Hebert, 2007). Coupling the use of short and informative genetic markers with High Throughput Sequencing technology, known as DNA metabarcoding (hereafter metabarcoding), researchers can potentially generate community-level biodiversity studies for highly diverse regions in which traditional monitoring protocols are prohibitively time-consuming, costly, or even impossible.

The field of metabarcoding is gaining traction at a pivotal time when species discovery and monitoring need to outpace the most serious effects of anthropogenic stressors. Regional and global metabarcoding-based studies are currently widely applied. Metabarcoding of pollen loads from pollinators has allowed us to identify the plant species they visit (Bell et al., 2017; Tommasi et al., 2021), as well as to identify terrestrial arthropods from bulk soil samples (Basset et al., 2022; Clarke et al., 2021; Kirse et al., 2021). Further, some applications yield trophic information (Šigut et al., 2017; Toju & Baba, 2018), reveal migratory patterns (Suchan et al., 2019), and even facilitate arthropod monitoring in different tree species based on DNA from rainwater percolating from them (Macher et al., 2023). Nevertheless, how to confidently link a barcode sequence to a species remains one of the biggest challenges in metabarcoding studies (Keck et al., 2023; Porter & Hajibabaei, 2020). Depending on the desired level of taxonomic resolution, a barcode reference database is the most important aspect because it determines the accuracy of species detection (Magoga et al., 2022; Steinke et al., 2022). Although multiple tools and data processing pipelines exist (Liu et al., 2020), the mBRAVE platform relies on the Barcode Index Number (BIN) system, an indexed OTU equivalent generated for the barcode sequences of the Barcode of Life (BOLD) systems database where 99.7% of its 1.81 million records have an associated voucher specimen and required metadata (Ratnasingham & Hebert, 2013).

The Arthropod Program of the Smithsonian Tropical Research Institute (STRI), active within the permanent forest dynamic plot of Barro Colorado Island in Panama (Anderson-Teixeira et al., 2015; Lamarre et al., 2020), is one of the few ongoing arthropod monitoring programs in the tropics. To date, the program has recorded more than half a million specimens representing over 2300 species (Table S1) with 14 years of continuous data including seasonal replicates (Lamarre et al., 2020). Additionally, these data are publicly available in the BOLD database having generated over 2850 arthropod BINs from 11,171 DNA sequences for the island.

The long-term goal of our study is to incorporate routine metabarcoding of bulk arthropod samples into the Arthropod Initiative monitoring program. To achieve this goal, the present study assesses how well the Arthropod Program has captured the arthropod diversity of Barro Colorado. Specifically, we (1) compare, for a limited number of light trap-nights (see Materials And Methods), the diversity of insects detected through metabarcoding, to that identified via manual sorting. Expecting a higher number of species detected through metabarcoding, we also (2) compare the overlap between species identified through metabarcoding and the 14-year-long data of the Arthropod Program. These comparisons allow us to assess how our metabarcoding sampling effort captures known insect diversity. It also enables us to identify underrepresented groups in both surveying methods, highlighting the importance of local and regional barcode libraries for biodiversity monitoring. Finally, we (3) evaluate seasonal and site variation of insect communities detected through metabarcoding, since detecting fine-scale spatiotemporal patterns of communities is a prerequisite for any successful long-term monitoring scheme.

2 | MATERIALS AND METHODS

2.1 | Study site

Samples were collected on Barro Colorado Island (BCI, 9.159.15° N, 79.85° W; 120–160 m above sea level) in Panama. BCI is a 1542-hectare lowland tropical forest reserve created ca.1910 during the flooding of the Chagres River to fill the Panama Canal. Sampling was carried out within and near the 50 ha ForestGEO vegetation dynamics plot described in Anderson-Teixeira et al. (2015). We sampled along the same trails used for the long-term monitoring scheme in the Arthropod Program, these have been active since 2009 and described in Basset et al. (2013).

2.2 | Arthropod sampling

Arthropod samples were collected using Robinson light traps, a standard method for bulk collections of nocturnal insects (Kitching et al., 2001). Robinson light traps are seldom used in metabarcoding studies, unlike Malaise traps or soil surveys (Geiger et al., 2016; Kirse et al., 2021), but see Ji et al. (2013). To compare the efficacy of metabarcoding, two sets of collections were conducted during the wet season in May 2019 and the dry season in March 2021. One set of samples was used for DNA metabarcoding (hereafter "metabarcoding samples") while the second set of samples was manually sorted and identified following standard Arthropod Program protocols (hereafter "parallel samples"). In both cases, insects were collected with 10 automated bucket-type Robinson traps fitted with a 10-watt automated black-light bulb (F10T9BL) running on 12-volt DC batteries, fitted with acrylic intercept panes, and an acrylic roof to protect the catch from the rain (Basset et al., 2020; Lucas et al., 2016). Traps were set at chest height along the forest understory a minimum of 300 m apart; the lights were switched on with an automatic timer at 6:00 p.m. and ran all night until 6:00 a.m. Sampling was performed at 10 different locations for two non-consecutive new moon nights for a total of 20 sampling nights for each season and for each sampling protocol, totaling 40 metabarcoding and 40 parallel samples. Traps used for the metabarcoding sampling were further modified to collect arthropods directly into 95% ethanol to preserve DNA for extraction and sequencing. All traps were thoroughly cleaned and sterilized between uses with a 10% commercial bleach (Clorox, Clorox de Centroamerica) and distilled water solution, or using bleach wipes. All instruments were rinsed with distilled water to avoid any bleach residue. Ethanol from the traps was replaced immediately after collection with a total of 400 mL of 95% ethanol and samples were stored at −20°C until further manipulation. All instrument manipulation was carried out using sterilized latex gloves to avoid any cross contamination from handling the equipment.

2.3 | Sample preparation

Metabarcoding samples collected during each season were treated differently following Basset et al. (2020). Insects collected during the wet season (May 2019) were kept whole, except those larger than 2.0 cm, in which case one leg was removed and returned to the sample bottle while the rest of the body was discarded. The sample was reduced to a total volume of 50 mL of 95% ethanol. For the dry season collections (March 2021), larger insects (>2.0 cm) were treated similarly; however, the sample was preserved in 400 mL of 95% ethanol and thoroughly homogenized using a sterilized hand blender (Better Chef model IM-848, Los Angeles, USA). The homogenized sample was then separated into four 100 mL aliquots, and after leaving the sample to settle (10–15 min), the aliquot with the denser insect content was selected for subsequent DNA extraction and sequencing. Samples were sent to the Canadian Centre for DNA Barcoding (CCDB) in Guelph, Canada. DNA extraction, purification, PCR, and sequencing followed standard protocols from the CCDB (see Steinke et al., 2022). Samples were filtered to remove all ethanol and lysed with a volume of lysis buffer based on the sample's wet weight. Four replicates from each sample were transferred to separate wells of a standard 96-well microplate for DNA extraction. Each plate also included eight positive and eight negative controls for quality control.

Parallel samples were sorted manually and identified using STRI's Arthropod Program reference collection. In cases where morphological identification suggested different morphospecies, identification was verified by DNA sequencing of the COI marker following standard Arthropod Program protocols. Barcode data was deposited in BOLD while pinned specimens were deposited in the Arthropod Program collection. Identified species were matched to their BOLD BIN based on the Arthropod Program's barcode reference library which includes all specimens sequenced

to date, even though some do not belong to the focal monitoring groups (see Table S1). Species without a BIN were not included in our analyses.

2.4 | DNA extraction, COI amplification, and sequencing

A two-stage PCR following CCDB protocols was performed to generate amplicon libraries for sequencing. For the first PCR, purified extracts were used to amplify a 462 base-pair (bp) fragment of the Cytochrome *c* Oxidize subunit I (COI) barcode region using the primer combination *AncientLepF3* (Prosser et al., 2016) and *C_LepFolR* (Hernández-Triana et al., 2014). For the second PCR, platform-specific unique molecular identifies (UMIs) were attached to all samples and pooled for single-end sequencing on an Ion Torrent S5 high-throughput sequencer (Thermo Fischer Scientific, Waltham, Massachusetts, USA) following the manufacturer's instructions.

Sequence reads were associated to their source sample using their UMIs, and were uploaded to the mBRAVE platform (Ratnasingham, 2019) under the projects MBR-CYB001 and MBR-CYB002 (for wet and dry seasons, respectively). All reads were initially filtered to a minimum length of 100 bp and a maximum of 800 bp. Only reads with a minimum quality value (QV) of 20—allowing a maximum 15% bp with low QV (<20) and 1% bp with ultra-low QV (<10)—were retained. Reads were then trimmed by 25 bp at each end to remove primers and all reads were trimmed to a maximum of 500 bp. Retained reads were classified according to BINs (Ratnasingham & Hebert, 2013) by comparing them against four mBRAVE libraries. Firstly, we used the custom-made reference dataset for arthropods from BCI (DS-BCIARTH, 2850 BINs in 11,171 sequences), followed by three system libraries: SYS-CRLINSECTA (695,769 BINs in 978,611 sequences) and SYS-CRLNONINSECTARTH (78,275 BINs in 108,260 sequences) for all insects and non-insect arthropods, and SYS-MBRAVEC (80 BINs in 2225 sequences) to screen for potential standard contaminants. BINs were assigned conservatively with an ID distance threshold of 2%. Reads that did not match any sequence in the reference libraries were clustered into OTUs based on a minimum OTU size of 5 reads per cluster, and a maximum within OTU distance of 2%. Reads were excluded from the OTU threshold of 1% if sequencing errors produced spurious haplotypes or chimeras. This parameterization is largely based on benchmarking of the classification algorithm on mBrave (Steinke et al., 2022). OTUs that did not match any known BINs at the 2% similarity threshold were excluded from further analyses. These amounted to ~419,000 reads across 482 OTUs available in the Appendix S1. Admittedly, some of these OTU clusters may correspond to species awaiting a BIN or match to an existing BIN at a lower percentage than our conservative threshold. Nevertheless, our main objective in the current study is to assess the implementation of monitoring insect diversity in BCI through metabarcoding of bulk DNA samples. As such, a reference library of known existing species is of greater importance than including every recovered OTU. Using mBRAVE we generated BIN tables for all samples, including positive and negative controls and four replicates for each sample. All read counts for BINs detected in negative controls were subtracted from non-control samples using the R package "microDecon" (McKnight et al., 2019), which relies on proportions of contaminant OTUs in the blanks rather than simply removing the raw number of reads, removing contaminant reads rather than entire OTUs. This allowed us to keep BINs with low number of reads as long as they were not found within the contamination controls. Given the methodological and analytical challenges associated with obtaining accurate species abundance from DNA-based data (Luo et al., 2023) all reads were converted to presence absence data for downstream analyses.

2.5 | Sequence data analysis

Since taxonomic identification remains a challenge, particularly in the tropics, several BINs included multiple species names given that users may upload barcode sequences and name them according to their morphological identification, or placeholder names. The BIN table was manually inspected to ensure that in cases where BINs had more than one species name or had a different spelling, we enforced Arthropod Program names (public data available at https://fgeoarthropods.si.edu/), and in cases where there was no record we selected the name with the highest number of matches to that BIN in the BOLD database. Additionally, all BINs were further inspected to remove any species and/or genera which do not occur in the region. All analyses were performed using R v.4.1.1 (R Core Team, 2022) figures were generated with "ggplot2" (Wickham, 2016). Datasets were mainly analyzed using the "MicrobiotaProcess" (MBP; Xu et al., 2022) and the "Metacoder" (Foster et al., 2017) R packages, both of which implement functions from the "Phyloseq" (McMurdie & Holmes, 2013) and "Vegan" (Oksanen, 2007) packages. All data and scripts are available on GitHub (https://github.com/DanielSoutoV/metabarcoding).

2.6 | Metabarcoding sample analyses

To evaluate seasonal differences in species diversity for metabarcoding samples, rarefaction curves, and the Shannon diversity index were generated using the *get_rarecurve* and *get_alphaindex* functions of MBP (Xu et al., 2022). To compare differences among sampling sites and days, the matrix was filtered among seasons and transformed into a Bray–Curtis dissimilarity matrix. We then performed a non-metric multidimensional scaling (NMDS) analysis using the function *metaNMDS* of Vegan (Oksanen, 2007) and plotted using the *ordihull and ordispider* functions (see Results). Non-metric multidimensional scaling analysis allows us to visualize the differences, which were statistically verified through a permutational multivariate analysis of variance (PERMANOVA), as implemented through the

adonis2 function of the same package. We ran the same analysis for the combined dataset to compare differences between seasons. To visualize the number of BINs belonging to the most prevalent orders and families during both seasons, we generated percent abundance plots using the *mp_plot_abundance* function of MBP. To refine the taxonomic resolution of these differences, data were filtered to focus on the most common orders collected in the light traps (i.e., Lepidoptera, Diptera, Coleoptera, Hemiptera, Hymenoptera, and Blattodea). The *diff_analysis* function of MBP was used to evaluate different taxa prevalence across seasons using a Kruskal–Wallis test based on linear discriminant analysis (LDA) and Wilcoxon test with the Benjamini–Hochberg correction accounting for false discovery rate (FDR; Benjamini & Hochberg, 1995). This difference was visualized using the function *ggdiffbox*. "Heat trees" built with the package "Metacoder," are hierarchical taxonomic plots that allow visual comparisons between samples. Samples were plotted showing pairwise comparisons between wet and dry season sampling for all observed BINs and further filtered by the most abundant orders as indicated above. These trees indicate differences for a specific branch based on the mean difference of BINs present at each taxonomic rank using the function *compare_groups* of Metacoder (Foster et al., 2017).

2.7 | Metabarcoding and traditional classification sample analyses

We compared the BINs detected by metabarcoding across 40 light trapping nights to the 40 parallel samples and to the 14 years of arthropod sampling on BCI (1120 light trapping nights). Focusing on the BINs produced by the Arthropod Program (Table S1), we merged these three datasets (metabarcoding, parallel, and Arthropod Program sampling), and generated Metacoder heat trees to visualize the differences in BINs detected by traditional and metabarcoding classification methods. Venn diagrams were generated to visualize the distinct and shared BINs between metabarcoding samples, parallel samples, and the long-term sampling effort on BCI using the R package "VennDiagram" (Chen & Boutros, 2011). Finally, to compare species accumulation between datasets (metabarcoding, parallel samples, and the 14 years of the Arthropod Program light trapping observations), we calculated asymptotic estimates of species richness using the R package "iNEXT" (Hsieh et al., 2016).

3 | RESULTS

3.1 | Metabarcoding samples and spatiotemporal variation

Sequencing produced a total of 14.05 and 8.8 million reads for the wet and dry season sampling, respectively. After stringent filtering and classification, over 1.4 million reads for the wet season and over 525 thousand reads for the dry season were retained and assigned to 4402 arthropod species, based on their BIN. During the wet season, we detected a total of 2974 BINs while during the dry season our analysis revealed 2464 BINs. The mean number of reads per sample after filtering was 722,336 ± 3945 (SE) for the wet season, and 29,589 ± 3385 for the dry season with each sample trap-night containing an average of 633 BINs ± 17 during the wet season, and 365 BINs ± 1 in the dry season. Accumulation curves (Figure 1a) indicated that after approximately 20,000 reads the number of new species began to level off in both seasons, suggesting that despite our conservative filtering steps, our sampling had enough coverage to capture most diversity. Observed diversity and Shannon Diversity index differed significantly between seasons (Figure 1b). The relative abundance of the ten most frequent orders can be visualized in Figure 1c. As is often observed in light traps, the most abundant order was Lepidoptera while during the wet season there was a clear increase in the abundance of Diptera and Hemiptera. Figure 1d shows the ordination analysis in which the first PCoA axis accounts for ~24% of the difference between seasons. Permutational analysis of variance (PERMANOVA) revealed significant differences between sampling sites during the wet season ($R^2 = 0.499$, $F = 1.109$, p-value = 0.009; Figure S1), but not during the dry ($R^2 = 0.457$, $F = 0.937$, p-value = 0.933; Figure S2). Contrastingly, communities were significantly different between sampling days in the dry season ($R^2 = 0.0611$, $F = 1.171$, p-value = 0.052), but not in the wet season ($R^2 = 0.0532$, $F = 1.013$, p-value = 0.401).

It is important to note that not every recovered BIN was identified to species level and in some cases, they could only be classified to order. For instance, for Lepidoptera, one of the best-studied groups, we detected 1834 BINs of which 1427 (~77%) were classified to genus level, 602 (~32%) had a species name, while 166 (~9%) BINs remained at the order level. For other less well-studied groups, the percentage of BINs assigned to species level was much lower. Out of 1496 Diptera BINs detected, only ~18% were identified to genus and only 5% of BINs had a species name. For Coleoptera, 26% out of 580 detected BINs were classified to genus level and only 13% had a species name. The situation for Hemiptera and Hymenoptera is similar with only a fraction of BINs identified to species (22% of described species out of 240 and 25% named species out of 246 BINs, respectively). Similarly, several groups were detected—although in much lower proportions—that are not particularly attracted to light traps but probably carried by larger insects such as terrestrial mites (e.g., Mesostigmata (10 BINs), Trombidiformes (8 BINs), or parasites (e.g., Strepsiptera, 1 BIN)), further strengthening the robustness of metabarcoding for detecting often unseen and/or rare organisms. For a complete list of detected BINs, refer to Appendix S1.

Focusing on the orders with the highest number of BINs in our metabarcoding samples, we can detect differences between BINs belonging to each taxonomic rank between seasons. Notably, there is an increase of Dipteran BINs in the wet season particularly those belonging to Cecidomyiidae and Phoridae, while during the dry season, there appears to be an increase of Curculionid beetles and Geometridae and Erebidae moths (Figure 2). Metacoder heat trees (Figure 2) visually represent these differences highlighting which individual branches differ between seasons.

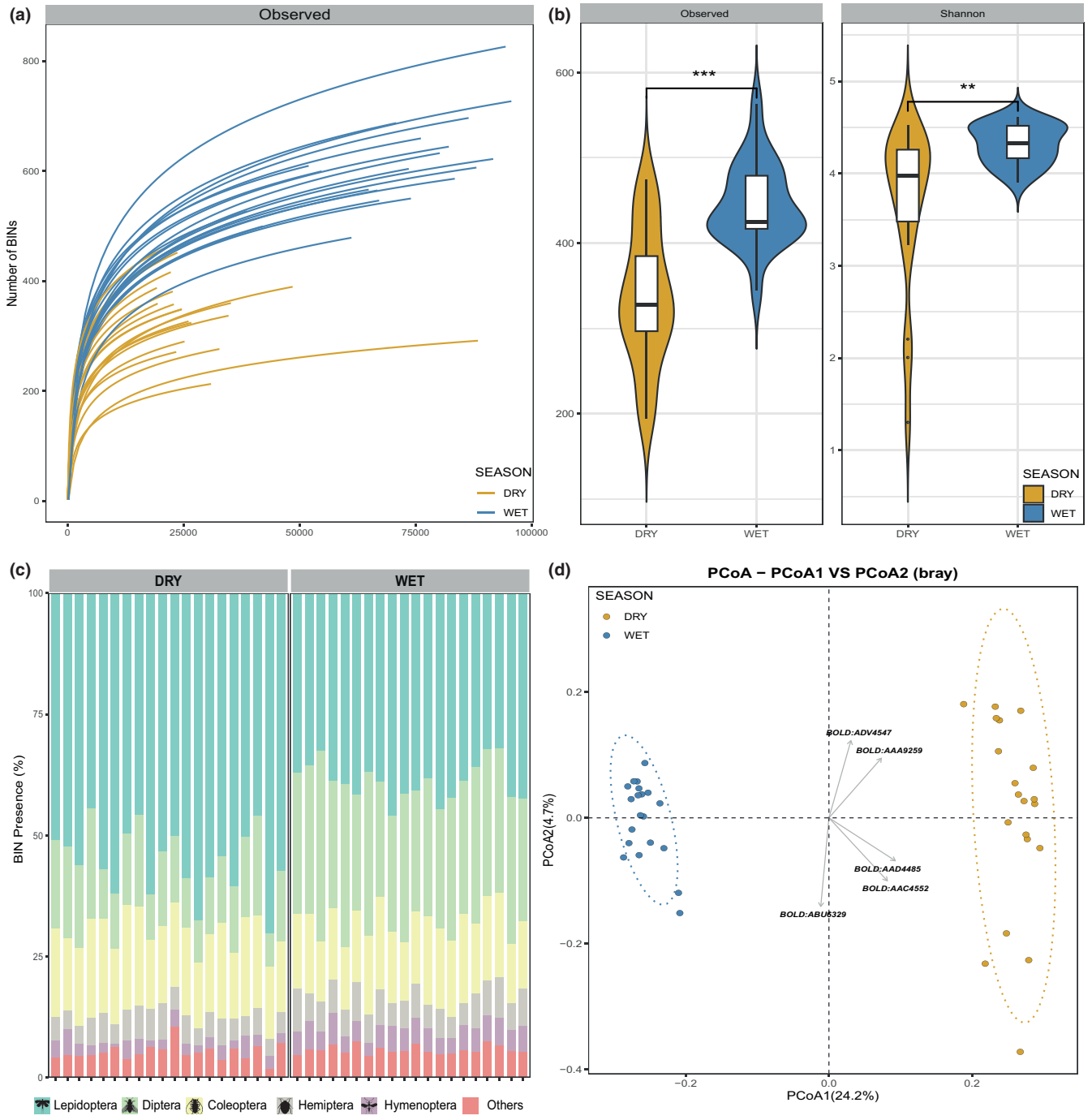

FIGURE 1 Metabarcoding samples and seasonal variation. (a) Rarefaction curves showing the number of BINs relative to total sample reads, each curve represents one light trap-night for wet (blue) and dry (orange) seasons; (b) Diversity indices—Observed diversity and Shannon index for wet (blue) and dry (orange) seasons. $^{**}p$-value < 0.05, $^{***}p$-value < 0.001; (c) BIN presence (%) of the most frequent insect orders in both seasons, each bar representing one light trap-night. Inset insect figures downloaded from phylopic.com; (d) Principal Coordinate Analysis plot based on a Bray-Curtis distance matrix for samples collected in wet (blue) and dry (orange) seasons. Arrows indicate notable BINs (see Table S2) driving the differences between seasonal surveys based on presence/absence data for each season.

3.2 | Metabarcoding and parallel samples

Filtering of the metabarcoding samples allowed us to compare both classification approaches by focusing on the Arthropod Program focal groups. Manual sorting and classification of the parallel samples recovered 538 species with an associated BIN. During the wet season, the mean number of focal species per trap was 68 ± 3 while during the dry season we recovered an average of 65 ± 2 focal species. As in the metabarcoding samples, the most abundant orders were Lepidoptera followed by Coleoptera and Hemiptera (Figure 3a). Figure 3b displays an ordination analysis showing the community (dis)similarity of arthropods between methods and seasons. As

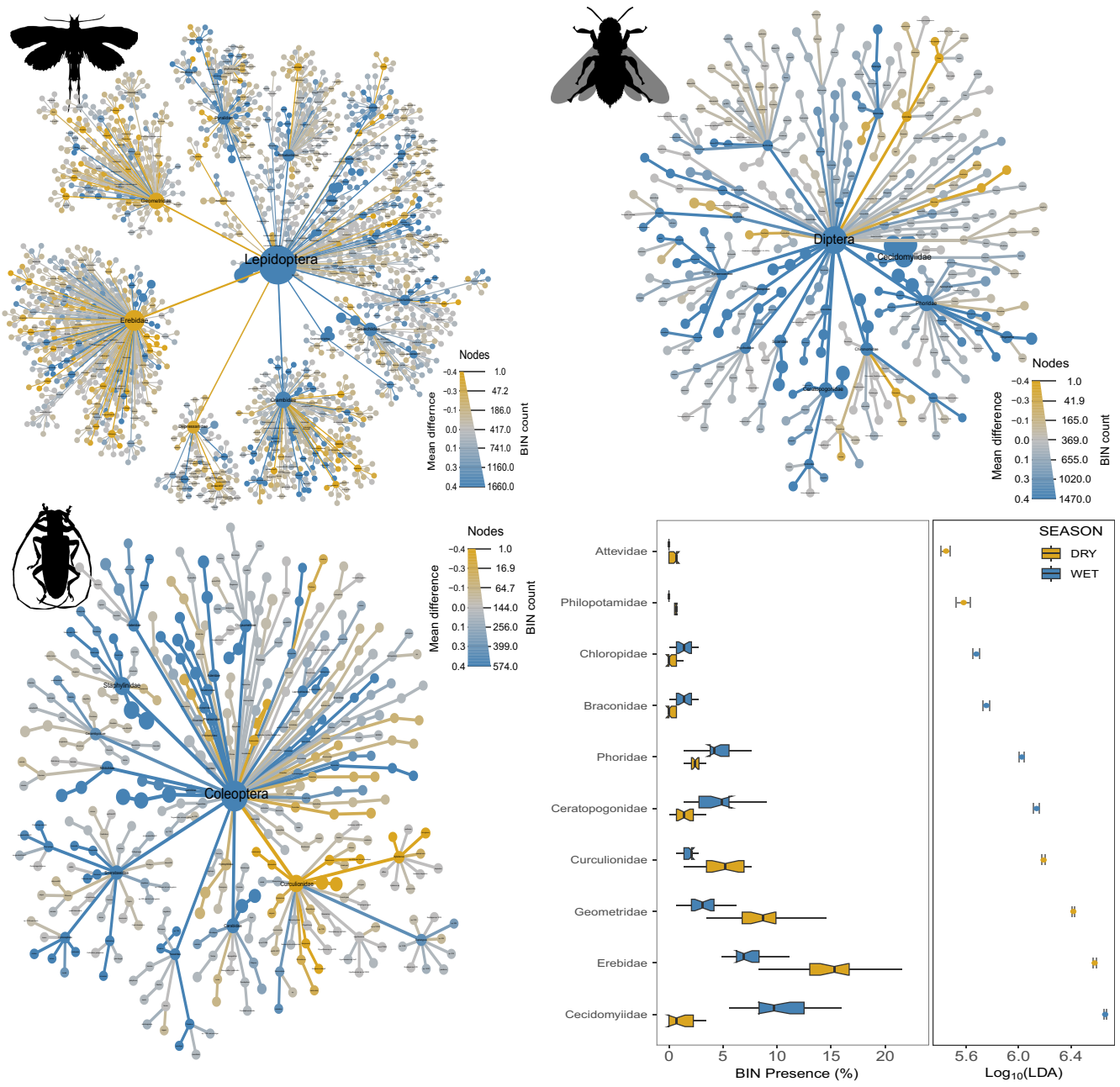

FIGURE 2 Seasonal differences in metabarcoding samples focusing on the three most abundant orders sampled via light-traps: Lepidoptera, Diptera, and Coleoptera. "Metacoder" heat trees indicate significant differences between presence/absence of BINs detected in each season. Terminal nodes of these trees represent a single BIN, while the size of internal nodes represents the number of BINs belonging to a given rank (e.g., family). Blue branches indicate more BINs during the wet season, orange branches indicate more BINs during the dry season. Gray indicates no significant difference between seasons. Bottom left panel shows the relative abundance of BINs for the main recovered families and a log-transformed Linear Discriminant Analysis (LDA) score for the most abundant families in each season in terms of BINs. Inset figures representing different orders were downloaded from phylopic.com.

expected, metabarcoding detected many more BINs, but when focusing on focal species, there was a considerable overlap between parallel and metabarcoding samples (382 BINs out of 538 ~70% overlap; Table 1). Despite the fact that manual sorting detected several focal species BINs missed by the metabarcoding approach, the sheer quantity of BINs detected, from focal and non-focal groups, is a clear advantage of metabarcoding, with this overlap representing ~30% of the total detection of focal samples only (Table 1).

3.3 | Metabarcoding and Arthropod Program monitoring data

The first axis of the PCoA comparing classification methods between all of the Arthropod Program light-trap data (1120 trap-nights) and metabarcoding data (40 trap-nights) accounted for 15.63% of the variance (Figure S4) with approximately 50% overlap of the 1263 BINs recorded through manual classification (Table 2). Excluding

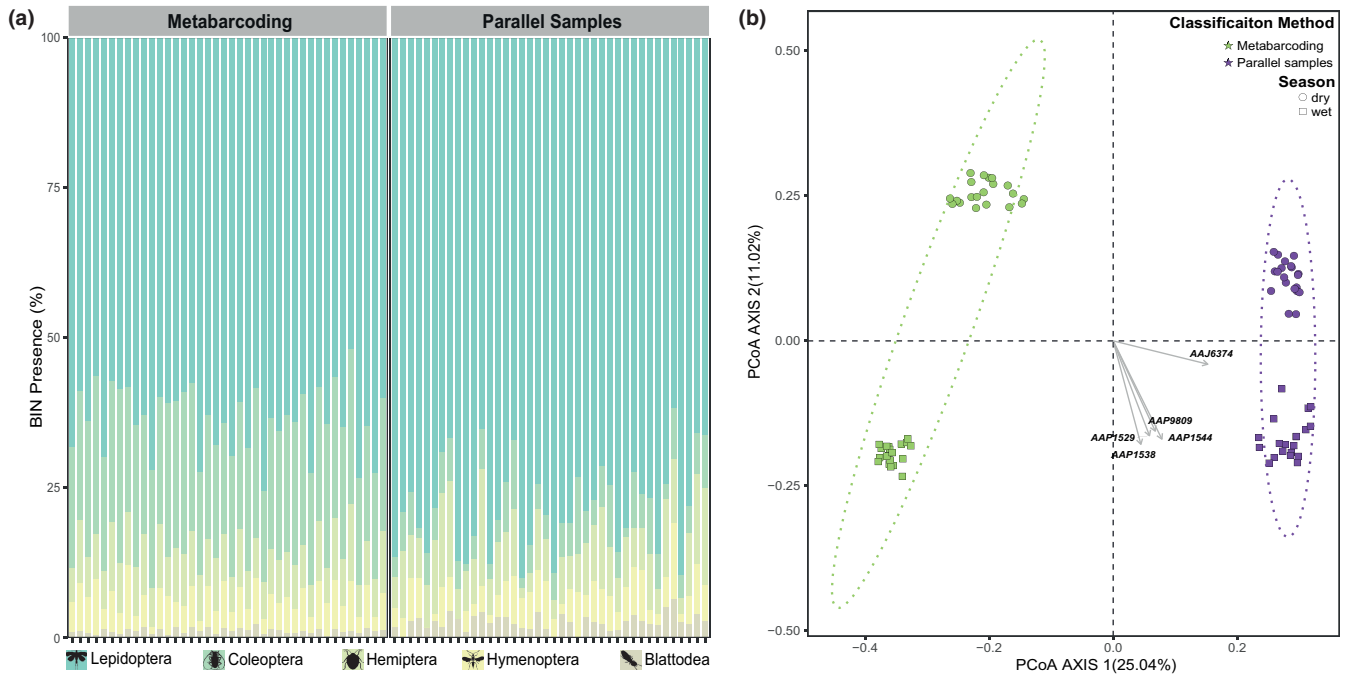

FIGURE 3 Differences between metabarcoding (green) and parallel (purple) samples. (a) Relative abundance of the most frequent orders in both classification methods. Each bar represents a single night trap. Inset insect figures downloaded from phylopic.com; (b) Principal Coordinate Analysis plot based on a Bray-Curtis distance matrix for samples classified using both methods and during both seasons. Arrows represent BOLD BINs which drive the difference between sampling methods.

non-focal groups from the analysis revealed that in most cases, even the limited number of metabarcoding trap-nights were able to detect several species not previously recorded by the traditional monitoring data (Table 2). There were notable exceptions, such as Geometridae and Crambidae, where although there was a high species overlap between both methods, the traditional approach detected more species than the metabarcoding approach (Table 2). PERMANOVA analyses reveal a clearly significant difference in community composition according to sampling methods; however, they do not reveal a significant difference between sampling sites during either season ($R^2 = 0.165$, $F = 0.878$, p-value $= 0.794$; $R^2 = 0.173$, $F = 0.930$, p-value $= 0.736$ for the wet and dry seasons respectively). NMDS ordination plots visualizing these similarities are available as Figures S5 and S6.

4 | DISCUSSION

Our study detected 4402 arthropod species from bulk collections using automated light-traps in a seasonal tropical forest. We were able to detect differences in species diversity between wet and dry seasons as well as sampling sites, and the taxonomic resolution obtained allowed us to explore which families within the most abundant orders drove the observed seasonal turnover. Metabarcoding has been successfully used to detect arthropod species richness and diversity across different soil layers (Porter et al., 2019), has demonstrated the complementarity of different sampling methods between above- and below-ground habitats (Kirse et al., 2021), and has been used to estimate differences in arthropod diversity between host trees, as detected with eDNA extracted from rainwater (Macher et al., 2023). It has been hailed as a reliable and cost-effective tool for biomonitoring diverse ecosystems (deWaard et al., 2019; Ji et al., 2013; Taberlet et al., 2012). However, classification of barcodes fundamentally relies on the reference databases used to link barcodes to taxonomic names (Keck et al., 2023) and it has been shown to be most effective in cases where the biodiversity is well known (Ji et al., 2013). Many metabarcoding studies fail to realize this perceived potential since, for the most part, they lack taxonomic resolution beyond order (Zenker et al., 2020), particularly for lesser-studied groups such as invertebrates (e.g., Diptera; Chimeno et al., 2022; Sun et al., 2019).

Taxonomic limitations continue to be a major challenge for global biodiversity assessments and monitoring. With taxonomists themselves facing the risk of extinction, these limitations will continue to undermine biodiversity estimates, particularly of invertebrates (Hochkirch et al., 2022). The Barcode of Life (Ratnasingham & Hebert, 2007) database contains over one million sequences and is unique among sequence repositories due in part to the stringent procedures required for sequence deposition. The mBrave platform is a user-friendly platform that seamlessly integrates BOLD databases and facilitates the construction of custom-made local and regional reference databases for barcode classification. Although the exclusive use of BINs may underestimate the overall diversity recovered with OTUs and their similarity to known taxa, we justify this approach on the basis that it allows us to confidently assign each recovered species to a vouchered BIN. Given our intention of implementing metabarcoding as a long-term monitoring strategy, the certainty of recovering common species, which serve as indicator

TABLE 1 BIN count for species identified through metabarcoding and through manual sorting.

Taxon	Metabarcoding	BIN overlap	Parallel samples
Lepidoptera	749	318	121
Erebidae	268	93	33
Geometridae	117	75	15
Pyralidae	86	18	2
Crambidae	118	75	18
Noctuidae	41	7	4
Notodontidae	51	26	10
Lycaenidae	6	1	2
Saturniidae	4	15	24
Hesperiidae	4	1	5
Nymphalidae	3	0	0
Others	51	7	8
Coleoptera	100	25	4
Curculionidae	74	8	0
Scarabaeidae	26	8	3
Passalidae	0	9	1
Hemiptera	11	31	14
Reduviidae	6	18	11
Flatidae	5	13	3
Hymenoptera	48	6	11
Formicidae	42	3	6
Apidae and Halictidae	6	3	5
Blattodea (termites)	21	2	6
Total	929	382	156

Note: BIN overlap indicates the number of BINs that were detected through both methods. Counts include Arthropod Program focal groups collected with light traps but not all subfamilies within these represent focal groups.

species for global change, outweighs the need to include every possible OTU.

We stress the centrality of establishing local and/or regional barcode reference libraries in maximizing the information gathered through high-throughput sequencing methods, particularly relevant for highly diverse and poorly known ecosystems. As is evident from the large number of BINs detected beyond the Arthropod Program's focal groups, our results depended on regional barcoding efforts such as the BioAlfa program from the Guanacaste Conservation Area in Costa Rica (Janzen & Hallwachs, 2016) and global barcoding projects such as the Global Malaise Program (Arribas et al., 2022; Geiger et al., 2016).

4.1 | Seasonal and spatial variation in metabarcoding data

Incorporating inter-season sampling for monitoring protocols significantly increases sampling effort and costs (Basset et al., 2015). Scaling up sampling with metabarcoding to multiple sampling nights and locations remains costly but with sequencing costs steadily decreasing, it will prove advantageous in the long term. Despite our study being limited to 20 light-trapping nights during each season, our results were robust enough to detect seasonal and spatial variation in arthropod communities. Insect diversity and abundance are directly related to temporal variation in temperature and resource availability and most tropical insects have their seasonal peak during the wet season (Newell et al., 2023; Richards & Windsor, 2007; Wolda, 1980, 1988). Our results not only detected a greater number of species during the wet season, but we were also able to identify significant differences between sampling locations, and sampling days, further supporting metabarcoding as a cost-effective strategy to detect fine-scale differences in arthropod communities. Adding environmental variables to our analyses such as canopy cover, forest debris, and overall forest structure would allow us to identify which variables drive these local differences.

Accurate documentation of ecosystem dynamics requires well-timed surveys. Long-term monitoring efforts rely on frequent sampling over seasons and years in order to capture phenological variation of arthropod emergence (Novais et al., 2016; Richards & Windsor, 2007; Wolda, 1980, 1988). Our sampling was performed over a single new-moon event for each season, and this represents

TABLE 2 BIN count for species identified through metabarcoding and through the 14-year Arthropod Program light-trapping monitoring protocol.

Taxon	Metabarcoding	BIN overlap	Arthropod program data
Lepidoptera	542	525	492
Erebidae	226	135	120
Geometridae	57	135	95
Pyraliidae	74	30	6
Crambidae	48	145	131
Noctuidae	32	16	9
Notodontidae	49	28	19
Lycaenidae	5	2	8
Saturniidae	4	15	24
Hesperiidae	2	3	44
Nymphalidae	2	1	13
Others	43	15	23
Coleoptera	91	34	18
Curculionidae	71	11	2
Scarabaeidae	20	14	15
Passalidae	0	9	1
Hemiptera	5	37	34
Reduviidae	3	21	26
Flatidae	2	16	8
Hymenoptera	39	15	47
Formicidae	34	11	35
Apidae and Halictidae	5	4	12
Blattodea (Termites)	19	4	17
TOTAL	696	615	608

Note: BIN overlap indicates the number of BINs that were detected through both methods for each focal group. Arthropod Program counts include all BINs collected by the light-trapping sampling efforts though these do not all belong to the monitoring focal groups.

a rather short sampling window. By restricting our sampling frequency and periodicity, we were likely to miss much of the diversity sampled in previous years through the Arthropod Program surveys. Indeed, we miss approximately 50% more if we focus exclusively on the monitoring program's focal groups. Repeated sampling, along with long-term monitoring protocols, which provide information on population dynamics (Lamarre et al., 2022), is indispensable for a proper metabarcoding-based biodiversity assessment at any locality. Additionally, fine-scale taxonomic resolution, possible only by sequence association to known BINs, strengthens the use of metabarcoding data for monitoring protocols as an alternative to manual sorting and identification since it integrates local, regional, and global databases. As barcode libraries continue to grow, the number of BINs recovered through long-term monitoring efforts will also increase, particularly when focusing on the less well-studied groups. By filtering our metabarcoding data to focal groups, it appears that the Arthropod Program has registered a comprehensive inventory of Barro Colorado Island. Continued efforts to manually sort and barcode new discoveries remain indispensable; however, the greatest strength in metabarcoding is that it allows us to expand our range of focal taxa.

4.2 | Unexpected guests

Even though light traps are mainly used for collecting nocturnal moths and beetles (Basset et al., 2020; Kitching et al., 2001), our results show that these traps were effective at collecting other orders, such as Diptera and Hemiptera. Although read number cannot be used as a proxy for taxon abundance (Ji et al., 2013; Luo et al., 2023; Yu et al., 2012), it is interesting to note that at one location during the dry season, the highest proportion of reads (~75%) belonged to *Apoica pallens*, a nocturnal eusocial wasp known for its swarming behavior, readily attracted to light traps (Warrant et al., 2006). The high number of reads recovered may be due to multiple wasps falling into the trap located near a nest (Filonila Perez, Yacksecari Lopez, Ricardo Bobadilla & José Alejandro Ramírez Silva, pers. obs.). Focusing on presence absence data alone, our analyses also detected several flightless species, which are probably hitchhiking on the bodies of other insects, which is likely the case for acari. Metabarcoding has been shown to detect prey sequences from within sampled insects (Toju & Baba, 2018). The presence of the little known parasitic insect genus *Myrmecolax* (Strepsiptera: Mymecolacidae) suggests that

we recovered DNA from either an adult free living male (which are usually rare), a juvenile male parasitizing an ant, or an endoparasitic female from within one of its potential hosts, known to belong to multiple orders (Kathirithamby et al., 2010). The presence of ants and termites is explained by alates attracted to the light trap. These examples, although conjectural, demonstrate that even though light-traps were broadly designed for collecting specific groups attracted to light, metabarcoding not only revealed the possibility of specimen by-catch but could potentially reveal predator–prey and host-parasitoid interactions (Šigut et al., 2017; Sow et al., 2019, 2020; Toju & Baba, 2018). Including additional sampling methods will undoubtedly complement these results, particularly when implementing soil sampling protocols (e.g., Winkler or pitfall traps) and flight-intercept traps (e.g., Malaise traps). Winkler and Malaise traps capture soil inhabitants or immature flying arthropods and diurnal flying species respectively.

4.3 | Metabarcoding and traditional monitoring

As expected, parallel sampling yielded fewer species than metabarcoding samples, even when focusing on the Arthropod Program's focal groups. This is evident when considering the total species of microlepidoptera collected by either method, which are often damaged in light-traps and thus difficult to identify through morphology. Partly, the mismatch between methods appears to be associated with species of the same genus (*Talara*, *Diaphania*) or groups notoriously difficult to identify (e.g., *Semaeopus* in Sterrhinae, other genera within Chrysauginae and Spilomelinae; see Appendix S1). These sources of error have been observed in similar studies comparing metabarcoding and traditional sorting of insect communities collected with light-traps (Mata et al., 2021). Similarly, the presence of unidentified cryptic complexes within the Arthropod Program collections may underestimate the true number of species that are readily detected through metabarcoding (Hebert et al., 2004; Ji et al., 2013; Lin et al., 2021).

Combining these metabarcoding data with total monitoring data from the Arthropod Program revealed a considerable overlap of both methods. Traditional sampling detected fewer species, accounting for more than 1000 light-trapping nights over the course of 14 years. On the other hand, 20 light-trapping nights per season over a single new-moon event yielded almost the same number of focal group BINs, with a ~50% overlap with the Arthropod Program long-term collections. This suggests that repeated sampling using metabarcoding would capture the diversity detected by the Program as well as additional species that may previously have been overlooked. As mentioned above, light trapping is an ideal sampling method for nocturnal moths, which include some of the best studied groups in the region (Hausmann et al., 2020; Janzen & Hallwachs, 2016; Murillo-Ramos et al., 2019). If we consider Geometridae, the Arthropod Program has recorded 230 BINs on BCI while metabarcoding detected 29 additional BINs. This suggests that, at least for this group, sampling on BCI has so far been relatively comprehensive. When we focus on Saturniidae, metabarcoding identified four BINs, which so far have not been recorded by the Arthropod Program. This family includes large moths, which due to their size are not readily collected by automatic bucket traps, and so it is likely that manual sorting may have missed these species (Basset et al., 2017). On the other hand, DNA metabarcoding is likely to detect these species as scales or wing fragments may have fallen into the trap when individuals attracted to the light bounce into the intercept panes (Patzold et al., 2020; Rose et al., 1994).

5 | CONCLUSIONS

Our study demonstrates the standalone capacity of metabarcoding to detect seasonal and spatial variation in arthropod diversity based on a low sampling effort; in our case 20 non-consecutive light traps in a single new-moon event per season in 10 different localities within the ForestGEO's 50-hectare plot. Detecting spatiotemporal differences in species diversity is one of the minimum requirements for a successful monitoring program (Montgomery et al., 2021). Our study successfully detected seasonal and site variation in species diversity and allowed us to identify which families were most prominent during each. Additionally, we were able to detect differences of arthropod communities between sampling localities during the wet season, and sampling nights during the dry. It is well known that microhabitat conditions influence community composition, and detecting these fine-scale differences greatly strengthens the application of metabarcoding to biomonitoring programs. Perhaps one of the biggest bottlenecks for arthropod monitoring relates to the time and effort spent in the sorting and identification steps, which significantly increases for every additional sample and replicate. The scalability of metabarcoding allows for a higher number of yearly replicates without the temporal and financial burden of manual sorting and identification. We detected a considerable number of species based on Barcode Index Numbers, which permitted a level of taxonomic resolution not often seen in DNA metabarcoding studies in highly diverse tropical forests. This level of resolution, however, was only possible by making use of the extensive international barcode data that is available. As metabarcoding studies gain popularity, and considering the thousands of species awaiting description, it is of utmost importance that the scientific community continues its ongoing efforts to populate barcode databases. These databases are essential in delivering on the promise of metabarcoding as a cost-effective tool for biodiversity monitoring and assessment.

AUTHOR CONTRIBUTIONS

Y.B., D.S.V., and G.P.A.L conceived the research, G.P.A.L., P.B., A.S., R.C., F.P., Y.L., R.B., and J.A.R. participated in fieldwork and sample preparation for subsequent DNA extraction and sequencing, Y.B. and B.L-H. performed data curation, D.S.V. analyzed the data with extensive input from S.T.S., E.N.V., A.C.Z., and Y.C. D.S.V. wrote the manuscript. All authors approved the final version of the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank Lovisa Dück for help during fieldwork and Jeremy deWaard for valuable discussions regarding metabarcoding protocols and data interpretation. We thank Douglas Yu for an in-depth and thoughtful review of a previous version of this manuscript. The project has been supported by the Czech Funding agency GAČR grant No. 20-31295S to Y.B. and 19-15645Y to G.P.A.L. Y.B., G.P.A.L., and J.A.R. were supported by the Sistema Nacional de Investigación, SENACYT, Panama.

CONFLICT OF INTEREST STATEMENT

No conflicts of interest.

DATA AVAILABILITY STATEMENT

Raw sequence data is available in the mBrave platform and upon request. Final datasets and scripts are available online as a GitHub repository at: https://github.com/DanielSoutoV/metabarcoding.

ORCID

Daniel Souto-Vilarós https://orcid.org/0000-0002-8803-5173
Yves Basset https://orcid.org/0000-0002-1942-5717
Petr Blažek https://orcid.org/0000-0002-0901-4578
Benita Laird-Hopkins https://orcid.org/0000-0003-0270-1241
Simon T. Segar https://orcid.org/0000-0001-6621-9409
Eduardo Navarro-Valencia https://orcid.org/0000-0002-6898-7310
Richard Čtvrtečka https://orcid.org/0000-0002-0599-6850
José Alejandro Ramírez Silva https://orcid.org/0000-0001-6799-5207
Greg P. A. Lamarre https://orcid.org/0000-0002-7645-985X

REFERENCES

Anderson-Teixeira, K. J., Davies, S. J., Bennett, A. C., Gonzalez-Akre, E. B., Muller-Landau, H. C., Wright, S. J., Abu Salim, K., Almeyda Zambrano, A. M., Alonso, A., Baltzer, J. L., Basset, Y., Bourg, N. A., Broadbent, E. N., Brockelman, W. Y., Bunyavejchewin, S., Burslem, D. F., Butt, N., Cao, M., Cardenas, D., … Zimmerman, J. (2015). CTFS-ForestGEO: A worldwide network monitoring forests in an era of global change. *Global Change Biology*, *21*, 528–549.

Arribas, P., Andújar, C., Bohmann, K., de Waard, J. R., Economo, E. P., Elbrecht, V., Geisen, S., Goberna, M., Krehenwinkel, H., Novotny, V., Zinger, L., Creedy, T. J., Meramveliotakis, E., Noguerales, V., Overcast, I., Morlon, H., Papadopoulou, A., Vogler, A. P., & Emerson, B. C. (2022). Toward global integration of biodiversity big data: A harmonized metabarcode data generation module for terrestrial arthropods. *GigaScience*, *11*, giac065.

Basset, Y., Cizek, L., Cuénoud, P., Didham, R. K., Guilhaumon, F., Missa, O., Novotny, V., Ødegaard, F., Roslin, T., Schmidl, J., Tishechkin, A. K., Winchester, N. N., Roubik, D. W., Aberlenc, H.-P., Bail, J., Barrios, H., Bridle, J. R., Castaño-Meneses, G., Corbara, B., … Leponce, M. (2012). Arthropod diversity in a tropical Forest. *Science*, *338*, 1481–1484.

Basset, Y., Cizek, L., Cuénoud, P., Didham, R. K., Novotny, V., Ødegaard, F., Roslin, T., Tishechkin, A. K., Schmidl, J., Winchester, N. N., Roubik, D. W., Aberlenc, H.-P., Bail, J., Barrios, H., Bridle, J. R., Castaño-Meneses, G., Corbara, B., Curletti, G., Duarte da Rocha, W., … Leponce, M. (2015). Arthropod distribution in a tropical rainforest: Tackling a four dimensional puzzle. *PLoS One*, *10*, e0144110.

Basset, Y., Donoso, D. A., Hajibabaei, M., Wright, M. T. G., Perez, K. H. J., Lamarre, G. P. A., León, L. F. D., Palacios-Vargas, J. G., Castaño-Meneses, G., Rivera, M., Perez, F., Bobadilla, R., Lopez, Y., Ramirez, J. A., & Barrios, H. (2020). Methodological considerations for monitoring soil/litter arthropods in tropical rainforests using DNA metabarcoding, with a special emphasis on ants, springtails and termites. *Metabarcoding and Metagenomics*, *4*, e58572.

Basset, Y., Eastwood, R., Sam, L., Lohman, D. J., Novotny, V., Treuer, T., Miller, S. E., Weiblen, G. D., Pierce, N. E., Bunyavejchewin, S., Sakchoowong, W., Kongnoo, P., & Osorio-Arenas, M. A. (2013). Cross-continental comparisons of butterfly assemblages in tropical rainforests: Implications for biological monitoring. *Insect Conservation and Diversity*, *6*, 223–233.

Basset, Y., Hajibabaei, M., Wright, M. T. G., Castillo, A. M., Donoso, D. A., Segar, S. T., Souto-Vilarós, D., Soliman, D. Y., Roslin, T., Smith, M. A., Lamarre, G. P. A., De León, L. F., Decaëns, T., Palacios-Vargas, J. G., Castaño-Meneses, G., Scheffrahn, R. H., Rivera, M., Perez, F., Bobadilla, R., … Barrios, H. (2022). Comparison of traditional and DNA metabarcoding samples for monitoring tropical soil arthropods (Formicidae, Collembola and Isoptera). *Scientific Reports*, *12*, 10762.

Basset, Y., & Lamarre, G. P. A. (2019). Toward a world that values insects. *Science*, *364*, 1230–1231.

Basset, Y., Lamarre, G. P. A., Ratz, T., Segar, S. T., Decaëns, T., Rougerie, R., Miller, S. E., Perez, F., Bobadilla, R., Lopez, Y., Ramirez, J. A., Aiello, A., & Barrios, H. (2017). The Saturniidae of Barro Colorado Island, Panama: A model taxon for studying the long-term effects of climate change? *Ecology and Evolution*, *7*, 9991–10004.

Bell, K. L., Fowler, J., Burgess, K. S., Dobbs, E. K., Gruenewald, D., Lawley, B., Morozumi, C., & Brosi, B. J. (2017). Applying pollen DNA metabarcoding to the study of plant–pollinator interactions1. *Applications in Plant Sciences*, *5*, apps.1600124.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*, 289–300.

Carbayo, F., & Marques, A. C. (2011). The costs of describing the entire animal kingdom. *Trends in Ecology and Evolution*, *26*, 154–155.

Chen, H., & Boutros, P. C. (2011). VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics*, *12*, 35.

Chimeno, C., Hausmann, A., Schmidt, S., Raupach, M. J., Doczkal, D., Baranov, V., Hübner, J., Höcherl, A., Albrecht, R., Jaschhof, M., Haszprunar, G., & Hebert, P. D. N. (2022). Peering into the darkness: DNA barcoding reveals surprisingly high diversity of unknown species of Diptera (Insecta) in Germany. *Insects*, *13*, 82.

Clarke, L. J., Suter, L., Deagle, B. E., Polanowski, A. M., Terauds, A., Johnstone, G. J., & Stark, J. S. (2021). Environmental DNA metabarcoding for monitoring metazoan biodiversity in Antarctic nearshore ecosystems. *PeerJ*, *9*, e12458.

Deutsch, C. A., Tewksbury, J. J., Huey, R. B., Sheldon, K. S., Ghalambor, C. K., Haak, D. C., & Martin, P. R. (2008). Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 6668–6672.

de Waard, J. R., Levesque-Beaudin, V., de Waard, S. L., Ivanova, N. V., McKeown, J. T. A., Miskie, R., Naik, S., Perez, K. H. J., Ratnasingham, S., Sobel, C. N., Sones, J. E., Steinke, C., Telfer, A. C., Young, A. D., Young, M. R., Zakharov, E. V., & Hebert, P. D. N. (2019). Expedited assessment of terrestrial arthropod diversity by coupling malaise traps with DNA barcoding 1. *Genome*, *62*, 85–95.

Didham, R. K., Basset, Y., Collins, C. M., Leather, S. R., Littlewood, N. A., Menz, M. H. M., Müller, J., Packer, L., Saunders, M. E., Schönrogge, K., Stewart, A. J. A., Yanoviak, S. P., & Hassall, C. (2020). Interpreting insect declines: Seven challenges and a way forward. *Insect Conservation and Diversity*, *13*, 103–114.

Donkersley, P., Ashton, L., Lamarre, G. P. A., & Segar, S. (2022). Global insect decline is the result of wilful political failure: A battle plan for entomology. *Ecology and Evolution*, *12*, e9417.

Engel, M. S., Ceríaco, L. M. P., Daniel, G. M., Dellapé, P. M., Löbl, I., Marinov, M., Reis, R. E., Young, M. T., Dubois, A., Agarwal, I., Lehmann, A. P., Alvarado, M., Alvarez, N., Andreone, F., Araujo-Vieira, K., Ascher, J. S., Baêta, D., Baldo, D., Bandeira, S. A., … Zacharie, C. K. (2021). The taxonomic impediment: A shortage of taxonomists, not the lack of technical approaches. *Zoological Journal of the Linnean Society*, *193*, 381–387.

Foster, Z. S. L., Sharpton, T. J., & Grünwald, N. J. (2017). Metacoder: An R package for visualization and manipulation of community taxonomic diversity data. *PLoS Computational Biology*, *13*, e1005404.

Geiger, M., Moriniere, J., Hausmann, A., Haszprunar, G., Wägele, W., Hebert, P., & Rulik, B. (2016). Testing the global malaise trap program – How well does the current barcode reference library identify flying insects in Germany? *Biodiversity Data Journal*, *4*, e10671.

Hallmann, C. A., Ssymank, A., Sorg, M., de Kroon, H., & Jongejans, E. (2021). Insect biomass decline scaled to species diversity: General patterns derived from a hoverfly community. *Proceedings of the National Academy of Sciences of the United States of America*, *118*, e2002554117.

Hallmann, C. A., Zeegers, T., van Klink, R., Vermeulen, R., van Wielink, P., Spijkers, H., van Deijk, J., van Steenis, W., & Jongejans, E. (2020). Declining abundance of beetles, moths and caddisflies in The Netherlands. *Insect Conservation and Diversity*, *13*, 127–139.

Hamilton, A. J., Basset, Y., Benke, K. K., Grimbacher, P. S., Miller, S. E., Novotný, V., Samuelson, G. A., Stork, N. E., Weiblen, G. D., & Yen, J. D. (2010). Quantifying uncertainty in estimation of tropical arthropod species richness. *The American Naturalist*, *176*(1), 90–95.

Harvey, J. A., Tougeron, K., Gols, R., Heinen, R., Abarca, M., Abram, P. K., Basset, Y., Berg, M., Boggs, C., Brodeur, J., Cardoso, P., de Boer, J. G., De Snoo, G. R., Deacon, C., Dell, J. E., Desneux, N., Dillon, M. E., Duffy, G. A., Dyer, L. A., … Chown, S. L. (2023). Scientists' warning on climate change and insects. *Ecological Monographs*, *93*, e1553.

Hausmann, A., Diller, J., Moriniere, J., Höcherl, A., Floren, A., & Haszprunar, G. (2020). DNA barcoding of fogged caterpillars in Peru: A novel approach for unveiling host-plant relationships of tropical moths (Insecta, Lepidoptera). *PLoS One*, *15*, e0224188.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 14812–14817.

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*, S96–S99.

Hernández-Triana, L. M., Prosser, S. W., Rodríguez-Perez, M. A., Chaverri, L. G., Hebert, P. D. N., & Ryan Gregory, T. (2014). Recovery of DNA barcodes from blackfly museum specimens (Diptera: Simuliidae) using primer sets that target a variety of sequence lengths. *Molecular Ecology Resources*, *14*, 508–518.

Hochkirch, A., Casino, A., Penev, L., Allen, D., Tilley, L., Georgiev, T., Gospodinov, K., & Barov, B. (2022). *European red list of insect taxonomists*. Publications Office of the European Union.

Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: An R package for rarefaction and extrapolation of species diversity (hill numbers). *Methods in Ecology and Evolution*, *7*, 1451–1456.

Janzen, D. H., & Hallwachs, W. (2016). DNA barcoding the Lepidoptera inventory of a large complex tropical conserved wildland, Area de Conservacion Guanacaste, northwestern Costa Rica. *Genome*, *59*, 641–660.

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P. M., Woodcock, P., Edwards, F. A., Larsen, T. H., Hsu, W. W., Benedick, S., Hamer, K. C., Wilcove, D. S., Bruce, C., Wang, X., Levi, T., Lott, M., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*, 1245–1257.

Kathirithamby, J., Hayward, A., Mcmahon, D. P., Ferreira, R. S., Andreazze, R., Almeida Andrade, H. T. D., & Fresneau, D. (2010). Conspecifics of a heterotrophic heteronomous species of Strepsiptera (Insecta) are matched by molecular characterization. *Systematic Entomology*, *35*, 234–242.

Kauserud, H., Mathiesen, C., & Ohlson, M. (2008). High diversity of fungi associated with living parts of boreal forest bryophytes. *Botany*, *86*, 1326–1333.

Keck, F., Couton, M., & Altermatt, F. (2023). Navigating the seven challenges of taxonomic reference databases in metabarcoding analyses. *Molecular Ecology Resources*, *23*, 742–755.

Kirse, A., Bourlat, S. J., Langen, K., & Fonseca, V. G. (2021). Metabarcoding malaise traps and soil eDNA reveals seasonal and local arthropod diversity shifts. *Scientific Reports*, *11*, 10498.

Kitching, R. L., Li, D., & Stork, N. E. (2001). Assessing biodiversity "sampling packages": How similar are arthropod assemblages in different tropical rainforests? *Biodiversity and Conservation*, *10*, 793–813.

Lamarre, G. P. A., Fayle, T. M., Segar, S. T., Laird-Hopkins, B. C., Nakamura, A., Souto-Vilarós, D., Watanabe, S., & Basset, Y. (2020). Chapter eight - monitoring tropical insects in the 21st century. In A. J. Dumbrell, et al. (Eds.), *Advances in ecological research. Tropical ecosystems in the 21st century* (pp. 295–330). Academic Press.

Lamarre, G. P. A., Pardikes, N. A., Segar, S., Hackforth, C. N., Laguerre, M., Vincent, B., Lopez, Y., Perez, F., Bobadilla, R., Silva, J. A. R., & Basset, Y. (2022). More winners than losers over 12 years of monitoring tiger moths (Erebidae: Arctiinae) on Barro Colorado Island, Panama. *Biology Letters*, *18*, 20210519.

Lin, X.-L., Mo, L., Bu, W.-J., & Wang, X.-H. (2021). The first comprehensive DNA barcode reference library of Chinese Tanytarsus (Diptera: Chironomidae) for environmental DNA metabarcoding. *Diversity and Distributions*, *27*, 1932–1941.

Lister, B. C., & Garcia, A. (2018). Climate-driven declines in arthropod abundance restructure a rainforest food web. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, E10397–E10406.

Liu, M., Clarke, L. J., Baker, S. C., Jordan, G. J., & Burridge, C. P. (2020). A practical guide to DNA metabarcoding for entomological ecologists. *Ecological Entomology*, *45*, 373–385.

Lucas, M., Forero, D., & Basset, Y. (2016). Diversity and recent population trends of assassin bugs (Hemiptera: Reduviidae) on Barro Colorado Island, Panama. *Insect Conservation and Diversity*, *9*, 546–558.

Luo, M., Ji, Y., Warton, D., & Yu, D. W. (2023). Extracting abundance information from DNA-based data. *Molecular Ecology Resources*, *23*, 174–189.

Macher, T.-H., Schütz, R., Hörren, T., Beermann, A. J., & Leese, F. (2023). It's raining species: Rainwash eDNA metabarcoding as a minimally invasive method to assess tree canopy invertebrate diversity. *Environmental DNA*, *5*, 3–11.

Magoga, G., Forni, G., Brunetti, M., Meral, A., Spada, A., De Biase, A., & Montagna, M. (2022). Curation of a reference database of COI sequences for insect identification through DNA metabarcoding: COins. *Database*, *2022*, baac055.

Mata, V. A., Ferreira, S., Campos, R. M., da Silva, L. P., Veríssimo, J., Corley, M. F. V., & Beja, P. (2021). Efficient assessment of nocturnal flying insect communities by combining automatic light traps and DNA metabarcoding. *Environmental DNA*, *3*, 398–408.

McKnight, D. T., Huerlimann, R., Bower, D. S., Schwarzkopf, L., Alford, R. A., & Zenger, K. R. (2019). microDecon: A highly accurate read-subtraction tool for the post-sequencing removal of contamination in metabarcoding studies. *Environmental DNA*, *1*, 14–25.

McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*, *8*, e61217.

Montgomery, G. A., Belitz, M. W., Guralnick, R. P., & Tingley, M. W. (2021). Standards and best practices for monitoring and benchmarking insects. *Frontiers in Ecology and Evolution*, *8*, 579193.

Murillo-Ramos, L., Brehm, G., Sihvonen, P., Hausmann, A., Holm, S., Ghanavi, H. R., Õunap, E., Truuverk, A., Staude, H., Friedrich, E., Tammaru, T., & Wahlberg, N. (2019). A comprehensive molecular phylogeny of Geometridae (Lepidoptera) with a focus on enigmatic small subfamilies. *PeerJ*, *7*, e7386.

Newell, F. L., Ausprey, I. J., & Robinson, S. K. (2023). Wet and dry extremes reduce arthropod biomass independently of leaf phenology in the wet tropics. *Global Change Biology*, *29*, 308–323.

Novais, S. M. A., Evangelista, L. A., Reis-Júnior, R., & Neves, F. S. (2016). How does dung beetle (Coleoptera: Scarabaeidae) diversity vary along a rainy season in a tropical dry Forest? *Journal of Insect Science*, *16*, 81.

Oksanen, J. (2007). Vegan: community ecology package. R package version 1.8-5. http://www.cran.r-project.org

Paknia, O., Rajaei Sh, H., & Koch, A. (2015). Lack of well-maintained natural history collections and taxonomists in megadiverse developing countries hampers global biodiversity exploration. *Organisms, Diversity and Evolution*, *15*, 619–629.

Patzold, F., Zilli, A., & Hundsdoerfer, A. K. (2020). Advantages of an easy-to-use DNA extraction method for minimal-destructive analysis of collection specimens. *PLoS One*, *15*, e0235222.

Porter, T. M., & Hajibabaei, M. (2020). Putting COI metabarcoding in context: The utility of exact sequence variants (ESVs) in biodiversity analysis. *Frontiers in Ecology and Evolution*, *8*, 248.

Porter, T. M., Morris, D. M., Basiliko, N., Hajibabaei, M., Doucet, D., Bowman, S., Emilson, E. J. S., Emilson, C. E., Chartrand, D., Wainio-Keizer, K., Séguin, A., & Venier, L. (2019). Variations in terrestrial arthropod DNA metabarcoding methods recovers robust beta diversity but variable richness and site indicators. *Scientific Reports*, *9*, 18218.

Prosser, S. W. J., de Waard, J. R., Miller, S. E., & Hebert, P. D. N. (2016). DNA barcodes from century-old type specimens using next-generation sequencing. *Molecular Ecology Resources*, *16*, 487–497.

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Ratnasingham, S. (2019). mBRAVE: The multiplex barcode research and visualization environment. *Biodiversity Information Science and Standards*, *3*, e37986.

Ratnasingham, S., & Hebert, P. D. N. (2007). bold: The barcode of life data system (http://www.barcodinglife.org). *Molecular Ecology Notes*, *7*, 355–364.

Ratnasingham, S., & Hebert, P. D. N. (2013). A DNA-based registry for all animal species: The barcode index number (BIN) system. *PLoS One*, *8*, e66213.

Richards, L. A., & Windsor, D. M. (2007). Seasonal variation of arthropod abundance in gaps and the understorey of a lowland moist Forest in Panama. *Journal of Tropical Ecology*, *23*, 169–176.

Rose, O. C., Brookes, M. I., & Mallet, J. L. B. (1994). A quick and simple nonlethal method for extracting DNA from butterfly wings. *Molecular Ecology*, *3*, 275.

Sánchez-Bayo, F., & Wyckhuys, K. A. G. (2019). Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation*, *232*, 8–27.

Šigut, M., Kostovčík, M., Šigutová, H., Hulcr, J., Drozd, P., & Hrček, J. (2017). Performance of DNA metabarcoding, standard barcoding, and morphological approach in the identification of host–parasitoid interactions. *PLoS One*, *12*, e0187803.

Sow, A., Brévault, T., Benoit, L., Chapuis, M.-P., Galan, M., Coeur d'acier, A., Delvare, G., Sembène, M., & Haran, J. (2019). Deciphering host-parasitoid interactions and parasitism rates of crop pests using DNA metabarcoding. *Scientific Reports*, *9*, 3646.

Sow, A., Haran, J., Benoit, L., Galan, M., & Brévault, T. (2020). DNA metabarcoding as a tool for disentangling food webs in agroecosystems. *Insects*, *11*, 294.

Stackebrandt, E., & Goebl, B. M. (1994). Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *International Journal of Systematic and Evolutionary Microbiology*, *44*, 846–849.

Steinke, D., de Waard, S. L., Sones, J. E., Ivanova, N. V., Prosser, S. W. J., Perez, K., Braukmann, T. W. A., Milton, M., Zakharov, E. V., de Waard, J. R., Ratnasingham, S., & Hebert, P. D. N. (2022). Message in a bottle—Metabarcoding enables biodiversity comparisons across ecoregions. *GigaScience*, *11*, giac040.

Stork, N. E., McBroom, J., Gely, C., & Hamilton, A. J. (2015). New approaches narrow global species estimates for beetles, insects, and terrestrial arthropods. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 7519–7523.

Suchan, T., Talavera, G., Sáez, L., Ronikier, M., & Vila, R. (2019). Pollen metabarcoding as a tool for tracking long-distance insect migrations. *Molecular Ecology Resources*, *19*, 149–162.

Sun, Z., Majaneva, M., Sokolova, E., Rauch, S., Meland, S., & Ekrem, T. (2019). DNA metabarcoding adds valuable information for management of biodiversity in roadside stormwater ponds. *Ecology and Evolution*, *9*, 9712–9722.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*, 2045–2050.

Toju, H., & Baba, Y. G. (2018). DNA metabarcoding of spiders, insects, and springtails for exploring potential linkage between above- and below-ground food webs. *Zoological Letters*, *4*, 4.

Tommasi, N., Ferrari, A., Labra, M., Galimberti, A., & Biella, P. (2021). Harnessing the power of Metabarcoding in the ecological interpretation of plant-pollinator DNA data: Strategies and consequences of filtering approaches. *Diversity*, *13*, 437.

Wagner, D. L., Grames, E. M., Forister, M. L., Berenbaum, M. R., & Stopak, D. (2021). Insect decline in the Anthropocene: Death by a thousand cuts. *Proceedings of the National Academy of Sciences of the United States of America*, *118*, e2023989118.

Warrant, E. J., Kelber, A., Wallén, R., & Wcislo, W. T. (2006). Ocellar optics in nocturnal and diurnal bees and wasps. *Arthropod Structure and Development*, *35*, 293–305.

Weisser, W. W., & Siemann, E. (2008). The various effects of insects on ecosystem functioning. In W. W. Weisser & E. Siemann (Eds.), *Insects and ecosystem function. Ecological studies* (pp. 3–24). Springer.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag.

Wolda, H. (1980). Seasonality of tropical insects. *Journal of Animal Ecology*, *49*, 277–290.

Wolda, H. (1988). Insect seasonality: Why? *Annual Review of Ecology and Systematics*, *19*, 1–18.

Xu, S., Li, Z., Tang, W., Dai, Z., Zhou, L., Feng, T., Chen, M., Liu, S., Fu, X., Wu, T., Hu, E., & Yu, G. (2022). MicrobiotaProcess: A comprehensive R package for managing and analyzing microbiome and other ecological data within the tidy framework. in press.

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*, 613–623.

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*, 2352–2366.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Souto-Vilarós, D., Basset, Y., Blažek, P., Laird-Hopkins, B., Segar, S. T., Navarro-Valencia, E., Zamora, A. C., Campusano, Y., Čtvrtečka, R., Savage, A. F., Perez, F., Lopez, Y., Bobadilla, R., Ramírez Silva, J. A., & Lamarre, G. P. A. (2024). Illuminating arthropod diversity in a tropical forest: Assessing biodiversity by automatic light trapping and DNA metabarcoding. *Environmental DNA*, *6*, e540. https://doi.org/10.1002/edn3.540