Toward repurposing global passive air sampling networks for insect monitoring: Promises and pitfalls of airborne eDNA

Jordi Vilanova[1*] https://orcid.org/0009-0009-3685-4343

Luis Rodrigo Arce-Valdés[1] https://orcid.org/0000-0001-6445-7534

Xianming Zhang[2] https://orcid.org/0000-0002-5301-7899

Julia J. Mlynarek[3, 4] https://orcid.org/0000-0002-1569-9403

Rassim Khelifa[1] https://orcid.org/0000-0001-6632-8787

[1] Biology Department, Concordia University, 7141 Sherbrooke St. W., Montreal, QC H4B 1R6, Canada. j.vilanovabroto@gmail.com; bio.l.rodrigo.arce@gmail.com; rassim.khelifa@concordia.ca

[2] Department of Chemistry and Biochemistry, Concordia University, 7141 Sherbrooke St. W., Montreal, QC H4B 1R6, Canada. xianming.zhang@concordia.ca

[3] Insectarium de Montréal, 4581, Sherbrooke St. E., Montreal, QC, H1X 2B3, Canada. julia.mlynarek@montreal.ca

[4] Institut de recherche en biologie végétal, Département des sciences biologiques, Université de Montréal. 4101 Sherbrooke St. E., Montreal, QC, H1X 2B2, Canada

Corresponding author (*): j.vilanovabroto@gmail.com

Abstract

1. Polyurethane foam passive air samplers (PUF-PAS) are widely deployed to monitor environmental pollutants, yet their capacity to capture biological signals such as environmental DNA (eDNA) remains largely unexplored. Recent advances in airborne eDNA research create a timely opportunity to evaluate PUF-PAS as a tool for biodiversity monitoring and to leverage existing global sampling networks for conservation. Focusing on insects, we assess whether PUF-PAS can provide eDNA-based biodiversity data comparable to those obtained with traditional sampling methods.

2. We compared the taxonomic richness, community composition, and detection biases of PUF-PAS to a conventional method for sampling flying insects, Malaise traps. We deployed both sampling approaches across five sites in Montreal, Canada, during the entire flight season, and collected samples biweekly. Insect communities were characterised at the order level using morphological identification (Malaise traps) and DNA metabarcoding (PUF-PAS).

3. Insect communities collected with PUF-PAS and Malaise traps showed substantial overlap, with 73.3% (n = 11) of insect orders detected by both methods; 20% (n = 3) were unique to Malaise traps and 6.7% (n = 1) to PUF-PAS. On average, estimates of insect abundance from Malaise traps were strongly positively associated with PUF-PAS detection probability and read counts. However, PUF-PAS detection probability and read counts were best explained by Malaise trap counts measured several weeks prior and integrated over extended temporal windows, indicating a delayed biological signal.

4. Our results suggest that PUF-PAS is better suited for interannual monitoring of common insect taxa, particularly for detecting broad temporal trends rather than short-term dynamics. Given the availability of global passive air sampling infrastructure, this approach not only has strong

potential for large-scale biodiversity monitoring, but also could foster interdisciplinary collaborations between environmental toxicologists and biodiversity scientists.

Keywords: airborne eDNA, metabarcoding, global monitoring, insects, Passive air sampler

1. Introduction

Globally, substantial changes in insect populations have been documented, with declines reported in many regions and ecosystems (Didham et al., 2020; Halley & Pimm, 2023; Harvey et al., 2022). For instance, long-term monitoring programs have reported large reductions in the biomass and abundance of flying insects, in some cases exceeding 70% (Dalton et al., 2023; Hallmann et al., 2017; Van Klink et al., 2024). Documenting such trends is essential for understanding ecosystem functioning and predicting the consequences of environmental changes. However, detecting and tracking changes in insect biodiversity remains challenging because most traditional monitoring approaches rely on lethal, labor-intensive sampling methods and extensive taxonomic expertise. As a result, there is growing interest in developing alternative approaches that are standardized, scalable, non-invasive, and cost-effective to monitor insect communities and inform decision makers for effective conservation.

Environmental DNA (eDNA) approaches try to overcome the limitations of traditional sampling methods by capturing DNA directly from the environment, including soil (Semenov, 2021), water (De Sousa et al., 2019), or air (Roger et al., 2022). Most eDNA applications developed over the past decades have focused on aquatic environments (Yamahara et al., 2025). However, more recently, airborne eDNA has emerged as a promising tool to monitor terrestrial species and the ecological changes they undergo. Additionally, airborne eDNA could be particularly well suited to monitor insect communities as many studies have already successfully identified insect DNA from sampled air (*e.g.,* Craine et al., 2025a; Roger et al., 2022; Tournayre et al., 2025), with two studies finding arthropods as the most dominant metazoan group recovered from airborne eDNA samples (Nousias et al., 2025; Sullivan et al., 2023). A better understanding of the biases of

71 airborne eDNA detection compared to traditional trapping techniques is needed to establish the
72 method for biodiversity monitoring.

73 Polyurethane foam passive air samplers (PUF-PAS) have been established on the global scale to
74 track aerosols and atmospheric pollutants passively (Harner et al., 2024; Liu et al., 2021; Saini et
75 al., 2020; Schuster et al., 2015). This tool shows promise for airborne eDNA sampling, as it can
76 be operated without power source, is cost-effective, and is used worldwide in a network designed
77 to sample air pollutants (Figure 1A) (Pozo et al., 2006). Recent studies have recovered eDNA from
78 the samplers (e.g., fungi, bacteria, animals), showing promising signs of their use for biodiversity
79 monitoring (Kalisa et al., 2024; Sanchez et al., 2025). Nevertheless, as with any eDNA technique,
80 it is crucial to analyse its biases, determine how well airborne eDNA reflects actual species
81 communities, and assess whether detected DNA signals are linked to organismal abundance or
82 biomass.

83 In this study, we investigate the representativeness of insect communities sampled by PUF-PAS,
84 by comparing them with a traditional generalist sampling method, Malaise traps (Malaise, 1937;
85 Santos & Fernandes, 2021). Specifically, we assess the extent to which insect detection by PUF-
86 PAS reflects insect community composition and abundance in five sites in the Montreal region,
87 Canada (Figure 1B). We further investigate the temporal relationship between insect abundance
88 and detection by PUF-PAS, considering both potential lag effects and the integration of signals
89 over time. We expect a positive relationship between insect biomass and the probability of
90 detection and the number of DNA reads recovered by PUF-PAS (Yates et al., 2025). We expect
91 no lag effect, with eDNA detection by PUF-PAS positively correlated with instantaneous insect
92 abundance.
93

2. Materials and Methods

2.1. Sampling tools

2.1.1. Polyurethane Foam (PUF) Disk-based Passive Air Sampler (PUF-PAS). PUF-PAS are passive air sampling devices consisting of a polyurethane foam (PUF) disk housed within a double-dome stainless steel chamber, following standard designs used in atmospheric monitoring (Shoeib & Harner, 2002) (Tisch Environmental, Inc., USA). The chamber is composed of two opposing hemispherical bowls (upper dome: 25 cm diameter × 8 cm tall; lower dome: 20 cm diameter × 7 cm tall) overlapping vertically by approximately 1 cm, with a side gap of 2 cm and seven holes of 0.65 mm on the underside in a circular distribution. The PUF disks (14 cm diameter × 1.35 cm thickness) were autoclaved prior to deployment and transported to the field in sterile containers. Disks were positioned centrally within the chamber and secured on a support rod (Figures 1C, 1D). To minimize contamination, disks were handled using sterile forceps and gloves at all stages of deployment and retrieval. Samplers were deployed in open air without active airflow and exposed continuously for two weeks.

2.1.2. Malaise trap. We used Malaise Trap II, Townes Style (Bugdorm, MegaView Co., Ltd., Taiwan). This sampling device is a modular flight interception trap, 165 cm in length, 115 cm in width, 200 cm high for the tall end, 110 cm high for the short end. The intercept panels have a 680 μm aperture, and the central mesh panel a 470 μm aperture. The collection bottles were filled with 90% ethanol to minimize individual degradation.

2.2. Sampling protocol

We set up each sampler in five different locations around Montreal, Canada (Figure 1B; Table S1). Malaise traps were placed with the tall end south-facing. PUF-PAS were placed 1.5 m above ground, two meters from the Malaise trap, south-facing (Figure 1C, D). From 13 May 2024 to 9 October 2024, we visited each site every two weeks to change the Malaise trap bottle and the PUF disk (Table S2). Malaise trap bottles were refilled with 90% ethanol, and the PUF disks were stored in sterile 50 mL falcon tubes. Both were placed in coolers for transportation from the field to the laboratory, where they were stored in a -20 °C freezer to minimize degradation of the individuals and DNA.

2.3. Sample processing

2.3.1. Malaise trap. Samples collected from Malaise traps were sorted manually and identified to the order level. To estimate the biomass of each order, we followed the protocol by Dunn et al. (2023), by straining the organisms and leaving them to dry for one hour at room temperature. To estimate the number of individuals, we placed the individuals on a flat surface and photographed them using a Canon EOS 5D Mark III camera. Individuals were then manually counted from the images using ImageJ 1.54 (Rueden et al., 2017).

2.3.2. PUF-PAS analysis. Following the work of Kalisa et al. (2024), we extracted metagenomic DNA from the PUF discs using the DNeasy PowerMax Soil Kit (Qiagen, Venlo, Limburg, The Netherlands) with a slightly modified version of the manufacturer instructions to accommodate for the size of the PUF disk. We included four negative controls (two sterile PUF disks and two DNA-free PCR samples) to evaluate possible contamination during DNA extractions and sequencing. Additionally, we designed and sequenced two replicated PUF discs using a mock community as positive control (Marinchel et al., 2023; Smith et al., 2017). We used PCR to amplify a 313 bp fragment of the COI gene using the “Leray” set of primers for insect metabarcoding (Morrill et al.,

2021). Then, we distributed the samples between two libraries (Table S2) and sequenced them on an Illumina Miseq (600-cycles, Illumina, San Diego, USA). We performed all bioinformatic processing within the QIIME2 v2025.4.0 environment (Bolyen et al., 2019) and employed the COInr database as our reference database for taxonomic assignment (Meglécz, 2023). We employed two different libraries, two denoising algorithms, and four quality filtering stringencies, which resulted in 16 ASVs × sample matrices. Finally, we filtered each of these matrices using a standardized protocol based on negative controls cleaning and minimum number of reads thresholds (Drake et al., 2022). Full details on library construction, mock community design, bioinformatic processing, and filtering choices are provided in Supplementary Text S1.

2.4. Statistical analyses

All analyses were conducted using R 4.5.2 (R Core Team, 2026) and RStudio v.2026.01.1+403 (RStudio Team, 2026). Due to the strong positive correlation between insect biomass and insect count (Bayesian Regression Model: Bayes $R^2 = 0.82$, Est.Error = 0.005, Est. Slope = 1.31) (Figure S1), we used only insect count as explanatory variable. Bayesian regression models (BRM) were fitted using the *brm* function in the *brms* package (Bürkner, 2017), using four chains, 4000 iterations, 2000 burn-in, target average acceptance probability for the No-U-Turn Sample of 0.999 and tree depth of 15. The MCMC convergence was evaluated using Rhat and Bulk and Tail ESS values (Table S3). To compare alpha diversity between sampling methods, we fitted a Bayesian regression model with order richness (i.e., the number of insect orders per sampling event) as the response variable, sampling method as a fixed effect, and sample nested within site as random effects, using default priors. To assess sampling coverage, we computed rarefaction and extrapolation curves of taxonomic order richness using the iNEXT package (Chao et al., 2014). To compare the community composition at the order level between sampling methods, we

160 computed the Jaccard Index using the function *vegdist* from the package *vegan* (Dixon, 2003). To
161 differentiate between the nestedness and turnover components of the Jaccard Index, we used the
162 function *beta.pair* from the *betapart* package (Baselga & Orme, 2012). We then computed the
163 mean Jaccard Index value and percentage of nestedness component by fitting an intercept-only
164 BRM for each response variable, with site as random effect and default priors. All BRMs were
165 fitted with weakly informative priors for all parameters (Table S4) for higher stability, no
166 divergences, and reducing high Pareto k-values, while still allowing the data to inform the posterior
167 (Gelman, 2006; Gelman et al., 2008).

168 We used Spearman rank correlations to test associations between (i) detection probabilities
169 derived from the two sampling methods, (ii) insect abundance from Malaise traps, and (iii)
170 sequencing read counts from PUF-PAS. Detection probabilities by PUF-PAS and Malaise traps,
171 predicted number of sequenced reads and predicted insect count were derived from four BRMs
172 which included presence and absence of insect order by both sampling methods, number of
173 sequenced reads and insect count as a response variables, insect order as explanatory variable
174 and sampling date nested within site as random effects.

175 To characterize potential lag in the relationship between PUF-PAS signals of insect detection with
176 those obtained with the Malaise traps, we considered both lag effects (delay between insect
177 presence and detection) and temporal integration windows (duration over which abundance
178 contributes to detection). We defined a set of candidate lag durations (0, 14, 28, and 42 days) and
179 integration window lengths (14, 28, 42, and 56 days), corresponding to multiples of our bi-weekly
180 sampling interval and allowing integration of signal across successive sampling occasions. To
181 understand whether insect detection with PUF-PAS was driven by recent abundance, lagged

abundance, or by cumulative abundance over extended periods, we fitted a model for each of the 16 temporal windows. Windows requiring data prior to the start of the sampling period were excluded. Using a sliding window approach, we calculated the mean count (log-transformed) of insects identified with the Malaise trap for each site and order, then determined the temporal window that best explains the eDNA signal detected with the PUF-PAS (presence or absence as well as the number of sequence reads of an insect order) using Akaike's Information Criterion (AIC). For both response variables, we used the standardized log-transformed mean Malaise-trap count within each lag–window combination as the predictor, and site and order as random intercepts. For the detection probability in PUF-PAS, we computed a binomial generalized linear mixed-effects model, whereas for the number of sequencing reads, we used a hurdle modelling framework to account for the excess of zero observations and overdispersion in the data. Specifically, we fitted generalized linear mixed-effects models using a truncated negative binomial distribution for positive counts and a separate zero process. Residual diagnostics (e.g., simulated residuals) were examined to assess dispersion, uniformity, and potential model misspecification. Model adequacy was assessed using simulation-based residual diagnostics using the *DHARMa* package (Hartig, 2022). When deviations from model assumptions were detected, we refined the functional form of the predictor by introducing higher-order polynomial terms (e.g., quadratic and cubic).

3. Results

3.1. Richness and composition

Rarefaction curves showed clear differences in taxonomic richness between sampling methods (Figure 2A). Malaise traps consistently detected a higher number of insect orders than PUF-PAS (Figure 2B, Figure S2). For instance, at comparable sampling intensity (48 samples), Malaise traps reached an estimated richness of 14 orders, whereas PUF-PAS reached 12 orders (Figure 2C). The order accumulation curve approached an asymptote for PUF-PAS samples, indicating lower total detectable diversity.

Malaise traps detected an average of 8.03 orders [95% CI 7.47-8.69], while PUF-PAS detected 4.50 [95% CI 3.89-5.07], with an average difference estimate of 3.53 [95% CI: 3.02-4.05].

Jaccard dissimilarity between sampling methods had a mean value of 0.46 [95% CI 0.36-0.56], indicating moderate difference in the community composition. A further beta partition analysis shows that a mean of 98.2 % [95% CI 88.6-100] of the Jaccard Index stems from nestedness, showing that the communities sampled by PUF-PAS represent a subset of those sampled by Malaise traps.

3.2. Relationship between detection probability in PUF-PAS and insect counts

Across the entire season, there was a strong positive correlation in the detection probability of each order between Malaise and PUF-PAS (Spearman correlation: $S = 37.7$, $rho = 0.93$, $P < 0.0001$), and between the detection probability by PUF-PAS and insect counts of each order ($S = 10.6$, $rho = 0.98$, $P < 0.0001$). Malaise traps consistently exhibited higher detection probabilities than PUF-PAS (Table S5).

The relationship between Malaise trap abundance and PUF-PAS detection depended strongly on the temporal lag and integration window used to summarize insect counts. Across all candidate

models, the best-supported models used a 42-day lag with a minor effect of integration window (ΔAIC < 2, Table S6). The model explained a substantial proportion of variation, with a marginal R^2 of 0.49 (conditional R^2 = 0.70). Model support declined sharply for shorter lags (ΔAIC for 28-day lag = 36.78-38.77; ΔAIC for 14-day lag = 88.06-92.78; ΔAIC for 0-day lag = 112.86-126.10, Figure 3a).

3.4. Relationship between number of sequenced reads and insect counts

The number of sequenced reads was strongly positively correlated with insect count (Spearman correlation: S = 61.7, rho = 0.88, P < 0.0001). Orders like Diptera, Hymenoptera, Coleoptera, and Trichoptera were disproportionately represented in the Malaise trap relative to the PUF-PAS (Table S5).

The relationship between Malaise trap abundance and PUF-PAS read counts was also strongly influenced by the temporal lag and integration window used to summarize insect counts. The best-supported model included a 42-day lag and a 56-day or 42-day integration window (ΔAIC < 2, Figure 3b). The top model showed a marginal R^2 of 0.32 (conditional R^2 = 0.72). Model support declined sharply for shorter lags (Table S7).

4. Discussion

In our study, we hypothesized that insect sampling with PUF-PAS and Malaise traps would yield similar community composition. However, our results provide evidence that PUF-PAS sampling captured only a subset of the insect community detected by Malaise traps. More specifically, order richness was consistently higher in Malaise traps across all sites, and the richness accumulation curve reached a plateau in PUF-PAS but not in Malaise traps. The number of DNA sequencing reads detected with the PUF-PAS was strongly positively associated with insect counts recorded in Malaise traps. Interestingly, there was a strong lag effect, with PUF-PAS detection probability best explained by Malaise abundance measured several weeks prior and integrated over extended temporal windows. Here, we discuss potential sources of difference in sampling between the two methods and present advantages and pitfalls of PUF-PAS in insect monitoring.

The insect orders sampled with PUF-PAS were concordant with other airborne eDNA studies that amplified insect DNA (*e.g.,* Craine et al., 2025; Roger et al., 2022; Tournayre et al., 2025), and corresponded to the most abundant and widespread taxa typically sampled in entomological surveys across Canada (Langor, 2019). However, species richness was higher in the Malaise trap (14 vs. 12 orders), and its accumulation curve did not reach an asymptote, contrary to the PUF-PAS, which plateaued. Importantly, some insect orders (Mantodea, Mecoptera, and Plecoptera) were detected by Malaise traps but not by PUF-PAS, suggesting that their absence in PUF-PAS samples is more likely due to methodological limitations rather than geographic or ecological constraints. In contrast, one order (Psocoptera) was detected by PUF-PAS but not by Malaise traps, highlighting a slight advantage of PUF-PAS over Malaise trap for this taxonomic group.

Across the entire season, we found strong positive relationships between detection probabilities obtained by both methods as well as between insect counts from Malaise traps and detection

probabilities from PUF-PAS. This is consistent with the general expectation that higher organismal abundance and biomass increase eDNA production and thus detection probability. However, the non-linear relationship indicates that airborne eDNA was detected only when a threshold of insect abundance was reached. This is likely because airborne eDNA detection is governed by a nonlinear chain of processes, involving DNA shedding (Andruszkiewicz Allan et al., 2021), atmospheric transport (Strickler et al., 2015; Tzafesta & Shokri, 2025), particle capture (Gloor et al., 2017), and molecular amplification (Berelson et al., 2025; Moreno et al., 2026). PUF-PAS seems to behave as a threshold-sensitive sampler that is more likely to detect common taxa.

Our results reveal a strong lagged and temporally integrated relationship between insect abundance and airborne eDNA detection, with the best-supported models indicating a delay of up to six weeks and integration over eight weeks. This finding highlights an important feature of PUF-PAS, that is, rather than reflecting instantaneous community composition, it integrates biological signals over time, similar to the dynamics in aquatic systems (Troth et al., 2021). Airborne eDNA originates from multiple sources, including tissue fragments, carcasses, and fecal material, which may be released or resuspended after peak organismal abundance (Andruszkiewicz Allan et al., 2021). In addition, DNA particles can remain suspended in the atmosphere or be transported across space before deposition, further decoupling detection from real-time local abundance (Garrett et al., 2025). Importantly, although the identification of a "best" lag–window combination provides one explanation for the observed pattern, it also reflects the fact that later-season observations are inherently easier to predict from insect abundance than early-season observations, as larger lag–window combinations rely on a subset of data with sufficient temporal history and reduced early-season variability. To our knowledge,

this is the first study that explicitly shows such pronounced temporal decoupling between insect abundance and airborne eDNA signals. Further work is needed to better understand the mechanisms responsible for these extended lags, such as *residence time, resuspension, and degradation dynamics*.

Study limitations. This study has some limitations that should be considered when interpreting the results. First, the analyses were limited to the order level, which may mask important variation at finer taxonomic scales (e.g., family or species). Nevertheless, meaningful differences among sampling methods were still detected at this taxonomic level, indicating that order-level resolution can capture broad community patterns. Secondly, we used a single primer set that, although has good coverage for insects, may miss certain taxa. Despite these limitations, our approach allowed us to evaluate the potential of existing PUF-PAS monitoring networks for airborne eDNA applications, highlighting both their limitations and their promise for large-scale biodiversity surveillance.

5. Future directions: advancing airborne eDNA monitoring of insects

This study provides a foundation for integrating passive airborne eDNA sampling into insect biodiversity monitoring frameworks. By demonstrating that PUF-PAS can reliably detect widespread and common insect orders, our results highlight the potential to leverage existing passive air sampling networks, originally developed for contaminant monitoring, for large-scale biodiversity surveillance.

A key finding of this study is that airborne eDNA signals are delayed relative to local biological dynamics, which has important implications for their interpretation. From a monitoring perspective, this temporal integration suggests that PUF-PAS may be particularly well suited for capturing broad temporal trends, such as interannual variation in insect occurrence and relative abundance, where delayed signals are less problematic. In contrast, the method may be less appropriate for resolving fine-scale seasonal dynamics, unless the mechanisms underlying the lag (e.g., atmospheric residence time, resuspension, and DNA degradation) are better understood and explicitly accounted for (Barnes & Turner, 2016; Garrett et al., 2025).

The current design of the PUF-PAS can be enhanced to improve eDNA sampling performance. Optimizing filter materials and particle capture efficiency, as well as systematically comparing passive and active air samplers under standardized conditions, will help identify approaches that maximize taxonomic recovery. In particular, active air samplers have shown higher detection efficiency than passive systems (Newton et al., 2026), with the potential to approach the performance of conventional trapping methods such as Malaise traps (Craine et al., 2025) Refining deployment strategies, including shorter sampling intervals, will also be essential to reduce temporal averaging and better resolve temporal dynamics in eDNA signals. Beyond

methodological improvements, accounting for atmospheric transport and DNA decay will be necessary to translate eDNA signals into ecologically meaningful metrics. As these developments progress, passive airborne eDNA sampling has the potential to become a powerful component of global biodiversity monitoring systems, particularly for detecting broad patterns of insect occurrence across large spatial and temporal scales.

Acknowledgements

We thank Dr. Elizabeth Clare (York University) for offering helpful advice on the project and Tom Harner (Environment and Climate Change Canada) for providing the PUF-PAS and the GAPS world map. We are thankful to Marie Deumeland, Alin Buruiana and Alma Karina Moncivais Gómez and other volunteers for support in field sampling. Thanks to Sépaq for providing a permit to conduct our research (Authorization No. PNMSB 2024-01). This study was funded by an NSERC CRC Tier 2 (CRC-2022-00134), NSERC Discovery Grant (RGPIN-2024-04564), and NSERC Alliance (ALLRP 597482-2024).

Authors' contributions

J.V.: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing; L.R.A.V.: Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing; X.Z.: Conceptualization, Methodology, Resources, Writing – review & editing; J.J.M.: Funding acquisition, Writing – review & editing; R.K.: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing

Data availability statement

The data and code used in this study are available from the Figshare Repository: https://figshare.com/s/01a97cef04a571adfeea.

The raw sequences can be found in the NCBI-SRA Repository: https://dataview.ncbi.nlm.nih.gov/object/PRJNA1436621?reviewer=8a49g31plkjmm5jotqhjac0kd2

Conflict of interest declaration

The authors declare that they have no competing interests.

Figures

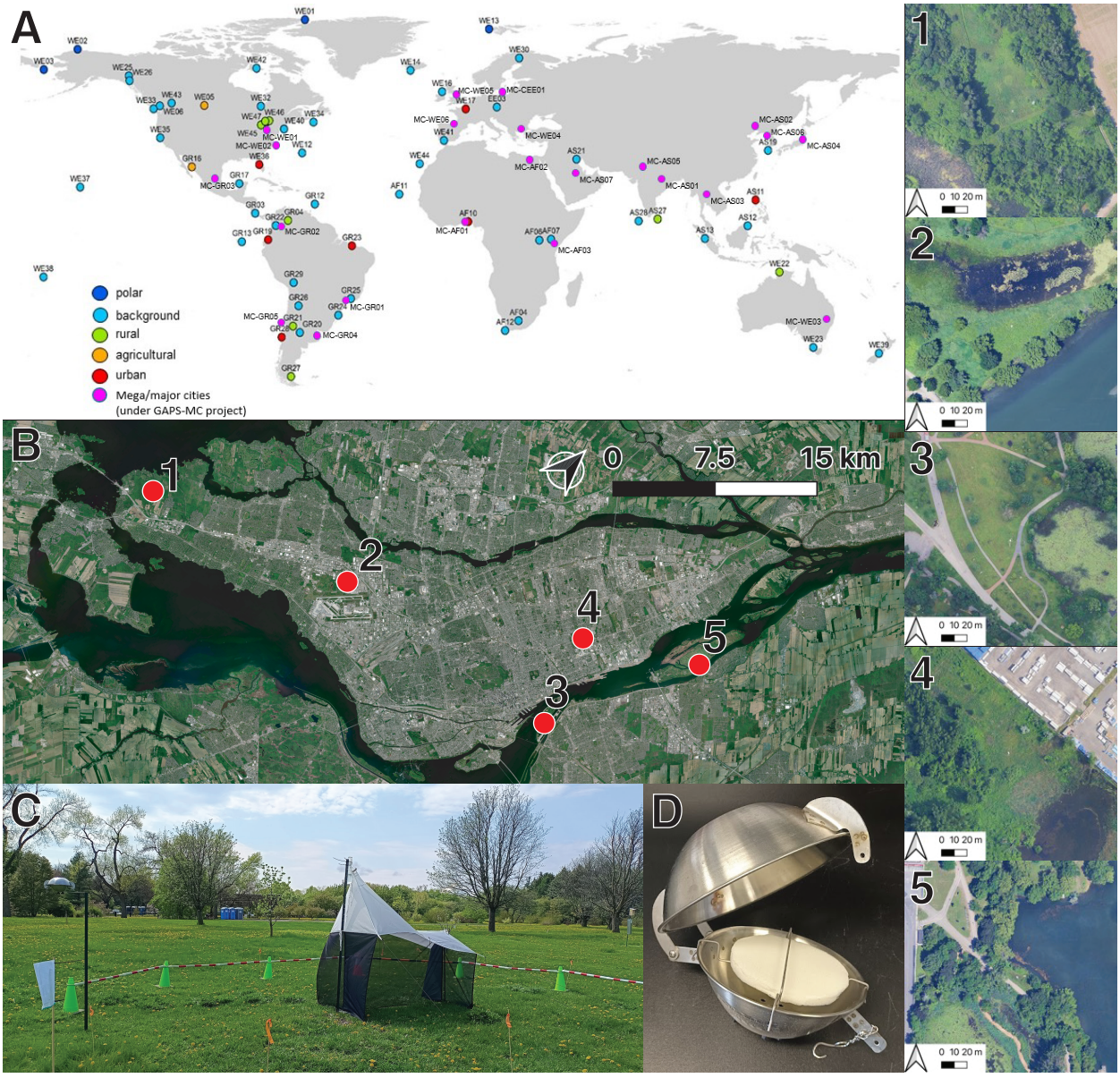
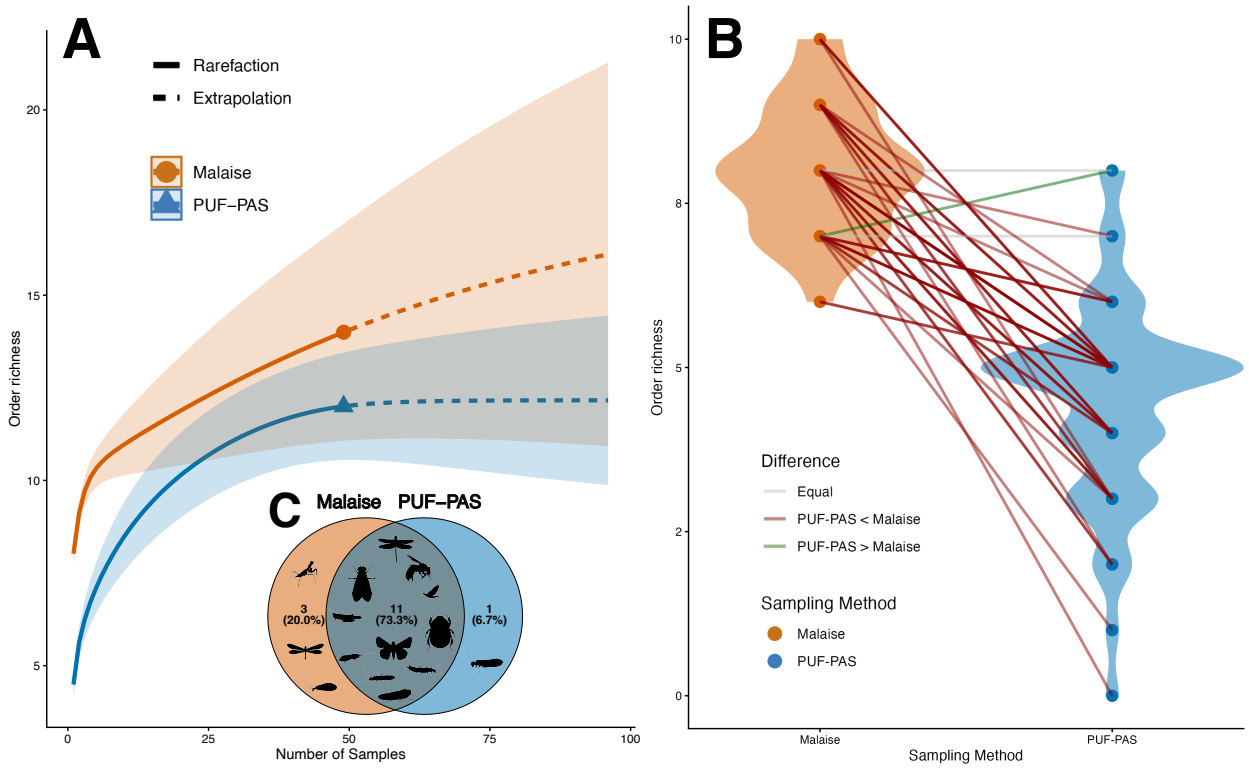


Figure 1. Sampling methods and locations. (A) Global Atmospheric Passive Sampling Network (GAPS) locations. (B) Study area indicating the five sampling sites, (1) McGill Bird Observatory, (2) Parc National des Îles-de-Boucherville, (3) Jardin Botanique de Montréal, (4) Marais IPEX, (5) Parc Jean-Drapeau. (C) Malaise trap with PUF-PAS. (D) PUF-PAS showing the position of the polyurethane disk. GAPS map provided by Tom Harner.



363

364 **Figure 2.** Order richness for each sampling method. (A) Rarefaction (solid lines) and extrapolation
365 (dashed lines) curves of insect order richness for Malaise traps (orange) and PUF-PAS (blue).
366 Shaded areas represent 95% credible intervals. (B) Order richness for each sampling method. Lines
367 connect the same sampling event, representing higher order richness in Malaise traps (red), same
368 order richness sampled (gray), and higher order richness in PUF-PAS (green). The orders
369 *Mantodea*, *Mecoptera*, and *Plecoptera* were detected in the Malaise traps only, whereas
370 *Psocoptera* was detected only in the PUF-PAS. (C) Venn diagram with the communities sampled
371 by each method, with silhouettes of the orders and percentages of convergence and divergence.
372 Silhouettes were downloaded from PhyloPic.

373

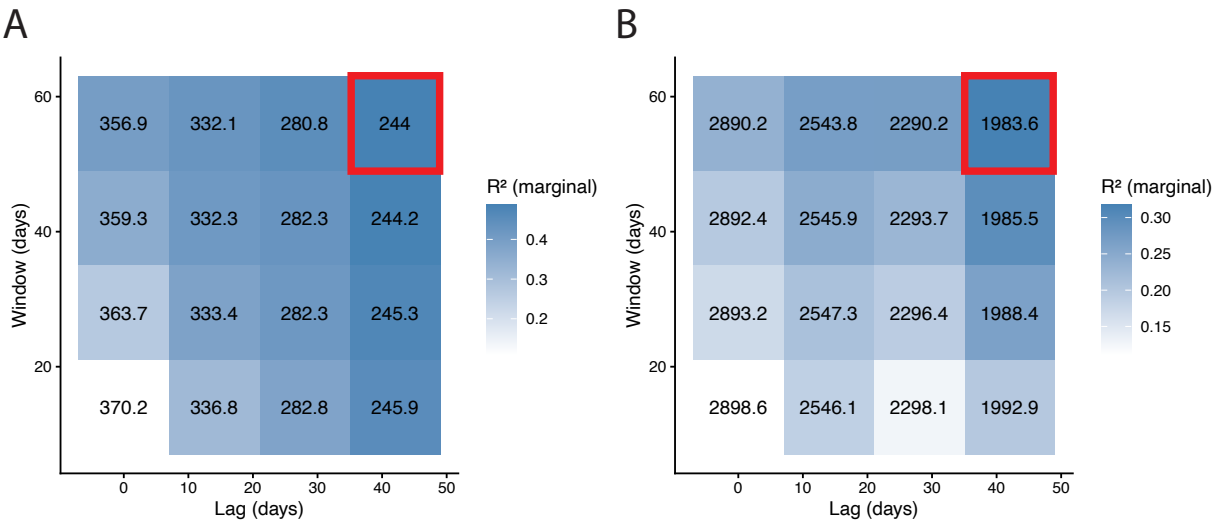


Figure 3. Model performance across temporal lag–window combinations for (a) insect detection probability and (b) eDNA read abundance measured by PUF-PAS, each explained by insect counts from Malaise traps. Colors represent the proportion of variance explained (R^2), while values within cells indicate Akaike Information Criterion (AIC). Temporal windows vary in both duration and lag relative to the sampling date. The red square highlights the model with the lowest AIC, indicating the best-supported model.

References

Andruszkiewicz Allan, E., Zhang, W. G., C. Lavery, A., & F. Govindarajan, A. (2021). Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. *Environmental DNA*, *3*(2), 492–514. https://doi.org/10.1002/edn3.141

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, *17*(1), 1–17. https://doi.org/10.1007/s10592-015-0775-4

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

Berelson, M. F. G., Heavens, D., Nicholson, P., Clark, M. D., & Leggett, R. M. (2025). From air to insight: The evolution of airborne DNA sequencing technologies. *Microbiology*, *171*(5), 001564. https://doi.org/10.1099/mic.0.001564

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bürkner, P.-C. (2017). **brms**: An *R* Package for Bayesian Multilevel Models Using *Stan*. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., & Ellison, A. M. (2014). Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs*, *84*(1), 45–67. https://doi.org/10.1890/13-0133.1

Craine, J. M., Schulte, N., Devitt, J., Leopold, D., Saltonstall, K., & Fierer, N. (2025). Surveying tropical faunal diversity via airborne DNA analyses. *Scientific Reports*, *15*(1), 36701. https://doi.org/10.1038/s41598-025-20767-3

Dalton, R. M., Underwood, N. C., Inouye, D. W., Soulé, M. E., & Inouye, B. D. (2023). Long-term declines in insect abundance and biomass in a subalpine habitat. *Ecosphere*, *14*(8), e4620. https://doi.org/10.1002/ecs2.4620

De Sousa, L. L., Silva, S. M., & Xavier, R. (2019). DNA metabarcoding in diet studies: Unveiling ecological aspects in aquatic and terrestrial ecosystems. *Environmental DNA*, *1*(3), 199–214. https://doi.org/10.1002/edn3.27

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*(2), 103–114. https://doi.org/10.1111/icad.12408

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, *14*(6), 927–930. https://doi.org/10.1111/j.1654-1103.2003.tb02228.x

Drake, L. E., Cuff, J. P., Young, R. E., Marchbank, A., Chadwick, E. A., & Symondson, W. O. C. (2022). An assessment of minimum sequence copy thresholds for identifying and

reducing the prevalence of artefacts in dietary metabarcoding data. *Methods in Ecology and Evolution*, *13*(3), 694–710. https://doi.org/10.1111/2041-210X.13780

Garrett, N. R., Tournayre, O. R., Littlefair, J. E., Ivanova, N. V., Mei, G., Jedrecka, T., Briscoe, A., Naaum, A., Simmons, N. B., & Clare, E. (2025). *Sampling intensity and temporal persistence of airborne eDNA in partially enclosed spaces* (p. 2025.07.14.664745). bioRxiv. https://doi.org/10.1101/2025.07.14.664745

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, *1*(3). https://doi.org/10.1214/06-BA117A

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, *2*(4). https://doi.org/10.1214/08-AOAS191

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., & Egozcue, J. J. (2017). Microbiome Datasets Are Compositional: And This Is Not Optional. *Frontiers in Microbiology*, *8*. https://doi.org/10.3389/fmicb.2017.02224

Halley, J. M., & Pimm, S. L. (2023). The rate of species extinction in declining or fragmented ecological communities. *PLOS ONE*, *18*(7), e0285945. https://doi.org/10.1371/journal.pone.0285945

Hallmann, C. A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T., Goulson, D., & de Kroon, H. (2017). More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLOS ONE*, *12*(10), e0185809. https://doi.org/10.1371/journal.pone.0185809

449 Harner, T., Saini, A., Shahpoury, P., Eng, A., Schuster, J. K., Kalisa, E., & Mastin, J. (2024).
450 Cross-cutting research and future directions under the GAPS networks. *Environmental*
451 *Science: Advances*, *3*(6), 798–807. https://doi.org/10.1039/D4VA00034J

452 Hartig, F. (2022). *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed)*
453 *Regression Models* (Version 0.4.5) [R]. http://florianhartig.github.io/DHARMa/

454 Harvey, J. A., Tougeron, K., Gols, R., Heinen, R., Abarca, M., Abram, P. K., Basset, Y., Berg,
455 M., Boggs, C., Brodeur, J., Cardoso, P., de Boer, J. G., De Snoo, G. R., Deacon, C., Dell,
456 J. E., Desneux, N., Dillon, M. E., Duffy, G. A., Dyer, L. A., … Chown, S. L. (2022).
457 Scientists' warning on climate change and insects. *Ecological Monographs*.
458 https://doi.org/10.1002/ecm.1553

459 Kalisa, E., Saini, A., Lee, K., Mastin, J., Schuster, J. K., & Harner, T. (2024). Capturing the
460 Aerobiome: Application of Polyurethane Foam Disk Passive Samplers for Bioaerosol
461 Monitoring. *ACS ES&T Air*, *1*(5), 414–425. https://doi.org/10.1021/acsestair.3c00107

462 Liu, Q., Li, L., Zhang, X., Saini, A., Li, W., Hung, H., Hao, C., Li, K., Lee, P., Wentzell, J. J. B.,
463 Huo, C., Li, S.-M., Harner, T., & Liggio, J. (2021). Uncovering global-scale risks from
464 commercial chemicals in air. *Nature*, *600*(7889), 456–461.
465 https://doi.org/10.1038/s41586-021-04134-6

466 Malaise, R. (1937). A new insect-trap. *Entomologisk Tidskrift*, *58*, 148–160.

467 Marinchel, N., Marchesini, A., Nardi, D., Girardi, M., Casabianca, S., Vernesi, C., & Penna, A.
468 (2023). Mock community experiments can inform on the reliability of eDNA
469 metabarcoding data: A case study on marine phytoplankton. *Scientific Reports*, *13*,
470 20164. https://doi.org/10.1038/s41598-023-47462-5

Meglécz, E. (2023). COInr and mkCOInr: Building and customizing a nonredundant barcoding reference database from BOLD and NCBI using a semi-automated pipeline. *Molecular Ecology Resources*, *23*(4), 933–945. https://doi.org/10.1111/1755-0998.13756

Moreno, J., Jackson, N., & Frere, C. H. (2026). *A comparison of passive airDNA collection and extraction methods on the recovered terrestrial vertebrate assemblage*. Preprints. https://doi.org/10.22541/au.177153466.61633211/v1

Morrill, A., Kaunisto, K. M., Mlynarek, J. J., Sippola, E., Vesterinen, E. J., & Forbes, M. R. (2021). Metabarcoding prey DNA from fecal samples of adult dragonflies shows no predicted sex differences, and substantial inter-individual variation, in diets. *PeerJ*, *9*, e12634. https://doi.org/10.7717/peerj.12634

Newton, J. P., Allentoft, M. E., Bateman, P. W., & Nevill, P. (2026). Airborne DNA and Spider Webs Outperform Other eDNA Sources for Monitoring Terrestrial Vertebrates. *Molecular Ecology Resources*, *26*(1), e70067. https://doi.org/10.1111/1755-0998.70067

Nousias, O., McCauley, M., Stammnitz, M. R., Farrell, J. A., Koda, S. A., Summers, V., Eastman, C. B., Duffy, F. G., Duffy, I. J., Whilde, J., & Duffy, D. J. (2025). Shotgun sequencing of airborne eDNA achieves rapid assessment of whole biomes, population genetics and genomic variation. *Nature Ecology & Evolution*, *9*(6), 1043–1060. https://doi.org/10.1038/s41559-025-02711-w

Pozo, K., Harner, T., Wania, F., Muir, D. C. G., Jones, K. C., & Barrie, L. A. (2006). Toward a Global Network for Persistent Organic Pollutants in Air: Results from the GAPS Study. *Environmental Science & Technology*, *40*(16), 4867–4873. https://doi.org/10.1021/es060447t

R Core Team. (2026). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Roger, F., Ghanavi, H. R., Danielsson, N., Wahlberg, N., Löndahl, J., Pettersson, L. B., Andersson, G. K. S., Boke Olén, N., & Clough, Y. (2022). Airborne environmental DNA metabarcoding for the monitoring of terrestrial insects—A proof of concept from the field. *Environmental DNA*, *4*(4), 790–807. https://doi.org/10.1002/edn3.290

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., & Eliceiri, K. W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, *18*(1), 529. https://doi.org/10.1186/s12859-017-1934-z

Saini, A., Harner, T., Chinnadhurai, S., Schuster, J. K., Yates, A., Sweetman, A., Aristizabal-Zuluaga, B. H., Jiménez, B., Manzano, C. A., Gaga, E. O., Stevenson, G., Falandysz, J., Ma, J., Miglioranza, K. S. B., Kannan, K., Tominaga, M., Jariyasopit, N., Rojas, N. Y., Amador-Muñoz, O., … Shoeib, T. (2020). GAPS-megacities: A new global platform for investigating persistent organic pollutants and chemicals of emerging concern in urban air. *Environmental Pollution*, *267*, 115416. https://doi.org/10.1016/j.envpol.2020.115416

Sanchez, D. E., Walker, F. M., Marriott, S. J., Riley, A. L., Stankavich, S., Adams, A. M., Solick, D., Bradley, D., & Newman, C. (2025). Out in the Open: Investigating Passive Airborne EDNA Detection of Bats at Artificial Feeding Stations. *Environmental DNA*, *7*(3), e70108. https://doi.org/10.1002/edn3.70108

Santos, J. C., & Fernandes, G. W. (Eds.). (2021). *Measuring Arthropod Biodiversity: A Handbook of Sampling Methods*. Springer International Publishing. https://doi.org/10.1007/978-3-030-53226-0

Schuster, J. K., Harner, T., Fillmann, G., Ahrens, L., Altamirano, J. C., Aristizábal, B., Bastos, W., Castillo, L. E., Cortés, J., Fentanes, O., Gusev, A., Hernandez, M., Ibarra, M. V., Lana, N. B., Lee, S. C., Martínez, A. P., Miglioranza, K. S. B., Puerta, A. P., Segovia, F., … Tominaga, M. Y. (2015). Assessing Polychlorinated Dibenzo- *p* -dioxins and Polychlorinated Dibenzofurans in Air across Latin American Countries Using Polyurethane Foam Disk Passive Air Samplers. *Environmental Science & Technology*, *49*(6), 3680–3686. https://doi.org/10.1021/es506071n

Semenov, M. V. (2021). Metabarcoding and Metagenomics in Soil Ecology Research: Achievements, Challenges, and Prospects. *Biology Bulletin Reviews*, *11*(1), 40–53. https://doi.org/10.1134/S2079086421010084

Shoeib, M., & Harner, T. (2002). Characterization and Comparison of Three Passive Air Samplers for Persistent Organic Pollutants. *Environmental Science & Technology*, *36*(19), 4142–4151. https://doi.org/10.1021/es020635t

Smith, K. F., Kohli, G. S., Murray, S. A., & Rhodes, L. L. (2017). Assessment of the metabarcoding approach for community analysis of benthic-epiphytic dinoflagellates using mock communities. *New Zealand Journal of Marine and Freshwater Research*, *51*(4), 555–576. https://doi.org/10.1080/00288330.2017.1298632

Strickler, K. M., Fremier, A. K., & Goldberg, C. S. (2015). Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biological Conservation*, *183*, 85–92. https://doi.org/10.1016/j.biocon.2014.11.038

Sullivan, A. R., Karlsson, E., Svensson, D., Brindefalk, B., Villegas, J. A., Mikko, A., Bellieny, D., Siddique, A. B., Johansson, A.-M., Grahn, H., Sundell, D., Norman, A., Esseen, P.-

A., Sjödin, A., Singh, N. J., Brodin, T., Forsman, M., & Stenberg, P. (2023). *Airborne eDNA captures three decades of ecosystem biodiversity*. Ecology. https://doi.org/10.1101/2023.12.06.569882

Tournayre, O., Littlefair, J. E., Garrett, N. R., Allerton, J. J., Brown, A. S., Cristescu, M. E., & Clare, E. L. (2025). First national survey of terrestrial biodiversity using airborne eDNA. *Scientific Reports*, *15*(1), 19247. https://doi.org/10.1038/s41598-025-03650-z

Troth, C. R., Sweet, M. J., Nightingale, J., & Burian, A. (2021). Seasonality, DNA degradation and spatial heterogeneity as drivers of eDNA detection dynamics. *Science of The Total Environment*, *768*, 144466. https://doi.org/10.1016/j.scitotenv.2020.144466

Tzafesta, E., & Shokri, M. (2025). The combined negative effect of temperature, UV radiation and salinity on eDNA detection: A global *meta*-analysis on aquatic ecosystems. *Ecological Indicators*, *176*, 113669. https://doi.org/10.1016/j.ecolind.2025.113669

Van Klink, R., Bowler, D. E., Gongalsky, K. B., Shen, M., Swengel, S. R., & Chase, J. M. (2024). Disproportionate declines of formerly abundant species underlie insect loss. *Nature*, *628*(8007), 359–364. https://doi.org/10.1038/s41586-023-06861-4

Yamahara, K. M., Allan, E. A., Robidart, J., Wilson, W. H., Birch, J. M., Craw, P., Edson, E., Engstrom, I. B., Fukuba, T., Govindarajan, A. F., Martins, A. M., Parsons, K. M., Sieben, V. J., Thomas, A., Wilson, I., & Scholin, C. A. (2025). A State-Of-The-Art Review of Aquatic eDNA Sampling Technologies and Instrumentation: Advancements, Challenges, and Future Prospects. *Environmental DNA*, *7*(4), e70170. https://doi.org/10.1002/edn3.70170

Yates, M. C., Wilcox, T. M., Kay, S., Peres-Neto, P., & Heath, D. D. (2025). A Framework to Unify the Relationship Between Numerical Abundance, Biomass, and Environmental DNA. *Environmental DNA*, *7*(2), e70073. https://doi.org/10.1002/edn3.70073

Supplementary materials

Text S1: PUF-PAS analysis

DNA extractions

We performed DNA extractions, PCR, indexing, and sequencing at the Genomics platform of the CERMO-FC Center of UQÀM (Quebec, Canada). Following a previously published work of metabarcoding analyses of PUF-PAS (Kalisa et al. 2024), we extracted metagenomic DNA from the PUF discs using the DNeasy PowerMax Soil Kit (Qiagen, Venlo, Limburg, The Netherlands) with a slightly modified version of the manufacturer instructions. Specifically, we employed each whole PUF disc as substrate during the lysis step. Each PUF disc was allowed to absorb the lysis solution mix (15 mL of PowerBead Solution + 1.2 mL of C1 Solution) inside the PowerMax Bead Pro tubes. After digestion, the supernatant was extracted from the PUF using the centrifugation step (2500 g × 3 min), i.e., the supernatant flowed to the bottom of the tube. All following steps were performed as described in the manufacturer's instructions. After extractions, DNA concentrations were measured across all samples, as well as the positive and negative controls, using a Qubit fluorometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

To evaluate possible contamination during DNA extractions, we included two negative controls consisting of autoclaved PUF discs. Additionally, since the inclusion of mock communities has been considered a standard control in metabarcoding studies to reveal biases and test bioinformatic pipelines (Smith et al. 2017; Marinchel et al. 2023), we included two replicated PUF discs that we used to sample our mock community. This mock community consisted of eight exotic species that are not found naturally in Canada and whose samples were donated by the Montreal Insectarium. These species consisted of the lepidopterans *Caligo eurilochus*,

Cethosia biblis, *Heliconius doris*, *Papilio demoleus, Papilio lowii*, *Papilio rumanzovia*, and *Parthenos sylvia* and the mantid *Hymenopus coronatus*. We performed five smears across the two positive PUF discs using one frozen-and-thawed individual of each species.

PCR amplification and sequencing

We used PCR to amplify a 313 bp fragment of the COI gene using the "Leray" set of primers for insect metabarcoding (Morrill et al. 2021). This pair of primers consisted of the forward primer mlCOIintF: 5'-GGWACWGGWTGAACWGTWTAYCCYCC-3' (Leray et al. 2013) and the reverse primer jgHCO2198: 5'-TAIACYTCIGGRTGICCRAARAAYCA-3' (Geller et al. 2013). Both primers were modified to include standard Nextera adapter sequences. We performed PCR using 1X of BioBasic Inc. Taq PCR Master Mix, 10 mM of dNTPs, 0.4 µM of each primer, 0.5U of BioBasic Inc. Taq DNA polymerase, and 4 µL of extracted DNA. Thermal conditions during PCR consisted of 95 ºC for 2 min, followed by 35 cycles of 95 ºC for 20 s, 46 ºC for 30 s, and 72 ºC for 30 s, with a final extension period of 72 ºC for 5 min. We included two negative PCR controls using DNA-free mixes to assess contamination during the PCR steps. Then, we performed a second PCR to add Illumina sequencer adapters with dual indexes for sample demultiplexing. PCR products were assessed on 2% agarose gels and normalized using the Just-a-plate PCR normalization and purification kit (CharmBiotech, San Diego, California, USA). We pooled the samples across two libraries (Table S2) and purified each of them using the AMPure XP beads kit using the manufacturer protocol (Beckman Coulter, Brea, California, USA). The pooled libraries were quantified using the QubitTM dsDNA HS Assay Kit (Thermo Fisher Scientific) and average fragment sizes were determined using a High sensitivity D1000 ScreenTape on a TapeStation 4200 instrument (Agilent Technologies, Santa Clara, California,

609 USA). Sequencing was performed on a Miseq using a MiSeq reagent kit v3 (600-cycles;
610 Illumina, San Diego, California, USA). Before sequencing, Phix control library (Illumina) was
611 spiked into each library to improve unbalanced base composition.
612
613 *Reference database*
614 We employed the COInr database as our reference database for bioinformatic processing
615 (Meglécz 2023). The COInr database is a freely available, comprehensive and dereplicated
616 database of COI sequences extracted both from BOLD (Ratnasingham et al. 2024) and GenBank
617 (Benson et al. 2012). We downloaded the 2025-05-23 version of COInr
618 (https://zenodo.org/records/15515860). Then, we used the *select_region.pl* perl script found
619 within the mkCOInr-0.5.0 tool (Meglécz 2023) to perform *in silico* PCR on the whole database
620 for the Leray set of primers. This allowed us to recover a final dataset consisting of 5,748,708
621 unique taxonomically identified sequences for the specific COI region that we amplified.
622
623 *Bioinformatic analyses*
624 We performed all bioinformatic processing within the QIIME2 v2025.4.0 environment (Bolyen
625 et al. 2019). Samples were processed independently based on the library at which they were
626 sequenced. Excluding positive and negative controls, we recovered 2,344,638 paired raw reads
627 between the two libraries with an average of 47,850 (SD=28,068) paired reads per sample. First,
628 we employed cutadapt (Martin 2011) to remove primers both from the R1 and R2 reads. Then,
629 we performed quality-filtering and merged the R1 and R2 reads using VSEARCH merge-pairs
630 function (Rognes et al. 2016). We tested multiple combinations of quality score thresholds to
631 perform sequence truncation (*--p-truncqual*; t) and the minimum accepted overlap after

truncation to merge R1 and R2 reads (*--p-minovlen*; l). However, based on the low number of reads per sample, we decided to keep four filtering schemes: i) relaxed filtering (t=0; l=10); ii) moderate filtering (t=20; l=10); iii) strict filtering based on sequence overlap (t=20; l=50); and iv) strict filtering based on quality trimming (t=26; l=10). Additionally, we filtered the database for merged reads within the expected amplicon length (313 bp) and with a maximum number of expected errors in the merged read of 1 *(--p-maxee* 1). Then we denoised the reads of each filtering scheme using the *denoise-other* function in Deblur. Finally, to test an alternative denoising algorithm, we employed the *denoise-paired* function of the DADA2 software (Callahan et al. 2016) by applying it on the primer-filtered reads of both libraries (after cutadapt). We applied the same four filtering stringencies that we kept with VSEARCH. These analyses produced a final number of 16 ASVs × sample matrices (2 libraries × 2 denoising algorithms × 4 quality filtering stringencies). Finally, we performed the taxonomic assignment of each of the 16 ASVs databases at the species level using a naive bayes algorithm that we trained using the *feature-classifier classify-sklearn* function with Scikit-learn (Pedregosa et al. 2011) and our processed COInr database.

ASV filtering

To filter the 16 ASVs databases for contaminations and wrongly assigned reads we implemented a standardized protocol based on minimum number of reads thresholds (Drake et al. 2022). Specifically, first we removed across all samples the ASVs that appeared in any of our four negative controls (2 extraction controls + 2 PCR controls) and that presented a read count equal or below the values seen in the negative controls (Max contamination method). Then, we removed ASVs within each sample that had a relative frequency below 8 stringency thresholds

(0.025%, 0.05%, 0.075%, 0.1%, 0.25%, 0.5%, 0.75% and 1%; Sample % method). For each ASV × sample matrix and filtering scheme we estimated the final number of ASVs, percentage of samples retained, number of correctly identified taxa within the positive controls, number of incorrect taxa within the positive controls, and number of samples that included ASVs assigned to species from the mock community (tag-jumping).

To limit contamination and non-local signals while retaining low-abundance detections, we employed the Deblur denoising algorithm with moderate quality filtering (t=20; l=10) and negative control cleaning, but no filtering by relative frequency threshold (Supplementary data 1). Additionally, further analyses were performed after limiting the taxonomic identification depth of the selected filtered ASV × sample matrix to the order level to compare it to the Malaise traps, which further limited taxonomic misidentification (Somervuo et al., 2017; Sotnikov et al., 2025) and the number of false positives.

680 Table S1: Geographic coordinates of the study sites where Malaise traps and PUF-PAS were

681 placed.

Site	Latitude	Longitude
McGill Bird Observatory	45.42908	-73.93803
Technoparc IPEX	45.48299	-73.74932
Parc Jean-Drapeau	45.50213	-73.52686
Jardin Botanique de Montreal	45.55881	-73.56339
Parc national des Îles-de-Boucherville	45.60498	-73.46430

682

683

684 Table S2: PUF-PAS metadata with unique ID, site, date of sampling start, date of sampling end

685 and sequencing library

ID	site	from	to	library
PUF_BOT_3	Botanical Garden	14/06/2024	28/06/2024	first library
PUF_BOT_10	Botanical Garden	20/09/2024	04/10/2024	first library
PUF_BOU_7	Boucherville	19/08/2024	02/09/2024	first library
PUF_BOU_9	Boucherville	16/09/2024	30/09/2024	first library
PUF_IPEX_1	IPEX	15/05/2024	29/05/2024	first library
PUF_IPEX_3	IPEX	12/06/2024	26/06/2024	first library
PUF_IPEX_8	IPEX	21/08/2024	04/09/2024	first library
PUF_JD_4	Jean-Drapeau	03/07/2024	17/07/2024	first library
PUF_JD_10	Jean-Drapeau	25/09/2024	09/10/2024	first library
PUF_MBO_4	Bird Observatory	01/07/2024	15/07/2024	first library
PUF_BOT_1	Botanical Garden	17/05/2024	31/05/2024	second library
PUF_BOT_2	Botanical Garden	31/05/2024	14/06/2024	second library
PUF_BOT_4	Botanical Garden	28/06/2024	12/07/2024	second library
PUF_BOT_5	Botanical Garden	12/07/2024	26/07/2024	second library
PUF_BOT_6	Botanical Garden	26/07/2024	09/08/2024	second library
PUF_BOT_7	Botanical Garden	09/08/2024	23/08/2024	second library
PUF_BOT_8	Botanical Garden	23/08/2024	06/09/2024	second library
PUF_BOT_9	Botanical Garden	06/09/2024	20/09/2024	second library
PUF_BOU_1	Boucherville	27/05/2024	10/06/2024	second library
PUF_BOU_2	Boucherville	10/06/2024	24/06/2024	second library
PUF_BOU_3	Boucherville	24/06/2024	08/07/2024	second library
PUF_BOU_4	Boucherville	08/07/2024	22/07/2024	second library
PUF_BOU_5	Boucherville	22/07/2024	05/08/2024	second library
PUF_BOU_6	Boucherville	05/08/2024	19/08/2024	second library
PUF_BOU_8	Boucherville	02/09/2024	16/09/2024	second library
PUF_IPEX_2	IPEX	29/05/2024	12/06/2024	second library
PUF_IPEX_4	IPEX	26/06/2024	10/07/2024	second library
PUF_IPEX_5	IPEX	10/07/2024	24/07/2024	second library
PUF_IPEX_6	IPEX	24/07/2024	07/08/2024	second library
PUF_IPEX_7	IPEX	07/08/2024	21/08/2024	second library
PUF_IPEX_9	IPEX	04/09/2024	18/09/2024	second library
PUF_IPEX_10	IPEX	18/09/2024	02/10/2024	second library
PUF_JD_1	Jean-Drapeau	22/05/2024	05/06/2024	second library
PUF_JD_2	Jean-Drapeau	05/06/2024	19/06/2024	second library

PUF_JD_3	Jean-Drapeau	19/06/2024	03/07/2024	second library
PUF_JD_5	Jean-Drapeau	17/07/2024	31/07/2024	second library
PUF_JD_6	Jean-Drapeau	31/07/2024	14/08/2024	second library
PUF_JD_7	Jean-Drapeau	14/08/2024	28/08/2024	second library
PUF_JD_8	Jean-Drapeau	28/08/2024	11/09/2024	second library
PUF_JD_9	Jean-Drapeau	11/09/2024	25/09/2024	second library
PUF_MBO_1	Bird Observatory	20/05/2024	03/06/2024	second library
PUF_MBO_2	Bird Observatory	03/06/2024	17/06/2024	second library
PUF_MBO_3	Bird Observatory	17/06/2024	01/07/2024	second library
PUF_MBO_5	Bird Observatory	15/07/2024	29/07/2024	second library
PUF_MBO_6	Bird Observatory	29/07/2024	12/08/2024	second library
PUF_MBO_7	Bird Observatory	12/08/2024	26/08/2024	second library
PUF_MBO_8	Bird Observatory	26/08/2024	09/09/2024	second library
PUF_MBO_9	Bird Observatory	09/09/2024	23/09/2024	second library
PUF_MBO_10	Bird Observatory	23/09/2024	07/10/2024	second library

686

687 Table S3: Summary of MCMC convergence diagnostics for all Bayesian Regression Models in

688 this study

Model	Rhat	Min Bulk ESS	Min Tail ESS
log(biomass_mg + 1) ~ log(count + 1) + (1|site/ID)	1.003	2022	2456
Richness ~ source + (1|site/Sample)	1.002	2790	3514
	1.003	1898	2580
Jaccard ~ 1 + (1|site)	1.002	1560	1678
Percentage Nestedness ~ 1 + (1|site)	1.002	1666	1782
Insect Count ~ order + (site/date)	1.003	1280	2393
No. sequenced reads ~ order + (site/date)	1.003	1518	1742
Detection by PUF ~ order + (site/date)	1.003	1903	2392
Detection by Malaise ~ order + (site/date)	1.004	2671	2647

689

690

Table S4: Priors used for the detection of orders by sampling method, detection in PUF-PAS and number of reads models

Model	Family	Parameter	Prior distribution
Detection by PUF ~ order + (site/date)	Bernouilli	Fixed effects intercept	Normal(0, 3)
	Bernouilli	Fixed effects slopes	Normal(0, 2)
	Bernouilli	Group-level SDs	Student-t(3, 0, 1)
Detection by Malaise ~ order + (site/date)	Bernouilli	Fixed effects intercept	Normal(0, 3)
	Bernouilli	Fixed effects slopes	Normal(0, 2)
	Bernouilli	Group-level SDs	Student-t(3, 0, 1)
# sequenced reads ~ order + (site/date)	Gaussian	Fixed effects intercept	Normal(0, 10)
	Gaussian	Fixed effects slopes	Normal(0, 10)
	Gaussian	Group-level SDs	Student-t(3, 0, 5)
	Gaussian	Residual SD	Student-t(3, 0, 5)
Insect Count ~ order + (site/date)	Gaussian	Fixed effects intercept	Normal(0, 10)
	Gaussian	Fixed effects slopes	Normal(0, 10)
	Gaussian	Group-level SDs	Student-t(3, 0, 5)
	Gaussian	Residual SD	Student-t(3, 0, 5)

Table S5. Detection probabilities in malaise traps and PUF-PAS, mean number of sequenced reads and mean insect count, with 95% Bayesian credible intervals, per order

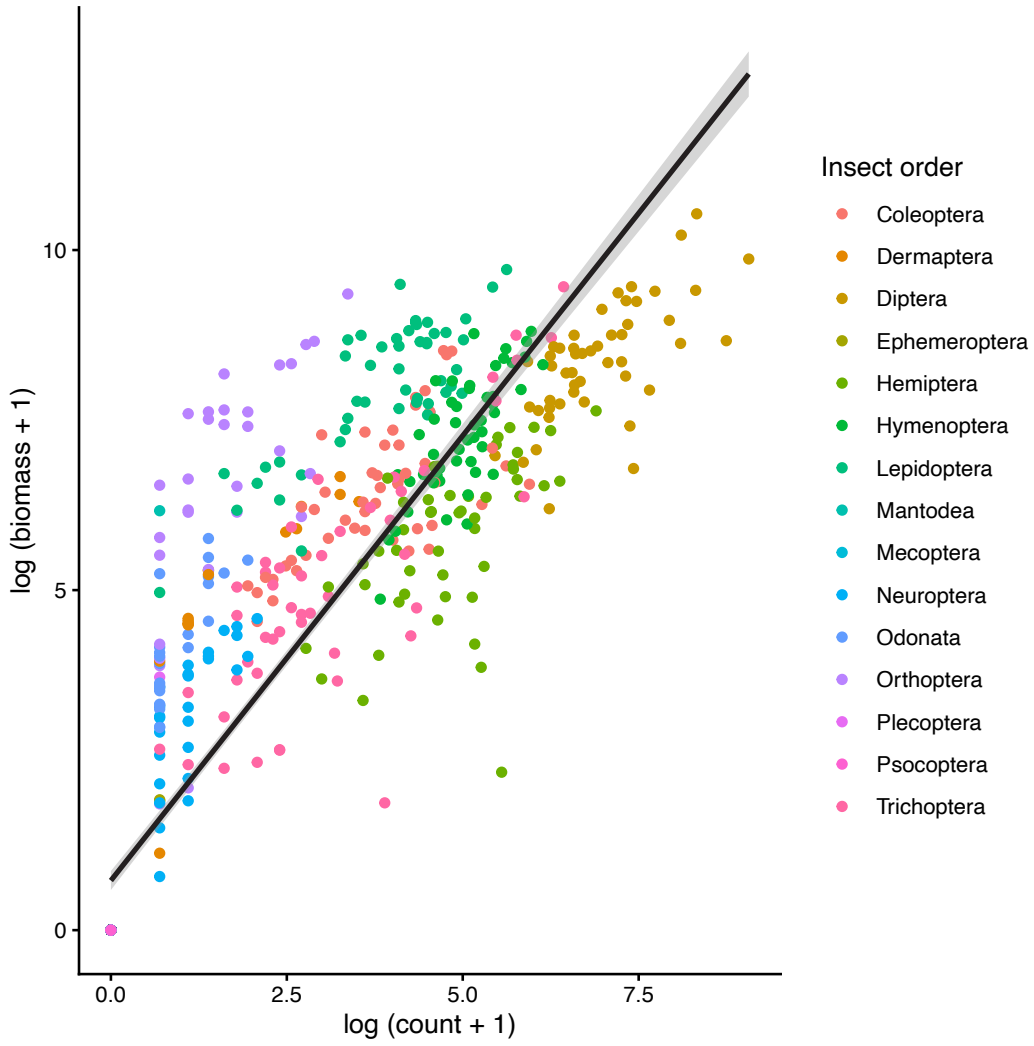
	Malaise				PUF-PAS			
	Probability of detection		Insect count		Probability of detection		Number of reads	
Order	**mean**	**CI**	**mean**	**CI**	**mean**	**CI**	**mean**	**CI**
Coleoptera	0.94	[0.86, 0.99]	41.57	[32.04, 53.56]	0.42	[0.21, 0.67]	2.15	[0.99, 3.82]
Dermaptera	0.32	[0.15, 0.49]	0.81	[0.38, 1.32]	0.02	[0, 0.08]	0.08	[-0.31, 0.71]
Diptera	1	[0.97, 1]	1036.41	[789.86, 1317.49]	0.79	[0.61, 0.94]	11.92	[7.41, 19.31]
Ephemeroptera	0.04	[0, 0.11]	0.03	[-0.2, 0.34]	0.01	[0, 0.05]	0.06	[-0.31, 0.65]
Hemiptera	1	[0.97, 1]	124.82	[97.18, 163.73]	0.99	[0.95, 1]	478.07	[309.24, 749.33]
Hymenoptera	1	[0.97, 1]	150.23	[115.98, 194.03]	0.91	[0.8, 0.99]	10.6	[6.66, 17.38]
Lepidoptera	1	[0.97, 1]	47.48	[36.65, 61.44]	0.96	[0.88, 1]	48.41	[30.84, 75.78]
Mantodea	0.03	[0, 0.08]	0.01	[-0.21, 0.32]	0	[0, 0.03]	0	[-0.35, 0.56]
Mecoptera	0.03	[0, 0.08]	0.01	[-0.21, 0.3]	0	[0, 0.03]	0	[-0.35, 0.58]
Neuroptera	0.54	[0.34, 0.72]	0.85	[0.42, 1.4]	0.04	[0, 0.11]	0.23	[-0.22, 0.88]
Odonata	0.47	[0.29, 0.67]	0.58	[0.21, 1.03]	0.02	[0, 0.08]	0.12	[-0.28, 0.76]
Orthoptera	0.62	[0.42, 0.79]	1.7	[1.08, 2.5]	0.08	[0.01, 0.2]	0.28	[-0.19, 0.98]
Plecoptera	0.03	[0, 0.08]	0.02	[-0.23, 0.3]	0	[0, 0.03]	0	[-0.35, 0.56]
Psocoptera	0.01	[0, 0.05]	0	[-0.23, 0.28]	0.02	[0, 0.08]	0.16	[-0.26, 0.79]
Trichoptera	0.98	[0.93, 1]	20.88	[16.05, 27.22]	0.13	[0.03, 0.3]	0.72	[0.13, 1.76]

Table S6. Model comparison of generalized linear mixed models relating PUF-PAS detection probability (presence/absence of insect orders) to insect abundance measured by Malaise traps across combinations of temporal lags and integration windows. Lag (days) represents the delay between insect occurrence and eDNA detection, while window (days) defines the duration over which insect counts were aggregated. Models are ranked by Akaike Information Criterion (AIC), with ΔAIC indicating differences relative to the best-supported model. β denotes the fixed effect of insect count, and marginal (R^2m) and conditional (R^2c) coefficients of determination quantify variance explained by fixed effects alone and by both fixed and random effects, respectively. n indicates the number of observations.

Window code	Lag (days)	Window (days)	n	AIC	beta	R^2m	R^2c	ΔAIC
detection_ window15	42	56	435	244.04	2.30	0.49	0.70	0.0
detection_window11	42	42	435	244.17	2.27	0.49	0.69	0.1
detection_window7	42	28	435	245.26	2.19	0.47	0.68	1.2
detection_window3	42	14	435	245.92	2.11	0.45	0.67	1.9
detection_window14	28	56	510	280.82	2.20	0.45	0.69	36.8
detection_window10	28	42	510	282.26	2.11	0.43	0.68	38.2
detection_window6	28	28	510	282.27	2.06	0.42	0.68	38.2
detection_window2	28	14	510	282.82	1.95	0.38	0.67	38.8
detection_window13	14	56	585	332.10	2.06	0.43	0.67	88.1
detection_window9	14	42	585	332.26	2.01	0.41	0.66	88.2
detection_window5	14	28	585	333.44	1.92	0.39	0.65	89.4
detection_window1	14	14	585	336.82	1.69	0.32	0.64	92.8
detection_window12	0	56	660	356.91	2.02	0.40	0.68	112.9
detection_window8	0	42	660	359.28	1.88	0.36	0.67	115.2
detection_window4	0	28	660	363.70	1.60	0.27	0.66	119.7
detection_window	0	14	660	370.14	1.06	0.11	0.67	126.1

Table S7. Model comparison of generalized linear mixed models relating sequencing read counts from PUF-PAS to insect abundance measured by Malaise traps across combinations of temporal lags and integration windows. Lag (days) represents the delay between insect occurrence and eDNA signal, while window (days) defines the duration over which insect counts were aggregated. Models are ranked by Akaike Information Criterion (AIC), with ΔAIC indicating differences relative to the best-supported model. β denotes the fixed effect of insect count on read abundance, and marginal (R^2m) and conditional (R^2c) coefficients of determination represent variance explained by fixed effects alone and by both fixed and random effects, respectively. n indicates the number of observations.

Window code	Lag (days)	Window (days)	n	AIC	beta	R^2m	R^2c	ΔAIC
count_window15	42	56	435	1983.65	1.10	0.32	0.72	0.0
count_window11	42	42	435	1985.48	1.05	0.30	0.71	1.8
count_window7	42	28	435	1988.45	0.98	0.26	0.70	4.8
count_window3	42	14	435	1992.89	0.83	0.20	0.68	9.2
count_window14	28	56	510	2290.23	1.01	0.27	0.70	306.6
count_window10	28	42	510	2293.67	0.92	0.23	0.68	310.0
count_window6	28	28	510	2296.39	0.84	0.19	0.68	312.7
count_window2	28	14	510	2298.13	0.69	0.13	0.68	314.5
count_window13	14	56	585	2543.75	1.02	0.27	0.70	560.1
count_window9	14	42	585	2545.93	0.97	0.25	0.69	562.3
count_window1	14	14	585	2546.09	0.84	0.19	0.68	562.4
count_window5	14	28	585	2547.27	0.90	0.21	0.69	563.6
count_window12	0	56	660	2890.17	0.97	0.24	0.71	906.5
count_window8	0	42	660	2892.40	0.89	0.20	0.71	908.7
count_window4	0	28	660	2893.21	0.83	0.17	0.71	909.6
count_window	0	14	660	2898.58	0.68	0.11	0.71	914.9



733

734 Figure S1: Correlation plot between the individual abundance and biomass (both logarithmic
735 transformed + 1). The area around the line represents the standard error. To assess the possible
736 correlation between insect biomass and insect count, we fitted both variables to a Bayesian
737 Regression Model, with sampling date nested within site as random effects. We calculated the
738 Bayesian R2 for each model. We found a significant relationship between insect count and
739 biomass from the malaise traps (Bayes R2 = 0.82, Est.Error = 0.0051).

740

741

742

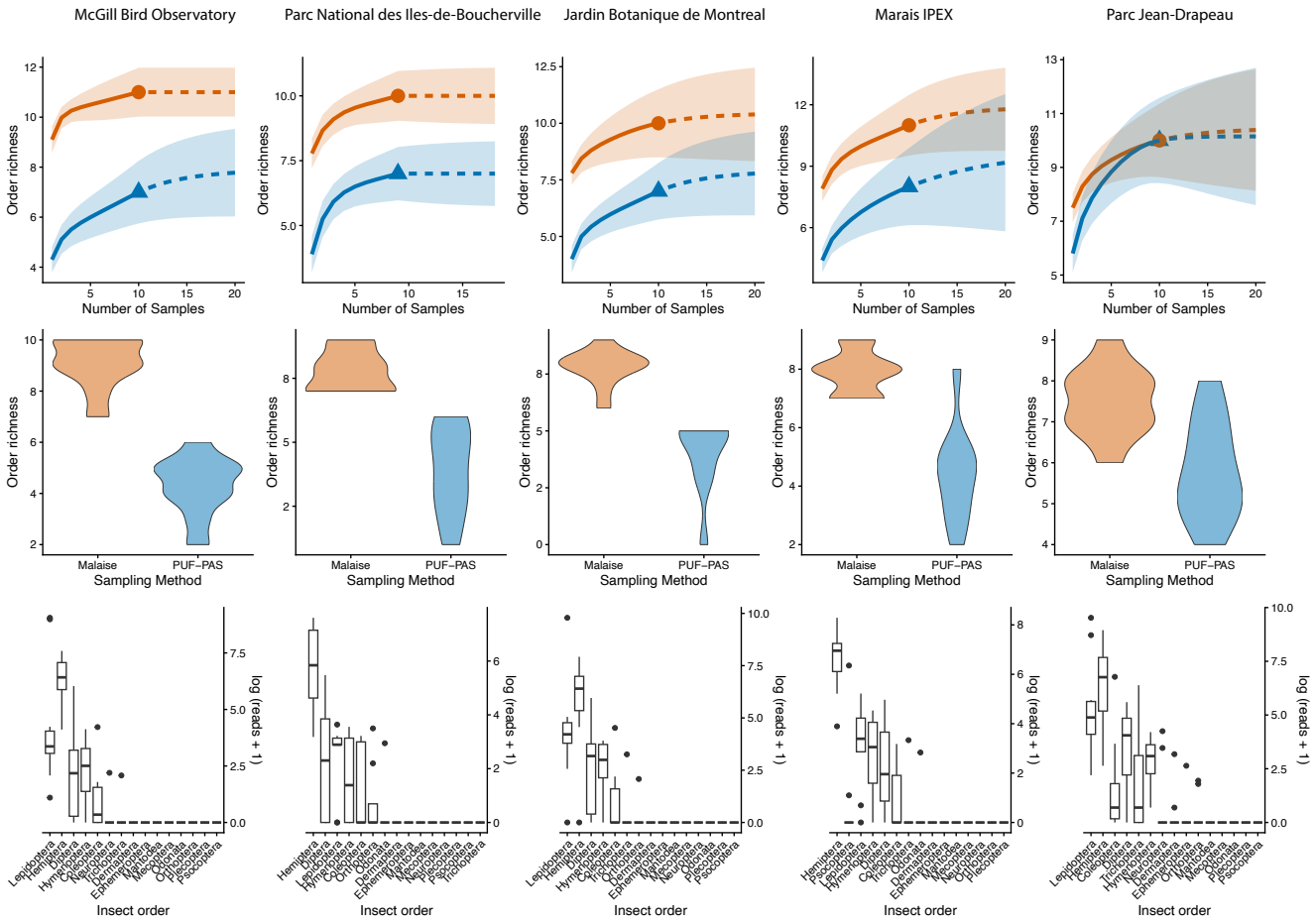


Figure S2: Rarefaction and extrapolation curves (top row), order richness plots (middle row) and boxplot of the number of reads (bottom row) per order subdivided by sites, including McGill Bird Observatory, Parc National des Îles-de-Boucherville, Jardin Botanique de Montréal, Marais IPEX and Parc Jean-Drapeau.

Bibliography

Benson DA et al. 2012. GenBank. Nucleic acids research. 41(D1):D36–D42

Bolyen E et al. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 37(8):852–857. https://doi.org/10.1038/s41587-019-0209-9

Callahan BJ et al. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 13(7):581–583. https://doi.org/10.1038/nmeth.3869

Drake LE et al. 2022. An assessment of minimum sequence copy thresholds for identifying and reducing the prevalence of artefacts in dietary metabarcoding data. Methods in Ecology and Evolution. 13(3):694–710. https://doi.org/10.1111/2041-210X.13780

Geller J, Meyer C, Parker M, Hawk H. 2013. Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. Molecular Ecology Resources. 13(5):851–861. https://doi.org/10.1111/1755-0998.12138

Kalisa E et al. 2024. Capturing the Aerobiome: Application of Polyurethane Foam Disk Passive Samplers for Bioaerosol Monitoring. ACS EST Air. 1(5):414–425. https://doi.org/10.1021/acsestair.3c00107

Leray M et al. 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(1):34. https://doi.org/10.1186/1742-9994-10-34

774 Marinchel N et al. 2023. Mock community experiments can inform on the reliability of eDNA
775 metabarcoding data: a case study on marine phytoplankton. Sci Rep. 13:20164.
776 https://doi.org/10.1038/s41598-023-47462-5

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

779 Meglécz E. 2023. COInr and mkCOInr: Building and customizing a nonredundant barcoding
780 reference database from BOLD and NCBI using a semi-automated pipeline. Molecular
781 Ecology Resources. 23(4):933–945. https://doi.org/10.1111/1755-0998.13756

782 Morrill A et al. 2021. Metabarcoding prey DNA from fecal samples of adult dragonflies shows
783 no predicted sex differences, and substantial inter-individual variation, in diets. PeerJ.
784 9:e12634. https://doi.org/10.7717/peerj.12634

785 Pedregosa F et al. 2011. Scikit-learn: Machine Learning in Python. Journal of Machine Learning
786 Research. 12(85):2825–2830

787 Ratnasingham S et al. 2024. BOLD v4: A Centralized Bioinformatics Platform for DNA-Based
788 Biodiversity Data. In: DeSalle R, editor. DNA Barcoding: Methods and Protocols.
789 Springer US; p 403–441 [accessed 2025 Sept 14]. https://doi.org/10.1007/978-1-0716-
790 3581-0_26. https://doi.org/10.1007/978-1-0716-3581-0_26

791 Rognes T et al. 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ.
792 4:e2584. https://doi.org/10.7717/peerj.2584

793 Smith KF, Kohli GS, Murray SA, Rhodes LL. 2017. Assessment of the metabarcoding approach
794 for community analysis of benthic-epiphytic dinoflagellates using mock communities.

New Zealand Journal of Marine and Freshwater Research. 51(4):555–576.
https://doi.org/10.1080/00288330.2017.1298632