

Characterizing a species-rich and understudied tropical insect fauna using DNA barcoding

David R. Hemprich-Bennett^{1,*}, Ezekiel Donkor², Bernard Adams², Naana Afua Acquah², Eva D. Ofori², Samuel Anie-Amoah², Abigail Bailey¹, Hugh Charles J. Godfray¹, Owen T. Lewis¹, Fred Aboagye-Antwi^{2,3}, Talya D. Hackett^{1,4}

¹Department of Biology, University of Oxford, Life and Mind Building, South Parks Road, Oxford, OX1 3EL, United Kingdom

²Department of Animal Biology and Conservation Science, School of Biological Science, College of Basic and Applied Sciences, University of Ghana P.O. Box LG 25, Legon Accra GA-490-0243 Ghana

³African Regional Postgraduate Programme in Insect Science, School of Biological Science, College of Basic and Applied Sciences, University of Ghana P.O. Box LG 25, Legon Accra GA-490-0243 Ghana

⁴School of Biology, Faculty of Biological Sciences, University of Leeds Irene Manton Building, Woodhouse Lane, Leeds LS2 9JT United Kingdom

*Correspondence address: David R. Hemprich-Bennett, Department of Biology, University of Oxford, South Parks Road, Oxford, OX1 3EL, United Kingdom.

E-mail: david.hemprich-bennett@biology.ox.ac.uk

Abstract

Background West Africa has high biodiversity that is relatively understudied, especially for insects. Studies of West African arthropod diversity can therefore help address important questions regarding conservation, ecosystem services, insecticide use, and other species-control interventions in agriculture and disease management. We intensively sampled arthropods in Ghana using complementary trapping methods, generated DNA barcodes, and classified sequences by Barcode Index Numbers (BINs), a species proxy. Using this dataset, we investigate assemblage composition, temporal activity patterns, and the state of regional biodiversity sampling.

Results Sequencing DNA from 95,996 individuals captured using Malaise, yellow pan, pitfall, Heath and Centre for Disease Control (CDC) traps, we identified 10,120 unique BINs. The rate of species accumulation did not approach an asymptote for any taxonomic group or trap type, indicating high biodiversity. The different trap types sampled different subsets of the local community, with the greatest similarity between yellow pan and pitfall traps. More insects and species (BINs) were trapped during the day than at night. Our dataset shared more BINs in the Barcode of Life Database with South Africa than with any other country, although this predominantly reflects the limited sampling and DNA sequencing campaigns in Africa.

Conclusion This study more than doubles the published BINs for West Africa, offering insights into the biodiversity of an ecologically important but understudied taxon and region. Using multiple trap types allowed a more complete assessment of the local arthropod assemblage. The public release of these data will support and stimulate further taxonomic and ecological work in the region.

Keywords DNA barcoding, Entomology, West Africa, Arthropods, Malaise traps, Entomology methods

Background

Tropical regions harbour ~50% of the world's described species [1], but only 1 million out of an estimated 7 million species of terrestrial arthropods (predominantly insects) have been formally described [2]. The global dry biomass of all insects is ~300 million metric tonnes [3], similar to the combined biomass of humanity and its livestock. Given their functional diversity and numerical dominance in many terrestrial ecosystems, arthropods play critical roles in global ecosystem functioning [4, 5].

Documenting the extraordinary diversity of insect taxa is increasingly important because of the threats that tropical

ecosystems are experiencing, including climate change, land-use change, introduced species, and pollution [6]. These threats are expected to affect many species supra-additively, and with uneven impacts even for closely related taxa [7]. The impacts of these changes are better understood for more heavily studied taxa such as vertebrates [8]. Low levels of baseline data and taxonomic challenges limit our ability to measure changes in insect populations and communities [9]. As a result, the few studies on insect declines tend to come from better-studied temperate faunas in Europe and North America [4], or from a few tropical sites where well-established field stations have facilitated relatively intensive and long-term study, e.g., [10].

Received: 30 September 2025. Revised: 13 February 2026. Accepted: 15 March 2026

© The Author(s) 2026. Published by Oxford University Press on behalf of GigaScience. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

A major impediment to tropical insect studies is that even a brief sampling campaign generates a huge number and diversity of specimens with concomitant logistical and taxonomic challenges [11].

These taxonomic and sampling impediments prevent our understanding of the impact of specific interventions, such as crop pest and disease vector control, on non-target insect populations and broader ecological communities. In particular, technological advances are generating new methods for species-specific targeting of pests and vectors, thereby avoiding widespread bio-cide use [12–14]. Regulators and policymakers require an understanding of the wider biodiversity impacts of these new technologies.

In the past two decades, DNA barcoding has emerged as a key technique to accelerate rates of sample characterization. DNA barcoding works by comparing specific regions of DNA from collected specimens to sequences in a reference database, enabling researchers to infer their likely identity [15]. A partial sequence of the cytochrome c oxidase subunit I (COI) gene has emerged as the near-universal barcode region for arthropods, being highly conserved within species, but variable between species [15, 16]. The primary repository for insect DNA barcodes is the Barcode of Life Data System (BOLD) [17], which stores publicly accessible sequences and associated metadata, and assigns a unique identifier to each sequence cluster, known as a Barcode Index Number (BIN) [18]. BINs are clusters of highly similar DNA sequences and so can act as a species proxy when a species has yet to be formally described. BINs are not 1:1 equivalents for species, as within-species variability in the standard DNA barcode segment is not uniform across broader taxonomic groups, and phenomena such as *Wolbachia* infections [19], between-species hybridization, and errors in data uploaded to BOLD [20] can lead to single BINs containing multiple species and *vice versa*. The addition of new sequencing data to BOLD can also prompt the revision of BIN assignments, making any individual assignment somewhat provisional [21]. Nevertheless, BINs can be useful in ecological studies where precise species information is not essential, e.g., in identifying areas of high biodiversity [22, 23], estimating trends in abundance and diversity [23, 24], or understanding ecological community structure [25]. DNA barcodes are also an invaluable tool for species discovery, especially in areas where coverage of described species is high [26], as uploading data to BOLD allows samples to be matched to previous records along with associated taxonomic, geographic and temporal metadata. In addition, these COI barcodes can be used in metabarcoding studies, acting as reference sequences for eDNA [27] or dietary analysis [28]. BOLD currently has 16.5 million public records, representing 1.2 million BINs. Here, we use BINs as operational taxonomic units (OTUs), where samples are classified into clusters based on shared or diverging sequences. In this form, BINs have been used to analyse ecological patterns in the tropics [29–32], while simultaneously revealing the presence of many previously unknown taxa.

In this paper, we describe the implementation and outcomes of a high-intensity arthropod-sampling campaign in a biodiversity-rich area in Ghana, West Africa. Like many tropical regions, West African forests and savannas have high levels of biodiversity that are threatened by human actions [33]. Compared even to other tropical regions, relatively few arthropod species in West Africa

have been described, and species from this region are under-represented in global DNA datasets such as BOLD, hampering ecological study, estimates of species distribution, extinction and more [34].

Our insect sampling and DNA barcoding campaign primarily took place to make a reference library for use in an ongoing dietary metabarcoding project investigating the position of the malarial mosquito vector, *Anopheles gambiae* [35], in its local ecological community and to assess the effects of different control strategies on non-target organisms. The resulting dataset—which is publicly available and analysed here—will also provide a rich resource for biologists concerned with documenting arthropod diversity in West Africa. Voucher specimens are currently stored at the Centre for Biodiversity Genomics site in Guelph, Canada, and are available for further analyses; in accordance with the Nagoya Protocol Material Transfer Agreement, all specimens and any remaining DNA extract will be returned to the University of Ghana upon request.

Specifically, we ask: (1) What fraction of the local arthropod assemblage does our sample of nearly 100 K individuals reveal? (2) How important is it to use multiple trap types to sample biodiversity? (3) How do insect activity patterns differ between day and night? (4) Which countries represented in the BOLD database share the most BINs with our dataset, and what does this tell us about the global completeness of the BOLD dataset? More broadly, the resulting data on rates of taxonomic discovery through time and across trapping methods will support the optimization of survey methods and strategies for other poorly studied but species-rich arthropod communities.

Methods

Field sites

Samples were collected at and adjacent to two villages (“sites”) in the Volta region of Ghana, Abutia Amegame in the Ho West District (6.209 N, 0.441 E) and Mafi Agove in the Central Tongu District (6.457 N, 0.316 E). Monthly sampling campaigns were conducted from February 2019 to March 2020 inclusive, and (following a pause imposed by the COVID-19 pandemic) from April to June 2021. The villages are small subsistence farming settlements (<1,000 people) within a matrix of grassland, cropland, and forest fragments, in a region characterized by pronounced annual dry and rainy seasons.

Sampling

We sampled within a 500 m radius of village centres, dividing each circle into four equal quadrants (NW, NE, SE, and SW). At each village, on every visit, we set four transects (one within each quadrant) at random, pre-determined start locations and directions (Supplementary Fig. S1, Supplementary Fig. S2). We placed a Malaise trap (Supplementary Fig. S3) at the start of a transect orientated in a direction most likely to intersect with arthropod flight paths based on the typical wind direction, topography and vegetation. At 10 m, 20 m, 30 m, and 40 m, we placed a yellow pan trap and a pitfall trap (Supplementary Fig. S4), both filled with soapy water, on alternating sides of the transect line. We set a Heath trap

(Supplementary Fig. S5) at 50 m and a CO₂-baited Centre for Disease Control (CDC) trap (Supplementary Fig. S6) at the 100 m point to avoid interference with other traps. Traps were left for 24 hours, and Malaise trap bottles were exchanged at 06:00, 12:00, 18:00, and 00:00 to capture temporal dynamics during the sampling period. The arthropods collected using each trap type on each transect on each date are referred to as a “Lot”.

DNA barcoding

We sub-sorted all Lots before they were sent for sequencing to maximize diversity. Individuals from Malaise, pan, pitfall, and CDC Lots were identified to taxonomic order, assigned a morphospecies identity based on visual inspection, and up to five individuals per morphospecies per Lot were selected for sequencing. Heath trap Lots had considerably higher arthropod abundance and richness than those from other trap types. The number of arthropods selected for sequencing per Lot was in proportion to their wet mass, determined as the weight after filtering off ethanol through Nitex mesh. Following a brief visual inspection, samples were selected to maximize the number of morphospecies and in approximate proportion to Lot contents. Due to logistical and financial constraints, only 34 out of 117 Malaise lots were fully sequenced. Araneae (spider) samples from pan and pitfall traps were removed for use in ongoing dietary analyses, and so Araneae are omitted from some analyses here.

Samples were DNA barcoded at the Canadian Centre for DNA Barcoding (CCDB), using their standard protocol: photographing each specimen before performing a non-destructive DNA extraction, PCR using the CO1 “Folmer” region [36], and sequencing on a PacBio Sequel. Through the BOLD data-management platform [17], samples were assigned provisional taxonomy (typically to family level, though more precise information was automatically assigned where possible) and a BIN [18]. Where species-level assignments were made, the resulting species list was queried against a custom database of pests of crops and human health, assembled from a manual search of the literature.

Data analyses

For final analyses, our data were downloaded from BOLD on 15 September 2025, and BOLD was at the same time queried for information on the publicly available BINs that matched BINs in our dataset. All analyses used R 4.5.1 [37], with plots created using the ggplot2 package [38].

Species richness and sampling completeness

Sampling completeness was calculated for each taxonomic order and trap type using the “iNEXT” R package [39]. The rate of BIN accumulation was used to estimate how many individual insects it would be necessary to sequence to document all BINs that would be captured by a given trap type. Calculations were restricted to combinations of taxonomic order and trap type where at least 20 BINs were detected.

Assemblage composition comparisons

To compare assemblages among trap types, a series of non-metric multidimensional scaling (NMDS) analyses were run using the “vegan” R package [40] at order, family, genus, and BIN levels. Taxonomic groups were only included in the analyses if they contained a minimum of 10 samples, with the raw abundance of each taxonomic group in a trap type being used. Each sampling location was categorized by both trap type and habitat type (forest, semi-natural, or village). CDC traps were not included for the genus-level analyses, as there were insufficient samples assigned to the level of genus.

To test for differences among the assemblages included in the NMDS analyses, we also ran a Permutational Multivariate Analysis of Variance using Bray–Curtis dissimilarity values and tested for the interaction between trap type and habitat.

Temporal analyses

To explore temporal changes in arthropod assemblages for the Malaise trap data, we ran two linear mixed effects models. We modelled the change in the number of insects captured (abundance) or the number of BINs detected (richness) with the time of day as a fixed effect and allowed a random intercept for the Lot.

To compare diurnal and nocturnal assemblages, we ran a set of NMDS and Permutational Multivariate Analysis of Variance analyses at order, family, genus, and BIN levels, as above.

We tested if numbers of arthropod individuals and BINs captured were more variable in the daytime than at night with Brown–Forsythe tests using the “onewaytests” package [41].

Geographic analyses

We queried BOLD for the available metadata corresponding to all BINs in our dataset that already had public matches from other studies. The resulting dataset was used to identify countries and continents which shared BINs with our dataset, and the taxonomic assignment of those BINs. Data for the 20 countries with the highest numbers of public BINs matching BINs in our dataset were queried on 15 September 2025 to investigate the extent to which similarity in arthropod assemblages is a result of geographic variations in sampling and sequencing effort versus proximity to Ghana. This linear regression analysis modelled the number of BINs shared with our dataset for each country as a function of the country’s geographic distance from Ghana and the number of public sequences for each country as fixed effects.

We expected the number of shared BINs to be more impacted by the country’s sampling effort than their distance from Ghana, so that, e.g., the relatively well-sampled Costa Rica might have more BINs in common with our dataset than the under-sampled but neighbouring Togo.

Results

Overview of the barcode library

Of the 95,996 samples analysed, sequences were obtained from 81,518 (mean sequence length 653.4 bp, s.d 14.4 bp; see Supplementary Fig. S7).

We recorded 10,120 unique BINs across the 95,996 samples sequenced (Table 1), of which 4,939 were newly recorded in our

Table 1 The number of samples sequenced and the number and diversity of BINs documented for each trap type. “BINs unique to trap type” denotes BINs that were not found in any other type of trap.

Trap type	Number of samples	Number of BINs	Number of BINs unique to the trap type	Shannon diversity
CDC	3,039	758	248	4.20
Heath	65,293	6,896	5,287	6.67
Malaise	11,975	3,233	1,810	6.40
Pitfall	8,296	856	232	4.20
Yellow pan	7,223	1,345	444	4.90

project. At the time of writing, this total represents ~0.8% of the total BINs for all taxa on BOLD. Notably, only 8.5% of public BINs on BOLD originate from Africa, meaning our contribution accounts for nearly 10% of all African taxa in the database. Excluding Ghana, only 4,418 public BINs are from West Africa, and our dataset alone contributes more than double that number. Heath traps, with the greatest number of sequenced samples, provided the most novel BINs (i.e., those not previously represented on BOLD) (2,850) (Table 1; Supplementary Fig. S8).

We recorded 31 taxonomic orders and 384 taxonomic families, with 92% of samples being Coleoptera, Diptera, Hemiptera, Hymenoptera, or Lepidoptera (Table 2; Supplementary Tables S1 and S2). A total of 583 samples (25 BINs) were known crop pest species (Supplementary Table S3), with 264 samples belonging to the families of Diptera which are known to blood-feed (Culicidae, Simuliidae, and Tabanidae); 192 were Culicidae, a haematophagous taxon of particular interest (Supplementary Table S4).

BIN accumulation

For the 13 most abundant taxonomic orders, the average sampling completeness (the percentage of BINs detected relative to those estimated to exist within the community) was 53.3%. Neuroptera had the lowest completeness at 13.5% (137 samples), while Trichoptera had the highest at 71.1% (571 samples). None of the trap types or taxonomic orders reached full completeness (Fig. 1), and some orders, such as Coleoptera, Diptera, and Lepidoptera, were estimated to have thousands of unsampled BINs present at our sites.

Trap complementarity

Overall, assemblages of arthropods caught in each trap type differed significantly at the level of order ($F_{4,2} = 171.8$, $P < 0.001$, $R^2 = 0.46$), family ($F_{4,2} = 106.3$, $P < 0.001$, $R^2 = 0.36$), genus ($F_{3,2} = 6.0$, $P < 0.001$, $R^2 = 0.17$) and BIN ($F_{4,2} = 22.9$, $P < 0.001$, $R^2 = 0.12$). The degree of overlap among trap types decreased with increasing taxonomic resolution (Fig. 2), showing the advantage of fine-scale taxonomic IDs (in this case BINs) when analysing diverse arthropod assemblages. Heath traps were distinct from the other trap types, even at the order level, but there was near-complete order-level overlap between pan and pitfall traps (Fig. 2a). However, at the BIN level, there was very little over-

Table 2 The percentage constitution by taxonomic order of samples sequenced for each trap type. For example, 57.85% of all samples sequenced from CDC traps were Diptera.

Order	Trap type				
	CDC (%)	Heath (%)	Malaise	Pitfall (%)	Yellow pan (%)
Araneae	0.49	0.21	0	0	0.03
Archaeognatha	0	0	0	0.05	0.06
Blattodea	0.07	0.6	0.33	1.25	0.8
Coleoptera	13.43	33.02	5.6	13.39	8.49
Dermaptera	0.07	0.11	0.04	0.13	0.08
Diptera	57.85	9.4	54.7	7.22	28.77
Embioptera	0	0.01	0.01	0	0
Entomobryomorpha	0.07	0.09	0.3	2.74	2.44
Ephemeroptera	0	0.33	0	0	0
Hemiptera	4.61	19.58	11.38	5.61	14.97
Hymenoptera	10.04	9.52	16.64	58.47	36.74
Isopoda	0	0	0	0.02	0
Ixodida	0	0	0	0.04	0
Lepidoptera	11.19	22.62	6.82	0.22	0.57
Mantodea	0.03	0.06	0	0.04	0.06
Mecoptera	0	0.02	0	0	0
Mesostigmata	0	0.09	0	0	0
Neuroptera	0	0.17	0.16	0.05	0.03
Odonata	0.07	0.01	0.03	0	0.03
Orthoptera	0.43	2.7	1.84	9.99	5.39
Phasmida	0	0	0	0.01	0
Plecoptera	0.1	0	0.01	0.01	0.01
Poduromorpha	0	0	0	0.01	0
Pseudoscorpiones	0	0.01	0	0	0
Psocodea	0.59	0.13	0.82	0.1	0.26
Strepsiptera	0	0.03	0	0	0
Symphyleona	0	0.01	0.03	0	0.04
Thysanoptera	0.07	0.21	0.03	0.1	0.75
Trichoptera	0	0.7	0.94	0.01	0.04

lap, with each trap type capturing a distinct arthropod assemblage (Fig. 2d), highlighting the importance of using multiple trap types when surveying arthropods to gain a more representative dataset. Due to the great richness of taxa captured, there were few clear trends of specific taxonomic groups particularly driving these differences, but in general the Culicidae were primarily captured in CDC traps, Formicidae in Heath and pitfall traps, Chlopidae and Muscidae in Malaise traps, and Dolichopodidae mostly in yellow pan traps. Reflecting their relative distinctness from the other trap types, several taxa were predominantly captured in Heath traps, including Termitidae, Carabidae, Chrysomelidae, Dysticidae, Scarabidae, Staphylinidae, Cicadellidae, Miridae, Rhy-parochromidae, Braconidae, Lepidoptera, Mantodea, Orthoptera, and Trichoptera.

Assemblages did not differ significantly among forest, semi-natural, or village habitat types at order ($F_{4,2} = 1.3$, $P = 0.2$, $R^2 = 0.002$) and family ($F_{4,2} = 1.4$, $P = 0.09$, $R^2 = 0.002$) levels; minor differences emerge at the levels of genus ($F_{3,2} = 1.4$, $P = 0.02$, $R^2 = 0.03$) and BIN ($F_{4,2} = 2.3$, $P < 0.001$, $R^2 = 0.006$) although the effect

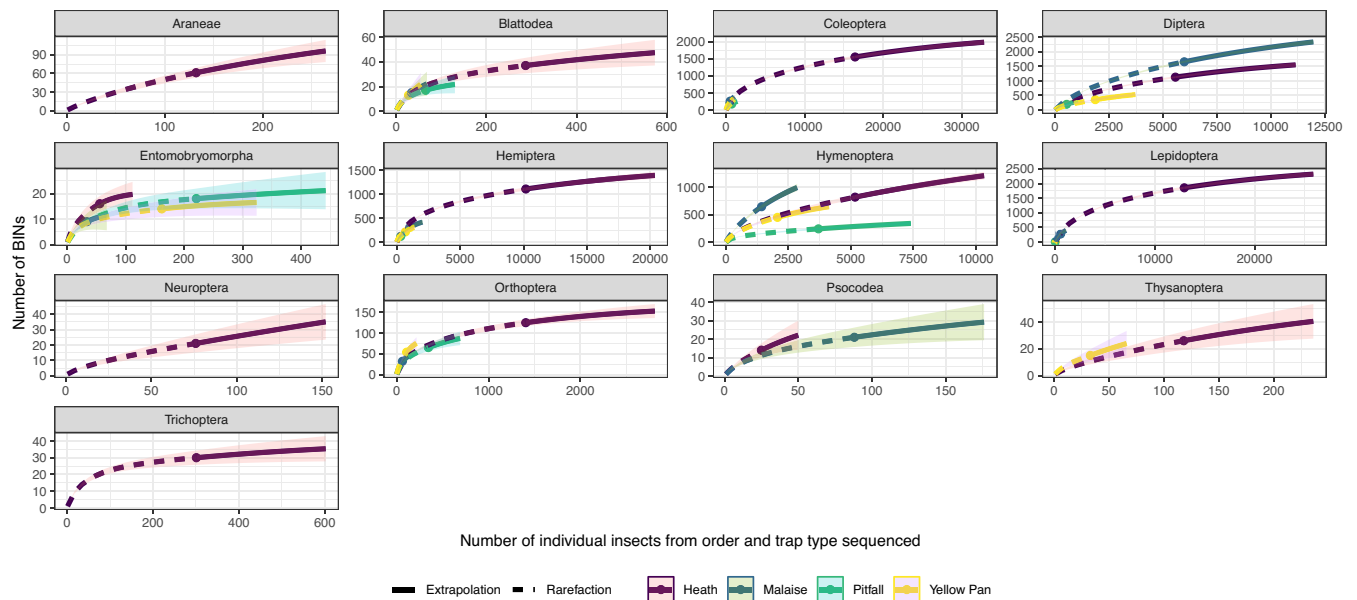


Figure 1 Type 1 iNEXT plots for different taxa showing the observed and extrapolated accumulation of BINs in relation to the number of individuals sequenced. Curves are plotted separately for each trapping method.

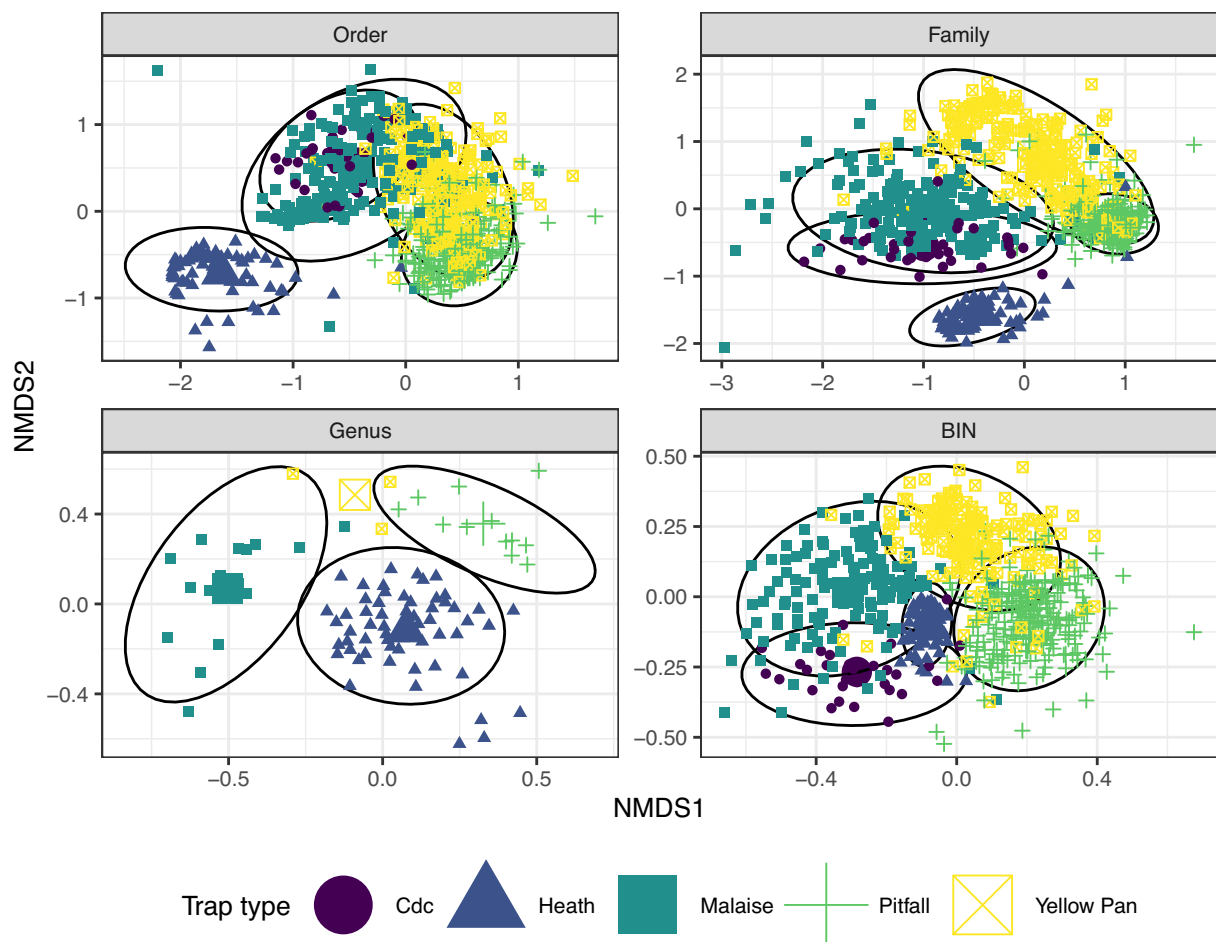


Figure 2 NMDS plots comparing insect assemblages among trap types. Each plot differs in terms of the taxonomic resolution at which samples were categorized. The larger points are the centroids for each trap type. There are fewer data points for the genus-level plot because many samples were not assigned to genus level, despite being identified to family level and assigned a BIN.

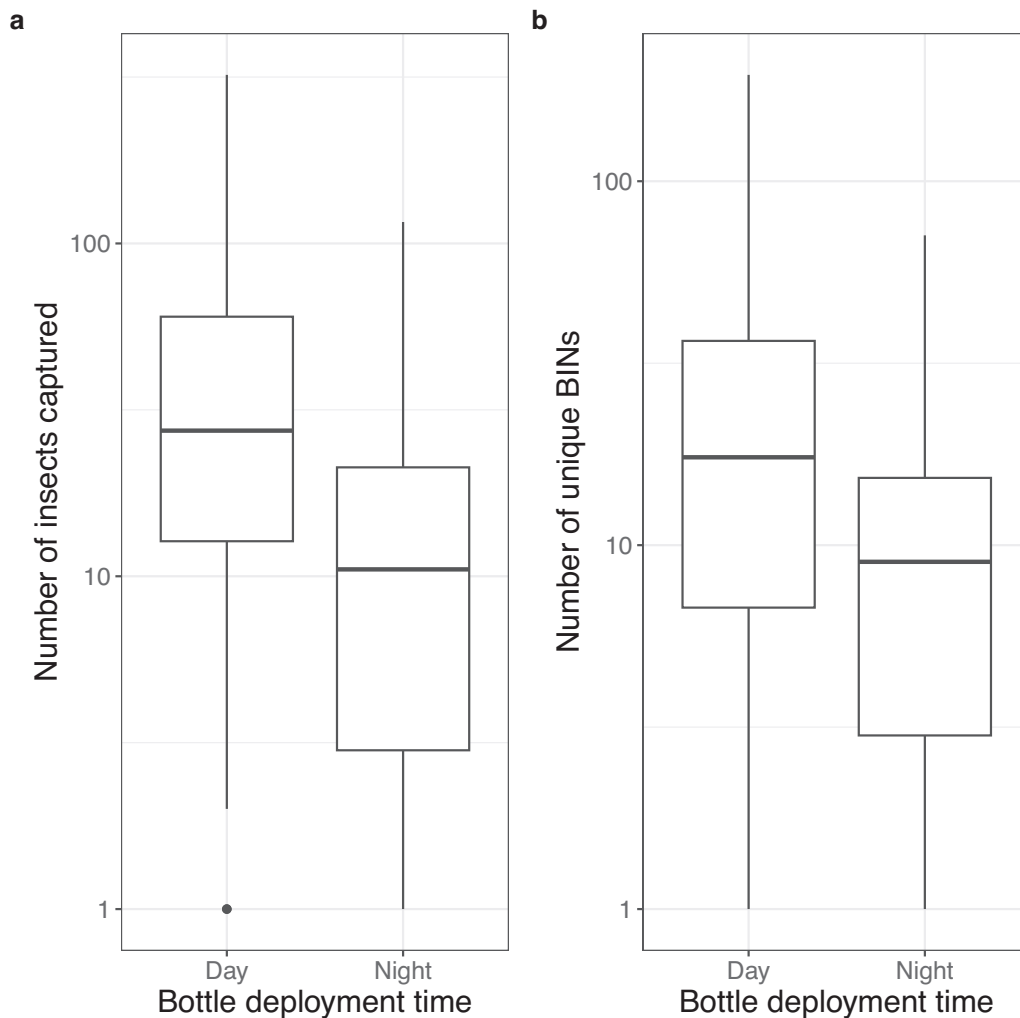


Figure 3 Boxplots showing a) the number of insects captured during diurnal and nocturnal Malaise trap deployment; b) the number of unique BINs captured during diurnal and nocturnal Malaise trap deployment. Y-axes are on a log scale. Boxes show the median, interquartile range and $1.5 \times$ the interquartile range.

sizes and R^2 indicate trivial effects at most (Supplementary Fig. S 9).

Temporal analyses

Malaise traps captured significantly more individual insects during the day ($\bar{x} = 49.1 \pm 63.9$) than at night ($\bar{x} = 20 \pm 25.8$; $F_{1,84} = 13.5$; $P < 0.001$). This trend was consistent across almost all taxa (Fig. 3a), although driven especially by Diptera and Hymenoptera. Many families were substantially more abundant during the day than at night (e.g., Chloropidae, Muscidae, Ceratopogonidae, Cecidomyiidae, and Formicidae; see Supplementary Table S6), while a few were more abundant at night than during the day (Crambidae, Aphrophoridae, Erebidae, Euteliidae, and Gracillariidae). The traps also captured a greater number of unique BINs during the day ($\bar{x} = 32.4 \pm 41$) than at night ($\bar{x} = 14.6 \pm 17.6$) ($F_{1,84} = 12.8$; $P < 0.005$) (Fig. 3b). The trend of greater variability in captures was significant both at the level of number of insects ($F_{1,78.7} = 10.6$; $P < 0.005$) and number of BINs ($F_{1,84} = 9.6$; $P < 0.005$). Ordinations of diurnal and nocturnal assemblages differed slightly at order (F

$_{1,2} = 9$, $P < 0.001$, $R^2 = 0.06$), family ($F_{1,2} = 8.4$, $P < 0.001$, $R^2 = 0.06$), and BIN ($F_{1,2} = 2.4$, $P < 0.01$, $R^2 = 0.02$) levels but not at the genus level ($F_{1,2} = 1$, $P = 0.4$, $R^2 = 0.07$) (see Supplementary Fig. S12).

Geographic analyses

Of the 10,120 unique BINs recorded in our project, 3,281 (32.3%) were found in the publicly available BOLD dataset. 2,625 BINs (25.9%) were present on BOLD but with no publicly available records (i.e., only found in private datasets). Records of these shared BINs came from 189 countries in seven geographic regions (Table 3). The number of shared BINs detected per taxonomic order was broadly consistent with the abundance of each order in our traps (Tables 2 and 3; Supplementary Table S5): Our captures were dominated by Coleoptera, Diptera, Hemiptera, Hymenoptera, and Lepidoptera, and these orders also had the most BINs detected in public datasets. Of the 200 most abundant BINs in our dataset, 104 were already publicly available, 86 had already been sequenced but no other representatives of that

Table 3 The number of BINs found in our dataset that have been found in each major geographic region.

Order name	East Asia & Pacific	Europe & Central Asia	Latin America & Caribbean	Middle East & North Africa	North America	South Asia	Sub-Saharan Africa	Already sequenced, no public sequences	Unique to our project
Araneae	2	2	3	1	1	3	25	9	34
Blattodea	0	0	1	0	0	0	13	19	15
Coleoptera	34	23	32	19	16	37	236	532	1057
Dermaptera	0	0	0	1	1	0	2	2	5
Diptera	97	34	62	116	28	116	971	761	935
Entomobryomorpha	3	0	3	2	1	2	2	12	11
Ephemeroptera	0	0	0	1	0	0	1	0	11
Hemiptera	46	27	32	50	23	51	309	408	591
Hymenoptera	30	22	24	48	14	24	415	376	854
Lepidoptera	75	81	30	99	32	75	1027	401	529
Mantodea	0	0	0	0	0	0	4	1	5
Mesostigmata	0	1	1	1	1	2	2	2	3
Neuroptera	0	0	0	1	0	0	7	7	9
Odonata	1	0	1	0	1	1	2	1	0
Orthoptera	2	9	1	6	1	3	46	63	61
Plecoptera	0	0	0	0	0	0	1	0	1
Poduromorpha	1	0	0	0	0	0	1	0	1
Psocodea	9	0	13	3	4	9	16	3	6
Thysanoptera	2	1	2	2	1	2	9	14	17
Trichoptera	0	0	0	0	0	0	16	3	10

BIN were publicly available on BOLD, and 10 were unique to our project. Only 8 of the 200 most abundant BINs were assigned binomial names by BOLD (*Carpophilus marginellus*, *Corynoptera forcipata*, *Euplatypus hintzi*, *Hycleus hermanniae*, *Microvelia pygmaea*, *Monolepta jacksoni*, *Nysius graminicola*, and *Peregrinus maidis*). Four of these are considered crop pest species (Supplementary Table S3). BINs were assigned binomial names in 355 cases, accounting for 2.6% of our total samples.

A total of 13 of the top 20 countries sharing the most public BINs with our project are in Africa (Table 4, Figs 4 and 5). In the overall model, the number of shared BINs decreased with geographic distance and increasing number of public BINs ($F_{2,166} = 3.76$; $P < 0.05$; $R^2 = 0.03$). However, when examining the model's two fixed effects independently, the number of shared BINs correlated positively with the number of samples barcoded in each country ($\beta = 2.12$, $P < 0.05$) (Fig. 4) but not with the distance between the country and Ghana ($\beta = -1.92$, $P = 0.56$) (Fig. 5).

Discussion

We found >10,000 unique BINs, more than doubling the number previously documented for West Africa. Despite this intense sampling effort, our species accumulation curves did not asymptote for any taxonomic group or trap type, highlighting the high biodiversity of the sites studied. There was greater overlap in BINs with South Africa than any other country, despite the large geographic distance between the two countries, highlighting the low taxonomic completeness from previous work for much of the region. The model had a low R^2 (0.03), indicating that other variables are likely important for the number of shared taxa, such as the country's climate, habitat types, or trap

Table 4 The 20 countries sharing the most publicly available BINs with our dataset.

Rank	Country	Geographic region	Number of shared BINs
1	South Africa	Sub-Saharan Africa	1,840
2	Tanzania	Sub-Saharan Africa	1,332
3	Gabon	Sub-Saharan Africa	1,293
4	Kenya	Sub-Saharan Africa	680
5	Mozambique	Sub-Saharan Africa	504
6	Cameroon	Sub-Saharan Africa	384
7	Nigeria	Sub-Saharan Africa	360
8	Ghana	Sub-Saharan Africa	356
9	Pakistan	South Asia	333
10	Australia	East Asia & Pacific	233
11	Egypt	Middle East & North Africa	233
12	Madagascar	Sub-Saharan Africa	223
13	Costa Rica	Latin America & Caribbean	204
14	Cote d'Ivoire	Sub-Saharan Africa	199
15	Ethiopia	Sub-Saharan Africa	193
16	Bangladesh	South Asia	168
17	United States	North America	154
18	Malaysia	East Asia & Pacific	149
19	India	South Asia	147
20	Central African Republic	Sub-Saharan Africa	143

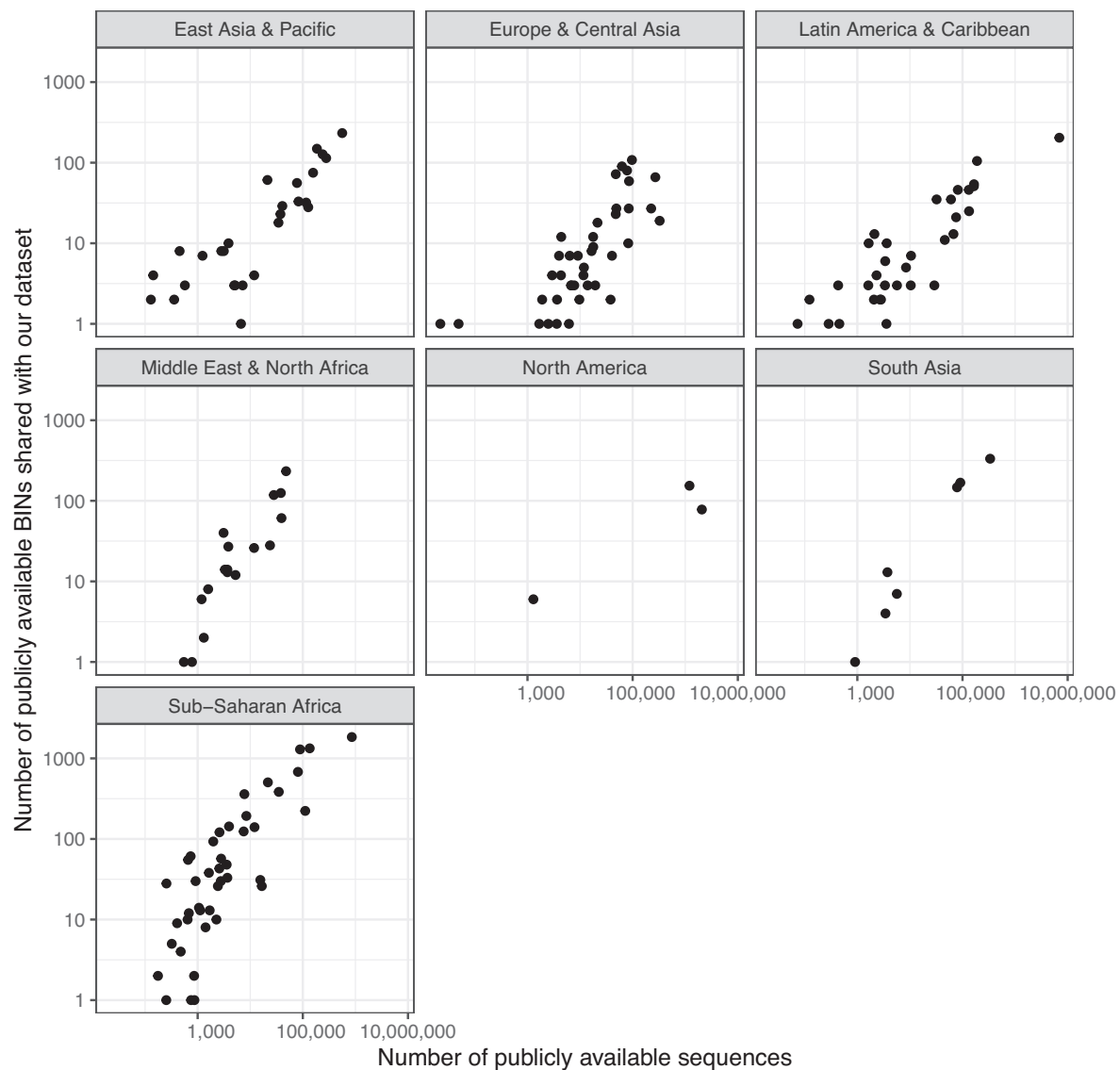


Figure 4 The total number of public sequences available from every country, and the number of BINs that each country shares with our dataset.

types most used in sampling. Our data also show the importance of using multiple trapping methods when documenting arthropod assemblages and highlight the distinct taxonomic groups that are captured by the trap types used, patterns that are especially clear when samples are identified to a fine taxonomic resolution.

Efficiency and complementarity of trapping methods

Each trap type captured a distinct subset of the arthropod assemblage [42], with our five trap types capturing 60.3% of the estimated total BINs present at the study sites (Supplementary F fig. S11). Effective biodiversity survey design requires an understanding of how trap complementarity and overlap influence sampling completeness, allowing effort and cost to be optimized for the question at hand [42]. Generally, Malaise trap samples are

consistently dominated by the same 20 insect families, primarily Diptera and Hymenoptera, across continents and biomes [43]. In our study, 14 of these globally dominant families were also among the top 20 most abundant families we recorded in Malaise traps, with five of the remaining six being Diptera (see Supplementary Table S5). While many large-scale insect survey campaigns rely solely on Malaise traps [29–32, 44], our results confirm that using a broader range of sampling methods will provide more comprehensive assessments than increased sampling effort with a single trap type [42, 45, 46]. Determining the species accumulation benefit of any given trap type or combination would require a dedicated study in an area with a well-characterized community (e.g., Wytham Woods, UK [47] or Zackenberg, Greenland [48]) but would be a valuable insight for study design.

Arthropod assemblages captured in pitfall and yellow pan traps had high order-level and family-level overlap, but those captured using other trap types were more distinct. The distinction between assemblages caught by yellow pan and pitfall traps is only clear

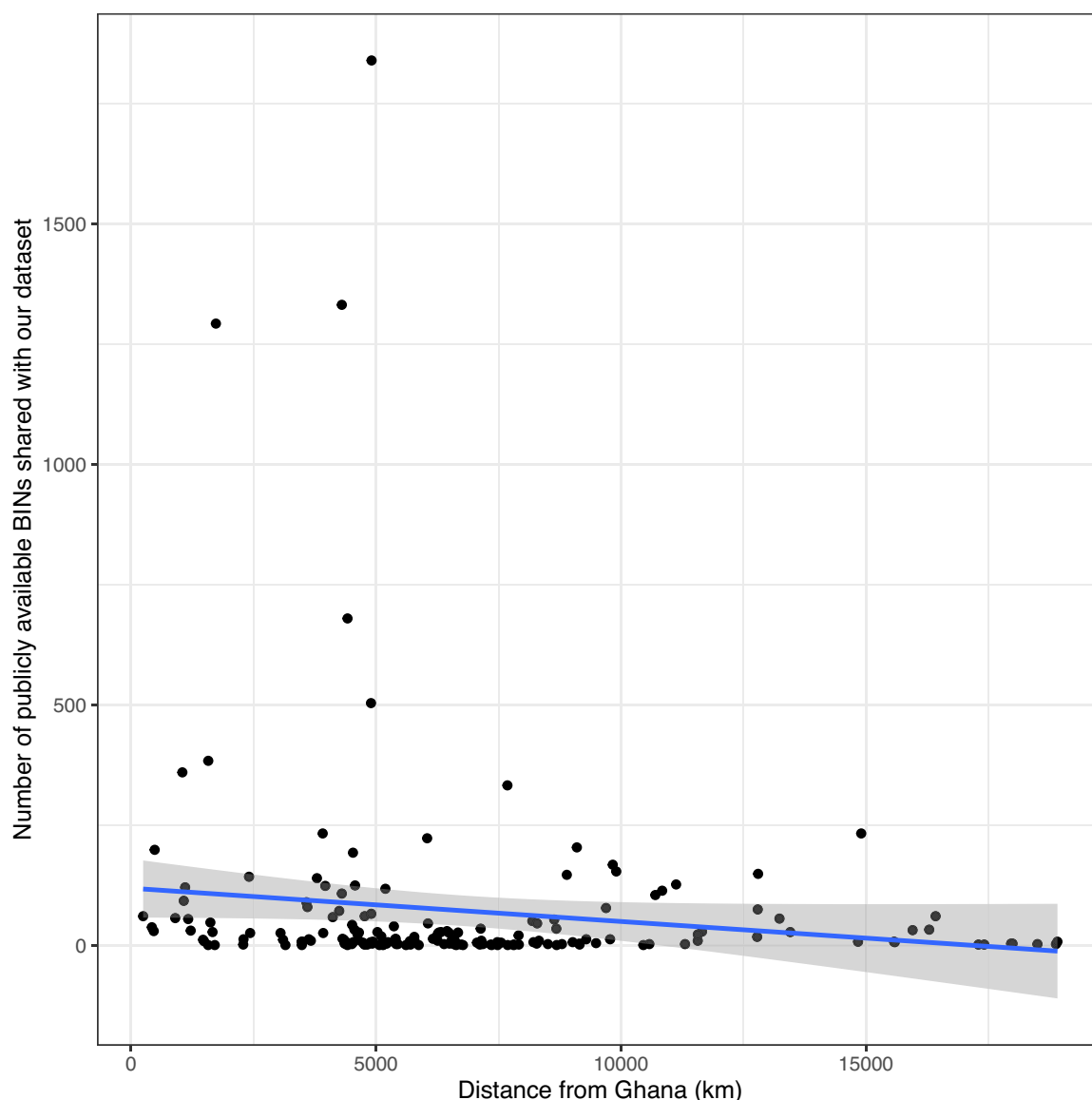


Figure 5 The number of public BINs shared with our dataset per country, and the shortest distance between that country's centroid and the centroid of Ghana.

at the BIN level (Figs 2 and 3; [Supplementary Table S2](#)). Both trap types are placed on the ground and so can capture cursorial arthropods. However, yellow pan traps generally target flying, flower-visiting insects that are attracted to the colour yellow [46]. It appears that our yellow pan traps functioned to capture these flower visitors (overlapping substantially with Malaise traps) as well as many of the cursorial species captured by pitfall traps. At the BIN level, a distinction between the two trap types becomes apparent. Taxa that were much more frequent in yellow pan traps than pitfall traps included Diptera families such as Chloropidae and Sarcophagidae, which are known to be flower visitors. Conversely, Sphaeroceridae were more frequent in pitfall traps than pan traps; many Sphaeroceridae are saprophagous, and they may have been attracted to pitfall traps as potential egg-laying sites, drawn by the decomposition smell from trapped invertebrates.

While sampling effort (number of insects sequenced) was greatest for Heath traps, the highest rate of BIN accumulation observed was for Malaise traps ([Supplementary Fig. S11](#)). Heath traps produced the greatest overall number of BINs for the study, and with a distinct subset of the community compared to other trap types, even at the family level. Heath traps (and other light traps) are commonly employed to target nocturnal Lepidoptera; our results suggest that they may also be an effective way to sample and discover a broad range of other taxa, especially if deployed alongside complementary trapping methods. Budget limitations precluded us from sequencing all trapped arthropods (likely >1 million individuals). These results might to some extent be biased by our sub-sorting approach, but we expect the impact to be minimal as sub-sorting was designed to capture the diversity of traps rather than abundance. Nevertheless, cryptic taxa may be under-detected due to this limitation, potentially impacting rates of BIN accumulation

or trap complementarity. Given that Malaise trap samples showed the highest rate of BIN accumulation (Supplementary Fig. S11), but unfortunately many Malaise trap Lots were unable to be sequenced, this will have likely reduced the overall contribution of this trap type.

Taxa of potential human importance

A very small fraction of all BINs (0.25%) and individuals (0.6%) corresponded to known crop pest species (Supplementary Table S3). The most common of these species included the Lepidoptera species *Thaumotobia leucotreta* and Hemiptera species *Nesidiocoris tenuis* and *Rhopalosiphum rufiabdominale*, known pests of peppers and tomatoes, both crops commonly grown at our study sites. Although it is documented to cause significant crop damage in the study areas [49], the important crop pest Fall Armyworm, *Spodoptera frugiperda* (Lepidoptera, Noctuidae), was only detected twice in our data set, and was only captured in Heath traps. Despite including CO₂-baited CDC traps specifically deployed to catch blood-feeding Diptera, only 3.1% of all BINs and 1.9% of all individuals trapped were from families that contain blood-feeding species (Supplementary Table S4). We caution that most of these BINs were not identified to species, and most families in question (e.g., the especially abundant family Ceratopogonidae) contain both blood-feeding and non-blood-feeding taxa. The three nearly exclusively haematophagous families most likely to be of concern to human health (Culicidae, Simuliidae, and Tabanidae) comprised 0.5% of all captured BINs and 0.28% of all captured insects. The commonest of these families was Ceratopogonidae (mainly caught in Heath traps). Tabanidae were mostly caught in Malaise traps, and Culicidae were mostly caught in CDC traps (Supplementary Table S4). Taken together, these results show the utility of different trapping methods for surveillance of economically important tropical insects, while highlighting their relatively minor contribution to the overall insect assemblage when compared to the range of numerous economically neutral or beneficial insects.

Diurnal activity patterns

In our Malaise traps, more insect individuals and more BINs were trapped during the day than at night. This pattern was driven largely by Diptera and Hymenoptera (Supplementary Table S6). Surprisingly, inspection of the order and family-level taxonomic composition of Malaise trap samples revealed few other taxa with strongly day-biased or night-biased activity. Wong and Didham [50] previously found overall global insect activity patterns to be higher at night than in the day, but the effect was influenced by both insect community composition and habitat type. Activity patterns were higher during the day in grasslands, savannahs, and forests [50], habitat types that are somewhat analogous to our matrix of grassland, cropland and forest fragments, and all habitats where there is a strong variation between daytime and nighttime temperatures. Activity patterns were more variable during daytime than at night, potentially reflecting higher between-day variance in thermal conditions, which strongly influence insect activity [51] during daylight hours. Taken together, these results suggest that, for passive trapping methods such as Malaise traps, efforts should be made to standard-

ise the extent to which deployments span diurnal and nocturnal periods.

Conclusions

Our dataset provides insights into our current knowledge of tropical arthropod biodiversity and highlights the need for extensive further research in the region and beyond. By publishing this manuscript and dataset, we provide a genetic and taxonomic resource for the scientific community, particularly for those studying tropical arthropod fauna in West Africa. While BINs are an invaluable first step in describing biodiversity, we encourage efforts towards the formal taxonomic description of the many unnamed taxa within this understudied arthropod assemblage.

Availability of source code and requirements

Project name: ghana_bins

Project homepage: https://github.com/hemprichbennett/ghana_bins

License: MIT

Programming language: R

Package management: R or Docker

Additional files

Supplementary Fig. S1. Map of random sampling points at the village Abutia Amegame.

Supplementary Fig. S2. Map of random sampling points at the village Mafi Agove.

Supplementary Fig. S3. Malaise trap.

Supplementary Fig. S4. Pitfall trap and yellow pan trap.

Supplementary Fig. S5. Heath trap.

Supplementary Fig. S6. CDC trap.

Supplementary Fig. S7. Jittered density plot showing the distribution of sequence lengths generated.

Supplementary Fig. S8. Public availability of BINs detected, by taxonomic order and trap-type.

Supplementary Fig. S9. NMDS plots comparing insect assemblages among different habitat types. Each plot differs in terms of the taxonomic resolution at which samples were categorised. The larger points are the centroids for each trap type.

Supplementary Fig. S10. Type 2 iNEXT plot, showing the estimated sample coverage for each taxonomic order and trap type.

Supplementary Fig. S11. Type 1 iNEXT plot showing the observed and extrapolated accumulation of BINs in relation to the number of individuals sequenced for each trap type.

Supplementary Fig. S12. NMDS plots comparing insect assemblages among day and night-deployed Malaise traps. Each plot differs in terms of the taxonomic resolution at which samples were categorised. The larger points are the centroids for each trap type. There are fewer data points for the genus-level plot because many samples were not assigned to genus level, despite being identified to family level and assigned a BIN.

Supplementary Table S1. The number of samples of a given order sequenced in each trap type.

Supplementary Table S2. The number of samples of a given family sequenced in each trap type.

Supplementary Table S3. The number of crop pest species detected in each trap type.

Supplementary Table S4. The number of individuals of Dipteran families known to contain blood-feeders detected in each trap type.

Supplementary Table S5. The most abundant 20 taxonomic families for each trap type, and whether they were found in Srivathsan et al.'s 2023 paper documenting the globally most abundant taxonomic families in Malaise trap data. For example, cecidomyiidae were the most abundant family in CDC data, and were found in Srivathsan et al., Drosophilidae were 11th most abundant family in Malaise trap data and were not found in Srivathsan et al.'s 20 most abundant families in Malaise trap data.

Supplementary Table S6. The numbers of individuals of different arthropod taxa collected in Malaise traps during daytime and nighttime.

Acknowledgments

This research was supported by members of the Target Malaria Ghana Stakeholder Engagement Team, especially Divine Dzokoto and Linda Mawutor Aboagye. Data collection was facilitated by village elders, members of the two study communities, and especially those who assisted with trap deployment and sample collections. Sample sorting protocols and DNA barcoding processing was supported by the Centre for Biodiversity Genomics, especially Michelle D'Souza and Jayme Sones. CDC trap samples were sorted by Ben Bellekom. All sampling data are stored on Earthcape Biodiversity Database Platform [54] with the support of Evgeniy Meyke. All authors were working as members of the Target Malaria Not-for-Profit Research Consortium [55], which receives core funding from the Gates Foundation and from Open Philanthropy.

Author contributions

David Hemprich-Bennett (Methodology [equal], Formal analysis [equal], Data curation [equal], Writing—original draft, Writing—review and editing [equal], Visualisation), Ezekiel Donkor (Data curation [equal], Investigation [equal], Resources [equal], Writing—review & editing [equal]), Bernard Adams (Investigation [equal], Resources [equal], Writing—review & editing [equal]), Naana Afua Acquah (Investigation [equal], Resources [equal], Writing—review & editing [equal]), Eva D. Ofori (Investigation [equal], Resources [equal], Writing—review & editing [equal]), Samuel Anie-Amoah (Investigation [equal], Resources [equal], Writing—review & editing [equal]), Abigail Bailey (Investigation [equal], Resources [equal], Writing—review & editing [equal]), H. Charles J. Godfray (Conceptualization [equal], Funding acquisition [equal], Methodology [equal], Project administration [equal], Supervision [equal], Writing—review & editing [equal]), Owen T. Lewis (Funding acquisition [equal], Methodology [equal], Project administration [equal], Supervision [equal], Writing—review & editing [equal]), Fred Aboagye-Antwi (Conceptualization [equal], Funding acquisition [equal], Methodology [equal], Project administration [equal], Supervision [equal], Writing—review & editing [equal]), and Talya D. Hackett (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Investigation [equal], Methodol-

ogy [equal], Project administration [equal], Supervision [equal], Writing—review & editing [equal])

Funding

All authors were working as members of the Target Malaria Research Consortium, which receives core funding from the Gates Foundation and from Open Philanthropy. DRHB is also supported by a grant from the Wellcome Trust (UNS1285360)

Data availability

Data for all samples are available from BOLDsystems [52] under the project IDs “GCEP”, “TMGHA”, and “TMGHB”. All additional Supplementary material are available in the GigaScience repository, GigaDB [53].

Competing interests

None of the authors have any competing interests to declare.

References

- Lewis SL, Edwards DP, Galbraith D. Increasing human dominance of tropical forests. *Science*. 2015;**349**:827–32. <https://doi.org/10.1126/science.aaa9932>.
- Stork NE. How many species of insects and other terrestrial arthropods are there on Earth? *Annu Rev Entomol*. 2018;**63**:31–45. <https://doi.org/10.1146/annurev-ento-020117-043348>.
- Rosenberg Y, Bar-On YM, Fromm A *et al*. The global biomass and number of terrestrial arthropods. *Sci Adv*. 2023;**9**:eabq4049. <https://doi.org/10.1126/sciadv.abq4049>.
- Blüthgen N, Dicks LV, Forister ML *et al*. Insect declines in the Anthropocene. *Nat Rev Earth Environ*. 2023;**4**:683–6. <https://doi.org/10.1038/s43017-023-00478-x>.
- Rodger JG, Bennett JM, Razanajatovo M *et al*. Widespread vulnerability of flowering plant seed production to pollinator declines. *Sci Adv*. 2021;**7**:eabd3524. <https://doi.org/10.1126/sciadv.abd3524>.
- Boyle MJW, Bonebrake TC, Dias da Silva K *et al*. Causes and consequences of insect decline in tropical forests. *Nat Rev Biodivers*. 2025;**1**:315–31. <https://doi.org/10.1038/s44358-025-00038-9>.
- Williamson J, Teh E, Jucker T *et al*. Local-scale temperature gradients driven by human disturbance shape the physiological and morphological traits of dung beetle communities in a Bornean oil palm–forest mosaic. *Funct Ecol*. 2022;**36**:1655–67. <https://doi.org/10.1111/1365-2435.14062>.
- Wagner DL, Grames EM, Forister ML *et al*. Insect decline in the Anthropocene: death by a thousand cuts. *Proc Natl Acad Sci USA*. 2021;**118**. <https://doi.org/10.1073/pnas.2023989118>.
- Cooke R, Outhwaite CL, Bladon AJ *et al*. Integrating multiple evidence streams to understand insect biodiversity change. *Science*. 2025;**388**:eadq2110. <https://doi.org/10.1126/science.adq2110>.
- Bonadies E, Lamarre GPA, Souto-Vilarós D *et al*. Population trends of insect pollinators in a species-rich tropical rainfor-

- est: stable trends but contrasting patterns across taxa. *Biol Lett*. 2024;**20**:20240170. <https://doi.org/10.1098/rsbl.2024.0170>.
11. Gotelli NJ. A taxonomic wish-list for community ecology. *Phil Trans R Soc Lond B*. 2004;**359**:585–97. <https://doi.org/10.1098/rstb.2003.1443>.
 12. Hammond A, Galizi R, Kyrrou K *et al*. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. *Nat Biotechnol*. 2016;**34**:78–83. <https://doi.org/10.1038/nbt.3439>.
 13. Kyrrou K, Hammond AM, Galizi R *et al*. A CRISPR-Cas9 gene drive targeting doublesex causes complete population suppression in caged *Anopheles gambiae* mosquitoes. *Nat Biotechnol*. 2018;**36**:1062–6. <https://doi.org/10.1038/nbt.4245>.
 14. Willis K, Burt A. Engineering drive–selection balance for localized population suppression with neutral dynamics. *Proc Natl Acad Sci USA*. 2025;**122**. <https://doi.org/10.1073/pnas.2414207122>.
 15. Hebert PDN, Cywinska A, Ball SL *et al*. Biological identifications through DNA barcodes. *Proc R Soc Lond B*. 2003;**270**:313–21. <https://doi.org/10.1098/rspb.2002.2218>.
 16. Hebert PDN, Ratnasingham S, De Waard JR. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc R Soc Lond B*. 2003;**270**:S96–9. <https://doi.org/10.1098/rsbl.2003.0025>.
 17. Ratnasingham S, Hebert PDN. Bold: the Barcode of life data system (<http://www.barcodinglife.org>). *Mol Ecol Notes*. 2007;**7**:355–64. <https://doi.org/10.1111/j.1471-8286.2007.01678.x>.
 18. Ratnasingham S, Hebert PDN. A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLoS One*. 2013;**8**:e66213. <https://doi.org/10.1371/journal.pone.0066213>.
 19. Smith MA, Bertrand C, Crosby K *et al*. Wolbachia and DNA barcoding insects: patterns, potential, and problems. *PLoS One*. 2012;**7**:e36514. <https://doi.org/10.1371/journal.pone.0036514>.
 20. Stein F, Gailing O. Identification of BOLD engine deficiencies and suggestions for improvement based on a curated *Tachina* (Diptera) record set. *PLoS One*. 2025;**20**:e0331216. <https://doi.org/10.1371/journal.pone.0331216>.
 21. Meier R, Blaimer BB, Buenaventura E *et al*. A re-analysis of the data in Sharkey *et al.*'s (2021) minimalist revision reveals that BINs do not deserve names, but BOLD systems needs a stronger commitment to open science. *Cladistics*. 2022;**38**:264–75. <https://doi.org/10.1111/cla.12489>.
 22. Young MR, Proctor HC, deWaard JR *et al*. DNA barcodes expose unexpected diversity in Canadian mites. *Mol Ecol*. 2019;**28**:5347–59. <https://doi.org/10.1111/mec.15292>.
 23. D'Souza ML, Van Der Bank M, Shongwe Z *et al*. Biodiversity baselines: tracking insects in Kruger National Park with DNA barcodes. *Biol Conserv*. 2021;**256**:109034. <https://doi.org/10.1016/j.biocon.2021.109034>.
 24. Vamosi JC, Gong Y-B, Adamowicz SJ *et al*. Forecasting polination declines through DNA barcoding: the potential contributions of macroecological and macroevolutionary scales of inquiry. *New Phytol*. 2017;**214**:11–18. <https://doi.org/https://doi.org/10.1111/nph.14356>.
 25. Hemprich-Bennett DR, Kemp VA, Blackman J *et al*. Altered structure of bat-prey interaction networks in logged tropical forests revealed by metabarcoding. *Mol Ecol*. 2021;**30**:5844–57. <https://doi.org/10.1111/mec.16153>.
 26. Meier R, Lawniczak MKN, Srivathsan A. Illuminating entomological dark matter with DNA barcodes in an era of insect decline, deep learning, and genomics. *Annu Rev Entomol*. 2024;**70**:185–204. <https://doi.org/10.1146/annurev-ento-040124-014001>.
 27. Taberlet P, Coissac E, Hajibabaei M *et al*. Environmental DNA. *Mol Ecol*. 2012;**21**:1789–93. <https://doi.org/10.1111/j.1365-294X.2012.05542.x>.
 28. Clare EL. Molecular detection of trophic interactions: emerging trends, distinct advantages, significant considerations and conservation applications. *Evol Appl*. 2014;**7**:1144–57. <https://doi.org/10.1111/eva.12225>.
 29. Ashfaq M, Sabir JSM, El-Ansary HO *et al*. Insect diversity in the Saharo-Arabian region: revealing a little-studied fauna by DNA barcoding. *PLoS One*. 2018;**13**:e0199965. <https://doi.org/10.1371/journal.pone.0199965>.
 30. Bukowski B, Ratnasingham S, Hanisch PE *et al*. DNA barcodes reveal striking arthropod diversity and unveil seasonal patterns of variation in the southern Atlantic Forest. *PLoS One*. 2022;**17**:e0267390. <https://doi.org/10.1371/journal.pone.0267390>.
 31. deWaard JR, Levesque-Beaudin V, deWaard SL *et al*. Expedited assessment of terrestrial arthropod diversity by coupling Malaise traps with DNA barcoding. *Genome*. 2019;**62**:85–95. <https://doi.org/10.1139/gen-2018-0093>.
 32. D'Souza ML, Hebert PDN. Stable baselines of temporal turnover underlie high beta diversity in tropical arthropod communities. *Mol Ecol*. 2018;**27**:2447–60. <https://doi.org/10.1111/mec.14693>.
 33. Norris K, Asase A, Collen B *et al*. Biodiversity in a forest-agriculture mosaic—the changing face of West African rainforests. *Biol Conserv*. 2010;**143**:2341–50. <https://doi.org/10.1016/j.biocon.2009.12.032>.
 34. Virgilio M, Jordaens K, Berman FC *et al*. Identifying insects with incomplete DNA barcode libraries, African fruit flies (Diptera: Tephritidae) as a test case. *PLoS One*. 2012;**7**:e31581. <https://doi.org/10.1371/journal.pone.0031581>.
 35. Sinka ME, Bangs MJ, Manguin S *et al*. The dominant *Anopheles* vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic précis. *Parasites Vectors*. 2010;**3**:117. <https://doi.org/10.1186/1756-3305-3-117>.
 36. Folmer O, Black M, Hoeh W *et al*. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit 1 from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol*. 1994;**3**:294–9.
 37. R Core Team. *R: a Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing; 2023.
 38. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag; 2016.
 39. Hsieh TC, Ma KH, Chao A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol Evol*. 2016;**7**:1451–6. <https://doi.org/10.1111/2041-210X.12613>.

40. Oksanen J, Simpson GL, Blanchet FG *et al.* *vegan*: Community ecology package 2022. <https://vegandevs.github.io/vegan/>. Accessed June 2025.
41. Dag O, Kasikci M, Dolgun A *et al.* *onewaytests*: One-way tests in independent groups designs. 2023. <https://journal.r-project.org/articles/RJ-2018-022/>. Accessed June 2025.
42. Hoffmann L, Stoll S. Catch effectiveness, complementarity and costs of five sampling techniques for flying insects across different land use types. *Insect Conserv Diversity*. 2025;**18**:970–81. <https://doi.org/10.1111/icad.12839>.
43. Srivathsan A, Ang Y, Heraty JM *et al.* Convergence of dominance and neglect in flying insect diversity. *Nat Ecol Evol*. 2023;**7**:1012–21. <https://doi.org/10.1038/s41559-023-02066-0>.
44. Steinke D, deWaard SL, Sones JE *et al.* Message in a Bottle—Metabarcoding enables biodiversity comparisons across ecoregions. *GigaScience* 2022;**11**:giac040. <https://doi.org/10.1093/gigascience/giac040>.
45. Noyes JS. A study of five methods of sampling Hymenoptera (Insecta) in a tropical rainforest, with special reference to the Parasitica. *J Nat Hist*. 1989;**23**:285–98. <https://doi.org/10.1080/00222938900770181>.
46. Missa O, Basset Y, Alonso A *et al.* Monitoring arthropods in a tropical landscape: relative effects of sampling methods and habitat types on trap catches. *J Insect Conserv*. 2009;**13**:103–18. <https://doi.org/10.1007/s10841-007-9130-5>.
47. Savill P, Perrins C, Kirby K, Fisher N (eds.) *Wytham Woods: Oxford's Ecological Laboratory*. Oxford: Oxford University Press, 2010.
48. Roslin T, Wirta H, Hopkins T *et al.* Indirect interactions in the high arctic. *PLoS One*. 2013;**8**:e67367. <https://doi.org/10.1371/journal.pone.0067367>.
49. Jordon MW, Hackett TD, Aboagye-Antwi F *et al.* Effects of distance from semi-natural habitat on fall armyworm (*Spodoptera frugiperda*, J. E. Smith) and its potential natural enemies in Ghana. *Bull Entomol Res*. 2022;**112**:343–53. <https://doi.org/10.1017/S0007485321000894>.
50. Wong MKL, Didham RK. Global meta-analysis reveals overall higher nocturnal than diurnal activity in insect communities. *Nat Commun*. 2024;**15**:3236. <https://doi.org/10.1038/s41467-024-47645-2>.
51. Williams CB. Studies in the effect of weather conditions on the activity and abundance of insect populations. *Philos Trans R Soc Lond B Biol Sci*. 1961;**244**:331–78. <https://doi.org/10.1098/rstb.1961.0011>.
52. Ratnasingham S, Wei C, Chan D *et al.* BOLD v4: a centralized bioinformatics platform for DNA-based biodiversity data. In: DeSalle R, ed. *DNA Barcoding: Methods and Protocols*. New York, NY: Springer US; 2024:403–41, Chapter 26.
53. Hemprich-Bennett DR, Donkor E, Adams B *et al.* Supporting data for “Characterising a species-rich and understudied tropical insect fauna using DNA barcoding.” GigaScience Database. 2026. <https://doi.org/10.5524/102812>.
54. Earthcape: Earthcape. <https://earthcape.com>. Accessed 12 Mar 2026.
55. Target Malaria: Target Malaria. www.targetmalaria.org. Accessed 12 Mar 2026.