

A DNA barcode survey of insect biodiversity in Pakistan

Muhammad Ashfaq¹, Arif M. Khan², Akhtar Rasool³, Saleem Akhtar⁴, Naila Nazir⁵, Nazeer Ahmed⁶, Farkhanda Manzoor⁷, Jayme Sones⁸, Kate Perez⁸, Ghulam Sarwar⁹, Azhar A. Khan¹⁰, Muhammad Akhter¹¹, Shafqat Saeed¹², Riffat Sultana¹³, Hafiz Muhammad Tahir¹⁴, Muhammad A. Rafi¹⁵, Romana Iftikhar¹⁶, Muhammad Tayyib Naseem¹⁷, Mariyam Masood¹⁸, Muhammad Tufail¹⁹, Santosh Kumar²⁰, Sabila Afzal²¹, Jaclyn McKeown⁸, Ahmed Ali Samejo¹³, Imran Khaliq¹⁹, Michelle L. D'Souza⁸, Shahid Mansoor²² and Paul D. N. Hebert¹

¹ Centre for Biodiversity Genomics & Department of Integrative Biology, University of Guelph, Guelph, Canada

² Department of Biotechnology, University of Sargodha, Sargodha, Pakistan

³ Centre for Animal Sciences and Fisheries, University of Swat, Mingora, Pakistan

⁴ Directorate of Entomology, Ayub Agricultural Research Institute, Faisalabad, Pakistan

⁵ Department of Entomology, University of Poonch, Rawalakot, Azad Kashmir, Pakistan

⁶ Faculty of Life Sciences and Informatics, Balochistan University of Information Technology, Engineering and Management Sciences, Quetta, Pakistan

⁷ Department of Zoology, Lahore College for Women University, Lahore, Pakistan

⁸ Centre for Biodiversity Genomics, University of Guelph, Guelph, Canada

⁹ Institute of Zoology, University of the Punjab, Lahore, Pakistan

¹⁰ College of Agriculture, Bahauddin Zakariya University Bahadur Campus, Layyah, Pakistan

¹¹ Pulses Research Institute, Ayub Agricultural Research Institute, Faisalabad, Pakistan

¹² Faculty of Agriculture and Environmental Sciences, MNS University of Agriculture, Multan, Pakistan

¹³ Department of Zoology, University of Sindh, Jamshoro, Pakistan

¹⁴ Department of Zoology, Government College University Lahore, Lahore, Pakistan

¹⁵ National Insect Museum, National Agricultural Research Center, Islamabad, Pakistan

¹⁶ Department of Plant Pathology, Washington State University, Pullman, WA, United States

¹⁷ Department of Biology, University of Copenhagen, Copenhagen, Denmark

¹⁸ Government College Women University Faisalabad, Faisalabad, Pakistan

¹⁹ Ghazi University, Dera Ghazi Khan, Pakistan

²⁰ Department of Zoology, Cholistan University of Veterinary and Animal Sciences, Bahawalpur, Pakistan

²¹ Department of Zoology, University of Narowal, Narowal, Pakistan

²² National Institute for Biotechnology and Genetic Engineering, Faisalabad, Pakistan

Submitted 30 December 2021

Accepted 23 March 2022

Published 25 April 2022

Corresponding author

Muhammad Ashfaq,

mashfaq@uoguelph.ca

Academic editor

Joseph Gillespie

Additional Information and
Declarations can be found on
page 11

DOI 10.7717/peerj.13267

© Copyright

2022 Ashfaq et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

ABSTRACT

Although Pakistan has rich biodiversity, many groups are poorly known, particularly insects. To address this gap, we employed DNA barcoding to survey its insect diversity. Specimens obtained through diverse collecting methods at 1,858 sites across Pakistan from 2010–2019 were examined for sequence variation in the 658 bp barcode region of the cytochrome *c* oxidase 1 (COI) gene. Sequences from nearly 49,000 specimens were assigned to 6,590 Barcode Index Numbers (BINs), a proxy for species, and most (88%) also possessed a representative image on the Barcode of Life Data System (BOLD). By coupling morphological inspections with barcode matches on BOLD, every BIN was assigned to an order (19) and most (99.8%) were placed to a family (362). However, just 40% of the BINs were assigned to a genus (1,375) and

21% to a species (1,364). Five orders (Coleoptera, Diptera, Hemiptera, Hymenoptera, Lepidoptera) accounted for 92% of the specimens and BINs. More than half of the BINs (59%) are so far only known from Pakistan, but others have also been reported from Bangladesh (13%), India (12%), and China (8%). Representing the first DNA barcode survey of the insect fauna in any South Asian country, this study provides the foundation for a complete inventory of the insect fauna in Pakistan while also contributing to the global DNA barcode reference library.

Subjects Ecology, Entomology, Genetics, Taxonomy, Population Biology

Keywords DNA barcoding, Cytochrome c oxidase I, Barcode index number, Biodiversity overlap, BOLD

INTRODUCTION

With an area of 882,000 km², Pakistan includes seven biomes (Deserts & Xeric Shrublands, Flooded Grasslands & Savannas, Mangroves, Montane Grasslands & Shrublands, Temperate Broadleaf & Mixed Forests, Temperate Conifer Forests, Tropical & Subtropical Coniferous Forests) and portions of the Indo-Malayan and Palearctic biogeographic realms (*Dinerstein et al., 2017*). Because of this physiographic and climatic variation, its faunal diversity is quite high (*Anwar, Jasra & Ahmad, 2008*). While its vertebrate fauna is well known (*Khan, 1976; Roberts, 1997; Grimmett, Roberts & Inskipp, 2008*), prior studies on other animal lineages have been restricted to specific taxa or regions (*Inayat et al., 2010; Iftikhar et al., 2016a; Manzoor, Khan & Shah, 2020*). Although over 5,000 insect species have been reported (*Government of Pakistan, 2000*), no taxonomically comprehensive or wide-ranging assessments have been undertaken due to the scarcity of taxonomists, and the fact that many insect species remain undescribed (*Ward & Lariviere, 2004*). Because conventional morphological approaches (*Pompeo et al., 2017*) are difficult to implement at scale (*Fattorini, 2013*), the species count for Pakistan remains uncertain as it does globally (*Scheffers et al., 2012*).

The effectiveness of DNA barcoding (*Hebert et al., 2003*) in both specimen identification and species discovery (*Hebert et al., 2004; Huemer et al., 2014; Kress et al., 2015*) has stimulated its rapid adaptation in modern biodiversity studies (*Ashfaq & Hebert, 2016; DeSalle & Goldstein, 2019*). This work has generated DNA barcode coverage for more than 760,000 animal species on the Barcode of Life Data System (BOLD) (www.boldsystems.org) (*Ratnasingham & Hebert, 2007*). The effectiveness of the Barcode Index Number (BIN) system (*Ratnasingham & Hebert, 2013*) as a species proxy (*Hausmann et al., 2013*) has made it possible to rapidly evaluate species diversity, enabling large-scale biotic inventories (*Hebert et al., 2016; Wirta et al., 2016*). Because BINs show close congruence with species boundaries established through morphological study (*Ortiz et al., 2017; Huemer et al., 2019*) they can be used to delineate newly encountered species (*Mitchell, Moeseneder & Hutchinson, 2020*), to discern cryptic species (*Zahiri et al., 2017; Zhou et al., 2019*), to plot species distributions (*Ren et al., 2018*), to estimate species richness in bulk samples (*Andújar et al., 2018; Braukmann et al., 2019*), to analyze museum

collections (*Pentinsaari et al., 2020*), and to assess faunal similarity at regional and global scales (*Ashfaq et al., 2017*).

The effectiveness of DNA barcoding coupled with advances in sequencing technology allow it to support large-scale biodiversity analysis (*Wilson et al., 2017*). However, the intensity of study has varied among regions (*Weigand et al., 2019*). For example, the BIN count (84,000) for Canada is $8\times$ that for Russia (11,000), although the latter nation is $1.7\times$ larger (www.boldsystems.org, accessed 7 September 2021). In a similar fashion, the BIN count for Germany (23,000) is $4\times$ that for India (5,800), although the latter nation is $9\times$ larger. The current study extends DNA barcode coverage for Pakistan to both advance knowledge of the taxonomic composition of its insect fauna and to develop a barcode reference library that supports routine eDNA and metabarcoding studies in the future.

MATERIALS AND METHODS

Sample collection and preparation

Insects were sampled at 1,858 sites across Pakistan (*Fig. 1*) from 2010–2019 using both active and passive collecting methods including sweep nets, hand collections, hostplant beating, light traps, Malaise traps, pitfall traps, and UV illuminated sheets. Plans for the specimen collections/sites were approved by the Director, National Institute for Biotechnology and Genetic Engineering, Faisalabad under the project HEC No. 20-1403/R& D/09. The specimens were identified to an order and, where possible, to lower taxonomic ranks. Large specimens were either pinned and preserved dry or placed in Matrix tubes with 95% ethanol. Small specimens were individually placed in a well containing 30 μl of 95% ethanol in 96-well microplates. Specimen metadata and an image (except for Malaise samples where only representative specimens of each BIN were imaged) were submitted to BOLD where the information can be accessed on both the specimen page and corresponding BIN page. Voucher specimens are archived at the National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan (with sample ID prefix NIBGE) or at the Centre for Biodiversity Genomics (CBG), Guelph, Canada (with ID prefix BIOUG).

DNA barcoding

A total of 60,273 insects were barcoded following standard protocols (*deWaard et al., 2019a, 2019b*). In brief, a leg was removed with sterile forceps from each large specimen and transferred to a well preloaded with 30 μl of 95% EtOH. As smaller specimens were already in plates, they were ready for analysis, but vouchers were recovered after DNA extraction (*Porco et al., 2010*). DNA extraction, PCR amplification, and sequencing were performed at the Canadian Centre for DNA Barcoding (CCDB) following established protocols (*Ivanova, deWaard & Hebert, 2006; Hebert et al., 2018; deWaard et al., 2019b*). PCR reactions were either 6 μl or 12 μl (*Hebert et al., 2013*). Three quarters (73%) of the specimens were Sanger sequenced while the rest were analyzed using SMRT sequencing on a Sequel platform (Pacific Biosciences, Menlo Park, CA, USA). Sanger sequencing employed BigDye Terminator Cycle Sequencing Kit (v3.1) on an Applied Biosystems 3730XL DNA Analyzer. Sequences were assembled, aligned and edited using CodonCode

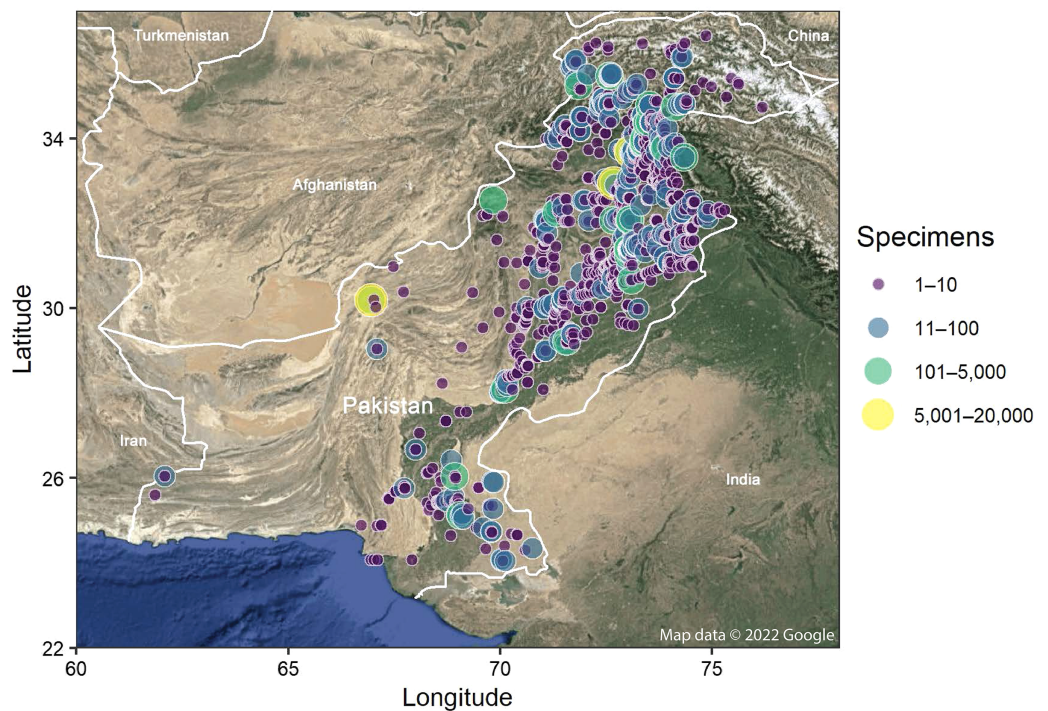


Figure 1 Map showing the collection sites for the insects examined in this study. The size and color of each site point indicate the number of specimens sampled. Map was generated in R using Google Maps satellite imagery. [Full-size !\[\]\(5f471a71b78d7676bc356df190b88ab4_img.jpg\) DOI: 10.7717/peerj.13267/fig-1](https://doi.org/10.7717/peerj.13267/fig-1)

Aligner before submission to BOLD. SMRT sequencing employed protocols described by [Hebert et al. \(2018\)](#). The resultant sequences were uploaded to mBRAVE (Multiplex Barcoding Research and Visualization Environment; www.mbrave.net) for editing (sequence trimming, quality filtering, de-replication), identification, and generation of operational taxonomic units (OTUs). The edited sequences were subsequently exported to BOLD for BIN assignment and reference library development. The specimen records, sequence data, electropherograms, and primer details are available in the dataset “DS-INSCTPAK” (dx.doi.org/10.5883/DS-INSCTPAK). All DNA extracts are stored within the DNA archive facility at the CBG.

Data analysis

The final dataset ($N = 50,592$) included 50,094 new barcode records and 498 public records on BOLD from specimens collected in Pakistan ([Table S1](#)). All records were assigned taxonomy and BINs following the workflow outlined by [deWaard et al. \(2019b\)](#). In brief, once the barcode data was on BOLD, each record went through a taxonomic assignment and verification workflow. Earlier studies ([Ashfaq et al., 2013](#); [Nazir et al., 2014](#); [Iftikhar et al., 2016b](#); [Akhtar et al., 2018](#); [Naseem et al., 2019](#)) on five taxa (antlions, aphids, butterflies, grasshoppers, thrips) coupled analysis of barcode results with detailed morphological study by taxonomic specialists. All sequences meeting the quality criteria were either assigned to an existing BIN or founded a new one ([Ratnasingham & Hebert, 2013](#)). Sequences founding a new BIN had to possess >500 bp of the barcode region with

<1% ambiguous bases and no stop codons. Shorter sequences (300–495) that met the latter two quality criteria and that were a close sequence match to an established BIN were assigned to it (*deWaard et al., 2019a*). The remaining short sequences (1,230) that failed to gain a BIN assignment were run through the stand-alone version of the RESL algorithm (*Ratnasingham & Hebert, 2013*) (using the function Cluster Sequences on BOLD) to estimate the number of additional OTUs among them. One representative from each OTU was then queried against the BOLD ID Engine to link them with known BINs (*deWaard et al., 2019b*). The BIN details with specimen records and representative images (where available) are accessible on BOLD (dx.doi.org/10.5883/DS-INSCTPAK).

Various statistical approaches were used to estimate the number of insect species in Pakistan (*Chao & Chiu, 2016*) including the parametric estimator Preston's log-normal as well as non-parametric estimators *Chao1*, and the *first-order* and *second-order jackknife*. A bias-corrected version of each non-parametric estimator, designed to improve performance under conditions of low sampling effort, was also included (*Lopez et al., 2012*). All estimates were calculated using the R packages *vegan* and *BAT*. In addition, a species accumulation curve was drawn based on a sample-size-based rarefaction and extrapolation to at most double the minimum observed sample size, guided by an estimated asymptote using the R package *iNEXT* (*Hsieh, Ma & Chao, 2016*).

RESULTS

DNA barcodes were recovered from 50,094 (83%) of the 60,273 specimens analyzed. The other 17% either failed to amplify or generated problematic sequences (*e.g.*, contamination, NUMTs, stop codons, endosymbionts) that were excluded from subsequent analysis. Considering orders with 100 or more specimens, sequence recovery ranged from a low of 63% for Blattodea to 95% for Lepidoptera. Sequence recovery for the other four major orders of insects showed considerable variation (Diptera: 91%, Coleoptera: 80%, Hymenoptera: 78%, Hemiptera: 69%).

All 50,592 insects with a barcode were assigned to one of 19 orders while 99.8% received an assignment to one of 362 families ([Table 1](#), [Table S1](#)). Five orders represented 92% of the specimens: Diptera (40%), Hymenoptera (21%), Lepidoptera (12%), Hemiptera (11%), and Coleoptera (8%) ([Fig. 2](#)). Six orders (Mantodea, Neuroptera, Odonata, Orthoptera, Psocodea, Thysanoptera) were each represented by >100 specimens while the remaining eight possessed fewer representatives ([Fig. 2](#), [Table S1](#)). Most of these sequences (98%) received a BIN assignment, leading to a total of 6,590 BINs. The other 1,230 barcode sequences did not meet the criteria for BIN assignment but included 629 OTUs when analyzed using "Cluster Sequences" function on BOLD. The BOLD ID Engine assigned 82 of these OTUs to known BINs, but the other 547 OTUs likely represent taxa new to BOLD. Many (57%) of the 6,590 BINs were represented by two or more sequences, but 43% were represented by just a single specimen. The ratio of these singletons was above 40% in all five major orders but was highest in Hymenoptera (48%). Most BINs (88%; $N = 5,754$) possessed an image of at least one voucher.

The percentage of records in each of the five major orders with a BIN assignment ranged from 93% (Coleoptera) to 99% (Diptera, Lepidoptera) with Hemiptera and Hymenoptera

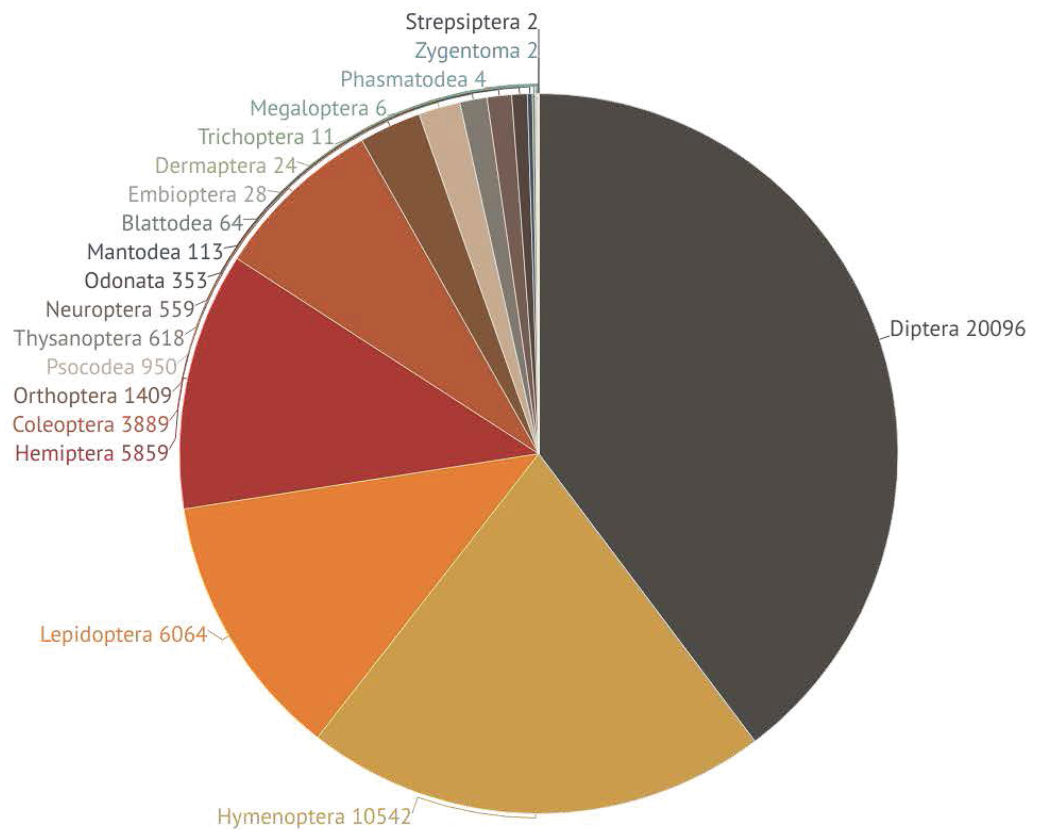


Figure 2 Pie chart showing the number of specimens barcoded from each of the 19 insect orders. Different colors represent different orders. Numbers next to each slice indicate the specimen count for the order.

Full-size DOI: [10.7717/peerj.13267/fig-2](https://doi.org/10.7717/peerj.13267/fig-2)

intermediate (96%) (Table 1). These five orders also contributed most of the BINs (92%) and families (81%) (Table 1, Figs. 3A, 3B). Only 40% of BINs were placed to a genus and 21% to a species, but this still led to records for 1,375 genera and 1,364 species (Table 1, Table S1). Among the five major orders, more BINs were identified to a genus (72%) and species (41%) in Lepidoptera than in the other four orders (Table 1). For example, just 13.8% of Diptera BINs and 10.2% of Hymenoptera BINs were assigned to a species (Table 1).

Specimen counts for the 362 families were highly variable as 15 families were each represented by >1,000 specimens while 38 had just one (Table S1). This pattern was also reflected in the number of BINs as 15 families had >100 BINs while 86 had just one. The Chironomidae ($N = 3,258$) and Braconidae ($N = 2,174$) were represented by the most specimens while Cecidomyiidae (238 BINs) and Platygastriidae (230 BINs) were most diverse. Figure 4 shows the BIN diversity and BIN:specimen ratio for the 15 families with >100 BINs. The ratio was highest (0.33) for Geometridae (Lepidoptera) and lowest (0.05) for Chironomidae. The species accumulation curve did not reach an asymptote indicating more species await detection (Fig. 5). Species estimates for the country ranged from 9,253 to 12,246 species suggesting that, on average, 40% of species remain to be sampled (Table 2).

Table 1 Number of specimens belonging to 19 insect orders from Pakistan with DNA barcode records. The number of families, genera, species, and BINs is reported for each order.

Order	Specimens with barcodes	Specimens assigned to BINs (%)	BINs recovered	OTUs without BIN*	Singleton BINs (%)	BINs assigned to family (%)	Families recovered	BINs assigned to genus (%)	Genera recovered	BINs assigned to species (%)	Species recovered
Blattodea	64	84	19	5	36.8	100	5	78.9	9	52.6	10
Coleoptera	3,889	93	819	123	45.2	100	56	21.9	118	13.3	119
Dermaptera	24	83.3	3	2	33.3	100	2	33.3	1	0.0	0
Diptera	20,095	99	1,684	94	40.1	99.0	68	29.8	212	13.8	222
Embioptera	28	96.4	7	1	14.3	100	2	14.3	1	14.3	1
Hemiptera	5,859	96.5	642	73	41.9	98.3	59	31.6	132	22.6	135
Hymenoptera	10,542	96	1,711	177	47.7	99.4	50	34.7	226	10.2	170
Lepidoptera	6,064	99.4	1,233	24	42.5	99.6	62	71.9	516	41.5	514
Mantodea	113	97.3	36	2	50.0	100	2	13.9	4	5.6	2
Megaloptera	6	100	1	0	0.0	100	1	100	1	100	1
Neuroptera	559	92.3	99	6	39.4	99.0	7	54.5	30	36.4	32
Odonata	353	92.6	51	11	21.6	100	12	92.2	30	88.2	47
Orthoptera	1,409	97.59	163	21	30.1	100	12	44.2	53	37.4	54
Phasmatodea	4	75	3	1	100.0	100	1	0.0	0	0.0	0
Psocodea	950	97.5	31	5	22.6	93.5	13	38.7	10	19.4	6
Strepsiptera	2	100	1	0	0.0	100	1	100	1	0.0	0
Thysanoptera	618	99.3	76	2	34.2	100	3	80.3	27	69.7	48
Trichoptera	11	100	10	0	90.0	100	6	60.0	4	40.0	4
Zygentoma	2	100	1	0	0.0	0.0	0	0.0	0	0.0	0
Total	50,592	97.6%	6,590	547	42.9%	99%	362	40%	1,375	21%	1,364

Note:

* For recognition as a new BIN, a sequence must include >500 bp of the barcode region (positions 70 bp to 700 bp in the BOLD alignment) and possess <1% ambiguous bases.

BOLD was searched to ascertain if the 6,590 insect BINs from Pakistan were known from other countries. This analysis showed that 2,684 BINs (41%) were shared with at least one of 199 other countries while the others (3,906) are so far only known from Pakistan. The percentage of shared species ranged from 0.02% to 13%. [Figure 6](#) shows the overlap values between Pakistan and countries with >1,000 BINs. BIN overlap was higher with nearby countries (Bangladesh: 13%, India: 12%, China: 8%) than for other regions. For example, Pakistan shared just 5% of its BINs with Australia, South Africa, and Germany ([Fig. 6](#)). The overlap between Canada and Costa Rica, both countries with >50,000 insect BINs, was only 4% and 1% respectively ([Fig. 6](#)).

DISCUSSION

Current estimates of the number of insect species which occur in Pakistan range from 5,000 (*Ministry of Climate Change, Pakistan, 2019*) to 20,000 species (*Hasnain, 1998*), but they are certainly too low (*Baig & Al-Subaiee, 2009*). The current study aimed to refine estimates of species richness by coupling DNA barcoding with the BIN system. With over 50,000 specimens sequenced, this study represents, by far, the largest effort to assemble a

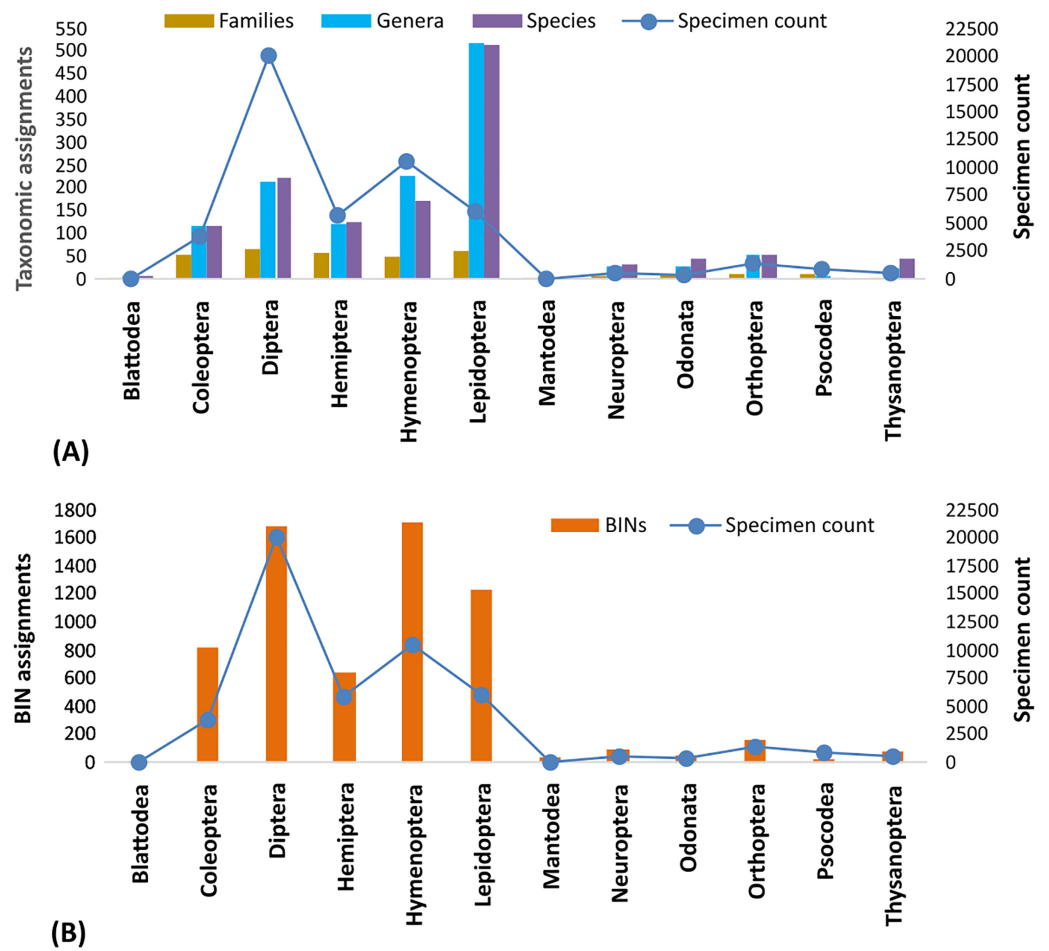


Figure 3 Taxonomic (A) and BIN assignments (B) for the 12 insect orders represented by >50 specimens. Species assignment in (A) is based on the assignment of barcode(s) to the named species on BOLD. [Full-size !\[\]\(5fd6ef84f97f42d7f8b34275f1b65312_img.jpg\) DOI: 10.7717/peerj.13267/fig-3](https://doi.org/10.7717/peerj.13267/fig-3)

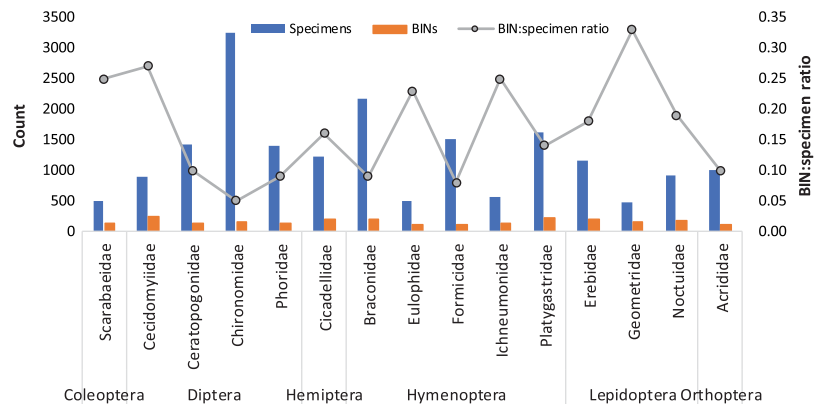


Figure 4 BIN diversity and BIN: specimen ratio for the 15 insect families represented by >100 BINs. [Full-size !\[\]\(b8ddfb9d90db8697d6b8ef7f72522b2e_img.jpg\) DOI: 10.7717/peerj.13267/fig-4](https://doi.org/10.7717/peerj.13267/fig-4)

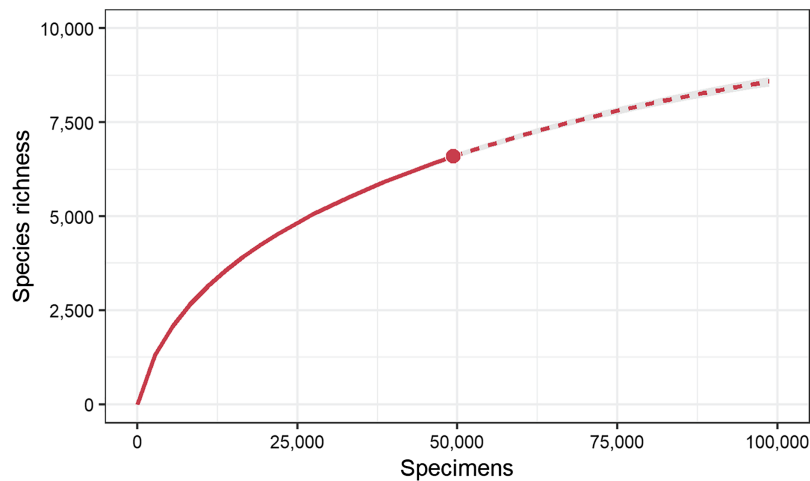


Figure 5 Sample-size-based rarefaction (solid line) and extrapolation (dashed line) sampling curves for 49,363 specimens with barcodes from Pakistan. Solid dots represent the observed richness of 6,590 species. The curve is estimated to reach an asymptote at 10,382 species.

Full-size DOI: 10.7717/peerj.13267/fig-5

Table 2 Species richness estimates based on the abundances of the 6,590 insect BINS encountered at 1,858 sites across Pakistan.

SPECIMENS	BINS	PRESTON	CHAO1	CHAO1P	JACK1AB	JACK1ABP	JACK2AB	JACK2ABP
49,363	6,590	9,253	10,377	12,285	9,416	11,147	11,189	12,246

Note:

Seven estimates were calculated: Preston's log-normal (PRESTON), Chao1 (CHAO1), first-order jackknife (JACK1AB), second-order jackknife (JACK2AB), and their bias-corrected complements (CHAO1P, JACK1ABP, JACK2ABP).

DNA barcode registry for the insect fauna of any South Asian country. While success (83%) in DNA barcode recovery was good, it varied considerably among orders from 63% for Blattodea to 95% for Lepidoptera. Similar variation in barcode recovery among different insect taxa has been reported in other studies (Geiger *et al.*, 2016; Pentinsaari *et al.*, 2020). For example, a study on the insect fauna of French Polynesia reported 91% recovery for Diptera vs 63% for Coleoptera (Ramage *et al.*, 2017). Similarly, a large-scale Canadian study revealed 95% recovery for Diptera vs 77% for Hemiptera and 74% for Hymenoptera (deWaard *et al.*, 2019a). Although DNA quantity and quality play an important role (Ballare *et al.*, 2019; Velasco-Cuervo *et al.*, 2019), failures in primer binding often underlie low sequence recovery (Hajibabaei *et al.*, 2005; Hebert *et al.*, 2016). Such failures can lead to the underestimation of species richness in insect groups where recovery is low (Hebert *et al.*, 2016). Other factors, such as co-amplification of pseudogenes (Song *et al.*, 2008), Wolbachia (Smith *et al.*, 2012), recent speciation (van Velzen *et al.*, 2012), or incomplete lineage sorting (Mallo & Posada, 2016) may also limit the efficacy of barcodes to delimitate species, consequently affecting the diversity estimates. Moreover, there are instances where the BIN system overestimated species diversity in certain insect groups, such as Chironomidae (Lin, Stur & Ekrem, 2015; Ekrem *et al.*, 2018).

The coupling of morphological inspection with barcode matches on BOLD (deWaard *et al.*, 2019a, 2019b) was very effective at placing BINs to an order (100%) and family

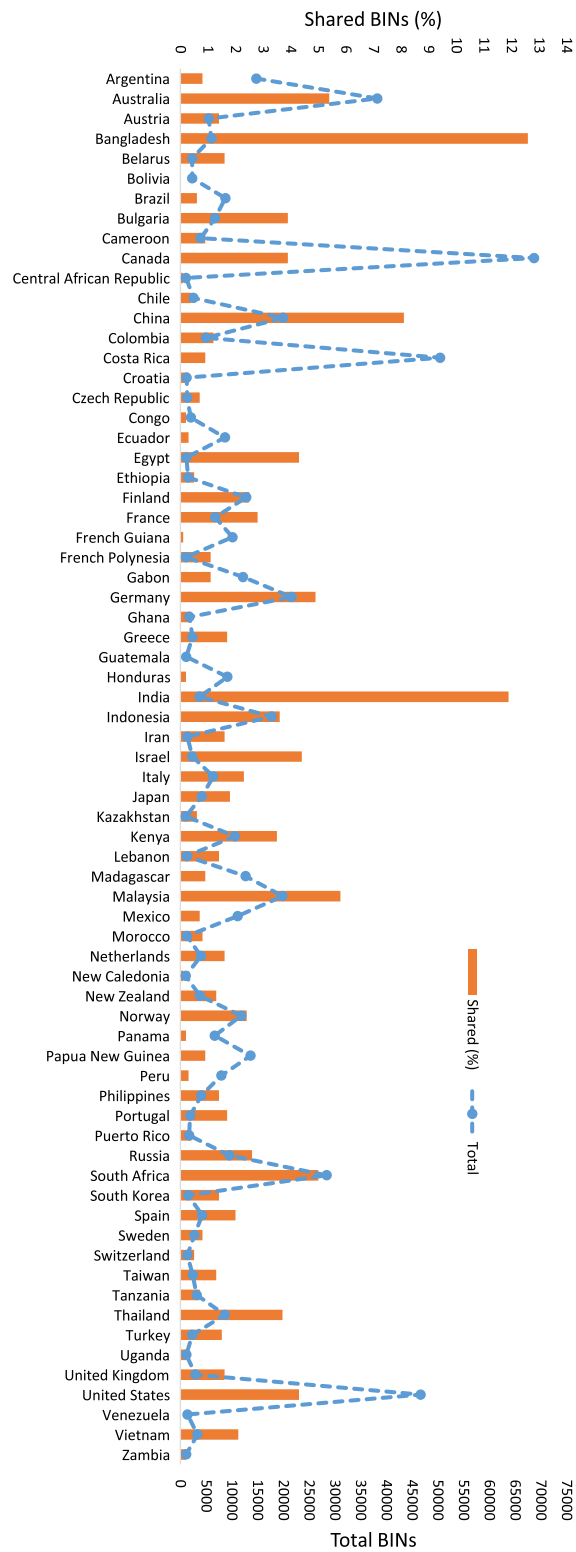


Figure 6 Percentage of insect BINs shared between Pakistan and the 70 other nations with >1,000 insect BINs on the Barcode of Life Data Systems (BOLD).

Full-size  DOI: 10.7717/peerj.13267/fig-6

(>99%). However, just 40% of the BINs could be assigned to a genus and 21% to a species indicating the need for better parameterization of the barcode reference library. This was particularly true for the three most diverse orders where species assignments were less than 15% (Diptera: 13.8%, Coleoptera: 13.3%, Hymenoptera: 10.2%). Considerably higher assignment success has been reported for Malaise samples from Germany (34%) and Canada (38%) (*Geiger et al., 2016; deWaard et al., 2019a*) reflecting the more comprehensive DNA barcode reference libraries available for these nations. Despite the limited reference database (*Virgilio et al., 2010*), the present analysis identified representatives from 1,375 genera and 1,364 species showing the value of the global reference library (BOLD) which far exceeds the results obtained by morphology alone (*Marshall, Paiero & Buck, 2009*). The present analysis revealed 6,590 BINs with species richness estimates indicating that the fauna of Pakistan certainly includes more than 10,000 species. As these estimates are based on specimens collected with uneven sampling and limited geographic coverage, they are likely to increase with more comprehensive efforts.

Although 19 insect orders were detected, five (Coleoptera, Diptera, Hemiptera, Hymenoptera, Lepidoptera) were dominant (92%), reinforcing prior results from morphological (*Stork, 2018*) and barcoding studies (*Ritter et al., 2019; Pentinsaari et al., 2020*). Malaise traps preferentially capture low-flying insects such as Diptera and Hymenoptera (*Cooksey & Barton, 1981; deWaard et al., 2019b*), the two orders that made 61% of the collections. Other studies have reported a similar pattern (*Brown, 2005; Karlsson et al., 2020*). For example, a Canadian study found that Diptera comprised 57% of the collections (*deWaard et al., 2019b*).

Fifteen of the 362 families dominated with 1,000 or more specimens and this pattern was also reflected in the BIN diversity. The uneven detection of families in the survey is supported by the fact that 38 families were represented by just one specimen and 88 by one BIN. Interestingly, nine of the 15 families with the most BINs were dipterans and hymenopterans with Cecidomyiidae and Ichneumonidae comprising the highest BIN: specimen ratio.

Because BOLD now hosts around nine million DNA barcode records for more than 760,000 animal species, it provides a good basis for assessing faunal overlap using BINs. Only 41% of the 6,590 insect BINs from Pakistan are currently known from other countries. As expected, BIN overlap was highest with neighboring countries. This result reflects the endemism of biodiversity (*Werneck et al., 2012*) and underscores the need to develop local biodiversity inventories. The current survey represents a first step towards building an inventory for the insect fauna of Pakistan.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This study was enabled by grant 106106-001 “Engaging Developing Nations in iBOL” from the International Development Research Centre (IDRC) in Canada and by grant HEC No. 20-1403/R& D/09 “Sequencing DNA Barcodes of Economically Important Insect Species

from Pakistan” from the Higher Education Commission of Pakistan. Sequence analysis was made possible by a grant from the Government of Canada through Genome Canada and Ontario Genomics in support of the International Barcode of Life (iBOL) project. This is a contribution to the ‘Food from Thought’ project supported by the Canada First Research Excellence Fund. Paul Hebert and Nazeer Ahmed were supported by the New Frontiers in Research Fund through its award to BIOSCAN. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

International Development Research Centre (IDRC) in Canada: 106106-001.

Higher Education Commission of Pakistan: 20-1403/R& D/09.

Government of Canada.

Canada First Research Excellence Fund.

New Frontiers in Research Fund.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Muhammad Ashfaq conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Arif M. Khan conceived and designed the experiments, performed the experiments, prepared figures and/or tables, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Akhtar Rasool conceived and designed the experiments, performed the experiments, prepared figures and/or tables, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Saleem Akhtar conceived and designed the experiments, performed the experiments, prepared figures and/or tables, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Naila Nazir conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Nazeer Ahmed conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, sample collections, and approved the final draft.
- Farkhanda Manzoor conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, sample collections, and approved the final draft.

- Jayme Sones conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Kate Perez conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Ghulam Sarwar conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Azhar A. Khan conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Muhammad Akhter conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Shafqat Saeed conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the paper, specimen identification, and approved the final draft.
- Riffat Sultana conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Hafiz Muhammad Tahir conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the paper, specimen identification, and approved the final draft.
- Muhammad A. Rafi conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the paper, specimen identification, and approved the final draft.
- Romana Iftikhar conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Muhammad Tayyib Naseem conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Mariyam Masood conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Muhammad Tufail conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Santosh Kumar conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Sabila Afzal conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, sample collections, and approved the final draft.

- Jaclyn McKeown performed the experiments, analyzed the data, authored or reviewed drafts of the paper, specimen imaging, and approved the final draft.
- Ahmed Ali Samejo conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Imran Khaliq conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Michelle L. D'Souza analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Shahid Mansoor conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, sample collections, and approved the final draft.
- Paul D. N. Hebert conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, specimen identifications, and approved the final draft.

Field Study Permissions

The following information was supplied relating to field study approvals (*i.e.*, approving body and any reference numbers):

Field experiments were approved by the Director National Institute for Biotechnology and Genetic Engineering Faisalabad (project number; HEC No. 20-1403/R& D/09).

Data Availability

The following information was supplied regarding data availability:

The sequences are available at Barcode of Life Data Systems: dx.doi.org/10.5883/DS-INSCTPAK.

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.13267#supplemental-information>.

REFERENCES

- Akhtar S, Ashfaq M, Zia A, Ali S, Ali GM, Zafar Y, Farhatullah. 2018. First report and redescription of five species of genus *Myrmeleon* (Neuroptera: Myrmeleontidae) from Pakistan. *Journal of Biodiversity and Environmental Sciences* **13**(1):180–190.
- Andújar C, Arribas P, Gray C, Bruce C, Woodward G, Yu DW, Vogler AP. 2018. Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill. *Molecular Ecology* **27**:146–166 DOI [10.1111/mec.14410](https://doi.org/10.1111/mec.14410).
- Anwar M, Jasra AW, Ahmad I. 2008. Biodiversity conservation status in Pakistan—a review. *Pakistan Journal of Forestry* **58**(1):39–48.
- Ashfaq M, Akhtar S, Khan AM, Adamowicz SJ, Hebert PDN. 2013. DNA barcode analysis of butterfly species from Pakistan points towards regional endemism. *Molecular Ecology Resources* **13**:832–843 DOI [10.1111/1755-0998.12131](https://doi.org/10.1111/1755-0998.12131).

- Ashfaq M, Akhtar S, Rafi MA, Mansoor S, Hebert PDN. 2017. Mapping global biodiversity connections with DNA barcodes: lepidoptera of Pakistan. *PLOS ONE* 12:e0174749 DOI 10.1371/journal.pone.0174749.
- Ashfaq M, Hebert PDN. 2016. DNA barcodes for bio-surveillance: regulated and economically important arthropod plant pests. *Genome* 59:933–945 DOI 10.1139/gen-2016-0024.
- Baig MB, Al-Subaiee FS. 2009. Biodiversity in Pakistan: key issues. *Biodiversity* 10:20–29 DOI 10.1080/14888386.2009.9712858.
- Ballare KM, Pope NS, Castilla AR, Cusser S, Metz RP, Jha S. 2019. Utilizing field collected insects for next generation sequencing: effects of sampling, storage, and DNA extraction methods. *Ecology and Evolution* 9:13690–13705 DOI 10.1002/ece3.5756.
- Braukmann TWA, Ivanova NV, Prosser SWJ, Elbrecht V, Steinke D, Ratnasingham S, de Waard JR, Sones JE, Zakharov EV, Hebert PDN. 2019. Metabarcoding a diverse arthropod mock community. *Molecular Ecology Resources* 19:711–727 DOI 10.1111/1755-0998.13008.
- Brown B. 2005. Malaise trap catches and the crisis in Neotropical Dipterology. *American Entomologist* 51:180–183 DOI 10.1093/ae/51.3.180.
- Chao A, Chiu C. 2016. Nonparametric estimation and comparison of species richness. In: *eLS*. Chichester: John Wiley & Sons, Ltd. DOI 10.1002/9780470015902.a0026329.
- Cooksey LM, Barton HE. 1981. Flying insect populations as sampled by Malaise trap on Crowley's Ridge in northeast Arkansas. *Journal of Arkansas Academy of Sciences* 35:29–32.
- DeSalle R, Goldstein P. 2019. Review and interpretation of trends in DNA barcoding. *Frontiers of Ecology and Evolution* 7:302 DOI 10.3389/fevo.2019.00302.
- deWaard JR, Levesque-Beaudin V, deWaard SL, Ivanova NV, McKeown JTA, Miskie R, Naik S, Perez KHJ, Ratnasingham S, Sobel CN, Sones JE, Steinke C, Telfer AC, Young AD, Young MR, Zakharov EV, Hebert PDN, Wilson JJ. 2019b. Expedited assessment of terrestrial arthropod diversity by coupling Malaise traps with DNA barcoding. *Genome* 62(3):85–95 DOI 10.1139/gen-2018-0093.
- deWaard JR, Ratnasingham S, Zakharov EV, Borisenko AV, Steinke D, Telfer AC, Perez KHJ, Sones JE, Young MR, Levesque-Beaudin V, Sobel CN, Abrahamyan A, Bessonov K, Blagoev G, deWaard SL, Ho C, Ivanova NV, Layton KKS, Lu L, Manjunath R, McKeown JTA, Milton MA, Miskie R, Monkhouse N, Naik S, Nikolova N, Pentinsaari M, Prosser SWJ, Radulovici AE, Steinke C, Warne CP, Hebert PDN. 2019a. A reference library for the identification of Canadian invertebrates: 1.5 million DNA barcodes, voucher specimens, and genomic samples. *Scientific Data* 6(1):308 DOI 10.1038/s41597-019-0320-2.
- Dinerstein E, Olson D, Joshi A, Vynne C, Burgess ND, Wikramanayake E, Hahn N, Palminteri S, Hedao P, Noss R, Hansen M, Locke H, Ellis EC, Jones B, Barber CV, Hayes R, Kormos C, Martin V, Crist E, Sechrest W, Price L, Baillie JEM, Weeden D, Suckling Kán, Davis C, Sizer N, Moore R, Thau D, Birch T, Potapov P, Turubanova S, Tyukavina A, de Souza N, Pintea L, Brito Jé C, Llewellyn OA, Miller AG, Patzelt A, Ghazanfar SA, Timberlake J, Klöser H, Shennan-Farpón Y, Kindt R, Lillesø JPB, van Breugel P, Graudal L, Voge M, Al-Shammari KF, Saleem M. 2017. An ecoregion-based approach to protecting half the terrestrial realm. *Bioscience* 67(6):534–545 DOI 10.1093/biosci/bix014.
- Ekrem T, Stur E, Orton MG, Adamowicz SJ. 2018. DNA barcode data reveal biogeographic trends in Arctic non-biting midges. *Genome* 61:787–796 DOI 10.1139/gen-2018-0100.
- Fattorini S. 2013. Regional Insect inventories require long time, extensive spatial sampling and good will. *PLOS ONE* 8:e62118 DOI 10.1371/journal.pone.0062118.
- Geiger M, Moriniere J, Hausmann A, Haszprunar G, Wägele W, Hebert PDN, Rulik B. 2016. Testing the global malaise trap program—how well does the current barcode reference library

- identify flying insects in Germany? *Biodiversity Data Journal* 4:e10671
DOI 10.3897/BDJ.4.e10671.
- Government of Pakistan. 2000.** Biodiversity action plan for Pakistan. Available at <https://portals.iucn.org/library/efiles/documents/2000-081.pdf>.
- Grimmett R, Roberts T, Inskipp T. 2008.** *Birds of Pakistan*. London: A & C Black Publishers Ltd., 255.
- Hajibabaei M, deWaard JR, Ivanova NV, Ratnasingham S, Dooh RT, Kirk SL, Mackie PM, Hebert PDN. 2005.** Critical factors for assembling a high volume of DNA barcodes. *Philosophical Transactions of the Royal Society London B: Biological Sciences* 360:1959–1967
DOI 10.1098/rstb.2005.1727.
- Hasnain T. 1998.** *Implementation of convention on biological diversity in Pakistan: policy brief series # 2*. Islamabad, Pakistan: Sustainable Development Policy Institute (SDPI).
- Hausmann A, Godfray HCJ, Huemer P, Mutanen M, Rougerie R, van Nieuwerkerken EJ, Ratnasingham S, Hebert PDN. 2013.** Genetic patterns in European geometrid moths revealed by the barcode index number (BIN) system. *PLOS ONE* 8:e84518
DOI 10.1371/journal.pone.0084518.
- Hebert PDN, Braukmann TWA, Prosser SWJ, Ratnasingham S, deWaard JR, Ivanova NV, Janzen DH, Hallwachs W, Naik S, Sones JE, Zakharov EV. 2018.** A Sequel to Sanger: amplicon sequencing that scales. *BMC Genomics* 19:219 DOI 10.1186/s12864-018-4611-3.
- Hebert PDN, Cywinska A, Ball SL, deWaard JR. 2003.** Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences* 270:313–321
DOI 10.1098/rspb.2002.2218.
- Hebert PDN, deWaard JR, Zakharov EV, Prosser SWJ, Sones JE, McKeown JTA, Mantle B, Salle JL. 2013.** A DNA Barcode Blitz: rapid digitization and sequencing of a natural history collection. *PLOS ONE* 8:e68535 DOI 10.1371/journal.pone.0068535.
- Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W. 2004.** Ten species in one: DNA barcoding reveals cryptic species in the Neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America* 101:14812–14817 DOI 10.1073/pnas.0406166101.
- Hebert PDN, Ratnasingham S, Zakharov EV, Telfer AC, Levesque-Beaudin V, Milton MA, Pedersen S, Jannetta P, deWaard JR. 2016.** Counting animal species with DNA barcodes: canadian insects. *Philosophical Transactions of the Royal Society B* 371:20150333
DOI 10.1098/rstb.2015.0333.
- Hsieh T, Ma K, Chao A. 2016.** iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution* 7:1451–1456
DOI 10.1111/2041-210X.12613.
- Huemer P, Mutanen M, Sefc KM, Hebert PDN. 2014.** Testing DNA barcode performance in 1000 species of European Lepidoptera: large geographic distances have small genetic impacts. *PLOS ONE* 9:e115774 DOI 10.1371/journal.pone.0115774.
- Huemer P, Wieser C, Stark W, Hebert PDN, Wiesmair B. 2019.** DNA barcode library of megadiverse Austrian Noctuoidea (Lepidoptera)-a nearly perfect match of Linnean taxonomy. *Biodiversity Data Journal* 7:e37734 DOI 10.3897/BDJ.7.e37734.
- Iftikhar R, Ashfaq M, Rasool A, Hebert PDN. 2016b.** DNA barcode analysis of thrips (Thysanoptera) diversity in Pakistan reveals cryptic species complexes. *PLOS ONE* 11:e0146014
DOI 10.1371/journal.pone.0146014.

- Iftikhar R, Ullah I, Diffie S, Ashfaq M. 2016a. Deciphering Thysanoptera: a comprehensive study on the distribution and diversity of thrips fauna in Pakistan. *Pakistan Journal of Zoology* 48(5):1233–1240.
- Inayat TP, Rana SA, Khan HA, Rehman K. 2010. Diversity of insect fauna in croplands of district Faisalabad. *Pakistan Journal of Agricultural Sciences* 47(3):245–250.
- Ivanova NV, deWaard JR, Hebert PDN. 2006. An inexpensive, automation-friendly protocol for recovering high quality DNA. *Molecular Ecology Notes* 6(4):998–1002
DOI 10.1111/j.1471-8286.2006.01428.x.
- Karlsson D, Hartop E, Forshage M, Jaschhof M, Ronquist F. 2020. The Swedish Malaise Trap Project: a 15 year retrospective on a countrywide insect inventory. *Biodiversity Data Journal* 8(3):e47255 DOI 10.3897/BDJ.8.e47255.
- Khan MS. 1976. An annotated checklist and key to the amphibians of Pakistan. *Biologia* 22(2):201–210.
- Kress WJ, García-Robledo C, Uriarte M, Erickson DL. 2015. DNA barcodes for ecology, evolution, and conservation. *Trends in Ecology and Evolution* 30(1):25–35
DOI 10.1016/j.tree.2014.10.008.
- Lin X, Stur E, Ekrem T. 2015. Exploring genetic divergence in a species-rich insect genus using 2790 DNA barcodes. *PLOS ONE* 10:e0138993 DOI 10.1371/journal.pone.0138993.
- Lopez LCS, de Fracasso MPA, Mesquita DO, Palma ART, Riul P. 2012. The relationship between percentage of singletons and sampling effort: a new approach to reduce the bias of richness estimates. *Ecological Indicators* 12:164–169 DOI 10.1016/j.ecolind.2011.07.012.
- Mallo D, Posada D. 2016. Multilocus inference of species trees and DNA barcoding. *Philosophical Transactions of the Royal Society B* 371:20150335 DOI 10.1098/rstb.2015.0335.
- Manzoor M, Khan SW, Shah SA. 2020. An annotated checklist of butterflies at elevated protected areas of Pakistan. *Journal of Bioresource Management* 7:41–52 DOI 10.35691/JBM.0202.0119.
- Marshall S, Paiero S, Buck M. 2009. Point Pelee National Park species list. Available at http://www.uoguelph.ca/debu/pelee_specieslist.htm (accessed 20 June 2020).
- Ministry of Climate Change, Pakistan. 2019. Pakistan's sixth national report to the United Nations Convention on Biological Diversity. Convention on Biological Diversity, Montreal, Canada. Available at <https://www.cbd.int/doc/nr/nr-06/pk-nr-06-en.pdf>.
- Mitchell A, Moeseneder CH, Hutchinson PM. 2020. Hiding in plain sight: DNA barcoding suggests cryptic species in all 'well-known' Australian flower beetles (Scarabaeidae: Cetoniinae). *PeerJ* 8:e9348 DOI 10.7717/peerj.9348.
- Naseem MT, Ashfaq M, Khan AM, Rasool A, Asif M, Hebert PDN. 2019. BIN overlap confirms transcontinental distribution of pest aphids (Hemiptera: Aphididae). *PLOS ONE* 14:e0220426 DOI 10.1371/journal.pone.0220426.
- Nazir N, Mehmood K, Ashfaq M, Rahim J. 2014. Morphological and molecular identification of acridid grasshoppers (Acrididae: Orthoptera) from Poonch division, Azad Jammu Kashmir, Pakistan. *Journal of Threatened Taxa* 6(3):5544–5552 DOI 10.11609/JoTT.o3507.5544-52.
- Ortiz A, Rubio R, Guerrero J, Garre M, Serrano J, Hebert P, Hausmann A. 2017. Close congruence between barcode index numbers (BINs) and species boundaries in the Erebiidae (Lepidoptera: Noctuoidea) of the Iberian Peninsula. *Biodiversity Data Journal* 5:e19840 DOI 10.3897/BDJ.5.e19840.
- Pentinsaari M, Blagojev GA, Hogg ID, Levesque-Beaudin V, Perez K, Sobel CN, Vandenbrink B, Borisenko A. 2020. DNA barcoding survey of an Arctic arthropod community: implications for future monitoring. *Insects* 11(1):46 DOI 10.3390/insects11010046.

- Pompeo PN, de Oliveira FLCI, Santos MAB, Mafra AL, Filho OK, Baretta D. 2017.** Morphological diversity of Coleoptera (Arthropoda: Insecta) in agriculture and forest systems. *Revista Brasileira de Ciência do Solo* **41**:e0160433 DOI [10.1590/18069657rbcs20160433](https://doi.org/10.1590/18069657rbcs20160433).
- Porco D, Rougerie R, Deharveng L, Hebert P. 2010.** Coupling non-destructive DNA extraction and voucher retrieval for small soft-bodied arthropods in a high-throughput context: the example of Collembola. *Molecular Ecology Resources* **10**(6):942–945 DOI [10.1111/j.1755-0998.2010.2839.x](https://doi.org/10.1111/j.1755-0998.2010.2839.x).
- Ramage T, Martins-Simoes P, Mialdea G, Allemand R, Duploux A, Rouse P, Davies N, Roderick GK, Charlat S. 2017.** A DNA barcode-based survey of terrestrial arthropods in the Society Islands of French Polynesia: host diversity within the SymbioCode Project. *European Journal of Taxonomy* **272**:1–13 DOI [10.5852/ejt.2017.272](https://doi.org/10.5852/ejt.2017.272).
- Ratnasingham S, Hebert PDN. 2007.** BOLD: the barcode of life data system (www.barcodinglife.org). *Molecular Ecology Notes* **7**:355–364 DOI [10.1111/j.1471-8286.2007.01678.x](https://doi.org/10.1111/j.1471-8286.2007.01678.x).
- Ratnasingham S, Hebert PDN. 2013.** A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLOS ONE* **8**:e66213 DOI [10.1371/journal.pone.0066213](https://doi.org/10.1371/journal.pone.0066213).
- Ren J, Ashfaq M, Hu X, Ma J, Liang F, Hebert PDN, Lin L, Germain JF, Ahmed MZ. 2018.** Barcode index numbers expedite quarantine inspections and aid the interception of nonindigenous mealybugs (Pseudococcidae). *Biological Invasions* **20**:449–460 DOI [10.1007/s10530-017-1546-6](https://doi.org/10.1007/s10530-017-1546-6).
- Ritter CD, Häggqvist S, Karlsson D, Sääksjärvi IE, Muasya AM, Nilsson RH, Antonelli A. 2019.** Biodiversity assessments in the 21st century: the potential of insect traps to complement environmental samples for estimating eukaryotic and prokaryotic diversity using high-throughput DNA metabarcoding. *Genome* **62**:147–159 DOI [10.1139/gen-2018-0096](https://doi.org/10.1139/gen-2018-0096).
- Roberts TJ. 1997.** *The mammals of Pakistan*. Revised Edition. Oxford: Oxford University Press, 525.
- Scheffers BR, Joppa LN, Pimm SL, Laurance WF. 2012.** What we know and don't know about earth's missing biodiversity. *Trends in Ecology & Evolution* **27**:501–510 DOI [10.1016/j.tree.2012.05.008](https://doi.org/10.1016/j.tree.2012.05.008).
- Smith MA, Bertrand C, Crosby K, Eveleigh ES, Fernandez-Triana J, Fisher BL, Gibbs J, Hajibabaei M, Hallwachs W, Hind K, Hrccek J, Huang D-W, Janda M, Janzen DH, Li Y, Miller SE, Packer L, Quicke D, Ratnasingham S, Rodriguez J, Rougerie R, Shaw MR, Sheffield C, Stahlhut JK, Steinke D, Whitfield J, Wood M, Zhou X. 2012.** Wolbachia and DNA barcoding insects: patterns, potential, and problems. *PLOS ONE* **7**:e36514 DOI [10.1371/journal.pone.0036514](https://doi.org/10.1371/journal.pone.0036514).
- Song H, Buhay JE, Whiting MF, Crandall KA. 2008.** Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proceedings of the National Academy of Science of the United States of America* **105**:13486–13491 DOI [10.1073/pnas.0803076105](https://doi.org/10.1073/pnas.0803076105).
- Stork NE. 2018.** How many species of insects and other terrestrial arthropods are there on Earth? *Annual Review of Entomology* **63**:31–45 DOI [10.1146/annurev-ento-020117-043348](https://doi.org/10.1146/annurev-ento-020117-043348).
- van Velzen R, Weitschek E, Felici G, Bakker FT. 2012.** DNA barcoding of recently diverged species: relative performance of matching methods. *PLOS ONE* **7**:e30490 DOI [10.1371/journal.pone.0030490](https://doi.org/10.1371/journal.pone.0030490).
- Velasco-Cuervo SM, Aguirre-Ramirez E, Gallo-Franco JJ, Obando RG, Carrejo N, Toro-Perea N. 2019.** Saving DNA from museum specimens: the success of DNA mini-barcodes in haplotype reconstruction in the genus *Anastrepha* (Diptera: Tephritidae). *Journal of Advanced Research* **16**:123–134 DOI [10.1016/j.jare.2018.11.001](https://doi.org/10.1016/j.jare.2018.11.001).

- Virgilio M, Backeljau T, Nevado B, De Meyer M. 2010.** Comparative performances of DNA barcoding across insect orders. *BMC Bioinformatics* **11**:206 DOI [10.1186/1471-2105-11-206](https://doi.org/10.1186/1471-2105-11-206).
- Ward DF, Lariviere M-C. 2004.** Terrestrial invertebrate surveys and rapid biodiversity assessment in New Zealand: lessons from Australia. *New Zealand Journal of Ecology* **28**(1):151–159.
- Weigand H, Beermann AJ, Čiampor F, Costa FO, Csabai Z, Duarte S, Geiger MF, Grabowski M, Rimet F, Rulik B, Strand M, Szucsich N, Weigand AM, Willassen E, Wyler SA, Bouchez A, Borja A, Čiamporová-Zaťovičová Z, Ekrem T. 2019.** DNA barcode reference libraries for the monitoring of aquatic biota in Europe: gap-analysis and recommendations for future work. *Science of the Total Environment* **678**:499–524 DOI [10.1016/j.scitotenv.2019.04.247](https://doi.org/10.1016/j.scitotenv.2019.04.247).
- Werneck FP, Nogueira C, Colli GR, Sites JW, Costa GC. 2012.** Climatic stability in the Brazilian Cerrado: implications for biogeographical connections of South American savannas, species richness and conservation in a biodiversity hotspot. *Journal of Biogeography* **39**:1695–1706 DOI [10.1111/j.1365-2699.2012.02715.x](https://doi.org/10.1111/j.1365-2699.2012.02715.x).
- Wilson JJ, Sing KW, Floyd RM, Hebert PDN. 2017.** DNA barcodes and insect biodiversity. In: Foottit RG, Adler PH, eds. *Insect Biodiversity: Science and Society*. Vol. 1. Second Edition. Oxford: Blackwell Publishing Ltd.
- Wirta H, Várkonyi G, Rasmussen C, Kaartinen R, Schmidt NM, Hebert PDN, Barták M, Blagoev G, Disney H, Ertl S, Gjelstrup P, Gwiazdowicz DJ, Huldén L, Ilmonen J, Jakovlev J, Jaschhof M, Kahanpää J, Kankaanpää T, Krogh PH, Labbee R, Lettner C, Michelsen V, Nielsen SA, Nielsen TR, Paasivirta L, Pedersen S, Pohjoismäki J, Salmela J, Villkamaa P, Väre H, von Tschirnhaus M, Roslin T. 2016.** Establishing a community-wide DNA barcode library as a new tool for arctic research. *Molecular Ecology Resources* **16**(3):809–822 DOI [10.1111/1755-0998.12489](https://doi.org/10.1111/1755-0998.12489).
- Zahiri R, Lafontaine JD, Schmidt BC, deWaard JR, Zakharov EV, Hebert PDN. 2017.** Probing planetary biodiversity with DNA barcodes: the Noctuoidea of North America. *PLOS ONE* **12**:e0178548 DOI [10.1371/journal.pone.0178548](https://doi.org/10.1371/journal.pone.0178548).
- Zhou Z, Guo H, Han L, Chai J, Che X, Shi F. 2019.** Singleton molecular species delimitation based on COI-5P barcode sequences revealed high cryptic/undescribed diversity for Chinese katydids (Orthoptera: Tettigoniidae). *BMC Evolutionary Biology* **19**(1):79 DOI [10.1186/s12862-019-1404-5](https://doi.org/10.1186/s12862-019-1404-5).