

## PREPRINT

*Author-formatted, not peer-reviewed document posted on 01/08/2023*

DOI: <https://doi.org/10.3897/arphapreprints.e110112>

# **The MOBIOSt: A FAIR (Findable, Accessible, Interoperable, and Reusable) Database for Mindanao's Terrestrial Biodiversity**

 Krizler Tanalgo, Kier Dela Cruz,  Angelo Agduma, Jeaneth Magelen Respicio, Sumaira Abdullah, Renee Jane Ele, Bona Abigail Hilario-Husain, Meriam Rubio, Sedra Murray, Lothy Casim, Athea Mohidda Pantog, Shiela Mae Balase, Rallyessa Mohann Abdulkasan, Chasty Andrea Aguirre, Nadjmussahar Banto, Sheila Mae Broncate, Ace Dimacaling, Gerald Vince Fabrero, Asraf Lidasan, Analiza Lingcob, Ariane Millondaga, Kathlene Faye Panilla, Crystal Queen Sinadjan, Norlaine Unte

# The MOBIOS<sup>+</sup>: A FAIR (Findable, Accessible, Interoperable, and Reusable) Database for Mindanao's Terrestrial Biodiversity

Krizler Cejuela Tanalgo<sup>‡</sup>, Kier Celestial Dela Cruz<sup>‡</sup>, Angelo Rellama Agduma<sup>‡</sup>, Jeaneth Magelen V. Respicio<sup>‡</sup>, Sumaira S. Abdullah<sup>‡</sup>, Renee Jane A. Ele<sup>‡</sup>, Bona Abigail Hilario-Husain<sup>‡</sup>, Meriam M. Rubio<sup>‡</sup>, Sedra A. Murray<sup>‡</sup>, Lothy F. Casim<sup>§</sup>, Athea Mohidda M. Pantog<sup>‡</sup>, Shiela Mae P. Balase<sup>‡</sup>, Rallyessa Mohann A. Abdulkasan<sup>‡</sup>, Chasty Andrea S. Aguirre<sup>‡</sup>, Nadjmussahar Banto<sup>‡</sup>, Sheila Mae M. Broncate<sup>‡</sup>, Ace D. Dimacaling<sup>‡</sup>, Gerald Vince N. Fabrerro<sup>‡</sup>, Asraf K. Lidasan<sup>‡</sup>, Analiza A. Lingcob<sup>‡</sup>, Ariane M. Millondaga<sup>‡</sup>, Kathlene Faye L. Panilla<sup>‡</sup>, Crystal Queen M. Sinadjan<sup>‡</sup>, Norlaine D. Unte<sup>‡</sup>

<sup>‡</sup> Ecology and Conservation Research Laboratory (Eco/Con Lab), Department of Biological Sciences, College of Science and Mathematics, University of Southern Mindanao, Kabacan, Philippines

<sup>§</sup> Molecular Parasitology Research Laboratory, Department of Biological Sciences, College of Science and Mathematics, University of Southern Mindanao, Kabacan 9407, Philippines

Corresponding author: Krizler Cejuela Tanalgo ([tkrizler@gmail.com](mailto:tkrizler@gmail.com))

## Abstract

## Background

Due to its complex geological history and the emergence of various biogeographic regions, the Philippines boasts an extraordinary array of flora and fauna. This unique combination has contributed to the country's exceptional density of terrestrial species, positioning it among the highest in the world. Mindanao, in the southern part of the Philippines, is the second largest group of islands and supports high biodiversity of unique and threatened species of flora and fauna. However, consolidated biodiversity records and information remain unavailable for the region. Filling these significant gaps would advance understanding biodiversity patterns in Mindanao across time and space. The Mindanao Open Biodiversity Information (MOBIOS<sup>+</sup>) database aims to consolidate 21st-century biodiversity data and establish an open-access database. The database we present here is the first of its kind and currently the most comprehensive attempt to establish the largest consolidated database for Mindanao biodiversity. With its vast collection of biodiversity data, this database will prove to be a valuable resource for advancing biodiversity research and analysis. It will facilitate the identification of species and areas that require immediate conservation prioritization and action, addressing the urgent challenges posed by our rapidly changing planet.

## New information

The MOBIO+ database is the first attempt to create a massive FAIR database aiming to collate biodiversity records in the Mindanao faunal region, South of the Philippines. The database currently includes 12,814 georeferenced species occurrences from 1,925 unique species or taxa from 10 animal classes from the Mindanao faunal region's terrestrial and freshwater bodies. We made all georeferenced species occurrences available in the Global Biodiversity Information Facility (GBIF) platform.

## Keywords

Biodiversity records, conservation, FAIR database, islands, Philippines

## Introduction

Current global biodiversity is at risk and faces an extinction rate higher than previous extinction events (Cowie et al. 2017). To tackle the pressing issue of biodiversity crisis, conservation biologists and experts must establish priorities to identify species and areas that require immediate attention and conservation efforts (Brum et al. 2017). Setting fine-scale prioritization analysis is essential to ensure it complements the local biodiversity context and reflects global priorities (Tanalgo and Hughes 2019, Knight et al. 2008). Effective biodiversity synthesis and prioritization on any scale require good data integration to accurately identify priorities based on data and evidence (Heberling et al. 2021, Dela Cruz et al. 2023).

The Philippine archipelago's more than 7,100 islands, has one of the highest levels of endemism of any place and is a hotspot for biodiversity conservation, directly related to the tropical environment of the country and the complicated geological past (Lohman et al. 2010, Rickart et al. 2003). The three largest faunal regions in the Philippines are the Luzon faunal region, the Mindanao faunal region, and the Palawan faunal region (Heaney 1986). This study mainly focused on the Mindanao faunal region, including islands from the Visayas, specifically the Eastern Visayas, which are Samar, Leyte, Bohol, and smaller islands such as Basilan, Dinagat, and Surigao. Although Mindanao is the second largest group of islands in the Philippines next to Luzon, biodiversity studies in the region have been limited in recent decades and are often challenged by low research funding allocation, a low number of experts, and the peace and order situation that hinder researchers in exploring many potential areas for biodiversity studies (Akduma et al. 2023, Alcalá 2004, Amoroso 2000). With the advent of new technological advances, increasing collaborations, and the emergence of diverse and young biodiversity researchers, the number of biodiversity research has increased in the early 21<sup>st</sup> century (Akduma et al. 2023, Abdullah et al. 2023, Dela Cruz et al. 2023). Integrating this unconsolidated biodiversity information is another new challenge that has a crucial impact on

understanding the large-scale biodiversity pattern in Mindanao. These gaps challenge the realization of practical, data-driven, and up-to-date conservation priorities, especially for the region's endemic and threatened species.

In this data paper, we introduce the Mindanao Open Biodiversity Information (MOBIOS<sup>+</sup>) database, which aims to create a FAIR (Findable, Accessible, Interoperable, and Reusable) platform by integrating publicly accessible biodiversity data from published studies within the Mindanao faunal region. (Fig. 1). By leveraging the integrated data within MOBIOS<sup>+</sup>, researchers can gain a comprehensive understanding of biodiversity distribution, patterns, and threats within Mindanao and the wider Philippines across different scales. The valuable information provided by the database will prove instrumental in the development of biodiversity synthesis, modeling, mapping, and prediction of environmental changes specifically in the Mindanao region.

## General description

**Purpose:** The MOBIOS<sup>+</sup> project has an overarching goal of establishing a biodiversity database for Mindanao following the FAIR (Findable, Accessible, Interoperable, and Reusable) principles to advance studies in biodiversity and develop a synthesis to identify current and future conservation priorities in the region. Furthermore, the MOBIOS<sup>+</sup> database aims to mainstream rich and important biodiversity in Mindanao. This is through making occurrence records of species from studies from various institutions in Mindanao centralised in an accessible platform (Fig. 1). To do this, we collated biodiversity records from published studies and organized them in a standard format, machine-readable, and digitally available through the Global Biodiversity Information Facility (GBIF) platform. This project aims to continuously update the species database, complementing on-ground biodiversity efforts in Mindanao. The database is housed at the Biodiversity Synthesis Centre of the Eco/Con Lab at the Department of Biological Sciences, University of Southern Mindanao.

## Sampling methods

**Description:** The MOBIOS<sup>+</sup> database currently contains information for 12,814 georeferenced species occurrences from 1,925 species from the Mindanao Faunal region (Mainland Mindanao and adjacent provinces in Visayas). The database represents at least ten taxonomic classes of terrestrial and freshwater fauna. This is the first database version that contains biodiversity records based on literature from the early 21<sup>st</sup> century for terrestrial fauna from the faunal region of Mindanao.

**Sampling description:** We applied the PRISMA (Preferred reporting items for systematic reviews and meta-analyses) approach (O'Dea et al. 2021) to collect and screen biodiversity studies from Mindanao from 2000-2023. We used the combinations of the following keywords: 'Biodiversity stud\*', 'Assessment\*', 'Survey', 'Terrestrial', 'Freshwater', and 'Mindanao' to search for published literature from Google Scholar. We also explore the

Biodiversity Literature Repository (BLR) (<https://biolitrepo.org/>) and the self-archiving ResearchGate (<https://www.researchgate.net/>) for additional published studies, particularly on new localities and natural history notes of species. We excluded the thesis and dissertation to standardise our database because most universities and academic institutions do not have online repositories to access their research. We then assigned teams to collect and analyze data from specific taxonomic groups. We sampled each literature for species occurrence from published articles and books using the Darwin Core Biodiversity Standard format (Wieczorek et al. 2012). We extracted information for the species list, taxonomic classification, conservation status, location of the species, and other relevant information.

**Quality control:** We counter-curated all listed species and their distribution in the database using the Integrated Taxonomic Information System (ITIS) database (<https://www.itis.gov/>). We excluded dubious species, such as those with problematic identification or species that naturally do not occur in the range without proper discussions or expert clarifications. According to the Catalogue of Life (CoL), taxonomy names were aligned and standardised (<https://www.catalogueoflife.org>). We retained the taxonomic classification of species with confusing arrangements. We plotted and mapped all species occurrence within the boundaries of the Mindanao faunal region using QGIS (v. 3.30) (QGIS team 2023) to curate species occurrence within the geographical range. Biodiversity records outside the range or within unusual locations were counter-checked with the original reference and corrected. We will update the MOBIO+ database in annual basis following the same screening procedure and standards. A release note will be published by the corresponding author and institution when new version of the database will be released.

**Step description:** The data included in the MOBIO+ database were extracted from published research articles publicly available online and organized using the GBIF standards for data publication. The following steps were taken to audit the accuracy of the data set in the MOBIO+ database (1) Collating and filtering published biodiversity studies from the Mindanao faunal region; (2) Reviewing the studies for suitability based on criteria; (3) Extracting species occurrence data and other relevant metadata from biodiversity studies in Mindanao faunal region; (4) Placing the species distribution and other metadata in Microsoft Office Excel format; (5) Curating species occurrence in Quantum GIS; (6) Organising of occurrence dataset following Darwin Core Standards.

## Geographic coverage

**Description:** The database contains data for terrestrial and freshwater ecosystems in the entire faunal region of Mindanao, south of the Philippines.

**Coordinates:** 4.7438 and 12.4646 Latitude; 126.6303 and 118.4747 Longitude.

## Taxonomic coverage

**Description:** The 12,814 georeferenced species occurrence is comprised of 1,925 species from 10 taxonomic classes (Fig. 2.) Invertebrates are represented by Gastropoda (27 spp., 1%), Bivalvia (4 spp., 0.20%), Arachnida (149 spp., 8%), Malacostraca (43 spp., 2%), and Insecta (866 spp., 45%) Fig. 3. Whilst, the vertebrates are represented by Actinopterygii (176 spp., 9%), Amphibia (65 spp., 3%), Reptilia (135 spp., 7%), Aves (382 spp., 20%) and Mammalia (78 spp., 4%).

### Taxa included:

Rank	Scientific Name	Common Name
kingdom	Animalia	Animals
class	Gastropoda	Gastropods
class	Bivalvia	Bivalves
class	Arachnida	Arachnids
class	Malacostraca	Crustaceans
class	Insecta	Insects
class	Actinopterygii	Fishes
class	Amphibia	Amphibians
class	Reptilia	Reptiles
class	Aves	Birds
class	Mammalia	Mammals

## Temporal coverage

**Data range:** 2000-1-01 - 2022-12-31.

## Usage licence

**Usage licence:** Open Data Commons Attribution License

## Data resources

**Data package title:** MOBIOSt Dataset version 1

**Resource link:** [https://ipt.pensoft.net/resource?r=mobios\\_data](https://ipt.pensoft.net/resource?r=mobios_data)

**Number of data sets: 1****Data set name:** The MOBIO+ database: a FAIR Mindanao Biodiversity**Data format:** Darwin Core Archive format**Description:** Our dataset contains 12,814 occurrence data for 1,925 taxonomic species from 10 classes recorded from Mindanao fauna region, Philippines.

Column label	Column description
occurrenceID	An identifier for the Occurrence (as opposed to a particular digital record of the occurrence). In the absence of a persistent global unique identifier, construct one from a combination of identifiers in the record that will most closely make the occurrenceID globally unique.
basisOfRecord	The specific nature of the data record.
occurrenceStatus	A statement about the presence or absence of a Taxon at a Location.
year	The four-digit year in which the Event occurred, according to the Common Era Calendar.
continent	The name of the continent in which the Location occurs.
countryCode	The standard code for the country in which the Location occurs.
stateProvince	The name of the next smaller administrative region than country (state, province, canton, department, region, etc.) in which the Location occurs.
county	The full, unabbreviated name of the next smaller administrative region than stateProvince (county, shire, department, etc.) in which the Location occurs.
municipality	The full, unabbreviated name of the next smaller administrative region than county (city, municipality, etc.) in which the Location occurs.
locality	The specific description of the place where the species was recorded.
decimalLatitude	The geographic latitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location.
decimalLongitude	The geographic longitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location.
coordinateUncertaintyInMeters	The horizontal distance (in meters) from the given decimalLatitude and decimalLongitude describing the smallest circle containing the whole of the Location.
geodeticDatum	The ellipsoid, geodetic datum, or spatial reference system (SRS) upon which the geographic coordinates given in decimalLatitude and decimalLongitude as based.
georeferencedBy	A list (concatenated and separated) of names of people, groups, or organizations who determined the georeference (spatial representation) for the Location.
scientificName	The full scientific name in lowest level taxonomic rank that can be determined.

kingdom	The full scientific name of the kingdom in which the taxon is classified.
phylum	The full scientific name of the phylum or division in which the taxon is classified.
class	The full scientific name of the class in which the taxon is classified.
family	The full scientific name of the family in which the taxon is classified.
genus	The full scientific name of the genus in which the taxon is classified.
specificEpithet	The name of the first or species epithet of the scientificName.
taxonRank	The taxonomic rank of the most specific name in the scientificName.
scientificNameAuthorship	The authorship information for the scientificName formatted according to the conventions of the applicable nomenclaturalCode.
taxonomicStatus	The status of the use of the dwc:scientificName as a label for a taxon.
associatedReferences	A list (concatenated and separated) of identifiers (publication, bibliographic reference, global unique identifier, URI) of literature associated with the Occurrence.

## Additional information

The MOBIOSt primary objective is to make biodiversity information from the literature more accessible, and ready and ensure the reusability of the data within our database for various biodiversity research purposes. To achieve this, we linked the MOBIOSt database deposited in GBIF IPT to the Catalogue of Life (CoL) (<https://www.catalogueoflife.org/>), and the original sources of the species records using the Digital Object Identifier (DOI) associated with each reference (see Fig. 1) (Table 1). The infrastructure utilizes the latest version of the unique identifier the updated Catalogue of Life (CoL) provides to establish links between species (Suppl. material 1).

Table 1

## References

- Abdullah S, Dela Cruz KC, Casim L, Agduma AR, Tanalgo KC (2023) Leaping Forward or Crawling Backward? Efforts and Biases in Amphibian and Reptile Research on a Megadiverse Faunal Region in the Philippines. *Conservation* 3 (3): 363-378. <https://doi.org/10.3390/conservation3030025>
- Agduma A, Garcia F, Pimentel J, Cabasan M, Ele RJ, Rubio M, Murray S, Husain-Hilario BA, Dela Cruz KC, Abdullah S, Balase SM, Tanalgo KC (2023) Overview of priorities, threats, and challenges to biodiversity conservation in the Southern Philippines. *Regional Sustainability* 4 (2): 203-213. <https://doi.org/10.1016/j.regsus.2023.05.003>
- Alcala A (2004) Biodiversity Research in the Philippines from 1998-2003. ASEAN Biodiversity

- Amoroso VB (2000) Long-term biodiversity research programme for Mindanao, Philippines. *The Korean Journal of Ecology* 23: 141-147.
- Brum FT, Graham CH, Costa GC, Hedges SB, Penone C, Radeloff VC, Rondinini C, Loyola R, Davidson AD (2017) Global priorities for conservation across multiple dimensions of mammalian diversity. *Proceedings of the National Academy of Sciences* 114: 7641-7646. <https://doi.org/10.1073/pnas.1706461114>
- Cowie R, Régnier C, Fontaine B, Bouchet P (2017) Measuring the Sixth Extinction: what do mollusks tell us? *The Nautilus* 131: 3-41.
- Dela Cruz KC, Abdullah S, Agduma AR, Tanalgo KC (2023) Early twenty-first century biodiversity data pinpoint key targets for bird and mammal conservation in Mindanao, Southern Philippines. *Biodiversity* 1-18. <https://doi.org/10.1080/14888386.2023.2210119>
- Heaney LR (1986) Biogeography of mammals in SE Asia: estimates of rates of colonization, extinction and speciation. *Biological Journal of the Linnean Society* 28: 127-165. <https://doi.org/10.1111/j.1095-8312.1986.tb01752.x>
- Heberling J, Miller J, Noesgaard D, Weingart S, Schigel D (2021) Data integration enables global biodiversity synthesis. *Proceedings of the National Academy of Sciences* 118 (6). <https://doi.org/10.1073/pnas.2018093118>
- Knight AT, Cowling RM, Rouget M, Balmford A, Lombard AT, Campbell BM (2008) Knowing but not doing: Selecting priority conservation areas and the research-implementation gap. *Conservation Biology* 22: 610-617. <https://doi.org/10.1111/j.1523-1739.2008.00914.x>
- Lohman DJ, Ingram KK, Prawiradilaga DM, Winker K, Sheldon FH, Moyle RG, Ng PK, Ong PS, Wang LK, Braile TM, Astuti D, Meier R (2010) Cryptic genetic diversity in “widespread” Southeast Asian bird species suggests that Philippine avian endemism is gravely underestimated. *Biological Conservation* 143: 1885-1890. <https://doi.org/10.1016/j.biocon.2010.04.042>
- O’Dea R, Lagisz M, Jennions M, Koricheva J, Noble DA, Parker T, Gurevitch J, Page M, Stewart G, Moher D, Nakagawa S (2021) Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension. *Biological Reviews* 96 (5): 1695-1722. <https://doi.org/10.1111/brv.12721>
- QGIS team (2023) QGIS Geographic Information System. Open Source Geospatial Foundation Project. Release date: 2023-5-11. URL: <https://qgis.org/en/site/>
- Rickart EA, Heaney LR, Tabaranza BR (2003) A new species of *Limnomys* (Rodentia: Muridae: Murinae) from Mindanao Island, Philippines. *Journal of Mammalogy* 84: 1443-1455. <https://doi.org/10.1644/BRP-112>
- Tanalgo KC, Hughes AC (2019) Priority-setting for Philippine bats using practical approach to guide effective species conservation and policy-making in the Anthropocene. *Hystrix* 30 (1): 74-83. <https://doi.org/10.4404/hystrix-00172-2019>
- Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D (2012) Darwin Core: An evolving community-developed biodiversity data standard. *PLOS One* 7 (1): e29715. <https://doi.org/10.1371/journal.pone.0029715>

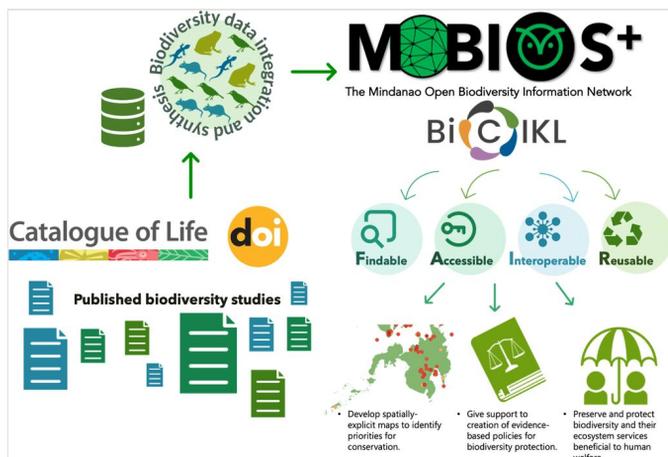


Figure 1.

A proposed framework for academic and research institutions to address and advance biodiversity knowledge gaps and strengthen the availability and access to biodiversity information that will facilitate effective conservation prioritization through FAIR data sharing.

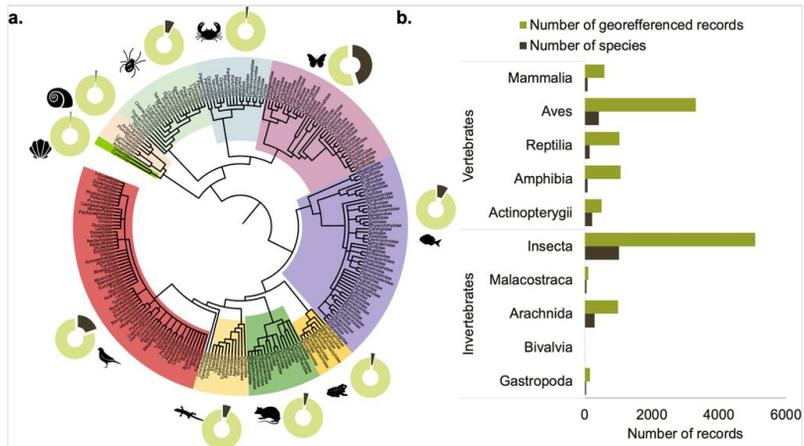


Figure 2.

Diversity and distribution of species occurrence records across taxonomic groups included in the first version of the MOBIOS+ database. The diversity of species (percentage,%) according to class compared to the overall number of species recorded in the MOBIOS+ database (a); and the total number of species and the number of georeferenced occurrences per animal class (b).

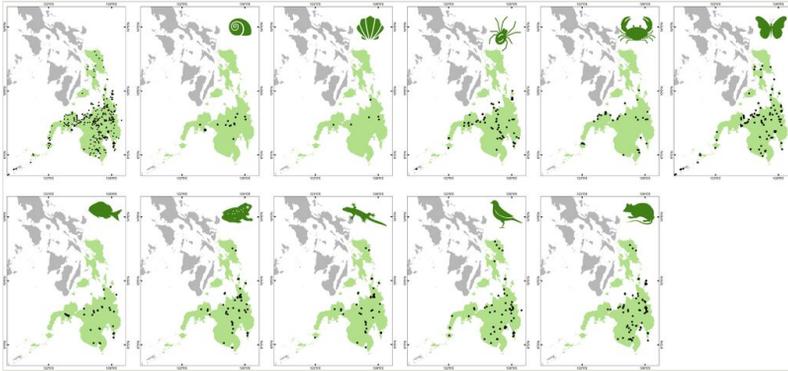


Figure 3.

Visualised georeferenced species occurrence in the Mindanao faunal region currently included in the MOBIO+ database (version 1).

Table 1.

Additional information for linked data to unique identifiers

Column	Definition
CoLIdentifiers	Catalogue of Life (CoL) identifiers; the Stable name identifiers for the Catalogue of Life (CoL).
CoLLinked	Linked to the CoL record.
DigitalObjectIdentifier	The permanent object identifier of the associatedReferences.

## Supplementary material

### **Suppl. material 1: Supplementary file for occurrence and species records from MOBIOS+ database**

**Authors:** Krizler C. Tanalgo and MOBIOS consortium

**Data type:** occurrences

**Brief description:** This supplementary file contains linked data from Catalogue of Life (CoL) and the DOI of the associated Reference

[Download file](#) (6.69 MB)