Versioned Archive and Review of Biotic Interactions and Taxon Names Found within globalbioticinteractions/pollinators-of-apocynaceae-database hash://md5/ba7bfe57d28af24d7dad244f9882dbc5

by Nomer, Elton and Preston, three naive review bots
review@globalbioticinteractions.org
https://globalbioticinteractions.org/contribute
https://github.com/globalbioticinteractions/pollinators-of-apocynaceae-database/issues

2025-07-08

Abstract

Life on Earth is sustained by complex interactions between organisms and their environment. These biotic interactions can be captured in datasets and published digitally. We present a review and archiving process for such an openly accessible digital interactions dataset of known origin and discuss its outcome. The dataset under review, named globalbioticinteractions/pollinators-of-apocynaceae-database, has fingerprint hash://md5/ba7bfe57d28af24d7dad244f9882dbc5, is 1.20MiB in size and contains 5,060 interaction with 3 unique types of associations (e.g., pollinatedBy) between 679 primary taxa (e.g., Asclepias syriaca) and 2,330 associated taxon (e.g., unidentified). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.

Contents

Introduction **2**
Data Review and Archive 2

Methods **3**

Results **4**
Files 4
Archived Dataset 14

Biotic Interactions 14
Interaction Networks 17
Taxonomic Alignment 19
Additional Reviews 23
GloBI Review Badge 23
GloBI Index Badge 24

Discussion **24**

Acknowledgements **25**

Author contributions **25**

References **25**

Introduction

Data Review and Archive

Data review and archiving can be a time-consuming process, especially when done manually. This review report aims to help facilitate both activities. It automates the archiving of datasets, including Darwin Core archives, and is a citable backup of a version of the dataset. Additionally, an automatic review of species interaction claims made in the dataset is generated and registered with Global Biotic Interactions (J. H. Poelen, Simons, and Mungall 2014).

This review includes summary statistics about, and observations about, the dataset under review:

> Ollerton J, Liede-Schumann S, Endress ME, Meve U, Rech AR, Shuttleworth A, Keller HA, Fishbein M, Alvarado-Cárdenas LO, Amorim FW, Bernhardt P, Celep F, Chirango Y, Chiriboga-Arroyo F, Civeyrel L, Cocucci A, Cranmer L, da Silva-Batista IC, de Jager L, Deprá MS, Domingos-Melo A, Dvorsky C, Agostini K, Freitas L, Gaglianone MC, Galetto L, Gilbert M, González-Ramírez I, Gorostiague P, Goyder D, Hachuy-Filho L, Heiduk A, Howard A, Ionta G, Islas-Hernández SC, Johnson SD, Joubert L, Kaiser-Bunbury CN, Kephart S, Kidyoo A, Koptur S, Koschnitzke C, Lamborn E, Livshultz T, Machado IC, Marino S, Mema L, Mochizuki K, Morellato LPC, Mrisha CK, Muiruri EW, Nakahama N, Nascimento VT, Nuttman C, Oliveira PE, Peter CI, Punekar S, Rafferty N, Rapini A, Ren ZX, Rodríguez-Flores CI, Rosero L, Sakai S, Sazima M, Steenhuisen SL, Tan CW, Torres C, Trøjelsgaard K, Ushimaru A, Vieira MF, Wiemer AP, Yamashiro T, Nadia T, Queiroz J, Quirino Z. The diversity and evolution of pollination systems in large plant clades: Apocynaceae as a case study. Ann Bot. 2019 Jan 23;123(2):311-325. doi: 10.1093/aob/mcy127. PMID: 30099492; PMCID: PMC6344220.

https://github.com/globalbioticinteractions/pollinators-of-apocynaceae-database/archive/c940b96cd5ecd4d26005eb6b65a1ea6b23f898db.zip 2025-07-05T03:48:51.573Z hash://md5/ba7bfe57d28af24d7dad244f9882dbc5

For additional metadata related to this dataset, please visit https://github.com/globalbioticinteractions/pollinators-of-apocynaceae-database and inspect associated metadata files including, but not limited to, *README.md*, *eml.xml*, and/or *globi.json*.

Methods

The review is performed through programmatic scripts that leverage tools like Preston (Elliott et al. 2025), Elton (Kuhn, Poelen, and Leinweber 2025), Nomer (Salim and Poelen 2025), globinizer (J. Poelen, Seltmann, and Mietchen 2024) combined with third-party tools like grep, mlr, tail and head.

Table 1: Tools used in this review process

tool name	version
preston	0.10.1
elton	0.15.13
nomer	0.5.17
globinizer	0.4.0
mlr	6.0.0
jq	1.6
yq	4.25.3
pandoc	3.1.6.1
duckdb	1.3.1

The review process can be described in the form of the script below [1].

```
# get versioned copy of the dataset (size approx. 1.20MiB) under review
elton pull globalbioticinteractions/pollinators-of-apocynaceae-database

# generate review notes
elton review globalbioticinteractions/pollinators-of-apocynaceae-database\
 > review.tsv

# export indexed interaction records
```

[1] Note that you have to first get the data (e.g., via elton pull globalbioticinteractions/pollinators-of-apocynaceae-database) before being able to generate reviews (e.g., elton review globalbioticinteractions/pollinators-of-apocynaceae-database), extract interaction claims (e.g., elton interactions globalbioticinteractions/pollinators-of-apocynaceae-database), or list taxonomic names (e.g., elton names globalbioticinteractions/pollinators-of-apocynaceae-database)

```
elton interactions globalbioticinteractions/pollinators-of-apocynaceae-database\
 > interactions.tsv

# export names and align them with the Catalogue of Life using Nomer
elton names globalbioticinteractions/pollinators-of-apocynaceae-database\
 | nomer append col\
 > name-alignment.tsv
```

or visually, in a process diagram.

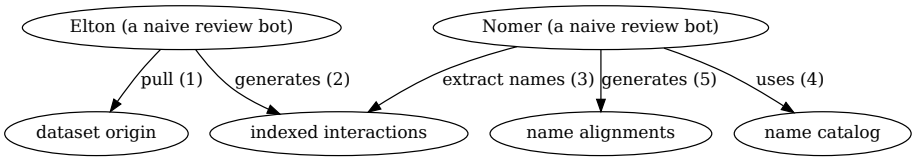

Figure 1: Review Process Overview

You can find a copy of the full review script at check-data.sh. See also GitHub and Codeberg.

Results

In the following sections, the results of the review are summarized [2]. Then, links to the detailed review reports are provided.

Files

The following files are produced in this review:

filename	description
biblio.bib	list of bibliographic reference of this review
check-dataset.sh	data review workflow/process as expressed in a bash script
data.zip	a versioned archive of the data under review
HEAD	the digital signature of the data under review
index.docx	review in MS Word format
index.html	review in HTML format
index.md	review in Pandoc markdown format
index.pdf	review in PDF format

[2]Disclaimer: The results in this review should be considered friendly, yet naive, notes from an unsophisticated robot. Please keep that in mind when considering the review results.

filename	description
indexed-citations.csv.gz	list of distinct reference citations for reviewed species interaction claims in gzipped comma-separated values file format
indexed-citations.html.gz	list of distinct reference citations for reviewed species interactions claims in gzipped html file format
indexed-citations.tsv.gz	list of distinct reference citations for reviewed species interaction claims in gzipped tab-separated values format
indexed-interactions-col-family-col-family.svg	network diagram showing the taxon family to taxon family interaction claims in the dataset under review as interpreted by the Catalogue of Life via Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024)
indexed-interactions-col-kingdom-col-kingdom.svg	network diagram showing the taxon kingdom to taxon kingom interaction claims in the dataset under review as interpreted by the Catalogue of Life via Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024)
indexed-interactions.csv.gz	species interaction claims indexed from the dataset under review in gzipped comma-separated values format
indexed-interactions.html.gz	species interaction claims indexed from the dataset under review in gzipped html format
indexed-interactions.tsv.gz	species interaction claims indexed from the dataset under review in gzipped tab-separated values format
indexed-interactions.parquet	species interaction claims indexed from the dataset under review in Apache Parquet format
indexed-interactions-sample.csv	list of species interaction claims indexed from the dataset under review in gzipped comma-separated values format
indexed-interactions-sample.html	first 500 species interaction claims indexed from the dataset under review in html format

filename	description
indexed-interactions-sample.tsv	first 500 species interaction claims indexed from the dataset under review in tab-separated values format
indexed-names.csv.gz	taxonomic names indexed from the dataset under review in gzipped comma-separated values format
indexed-names.html.gz	taxonomic names found in the dataset under review in gzipped html format
indexed-names.tsv.gz	taxonomic names found in the dataset under review in gzipped tab-separated values format
indexed-names.parquet	taxonomic names found in the dataset under review in Apache Parquet format
indexed-names-resolved-col.csv.gz	taxonomic names found in the dataset under review aligned with the Catalogue of Life as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-col.html.gz	taxonomic names found in the dataset under review aligned with the Catalogue of Life as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-col.tsv.gz	taxonomic names found in the dataset under review aligned with the Catalogue of Life as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-col.parquet	taxonomic names found in the dataset under review aligned with the Catalogue of Life as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format

filename	description
indexed-names-resolved-discoverlife.csv.gz	taxonomic names found in the dataset under review aligned with Discover Life bee species checklist as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-discoverlife.html.gz	taxonomic names found in the dataset under review aligned with Discover Life bee species checklist as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-discoverlife.tsv.gz	taxonomic names found in the dataset under review aligned with Discover Life bee species checklist as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-discoverlife.parquet	taxonomic names found in the dataset under review aligned with Discover Life bee species checklist as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-gbif.csv.gz	taxonomic names found in the dataset under review aligned with GBIF Backbone Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-gbif.html.gz	taxonomic names found in the dataset under review aligned with GBIF Backbone Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format

filename	description
indexed-names-resolved-gbif.tsv.gz	taxonomic names found in the dataset under review aligned with GBIF Backbone Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-gbif.parquet	taxonomic names found in the dataset under review aligned with GBIF Backbone Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-itis.csv.gz	taxonomic names found in the dataset under review aligned with Integrated Taxonomic Information System (ITIS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-itis.html.gz	taxonomic names found in the dataset under review aligned with Integrated Taxonomic Information System (ITIS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-itis.tsv.gz	taxonomic names found in the dataset under review aligned with Integrated Taxonomic Information System (ITIS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-itis.parquet	taxonomic names found in the dataset under review aligned with Integrated Taxonomic Information System (ITIS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format

filename	description
indexed-names-resolved-mdd.csv.gz	taxonomic names found in the dataset under review aligned with the Mammal Diversity Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-mdd.html.gz	taxonomic names found in the dataset under review aligned with Mammal Diversity Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-mdd.tsv.gz	taxonomic names found in the dataset under review aligned with Mammal Diversity Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-mdd.parquet	taxonomic names found in the dataset under review aligned with Mammal Diversity Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-ncbi.csv.gz	taxonomic names found in the dataset under review aligned with the NCBI Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-ncbi.html.gz	taxonomic names found in the dataset under review aligned with the NCBI Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format

filename	description
indexed-names-resolved-ncbi.tsv.gz	taxonomic names found in the dataset under review aligned with the NCBI Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-ncbi.parquet	taxonomic names found in the dataset under review aligned with the NCBI Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-pbdb.csv.gz	taxonomic names found in the dataset under review aligned with the Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-pbdb.html.gz	taxonomic names found in the dataset under review aligned with Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-pbdb.tsv.gz	taxonomic names found in the dataset under review aligned with Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-pbdb.parquet	taxonomic names found in the dataset under review aligned with Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format

filename	description
indexed-names-resolved-tpt.csv.gz	taxonomic names found in the dataset under review aligned with the Terrestrial Parasite Tracker (TPT) Taxonomic Resource as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-tpt.html.gz	taxonomic names found in the dataset under review aligned with the Terrestrial Parasite Tracker (TPT) Taxonomic Resource as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-tpt.tsv.gz	taxonomic names found in the dataset under review aligned with the Terrestrial Parasite Tracker (TPT) Taxonomic Resource as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-tpt.parquet	taxonomic names found in the dataset under review aligned with the Terrestrial Parasite Tracker (TPT) Taxonomic Resource as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-wfo.csv.gz	taxonomic names found in the dataset under review aligned with the World of Flora Online as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-wfo.html.gz	taxonomic names found in the dataset under review aligned with the World of Flora Online as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format

filename	description
indexed-names-resolved-wfo.tsv.gz	taxonomic names found in the dataset under review aligned with the World of Flora Online as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-wfo.parquet	taxonomic names found in the dataset under review aligned with the World of Flora Online as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-worms.csv.gz	taxonomic names found in the dataset under review aligned with the World Register of Marine Species (WoRMS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-worms.html.gz	taxonomic names found in the dataset under review aligned with the World Register of Marine Species (WoRMS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-worms.tsv.gz	taxonomic names found in the dataset under review aligned with the World Register of Marine Species (WoRMS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-worms.parquet	taxonomic names found in the dataset under review aligned with the World Register of Marine Species (WoRMS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-sample.csv	first 500 taxonomic names found in the dataset under review in comma-separated values format

filename	description
indexed-names-sample.html	first 500 taxonomic names found in the dataset under review in html format
indexed-names-sample.tsv	first 500 taxonomic names found in the dataset under review in tab-separated values format
interaction.svg	diagram summarizing the data model used to index species interaction claims
nanopub-sample.trig	first 500 species interaction claims as expressed in the nanopub format (Kuhn and Dumontier 2014)
nanopub.trig.gz	species interaction claims as expressed in the nanopub format (Kuhn and Dumontier 2014)
process.svg	diagram summarizing the data review processing workflow
prov.nq	origin of the dataset under review as expressed in rdf/nquads
review.csv.gz	review notes associated with the dataset under review in gzipped comma-separated values format
review.html.gz	review notes associated with the dataset under review in gzipped html format
review.tsv.gz	review notes associated with the dataset under review in gzipped tab-separated values format
review-sample.csv	first 500 review notes associated with the dataset under review in comma-separated values format
review-sample.html	first 500 review notes associated with the dataset under review in html format
review-sample.tsv	first 500 review notes associated with the dataset under review in tab-separated values format
review.svg	a review badge generated as part of the dataset review process
zenodo.json	metadata of this review expressed in Zenodo record metadata

Archived Dataset

Note that *data.zip* file in this archive contains the complete, unmodified archived dataset under review.

Biotic Interactions

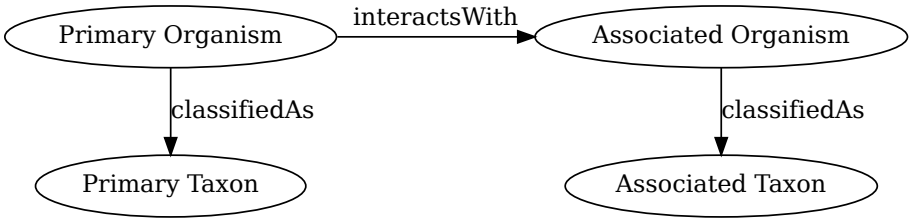

Figure 2: Biotic Interaction Data Model

In this review, biotic interactions (or biotic associations) are modeled as a primary (aka subject, source) organism interacting with an associate (aka object, target) organism. The dataset under review classified the primary/associate organisms with specific taxa. The primary and associate organisms The kind of interaction is documented as an interaction type.

The dataset under review, named globalbioticinteractions/pollinators-of-apocynaceae-database, has fingerprint hash://md5/ba7bfe57d28af24d7dad244f9882dbc5, is 1.20MiB in size and contains 5,060 interaction with 3 unique types of associations (e.g., pollinatedBy) between 679 primary taxa (e.g., Asclepias syriaca) and 2,330 associated taxon (e.g., unidentified).

An exhaustive list of indexed interaction claims can be found in gzipped csv, tsv and parquet archives. To facilitate discovery, a preview of claims available in the gzipped html page at indexed-interactions.html.gz are shown below.

The exhaustive list was used to create the following data summaries below.

Table 3: Sample of Indexed Interaction Claims

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Alstonia actinophylla	flowersVisitedBy	Pteropus scapulatus	Mickelburgh et al.1992
Alstonia boonei	pollinatedBy	Apis mellifera	Albers & Van der Maesen 1994
Alstonia boonei	pollinatedBy	Hypolimnus missipus	Albers & Van der Maesen 1994
Alstonia boonei	pollinatedBy	Eumenidae	Albers & Van der Maesen 1994

Table 4: Most Frequently Mentioned Interaction Types (up to 20 most frequent)

interactionTypeName	count
pollinatedBy	4794
flowersVisitedBy	248
interactsWith	18

Table 5: Most Frequently Mentioned Primary Taxa (up to 20 most frequent)

sourceTaxonName	count
Asclepias syriaca	222
Asclepias incarnata	182
Asclepias verticillata	161
Asclepias tuberosa	110
Catharanthus roseus	105
Asclepias curassavica	95
Oxypetalum balansae	88
Prestonia coalita	83
Ditassa burchellii	72
Vincetoxicum pycnostelma	68
Allamanda blanchetii	66
Gomphocarpus physocarpus	57
Araujia brachystephana	57
Gonolobus parviflorus	56
Orthosia congesta	56
Calotropis procera	53
Oxypetalum banksii ssp. corymbiferum	51
Hancornia speciosa	49
Periploca laevigata	47

Table 6: Most Frequently Mentioned Associate Taxa (up to 20 most frequent)

targetTaxonName	count
unidentified	410
unknown	189
Unidentified	140
Apis mellifera	100
Bees	74

targetTaxonName	count
Forcipomyia sp.	50
Zoophily	40
Unidentified 1	27
Unidentified 2	27
Sphingidae	25
Papilio demodocus demodocus	25
Musca domestica	22
Atrichelaphinis tigrina	20
Calliphora sp.	19
Entomophily	18
Megaselia sp.	18
Birds	17
Desmometopa inaurata	17
Polistes sp.	16

Table 7: Most Frequent Interactions between Primary and Associate Taxa (up to 20 most frequent)

sourceTaxonName	interactionTypeName	targetTaxonName	count
Asclepias syriaca	pollinatedBy	unknown	34
Catharanthus roseus	pollinatedBy	Papilio demodocus demodocus	22
Ceropegia gilgiana	pollinatedBy	unidentified	19
Orbea verrucosa	pollinatedBy	unidentified	15
Periploca laevigata	pollinatedBy	unknown	14
Ceropegia ampliata	pollinatedBy	unidentified	13
Aspidosperma parvifolium	pollinatedBy	?	12
Ceropegia aristolochioides ssp. deflersiana	pollinatedBy	unidentified	12
Chlorocyathus (Kappia) lobulata	pollinatedBy	unidentified	11
Oxypetalum balansae	pollinatedBy	Unidentified	11
Ditassa banksii	pollinatedBy	unidentified	10

sourceTaxonName	interactionTypeName	targetTaxonName	count
Oxypetalum banksii ssp. corymbiferum	pollinatedBy	unidentified	10
Ceropegia racemosa	pollinatedBy	Megaselia sp.	10
Hancornia speciosa	pollinatedBy	Sphingidae	9
Orthosia congesta	pollinatedBy	Unidentified	9
Aspidosperma quebracho-blanco	pollinatedBy	unidentified	8
Allamanda blanchetii	pollinatedBy	Bees	8
Tabernaemontana catharinensis	pollinatedBy	Zoophily	8
Ceropegia achtenii ssp. adolphii	pollinatedBy	Forcipomyia sp.	8

Interaction Networks

The figures below provide a graph view on the dataset under review. The first shows a summary network on the kingdom level, and the second shows how interactions on the family level. It is important to note that both network graphs were first aligned taxonomically using the Catalogue of Life. Please refer to the original (or verbatim) taxonomic names for a more original view on the interaction data.

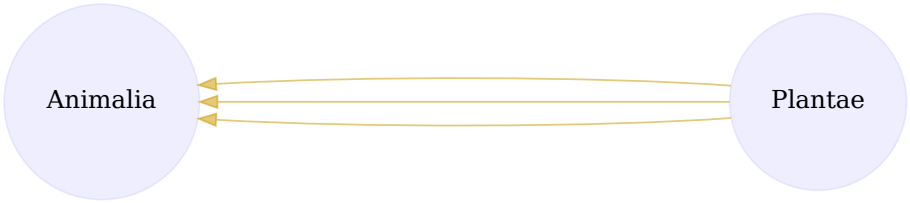

Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

You can download the indexed dataset under review at indexed-interactions.csv.gz. A tab-separated file can be found at indexed-interactions.tsv.gz

Learn more about the structure of this download at GloBI website, by opening a GitHub issue, or by sending an email.

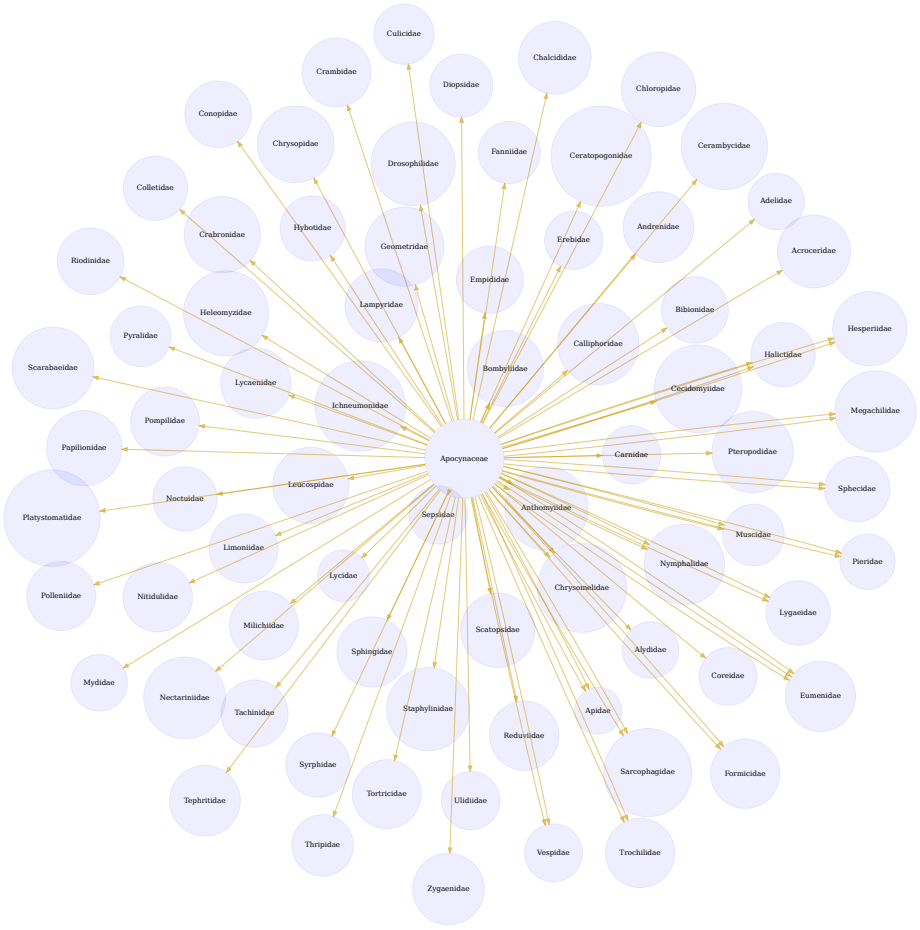

Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. download svg

Another way to discover the dataset under review is by searching for it on the GloBI website.

Taxonomic Alignment

As part of the review, all names are aligned against various name catalogs (e.g., col, ncbi, discoverlife, gbif, itis, wfo, mdd, tpt, pbdb, and worms). These alignments can help review name usage or aid in selecting of a suitable taxonomic name resource.

Table 8: Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
Abrostola triplasia	HAS_ACCEPTED_NAME	col	Abrostola triplasia
Abrostola triplasia	SYNONYM_OF	col	Abrostola triplasia
Cantharid beetle	NONE	col	Cantharid beetle
Acanthiophilus helianthi	HAS_ACCEPTED_NAME	col	Acanthiophilus helianthi

Table 9: Distribution of Taxonomic Ranks of Aligned Names by Catalog. Names that were not aligned with a catalog are counted as NAs. So, the total number of unaligned names for a catalog will be listed in their NA row.

resolvedCatalogName	resolvedRank	count
col	NA	788
col	class	2
col	family	37
col	genus	292
col	order	5
col	phylum	1
col	species	1508
col	subfamily	8
col	subgenus	6
col	suborder	1
col	subspecies	50
col	tribe	6
col	variety	3
discoverlife	NA	2451
discoverlife	species	244
gbif	NA	555
gbif	class	2

resolvedCatalogName	resolvedRank	count
gbif	family	39
gbif	form	1
gbif	genus	338
gbif	order	5
gbif	phylum	1
gbif	species	1646
gbif	subspecies	108
gbif	variety	18
itis	NA	1642
itis	class	2
itis	family	36
itis	genus	232
itis	order	6
itis	phylum	1
itis	species	714
itis	subfamily	9
itis	subgenus	1
itis	suborder	3
itis	subspecies	43
itis	superorder	1
itis	tribe	5
mdd	NA	2694
ncbi	NA	1089
ncbi	class	2
ncbi	cohort	1
ncbi	family	36
ncbi	genus	271
ncbi	order	5
ncbi	phylum	1
ncbi	species	1211
ncbi	subfamily	14
ncbi	subgenus	25
ncbi	suborder	2
ncbi	subspecies	48
ncbi	tribe	4
ncbi	varietas	2
pbdb	NA	2510
pbdb	class	2
pbdb	family	38
pbdb	genus	111
pbdb	infraclass	1
pbdb	order	5
pbdb	phylum	1
pbdb	species	10

resolvedCatalogName	resolvedRank	count
pbdb	subfamily	11
pbdb	suborder	3
pbdb	tribe	5
tpt	NA	2659
tpt	species	35
wfo	NA	2087
wfo	family	1
wfo	genus	41
wfo	species	556
wfo	subspecies	8
wfo	variety	4
worms	NA	2506
worms	class	1
worms	family	28
worms	genus	77
worms	order	5
worms	phylum	1
worms	species	71
worms	subclass	1
worms	subfamily	2
worms	suborder	1
worms	subspecies	1

Table 10: Name relationship types per catalog. Name relationship type "NONE" means that a name was not recognized by the associated catalog. "SAME_AS" indicates either a "HAS_ACCEPTED_NAME" or "SYNONYM_OF" name relationship type. We recognize that "SYNONYM_OF" encompasses many types of nomenclatural synonymies (ICZN 1999) (e.g., junior synonym, senior synonyms).

resolvedCatalogName	relationName	count
col	HAS_ACCEPTED_NAME	1898
col	SYNONYM_OF	512
col	NONE	860
discoverlife	NONE	2734
discoverlife	HAS_ACCEPTED_NAME	218
discoverlife	SYNONYM_OF	91
discoverlife	HOMONYM_OF	17
gbif	HAS_ACCEPTED_NAME	2246
gbif	NONE	615
gbif	SYNONYM_OF	724

resolvedCatalogName	relationName	count
itis	NONE	1753
itis	HAS_ACCEPTED_NAME	1131
itis	SYNONYM_OF	131
mdd	NONE	2976
mdd	HAS_ACCEPTED_NAME	2
ncbi	SAME_AS	1753
ncbi	NONE	1193
ncbi	SYNONYM_OF	121
pbdb	NONE	2687
pbdb	HAS_ACCEPTED_NAME	286
pbdb	SYNONYM_OF	21
tpt	NONE	2934
tpt	HAS_ACCEPTED_NAME	40
tpt	SYNONYM_OF	4
wfo	NONE	2348
wfo	SYNONYM_OF	108
wfo	HAS_ACCEPTED_NAME	560
wfo	HAS_UNCHECKED_NAME	19
worms	NONE	2706
worms	HAS_ACCEPTED_NAME	260
worms	SYNONYM_OF	22

Table 11: List of Available Name Alignment Reports

catalog name	alignment results
col	associated names alignments report in gzipped html, csv, and tsv)
ncbi	associated names alignments report in gzipped html, csv, and tsv)
discoverlife	associated names alignments report in gzipped html, csv, and tsv)
gbif	associated names alignments report in gzipped html, csv, and tsv)
itis	associated names alignments report in gzipped html, csv, and tsv)
wfo	associated names alignments report in gzipped html, csv, and tsv)
mdd	associated names alignments report in gzipped html, csv, and tsv)
tpt	associated names alignments report in gzipped html, csv, and tsv)

catalog name	alignment results
pbdb	associated names alignments report in gzipped html, csv, and tsv)
worms	associated names alignments report in gzipped html, csv, and tsv)

Additional Reviews

Elton, Nomer, and other tools may have difficulties interpreting existing species interaction datasets. Or, they may misbehave, or otherwise show unexpected behavior. As part of the review process, detailed review notes are kept that document possibly misbehaving, or confused, review bots. An sample of review notes associated with this review can be found below.

Table 12: First few lines in the review notes.

reviewDate	reviewCommentType	reviewComment
2025-07-08T06:28:03Z	summary	https://github.com/globalbioticinteractions/pollinators-of-apocynaceae-database/archive/c940b96cd5ecd4d26005eb6b65a1ea6b2
2025-07-08T06:28:03Z	summary	5060 interaction(s)
2025-07-08T06:28:03Z	summary	0 note(s)
2025-07-08T06:28:03Z	summary	5061 info(s)

In addition, you can find the most frequently occurring notes in the table below.

: Most frequently occurring review notes, if any.

For additional information on review notes, please have a look at the first 500 Review Notes in html format or the download full gzipped csv or tsv archives.

GloBI Review Badge

As part of the review, a review badge is generated. This review badge can be included in webpages to indicate the review status of the dataset under review.

Figure 5: Picture of a GloBI Review Badge [3]

Note that if the badge is green, no review notes were generated. If the badge is yellow, the review bots may need some help with interpreting the species interaction data.

[3] Up-to-date status of the GloBI Review Badge can be retrieved from the GloBI Review Depot

GloBI Index Badge

If the dataset under review has been registered with GloBI, and has been succesfully indexed by GloBI, the GloBI Index Status Badge will turn green. This means that the dataset under review was indexed by GloBI and is available through GloBI services and derived data products.

Figure 6: Picture of a GloBI Index Badge [4]

If you'd like to keep track of reviews or index status of the dataset under review, please visit GloBI's dataset index [5] for badge examples.

Discussion

This review and archive provides a means of creating citable versions of datasets that change frequently. This may be useful for dataset managers, including natural history collection data managers, as a backup archive of a shared Darwin Core archive. It also serves as a means of creating a trackable citation for the dataset in an automated way, while also including some information about the contents of the dataset.

This review aims to provide a perspective on the dataset to aid in understanding of species interaction claims discovered. However, it is important to note that this review does *not* assess the quality of the dataset. Instead, it serves as an indication of the open-ness[6] and FAIRness (Wilkinson et al. 2016; Trekels et al. 2023) of the dataset: to perform this review, the data was likely openly available, **F**indable, **A**ccessible, **I**nteroperable and **R**eusable. The current Open-FAIR assessment is qualitative, and a more quantitative approach can be implemented with specified measurement units.

This report also showcases the reuse of machine-actionable (meta)data, something highly recommended by the FAIR Data Principles (Wilkinson et al. 2016). Making (meta)data machine-actionable enables more precise procesing by computers, enabling even naive review bots like Nomer and Elton to interpret the data effectively. This capability is crucial for not just automating the generation of reports, but also for facilitating seamless data exchanges, promoting interoperability.

[4] Up-to-date status of the GloBI Index Badge can be retrieved from GloBI's API

[5] At time of writing (2025-07-08) the version of the GloBI dataset index was available at https://globalbioticinteractions.org/datasets

[6] According to http://opendefinition.org/: "Open data is data that can be freely used, reused and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike."

Acknowledgements

We thank the many humans that created us and those who created and maintained the data, software and other intellectual resources that were used for producing this review. In addition, we are grateful for the natural resources providing the basis for these human and bot activities. Also, thanks to https://github.com/zygoballus for helping improve the layout of the review tables.

Author contributions

Nomer was responsible for name alignments. Elton carried out dataset extraction, and generated the review notes. Preston tracked, versioned, and packaged, the dataset under review.

References

Elliott, Michael, Jorrit Poelen, Icaro Alzuru, Emilio Berti, and partha04patel. 2025. "Bio-Guoda/Preston: 0.10.5." Zenodo. https://doi.org/10.5281/zenodo.14662206.

ICZN. 1999. "International Code of Zoological Nomenclature." The International Trust for Zoological Nomenclature, London, UK. https://www.iczn.org/the-code/the-code-online/.

Kuhn, Tobias, and Michel Dumontier. 2014. "Trusty URIs: Verifiable, Immutable, and Permanent Digital Artifacts for Linked Data." In *The Semantic Web: Trends and Challenges*, edited by Valentina Presutti, Claudia d'Amato, Fabien Gandon, Mathieu d'Aquin, Steffen Staab, and Anna Tordai, 395–410. Cham: Springer International Publishing.

Kuhn, Tobias, Jorrit Poelen, and Katrin Leinweber. 2025. "Globalbioticinteractions/Elton: 0.15.1." Zenodo. https://doi.org/10.5281/zenodo.14927734.

Poelen, Jorrit H. (ed.). 2024. "Nomer Corpus of Taxonomic Resources Hash://Sha256/ B60c0d25a16ae77b24305782017b1a270b79b5d1746f832650F2027ba536e276 Hash://Md5/17f1363a277ee0e4ecaf1b91c665e47e." Zenodo. https://doi.org/10.5281/zenodo.12695629.

Poelen, Jorrit H., James D. Simons, and Chris J. Mungall. 2014. "Global Biotic Interactions: An Open Infrastructure to Share and Analyze Species-Interaction Datasets." *Ecological Informatics* 24 (November): 148–59. https://doi.org/10.1016/j.ecoinf.2014.08.005.

Poelen, Jorrit, Katja Seltmann, and Daniel Mietchen. 2024. "Globalbioticinteractions/Globinizer: 0.4.0." Zenodo. https://doi.org/10.5281/zenodo.10647565.

Salim, José Augusto, and Jorrit Poelen. 2025. "Globalbioticinteractions/Nomer: 0.5.15." Zenodo. https://doi.org/10.5281/zenodo.14893840.

Trekels, Maarten, Debora Pignatari Drucker, José Augusto Salim, Jeff Ollerton, Jorrit Poelen, Filipi Miranda Soares, Max Rünzel, Muo Kasina,

Quentin Groom, and Mariano Devoto. 2023. "WorldFAIR Project (D10.1) Agriculture-related pollinator data standards use cases report." Zenodo. https://doi.org/10.5281/zenodo.8176978.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3 (1). https://doi.org/10.1038/sdata.2016.18.