

# A Review of Biotic Interactions and Taxon Names Found in globalbioticinteractions/fmnh

By Nomer and Elton, two naive review bots.

2023-10-30

## 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 describe a review process of such an openly accessible digital interaction datasets of known origin, and discuss their outcome. The dataset under review (aka globalbioticinteractions/fmnh) contains 127,847 interactions with 8 (e.g., adjacentTo) unique types of associations between 22,308 primary taxa (e.g., *Trichobius joblingi* Wenzel, 1966) and 33,726 associated taxa (e.g., *Carollia perspicillata*). The report includes detailed summaries of interactions data as well as a taxonomic review from multiple perspectives.

## Contents

<b>Introduction</b>	<b>2</b>
Data Review . . . . .	2
<b>Methods</b>	<b>2</b>
<b>Results</b>	<b>3</b>
Biotic Interactions . . . . .	3
Interaction Networks . . . . .	7
Taxonomic Alignment . . . . .	7
Additional Reviews . . . . .	13
GloBI Review Badge . . . . .	14
GloBI Index Badge . . . . .	14
<b>Discussion</b>	<b>14</b>
<b>Acknowledgements</b>	<b>15</b>
<b>References</b>	<b>15</b>

# Introduction

## Data Review

Data review can be a time consuming process, especially when done manually. This review report aims to help facilitate data review of species interaction claims made in datasets registered with Global Biotic Interactions (Poelen, Simons, and Mungall 2014). The review includes summary statistics of, and observations about, the dataset under review:

Field Museum of Natural History (Botany) Pteridophyte Collection -  
Version 2.9 [https://fmipt.fieldmuseum.org/ipt/archive.do?r=fmnh\\_pteridophytes](https://fmipt.fieldmuseum.org/ipt/archive.do?r=fmnh_pteridophytes)  
2023-10-27T16:28:14.364Z 318f6e1904be7e0a8d13598f5e1cf5d2145322d946354170ee4faf24e1c297e2

## Methods

The review is performed through programmatic scripts that leverage tools like Preston, Elton, Nomer combined with third-party tools like grep, mlr, tail and head.

Table 1: Tools used in this review process

tool name	version
elton	
nomer	0.5.6
mlr	6.0.0
pandoc	3.1.6.1

The review process can be described in the form of a script:

```
# get versioned copy of the dataset under review
elton pull globalbioticinteractions/fmnh

# generate review notes
elton review globalbioticinteractions/fmnh\
  > review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/fmnh\
  > interactions.tsv

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

or visually, in a process diagram.

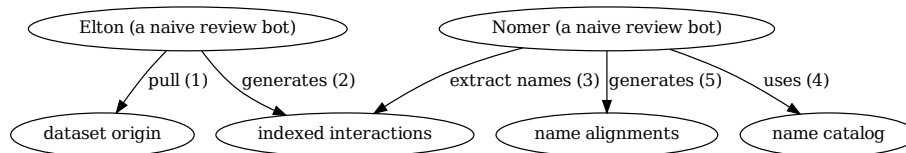


Figure 1: Review Process Overview

You can find a recent copy of the full review script at [check-data.sh](https://github.com/ncbi/ncbi-check-data.sh).

## Results

In the following sections, the results of the review are summarized <sup>1</sup>. Then, links to the detailed review reports are provided.

### Biotic Interactions

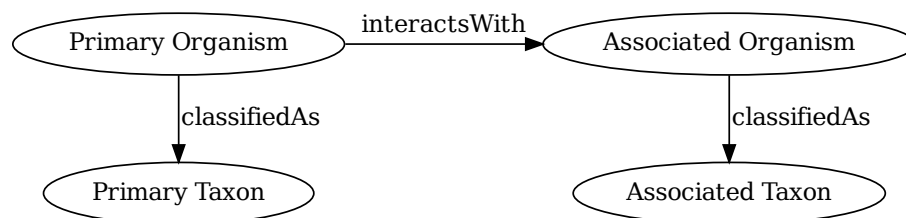


Figure 2: Biotic Interaction Data Model

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 (aka [globalbioticinteractions/fmnh](https://github.com/ncbi/globalbioticinteractions/fmnh)) contains 127,847 interactions with 8 (e.g., `adjacentTo`) unique types of associations between 22,308 primary taxa (e.g., *Trichobius joblingi* Wenzel, 1966) and 33,726 associated taxa (e.g., *Carollia perspicillata*).

An exhaustive list of indexed interaction claims can be found at [indexed-interactions \(csv/tsv/html\)](#). The list was used to create the following data summaries.

<sup>1</sup>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.

Table 2: Sample of Indexed Interaction Claims

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
4f24eb36-68f1-419a-a306-30552c6a02f4	eatenBy	Thunnus	Field Museum of Natural History (Zoology) Invertebrate Collection - Version 18.36. Accessed at <a href="https://fmipt.fieldmuseum.org/ipt/archive.do?r=fmnh_invertebrates">https://fmipt.fieldmuseum.org/ipt/archive.do?r=fmnh_invertebrates</a> on 30 Oct 2023.
Bradybaena similaris (Férussac, 1822)	adjacentTo	rotting logs	8f0fddb1-930a-490e-8367-ea627f02347d
Euglandina rosea (Férussac, 1821)	interactsWith	fields & hedges	4cd537ec-1061-4114-8bc8-4c858fa10346
Oxychilus draparnaudi (Beck, 1837)	adjacentTo	trunk & leaves of ???	e67f9f48-7a5d-4c61-b5d1-6fd9be5e8346

Table 3: Most Frequently Mentioned Interaction Types (up to 20 most frequent)

interactionTypeName	count
adjacentTo	91348
ectoparasiteOf	34986
interactsWith	779
parasiteOf	564
hasHost	172
eats	12
eatenBy	1
hostOf	1

Table 4: Most Frequently Mentioned Primary Taxa (up to 20 most frequent)

sourceTaxonName	count
Trichobius joblingi Wenzel, 1966	2639
Marchantiophyta Stotler & Crand.-Stotl.	1517
Megistopoda aranea (Coquill��t, 1899)	1287
Megistopoda proxima (S��guy, 1926)	1162
Trichobius parasiticus Gervais, 1844	1012
Strebla guajiro (Garcia & Casal, 1965)	1000
Strebla wiedemanni Kolenati, 1856	946
Aspidoptera phyllostomatis (Perty, 1833)	820
Speiseria ambigua Kessel, 1925	779
Aspidoptera falcata Wenzel, 1976	720
Hippoboscoidea	609
Paratrachobius longicrus (Miranda Ribeiro, 1907)	597
Trichobius costalimai Guimar��es, 1938	573
Fungus indet.	547
Nycterophilia coxata Ferris, 1916	543
Usnea Dill. ex Adans.	514
Ixodida Leach, 1815	505
Trichobius dugesii Townsend, 1891	492
Trichobioides perspicillatus (Pessoa & Galvao, 1937)	422

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

targetTaxonName	count
Carollia perspicillata	2692
Artibeus jamaicensis	1740
Desmodus rotundus	1564
Sturnira lilium	1527
ground	1397
tree	1221
rocks	1183
Phyllostomus discolor	1123
log	878
dead wood	802
Pteronotus parnellii	792
Carollia brevicauda	776
soil	763
trees	683
Carollia perspicillata azteca	649
Glossophaga soricina	629

targetTaxonName	count
Artiodactyla	618
leaf litter	599
Peromyscus maniculatus	567

Table 6: Most Frequent Interactions between Primary and Associate Taxa (up to 20 most frequent)

sourceTaxonName	interactionTypeNam	targetTaxonName	count
Trichobius joblingi Wenzel, 1966	ectoparasiteOf	Carollia perspicillata	1377
Megistopoda proxima (Séguy, 1926)	ectoparasiteOf	Sturnira lilium	906
Megistopoda aranea (Coquillétt, 1899)	ectoparasiteOf	Artibeus jamaicensis	793
Trichobius parasiticus Gervais, 1844	ectoparasiteOf	Desmodus rotundus	775
Strebla wiedemanni Kolenati, 1856	ectoparasiteOf	Desmodus rotundus	697
Strebla guajiro (Garcia & Casal, 1965)	ectoparasiteOf	Carollia perspicillata	570
Aspidoptera falcata Wenzel, 1976	ectoparasiteOf	Sturnira lilium	520
Aspidoptera phyllostomatis (Perty, 1833)	ectoparasiteOf	Artibeus jamaicensis	504
Trichobius costalimai Guimarães, 1938	ectoparasiteOf	Phyllostomus discolor	454
Speiseria ambigua Kessel, 1925	ectoparasiteOf	Carollia perspicillata	366
Paratrichobius longicrus (Miranda Ribeiro, 1907)	ectoparasiteOf	Artibeus lituratus	351

sourceTaxonName	interactionTypeNam	targetTaxonName	count
Trichobioides perspicillatus (Pessoa & Galvao, 1937)	ectoparasiteOf	Phyllostomus discolor	340
Trichobius joblingi Wenzel, 1966	ectoparasiteOf	Carollia perspicillata azteca	317
Trichobius caecus Edwards, 1918	ectoparasiteOf	Pteronotus parnellii	292
Strebla hertigi Wenzel, 1966	ectoparasiteOf	Phyllostomus discolor	284
Nycterophilia coxata Ferris, 1916	ectoparasiteOf	Leptonycteris curasoae	272
Trichobius dugesii Townsend, 1891	ectoparasiteOf	Glossophaga soricina	259
Trichobius joblingi Wenzel, 1966	ectoparasiteOf	Carollia brevicauda	256
Trichobius longipes (Rudow, 1871)	ectoparasiteOf	Phyllostomus hastatus	244

### 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. Note that both network graphs were first aligned taxonomically via the Catalogue of Life. Please refer to the original (or verbatim) taxonomic names for a more original view on the interaction data.

You can download the indexed dataset under review at [indexed-interactions.csv](#). A tab-separated file can be found at [indexed-interactions.tsv](#)

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

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 globi mdd tpt). These alignments may serve as

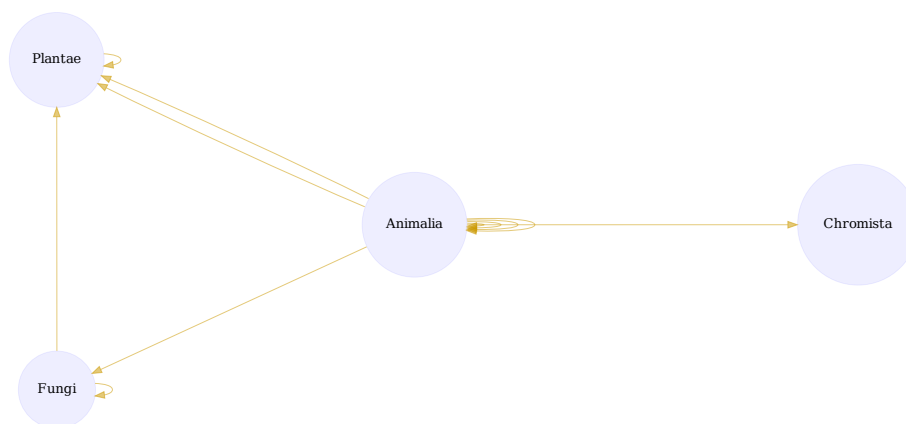


Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

a way to review name usage or aid in selecting of a suitable taxonomic name resource to use.

Table 7: Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
	NONE	ncbi	
	NONE	discoverlife	
	NONE	globi	
Abbreviata borneensis	NONE	col	Abbreviata borneensis

Table 8: 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
tpt	NA	50050
tpt	species	1039
tpt	genus	129
tpt	family	3
tpt	order	2
ncbi	NA	35927
ncbi	species	12950
ncbi	genus	1838
ncbi	subspecies	186
ncbi	family	151
ncbi	order	52
ncbi	8 varietas	35
ncbi	class	25
ncbi	phylum	14
ncbi	subclass	10
ncbi	subgenus	9
ncbi	clade	8
ncbi	subfamily	7
ncbi	infraorder	4



resolvedCatalogName	resolvedRank	count
ncbi	forma	1
ncbi	subtribe	1
mdd	NA	51223
itis	NA	41246
itis	species	7812
itis	genus	1317
itis	subspecies	381
itis	variety	210
itis	family	152
itis	order	52
itis	class	24
itis	phylum	13
itis	subclass	9
itis	division	8
itis	subfamily	6
itis	subgenus	5
itis	suborder	4
itis	kingdom	3
itis	superfamily	3
itis	infraorder	2
itis	subphylum	2
itis	superclass	1
globi	NA	41424
globi	species	21374
globi	genus	3328
globi	subspecies	1221
globi	variety	824
globi	subgenus	316
globi	family	175
globi	form	95
globi	order	56
globi	class	48
globi	phylum	34
globi	subphylum	15
globi	tribe	14
globi	subclass	12
globi	subfamily	11
globi	forma specialis	9
globi	series	6
globi	kingdom	5
globi	suborder	5
globi	infraclass	4
globi	infraorder	4
globi	subtribe	3

resolvedCatalogName	resolvedRank	count
globi	superclass	2
globi	superfamily	2
globi	section	2
globi	infrakingdom	1
globi	superorder	1
globi	parvorder	1
gbif	NA	25441
gbif	species	22545
gbif	genus	2014
gbif	subspecies	1066
gbif	variety	461
gbif	family	162
gbif	order	46
gbif	form	29
gbif	class	24
gbif	phylum	19
gbif	kingdom	4
discoverlife	NA	51224
col	NA	27643
col	species	20606
col	genus	1926
col	subspecies	859
col	variety	217
col	family	148
col	order	47
col	subgenus	45
col	class	22
col	phylum	18
col	subclass	6
col	subfamily	5
col	suborder	4
col	kingdom	3
col	infraorder	3
col	form	3
col	superfamily	3
col	section	1
col	gigaclass	1
col	superclass	1
col	parvorder	1
col	tribe	1

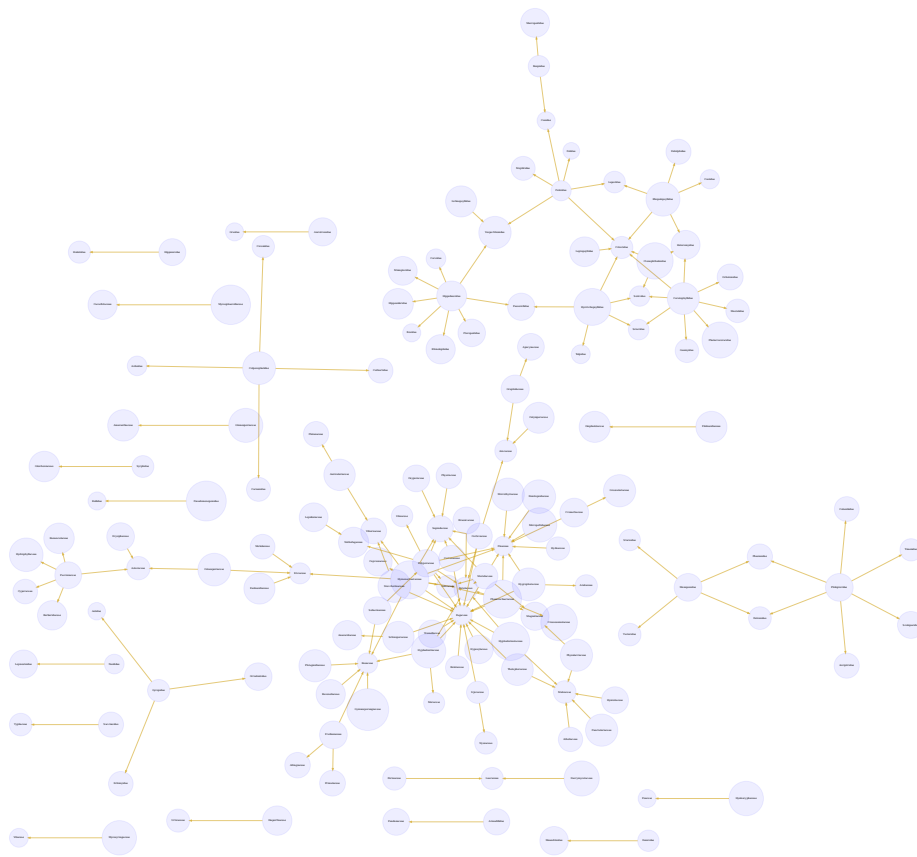


Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. [download svg](#)

Table 9: 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	NONE	29483
col	HAS_ACCEPTED_NAME	17455
col	SYNONYM_OF	11123
discoverlife	NONE	54209
gbif	NONE	27258
gbif	HAS_ACCEPTED_NAME	23535
gbif	SYNONYM_OF	14441
globi	NONE	27985
globi	SAME_AS	198927
itis	NONE	43325
itis	HAS_ACCEPTED_NAME	8731
itis	SYNONYM_OF	2569
mdd	NONE	53423
mdd	HAS_ACCEPTED_NAME	727
ncbi	NONE	37814
ncbi	SAME_AS	14041
ncbi	SYNONYM_OF	2375
ncbi	COMMON_NAME_OF	116
tpt	NONE	52944
tpt	HAS_ACCEPTED_NAME	1197
tpt	SYNONYM_OF	75

Table 10: List of Available Name Alignment Reports

catalog name	alignment results
col	associated names alignments (csv/tsv/html)
ncbi	associated names alignments (csv/tsv/html)
discoverlife	associated names alignments (csv/tsv/html)
gbif	associated names alignments (csv/tsv/html)
itis	associated names alignments (csv/tsv/html)

catalog name	alignment results
globi	associated names alignments (csv/tsv/html)
mdd	associated names alignments (csv/tsv/html)
tpt	associated names alignments (csv/tsv/html)

## 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 11: First few lines in the review notes.

reviewDate	reviewCommentType	reviewComment
2023-10-30T02:08:40Z	note	found unresolved reference [c8bdd05a-1fad-44c3-96fa-367fb96ec0de]
2023-10-30T02:08:40Z	note	found unresolved reference [https://arctos.database.museum/guid/CHAS:Inv:2017.13.1312]
2023-10-30T02:12:02Z	summary	https://github.com/globalbioticinteractions/fmnh/archiv
2023-10-30T02:12:02Z	summary	127863 interaction(s)

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

Table 12: Most frequently occurring review notes, if any.

reviewComment	count
found unresolved reference [c8bdd05a-1fad-44c3-96fa-367fb96ec0de]	1
found unresolved reference [https://arctos.database.museum/guid/CHAS:Inv:2017.13.1312]	1

For more exhaustive list of review notes, please have a look at the Review Notes (csv/tsv/html).

## 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: Sample of a GloBI Review Badge <sup>2</sup>

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.

## GloBI Index Badge

If the dataset under review has been registered with GloBI, and has been successfully 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: Sample of a GloBI Index Badge <sup>3</sup>

If you'd like to keep track of reviews or index status of the dataset under review, please visit GloBI's dataset index <sup>4</sup> for badge examples.

## Discussion

This review is intended to provide a perspective on the dataset to aid understanding of species interaction claims discovered. However, this review should *not* be considered as fitness of use or other kind of quality assessment. Instead, the review may be used as an indication of the open-ness<sup>5</sup> and FAIRness (Wilkinson et al. 2016; Trekels et al. 2023) of the dataset: in order to perform this review, the data was likely openly available, **F**indable, **A**ccessible, **I**nteroperable and **R**eusable. Currently, this Open-FAIR assessment is qualitative, and with measurement units specified, a more quantitative approach can be implemented.

---

<sup>2</sup>Up-to-date status of the GloBI Review Badge can be retrieved from the GloBI Review Depot

<sup>3</sup>Up-to-date status of the GloBI Index Badge can be retrieved from GloBI's API

<sup>4</sup>At time of writing (2023-10-30) the version of the GloBI dataset index was available at [<https://globalbioticinteractions.org/datasets>](<https://globalbioticinteractions.org/datasets>)

<sup>5</sup>According to <http://opendefinition.org/>: "Open data is data that can be freely used, re-used 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.

## References

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