Guide for extracting GloBI TPT data :)

Please add helpful suggestions and instructions for the how-to page here!!
We appreciate your help!

Helpful stuff to be added….
Helpful stuff to be added…

Or feel free to add to the appropriate sections below:

Guide for extracting GloBI TPT data :)

GloBI has a ton of useful data in it and can be used in many different ways. However, it can also be a bit overwhelming with so much data. This page offers some helpful links, hacks, and instructions for how to extract the information available in a useful format.

Introduction to using GloBI

Did you know there is a video tutorial on how to use GloBI?

- A Practical Exploration of Biotic Interaction Data Management and Information Retrieval through TPT and GloBI (video)

There are also detailed step-by-step instructions from the Species Interaction Data Workshop on how to extract information from GloBI.

1. Getting Interaction Data
2. Working with the Whole Dataset
3. Exploring Ixodes (tick) Records By Pointing and Clicking
4. Data Sources: Interaction Data Record Review
5. Data Sources: Taxonomic Name Review

Pre-compiled Datasets

TPT Specific.

All TPT interaction information in GloBI is regularly compiled, published, and available here: https://zenodo.org/record/5572874#.YhAKlfXMLFQ

- Scroll to the bottom of the page for downloadable data sets

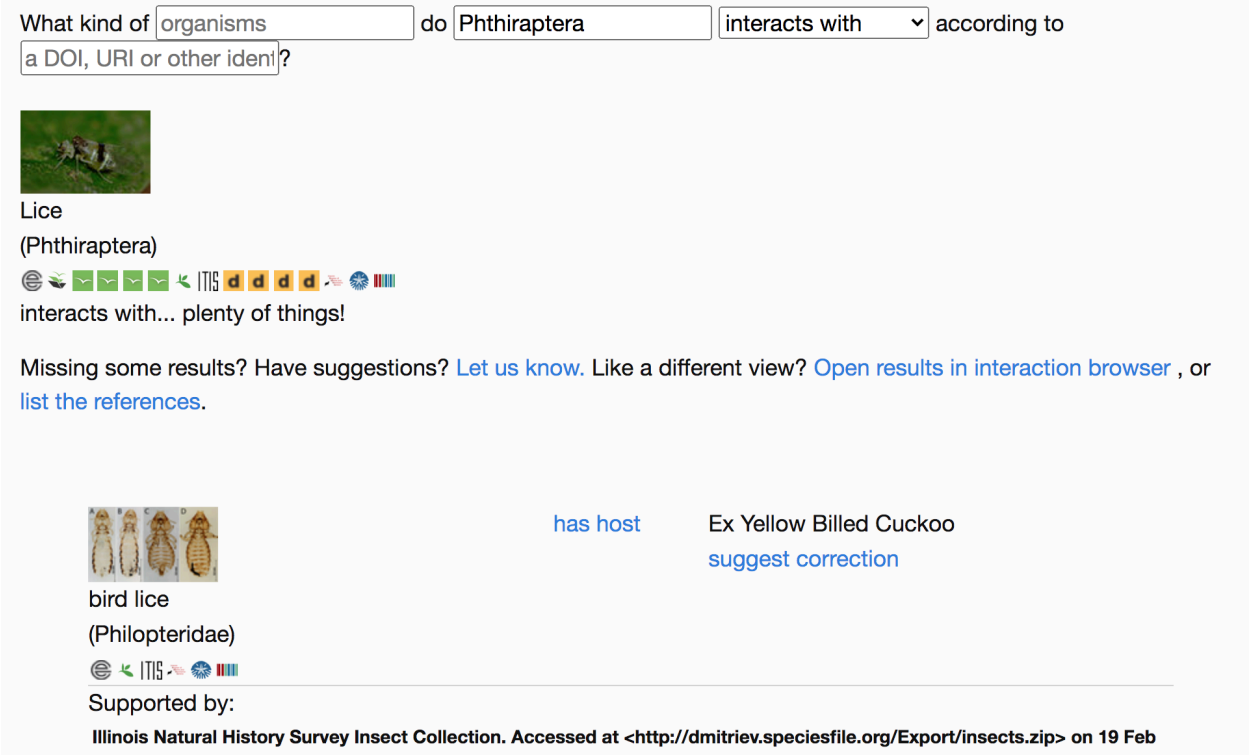
    - The “indexed_interactions_by_collection.tsv” data set gives you a breakdown of the different types of interactions each collection has and how many of each type of interaction.
    - View dataset in a spreadsheet viewing/manipulating program
- * Note: currently, the dataset includes all taxa interactions from collections in the TPT group. We are working on a way to only include taxa based on the TPT taxonomy lists.

Generally useful things

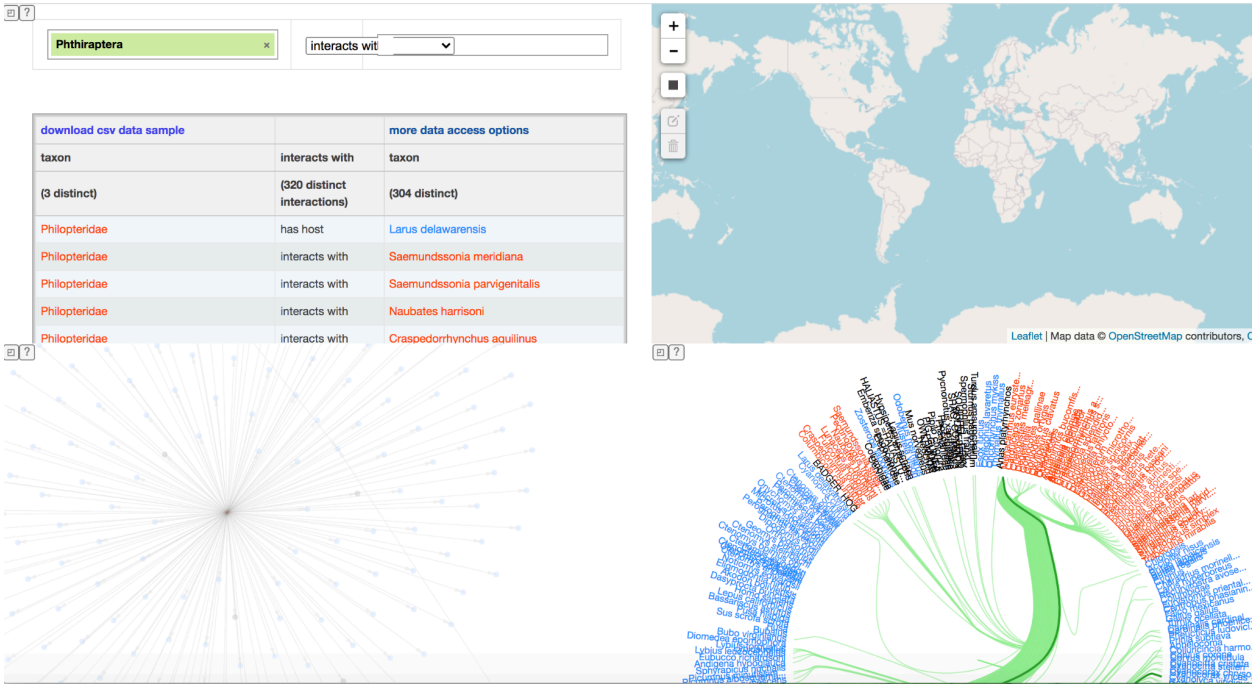
GUI Interaction searches.

To search for interactions, go to the main page https://www.globalbioticinteractions.org/

- Type in an organism in one or both search boxes
- Hit search



- For a downloadable list, click the “Open results in interaction browser” link
- This will bring up a map, two different interactive interaction graphs, and a list of interactions on the left.



- To download a list of the interactions, click the “download csv data sample” at the top of the list.
-

API based interaction searches and pre-made datasets.

A number of commonly searched datasets are premade and available here:
https://github.com/globalbioticinteractions/globalbioticinteractions/wiki/API#references-to-datasets
There are also detailed instructions on API based queries on this page

R based interaction searches.

For those that are comfortable using R, install and use rglobi for more precise datasets.

- Follow the instructions here:
  https://github.com/ropensci/rglobi
  - Detailed instructions on how to use rglobi are provided by Katja Seltmann’s **TPT-GloBI-R-Demo** here:
    https://github.com/ParasiteTracker/TPT-GloBI-R-Demo/blob/master/globi-to-graph2019.R
- To use GloBi data in bipartite interaction analyses, follow instructions here:
  https://www.globalbioticinteractions.org/deadwood2021/13-day-two-part-two/

Command line interaction searches.

For those that are comfortable using the command line on their computers:

- A GloBI command cheat sheet is here that can be modified for your terminal/command shell (just substitute appropriate taxa):

  https://www.globalbioticinteractions.org/deadwood2021/code/cheatsheet.txt
  Download dataset and run commands:

```
$ time cat data/stable/interactions.csv.gz\
  | gunzip\
  | mlr --csv filter '$sourceTaxonKingdomName == "Fungi"'\
  | mlr --csv filter '$targetTaxonGenusName == "Quercus"'\
  | mlr --csv cut -f sourceTaxonName,targetTaxonName\
  > data/oakfungi.csv
$ time cat data/stable/interactions.csv.gz\
  | gunzip\
  | mlr --csv filter '$targetTaxonKingdomName == "Fungi"'\
  | mlr --csv filter '$sourceTaxonGenusName == "Quercus"'\
  | mlr --csv cut -f targetTaxonName,sourceTaxonName\
  | tail -n+2\
  >> data/oakfungi.csv
```

- Download all data here: https://www.globalbioticinteractions.org/data

Matching Names.

To match names in GloBI to names in other databases such as ITIS or NCBI, use the tool Nomer:

https://github.com/globalbioticinteractions/nomer

SCAN specific things

To just look at interactions that are also in SCAN, use the command line.

- List of files available:

  https://depot.globalbioticinteractions.org/reviews/globalbioticinteractions/scan/README.txt

  https://depot.globalbioticinteractions.org/reviews/globalbioticinteractions/scan/indexed-interactions.csv.gz

  https://depot.globalbioticinteractions.org/reviews/globalbioticinteractions/scan/indexed-interactions-sample.csv

- Commands to use:

  cat interactions.csv.gz | gunzip | grep "globalbioticinteractions/scan" > globi-scan.csv

cat interactions.csv.gz | gunzip | grep "globalbioticinteractions/scan" | wc -l

"Ken's how-to-add-an-inaturalist-interaction document"
https://docs.google.com/document/d/12jFMA5a6EH1tqW2DcuNV8AQj2tuzx7ZvQoXtxaObaJ0/edit