Terrestrial Parasite Tracker Data and Collections-Based Research: Updates from the TPT Research Advisory Board and Other Collaborators

TPT Webinar March 31st
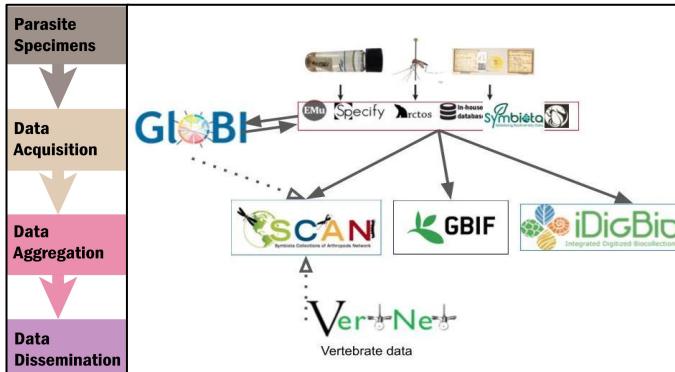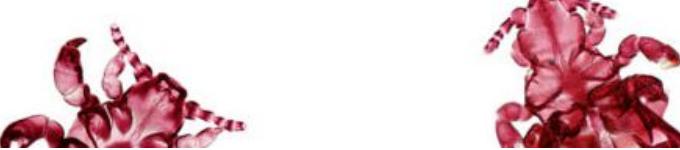2 pm - 3 pm EST

Terrestrial Parasite Tracker TCN

- Transcribe and georeference label data from **1.2+** million arthropod parasite specimens from 22 collections across North America (U.S. and territories)
- Document 500,000+ parasite-host associations via GloBI
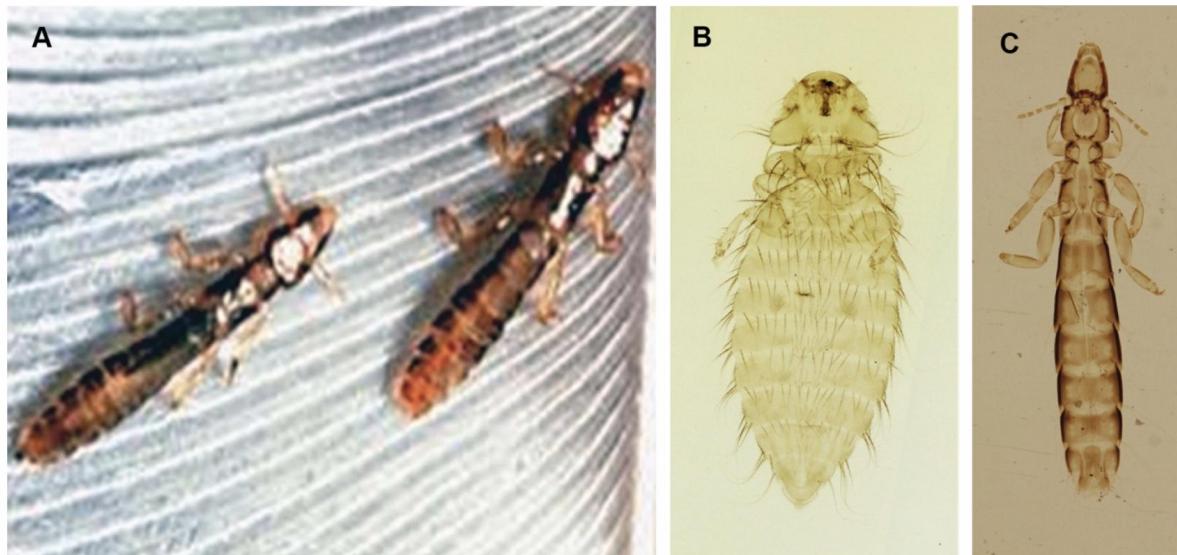
TPT Research Advisory Board

- RAB is comprised of researchers from academic and non-academic institutions
- Working to define project goals and conditions for collaboration (e.g., how to handle student data, data requested for proposals, publications, etc.)
- Developing digitization priorities for research publications (currently ticks, flies, lice)

Contents and archives available through www.bioone.org or www.jstor.org

Journal of Parasitology

CRITICAL COMMENT . . . . . .

A MISIDENTIFICATION CRISIS PLAGUES SPECIMEN-BASED RESEARCH: A CASE FOR GUIDELINES WITH A RECENT EXAMPLE (ALI ET AL., 2020)

Sarah E. Bush[1], Daniel R. Gustafsson[2], Vasyl V. Tkach[3], and Dale H. Clayton[1]
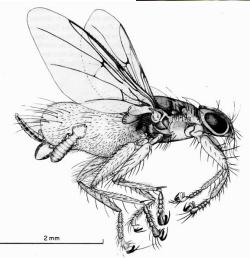
Better Understanding of Parasite Geographic Distributions

- Often assume that lice are <u>host specific</u> and <u>not geographically specific</u> - parasite distributions aren't that simple

- We <u>know very little about louse geographic distributions</u>

- Localities often poorly resolved on louse slide labels but maybe be can get <u>more detailed localities from associated host specimens</u>

- Digitization will help us to construct a picture of parasite geographic specificity with respect to hosts

Guimaraesiella cicchinoi
A widespread New World Trogon parasite

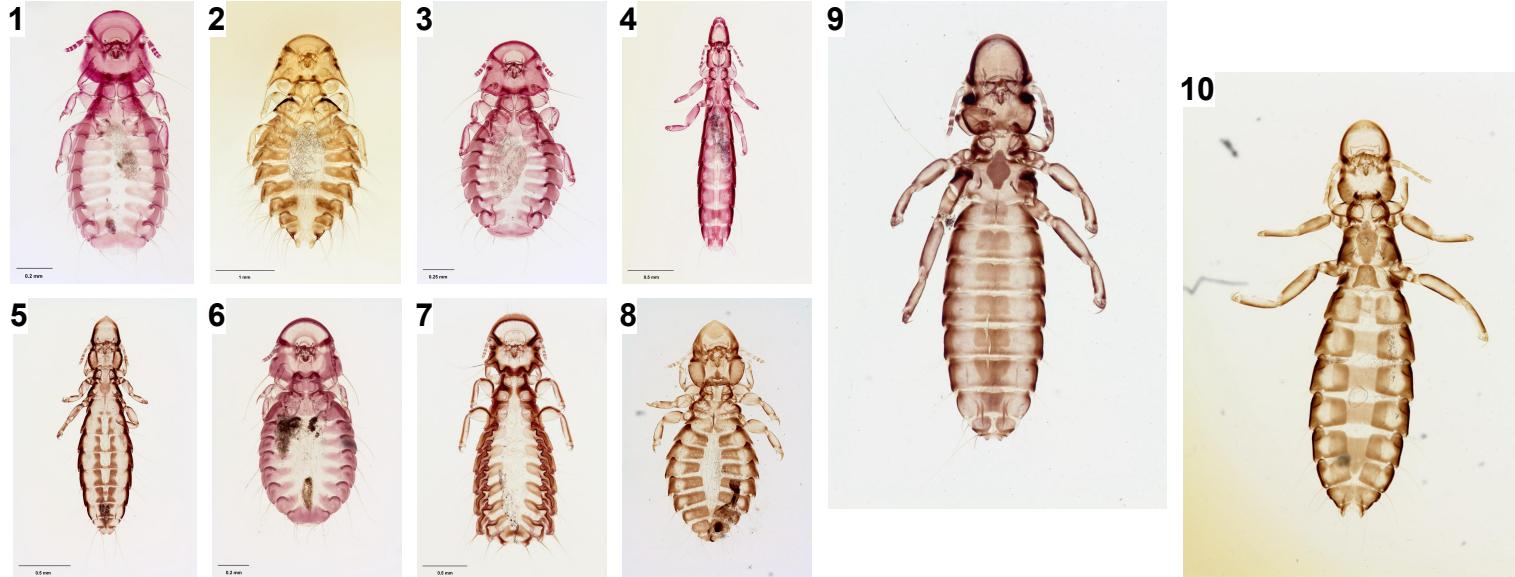# Using Artificial Intelligence to Develop a Taxonomy Guide

0.2 mm

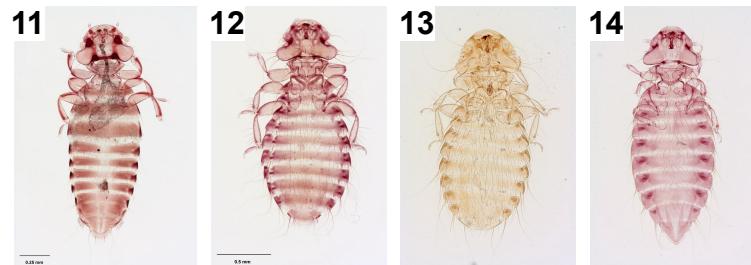# Creating an AI model with 14 louse genera found on poultry

Ischnocera:

1. *Campanulotes*
2. *Chelopistes*
3. *Coloceras*
4. *Columbicola*
5. *Cuclotogaster*
6. *Goniocotes*
7. *Goniodes*
8. *Lagopoecus*
9. *Lipeurus*
10. *Oxylipeurus*

Amblycera:

11. *Colpocephalum*
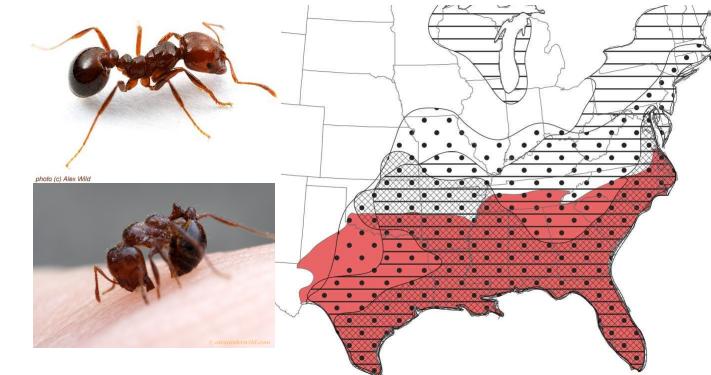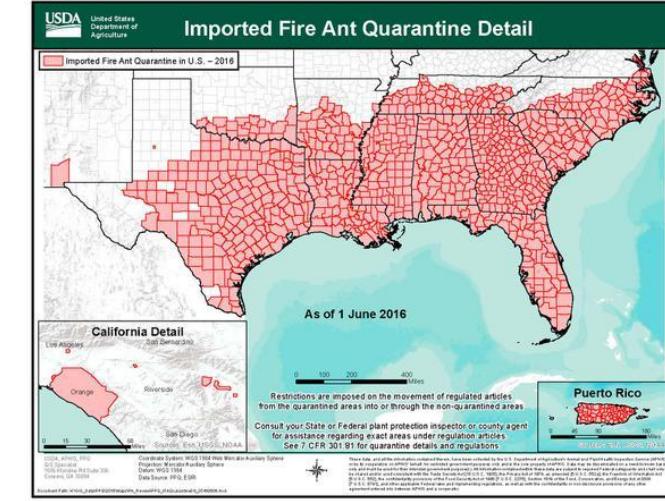12. *Hohorstiella*
13. *Menacanthus*
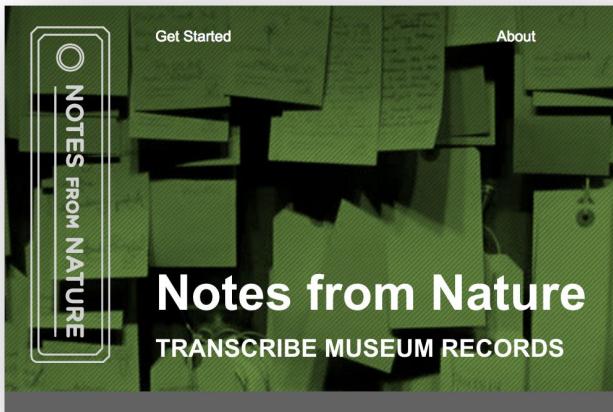14. *Menopon*

Photos: Ricardo Palma

Effect of Red Imported Fire Ants on Tick-borne pathogen vectors in Texas

Jessica Light

Dermacentor variabilis (American Dog ticks)
Larva / Nymph / Adult Male / Adult Female

Amblyomma americanum (Lone Star ticks)
Larva / Nymph / Adult Male / Adult Female

Amblyomma maculatum (Gulf Coast Tick)
Larva / Nymph / Adult Male / Adult Female

Ixodes scapularis (Blacklegged ticks or Deer ticks)
Larva / Nymph / Adult Male / Adult Female

- Ticks are vectors of:
  - Rocky Mountain Spotted Fever
  - Tularemia
  - Ehrlichiosis
  - Spotted fever rickettsiosis
  - Babesiosis
  - Lyme disease

Lyme Disease Bacteria
Borrelia burgdorferi
1 µm

Imported Fire Ant Quarantine Detail
Imported Fire Ant Quarantine in U.S. – 2016
As of 1 June 2016
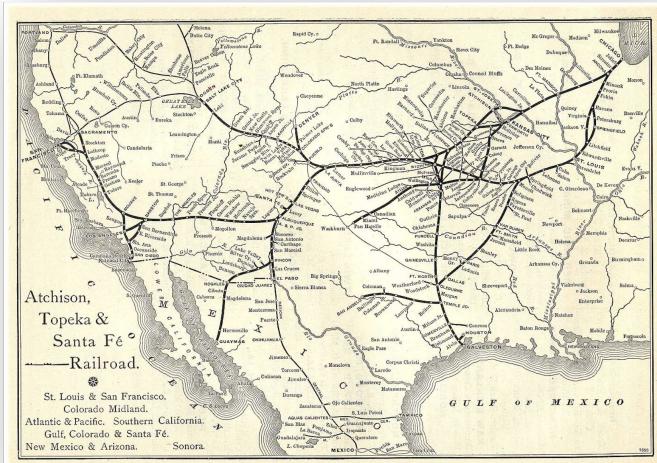California Detail
Puerto Rico

NSF funded Digi-Leap project
Goal: to increase the rate of
transcription and georeferencing

- Adding in machine learning steps to Notes from Nature
- Working collaboratively with volunteers to vet and quality control the data coming from OCR readers
- Information Extraction ML process to parse and identify

NSF funded Digi-Leap project
Goal: to increase the rate of transcription and *georeferencing*

- Millions of specimens have been georeferenced
- Many collectors visit the same areas - producing the same localities
- Building a gazetteer of all georeferenes from GBIF, iDigBio
- New localiaites can be compared with the gazetteer - reduce the number that need to be georeferenced

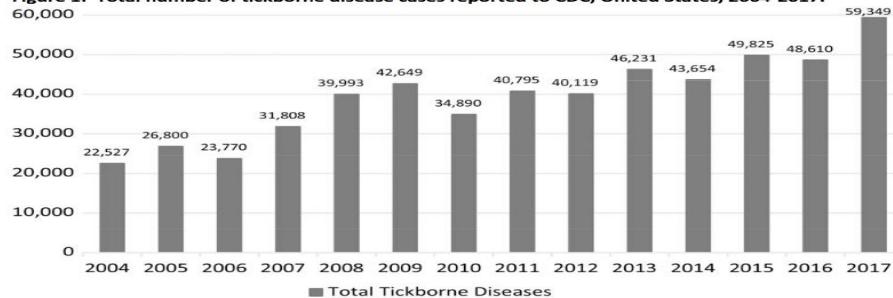# Leveraging Big Data to Improve Prediction of Tick-borne Disease Patterns and Dynamics
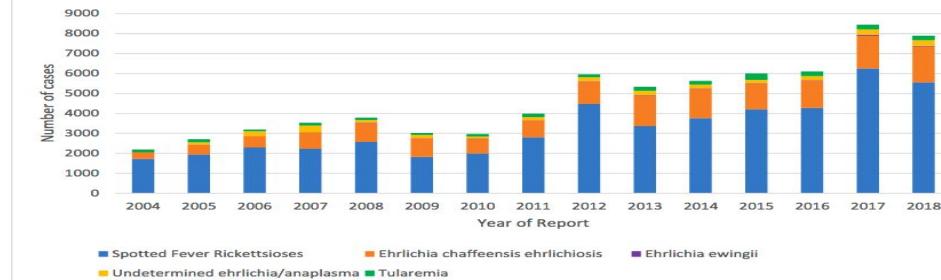
Mike Teglas, University of Nevada Reno

Guide to the Surveillance of Metastriate Ticks (Acari: Ixodidae) and their Pathogens in the United States, CDC

Figure 1. Total number of tickborne disease cases reported to CDC, United States, 2004-2017.

While the numbers of Lyme disease cases make up much of this chart, other tickborne diseases have also been on the rise. Of the nearly 50,000 cases of tickborne disease reported in 2018, nearly 8,000 cases were due to pathogens associated with metastriate ticks (Fig. 2).

Figure 2. Metastriate-transmitted Tickborne Diseases in the United States, 2004–2018

https://tickbase.net/

- Tick Crawler
  - Tick data acquisition from the Internet/Scientific Literature
- Field Collections
  - Project personnel, county and state agencies
- Historical tick data
  - Museums and Collections
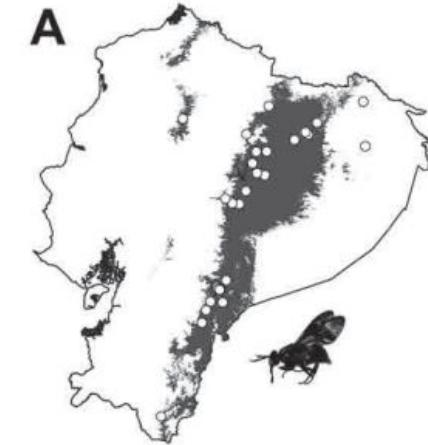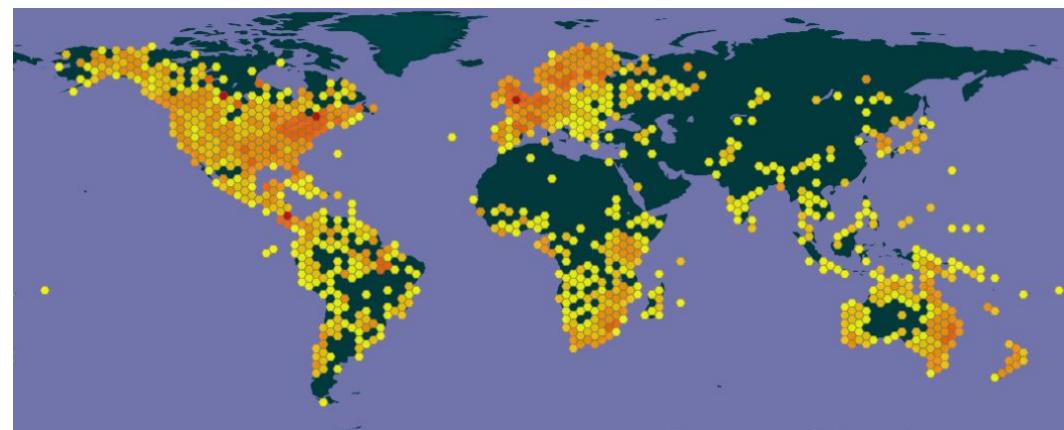
Species Distribution Modeling of North American Tabanids

Sophia Zaslow
Collaborators: Emily Sandall and Maureen Turcatel

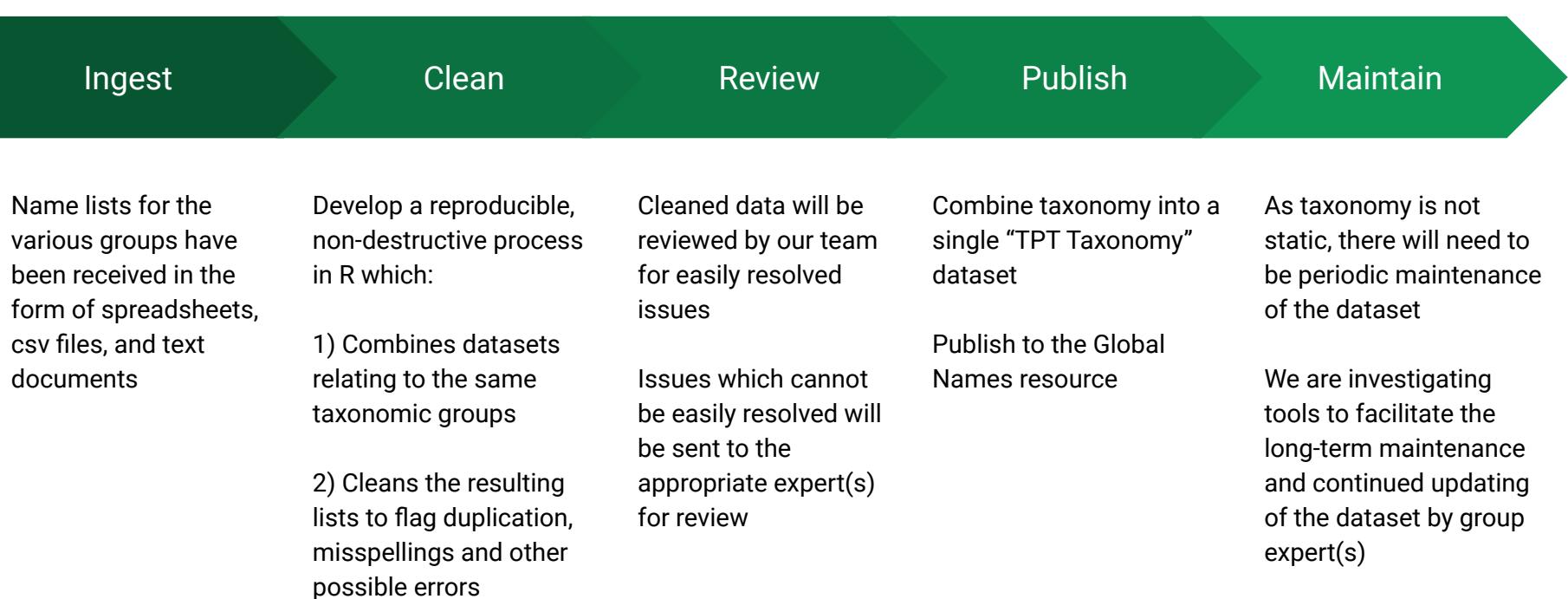Standard set up for Tabanid photographing

Map of Ecuador depicting the potential distribution of Chrysops varians var. tardus

Image of Tabanus rubidus

Map of Tabanid distribution recorded via GBIF https://www.gbif.org/species/6919

TPT Taxonomy Process Overview

Ingest	Clean	Review	Publish	Maintain

Name lists for the various groups have been received in the form of spreadsheets, csv files, and text documents

Develop a reproducible, non-destructive process in R which:

1) Combines datasets relating to the same taxonomic groups

2) Cleans the resulting lists to flag duplication, misspellings and other possible errors

Cleaned data will be reviewed by our team for easily resolved issues

Issues which cannot be easily resolved will be sent to the appropriate expert(s) for review

Combine taxonomy into a single "TPT Taxonomy" dataset

Publish to the Global Names resource

As taxonomy is not static, there will need to be periodic maintenance of the dataset

We are investigating tools to facilitate the long-term maintenance and continued updating of the dataset by group expert(s)

TPT Taxonomy Cleaning

01 Standardize
- Fix formatting of names and author strings (e.g., capitalization)
- Remove unresolvable names (sp, unidentified, probably, nov sp, etc.)
- Cast column headers into Darwin Core format

02 De-Duplicate
- Remove duplicate names
- Flag names with very similar and/or short epithets

03 Complete
- Add missing higher taxa
- Add author citation
- Add synonyms

04 Review
- TPT Team review of flagged names
- Taxonomic expert review of remaining flagged names

05 Finalize
- Combine resolved names into working taxonomy file
- Track remaining rejected names in a separate list
- Provide TPT taxonomy to Global Names

TPT Taxonomy Reconciliation
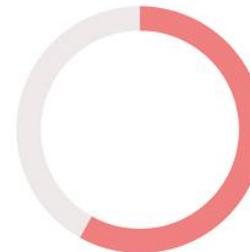
GBIF | Global Biodiversity Information Facility

Terrestrial Parasite Tracker
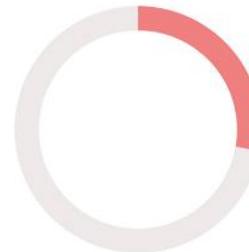
How do their contents differ?

Have we captured everything?
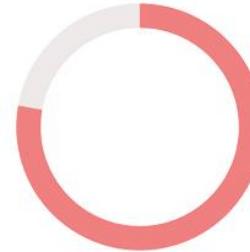
Lice
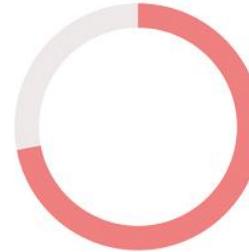Accepted: 5296
Invalid: 3824
Total: 9120

Ticks
Accepted: 854
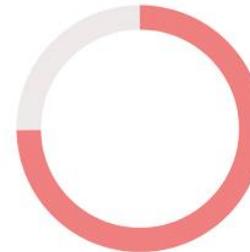Invalid: 2221
Total: 3075
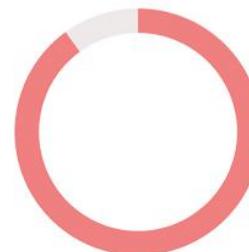
Fleas
Accepted: 2212
Invalid: 632
Total: 2844

Diptera
Accepted: 18664
Invalid: 7264
Total: 25928

Mites
Accepted: 31432
Invalid: 10598
Total: 42030

Hemiptera
Accepted: 7609
Invalid: 845
Total: 8454

Derived from GBIF backbone snapshot taken on March 26, 2021

Digitized Specimen Data = Entomological Intelligence

Information about what specimens are available for study, vouchered distribution records and host associations

Type locality:
"Carolina" (exact site and host unknown)

United States National Tick Collection (USNTC) Specimens:
Males: 1673
Females: 1055
Nymphs: 1940
Larvae: 4101
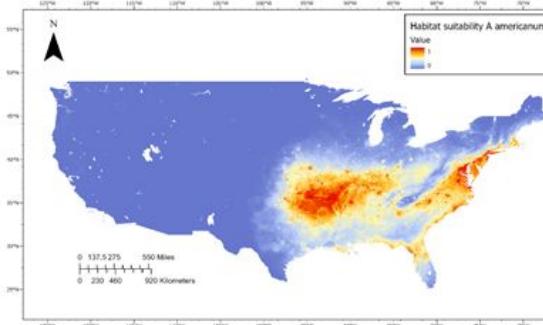Specimens documented from the following states:
Alabama, Arizona, Arkansas, District of Columbia, Florida, Georgia, Indiana, Kansas, Louisiana, Mississippi, Montana, New Mexico, North Carolina, Oklahoma, South Carolina, Texas, Texas, Virginia

Geo-referenced collection sites can drive ecological niche models, providing situational awareness

Amblyomma americanum (Linnaeus, 1758) Habitat suitability Model
Associated Pathogens: *Ehrlichia chaffeensis* and *E. ewingii* (ehrlichiosis), *Francisella tularensis* (tularemia), Heartland virus (Heartland virus disease), Bourbon virus (Bourbon virus disease), and Southern tick-associated rash illness (STARI). *Rickettsia rickettsia* (Rocky Mountain Spotted Fever), *Rickettsia parkeri*, *Coxiella burnetii* (Q fever), *Borrelia lonestari* (Suspected cause of southern tick-associated rash illness).

Amblyomma americanum (Lone star tick)
Adult female (dorsal view) Photo credit: WRBU

Making data available via EMU allows for integration with other DB

Smithsonian
National Museum of Natural History

WRBU

https://globalbioticinteractions.org/parasitetracker
Indexing and Reviewing Research Datasets

Creating Research Datasets: Directed Research in Data Science for Undergraduates

>>>>>>>GOAL

Understandable by people and computers

Katja Seltmann, seltmann@ucsb.edu

Link to taxonomic treatments

Bee Interaction & Tick Interaction Databases

Data review with GloBI

Wednesday, April 28, 2021 - 2:00pm to 4:00pm EDT

A Practical Exploration of Biotic Interaction Data Management and Information Retrieval through TPT and GloBI

Register at iDigBio:
https://www.idigbio.org/content/practical-exploration-biotic-interaction-data-management-and-information-retrieval-through

Terrestrial Parasite Tracker