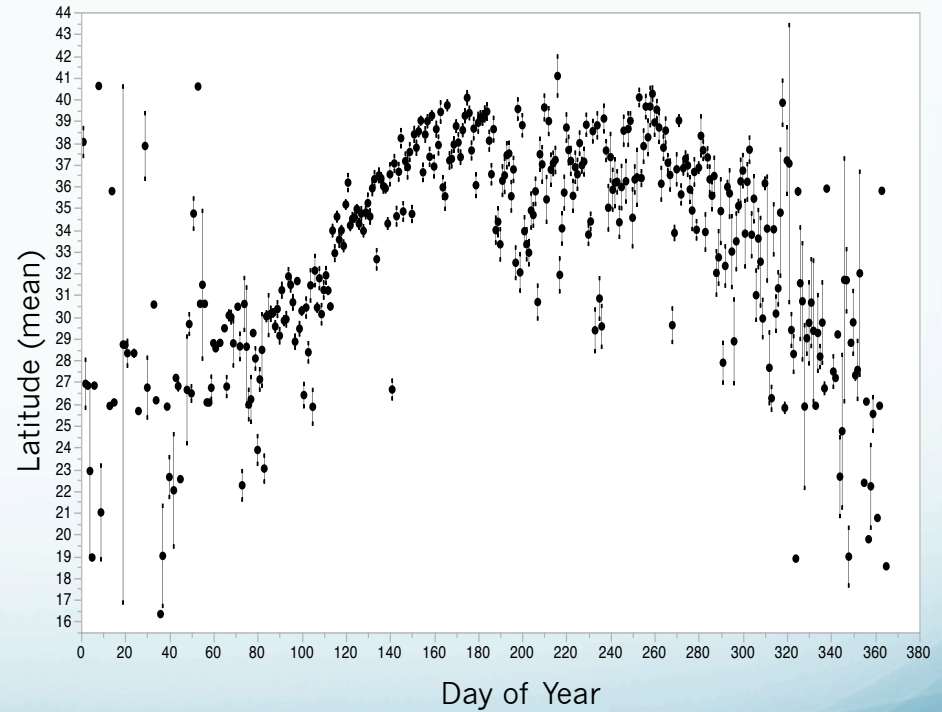


# From Museum Specimen Database<sup>1</sup> to Ecological Statement

Christine A. Johnson  
Division of Invertebrate Zoology



# Tri-trophic Digitization Team

## *Thematic Collections Network*

***Toby Schuh (Lead PI)***, Christine Johnson, ***Charles Bartlett***, John Heraty, Bob Magill, Rob Naczi, Benjamin Normark, ***Rich Rabeler***, Melissa Tulig, Kim Watson, Christiane Weirauch, ***Katja Seltmann (Project Manager)***, Collaborating Institutions, *Digitizers!*

### *Funding & Support*

NSF-ADBC

iDigBio

Barbara Thiers (NYBG Herbarium)

John Pickering (Discover Life)



# Tri-trophic Digitization

## *Thematic Collections Network*



Photo: S. Bauer



Illustration: W. J. Hooker



Photo: C. A. Johnson



Herbaria (14 + 3)  
Insect Collections (18)

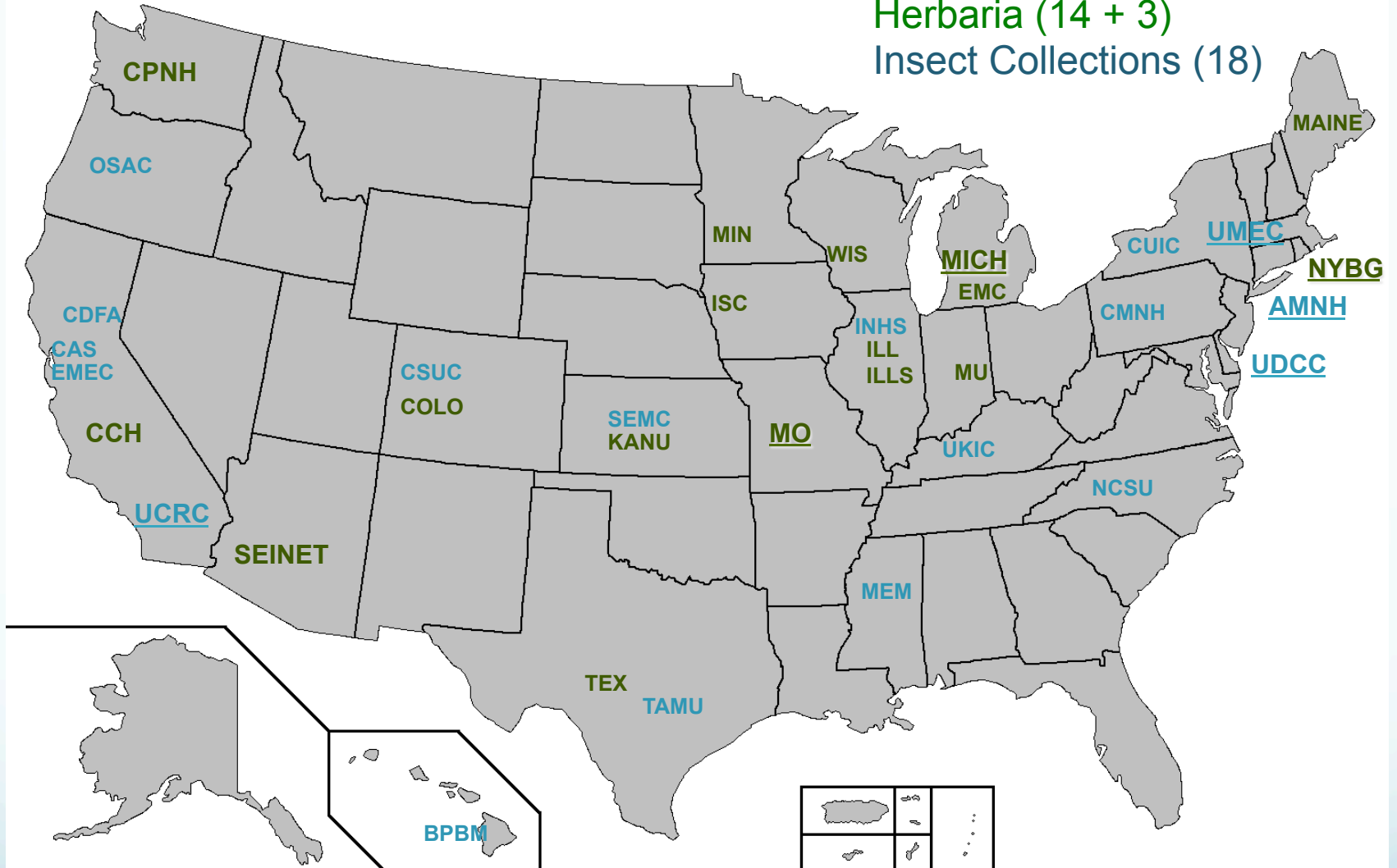


Image: Rich Rabeler

## Tri-trophic Digitization *Thematic Collections Network*

# Tri-trophic Digitization Progress

## *Thematic Collections Network*

- **Specimens Digitized**
  - Insects: 1,211,000 (cps = \$1.33)
  - Plants: 63,513 (complete); 424,747 (“skeletal”) (cps = .47)
- **Specimens Imaged**
  - Insects: ~10K
  - Plants: 1,019,764
- **Localities Georeferenced**
  - Insects: 54%
  - Plants: 16%
- **Submitted to iDigBio Aggregator**
  - Insects: 785,320
  - Plants: 281,127

## From Museum Specimen Database to Ecological Statement:

*Climate change has an effect on the timing of bug “emergence”*



Photo: Prud'homme et al. 2011



Photo: UFAL Campus – Maceió Michal Hoskovec



# From Museum Specimen Database to Ecological Statement:

## *Arthropod Easy Capture Database*

**PBI: Report Mode**

USI Prefix: AMNH\_IZC

[New sp.](#)  
[New sp. \[USI\]](#)

[Described sp.](#)  
[Described sp. \[USI\]](#)

[Coordinates](#)  
[Hosts](#)

[Labels \[by Host\]](#)  
[Labels \[by Locality\]](#)




[Queries](#)

[Log Out](#)

### Queries

<b>Family</b>	<input type="text" value="--Choose--"/>	<b>Subfamily</b>	<input type="text" value="--Choose--"/>	<b>Tribe</b>	<input type="text" value="--Choose--"/>		
<b>Genus</b>	<input type="text" value="--Choose--"/>	<b>Species</b>	<input type="text" value="--Choose--"/>	<b>Type</b>	<input type="text" value="--Choose--"/>	<b>Sex</b>	<input type="text" value="--Choose--"/>
<b>Country</b>	<input type="text" value="--Choose--"/>	<b>State/Prov.</b>	<input type="text" value="--Choose--"/>	<b>Subdivision</b>	<input type="text" value="--Choose--"/>		
<b>Collector</b>	<input type="text"/>	<b>Trip Code</b>	<input type="text"/>	<b>Depository</b>	<input type="text" value="--Choose--"/>		
<b>Det. By</b>	<input type="text" value="--Choose--"/>	<input type="text"/>		<b>Det. Year</b>	<input type="text"/>	<b>History</b>	<input type="text" value="--Choose--"/>
<b>Macro Habitat</b>	<input type="text" value="--Choose--"/>	<b>Micro Habitat</b>	<input type="text" value="--Choose--"/>	<b>Additional Info</b>	<input type="text" value="--Choose--"/>	<b>Specimen Notes</b>	<input type="text"/>
<b>Host Family</b>	<input type="text" value="--Choose--"/>	<b>Host Genus</b>	<input type="text" value="--Choose--"/>	<b>Host Species</b>	<input type="text" value="--Choose--"/>	<b>Host Subspecies</b>	<input type="text" value="--Choose--"/>
<b>Sort by:</b>	<input type="text" value="USI"/>	<input type="button" value="Search"/>		<input type="button" value="Count"/>		<input type="button" value="Download"/>	
							<input type="button" value="Clear"/>

---

# From Museum Specimen Database to Ecological Statement:

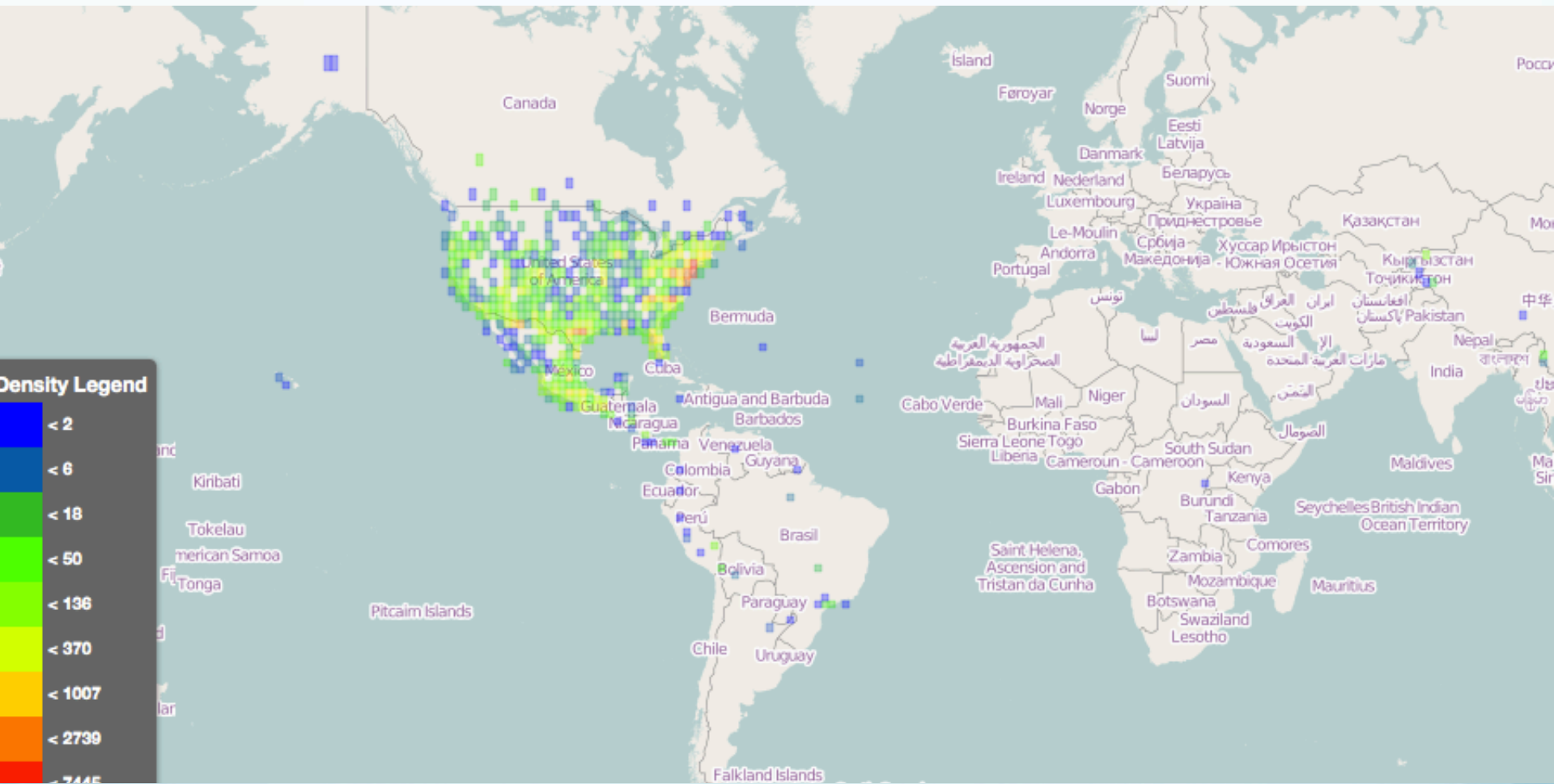
## *iDigBio Aggregator*

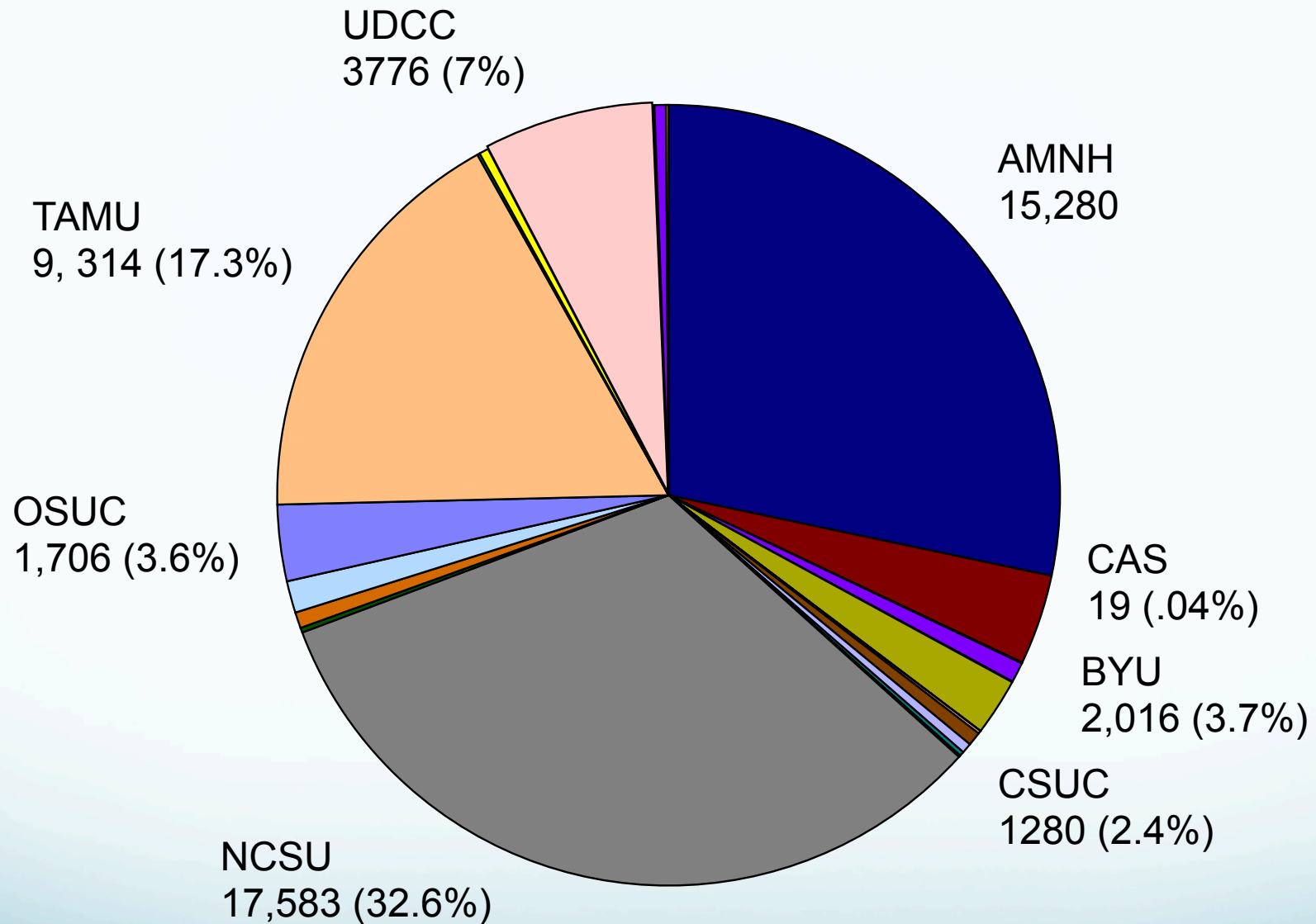
The screenshot displays the iDigBio Aggregator search interface. At the top, the iDigBio logo (Integrated Digitized Biocollections) is on the left, and navigation links for 'About iDigBio', 'Research', 'Technical Information', and 'Education' are on the right. Below the logo, a green navigation bar contains links for 'iDigBio Home', 'Portal Home', 'Search Records', 'Tutorial', 'Publishers', 'Research Tools', and 'Feedback'. The 'Search Records' section is active, showing a search query of 'Match all. Sort by Scientific' and a 'Clear' button. The results indicate 24,266,536 records with an approximate download time of 4 hours, 43 minutes, and 6 seconds. A 'Download' button is available for email-based retrieval. A map visualization shows a global distribution of records with a color-coded legend. The 'Advanced Search' section includes fields for 'Family' (dwc:family), 'Scientific Name' (dwc:scientificName), 'Genus' (dwc:genus), and 'Country' (dwc:country), each with 'Present' and 'Missing' radio buttons. A 'State/Province' field (dwc:stateProvince) is also present. At the bottom, there are controls for 'Add a field' (please select), 'Sort by' (Scientific Name), and 'Direction' (Ascending), along with 'Search', 'Clear', and 'Reset' buttons, and a 'Tips & Hints' link. A vertical 'FEEDBACK' button is located on the right side of the page.

# From Museum Specimen Database to Ecological Statement:

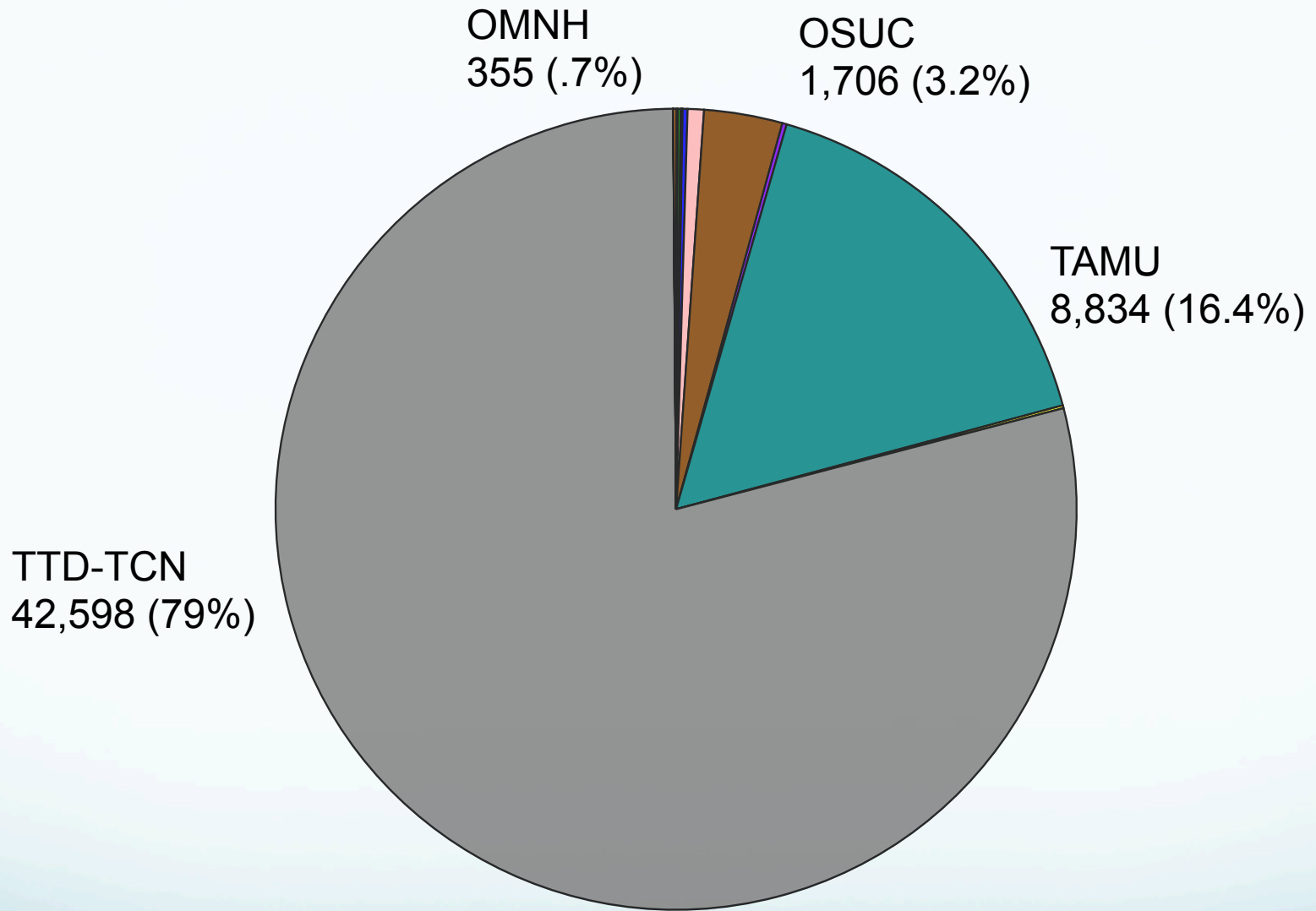
*Climate change has an effect on the timing of bug “emergence”*

- Downloaded Membracidae from iDigBio Portal (N = 53,959, no duplicate GUID/UUID)





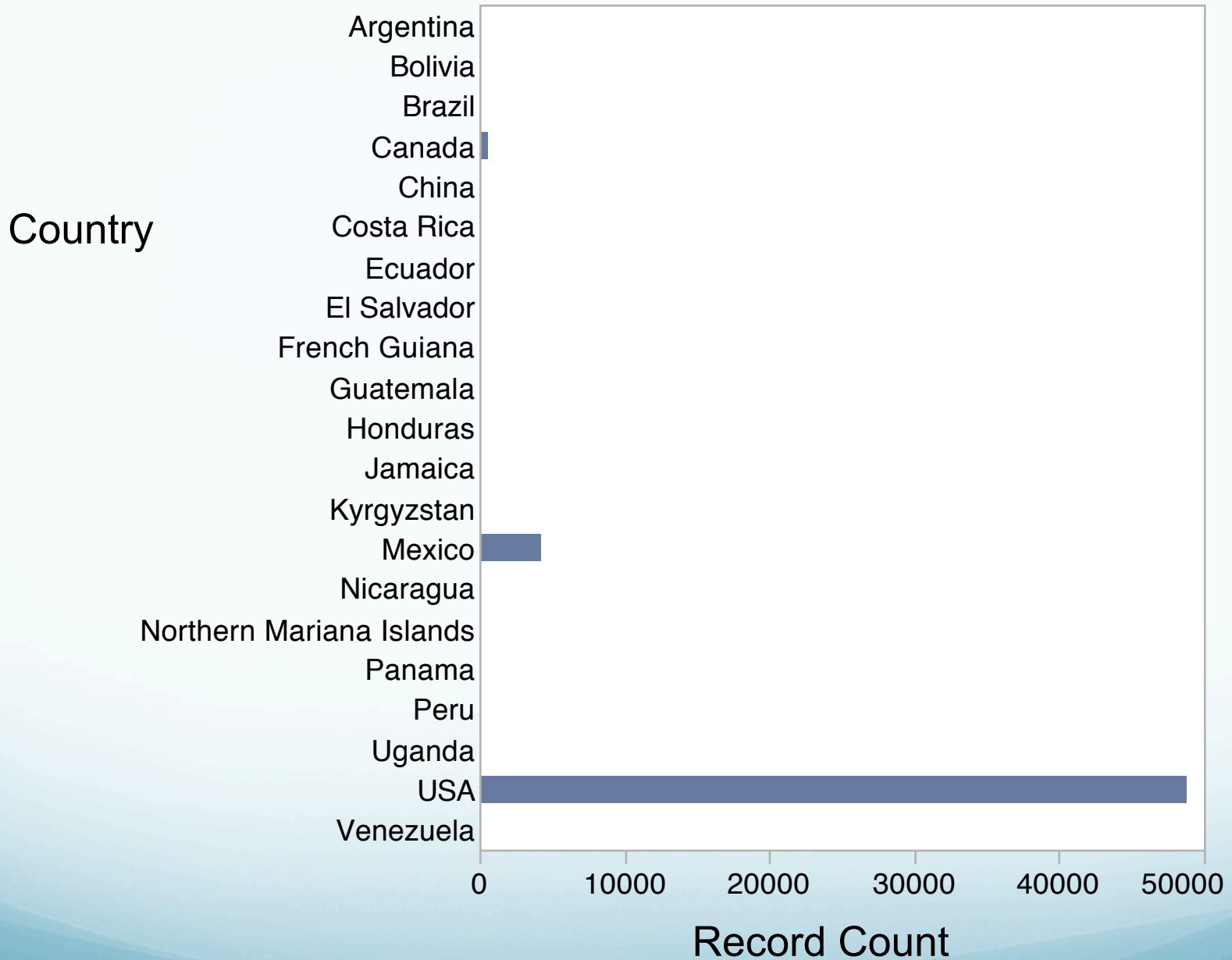
Membracids Data Records Submitted to iDigBio Aggregator By Institution



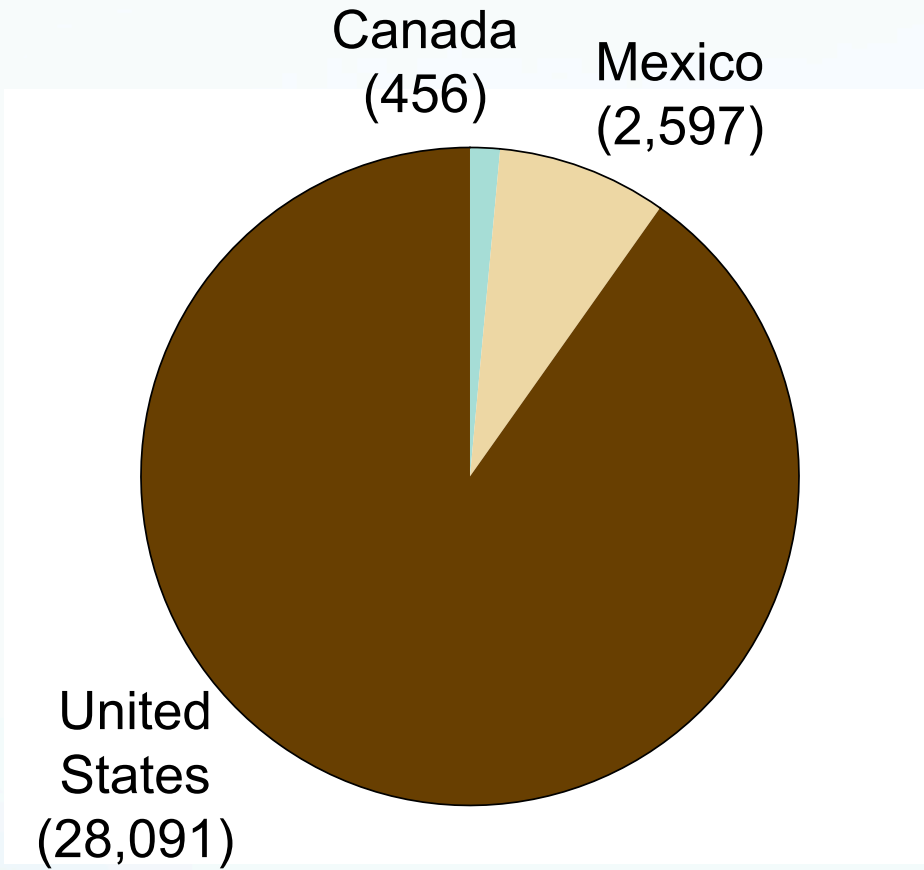
Membracidae Databased & Submitted to iDigBio By “Project” Name

## Data “Problems”

- No locality information
- Localities not georeferenced
- Skeletal taxonomy records - no information below subfamily or tribe level
- Duplicate information in different fields
- Some outdated taxonomy
- Little info about the host
- Outdated/error in host taxonomy also







“Complete” North American Records of Membracidae

Subfamily: Tribe: Genus species

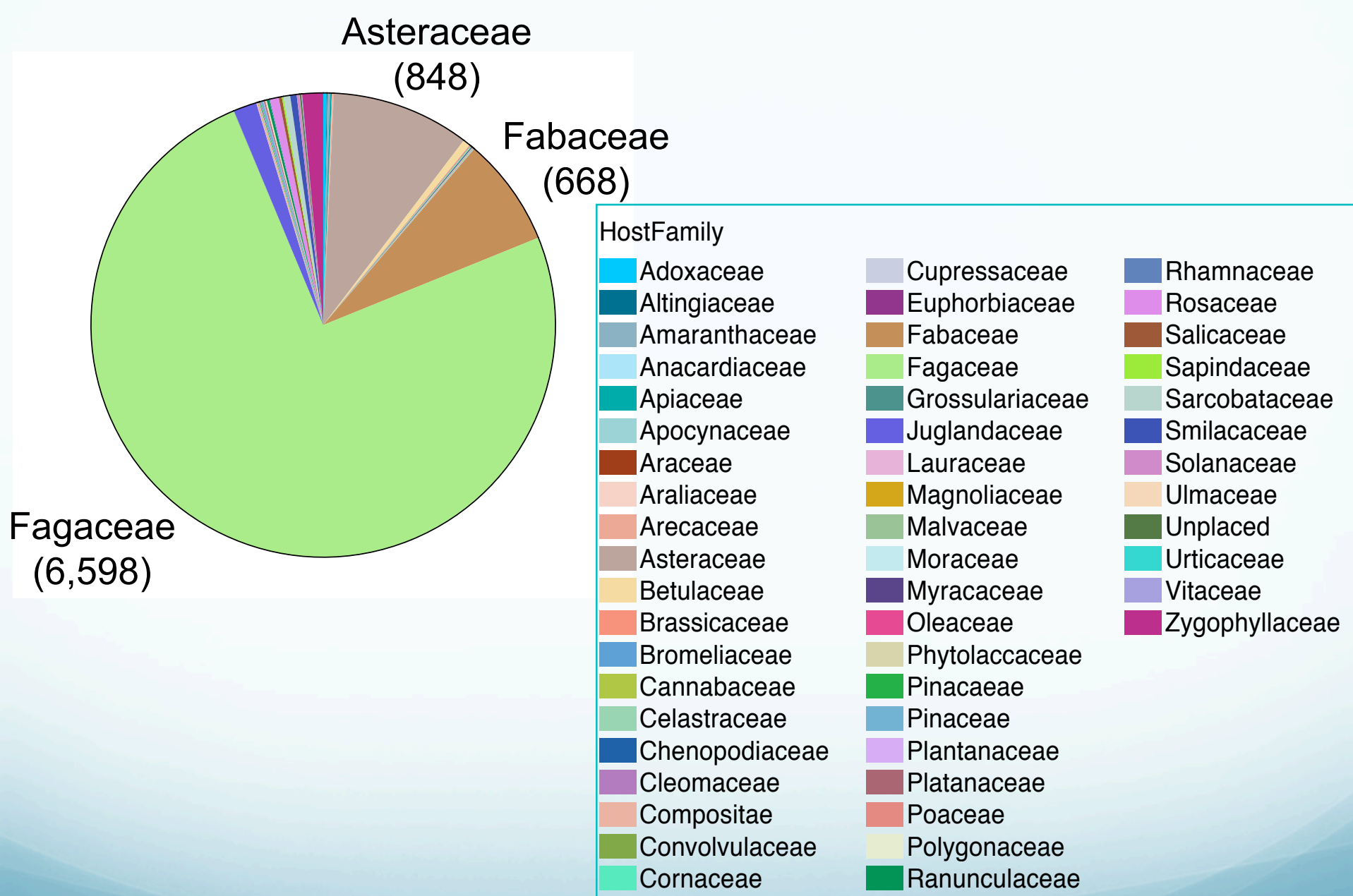
Number of Host Plants per Membracid Species



*Platycotis vittata* (37)  
Photo: E. Gofreed



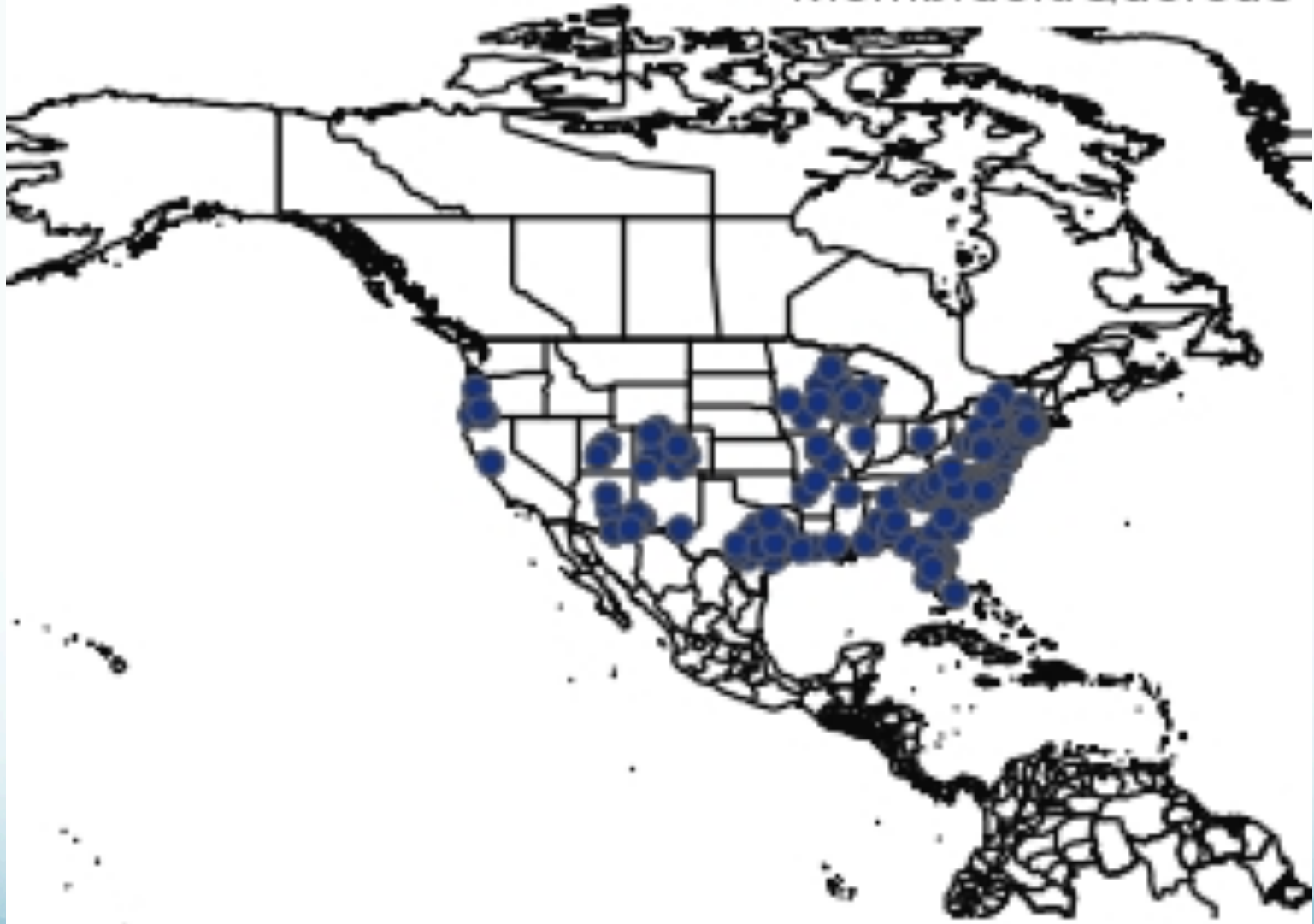
*Entylia carinata* (43)  
Photo: metriopectera



Host Families in U.S. Membracid Records



● MembracidQuercus

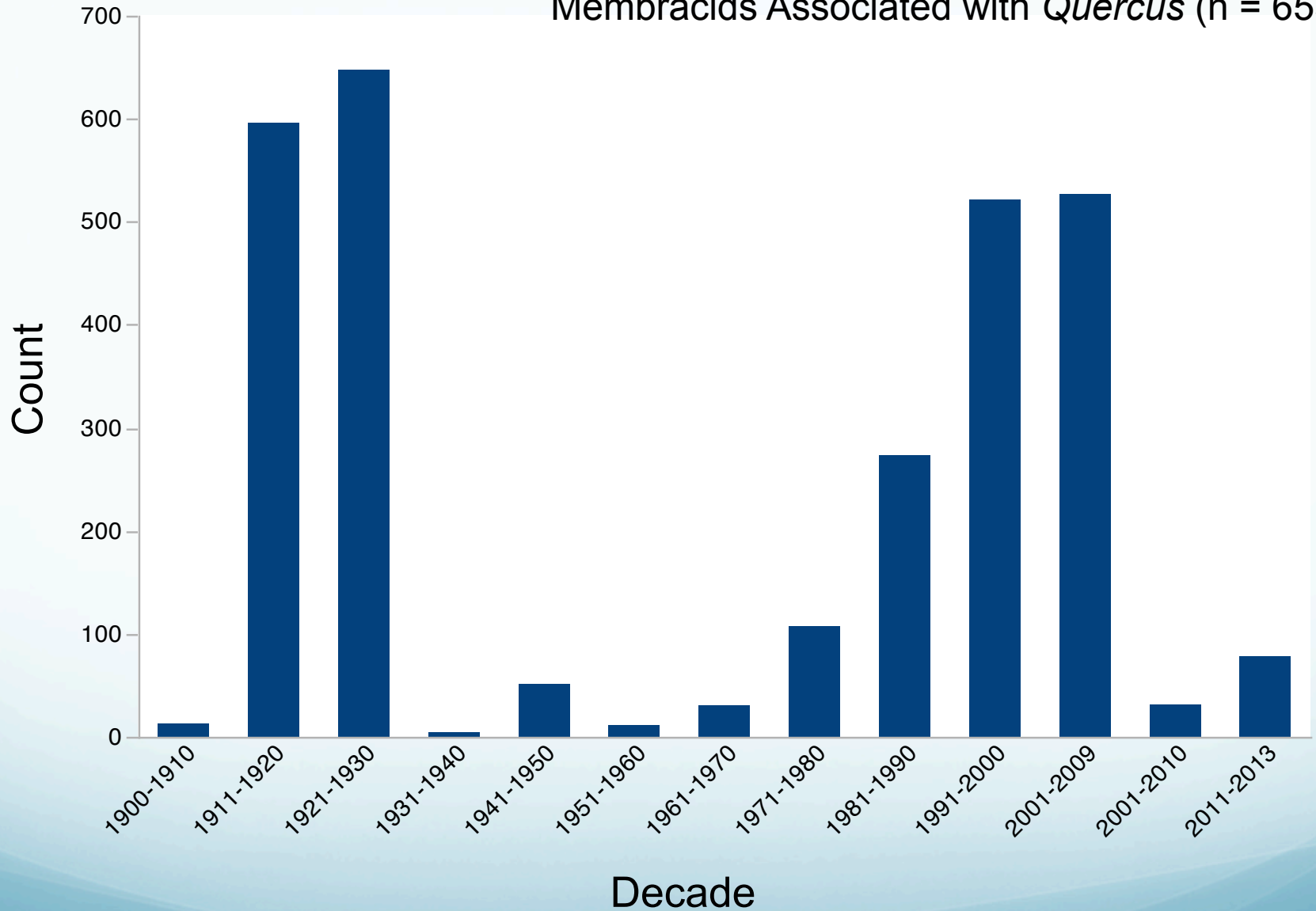


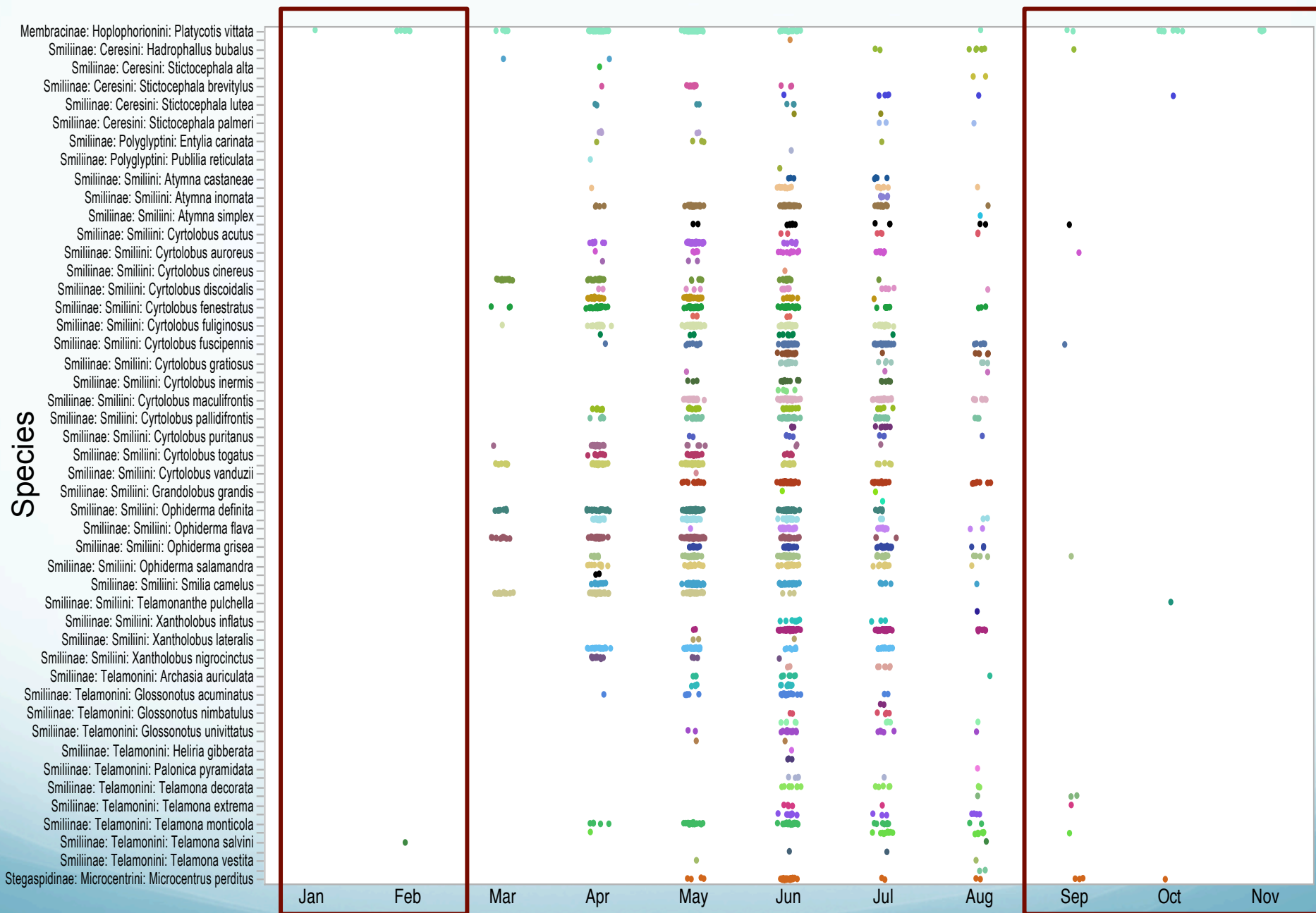
0 2400 4800 7200 km



© 2009 by Paul Huber

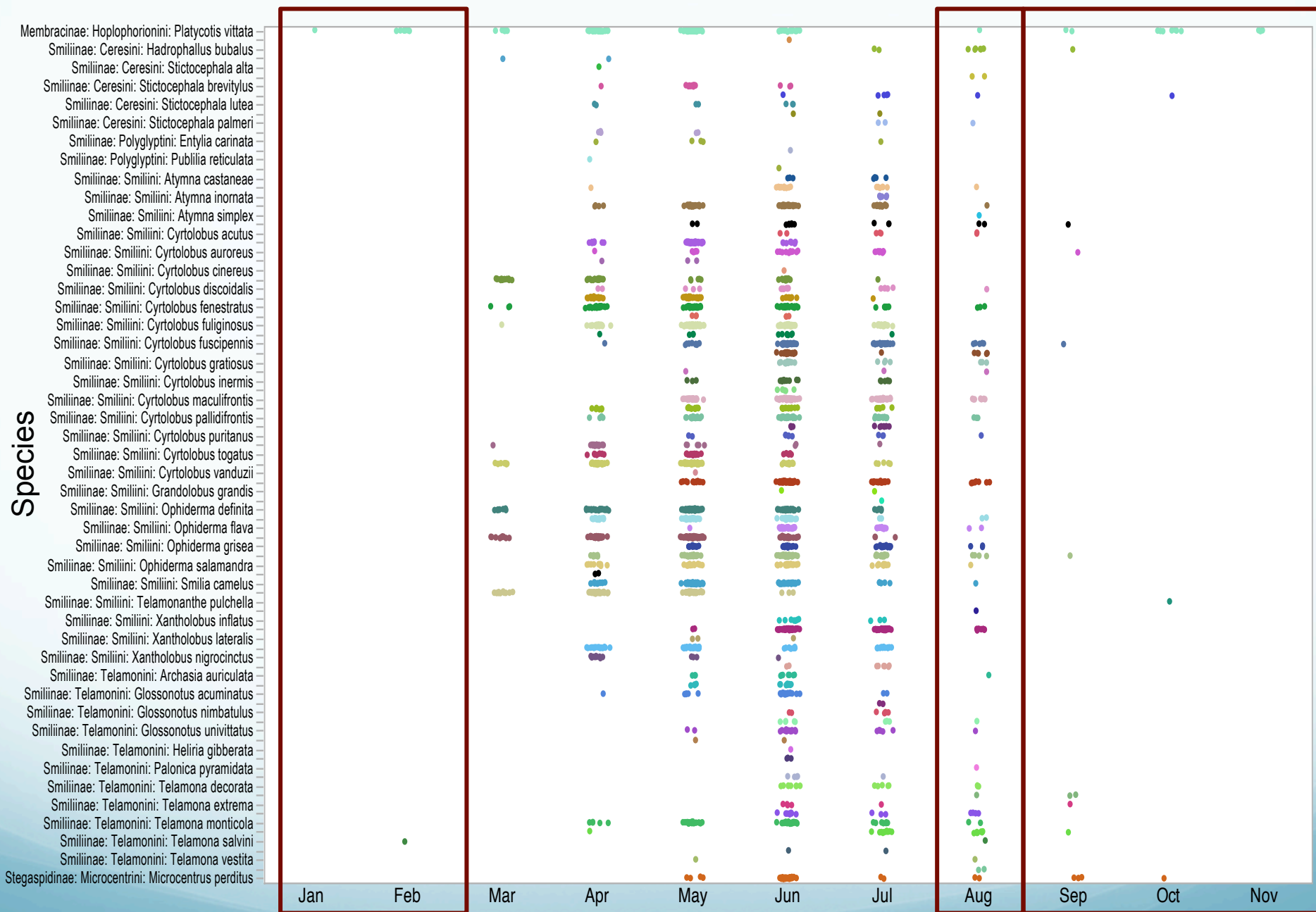
# Unique U.S. Collecting Events for Membracids Associated with *Quercus* (n = 6557)



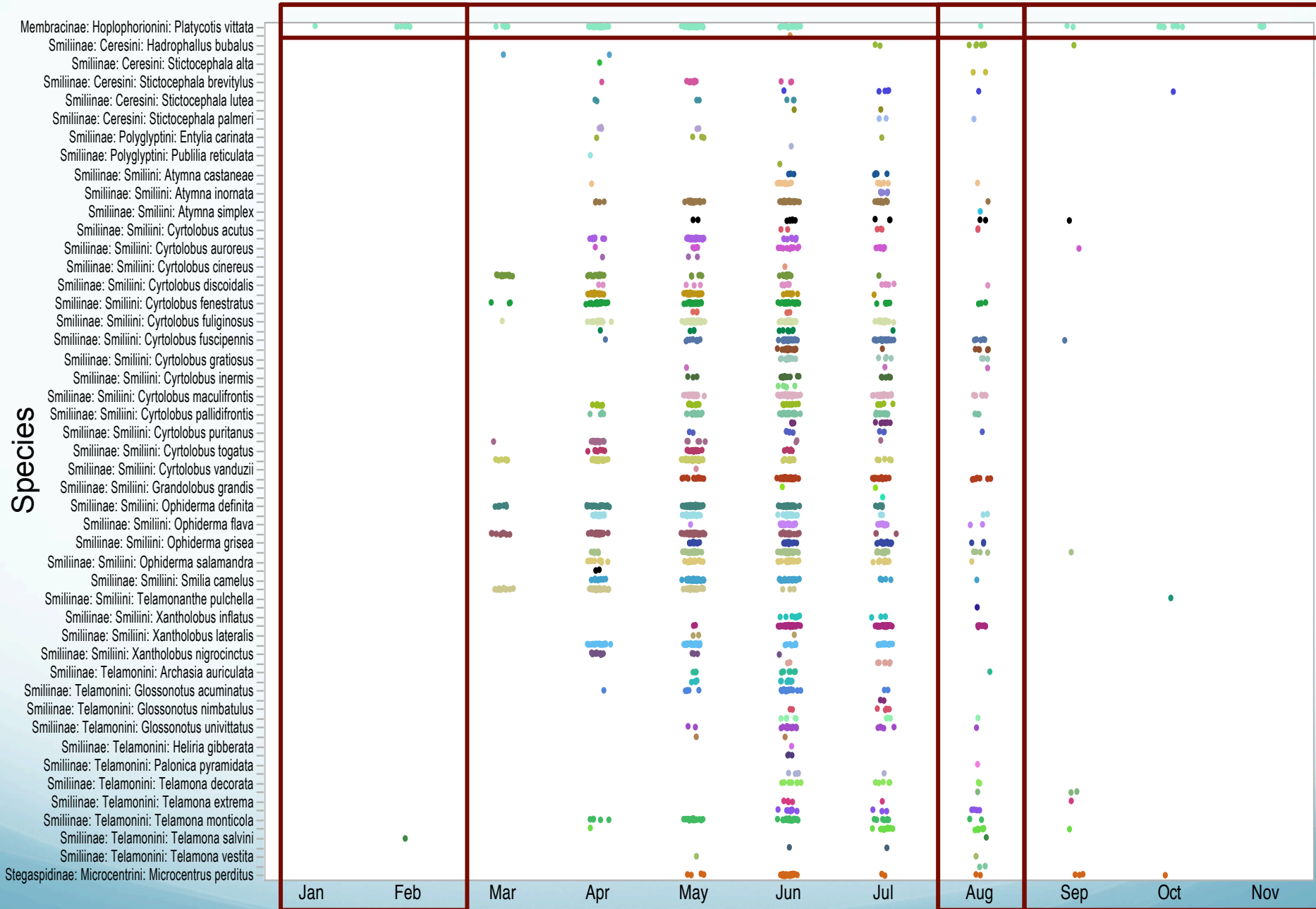


Month Species Collected (N = 2889; color = species)





Month Species Collected (N = 2889; color = species)

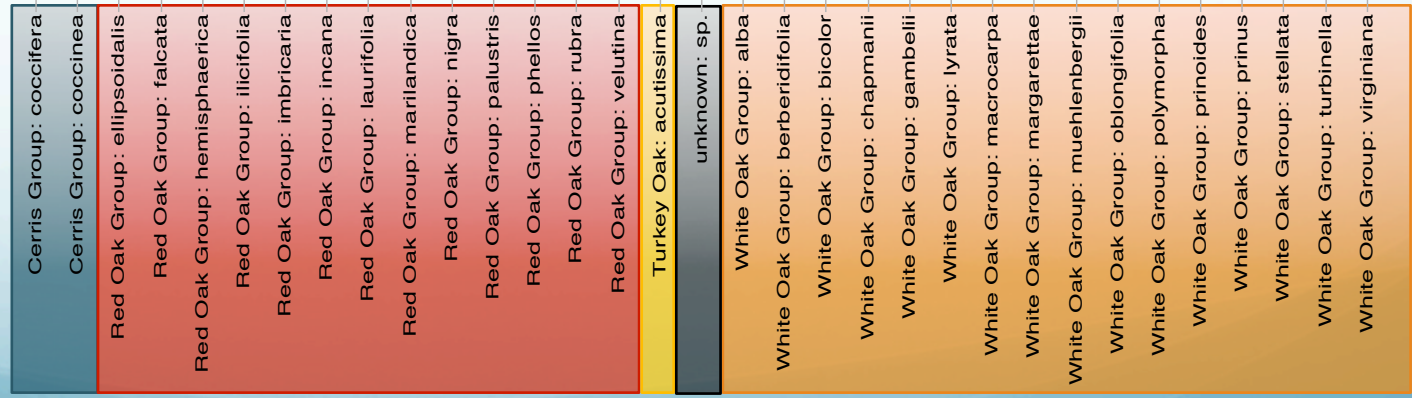
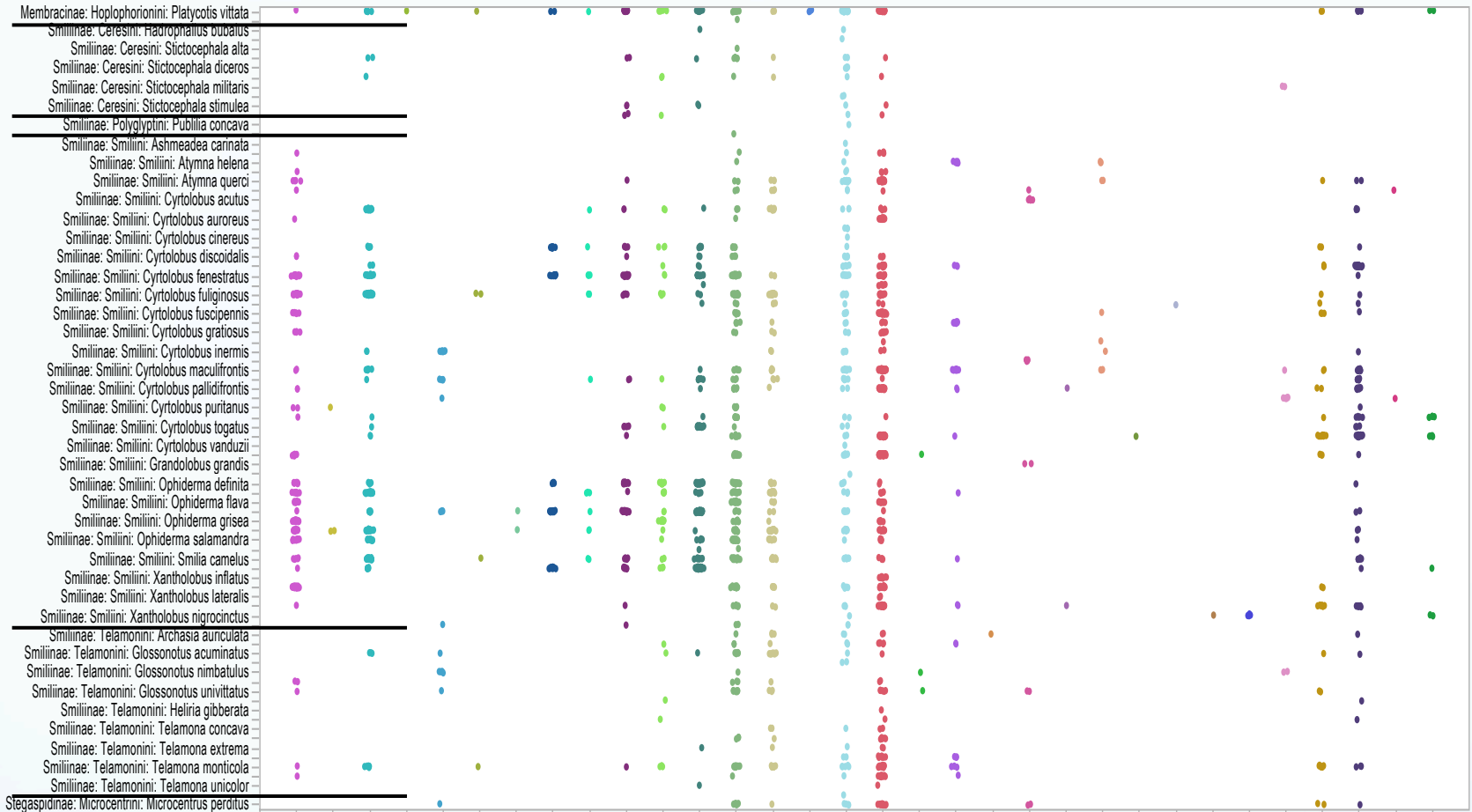


Month Species Collected (N = 2889; color = species)

# Quercus Host Group & Species

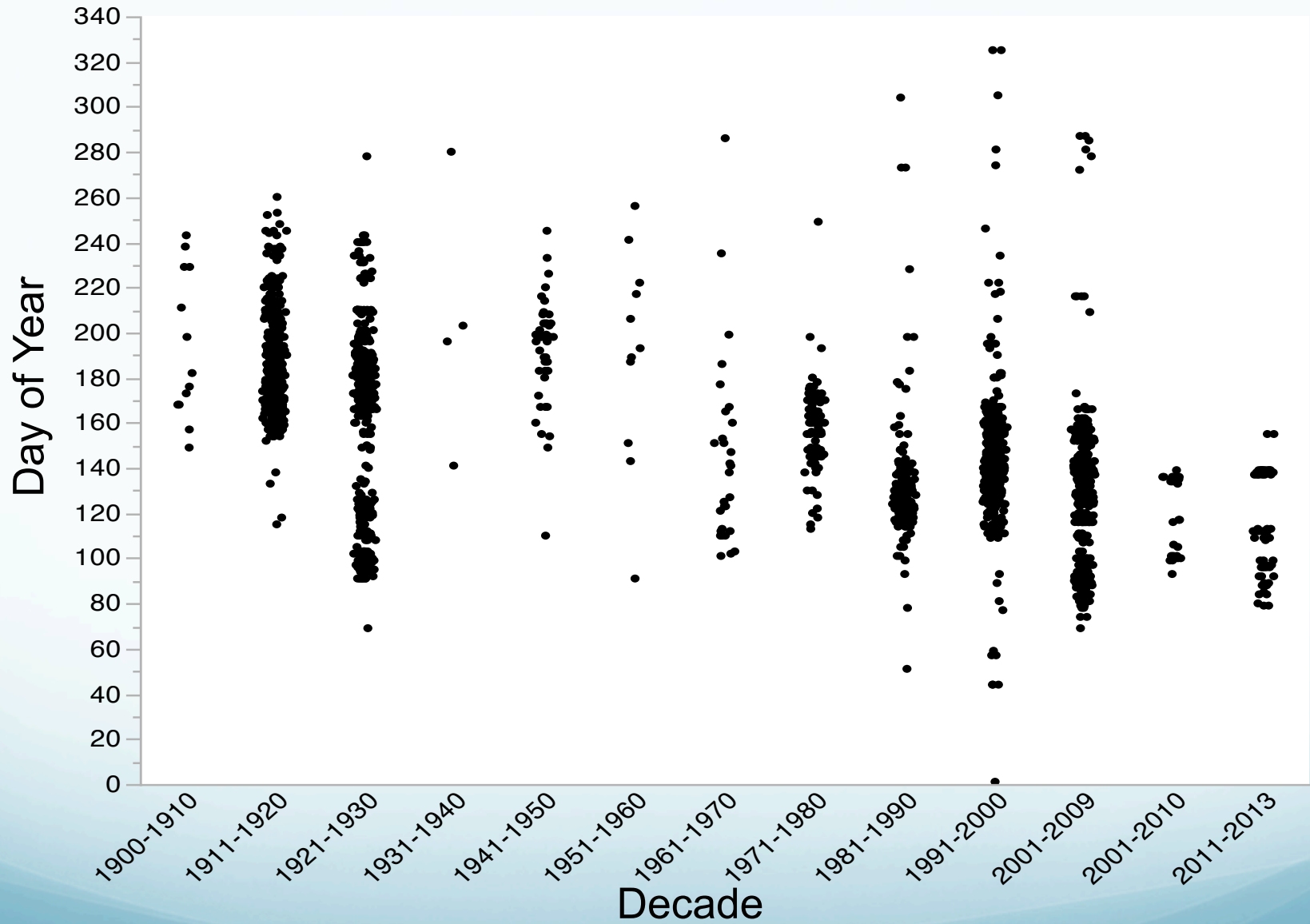
(n = 2679; color = host species)

# Membracid Subfamily, Tribe & Species



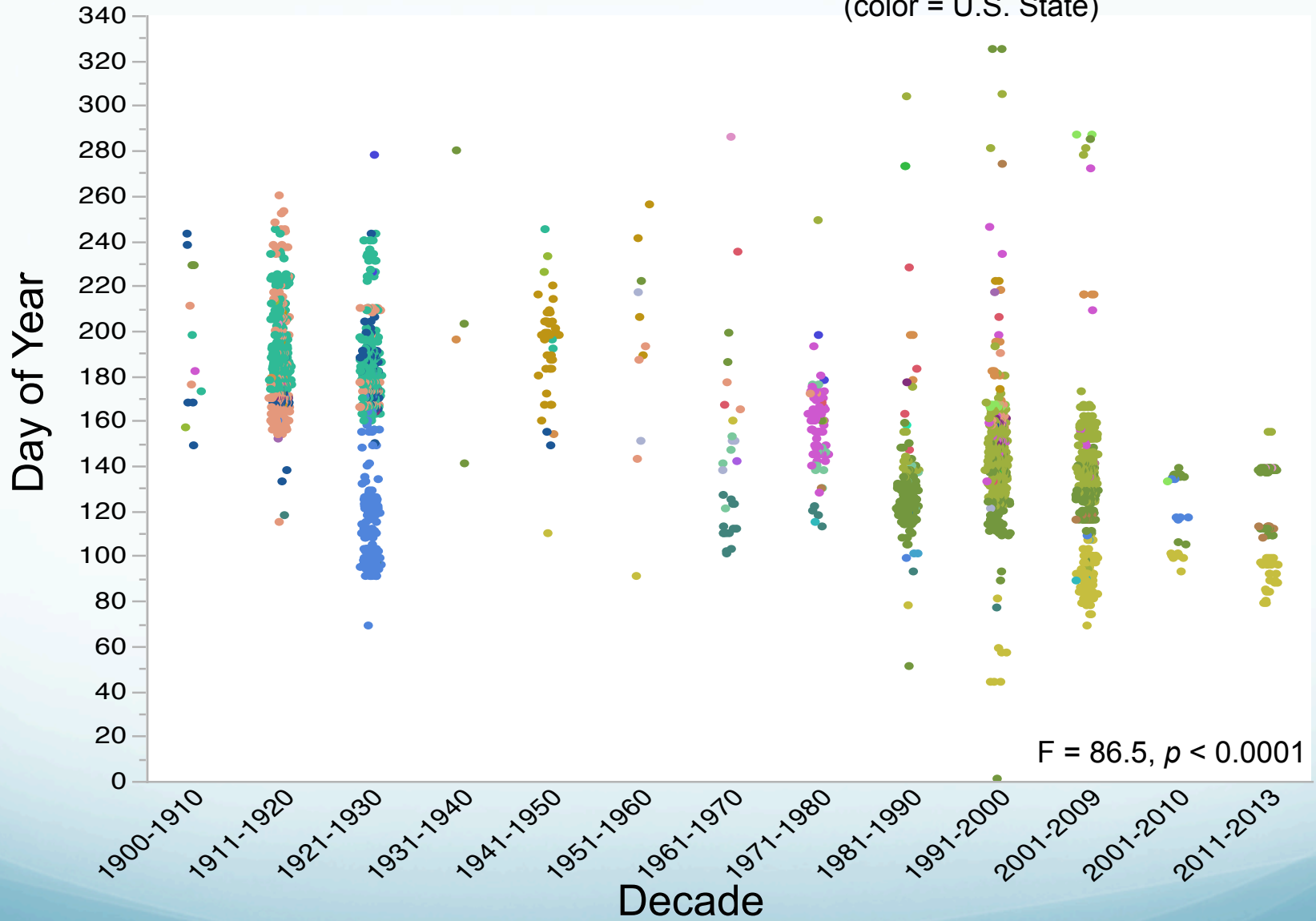
Cerris Group      Red Oak Group      Unknown sp.      White Oak Group

# Day of Year Specimens Collected



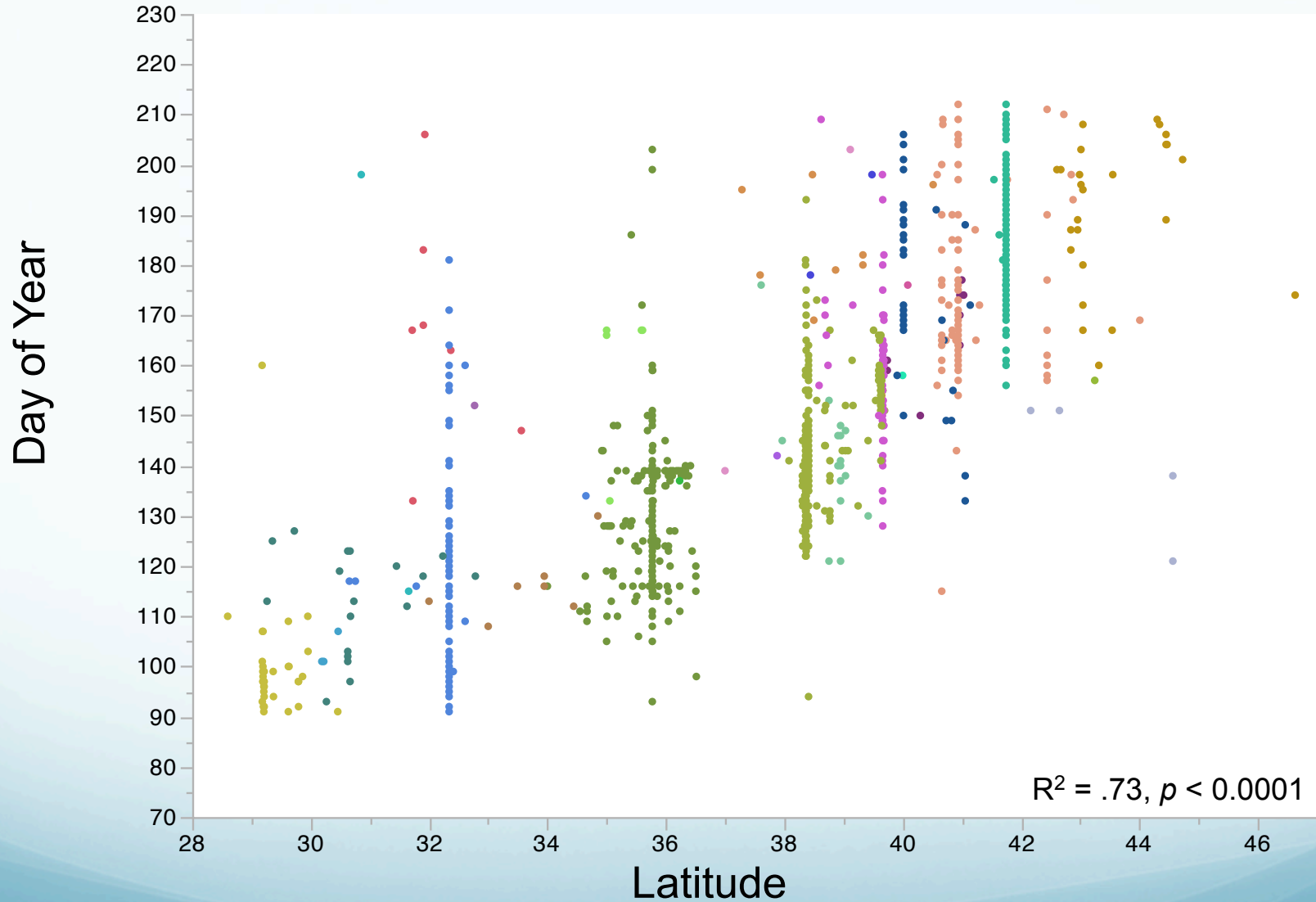
# Day of Year Specimens Collected

(color = U.S. State)

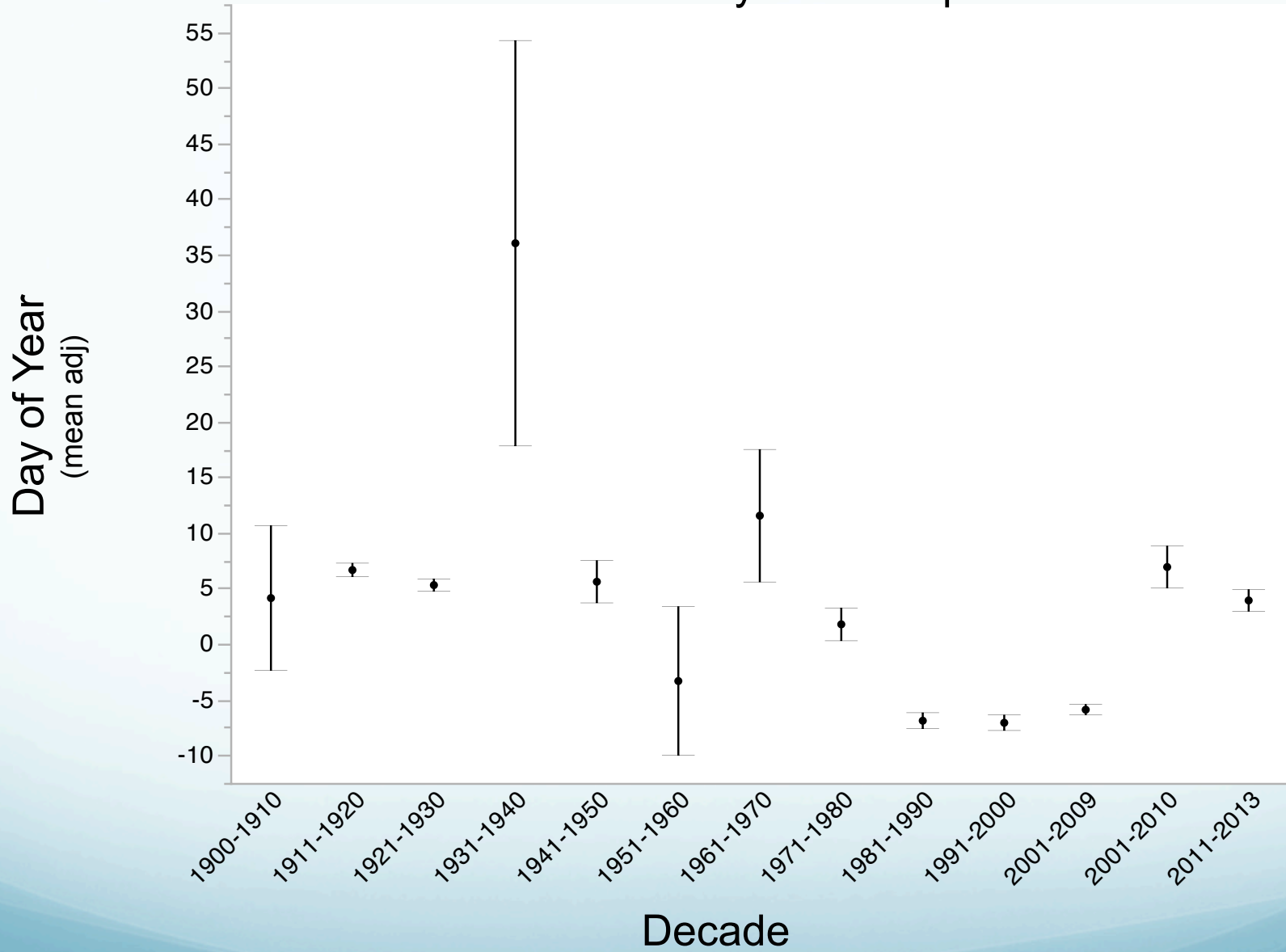


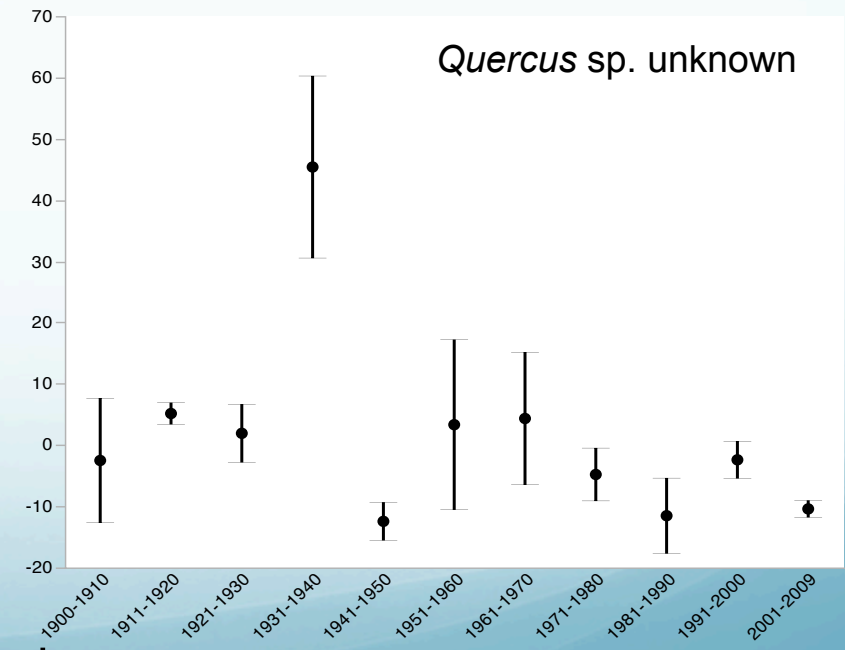
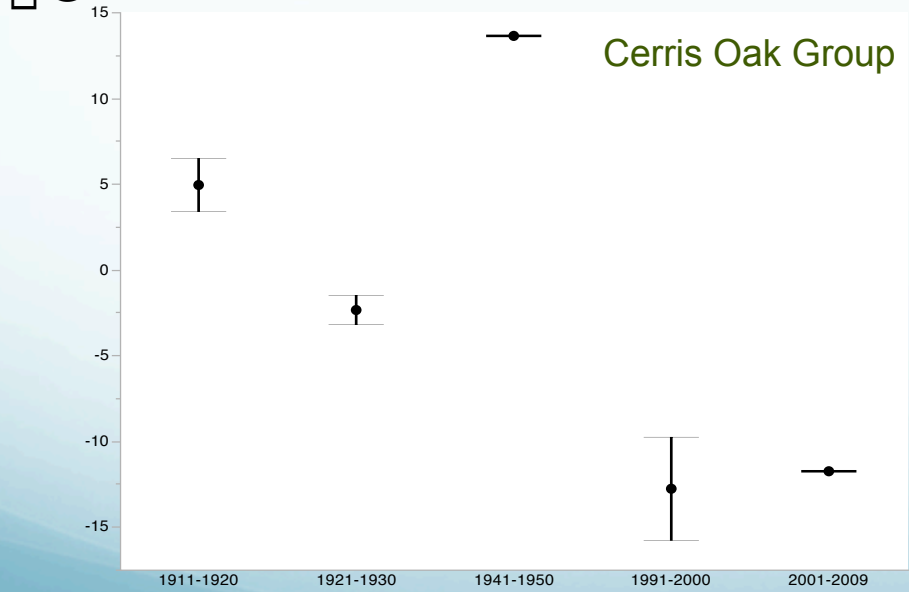
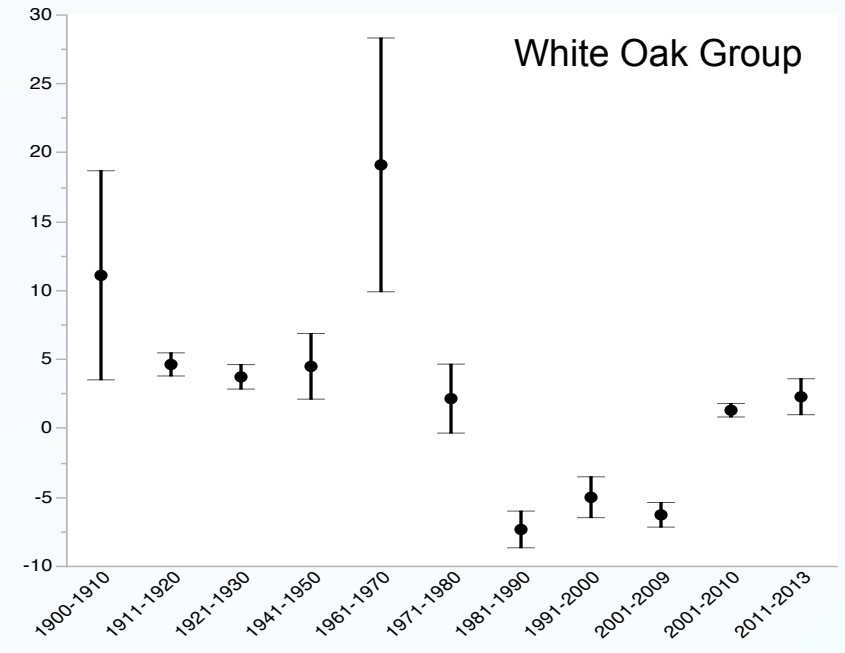
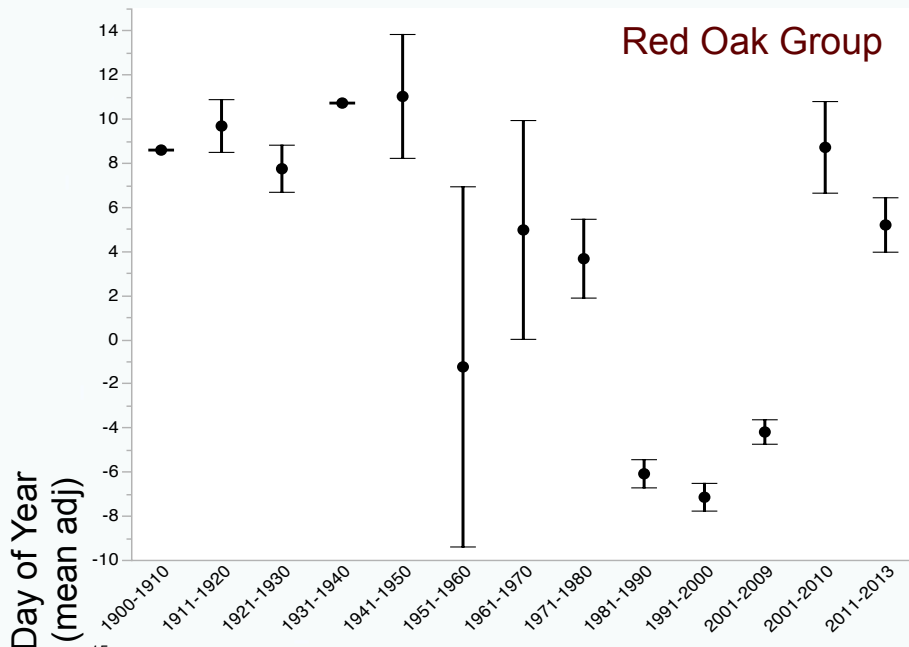
# Day of Year Specimens Collected

(color = U.S. State)



# Day of Year Specimens Collected

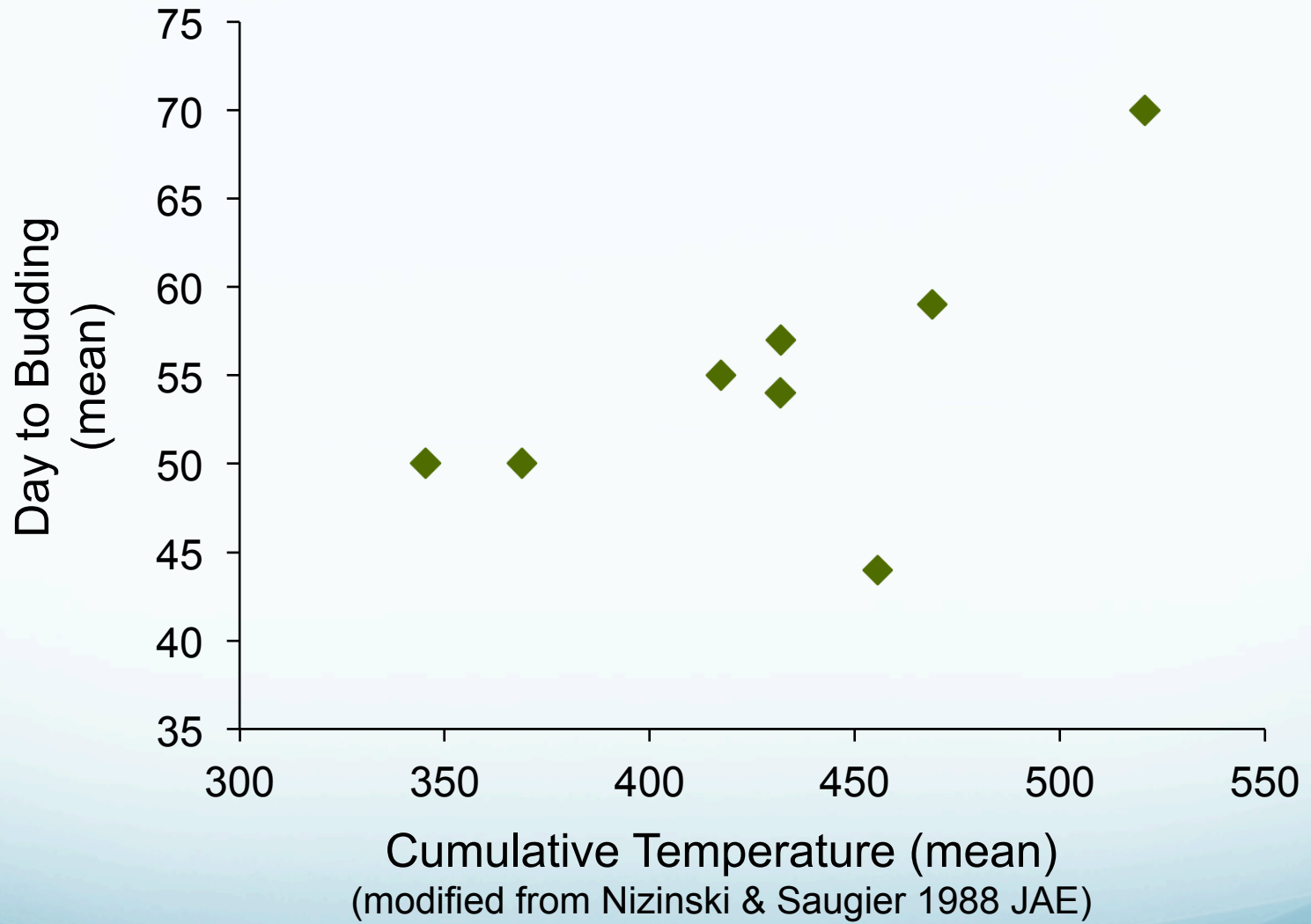




Decade



# Relationship between *Quercus* bud date & cumulative annual temperature



## *Conclusion:*

- *Historical records can provide valuable information about ecological trends associated with specimens*
  - *Seems to be an association among warming trend, oak budding & treehopper emergence*
- *BUT more data records & more complete data records are needed to fill in gaps*
- *Need data from large AND small institutions/collections*
- *Approach conclusions with caution*
- *Data label standards!*
- *Encourage deposition of some/all data from ecological research?*



Photo: Alex Wild

---

data

## Improving the use of information from museum specimens: Using Google Earth<sup>®</sup> to georeference Guiana Shield specimens in the US National Herbarium

Eduardo Garcia-Milagros and Vicki A. Funk

US National Herbarium, Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington DC, 20013-7012 USA

e-mail: [eduadogarmi@gmail.com](mailto:eduadogarmi@gmail.com); <http://botany.si.edu/bdg/index.html>

---

**Abstract.** Data found on labels of museum collections have been useful in a variety of biodiversity studies. However, the georeferenced data available are often hampered by poor interpretation of label information and as a result are not as accurate, and therefore useful, as they might be. We have used Google Earth<sup>®</sup> as a geographic information system to improve the georeferencing of the data. Its user interface allowed us to make use of all the label information and to represent the coordinates more accurately, thus producing a better quality and more reliable dataset to be used in our studies. The quality, defined as “fitness for use”, of the species-occurrence data generated, which is mostly affected by the values of accuracy and uncertainty associated to the coordinates, shows that uncertainty can be reduced. This method also allows us to show the power of examining georeferenced data from the stand point of ‘all collections from an expedition’ rather than ‘all collections from a single area.’ Type specimens housed at U.S. National Herbarium from the Guiana Shield were used in this work.

**Keywords.** museum collections, georeferencing, data quality, Google Earth, type specimen

---

### Introduction

The specimen collections housed in museums and

et al. 2000, Funk and Richardson 2002, Reddy and Davalos 2003). Although these studies suggest that collecting more data is necessary, the infor-



ORIGINAL  
ARTICLE

## Elevational gradient analyses and the use of historical museum specimens: a cautionary tale

Rebecca J. Rowe\*

---

*Committee on Evolutionary Biology, University of Chicago, Chicago, IL, USA and Division of Mammals, The Field Museum, Chicago, IL, USA*

### ABSTRACT

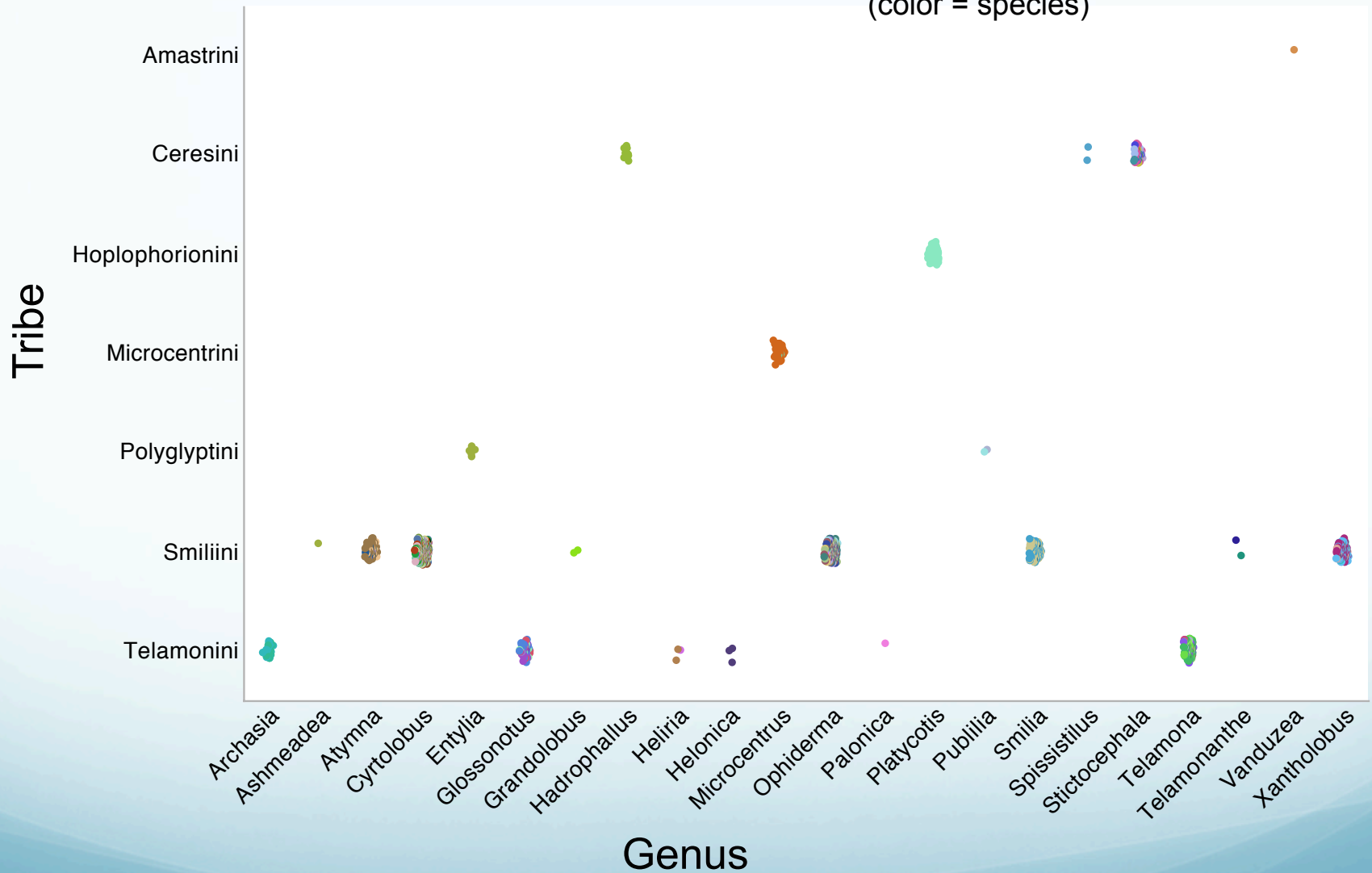
**Aim** The value of biodiversity informatics rests upon the capacity to assess data quality. Yet as these methods have developed, investigating the quality of the underlying specimen data has largely been neglected. Using an exceptionally large, densely sampled specimen data set for non-flying small mammals of Utah, I evaluate measures of uncertainty associated with georeferenced localities and illustrate the implications of uncritical incorporation of data in the analysis of patterns of species richness and species range overlap along elevational gradients.

**Location** Utah, USA, with emphasis on the Uinta Mountains.

**Methods** Employing georeferenced specimen data from the Mammal Networked Information System (MaNIS), I converted estimates of areal uncertainty into elevational uncertainty using a geographic information system (GIS). Examining patterns in both areal and elevational uncertainty measures, I

# Unique U.S. Collecting Events for Membracids Associated with *Quercus*

(color = species)



# Unique U.S. Collecting Events for Membracids Associated with *Quercus*

(color = species)

