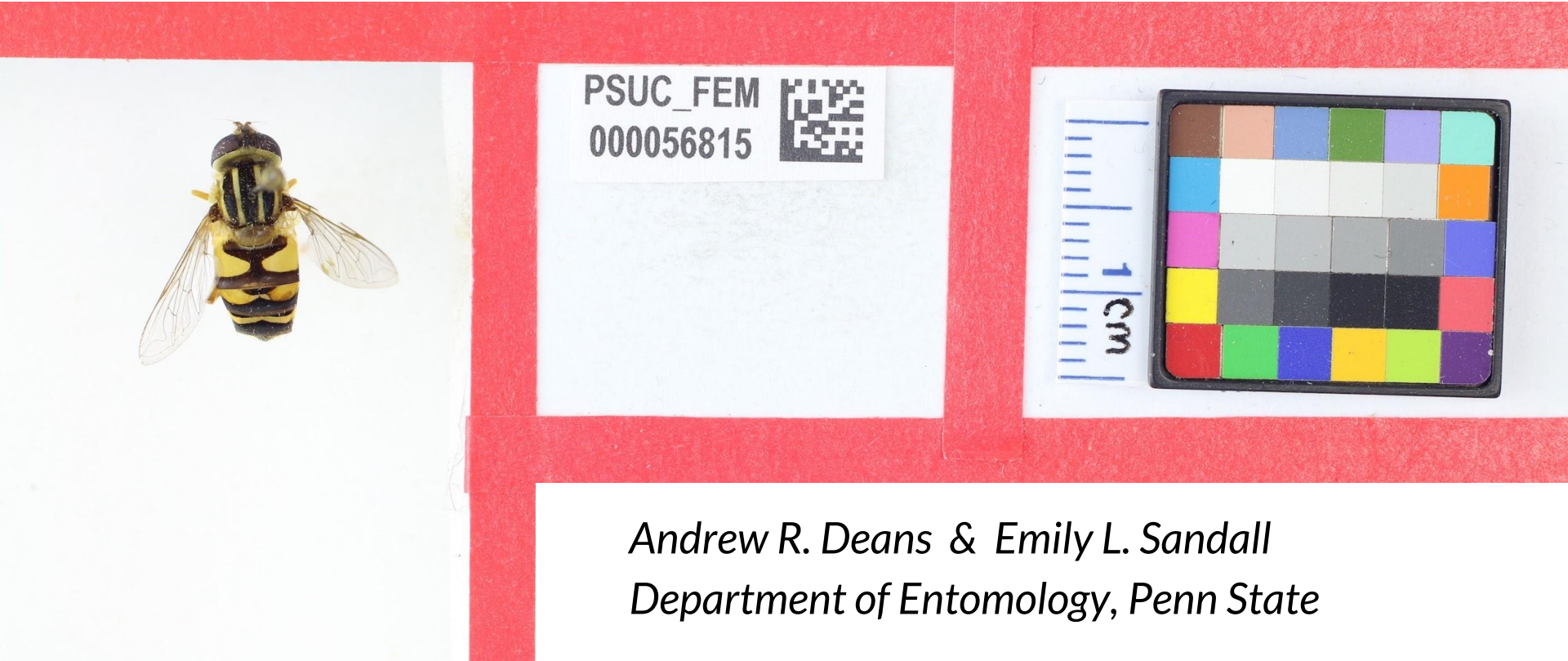
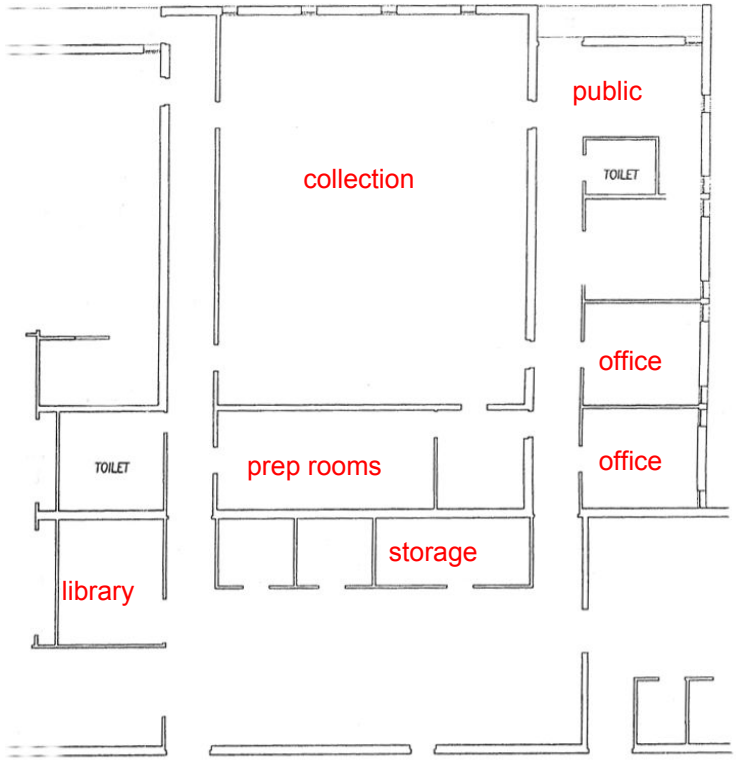
Digitizing the Frost Entomological Museum: Lessons learned and given



Andrew R. Deans & Emily L. Sandall
Department of Entomology, Penn State



~1,000,000(?) specimens

Invert · E · Base

InvertEBase Data Portal Homepage

- Search Collections
- Map Search
- Dynamic Checklist
- Image Library
- Search Images

Welcome Andy!

- My Profile
- Logout
- Sitemap

Welcome to InvertEBase: *reaching back to see the future: species-rich invertebrate faunas document causes and consequences of biodiversity shifts*

The rapid biodiversity change in North America has significant effects on essential ecosystem services, from impact on soil health and nutrient cycling, to agriculture, forestry and water quality. Effective monitoring of changes in biodiversity requires easy electronic access to historical specimen baseline information for temporal and regional species diversity comparisons, which can facilitate informed land management decisions. Vast amounts of specimen data are housed within the nation's natural history collections, but most of these data are not yet readily accessible as digital resources. The TCN "**InvertEBase**" is a 4-year collaborative effort to digitize specimen records from ten arthropod and mollusk collections housed at six major US museums in six states, three of them ranking among the top 10 collections in the world. They include the Field Museum of Natural History (Chicago, Il), Cleveland Museum of Natural History (CMNH, OH), Auburn University Museum of Natural History (AUMNH, AL), University of Michigan Museum of Zoology (MI), Delaware Museum of Natural History (DMNH, DE), and Frost Entomological Museum at Pennsylvania State University (PA). InvertEBase will digitize de novo ~ 2.4 million georeferenced specimen records as well as integrate and mobilize data for 3.9 million terrestrial and aquatic arthropod and invertebrate specimens with special focus on the United States fauna. InvertEBase will greatly expand the taxon and geographic coverage of existing TCNs, and include the phylum Mollusca for the first time; DMNH, AUMNH, and CMNH will serve all of their invertebrate specimen data online for the first time. This project will significantly automate specimen data capture by utilizing optical character and voice-recognition technologies. The digitized data from this project will be immediately deployed for habitat-based distribution modeling and analyses.

This project made possible by the National Science Foundation awards EF 14-02667 to P. Sierwald and R. Bieler (Field Museum), EF 14-00993 to A. Deans (Penn State), EF 14-02697 to E. Shea (Delaware Museum of Natural History), EF 14- 04964 to D. O'Foighil (Museum of Zoology, UMichigan), EF 14-01176 to J. Bond (Auburn University), and EF 14-02785 to G. Svenson (Cleveland Museum of Natural History), and EF 14-01450 to J. Hanken (Harvard University).These awards are made as part of the National Resource for Digitization of Biological Collections through the Advancing Digitization of Biological Collections program and all data resulting from this award will be available through the national resource (iDigBio.org).

NSF EF 14-00993


NSF DBI-1349356

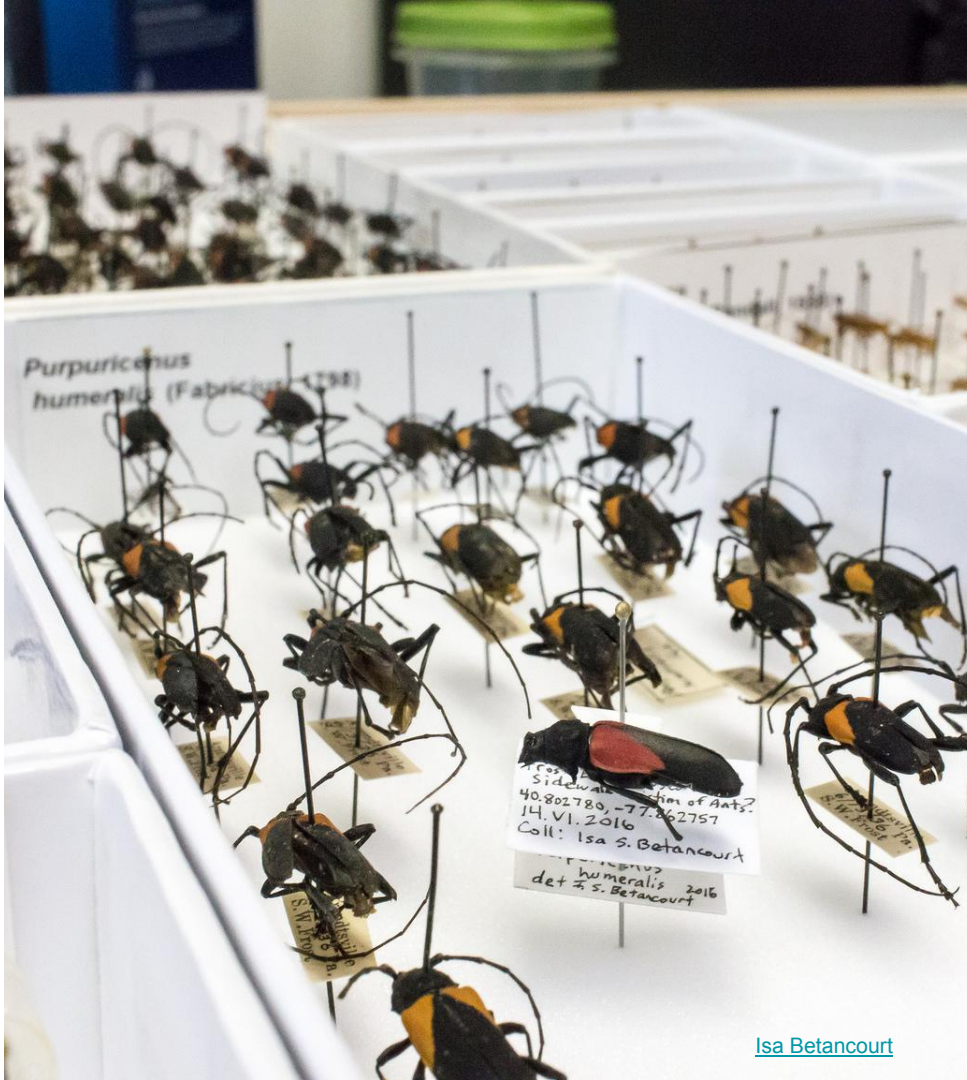
Lesson 1:

Digitization and collection improvement are best done *together*

Lesson 2:

Digitization of insects is *hard*

Troy University Herbarium



Isa Betancourt

Lesson 3:

Digitization of insects can be easy!

Pinned specimens => Darwin Core + spreadsheets

ENT432database - Google Sh

https://docs.google.com/spreadsheets/d/1aMhwl2u4APVix_NzjtoeCDOpymBeltOVhRhe8IyIhFs/edit#gid=0

ENT432database

File Edit View Insert Format Data Tools Add-ons Help Last edit was on January 9

Share

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	catalogNumber	institutionCode	basisOfRecord	countryCode	stateProvince	county	municipality	locality	decimalLatitude	decimalLongitude	geodeticDatum	samplingProtocol	eventDate	recordedBy	individualCount	lifeStage	identifiedB
2	ENT432_ARD_001	PSUC	PreservedSpecimen	US	Pennsylvania	Centre	University Park	loading dock area of Headhouse III	40.803	-77.862	WGS84	hand collected	2019-09-13	Andrew R. Deans	2	adult	Andrew R.
3	ENT432_ARD_002	PSUC	PreservedSpecimen	US	Pennsylvania	Centre	Rockspring	pond near Pasto Agri. Museum	40.701	-77.958	WGS84	D net	2019-08-24	Andrew R. Deans	1	larva	Andrew R.

Sheet1

Pinned specimens => Darwin Core + spreadsheets

ENT432database - Google Sh

https://docs.google.com/spreadsheets/d/1aMhwl2u4APVix_NzjtoeCDOpymBeltOVhRhe8IyIhFs/edit#gid=0

ENT432database

File Edit View Insert Format Data Tools Add-ons Help Last edit was on January 9

Share

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	catalogNumber	institutionCode	basisOfRecord	countryCode	stateProvince	county	municipality	locality	decimalLatitude	decimalLongitude	geodeticDatum	samplingProtocol	eventDate	recordedBy	individualCount	lifeStage	identifiedB
2	ENT432_ARD_001	PSUC	PreservedSpecimen	US	Pennsylvania	Centre	University Park	loading dock area of Headhouse III	40.803	-77.862	WGS84	hand collected	2019-09-13	Andrew R. Deans	2	adult	Andrew R.
3	ENT432_ARD_002	PSUC	PreservedSpecimen	US	Pennsylvania	Centre	Rockspring	pond near Pasto Agri. Museum	40.701	-77.958	WGS84	D net	2019-08-24	Andrew R. Deans	1	larva	Andrew R.

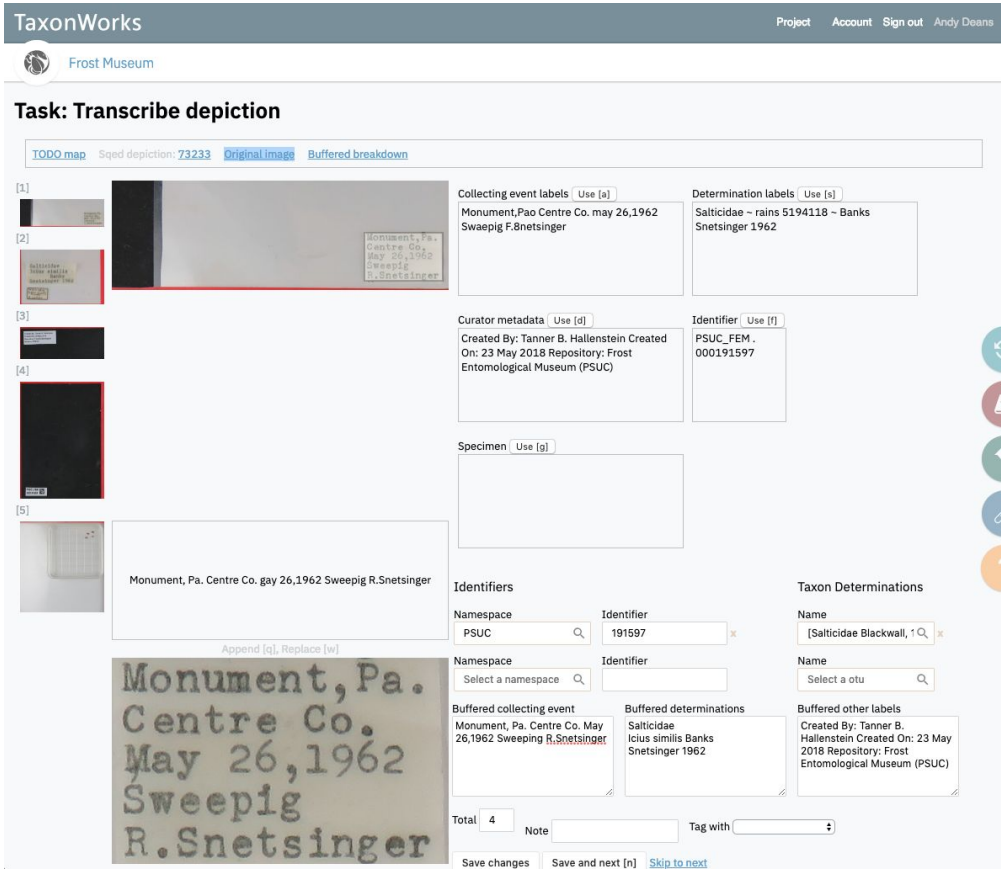
Sheet1

>127,000 occurrences @

SCAN

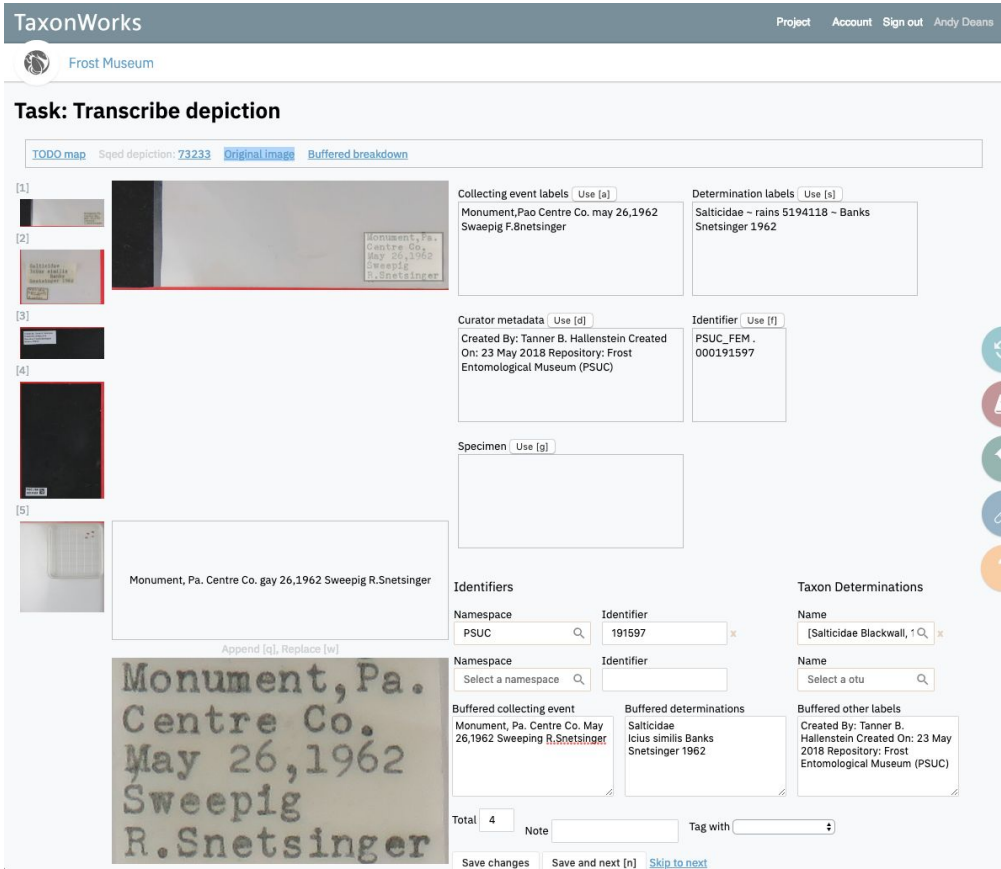
Symbiota Collections of Arthropods Network

Wet specimens => point and shoot + people power

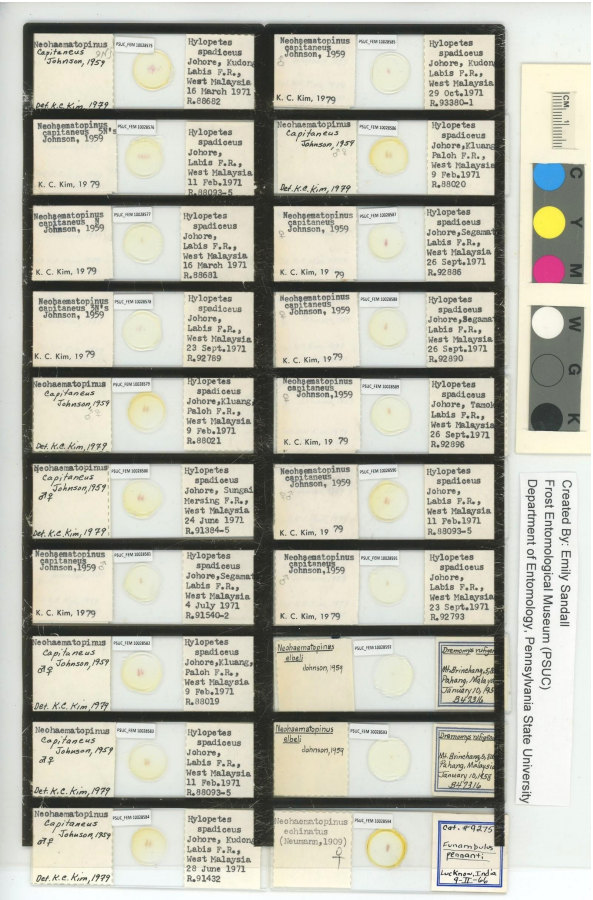


Wet specimens => point and shoot + people power

>48,000 lots/specimens



Slide specimens => scan + slice + people power



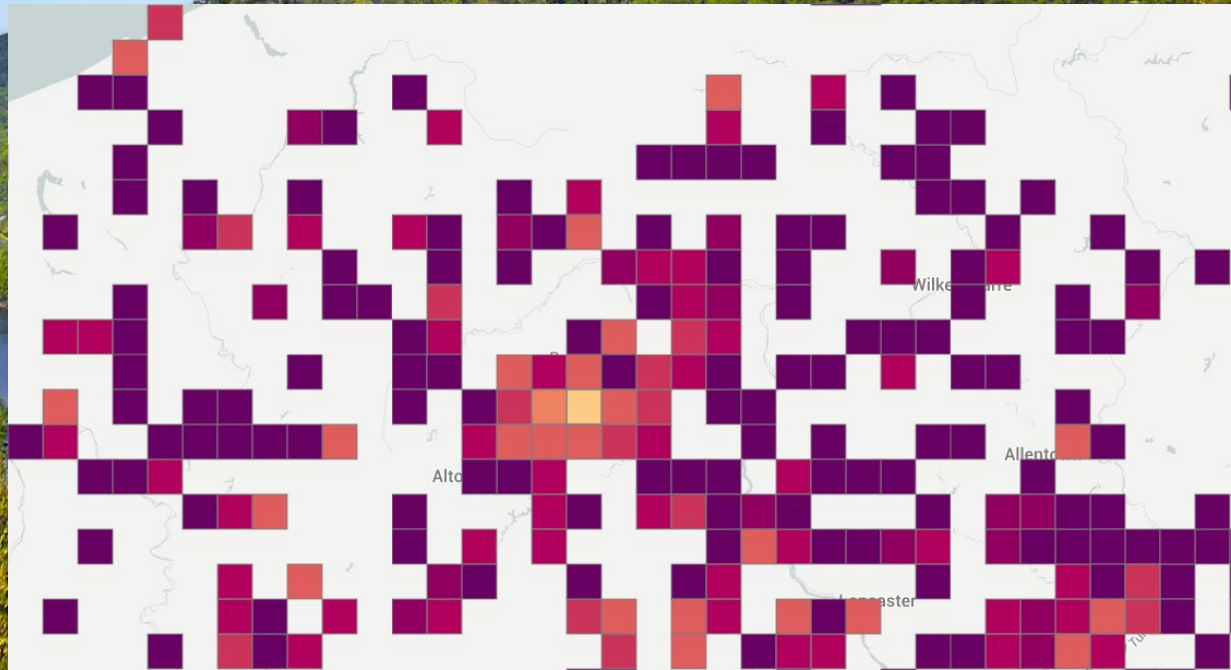
>15,000 slides

Lesson 4:

Digitization is the *best* way to learn about a collection and to set priorities

Nicholas A. Tonelli

Pennsylvania!



Nicholas A. Tonelli

Odonata!

Wayne Hodgkinson

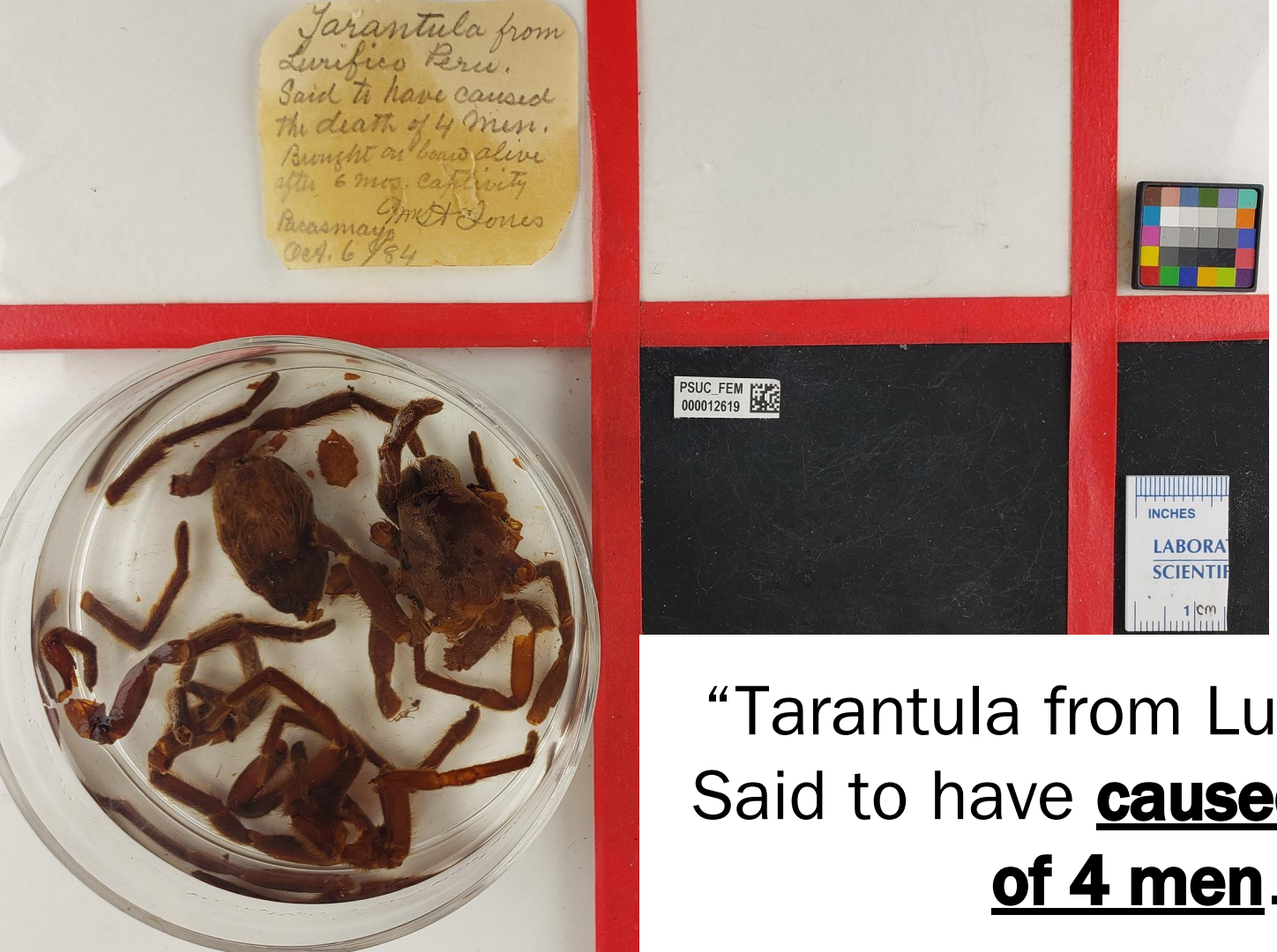
Odonata!

Mark Yokoyama

Wayne Hodgkinson

Araneae?!

Eric Kilby



"Tarantula from Lurifico Peru. Said to have **<u>caused the death of 4 men</u>**."

Lesson 5:

Digitization inspires and catalyzes research

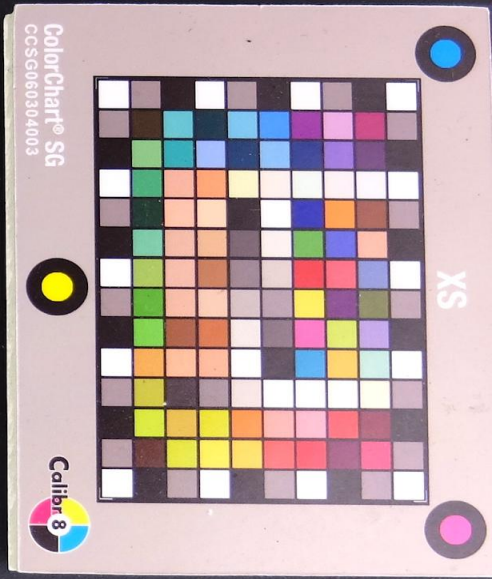
Created By: Kelsey VanHorn
Created On: 11 November 2016
Repository: Frost Entomological Museum (PSUC)

PSUC_FEM
000018962

ARGIA VIOLACEA Hagen ♀ IN COP. det. by G. H. Beatty, III

Darby Creek at Earle Estate
Radnor Twp., Delaware Co.,PENNA.
17 August,1946
coll. by G. H. Beatty, III

461635 COLLECTION OF G. H. & A. F. BEATTY



Reprinted from the *Proceedings of the Pa. Academy of Science*, Vol. 43, 1969

EDAPHIC FACTORS IN THE DISTRIBUTION OF PENNSYLVANIA ODONATA

G. H. & A. F. Beatty
P. O. Box 281
State College, Pennsylvania 16801

ABSTRACT

Pennsylvania is examined as a habitat for Odonata from the standpoint of physiography, glacial characteristics, and watersheds of river systems, all being features distinct from climate. Known distributions of most of the 170 species comprising the Pennsylvania Odonata fauna were plotted with respect to these features, and the results are summarized and discussed.

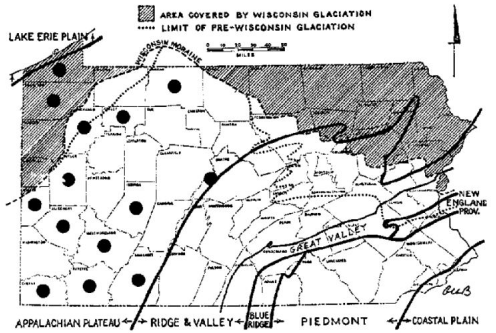


Figure 6. Distribution in Pennsylvania, by counties, of Enallagma antennatum.

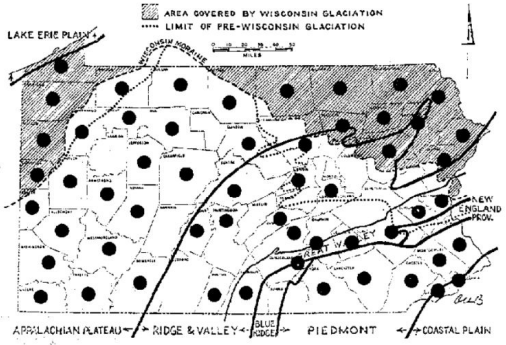


Figure 3. Distribution in Pennsylvania, by counties, of Ischnura verticalis.

Temporal differentiation in environmental niche modeling of Nearctic narrow-winged damselflies (Odonata: Coenagrionidae)

Emily L. Sandall[1] **and Andrew R. Deans**[1]

[1]**Frost Entomological Museum, Department of Entomology, Pennsylvania State University, University Park, PA**

Corresponding author:
Emily L. Sandall

Email address: els22@psu.edu

ABSTRACT

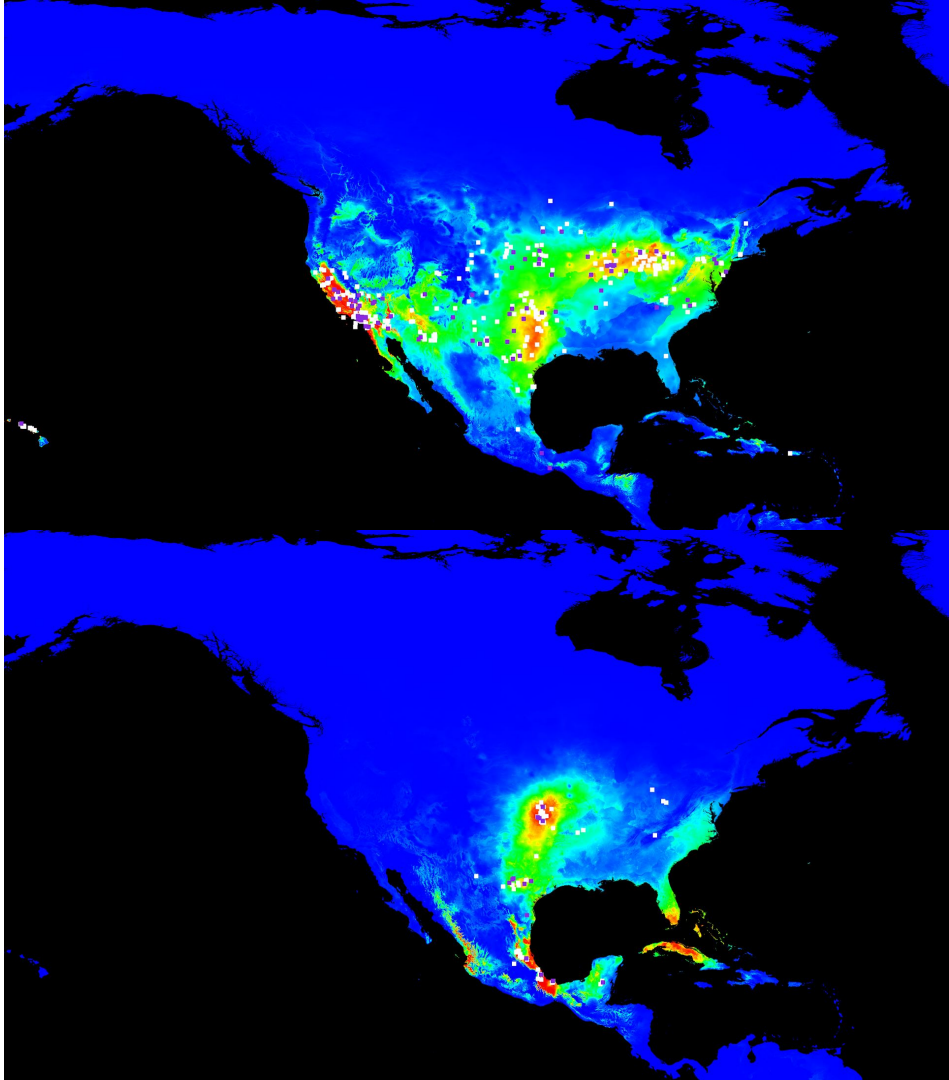
Narrow-winged damselflies (Odonata: Coenagrionidae) can be observed in a variety of habitats, by both professional collectors and amateur odonatologists. Their abundance and ease of recognition has resulted in a large amount of occurrence data, which can be used to establish species distribution maps through environmental niche modeling. Distributional models often aim to maximize the quantity of occurrence points and environmental variables to relate to the distribution, neglecting both the quality and overlap of these two datasets when generating the models. In order to examine the effects of temporal data and environmental variables influencing change in species distributions, we used occurrence data for twelve species of Coenagrionidae damselflies to generate niche models separated by time periods of specimen collection. Our study examines environmental niche models generated for four time periods for each of these coenagrionid species: *Amphiagrion abbreivatum* (Selys,1876), *Enallagma civile* (Hagen,1861), *Chromagrion conditum* (Hagen in Selys, 1876), *Nehalennia gracilis* Morse, 1895, *Enallagma hageni* (Walsh, 1863), *Hesperagrion heterodoxum* (Selys, 1868), *Nehalennia irene* (Hagen, 1861), *Argia moesta* (Hagen, 1861), *Ischnura ramburii* (Selys, 1850), *Argia tibialis* (Rambur, 1842), *Argia translata* Hagen in Selys, 1865, and *Argia vivida* Hagen in Selys, 1865. The best supported models in each analysis were generated with occurrences of specimens collected from the 1970s to 2000s, and we used occurrence data outside of this range, from the 1800s to 2017, to compare the consistency of model predictions based on specimens of different time periods. In this approach, combining traditional environmental niche modeling and analysis of the specimen records themselves, we have found that ranges for narrow-winged damselflies expand over time, with increase in distributional coverage and decrease in model strength without temporal overlap between occurrences and environmental variables.

INTRODUCTION

Linking current and past occurrences on a large scale, to document species distribution over time requires that specimens be identified and accessible. The digitization of natural history collections (Graham et al., 2004; Page et al., 2015) has liberated the occurrence and taxonomic determination data of millions of insect specimens. Some of these insects have captured the attention of naturalists for centuries, resulting in extensive datasets that allow for a thorough analysis of shifting distributions, outbreaks, or other population changes (Peterson et al., 2005; Estrada-Peña et al., 2013). Odonata, commonly known as dragonflies and damselflies, is one such order with extensive representation in natural history collections throughout time (Bybee et al., 2016). Maximal amounts of occurrence data are used to establish species distributions typically without thorough examination of individual specimen's temporal data. For many taxa, the quantity of digital data available may not be great enough to carry out such analyses. Taxa with widespread collection and digitization efforts enable analysis of data associated with the specimen occurrences in order to build environmental niche models of distribution. In doing so, we can identify how concept of species distribution changes as specimen occurrences are documented.

The addition of recent observations and citizen science projects to digital repositories extends the reach of occurrence data beyond the walls of a single museum or individual researcher (Graham et al., 2004;

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27261v1 | CC BY 4.0 Open Access | rec: 7 Oct 2018, publ: 7 Oct 2018



Sandall EL, Deans AR. (2018) Temporal differentiation in environmental niche modeling of Nearctic narrow-winged damselflies (Odonata: Coenagrionidae) *PeerJ Preprints* 6:e27261v1 https://doi.org/10.7287/peerj.preprints.27261v1

Lesson 6:

Digitization connects people to collections and to science

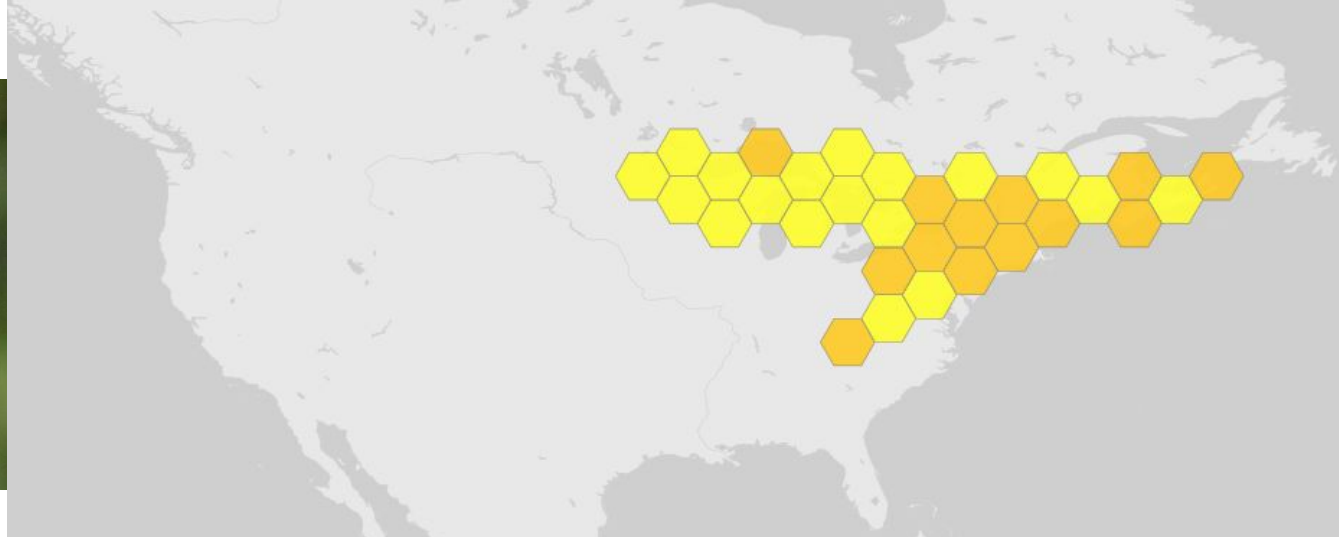
Spilomyia fusca Loew, 1864

Dolichovespula maculata (L., 1763)

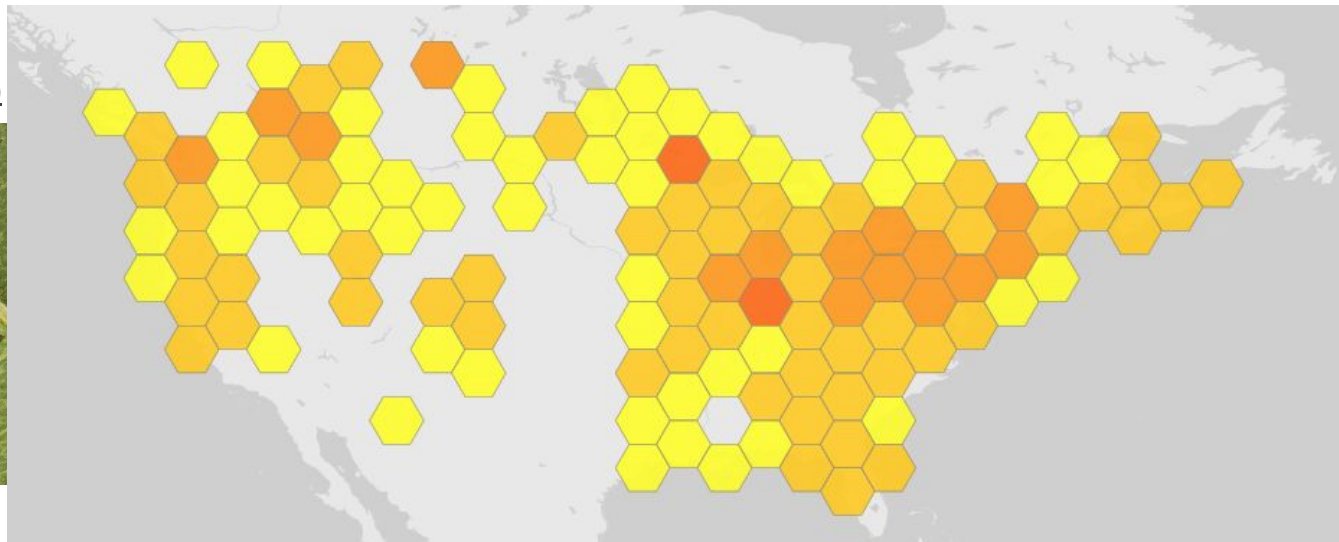
What can collection data tell us about mimicry?

- Batesian
- Müllerian
- imperfect
- aggressive

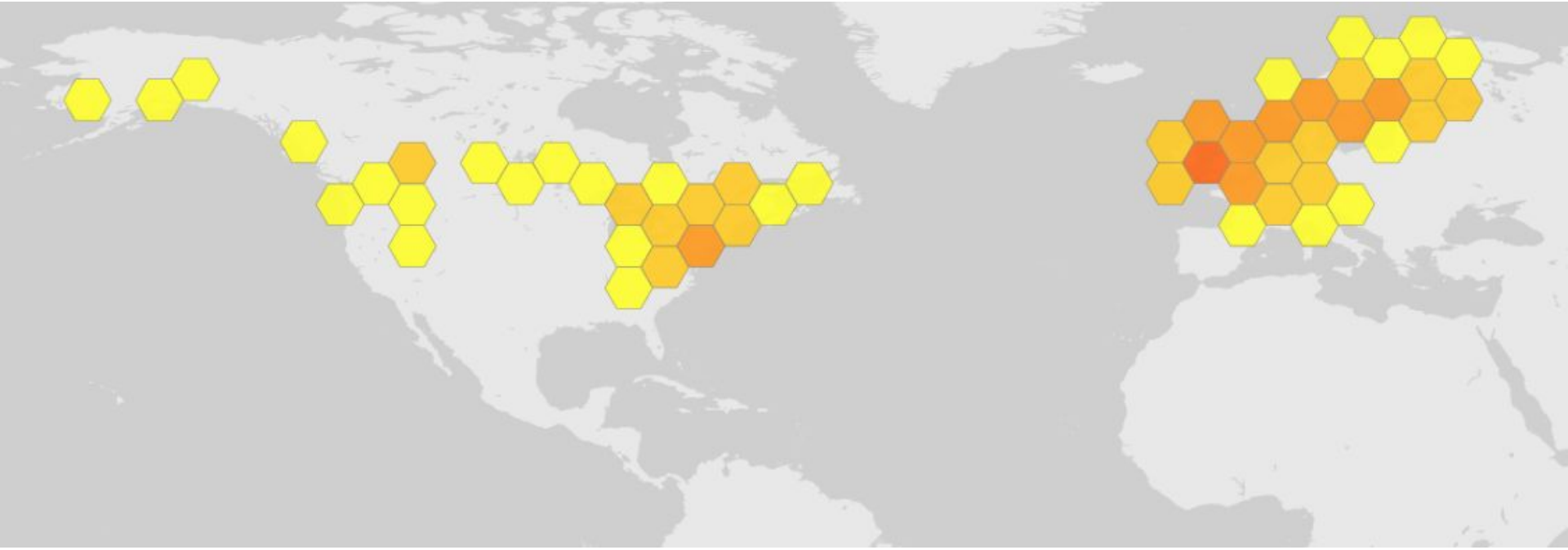
Spilomyia fusca Loew, 1864



Dolichovespula maculata (L., 1763)



Epistrophe grossulariae (Meigen, 1822)










Vouchering

Lesson 7:

Digitization connects people to urgent global problems

Little Tipple

Lesson 8:

Digitization is *transformative*

Collections research @ PennState College of Agricultural Sciences

- Pollinator ID tools
- Documenting pollinators
- Monitoring and understanding invasive species
- Diversity of beneficial insects in applied systems
- Roles of microbes in mediating plant-insect interactions







Acknowledgments

This material is based upon work supported by the U. S. National Science Foundation, under Grant Numbers DBI-1349356 and EF-1400993. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

We thank Sal Anzaldo, Karah Roof, Brian Reily, Angela Hoover, Rachel Davis, Isa Betancourt, and Rafa Toro for their substantial contributions to collection health and digitization.

Questions?

DOI: 10.26207/86jq-hq45