



AFTER DIGITIZATION...
TAXONOMY?



Global Biodiversity Information Facility

Free and open access to biodiversity data

417,944,819
OCCURRENCES

1,426,888
SPECIES

11,974
DATASETS

580
DATA PUBLISHERS

Sharing biodiversity data for re-use

- [Learn about GBIF](#)
- [Publish your data through GBIF](#)
- [Technical infrastructure](#)

Providing evidence for research and decisions

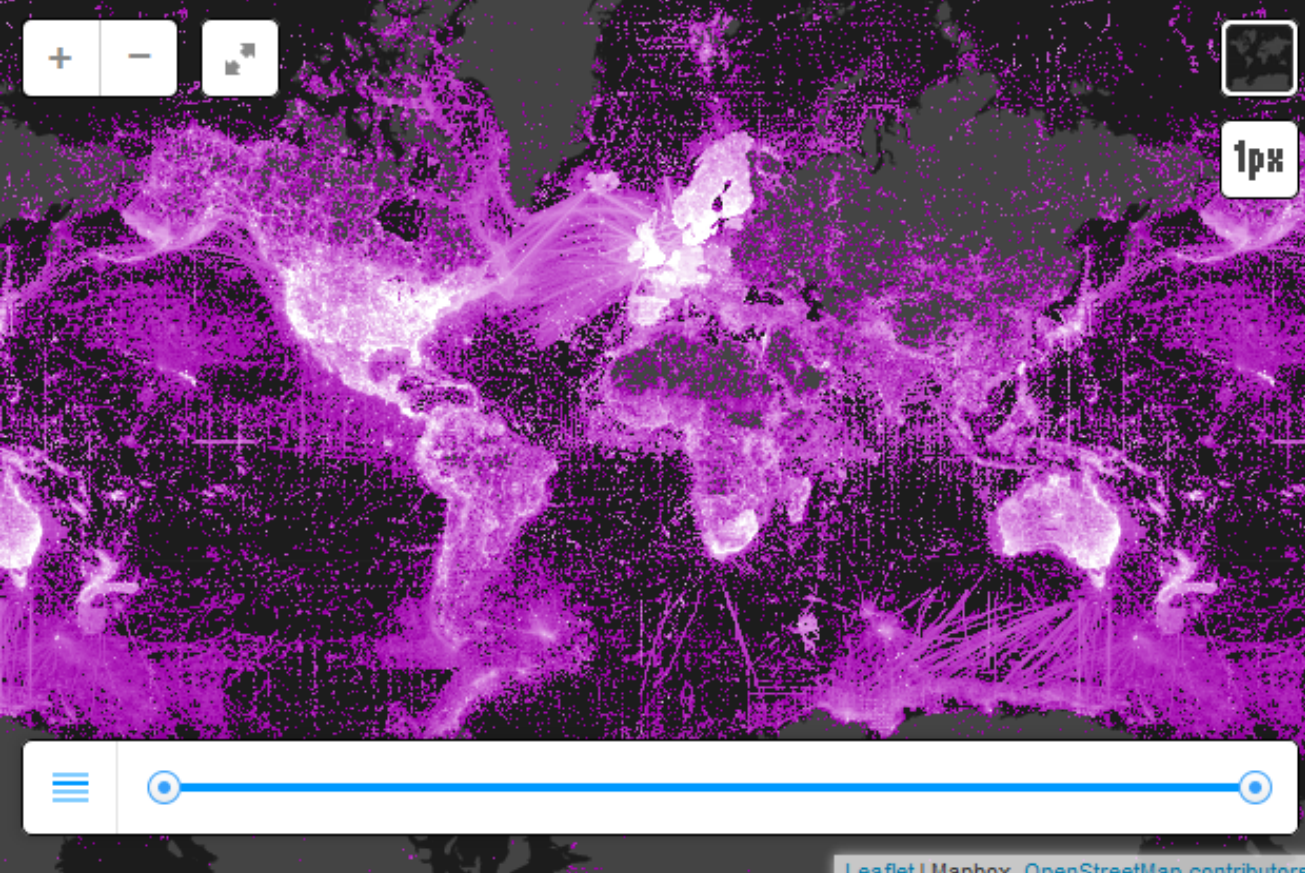
- [Using data through GBIF](#)
- [Enabling biodiversity science](#)
- [Supporting global targets](#)

Collaborating as a global community

- [Current Participants](#)
- [How GBIF is funded](#)
- [Enhancing capacity](#)

Search news items and information pages... Search

GBIF: Over 400 million occurrence data



Georeferenced data

VIEW RECORDS

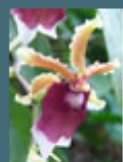
[All records](#) | [In viewable area](#)

ABOUT

This map shows the density of all 355,780,948 georeferenced occurrence records published through the GBIF network.

To explore the records, zoom into the map or click on the links above and add further filters to customize search results.

**GBIF: Over 350 million
georeferenced occurrence data**



Symbiota

Promoting
Bio-Collaboration



Main Menu

Symbiota Home

- Overview
- Active Data Portals
- Join Existing Portal
- Establish New Portal
- Support
- Help Pages
- Contact Us

Specimen Search

- Overview
- Data Interoperability
- Rare Species Protection

Floristic Research

- Research Species Lists
- Dynamic Species Lists

Identification Keys

- Overview

Image Library

Data Management

- Specimen Records
- Checklist Editor
- Voucher Editor
- Morphological Characters
- Mass Updater

News / Events

- Meetings / Presentations
- Symbiota Workshops

Active Symbiota Projects

The following biodiversity projects use the Symbiota Virtual Biota software package to establish biodiversity portals with specific regional and/or taxonomic scopes. Each portal represents the community of collaborating researchers that manage the core scientific data. If you are interested in becoming a data contributor for an already established data portal, contact the portal administrator for more information on gaining access to their data editing tools. For more information on what is involved in being a data provider, read the [Specimen Integration](#) page.

Quick Index

- [SEINet](#)
- [Intermountain Regional Herbarium Network \(IRHN\)](#)
- [Consortium of North American Bryophyte Herbaria \(CNABH\)](#)
- [Consortium of North American Lichen Herbaria \(CNALH\)](#)
- [Mycology Collections Data Portal \(MycPortal\)](#)
- [Consortium of Northeastern Herbaria Portal \(CNH\)](#)
- [Madrean Archipelago Biodiversity Assessment Project \(MABA\)](#)
- [Cooperative Taxonomic Resource for American Myrtaceae \(CoTRAM\)](#)
- [Herbario Virtual Austral Americano \(HVAA\)](#)

SEINet - Southwest Environmental Information Network

SEINet is a virtual flora that covers the southwestern region of North America with a particular interest in the desert regions.

- <http://swbiodiversity.org/seinet/index.php>
- [Active Projects](#)
- [Data Providers](#)
- Data digitization funding
 - Completion and Coordination of Databases of Arizona Vascular Plants at ASU, ARIZ, and ASC [BRC-0237418](#)
 - Collaborative Research: Expanding SEINet [HR-0847966](#)

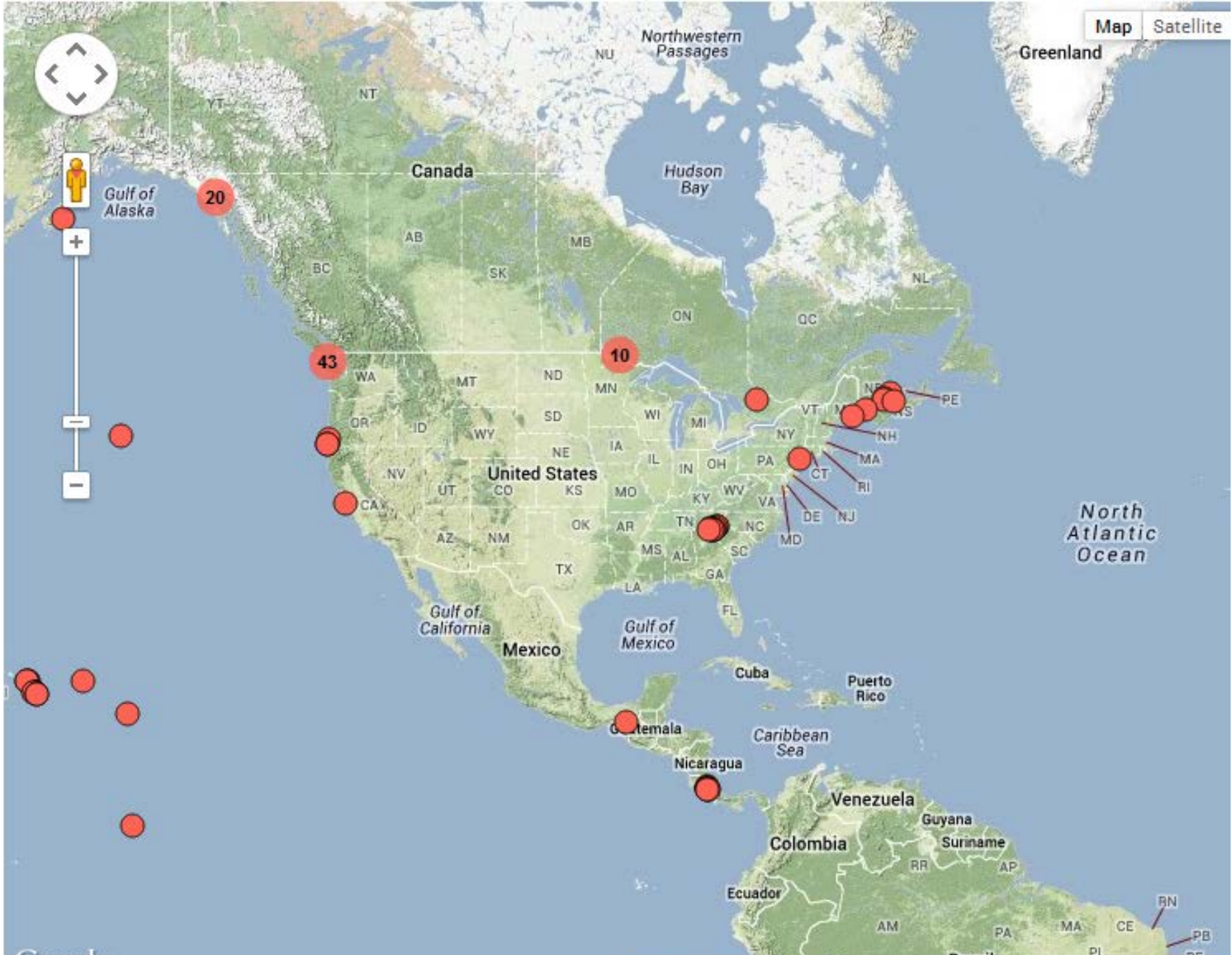
Login

Login as...

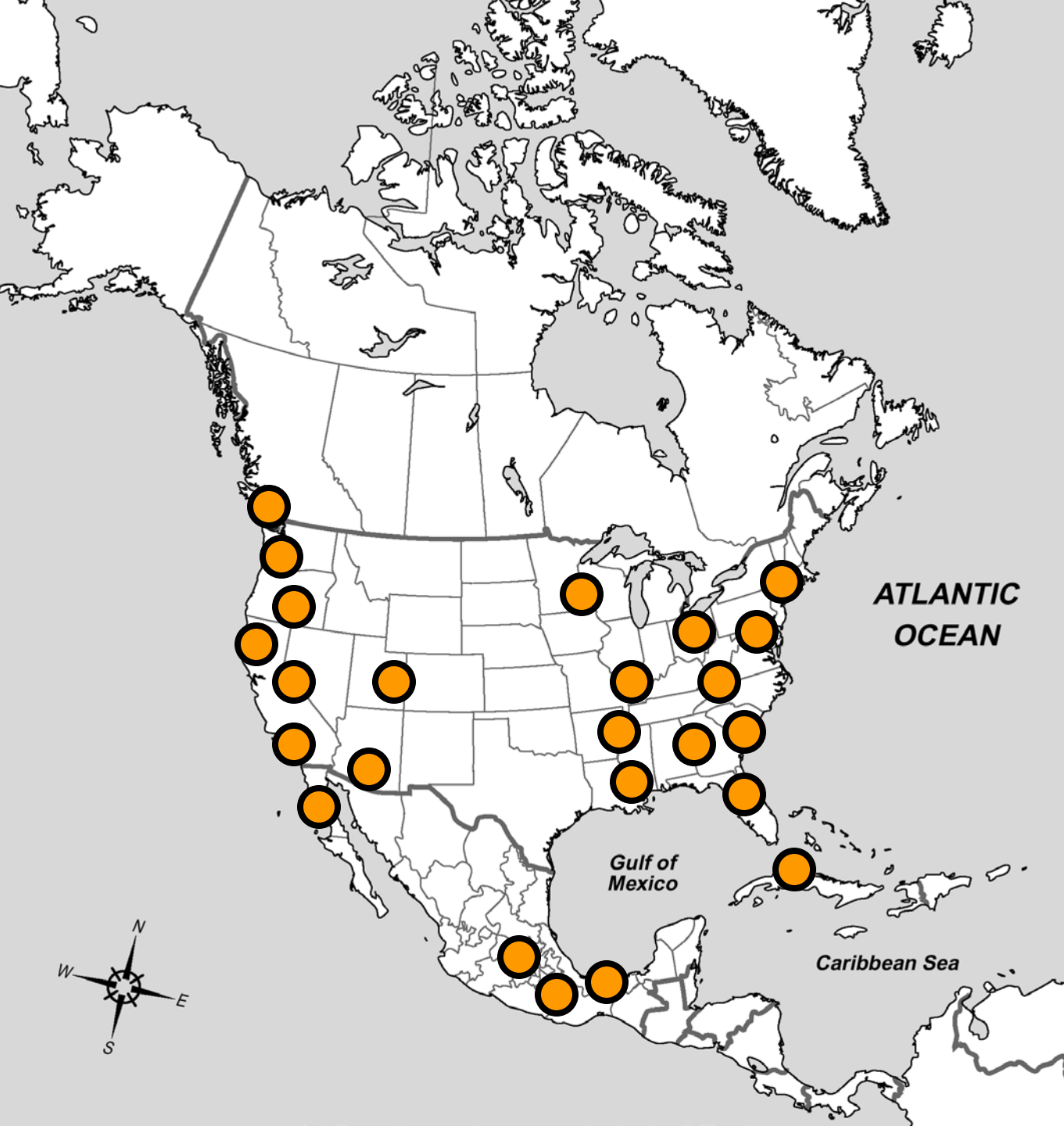
User:

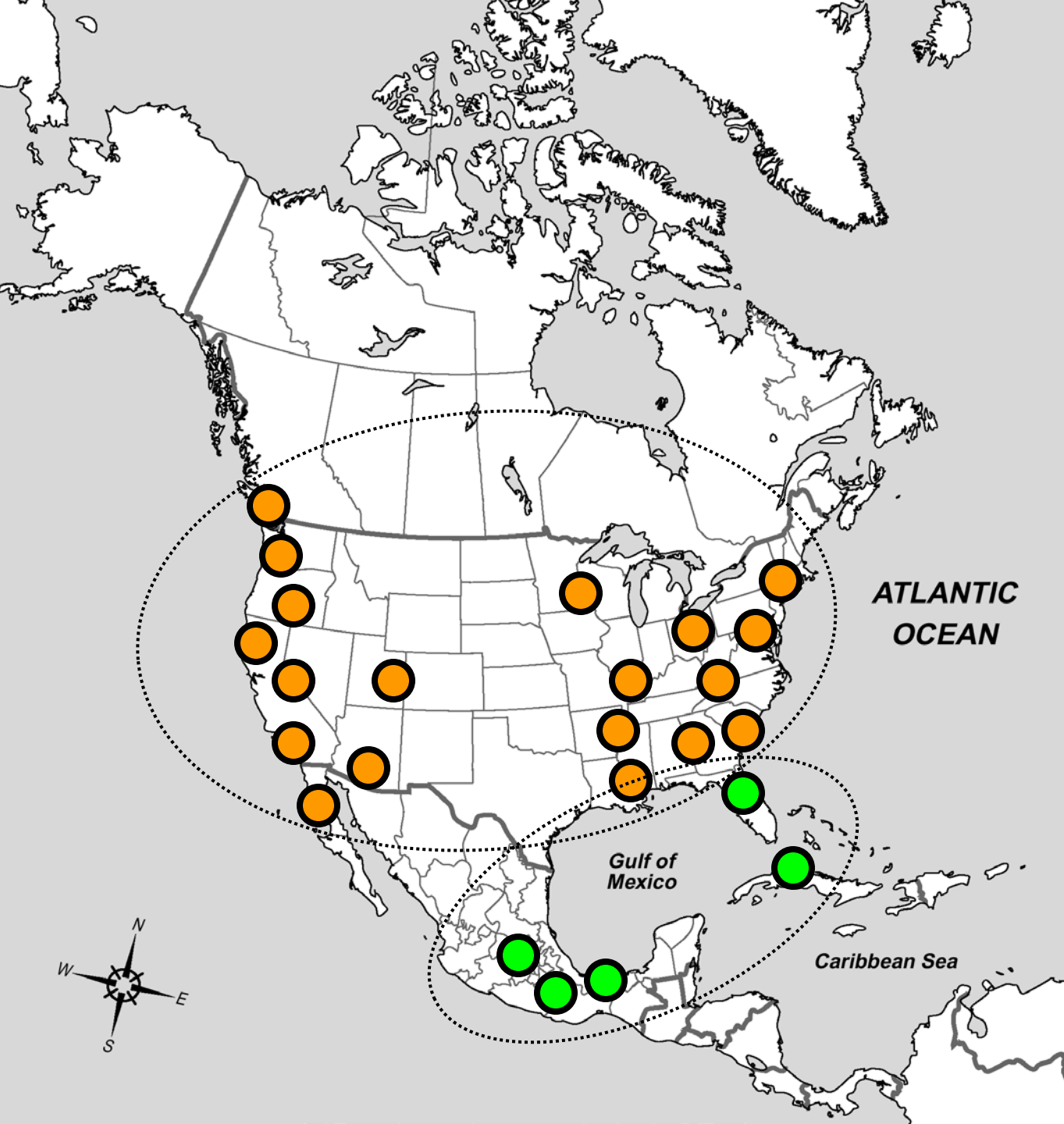
Password:

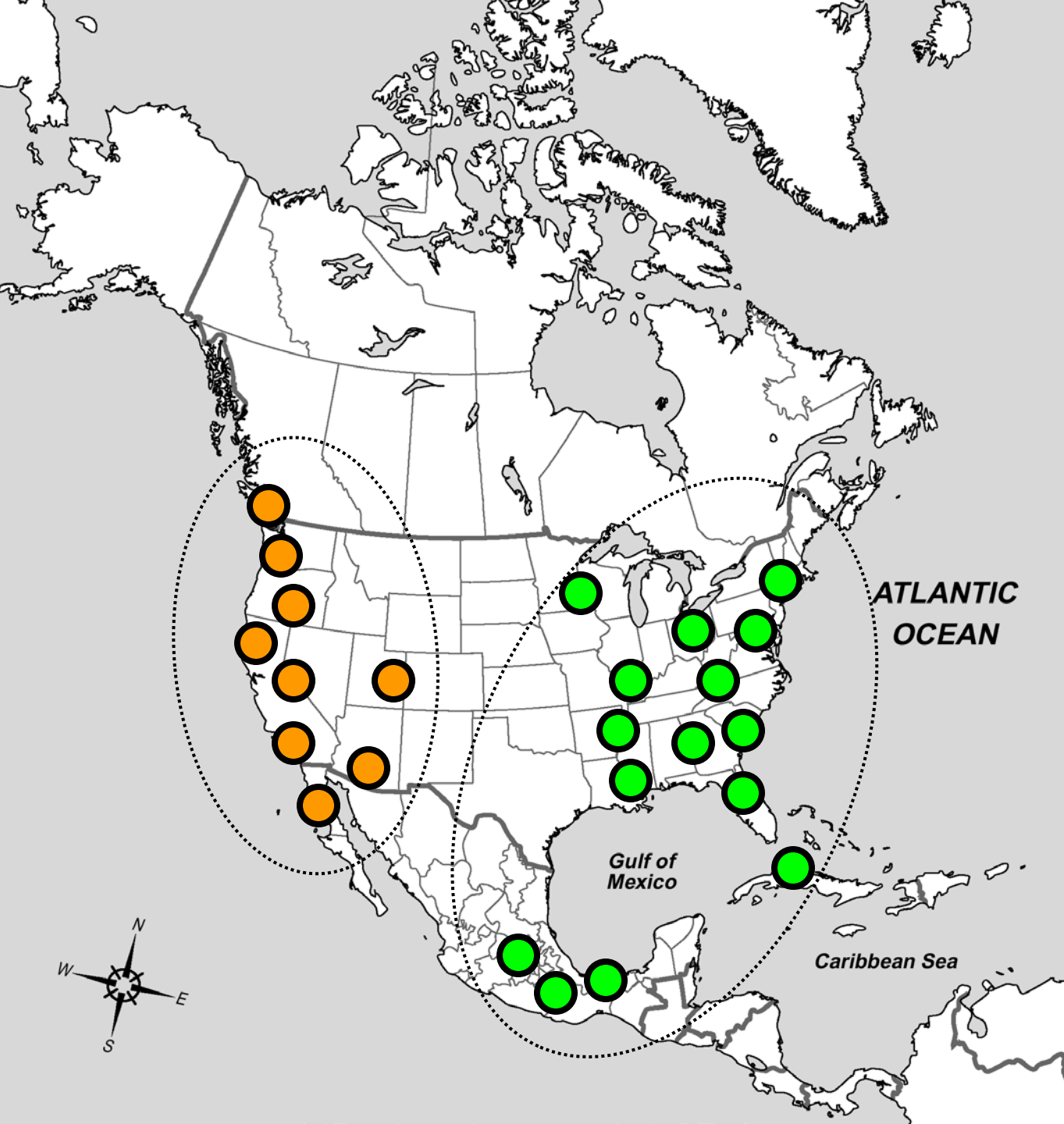
Login

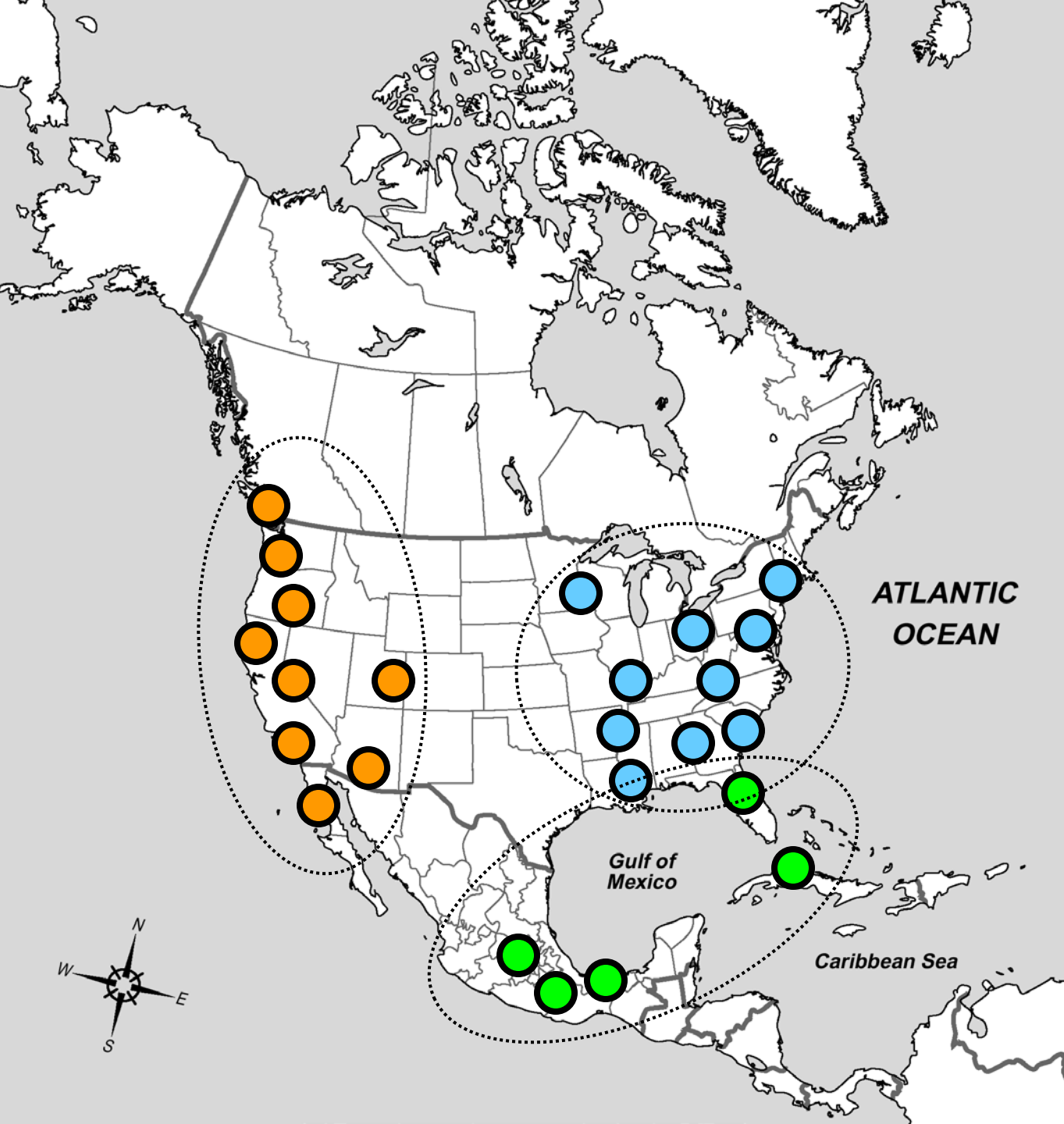


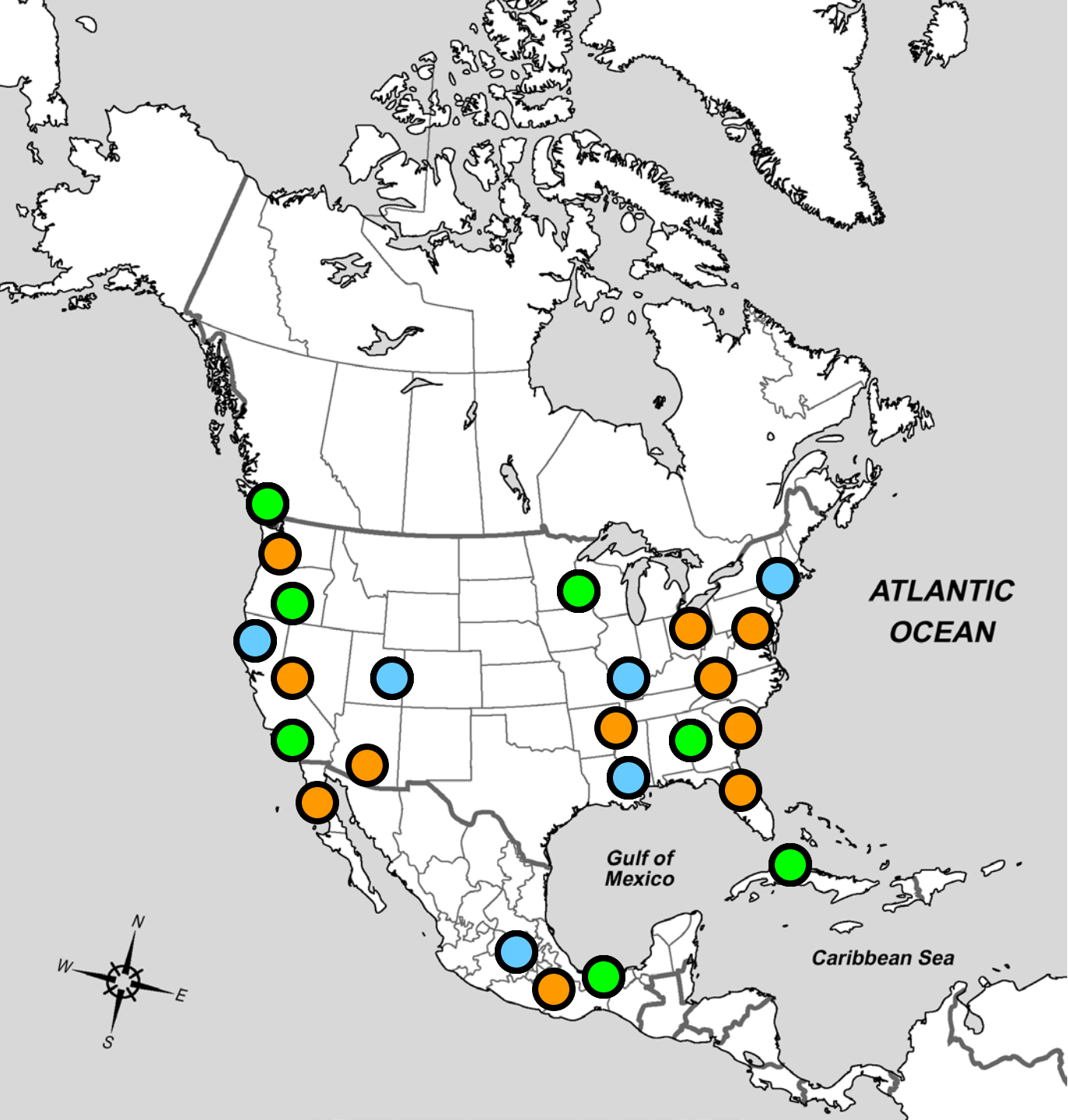
**Is this
taxonomy
real?**











Problems:

- 1. Changes in nomenclature
(genus names, epithets)**
- 2. Plain wrong identifications**
- 3. Changes in species concepts
(splitting, lumping)**

Problems:

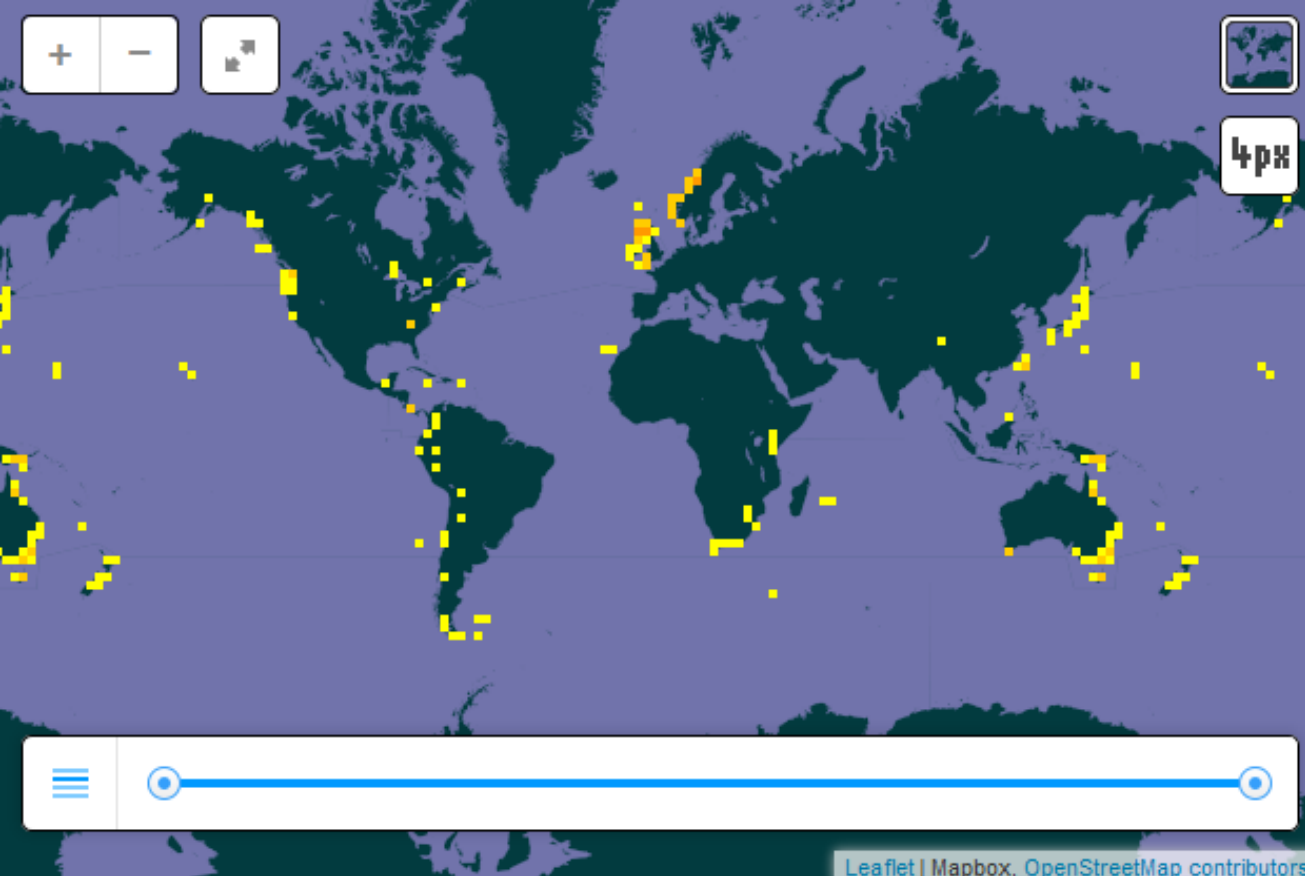
**1. Changes in nomenclature
(genus names, epithets)**

2. Plain wrong identifications

**3. Changes in species concepts
(splitting, lumping)**



Pseudocyphellaria crocata



Georeferenced data

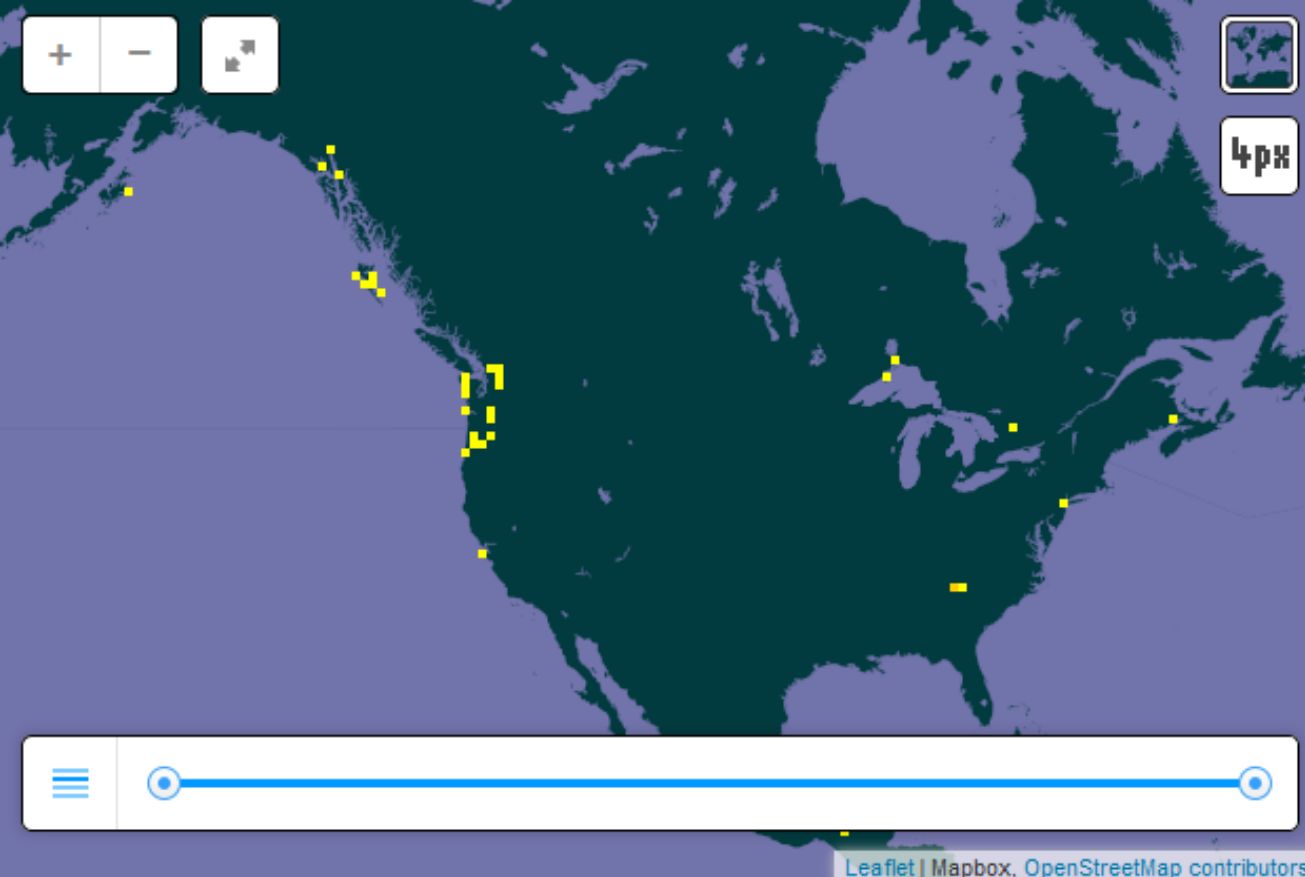
VIEW RECORDS

All 1,348 | [In viewable area](#)

DISTRIBUTIONS

Text based [distributions](#) present in some sources.

GBIF: Occurrence data *Pseudocyphellaria crocata* (global)



Georeferenced data

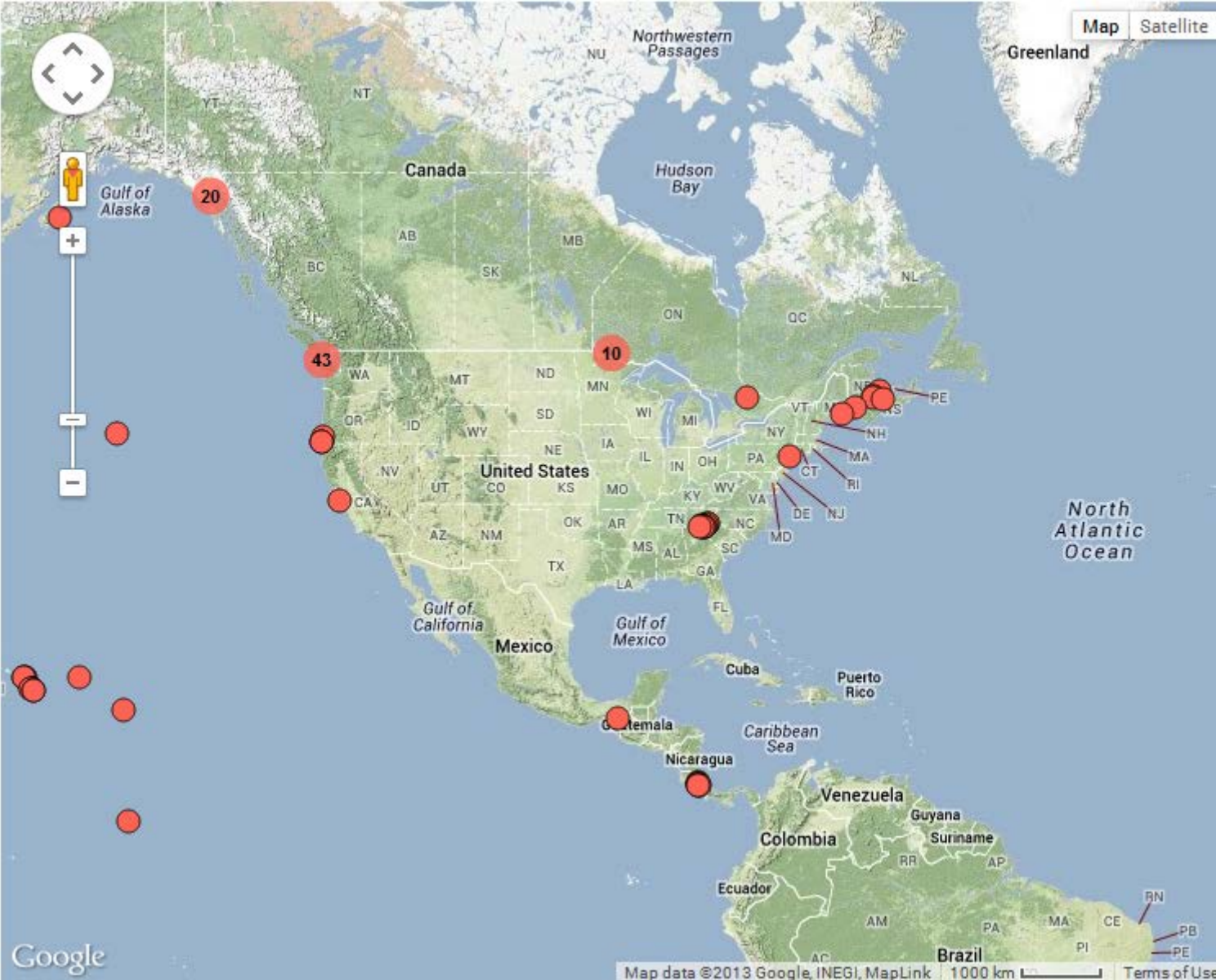
VIEW RECORDS

All 1,348 | [In viewable area](#)

DISTRIBUTIONS

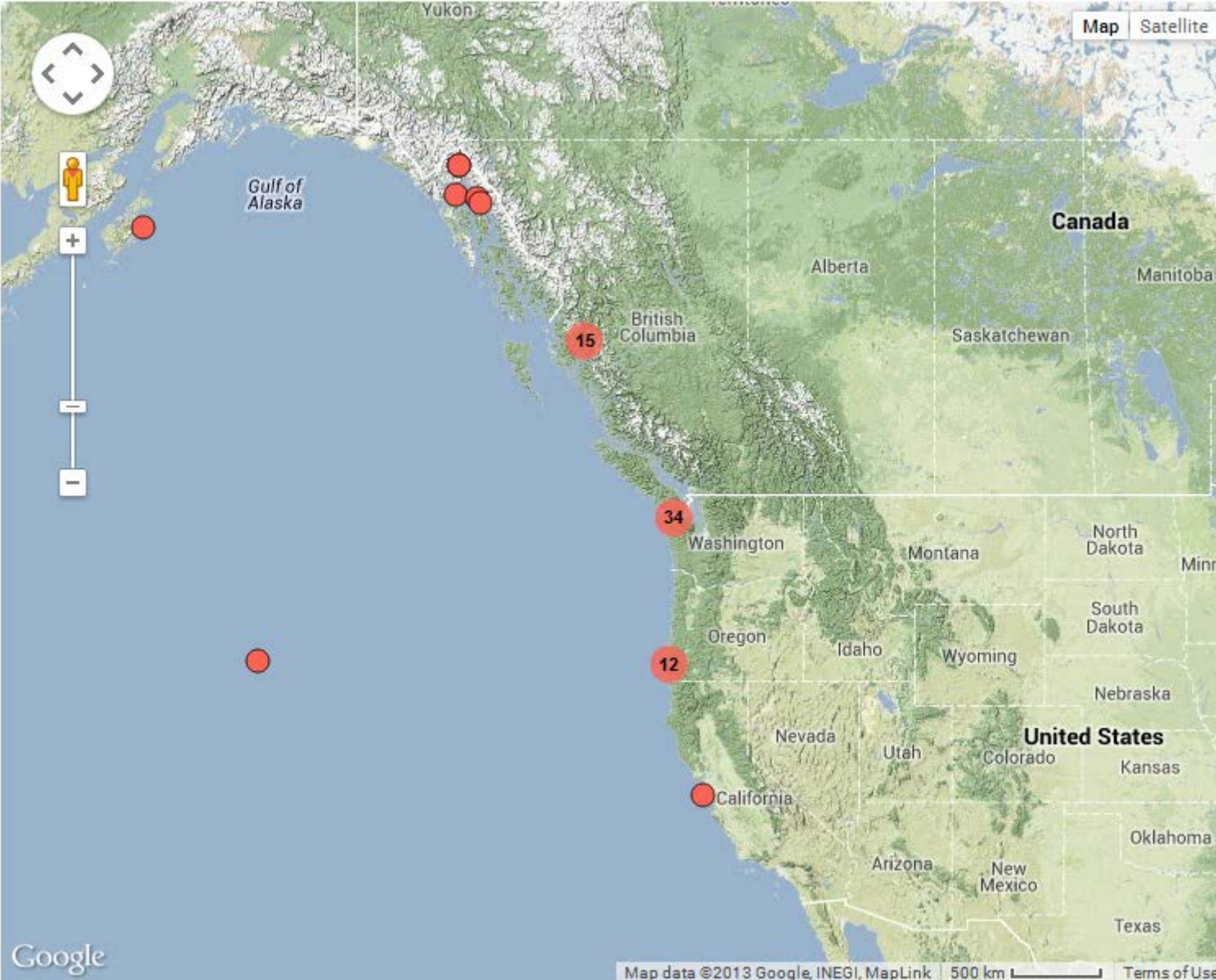
Text based [distributions](#) present in some sources.

GBIF: Occurrence data *Pseudocyphellaria crocata* (North America)



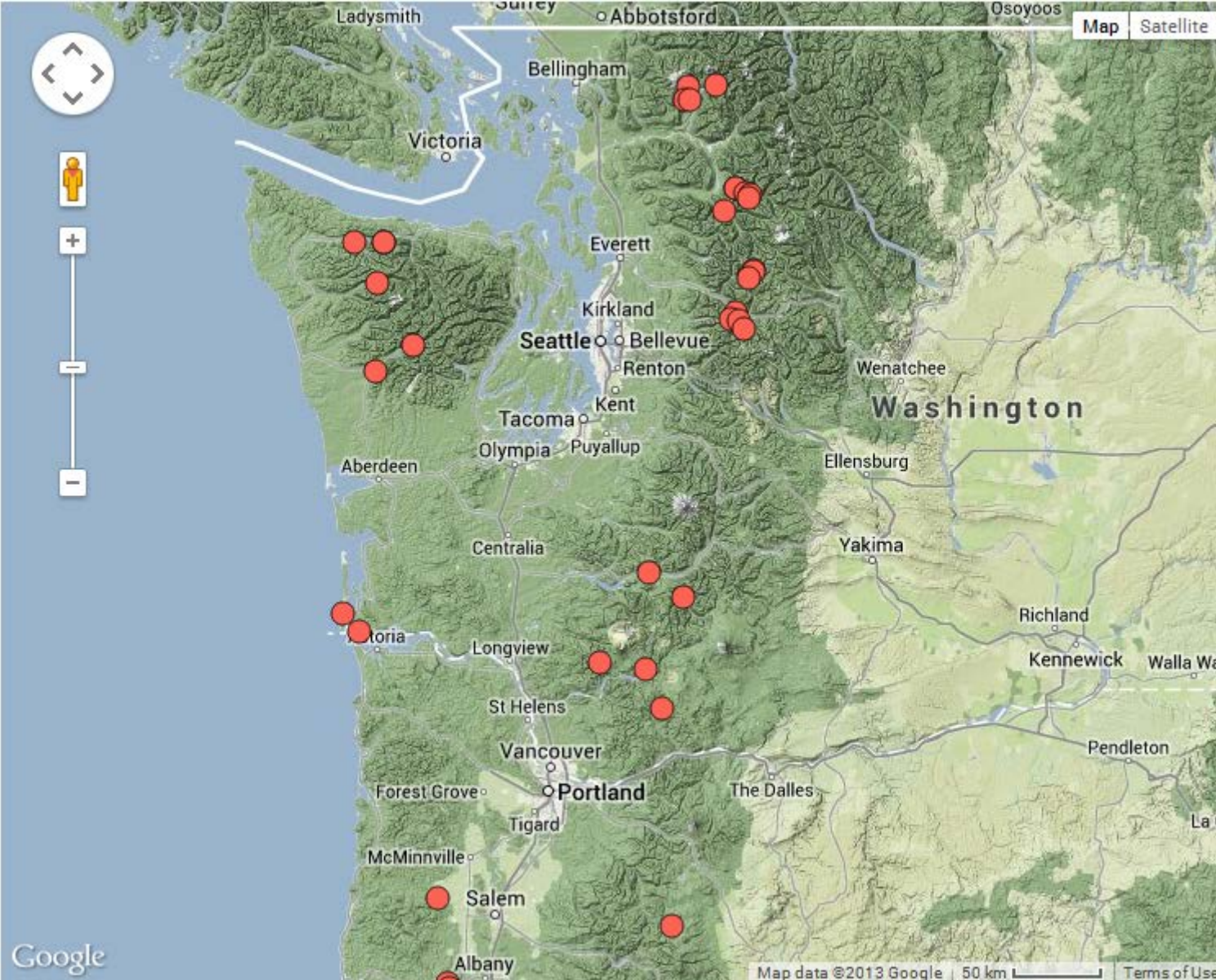
North America

SYMBIOTA: Occurrence data
Pseudocyphellaria crocata



Pacific North West

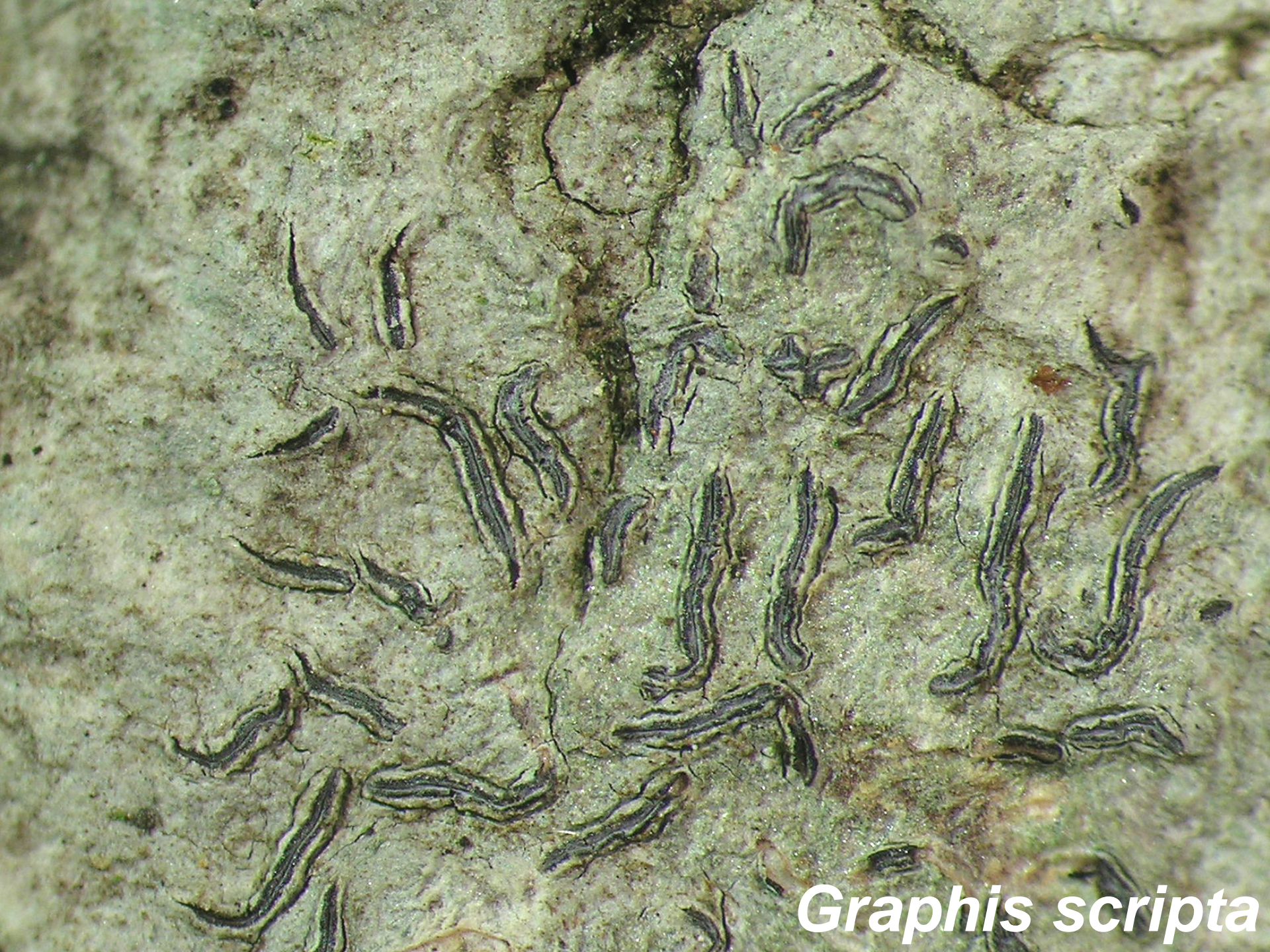
SYMBIOTA: Occurrence data
Pseudocyphellaria crocata



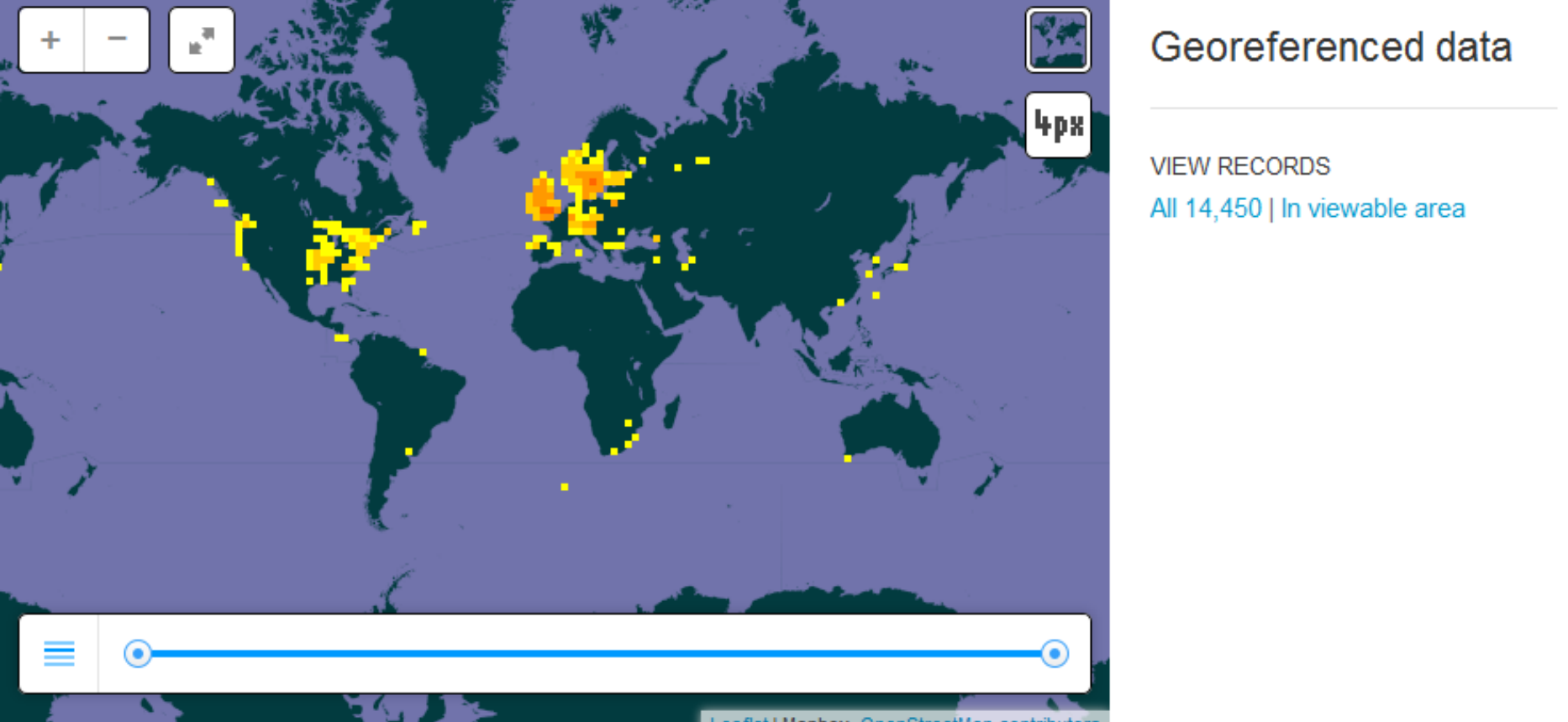
**Wash-
ington**

**Three
species!**

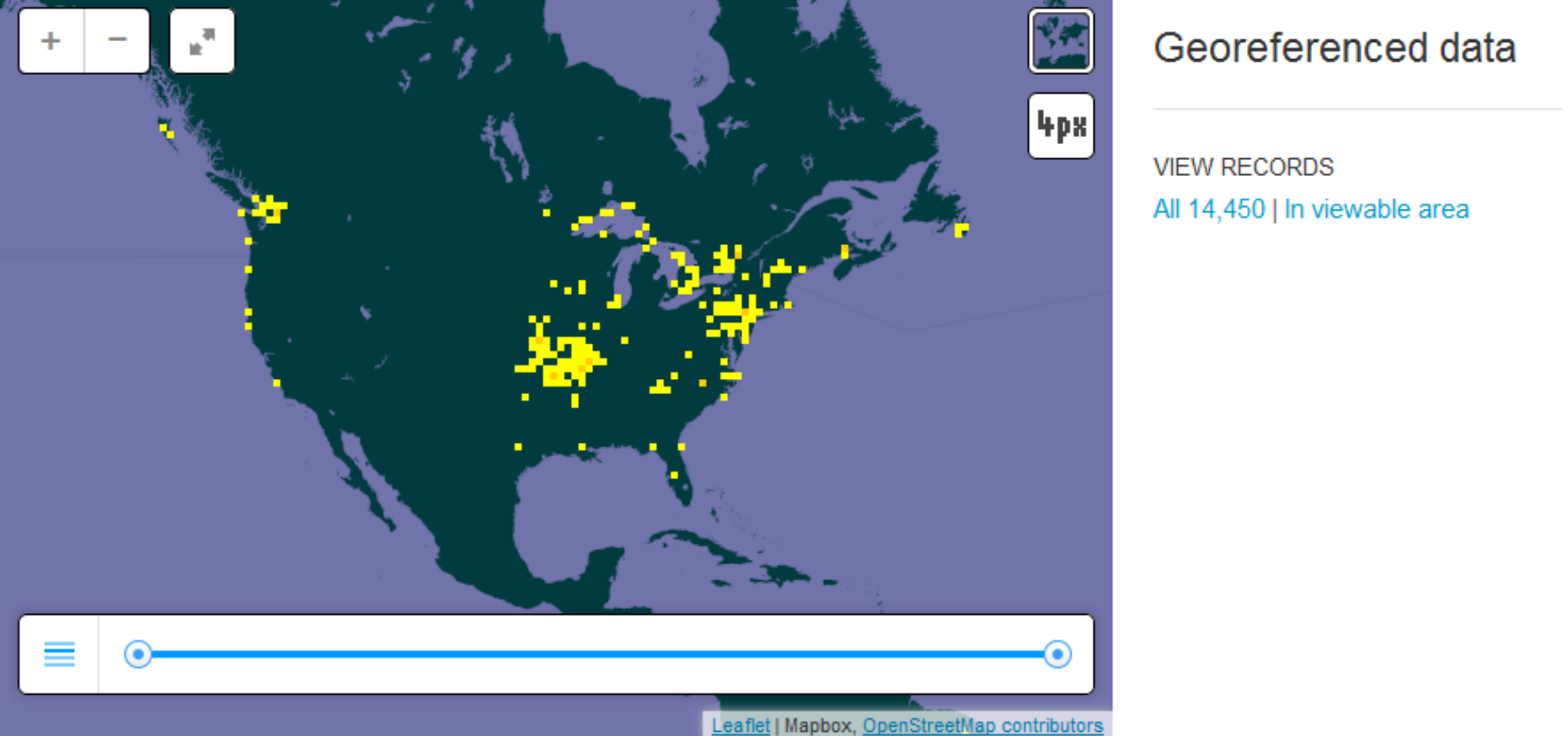
SYMBIOTA: Occurrence data
Pseudocyphellaria crocata



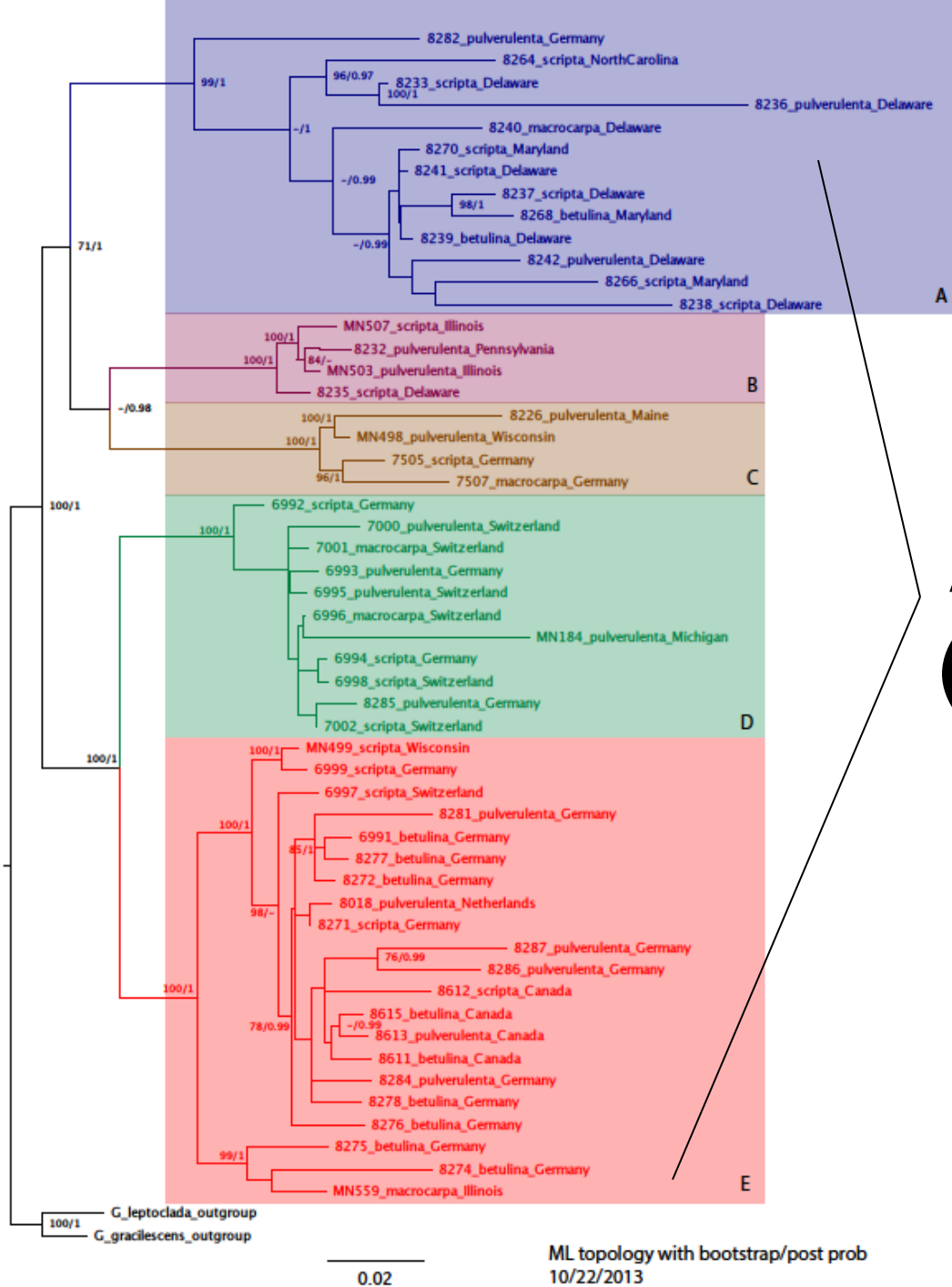
Graphis scripta



GBIF: Occurrence data *Graphis scripta* (global)



GBIF: Occurrence data *Graphis scripta* (North America)



**Up to 2011:
A single species
(*Graphis scripta*)**

**Now:
At least six
distinct species**

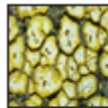


Cora glabrata



SYMBIOTA: Occurrence data

Cora glabrata



ASU

Arizona State University Lichen Herbarium

Catalog #: 522459

Taxon: *Dictyonema glabratum* (Spreng.) D. Hawksw.

Family: Corticiaceae

Determiner: T.L. Esslinger: Checklist of North America, 2002

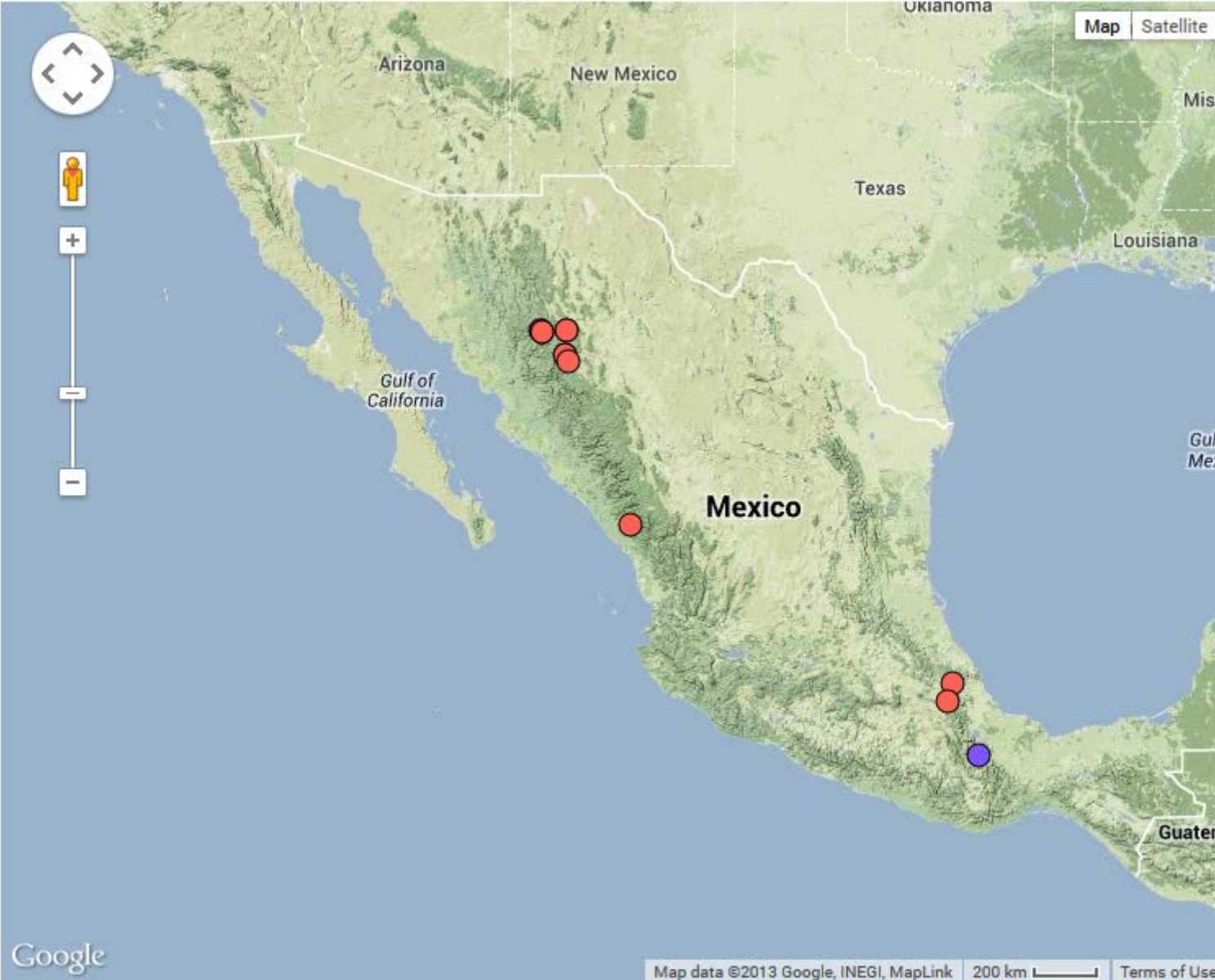
Collector: T.H. Nash III 29128

Date: 05 January 1990

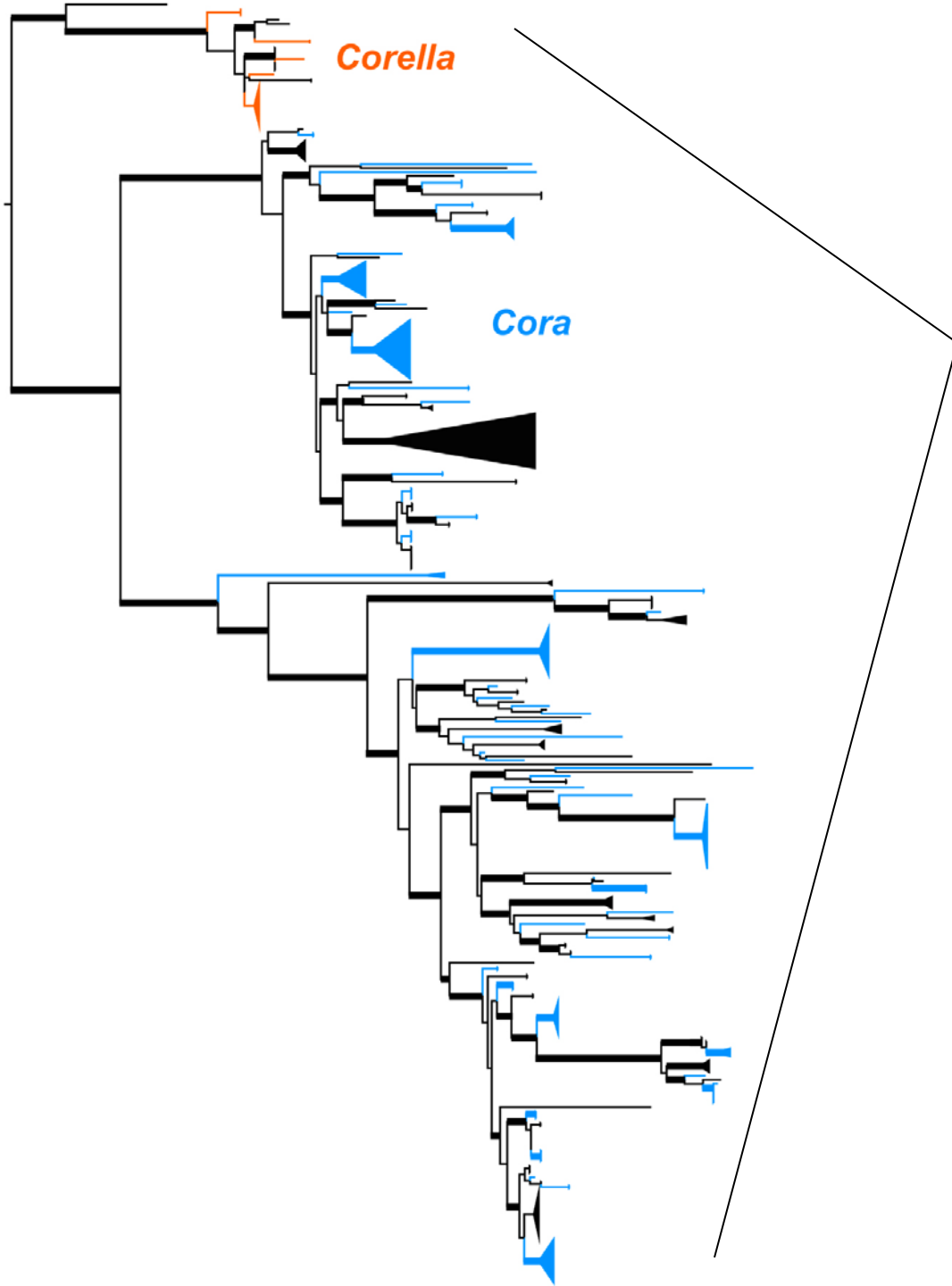
Locality: Venezuela, Mérida, , 17 km SE of Bailadores
80.216667 -71.875

Elevation: 2600 meters (8528ft)

Substrate: roadbank



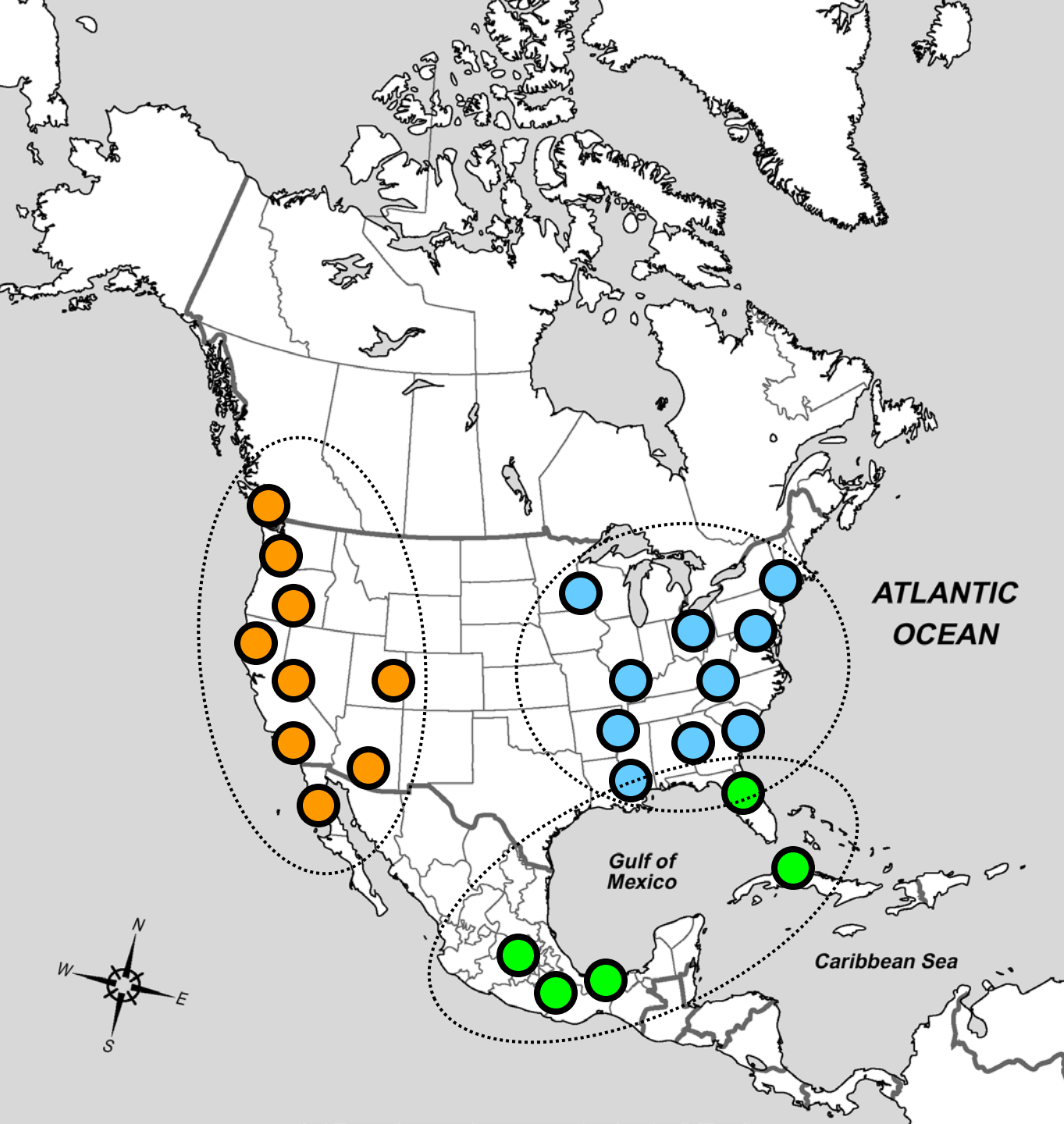
SYMBIOTA: Occurrence data
Cora glabrata



**Up to 2009:
A single species
(*Cora glabrata*)**

**Now:
Over 100
distinct species**

**How
can we
address
this
problem?**



In an ideal world:

New round of NSF Funding (ATRBC)

**"Advancing Taxonomic Revisions
of Biological Collections"**

Organize ourselves in TRNs

"Taxonomic Revisionary Networks"

In the real world:

**Effective and largely automated ways
to single out problematic records**

**Not only feed collections data
into research but also feed
research data into collections**

"Global automated annotations"

1. Apply reliability ("quality") scores to each identification using a combination of:

Date of most recent identification

Identification authority

Cited in monographic work?

Voucher for DNA studies?

Quality score ranging from 1 to 10

**2. Using automated georeferencing
and statistical tools
to single out and prioritize
problematic occurrence records
for taxonomic revision**

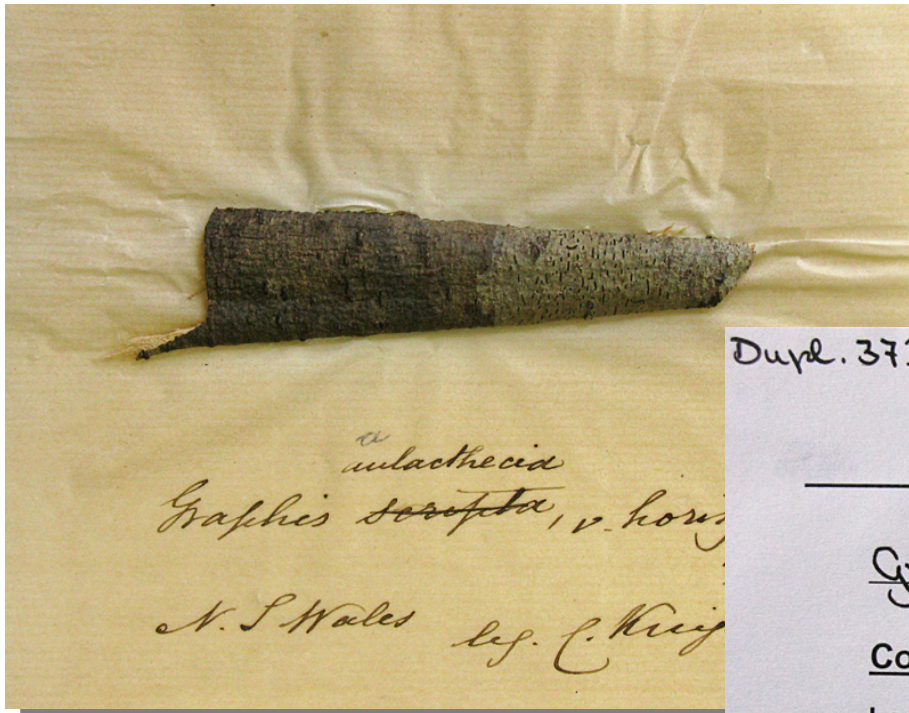
**3. Generate a network of taxonomic
experts and automatically
alert experts about specific
problematic occurrence records**

4. Connect published and unpublished research to collections data through taxonomic experts network

"Global automated annotations"

Automated annotations include indication of tools and level of expertise required to correctly identify species.

1. Apply reliability ("quality") scores



Score = 1

Score = 10

Dupl. 37370

Herbarium Klaus Kalb

Graphidactra multiformis (Mont. & Bosch)

Thor

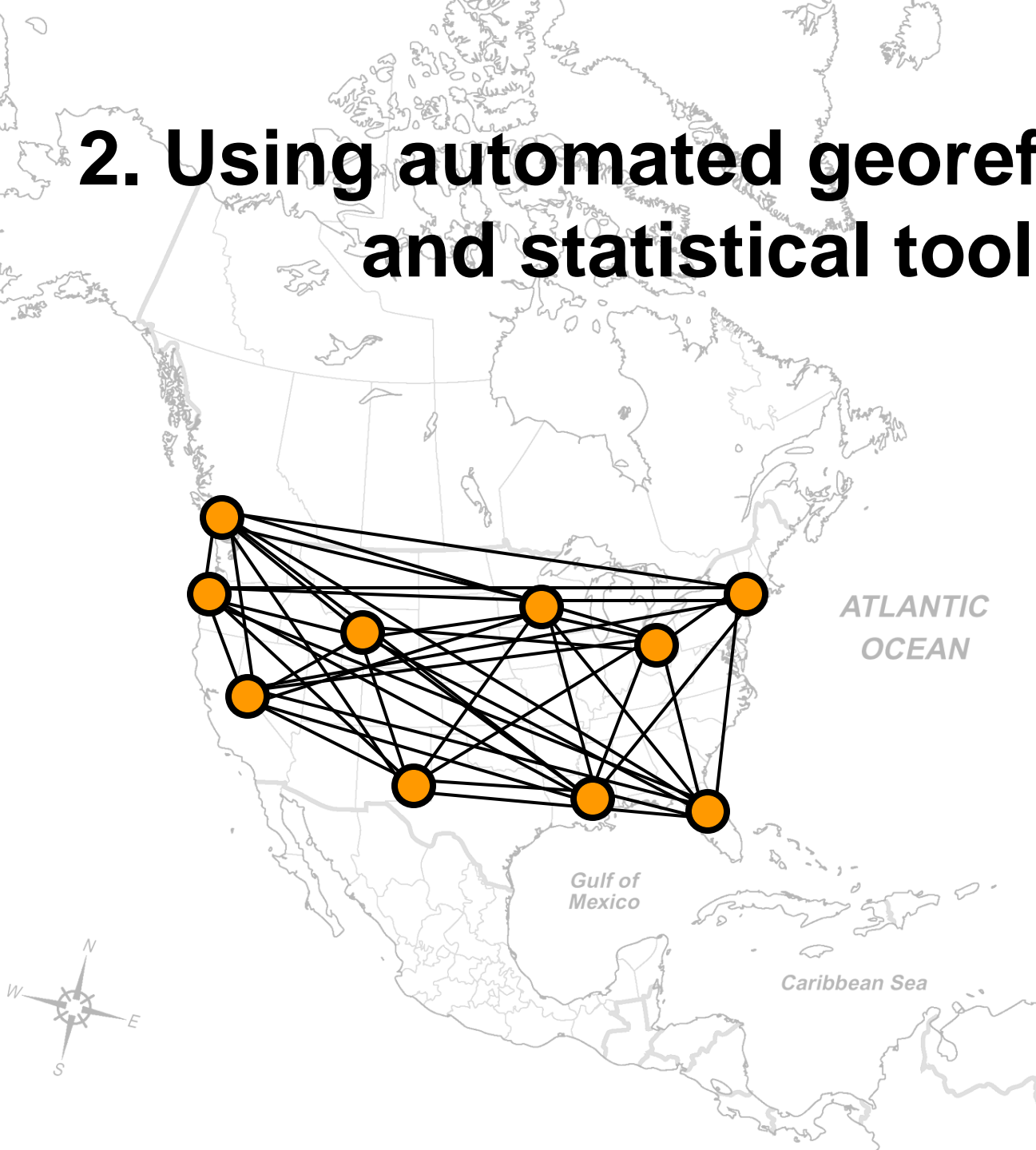
Country: Australia/Queensland

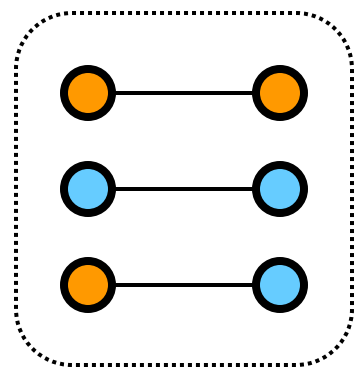
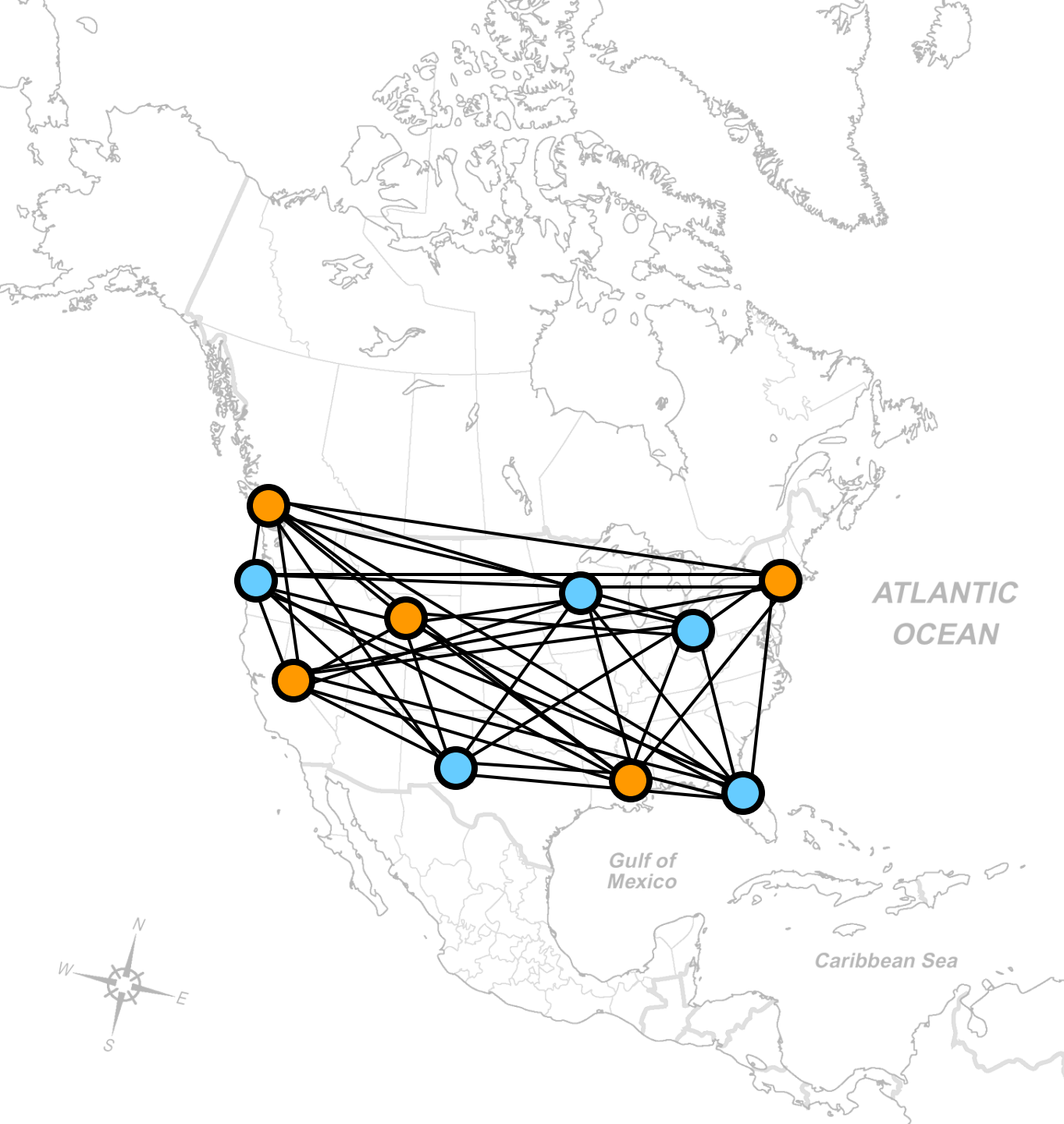
Locality: Daintree National Park, Cape Tribulation, c. 65 km N of Mossman; in a dense tropical rain-forest, 220 m. 16°14'43" S, 145°26'09" E.



Date: 27. VIII. 2008

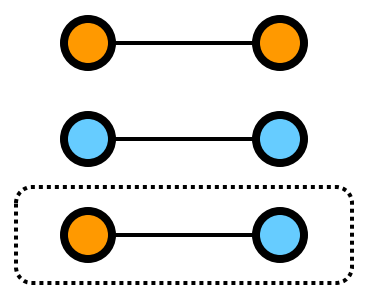
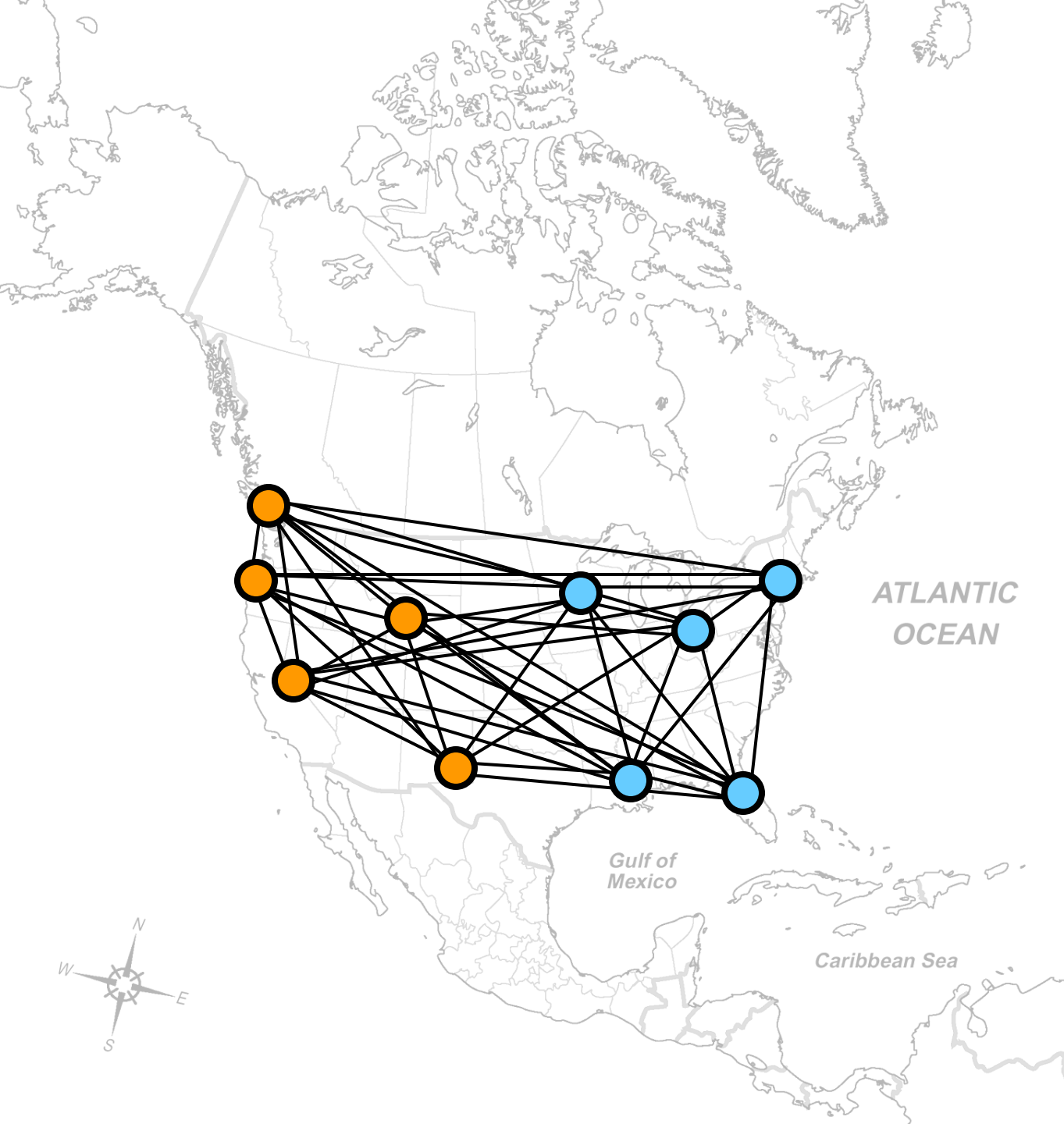
leg.: K. Kalb
det.: K. Kalb



2. Using automated georeferencing and statistical tools





-  High quality score
-  Low quality score



 **High quality score**
 **Low quality score**

3. Generate a network of taxonomic experts with automated alerts

Taxonomic experts database

**Indicate taxonomic groups
of expertise**

**Automatically send specimen
alerts to indicated expert**

4. Connect published and unpublished research to collections data through universal portals

**Annotated specimen:
"point" correction**

**Annotated name:
global "alert" for all specimens
identified with that name**

**"Global automated annotation"
of species name:**

Published reference and/or authority

**Alternative correct identifications
for a given name**

**Requirements to revise specimen
(macroscopic, microscopic, DNA)**

Magnitude of the problem?

Per institution:

- 1. Select 100 samples randomly**
- 2. Evaluate correct identification**
- 3. Score percentage of correct identification against:**

Date of most recent identification
Identification expertise