Welcome!
And a few logistical details

Wiki: https://www.idigbio.org/wiki/index.php/Fluid_Preserved_Invertebrate_Imaging
Adobe Connect (Kevin Love): http://idigbio.adobeconnect.com/paleo
Being broadcast and recorded
Be observant of remote audience; use microphone to make comments, ask questions
Chat box for remote participants
Efficiency: Starting on time; staying on track
Lunch: 1.25 hours/day
Meals: Breakfast and lunch provided in Kalamazoo Room. Dinner on your own.

Origin of this workshop
Working/interest group

This material is based upon work supported by the National Science Foundation under Cooperative Agreement EF-1115210. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Integrated Digitized Biocollections (iDigBio) An Introduction

Gil Nelson
Institute for Digital Information and Scientific Communication
Integrated Digitized Biocollections
Florida State University

Fluid-preserved Invertebrate and Microscopic Slide Imaging Workshop
16-18 September 2013
University of Michigan
Museum of Zoology

This material is based upon work supported by the National Science Foundation under Cooperative Agreement EF-1115210. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

The U.S. National Science Foundation estimates there may be as many as 1.8 billion biological and paleontological specimens stored in U. S. museums and academic institutions (perhaps as many as 3 billion worldwide). But, no one really knows!

In an effort to make these collections universally accessible to taxonomists, ecologists, researchers, and the general public, in 2011 NSF launched a $100 million, 10-year Advancing Digitization of Biodiversity Collections program and named Florida State University and University of Florida jointly as the national resource for digitization.

Advancing Digitization of Biodiversity Collections

Integrated Digitized Biocollections (iDigBio)
University of Florida
Florida State University
Florida Museum of Natural History

The goal is to digitize and make available via the Web at least 1 billion biological and paleontological records over the 10-year life of the project.

iDigBio
Integrated Digitized Biocollections

Mandate and Responsibility

- Provide/facilitate portal access to collections data
  - Make information available and discoverable
  - Label Data and images
- Enable digitization and research
  - Facilitate digitization workflows
  - Oversee implementation of standards and best practices for digitization
  - Allow for data discovery across organismal groups
- Be a client of digitization projects/networks
  - Actively seek partners and data sources
  - Respond to cyberinfrastructure needs
- Engage communities
  - Collections
  - Research
  - Citizen science and education
- Support ADBC goals
  - Access to information
  - Support for collections
  - Sustainability

Mandate and Responsibility

- Provide/facilitate portal access to collections data
  -
  -
- Enab
  -
  -
  -
- Be a
  -
  -
- Enga
  -
  - Research
  - Citizen science and education
- Support ADBC goals
  - Access to information
  - Support for collections
  - Sustainability

Grand Challenge

Develop a cloud computing infrastructure that links biological data from collections across the U.S. through one or more unified web interfaces to overcome the limitations of “data silos.”

Mandate and Responsibility

- Provide/facilitate portal access to collections data
  -
  -
- Enab
  -
  -
  -
- Be a
  -
  -
- Enga
  -
  - Research
  - Citizen science and education
- Support ADBC goals
  - Access to information
  - Support for collections
  - Sustainability

Grand Challenge

Devel ture
that li ons
across
unifie ne
limitations of "data silos."

More recently, we have been encouraged by NSF to enhance international collaboration and sharing.

The challenges being pursued by iDigBio are reflective of worldwide trends in digitization

- Global Biodiversity Informatics Facility (GBIF)
- OpenUp! (European Union)
- Atlas of Living Australia (ALA)
- SYNTHESYS (20 European natural history museums)

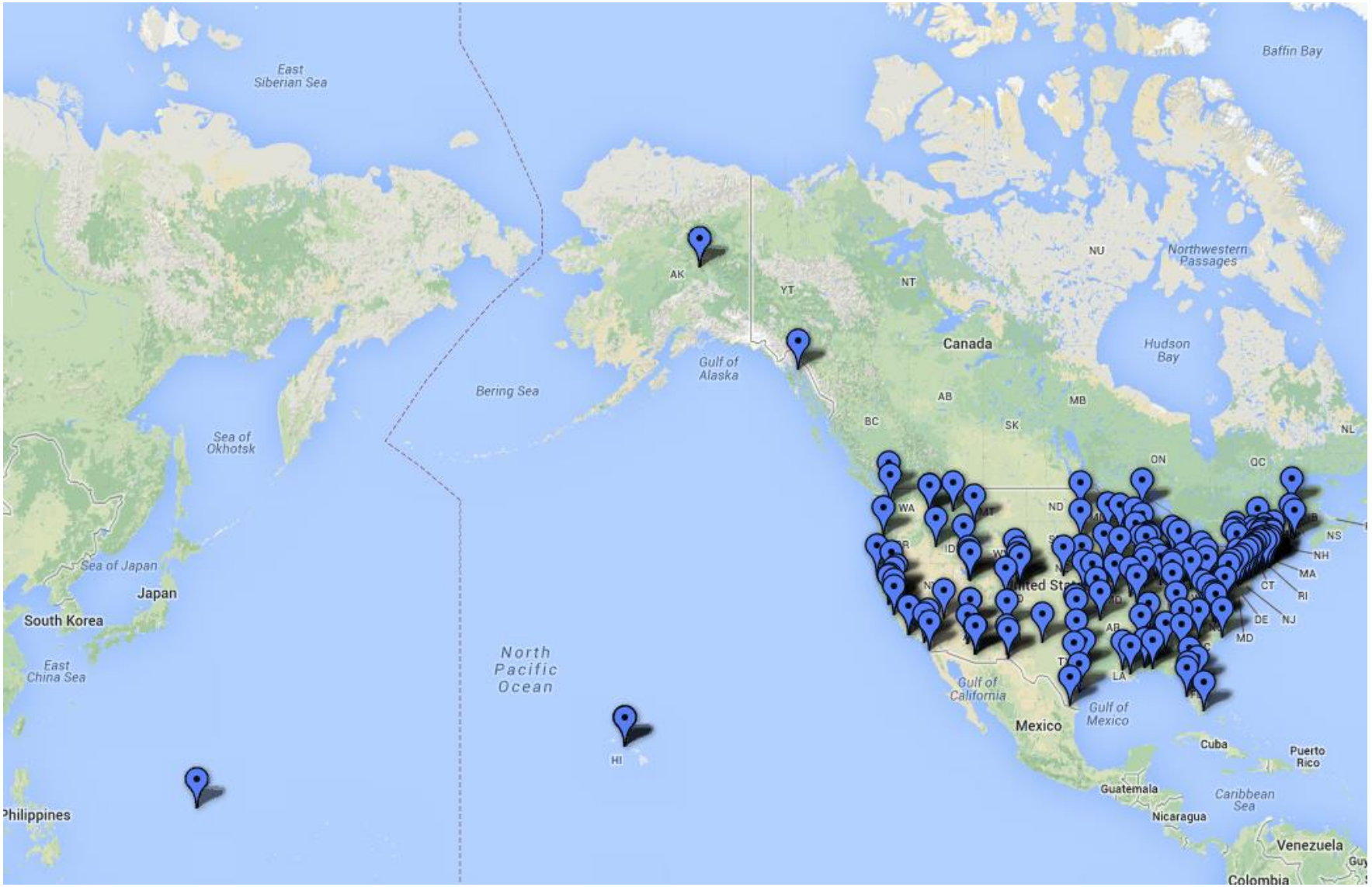
Ten Thematic Collections Networks (TCNs) plus 2 Partner to Existing Networks (PENs)

- InvertNet: An Integrative Platform for Research on Environmental Change, Species Discovery and Identification (*Illinois Natural History Survey, University of Illinois*) http://invertnet.org
- Plants, Herbivores, and Parasitoids: A Model System for the Study of Tri-Trophic Associations (*American Museum of Natural History*) http://tcn.amnh.org
- North American Lichens and Bryophytes: Sensitive Indicators of Environmental Quality and Change (*University of Wisconsin – Madison*) http://symbiota.org/nalichens/index.php http://symbiota.org/bryophytes/index.php (plus 2 PENs)
- Digitizing Fossils to Enable New Syntheses in Biogeography - Creating a PALEONICHES-TCN (*University of Kansas*)
- The Macrofungi Collection Consortium: Unlocking a Biodiversity Resource for Understanding Biotic Interactions, Nutrient Cycling and Human Affairs (*New York Botanical Garden*)
- Mobilizing New England Vascular Plant Specimen Data to Track Environmental Change (*Yale University*)
- Southwest Collections of Anthropods Network (SCAN): A Model for Collections Digitization to Promote Taxonomic and Ecological Research (*Northern Arizona University*) http://hasbrouck.asu.edu/symbiota/portal/index.php

New as of 1 July 2013

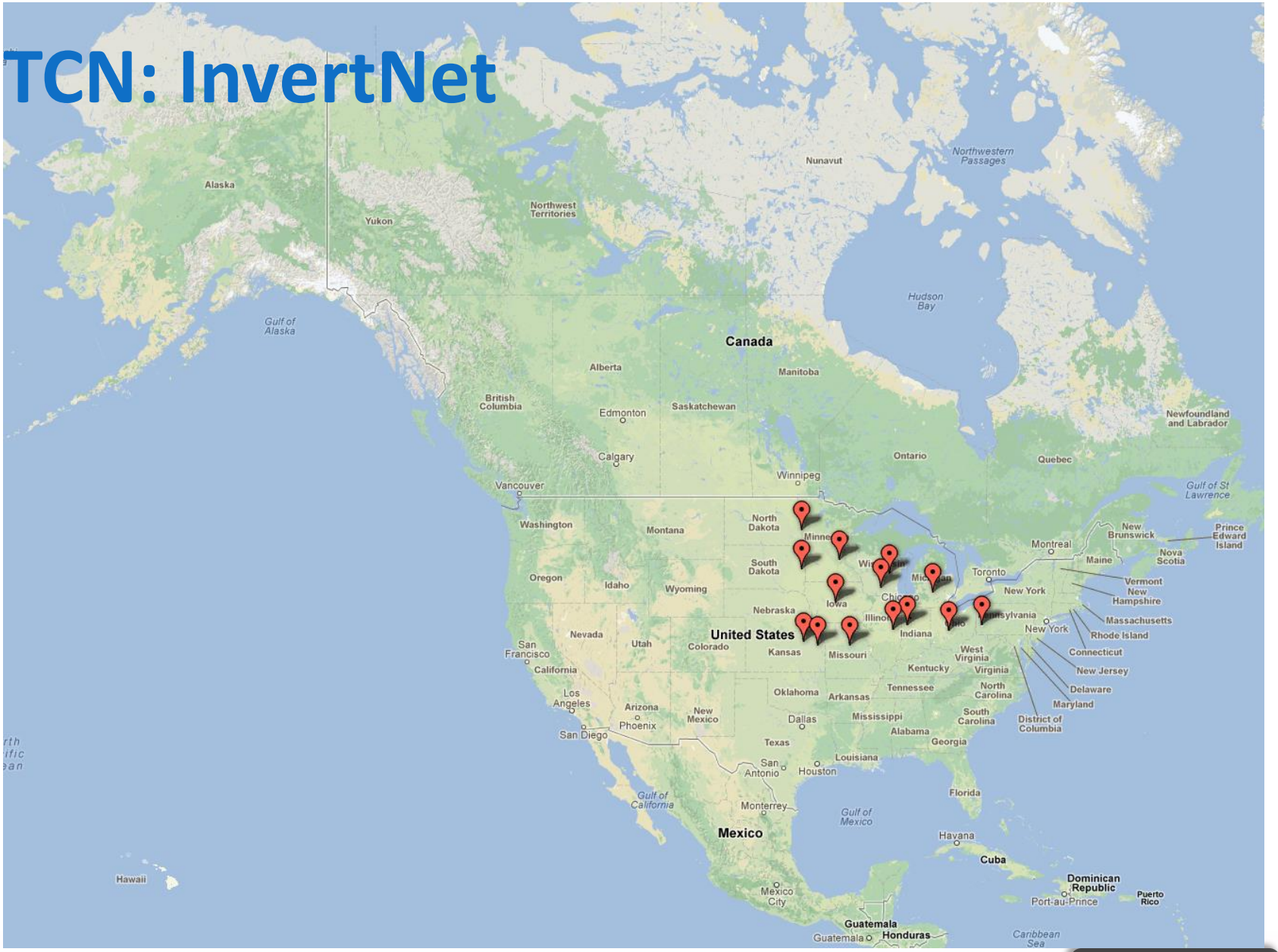
- iDigPaleo: Fossil Insect Collaborative: A Deep-Time Approach to Studying Diversification and Response to Environmental Change
- Developing a Centralized Digital Archive of Vouchered Animal Communication Signals
- The Macroalgal Herbarium Consortium: Accessing 150 Years of Specimen Data to Understand Changes in the Marine/Aquatic Environment

National Resource (iDigBio), Thematic Collection Networks (TCNs)

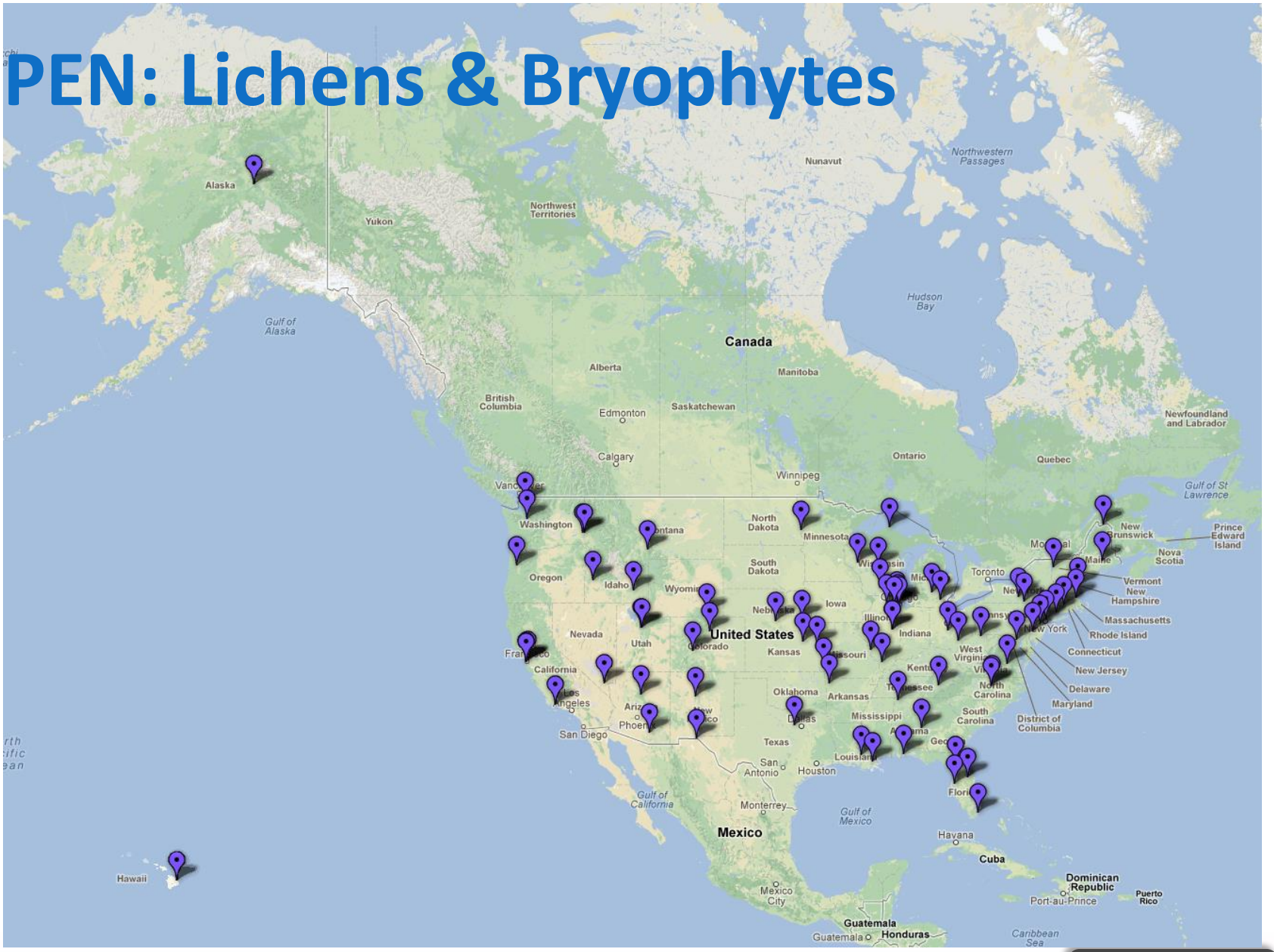


To date: 10 TCNs, 2 PENs, 160+ participating institutions, 49 states

TCN: InvertNet

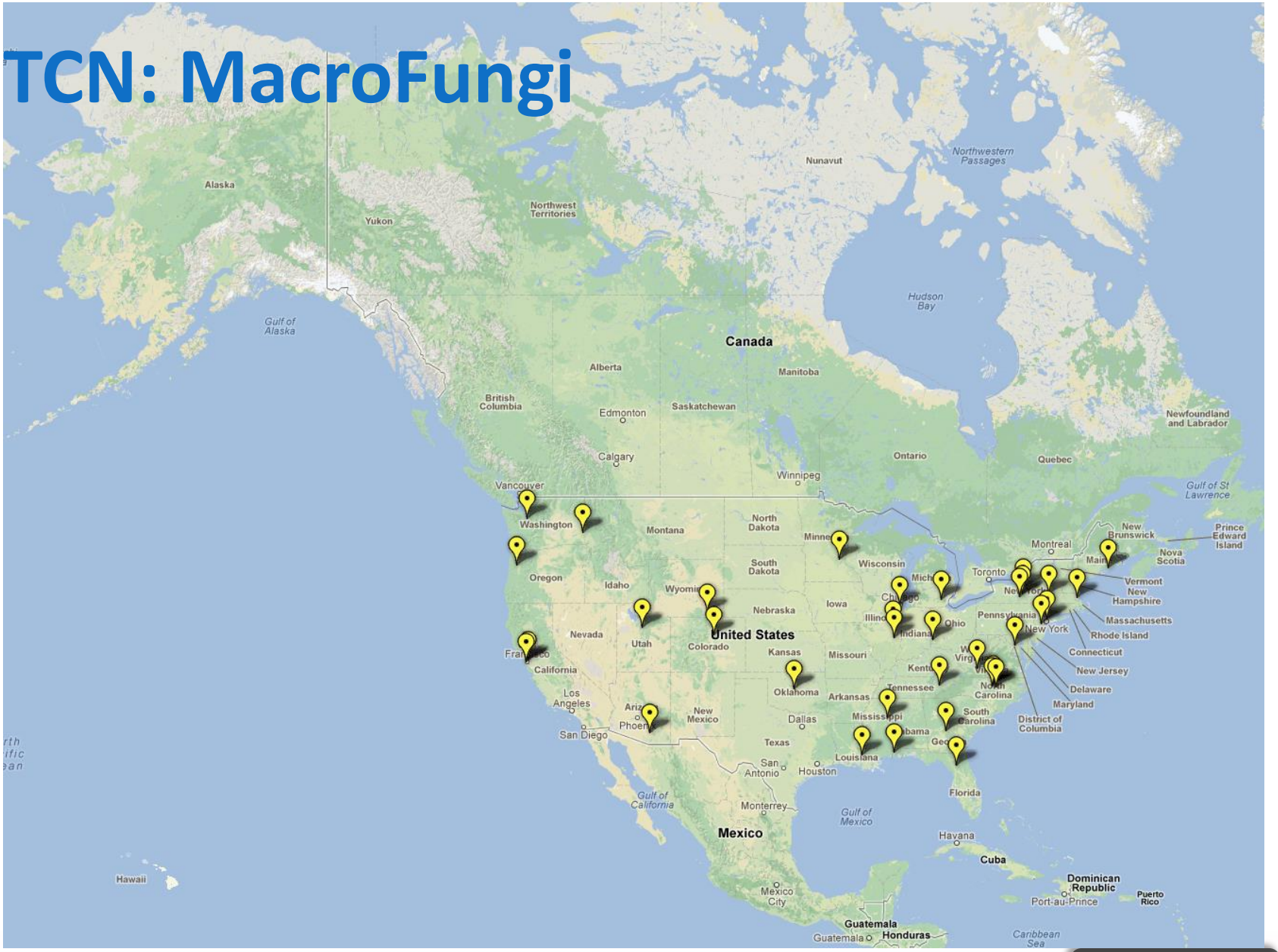


PEN: Lichens & Bryophytes



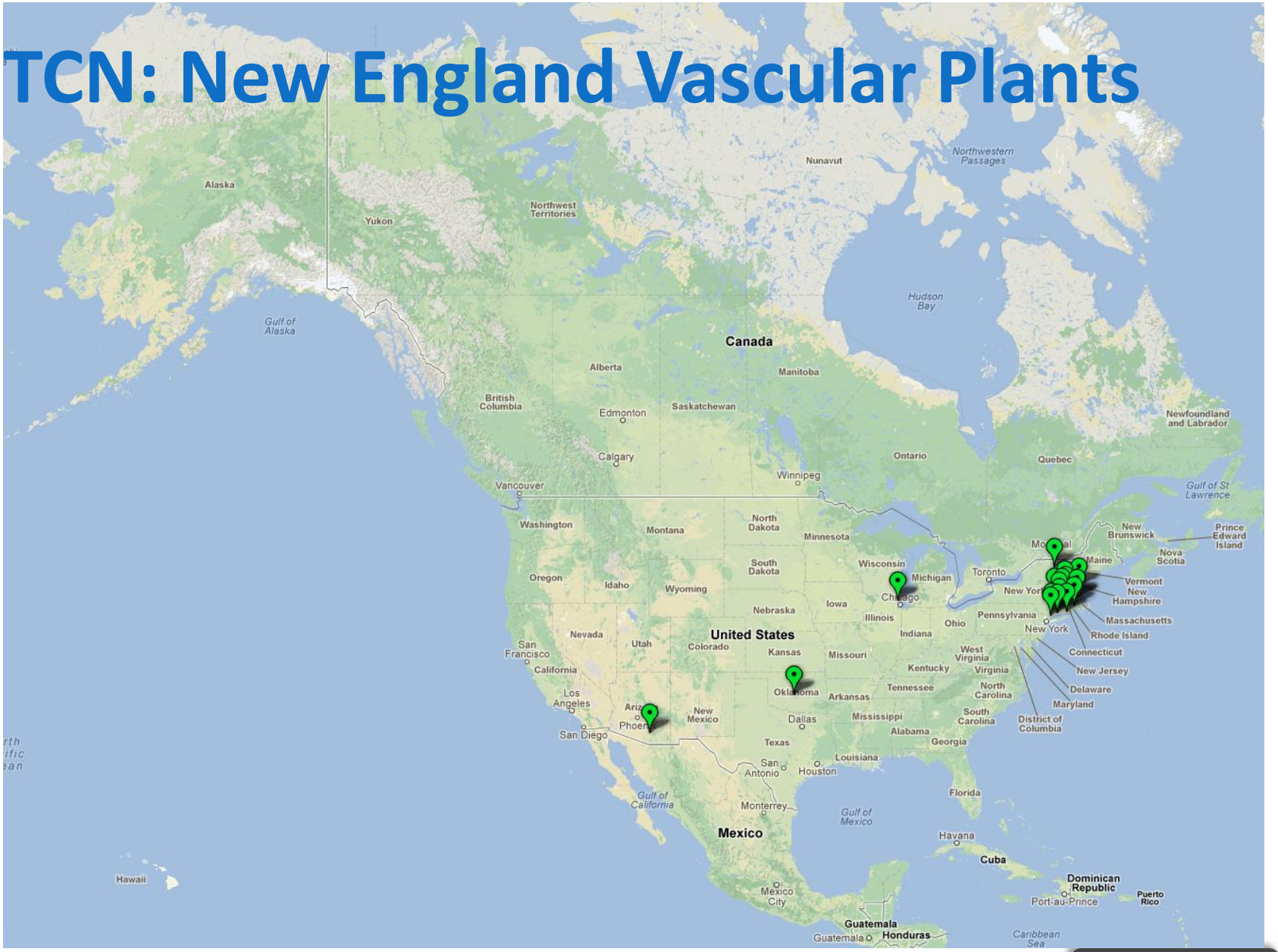
iDigBio
Integrated Digitized Biocollections

TCN: MacroFungi

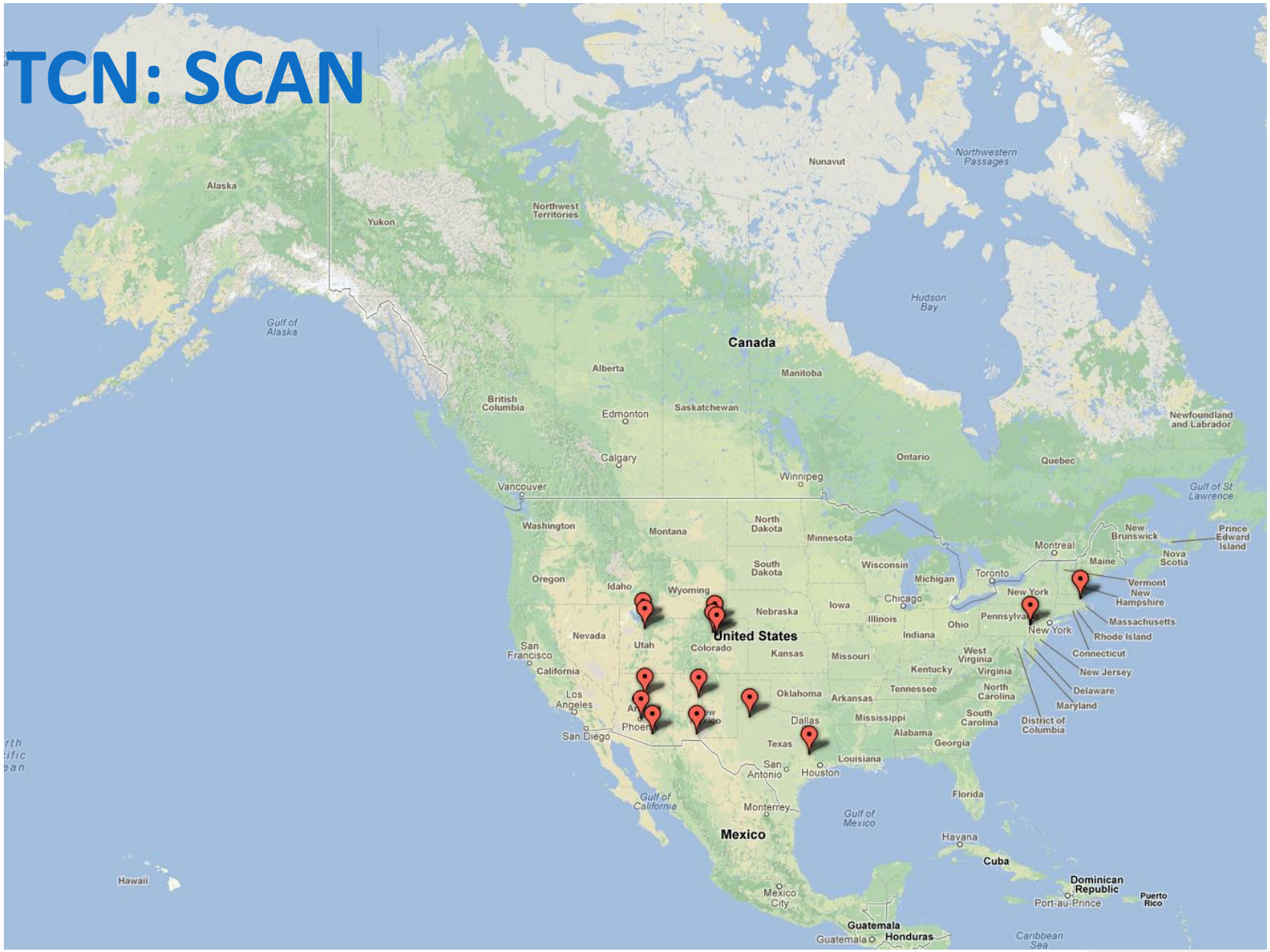


iDigBio
Integrated Digitized Biocollections

TCN: New England Vascular Plants



TCN: SCAN

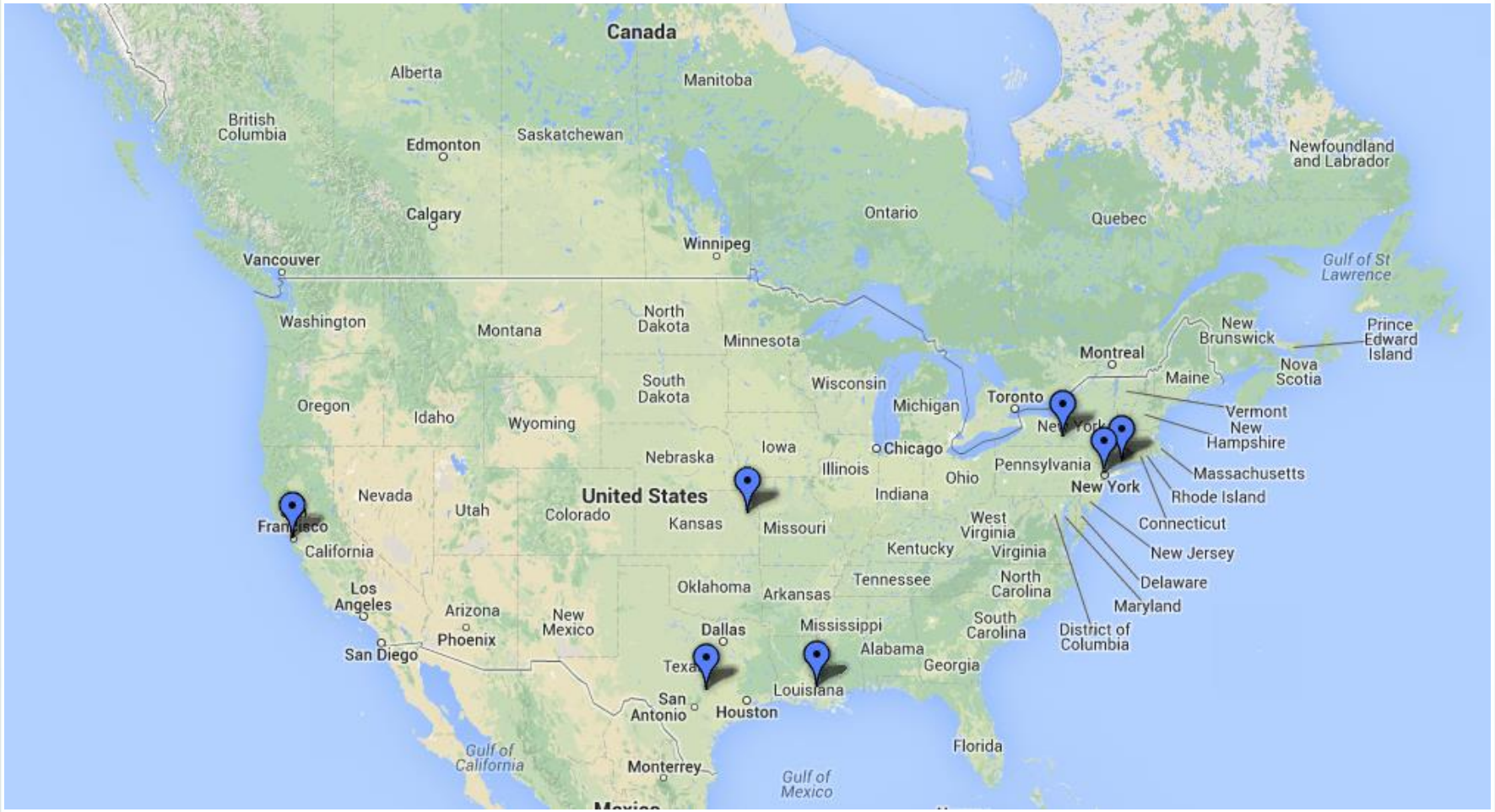

Alaska
Yukon
Northwest Territories
Nunavut
Northwestern Passages
Hudson Bay
Gulf of Alaska
Canada
Alberta
Manitoba
British Columbia
Edmonton
Saskatchewan
Calgary
Ontario
Quebec
Newfoundland and Labrador
Vancouver
Winnipeg
Gulf of St Lawrence
Washington
Montana
North Dakota
Minnesota
New Brunswick
Prince Edward Island
Montreal
Nova Scotia
Maine
Oregon
Idaho
Wyoming
South Dakota
Wisconsin
Michigan
Toronto
Vermont
New Hampshire
New York
Chicago
Iowa
Nebraska
Illinois
Pennsylvania
Massachusetts
Ohio
Nevada
Utah
Colorado
United States
Indiana
Rhode Island
San Francisco
Kansas
Missouri
West Virginia
Connecticut
California
Kentucky
Virginia
New Jersey
Los Angeles
Oklahoma
Arkansas
Tennessee
North Carolina
Delaware
Maryland
San Diego
Dallas
Mississippi
South Carolina
District of Columbia
Alabama
Georgia
Texas
Louisiana
San Antonio
Houston
Florida
Gulf of California
Monterrey
Gulf of Mexico
Mexico
Havana
Cuba
Hawaii
Dominican Republic
Port-au-Prince
Puerto Rico
Mexico City
Guatemala
Honduras
Caribbean Sea

TCN: PALEONICHES

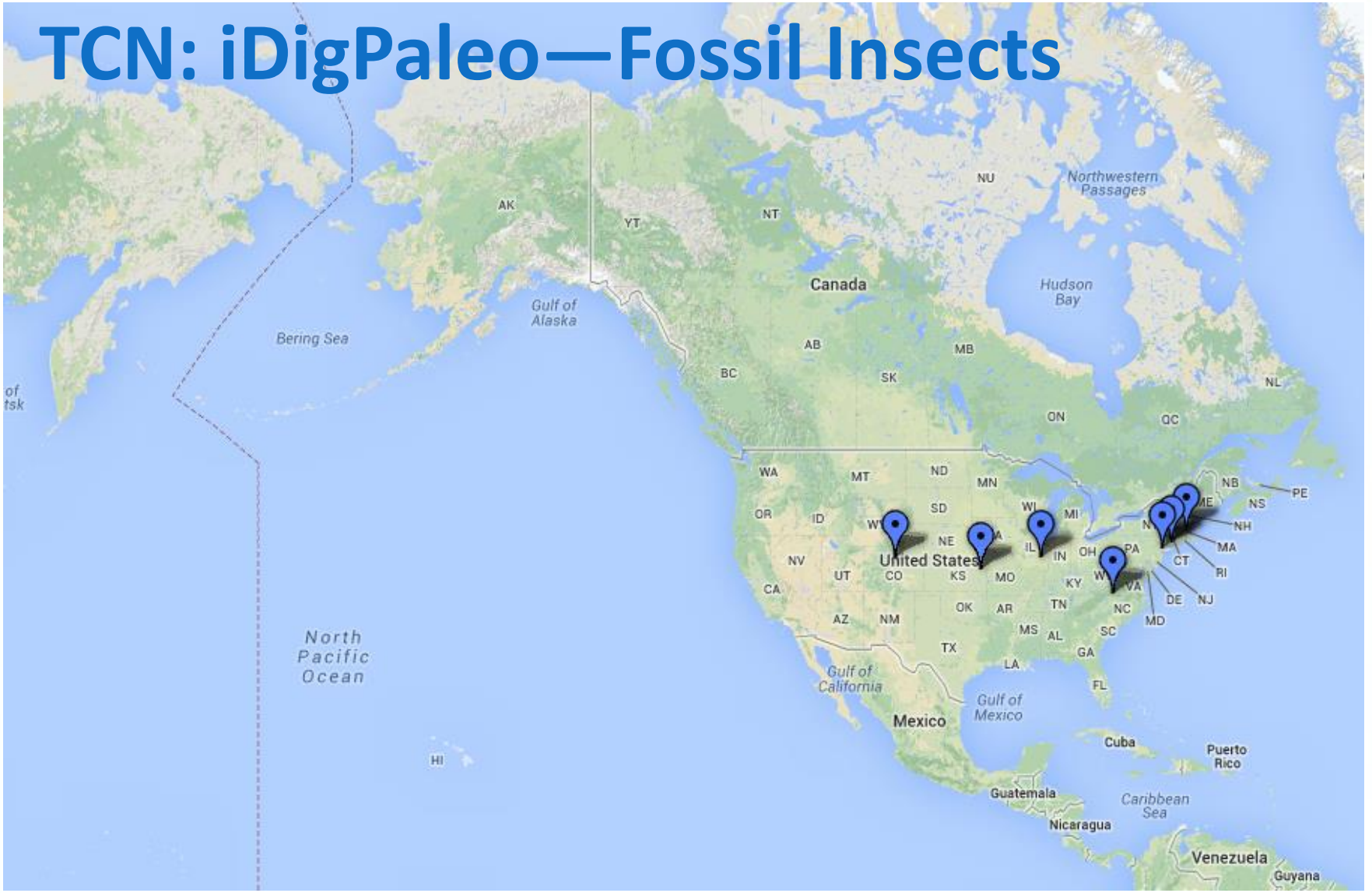
TCN: Tri-Trophic

iDigBio
Integrated Digitized Biocollections

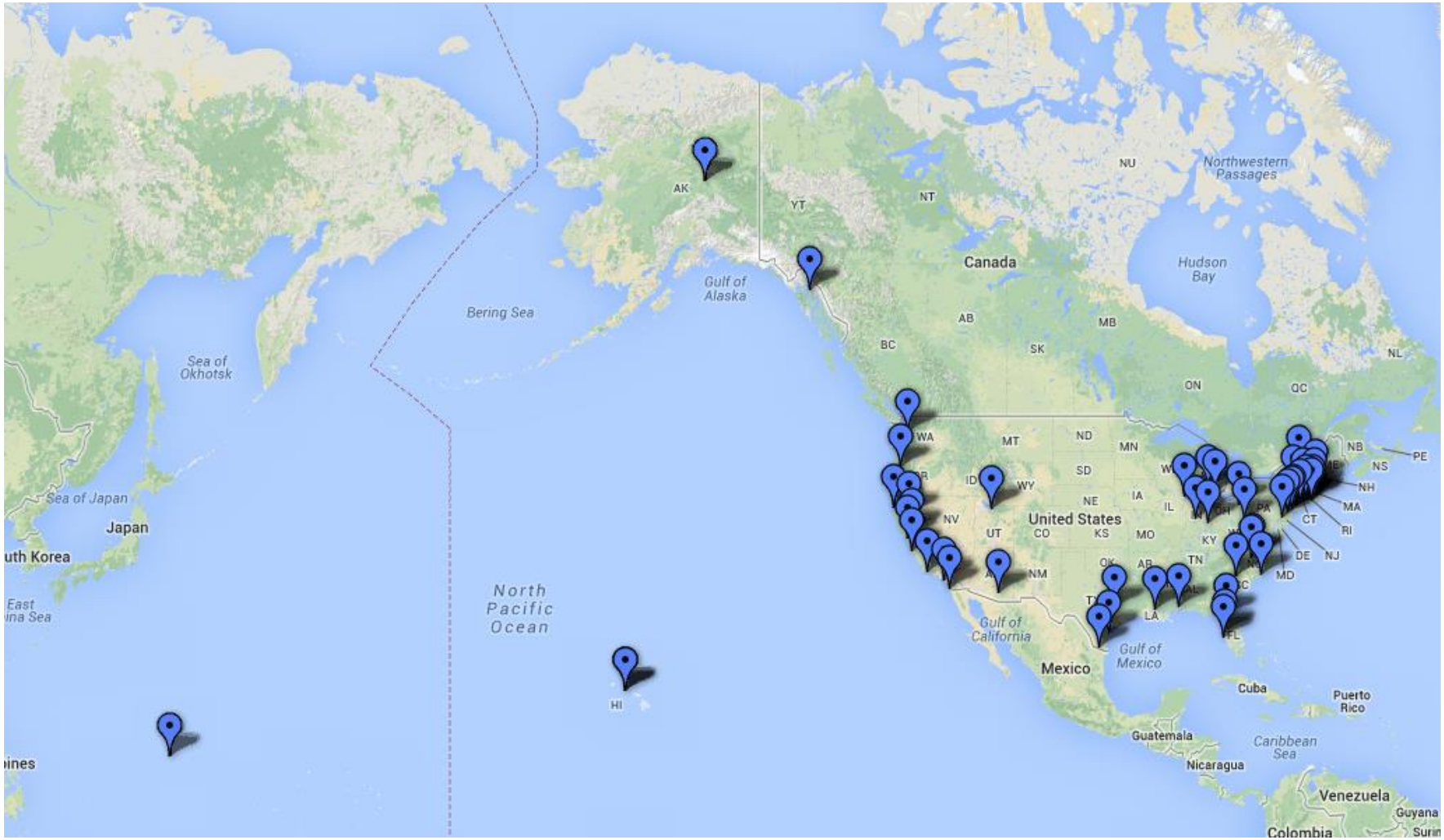
TCN: Animal Sounds



TCN: iDigPaleo—Fossil Insects

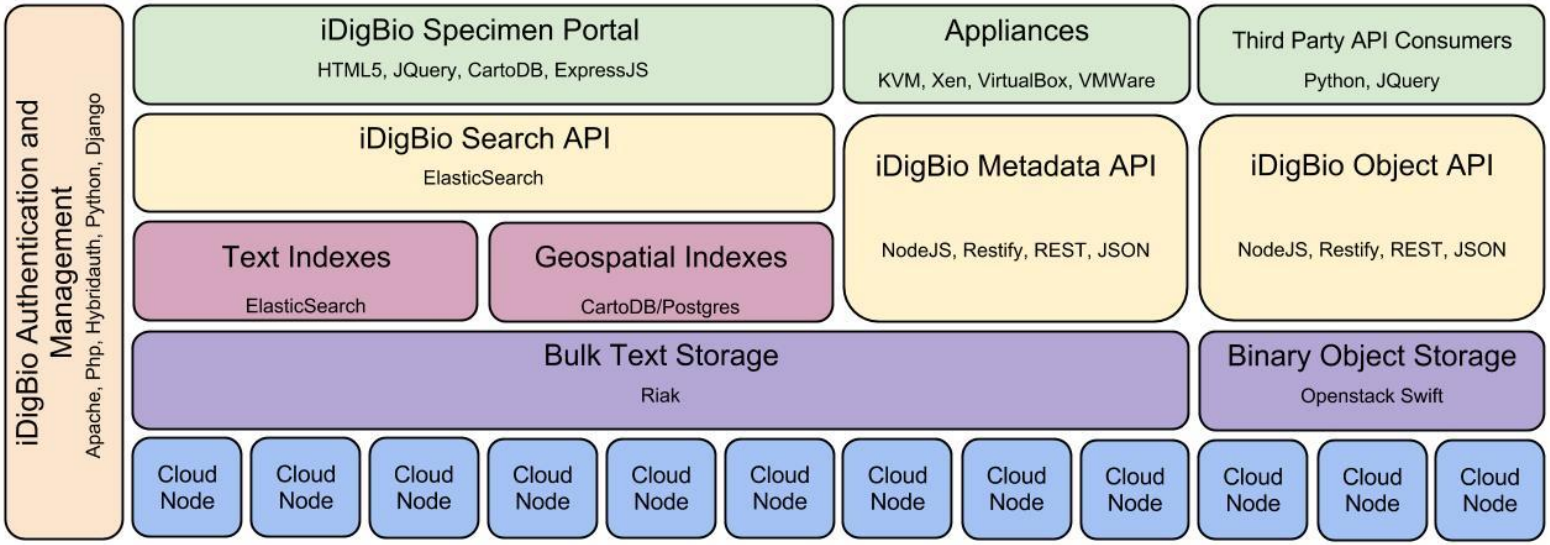


TCN: Macroalgal



Building the iDigBio Cloud

- Cloud-based strategy
  - Providing useful services/APIs (programmatic and web-based Application Programming Interface)
  - Federated scalable object storage and information processing
  - Digitization-oriented virtual appliances
  - Reliance on standards, proven solutions, and sustainable software
- Continuous consultation with stakeholders
  - Surveys, working groups, interest groups, workshops, person-to-person



Key Features of iDigBio

- Ingest all contributed data with emphasis on use of GUIDs, no restrictions
- Maintain persistent datasets and versioning, allowing new and edited records to be uploaded as needed while preserving existing records
- Ingest textual specimen records, plus associated still images, video, audio, and other media (or links to these resources as determined by the provider)
- Ingest linked documents and associated literature, including field notes, ledgers, monographs, related specimen collections, etc.
- Provide virtual annotation capabilities and track annotations back to the originating collection (collaborating with FilteredPush)
- Facilitate sharing and integration of data relevant to biodiversity research
- Provide computational services for biodiversity research

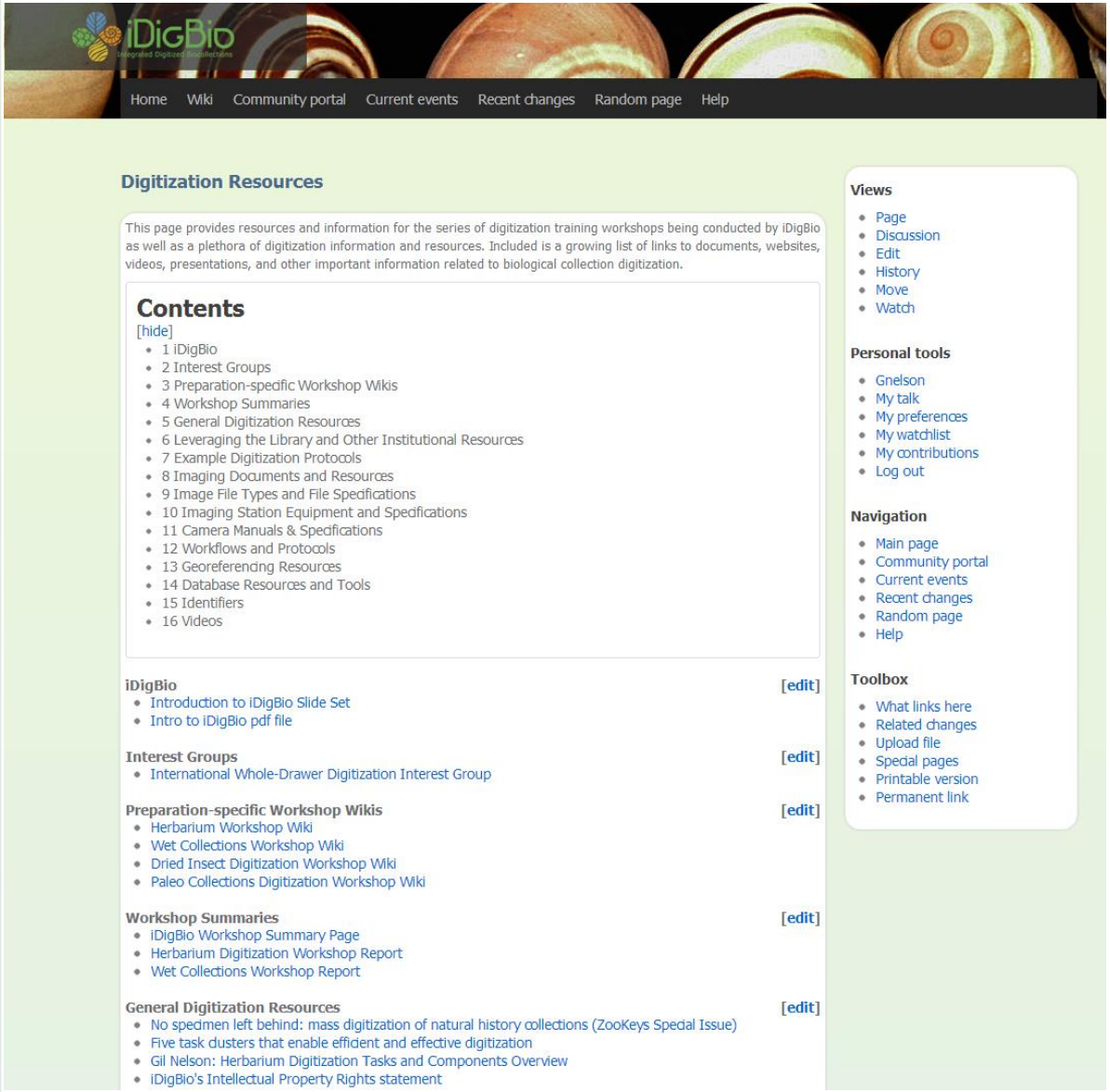
Recent, Ongoing, Upcoming Activities

- Assessment of common and effective digitization practices (paper in *ZooKeys*)
- Working groups
  - Minimum information for scientific collections working group (MISC)
  - Digitization workflows working groups
  - Georeferencing
  - Optical character recognition (OCR)
  - Biodiversity Informatics Manager working group
- Workshops - year 2:
  - > 150 institutions, 9 workshops, 3 symposia
  - 368 sponsored participants
  - Video archives on Vimeo, live streaming for remote participation
  - New model this year: train the trainer
  - Series of digitization training workshops (herbaria, wet collections, entomology, paleontology, fluid-preserved invertebrate imaging, small herbaria,)
- Server hosting: 8 virtual machines, TCN support
- Specimen data portal and website – continuous improvements
- Call for appliances, frequent opinion surveys

Digitization Workshops

In March 2012, the Steering Committee established a series of preparation-specific digitization training workshops focused on helping collections managers get started with and/or enhance local digitization programs, all to be held at host institutions.

- DROID (Developing Robust Object->Image->Data, May 2012)
- Herbarium digitization (Valdosta State, September 2012)
- Fluid-preserved collections digitization (U. Kansas, March 2013)
- Dried insect collections digitization (Field Museum, April 2013)
- Collections Digitization (West Virginia, ASB, April 2013)
- Imaging fluid-preserved invertebrates (U. Michigan, September 2013)
- Paleontology digitization (Yale Peabody Museum, September 2013)
- Small Herbarium Digitization (Florida State University, December 2013)
- Broadening Biodiversity in the Biodiversity Sciences (Atlanta, January, 2014)
- Original Source Materials Digitization (Yale Peabody Museum, March 2014)
- Digitization in the South Pacific (Honolulu, March 2014)
- Recruiting and Retaining Small Collections in Digitization (Mt. Pleasant, MI, April 2014)



Developed a community-oriented digitization resources wiki in support of our workshops and to serve digitization-related information across all preparation types.

Established a digitization list serv to promote workshop follow-up as well as community discussion and sharing.

iDigBio

Integrated Digitized Biocollections