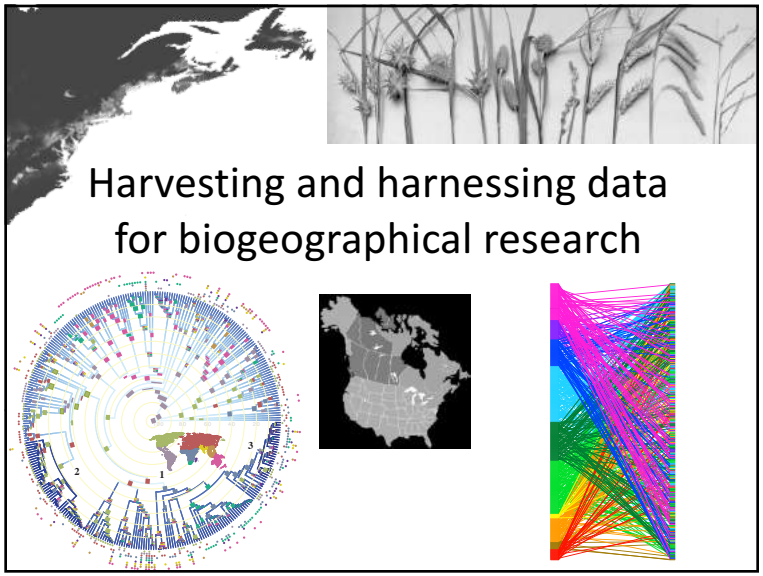
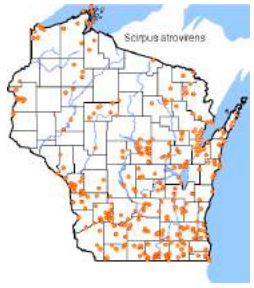


How do we know what grows where?

- inventories and surveys
  - natural areas, preserves, state forests, private properties
  - development permits
- herbarium/museum specimens
  - collectors often target specific places or taxa
  - records of what, where, when, and by whom

How do we know what grows where?

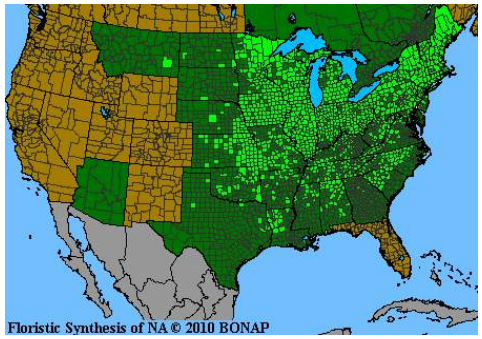
∑ Local observations = State Flora



black bulrush

How do we know what grows where?

∑ State Floras = National Flora



black bulrush

How do we know what grows where?

∑ National Floras = Global Flora

black bulrush

Who manages these data?

Wisconsin State Herbarium, University of Wisconsin

United States Environmental Protection Agency

U.S. Department of the Interior, March 3, 1849

Forest Service, U.S., Department of Agriculture

Missouri Botanical Garden

Michigan Natural Features Inventory

USDA

The Field Museum

Kew, Royal Botanic Gardens

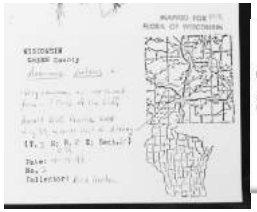
Wisconsin Dept. of Natural Resources

The New York Botanical Garden

Benefits of Digitization

- Preservation of rare/fragile specimens
- Easier to query a database than physical collections (e.g., WISFLORA)
- Efficient data sharing
- Collation yields powerful data repositories

Digitization of Distribution Data



Examples of data repositories

1. Wisconsin Flora – more on this later
2. North America Digital Flora
3. Biota of North America Project
4. Global Biodiversity Information Facility

North America Digital Flora

- Pros:
  - Great for generating species lists
  - Can query by species, attribute, or geography
  - Incorporates interesting GIS layers
- Cons:
  - Limited to wetland species
  - Geography search lacks precision

BONAP

- Pros:
  - Comprehensive of NA diversity
  - Great, easy to use distribution maps
- Cons:
  - Can't query the data
  - Can't download data
  - North America only
  - Poor resolution in Canada

GBIF

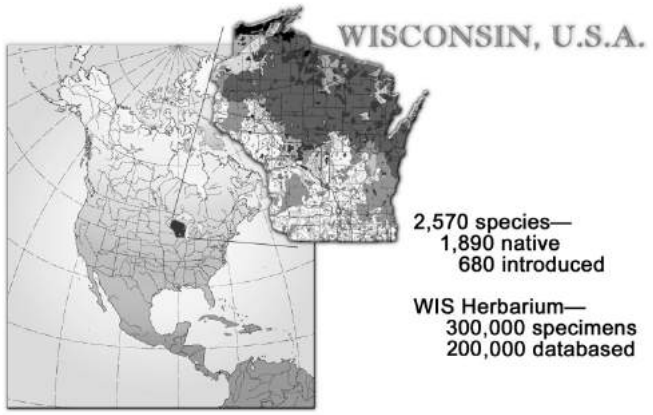
- Pros
  - Global data of plants, animals, fungi
  - Database is fully searchable
  - Data are downloadable
- Cons
  - Data are incomplete (WIS not in yet)
  - Some records are suspect

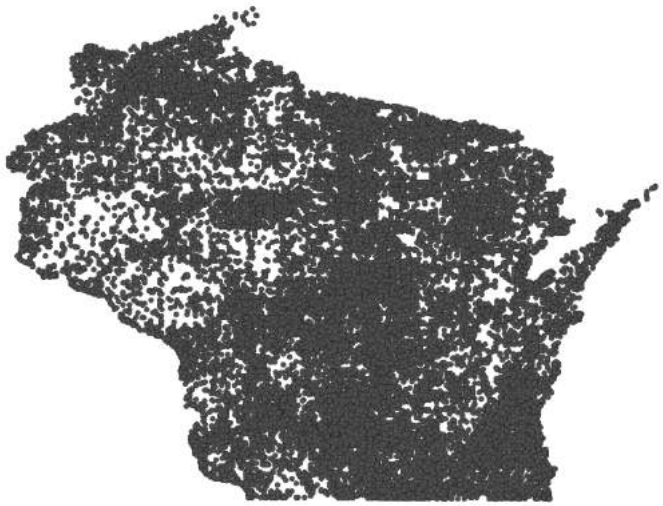
What can you do with these data?

- Test biogeographic patterns in Wiscosnin
- Historical biogeography
- Characterize species distributions
- Obtain species “climatic envelope”
- Identify areas exhibiting high levels of endemism
- Predict responses to global climate change
- Track and predict spread of invasive species

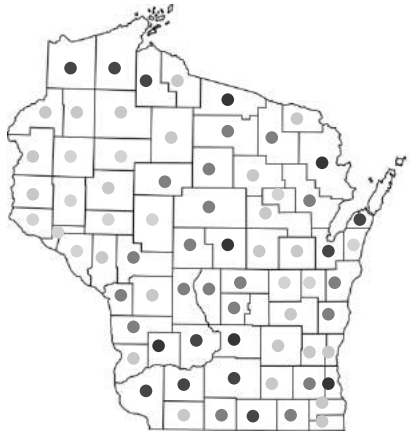
All without spending $1

Wisconsin Flora Mapping Project



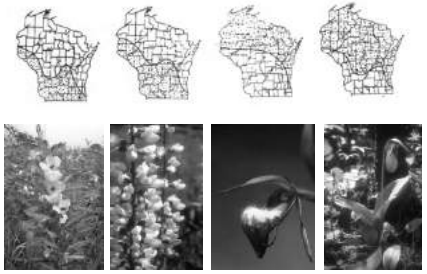


Wisconsin Flora Mapping Project

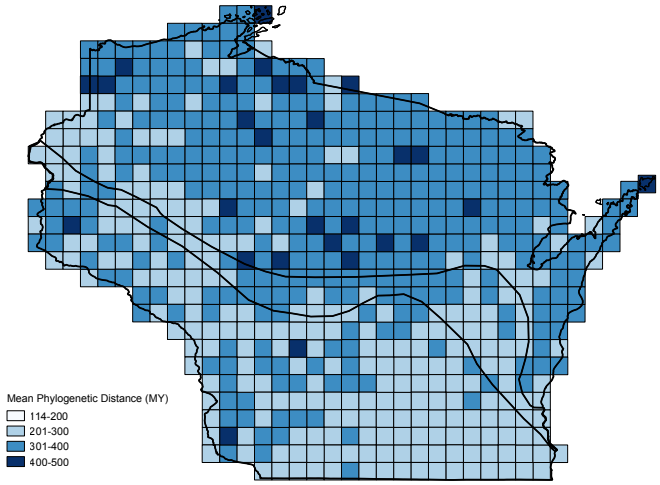


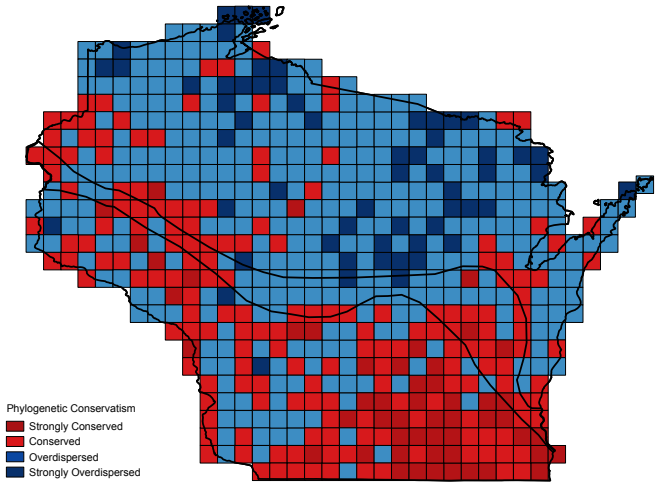
Wisconsin Flora Mapping Project

The tension zone separating the two provinces is based on the upper and lower limits of the southern and northern species, respectively.



Golden cassia | Wild indigo | Ram's head ladyslipper | Stemless ladyslipper





Historical Biogeography

- If you wanted, you could build a phylogeny of 300,000 organisms using 187 billion base pairs from Genbank

e.g., family Cyperaceae (sedges, bulrushes) – Spalink et al. 2016

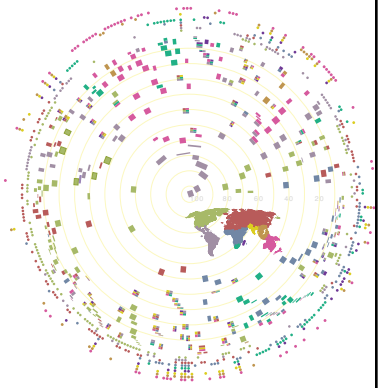
Historical Biogeography

- Build a phylogeny
  - Date the phylogeny

Data from Genbank , analyzed using BEAST

Historical Biogeography

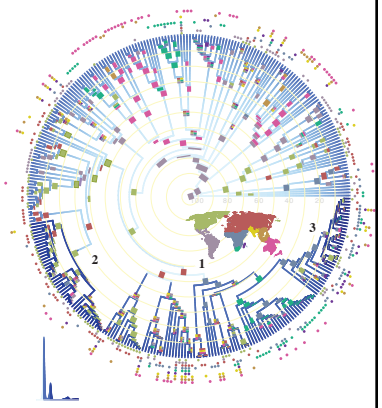
- Build a phylogeny
  - Date the phylogeny
  - Reconstruct areas inhabited by ancestors



Data from Genbank + eMonocot, analyzed using BEAST and BioGeoBEARS

Historical Biogeography

- Build a phylogeny
  - Date the phylogeny
  - Reconstruct areas inhabited by ancestors
  - Measure rates of diversification
  - Track evolution of morphological traits across time and space



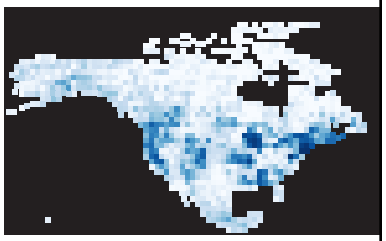
Data from Genbank + eMonocot, analyzed using BEAST,BioGeoBEARS and BAMM

High Resolution Species Distributions

- Characterize distributions of species for 1.45 million species using 460 million geo-referenced occurrences

High Resolution Species Distributions

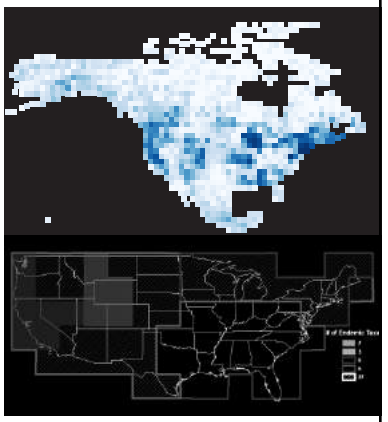
- Characterize distributions of species for 1.45 million species using 460 million geo-referenced occurrences
  - Identify species rich areas



Data from GBIF, analyzed in QGIS

High Resolution Species Distributions

- Characterize distributions of species for 1.45 million species using 460 million geo-referenced occurrences
  - Identify species rich areas
  - Define regions with high levels of rare or endemic taxa



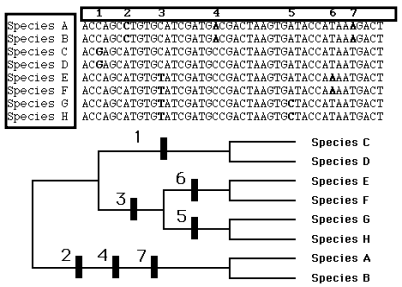
Data from GBIF, analyzed in QGIS and PAUP* - available in systematics lab

Parsimony Analysis of Endemicity (PAE)

Parsimony: typically used in phylogenetics…

Species

Characters



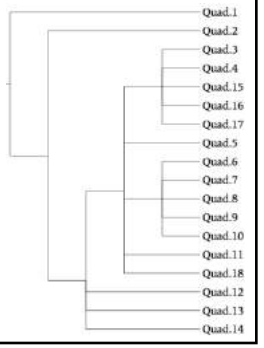
Example dataset from SR Downie

Parsimony Analysis of Endemicity (PAE)

…but, can be applied to PAE

	Species A	Species B	Species C	Species D	Species E	Species F	Species G	Species H
Area 1	1	0	1	0	0	0	0	0
Area 2	1	0	1	0	0	0	0	0
Area 3	0	1	0	1	0	0	0	0
Area 4	0	1	0	1	0	0	0	0
Area 5	0	0	0	0	0	0	0	0
Area 6	0	0	0	0	1	1	0	0
Area 7	0	0	0	0	1	1	0	0
Area 8	0	0	0	0	1	1	0	0
Area 9	0	0	0	0	1	1	0	0
Area 10	0	0	0	0	1	1	0	0
Area 11	0	0	0	0	0	0	0	0
Area 12	1	0	0	0	0	0	0	0
Area 13	1	0	0	0	0	0	0	0
Area 14	1	0	0	0	0	0	0	0
Area 15	0	1	0	1	0	0	0	0
Area 16	0	1	0	1	0	0	0	0
Area 17	0	1	0	1	0	0	0	0
Area 18	0	0	0	0	0	0	0	0

where geographic Areas are treated as taxa, and the presence/absence of the species serve as Characters

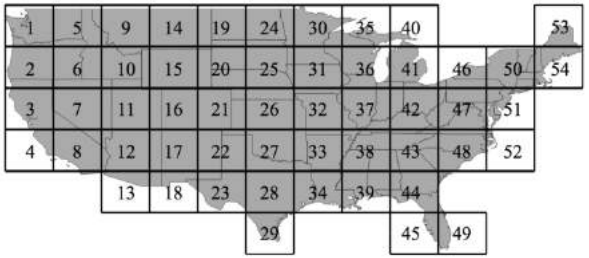


Case Study: finding areas of endemism in North American sedges

Step 1: Get distributional data of all species in N Am

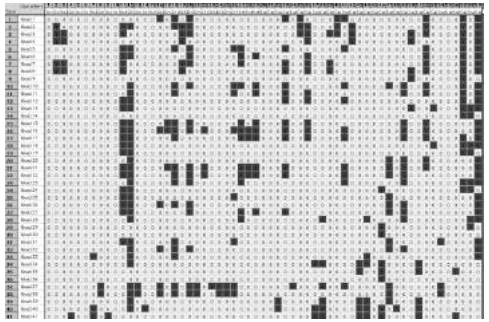
Case Study: finding areas of endemism in North American sedges

Step 2: Define geographic region and divide into quadrats



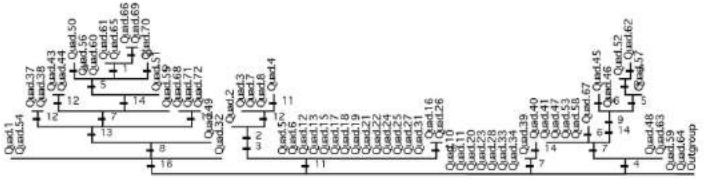
Case Study: finding areas of endemism in North American sedges

Step 3: Construct data matrix: absence = 0, presence = 1



Case Study: finding areas of endemism in North American sedges

Step 4: Conduct parsimony analysis



Case Study: finding areas of endemism in North American sedges

Step 5: Find clades consisting of 2+ endemic taxa with congruent distributions.

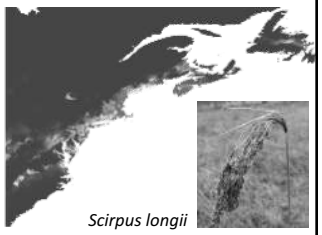
Ecological Niche Modeling

- Develop ecological niche models using waypoint and climate data (BioClim)

Scirpus longii

Ecological Niche Modeling

- Develop ecological niche models using waypoint and climate data
  - Identify suitable areas for endangered species

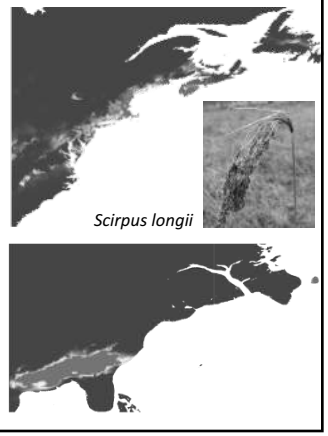


Scirpus longii

Data from GBIF and BioClim, analyzed with Maxent

Ecological Niche Modeling

- Develop ecological niche models using waypoint and climate data
  - Identify suitable areas for endangered species
  - Reconstruct distributions during Last Glacial Maximum

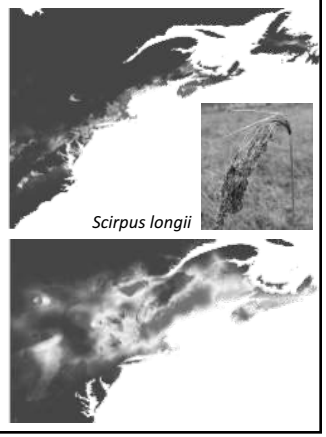


Scirpus longii

Data from GBIF and BioClim, analyzed with Maxent

Without spending $1, you can:

- Develop ecological niche models using waypoint and climate data
  - Identify suitable areas for endangered species
  - Reconstruct distributions during Last Glacial Maximum
  - Predict species distributions in 50 years

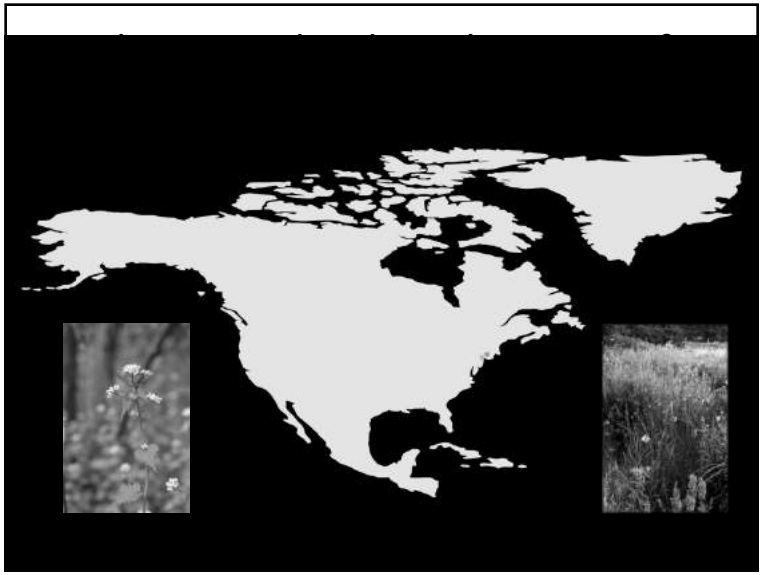


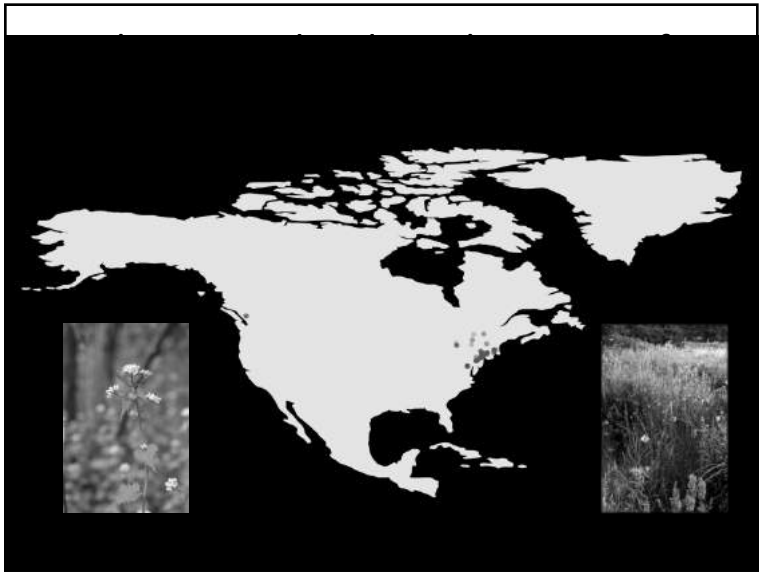
Scirpus longii

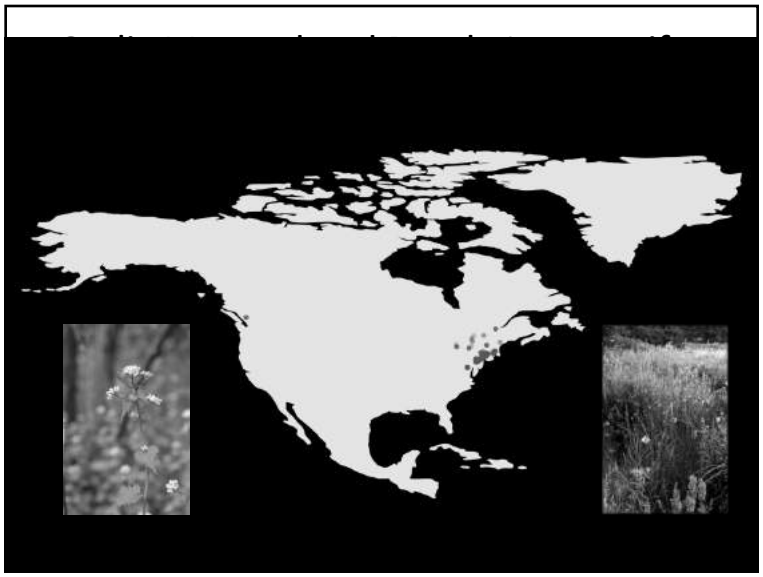
Data from GBIF and BioClim, analyzed with Maxent

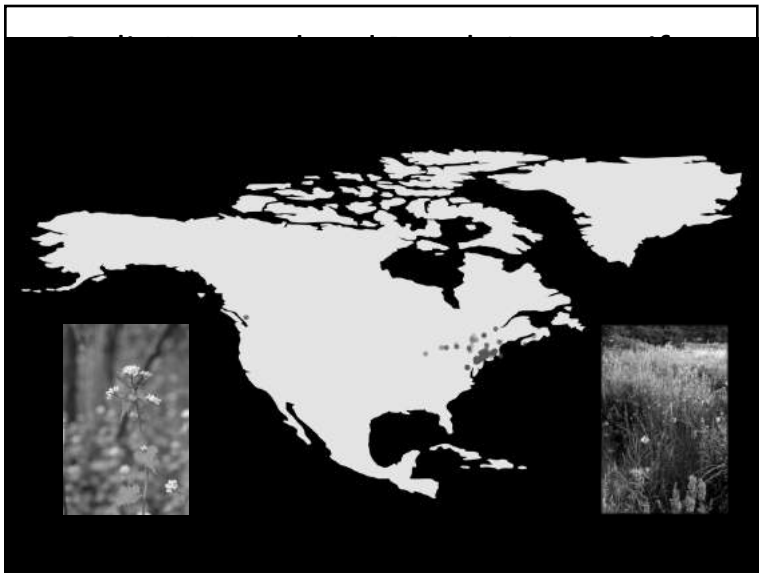
Track spread of invasive species

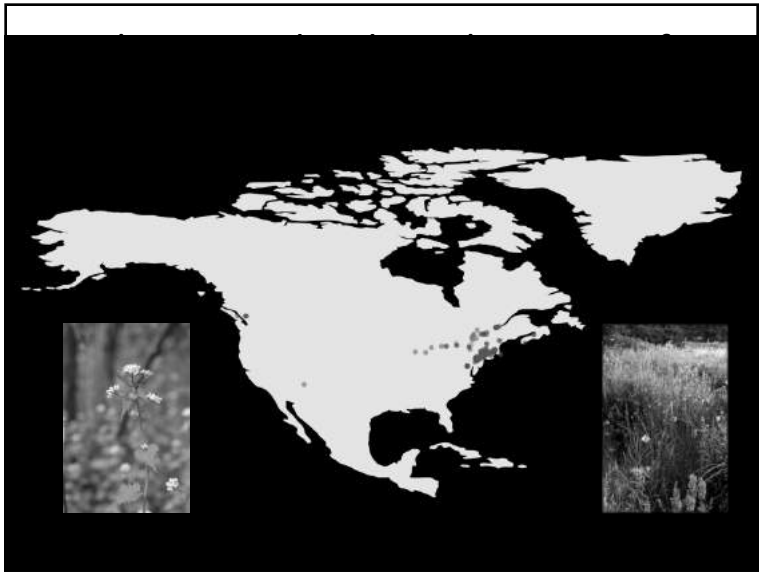
- Where did they originate?
- When did they get here?
- Where are they going?



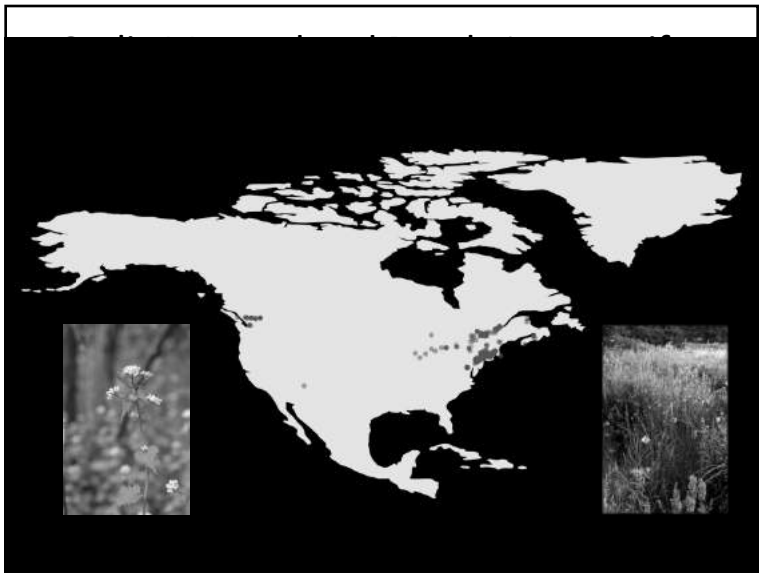


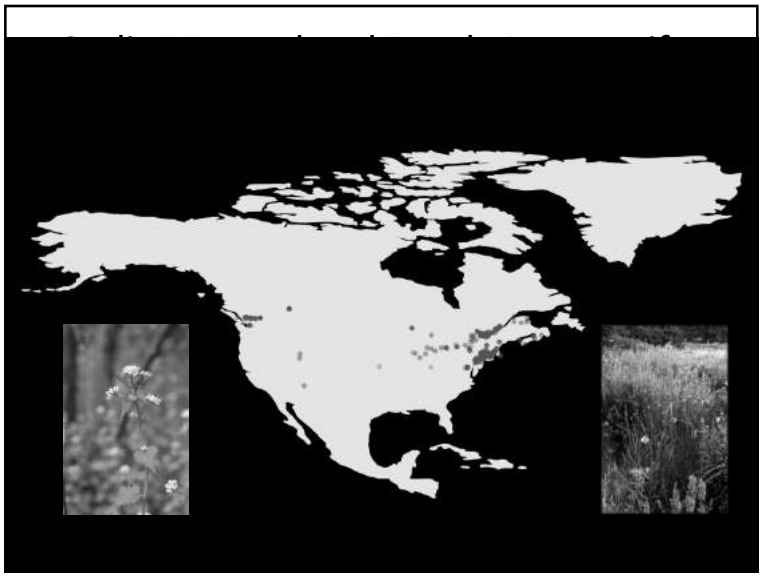


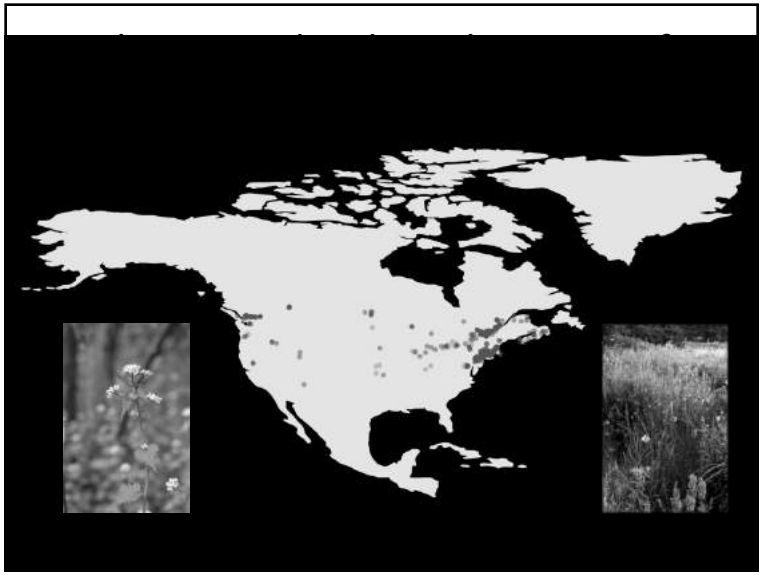


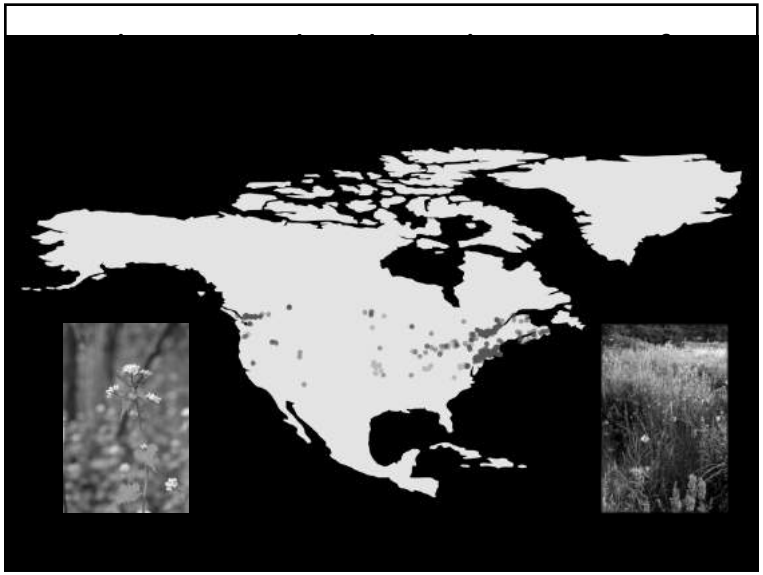


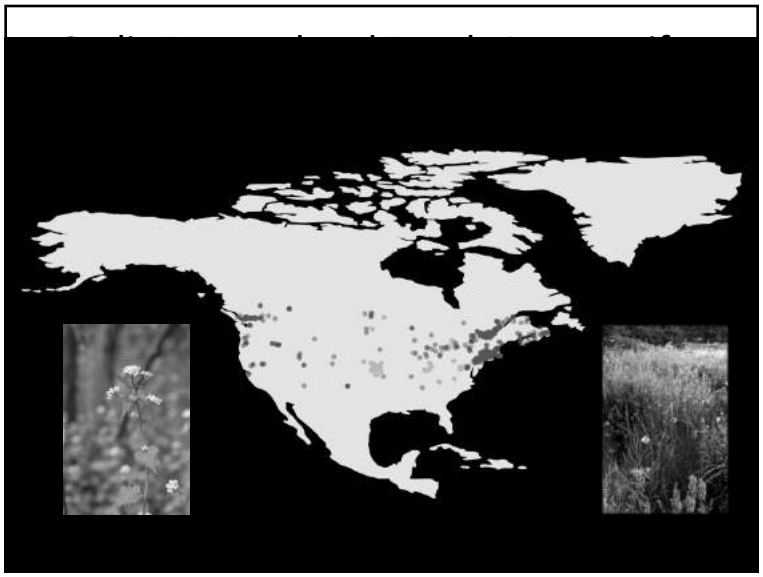
1927 Roadmap

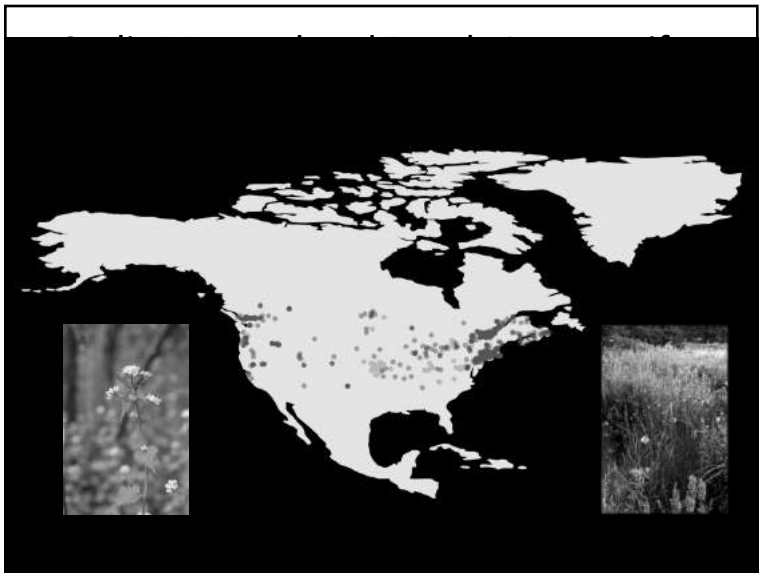


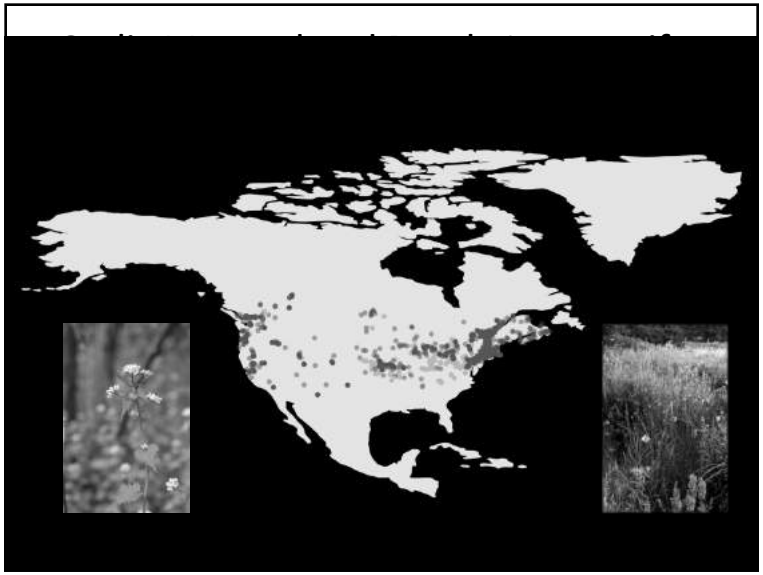








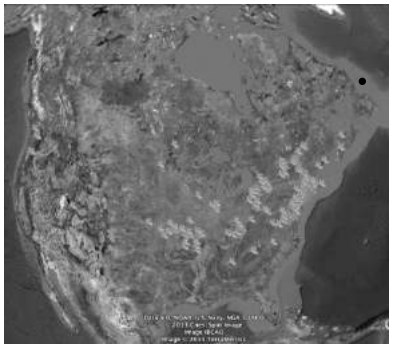




Some Cautions:

Distribution of *Alliaria petiolata* according to:

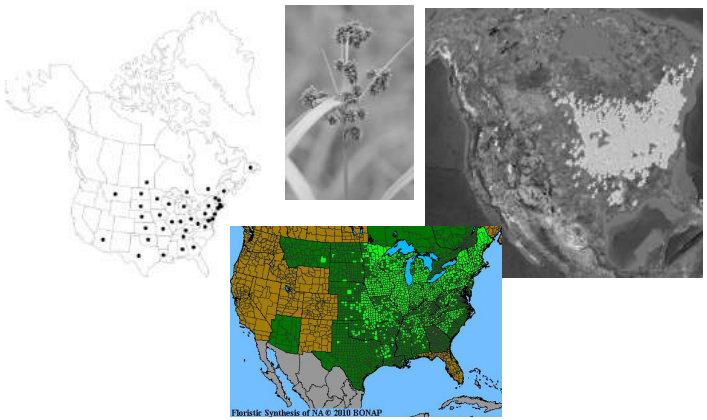
GBIF



WIS Flora

Some Cautions:

How do these maps of the distribution of *Scirpus atrovirens* differ? Why do they differ?



Additional Resources

- panbiogeography analysis (martitracks)
- simulate origin and spread of species (biogeosim)
- forecast species distributions (biomod)
- spatial analysis of diversity (biodiverse)
- list of free phylogenetic software here
- simulate historical island biogeography (shiba)
- spatial analysis in macroecology (SAM)
- predict and analyze distributions (GARP)