University *of* Wisconsin
Eau Claire

Detecting Introgression in *Helianthus* Populations in Wisconsin

Brandy Corwin and Nora Mitchell

Department of Biology
University of Wisconsin – Eau Claire

Introduction

Figure 1. Picture of sunflower sample.

- Sunflowers have incredible natural floral diversity and have their central diversity in North America.
- Introgression is the incorporation of genetic material from one species into another and can contribute to evolution.
- Rampant introgression can lead to the reduced fitness of hybridizing populations and can be a cause in their extinction. However, the introgression of a small number of advantageous alleles can increase population fitness. [1]
- Previous studies have detected introgression in other sunflower species in Wisconsin and there are both historical records of morphological intermediates and evidence for hybridization from crossing experiments for these species.
- There have also been many previous studies on different flora and fauna that have used ipyrad and TreeMix when looking at introgression. [2]
- The goal of this research was to detect if introgression occurred in the populations of three sunflower species in Wisconsin, *Helianthus giganteus*, *H. grosseserratus*, and *H. maximiliani*.

Methods

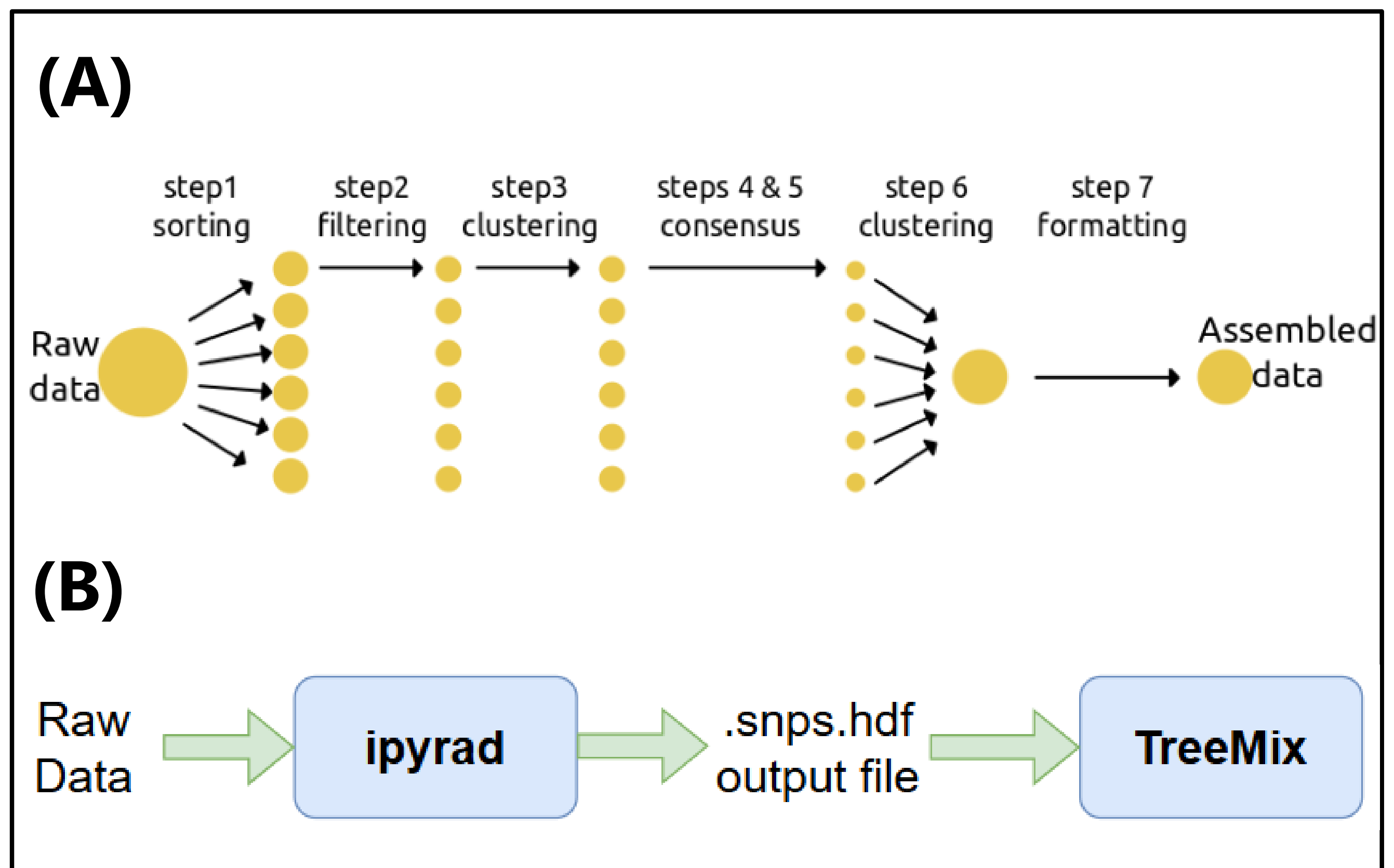

Figure 2. (A) Flowchart of ipyrad basic workflow. (B) Flowchart of tools used in determining introgression in sunflowers in this study.

- **ipyrad.** The ipyrad workflow filters and clusters raw genetic data to produce many outputs that are commonly useful. I used OnDemand to run ipyrad on raw RADseq data from 357 samples to clean, align, and detect SNPs across the dataset. This toolkit gives us many different output files, including a .snps.hdf file that was used as input for the TreeMix analysis. [3,4]
- **TreeMix.** I used OnDemand to run TreeMix on allele frequency data to infer population splits. The output is a maximum likelihood phylogenetic tree of the set of sunflower populations which is used to look for evidence of introgression. [3,4]

Results

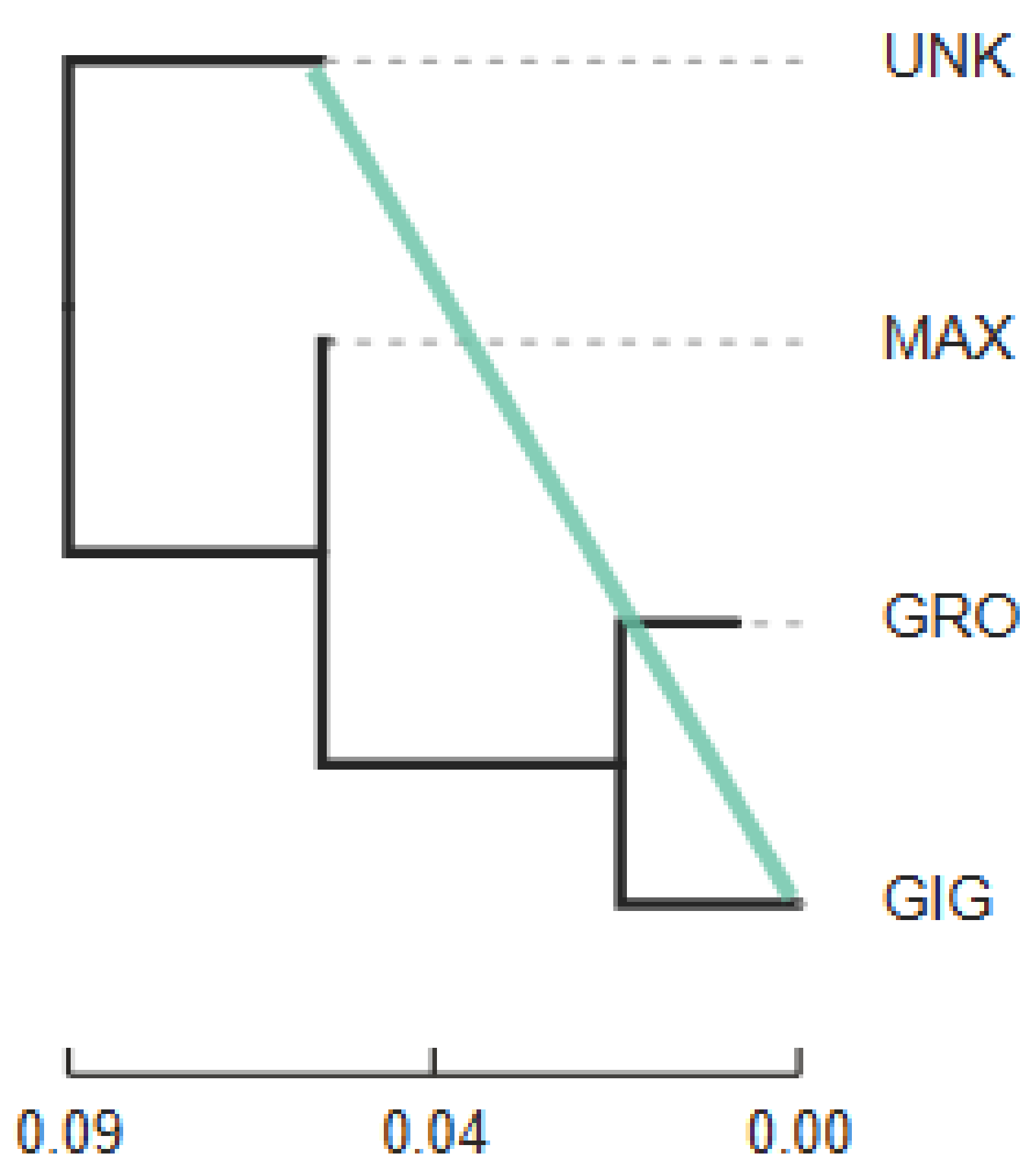

Figure 3. Resulting tree when TreeMix was run. Green line indicates a connection between the unknown samples and *Helianthus giganteus.*

Key Findings

- I found no detectable introgression among the sunflower populations in this study.
- It appears that the two samples of unknown populations are not of hybrid origin, and instead likely belong to *Helianthus giganteus.*

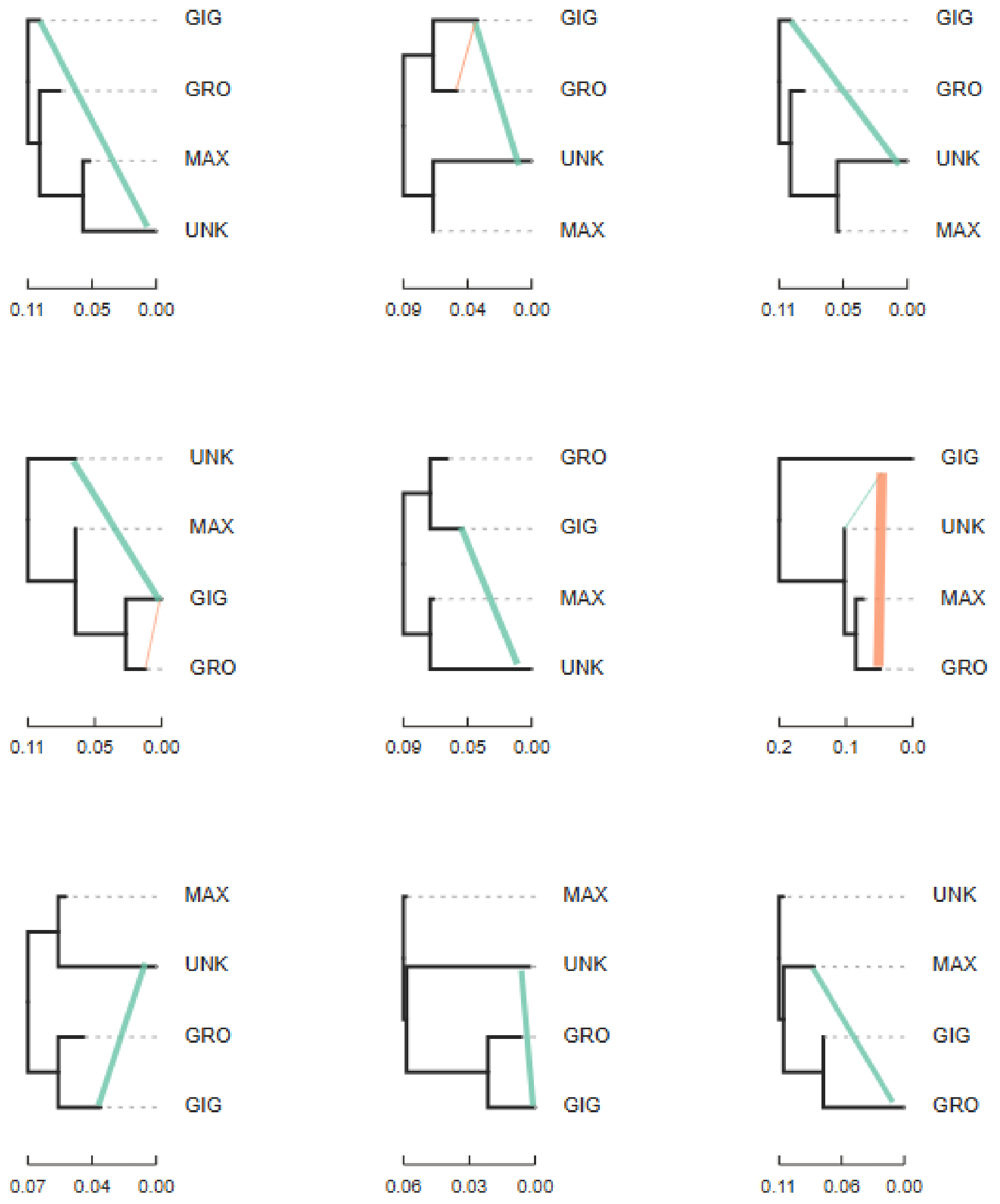

Figure 4. Tree graphs of nine iterations of subsamples of SNPs. Most of them are similar to the tree graph from the full data. The orange lines indicate a second connection between *H. giganteus* and *H. grosseserratus*, but as they are found in only three of the subsample iterations, there is less support for this connection.

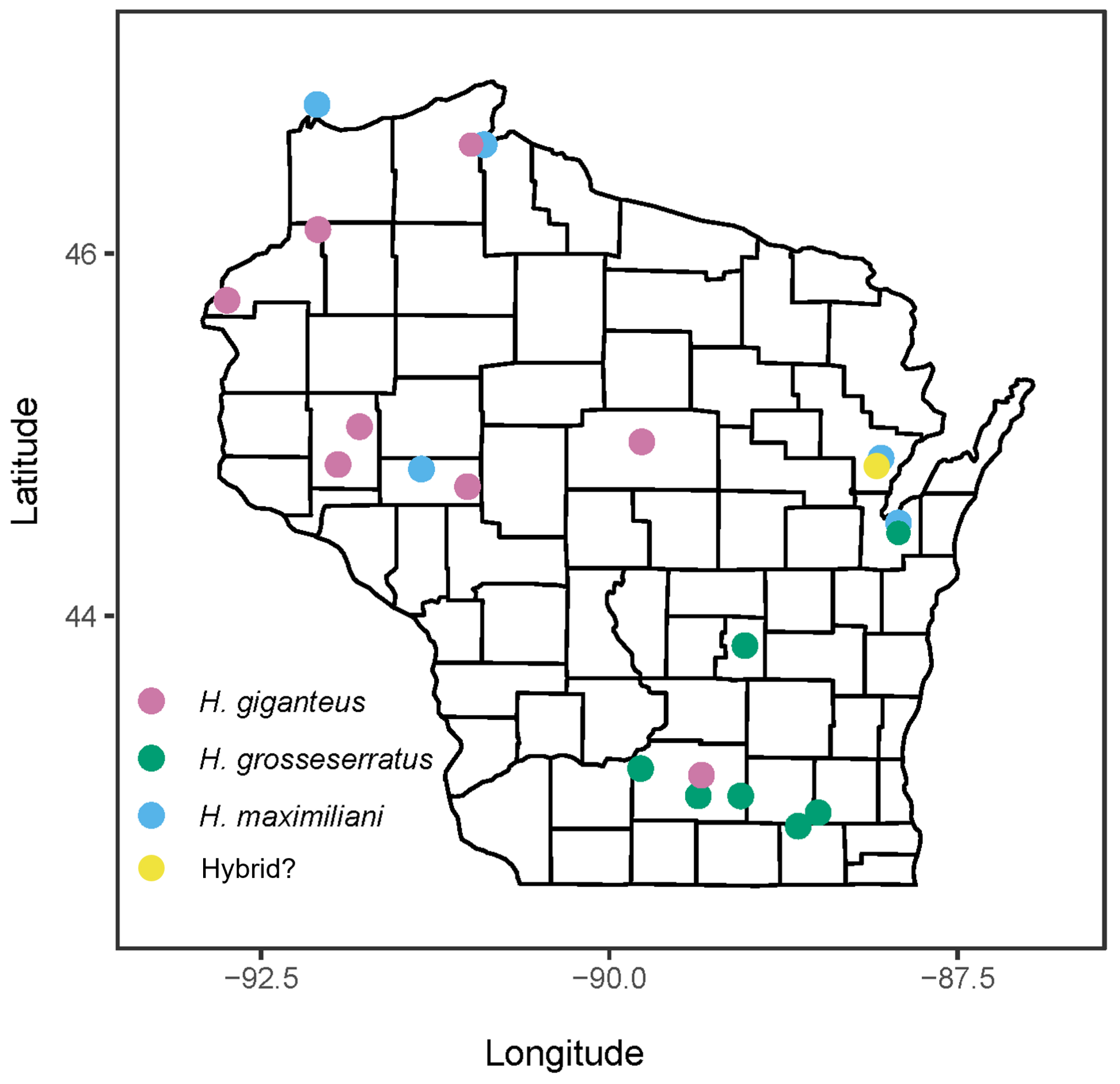

Figure 5. This map of Wisconsin shows the locations where the sunflower samples were collected.

Future Directions

- Run Admixture analyses using the program STRUCTURE to examine number of genetic clusters across these samples.
- In the future, we could also determine if the unknown samples are a part of *Helianthus giganteus*, or if they are a hybrid with a population not used in this study.

	total_filters	applied_order	retained_loci
total_prefiltered_loci	0	0	227306
filtered_by_rm_duplicates	3131	3131	224175
filtered_by_max_indels	1456	1456	222719
filtered_by_max_SNPs	703	452	222267
filtered_by_max_shared_het	124	97	222170
filtered_by_min_sample	215585	215585	6585
total_filtered_loci	220999	220721	6585

Figure 6. A table showing the number of loci caught by each filter.

References

1. Scascitelli M, Whitney K, Randell R et al. Genome scan of hybridizing sunflowers from Texas (*Helianthus annuus* and *H. debilis*) reveals asymmetric patterns of introgression and small islands of genomic differentiation. Molecular Ecology. 2010.
2. Thanou E, Kornilios P, Lymberakis P et al. Genomic and mitochondrial evidence of ancient isolations and extreme introgression in the four-lined snake. Current Zoology. 2020.
3. Hudak D et al. Open OnDemand: A web-based client portal for HPC centers. Journal of Open Source Software. 3(25), 622. 2018.
4. Eaton DAR & Overcast I. ipyrad: Interactive assembly and analysis of RADseq datasets. Bioinformatics. 2020.

Acknowledgements

The computational resources of the study were provided by the Blugold Center for High-Performance Computing under NSF grant CNS-1920220. Funding provided by the Milwaukee Public Museum Rundblad Fellowship. Thank you to University Printing Services for printing this poster.