phylogatR

PHYLOGEOGRAPHIC DATA AGGREGATION AND REPURPOSING

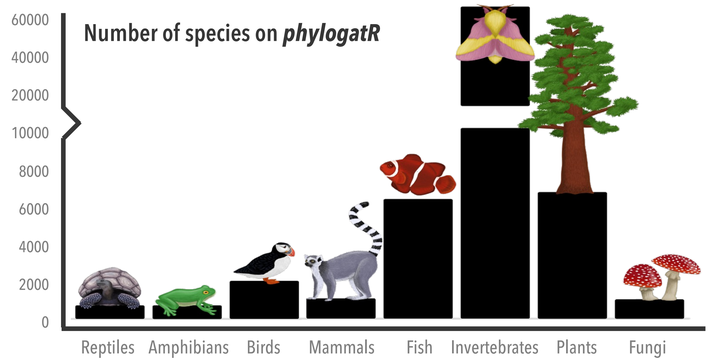
Danielle J. Parsons[1] | Sydney K. Decker[1] | Tara A. Pelletier[2] | Bryan C. Carstens[1] [1.Ohio State University 2.Radford University]

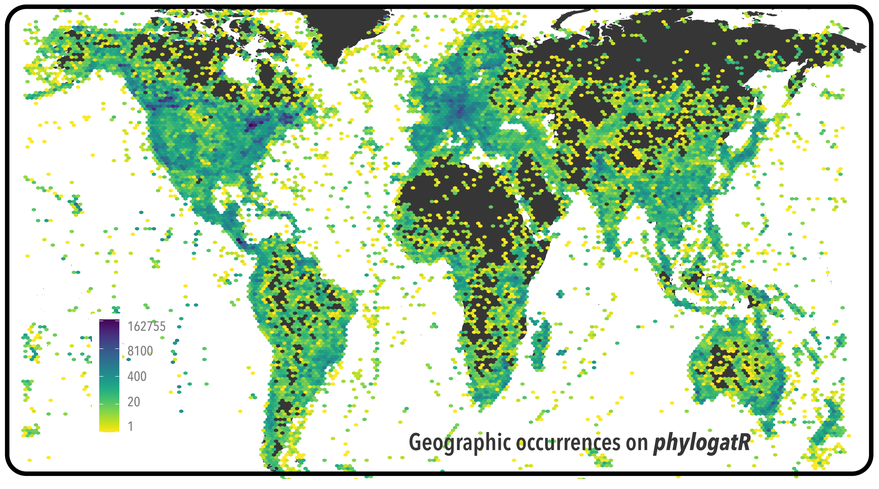
INTRODUCTION

Patterns of genetic diversity within species contain information about the history of that species, including how it responded to historical climate change. Researchers in many disciplines collect DNA sequence data from hundreds of samples and deposit these in open-source online data repositories, such as GenBank and BOLD. The existence of georeferenced DNA sequence data in databases can enable novel comparative analyses in ecology and evolution.

In order to facilitate these types of analyses on the largest possible scale from thousands of species, we developed software that parses data from several repositories of geographic and genetic data, organizes them under a taxonomic hierarchy, and produces data that are analysis ready.

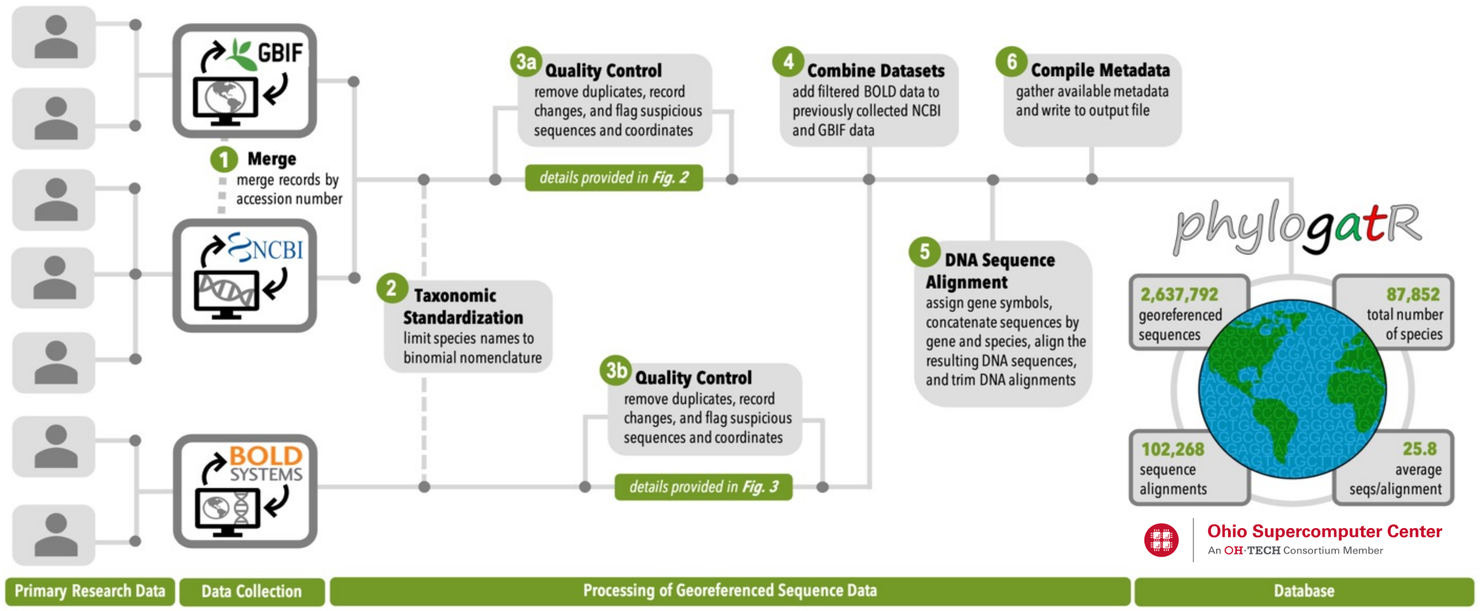
DATABASE AT A GLANCE





Data analysis can be conducted using R scripts or R Shiny apps from *phylogatR*. The database *phylogatR* is freely available via the Ohio Supercomputer Center

PIPELINE

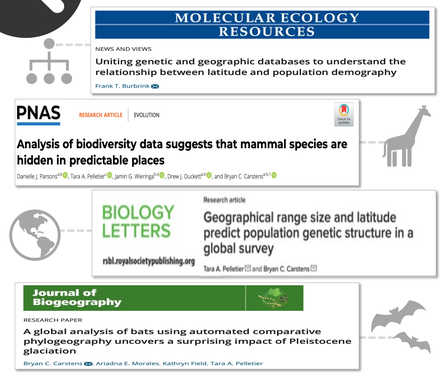


DATABASE GOALS

to 1) empower students to actively learn about computer code, genetics, and biodiversity science by repurposing genetic and geographic data and 2) enable basic scientific research in a discipline that is fundamentally about global change

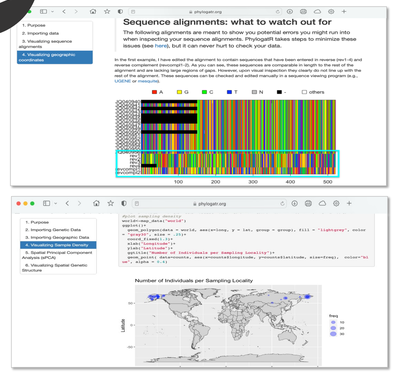
RESULTS

RESEARCH



Research using *phylogatR* is ongoing and includes several published projects.

EDUCATION



Educational modules and teaching resources for coding and data analysis are available.

OUTREACH

Interactive web applications are available to facilitate learning at all skill levels.