Conservation Genomics

Joel Sharbrough

July 28th, 2022

What is Conservation Genomics?

Applying genomics technologies to
species of conservation interest/concern

What is Conservation Genomics?

How can it help in preservation and protection of biodiversity?

Genomics as a tool for preserving biodiversity

- Monitoring/management
- Population profiling, threat assessments
- Characterizing disease transmission dynamics
- Genetic rescue as an intervention

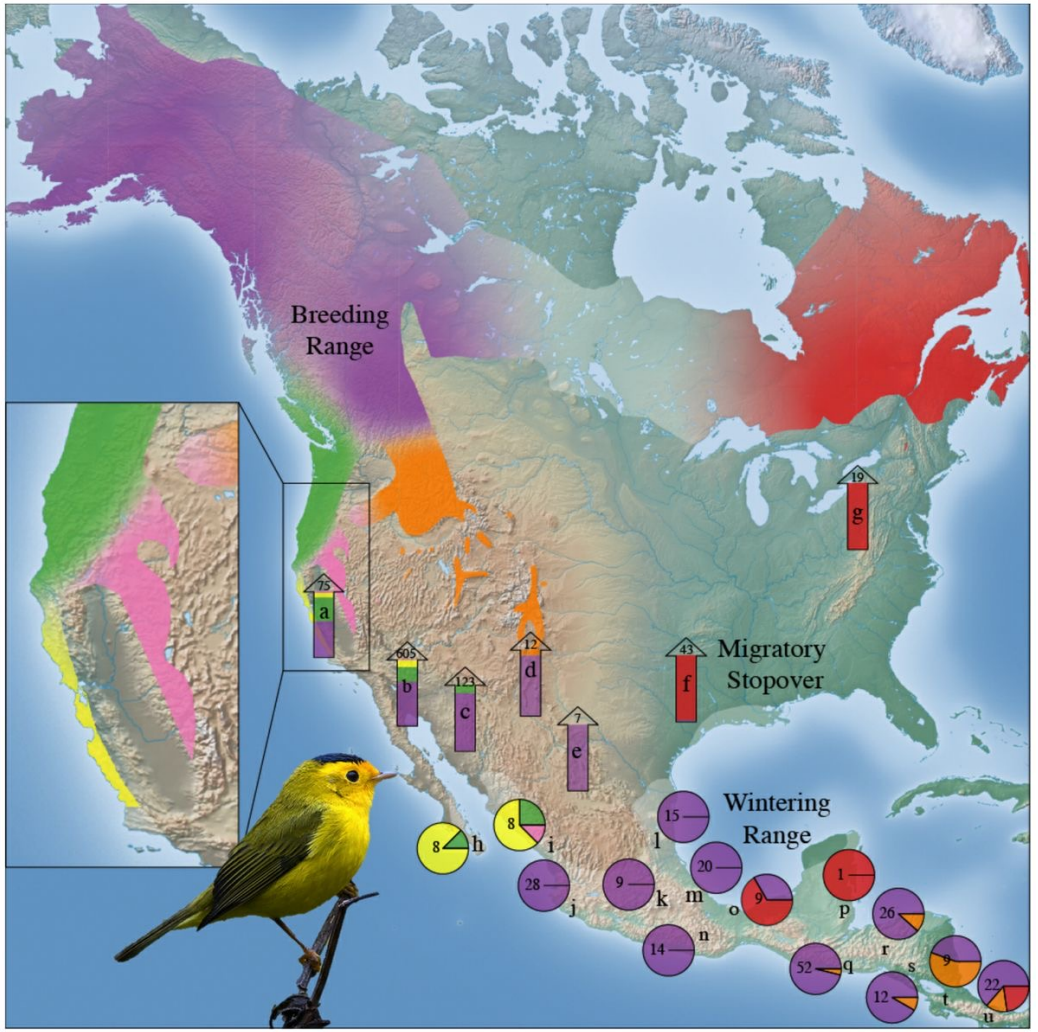
Monitoring/management

bird genoscape project

Mapping the Flyways of the Americas

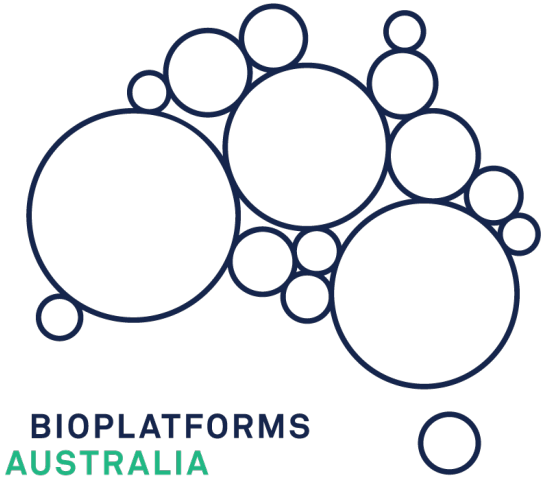
https://www.birdgenoscape.org/media/

https://www.theatlantic.com/science/archive/2019/09/america-has-lost-quarter-its-birds-fifty-years/598318/



Kristen Ruegg

Population profiling & threat assessment



- Assemble reference genomes for threatened species
- Collect/analyze population-level data
  - Genetic diversity
  - Functional gene repertoire

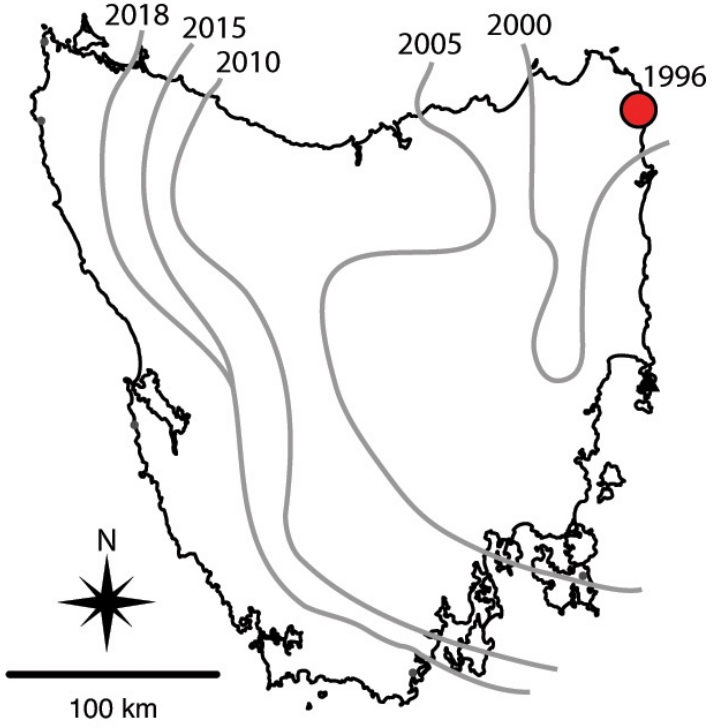
https://threatenedspeciesinitiative.com/

Characterizing disease transmission dynamics

A

B

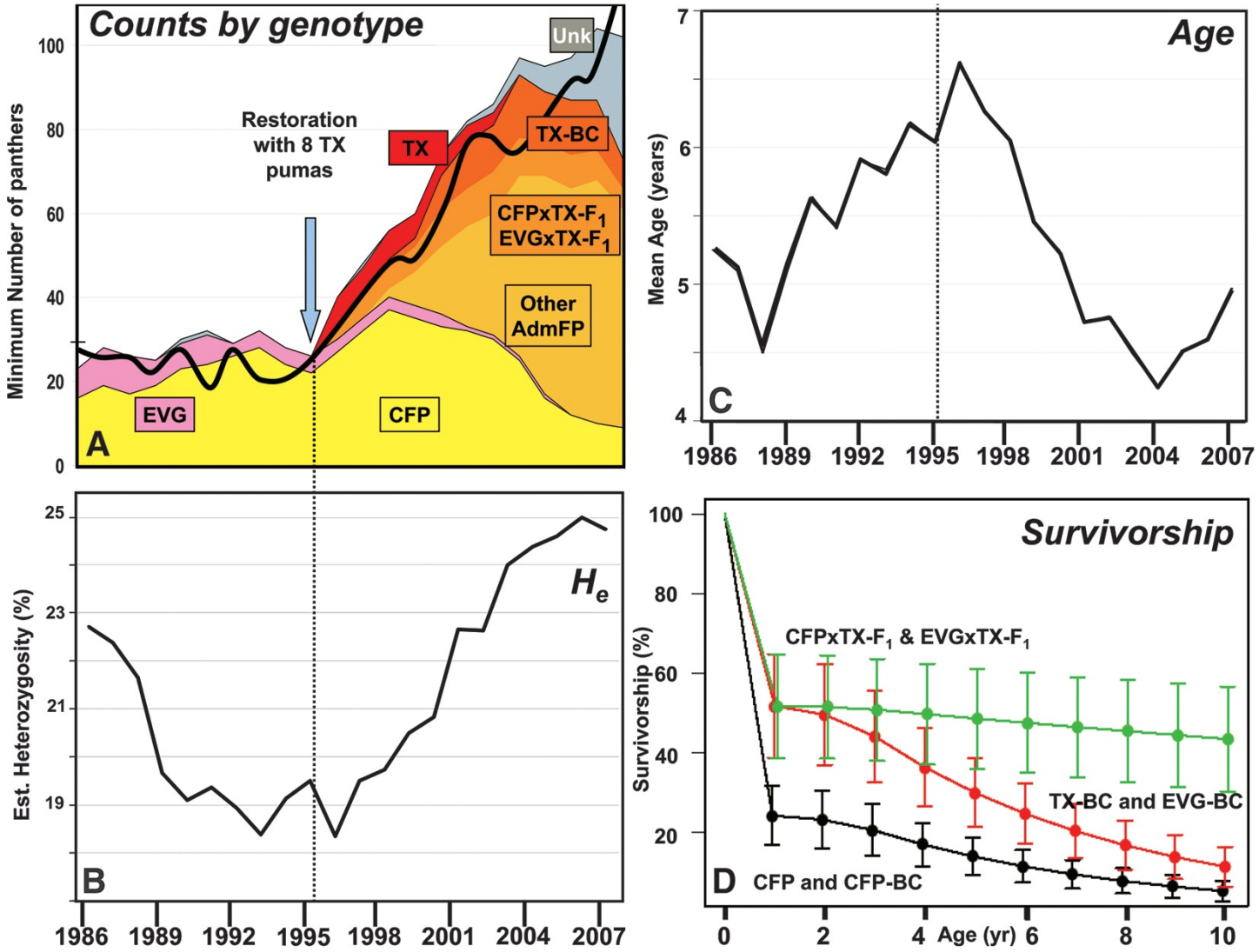
C



Storfer et al 2018. *PLoS Pathogens.*

Genetic rescue as an intervention

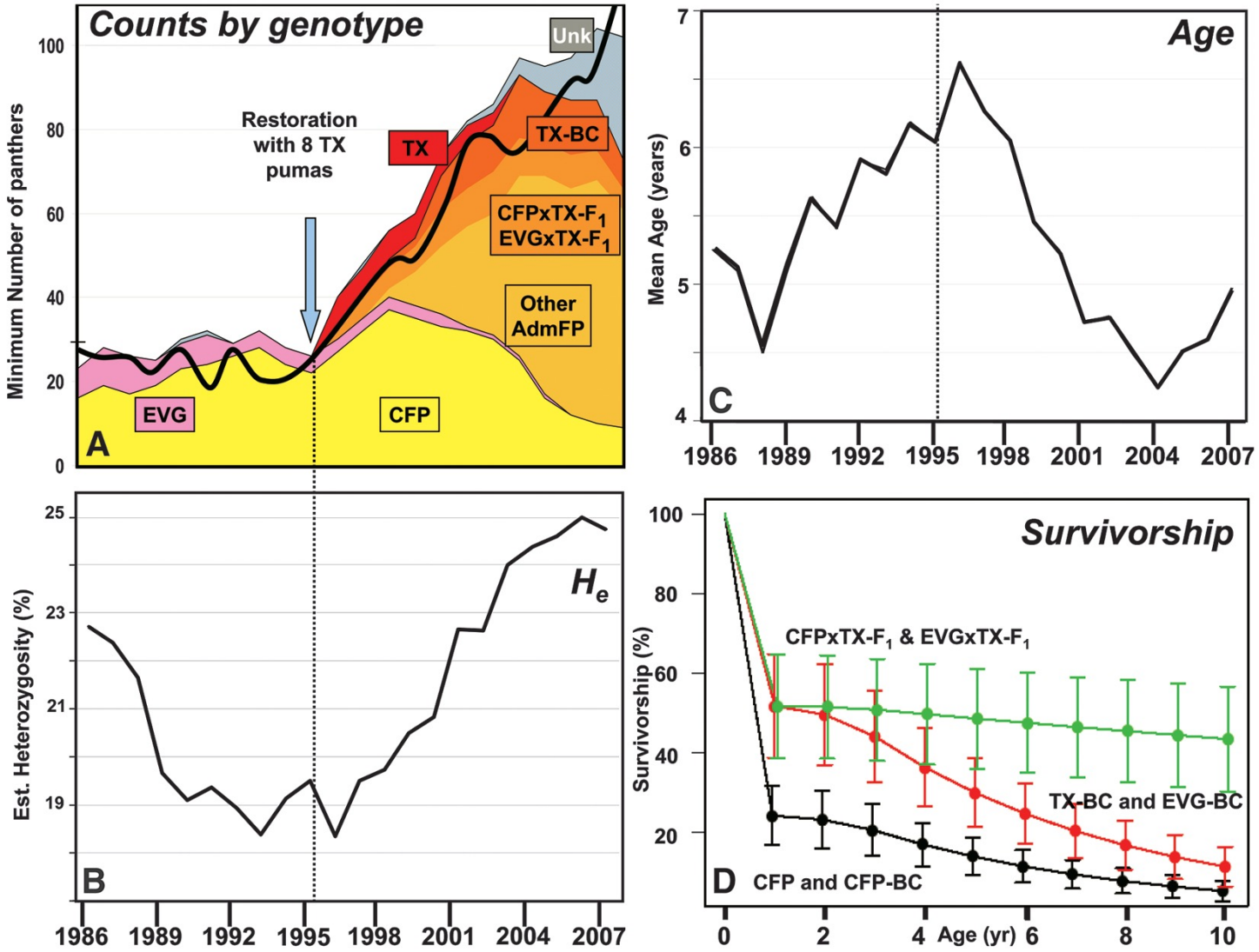
Florida Panther



Johnson et al., 2010. *Science*.

Genetic rescue as an intervention

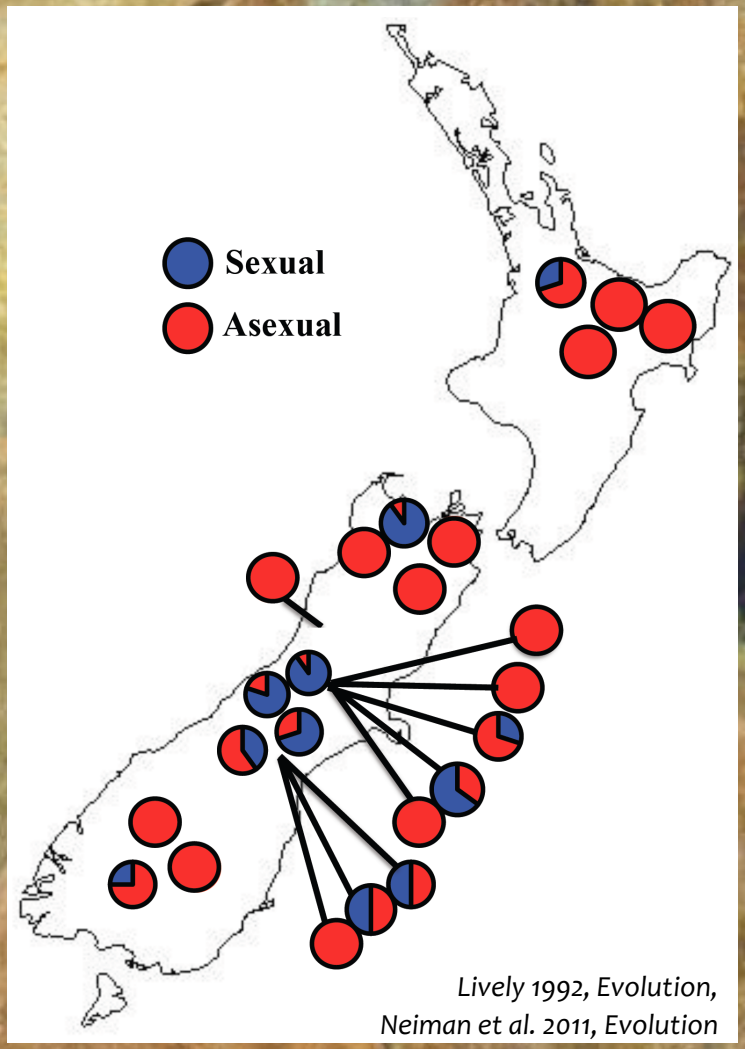
Florida Panther



Johnson et al., 2010. *Science*.

Naturalism as an entry point into science/genomics

New Zealand *Potamopyrgus antipodarum* ("Potamos")
A model system for sex & ploidy variation



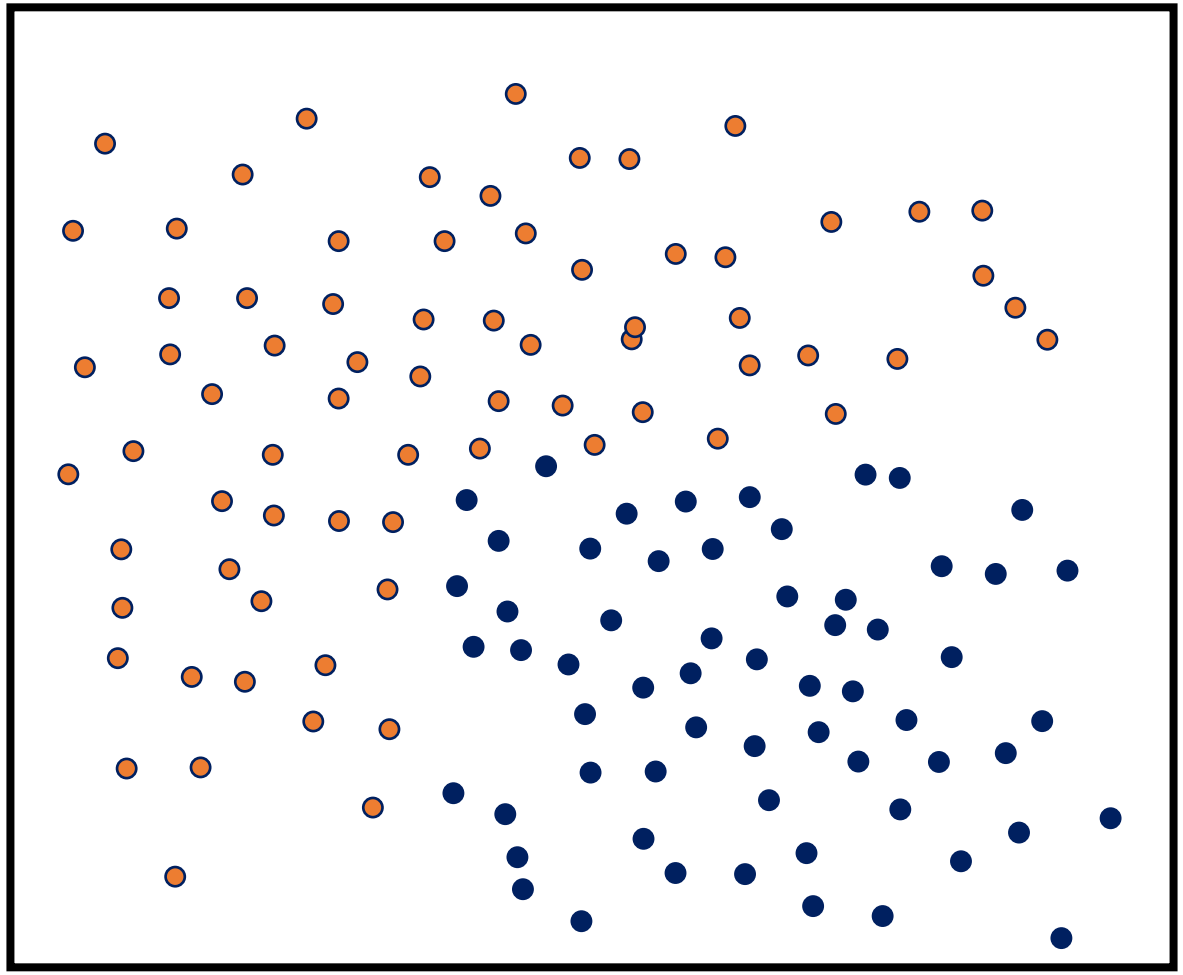
Lively 1992, Evolution,
Neiman et al. 2011, Evolution

Cotton (*Gossypium*)

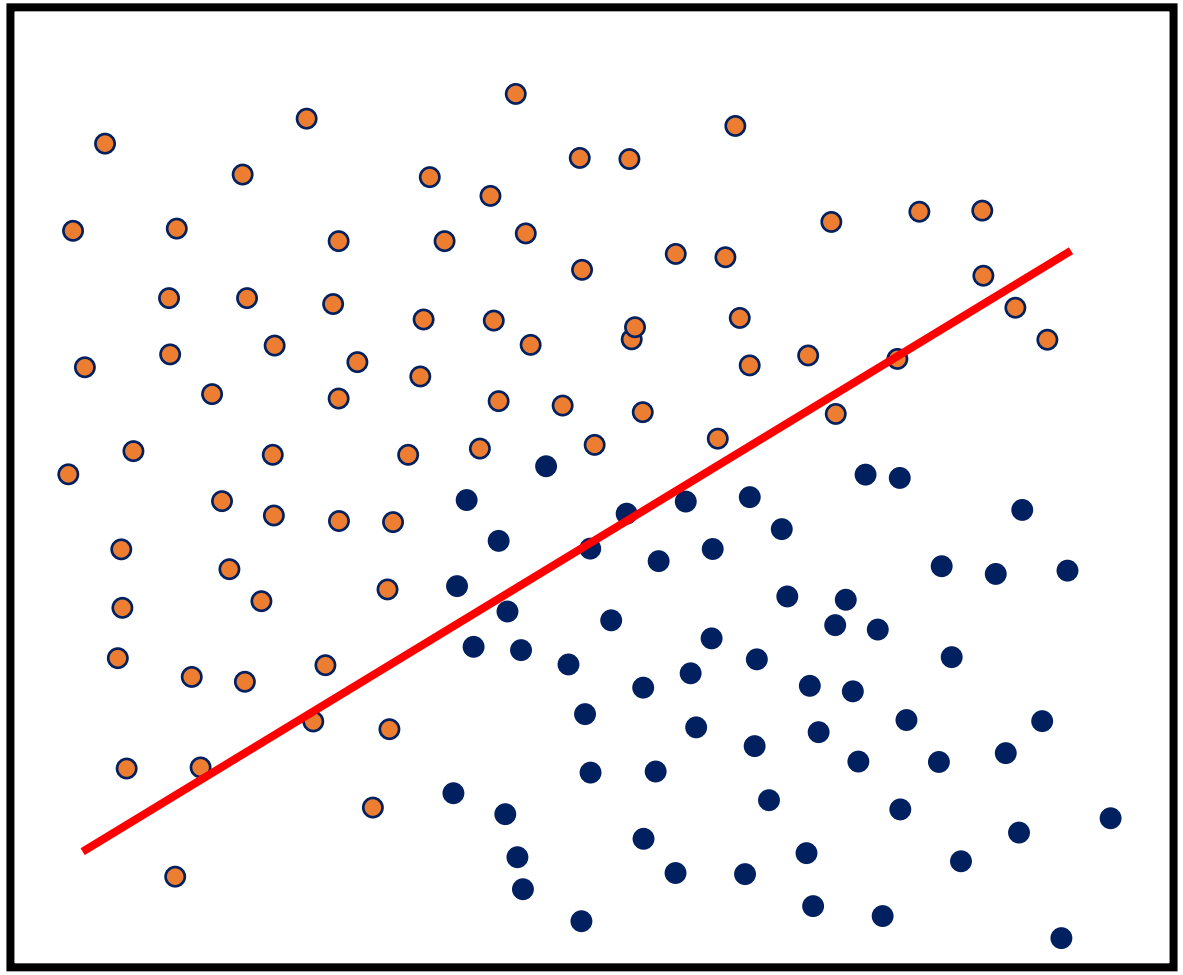
Using computational biology to promote naturalism


Always has been
Wait...it's just statistics?
Machine
Learning

Machine learning – classifying data



Machine learning – classifying data



Machine learning

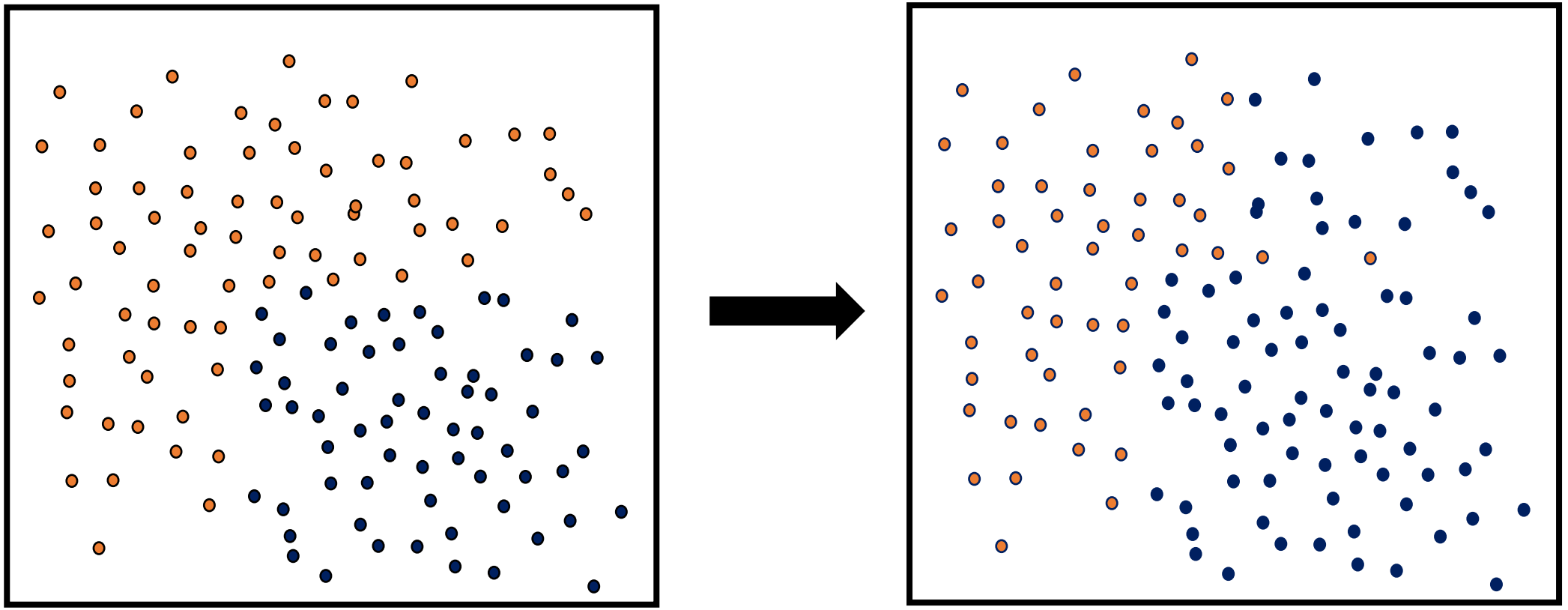
Unsupervised

- Trains on an unlabeled set of data
- Attempts to mimic the data, then compares to original for errors
- Clustering, principal component analysis, neural networks, etc.

Supervised

- Trains on a labeled set of data
- Attempts to classify unlabeled data based on what it learned
- Regression, Support Vector Machines, Decision Trees, Naïve Bayes, etc.

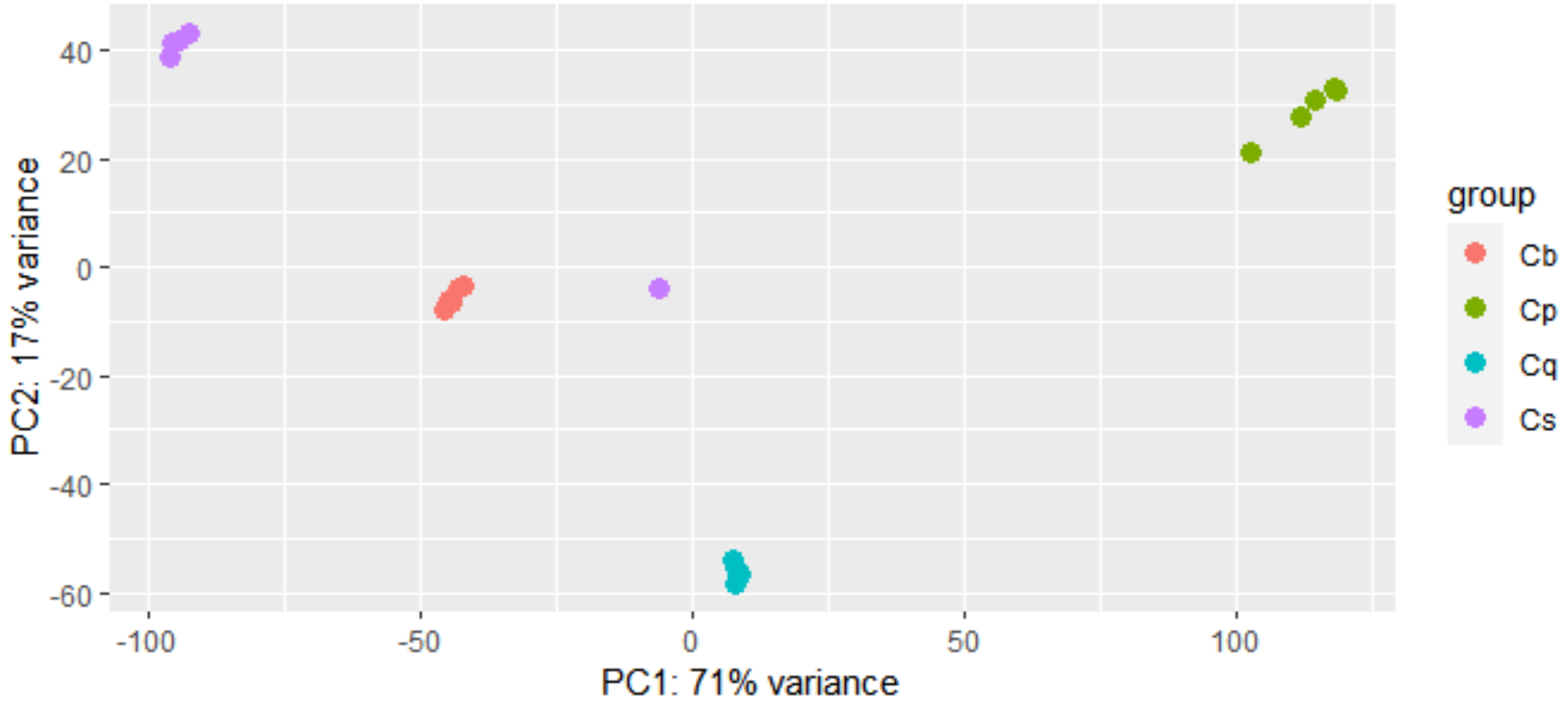
Unsupervised Machine Learning



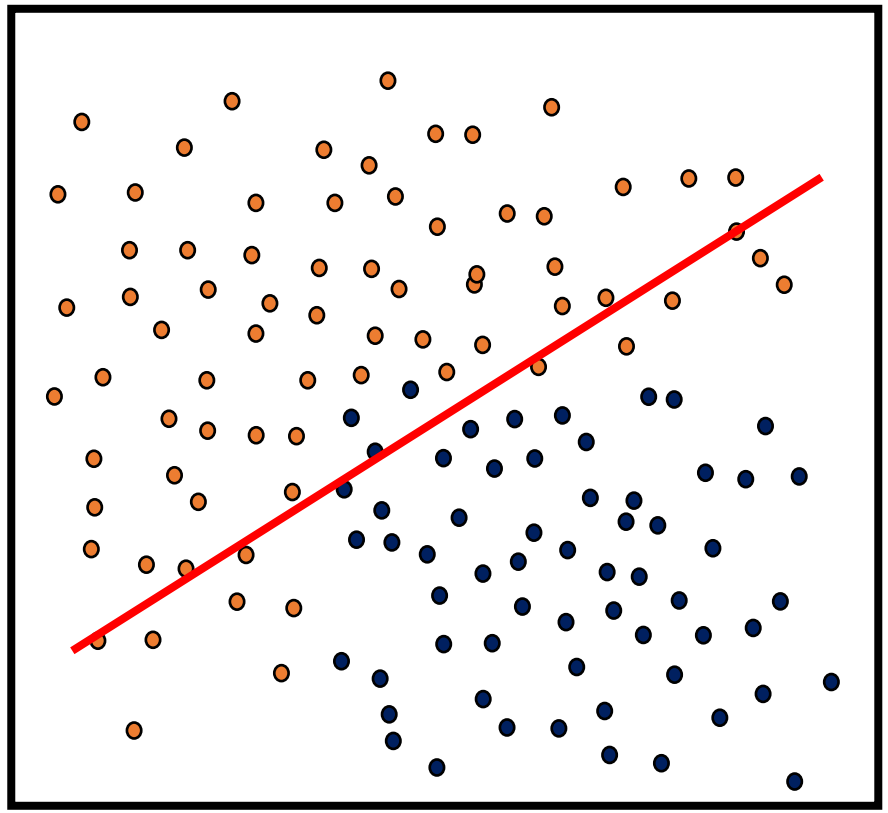
Training dataset

Mimicked dataset

Principle Components Analysis

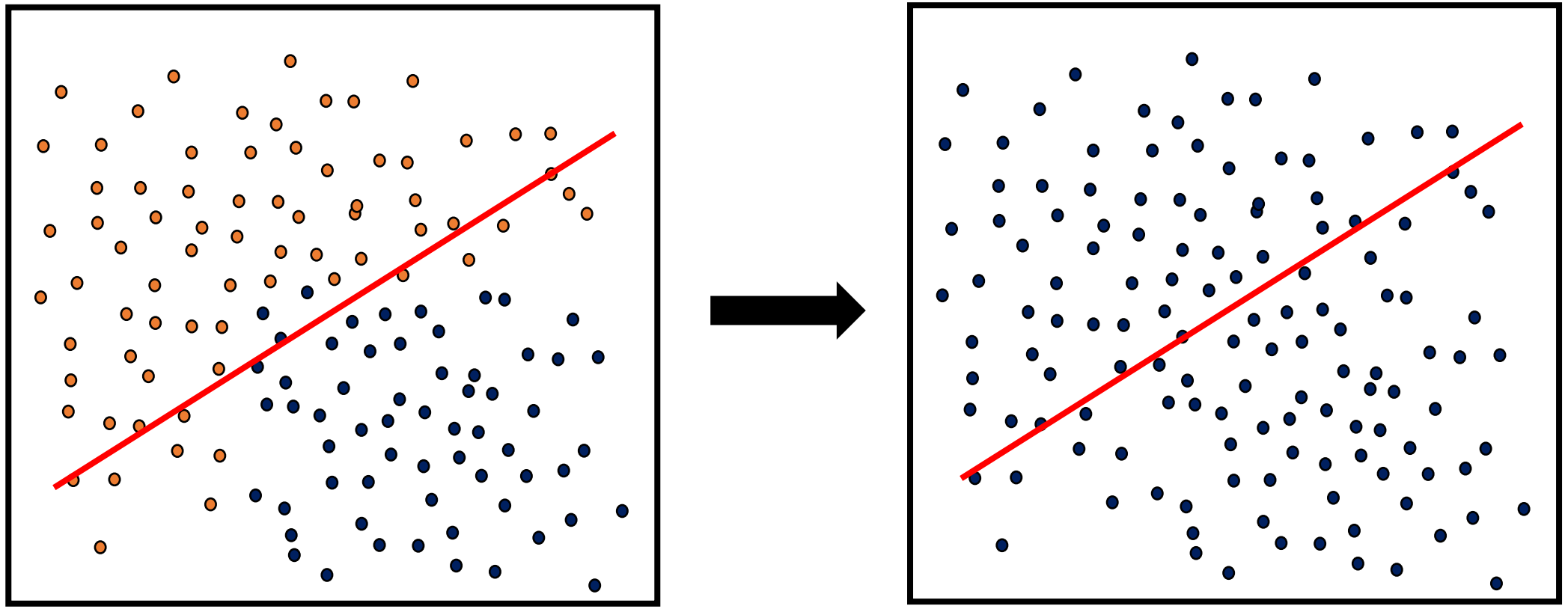


Supervised Machine Learning



Training dataset

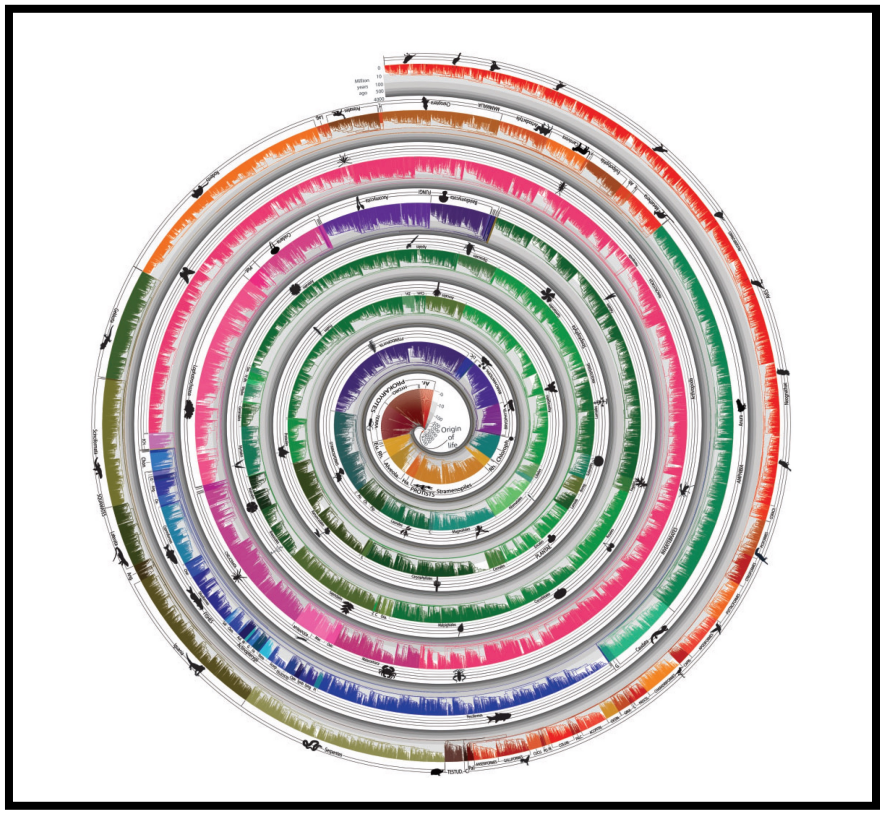
Supervised Machine Learning



Training dataset

Test dataset

Supervised Machine Learning – Seek app



Training dataset (Tree of Life)

On-the-fly
species identification!
Using Decision Tree

Sequoia sempervirens