

Introduction to phylogenetics

Emily Jane McTavish

January 22, 2018

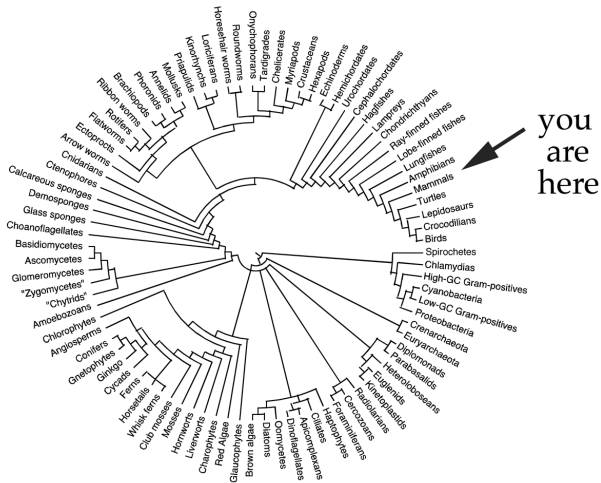


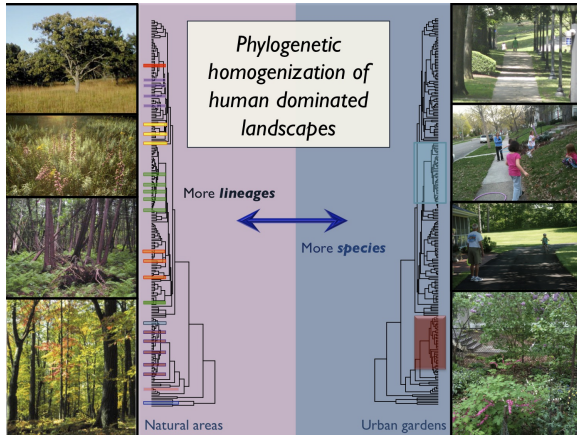
Image Ethan Hein

“Nothing in Biology Makes Sense Except in the Light of Evolution”
Theodosius Dobzhansky

“Nothing in Biology Makes Sense Except in the ~~Light of Evolution~~ Context
of Phylogenetics!”

Why use phylogenies?

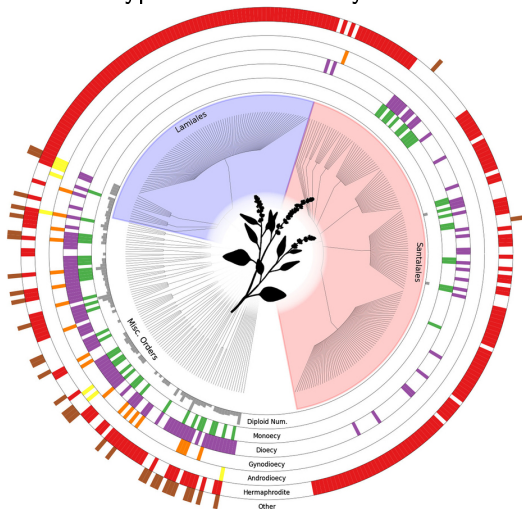
For an ecologist wants to know the relationships of the taxa in their study site



(example from the Cavender-Bares lab webpage)

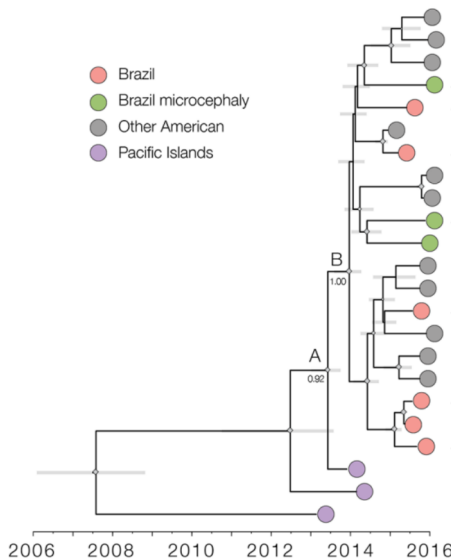
Why use phylogenies?

To understand rates and types of evolutionary transitions



The Tree of Sex Consortium.(2014) Scientific Data

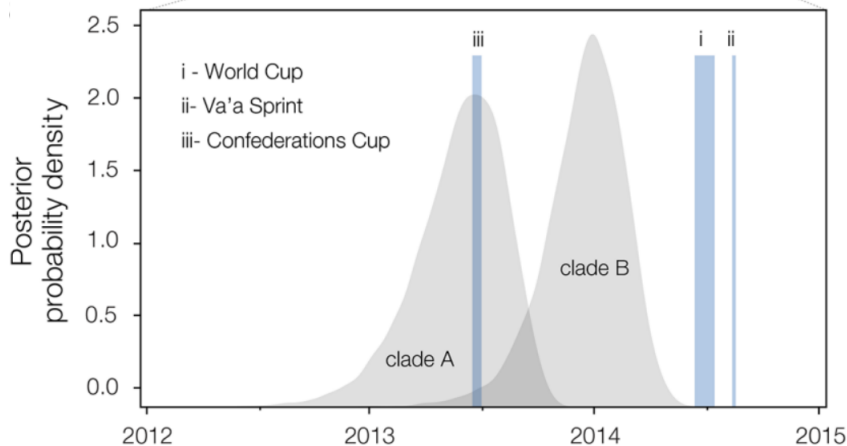
Tracing epidemics



Faria et al., Science. 2016

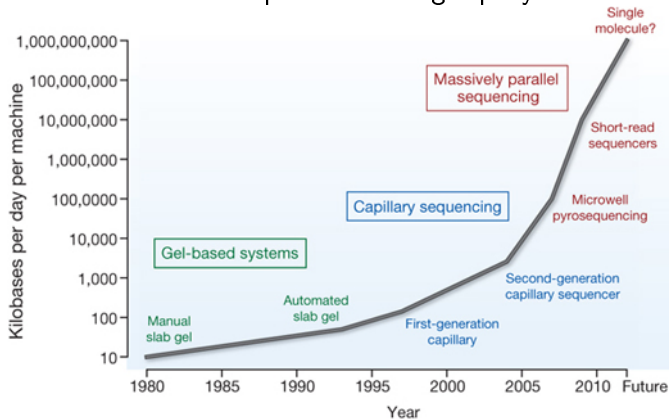
Introduction

Tracing epidemics, placing outbreaks in context



Faria et al., Science. 2016

The quantity of available sequence data for inferring evolutionary relationships is increasing rapidly



<http://genome.wellcome.ac.uk/>

Analysis of these data is computationally challenging

“With the advent of modern molecular biology, the ability to collect biological sequence data has out-paced the ability to adequately analyze these data”

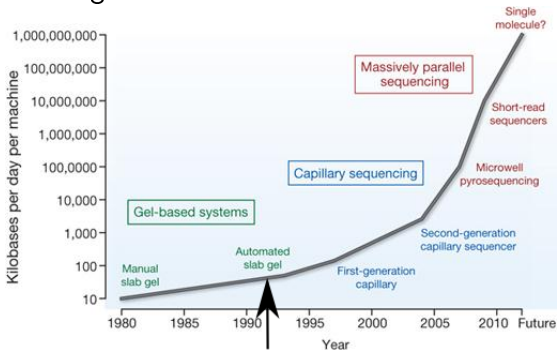
– Jeff Thorne (Evolutionary biologist)

Analysis of these data is computationally challenging

“With the advent of modern molecular biology, the ability to collect biological sequence data has out-paced the ability to adequately analyze these data”
– Jeff Thorne (Evolutionary biologist)

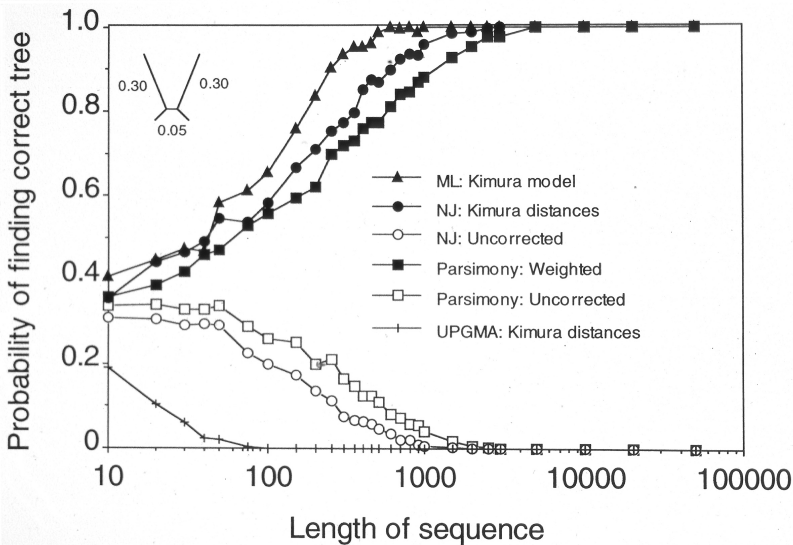
Thorne et al., Journal of Molecular Evolution. **1991**

Dramatic increase DNA sequence data generation has increased computational challenges



<http://genome.wellcome.ac.uk/>

Inappropriate statistical models lead to incorrect conclusions



Hillis, D. M., J. P. Huelsenbeck, and D. L. Swofford. 1994. Hobgoblin of Phylogenetics? Nature 369:363-364.

How can we take advantage of evolutionary information in large data sets,

in a way that is:

- biologically informed
- statistically rigorous
- computationally efficient?

Goals for this course:

Understand the theory and statistics underpinning phylogenetic analysis.

Apply this theory to computational phylogenetic analyses and learn to use remote computing resources (MERCED cluster)

Assess and discuss primary research papers with a phylogenetic focus