Phylogenetics

Adrian Castellanos
Oona Takano
Michelle Chrpa
Pedro Afonso Leite

Image from: http://tree.bio.ed.ac.uk/ Andrew Rambaut 2007
What is phylogenetics?

• The study of the evolutionary relationships of organisms
  – Uses various characters to find these relationships
  – “Family tree” for organisms across time and space
  – Often associated with taxonomy
Tree from Gündüz et al. 2007
Tree Anatomy

- Taxa
- Nodes
- Branches
- Clades
- Root
- Outgroup
What is phylogenetics used for?

- Evolutionary biology
- Community ecology
- Epidemiology
- Conservation
- Evolution of oral history
Sequence of the 1918 nonstructural gene (NS1) characterization of recent viruses bearing the 1918 pandemic H1N1 virus.

Christopher F. Basler*†, Ann H. Reid*†, Jody K. DeCesar†, Hongyong Zheng†, Mirella Salvatore†, Michael L. Follett-Pease†, Peter Palese†, and Jeffery K. Taubenberger†

Genetic variability of the human filarial parasite, Wuchereria bancrofti in South India

N. Pradeep Kumar *, K.P. Patra, S.L. Hoti, P.K. Das
Vector Control Research Centre (ICMR), Indira Nagar, Pondicherry 605 006, India

OPEN ACCESS Freely available online

The Phylogeny of Little Red Riding Hood

Jamshid J. Tehrani*

Department of Anthropology and Centre for the Coevolution of Biology and Culture, Durham University, Science Site, South Road, Durham, United Kingdom

Ecography 37: 614–625, 2014
doi: 10.1111/j.1600-0587.2013.00459.x
© 2013 The Authors. Ecography © 2013 Nordic Society Oikos
Subject Editor: Nate Swenson. Accepted 1 November 2013

Integrating phylogenetic community structure with species distribution models: an example with plants of rock barrens

Oluwatobi A. Oke, Stephen B. Heard and Jeremy T. Lundholm
But what really is phylogenetics?

- Organizing and grouping using various methods and computer programs
  - Many times, statistical methods are used
    - Maximum likelihood estimation
    - Bayesian inference
  - Supercomputer usage is going up
    - More characters, more taxa, more CPU needed
Phylogenetics in R

- R acts as grouping software for phylogenetics programs (similar to Mesquite)

- Capabilities:
  - Loading trees
  - Building phylogenies
  - Tree manipulation
  - Many more…
Parsimony

- Tree with lowest number of changes is most likely evolutionary hypothesis
  - Low computing power needed
  - Not frequently used anymore
- Common programs: PAUP*
- R packages: ape, phangorn
Maximum likelihood

• Uses a model of evolution to determine the most likely tree(s)
  – Models range from simple to complex
    • Can test using ModelTest, PartitionFinder, or R

• Common programs: RAXML, GARLI, PhyML

• R packages: ape, phangorn
Bayesian Methods

• Uses prior probabilities, a model of evolution, and statistical methods to find the most probable tree(s)
  – Complex sampling methods need high computing power (supercomputer recommended)

• Common programs: MrBayes, BEAST

• R packages: phyloch
  – Just a shell to run MrBayes
Other associated analyses

- Divergence estimation
  \textit{ape}
- Ancestral state reconstruction
  \textit{ape, phangorn}
- Trait evolution
  \textit{geiger}
- Phylogeography
  \textit{BioGeoBears}
Resources

• List of R phylogenetics packages
  http://cran.r-project.org/web/views/Phylogenetics.html
• Analysis of Phylogenetics and Evolution in R (book)
  – Free on the Springer website!
• Bodega Bay Phylogenetics Workshop Tutorials
  http://treethinkers.org/tutorials/
• EEB OSOS 2014 Workshop
  http://eeb.tamu.edu/open-source-open-science-workshop-fall-2014/
• Phylogenetic trees in R (demo)
• Phytools demo (tutorial)
  http://www.phytools.org/mpma/Exercise_2.1/
  http://www.phytools.org/eqg/Exercise_3.2/