History of Lineages Chapter 11

Jamie Oaks¹

¹Kincaid Hall 524 joaks1@gmail.com

April 11, 2014



© 2007 Boris Kulikov boris-kulikov.blogspot.com

Some of the slides and images were stolen (with permission) from Dr. Mark Holder.

Outline

What is phylogenetics?

Why is phylogenetics important?

Tree and character terminology and basics

History of phylogenetics

Phylogenetic Methods Molecular Evolution

Phylogeography

- Systematics The science devoted to the study of the diversity of organisms, and the relationships among them.
- Classification The ordering of organisms into named groups on the basis of their relationships.
- Phylogenetics The science of inferring the genealogical relationships between species.

Seen in the light of evolution, biology is, perhaps, intellectually the most satisfying and inspiring science. Without that light it becomes a pile of sundry facts some of them interesting or curious but making no meaningful picture as a whole.

 Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution. The American Biology Teacher 35:125–129.

... nothing in evolution makes sense except in the light of phylogeny ...

- Society of Systematic Biologists

We cannot understand biodiversity without its blueprint: the tree of life.



© 2007 Tree of Life Web Project tolweb.org

Bergmann's Rule



Bergmann's Rule



Bergmann's Rule



- Every field of biology studies organisms.
- These organisms are **not** independent.
- To analyze biological data correctly we need to account for shared history among organisms.

What is phylogenetics?

Why is phylogenetics important?

Tree and character terminology and basics

History of phylogenetics

Phylogenetic Methods Molecular Evolution

Phylogeography

Tree terminology



Terminal nodes Also called "leaves", "tips," or "taxa." These represent our observations (data). Depending on the study, this could be a species, population, individual organism, or a gene.

Internal nodes These represent ancestors of leaves. These are typically not observed. Again, these could be an ancestral species, population, organism, gene, etc.

- Root node The most recent common ancestor (MRCA) of all the tips. Sometimes the root node is not known or estimated, and so you will often see trees "unrooted."
 - Branches These represent topology, or the relationships among the nodes. Sometimes the length of the branches represent the amount of evolutionary change or duration of time.

Interpreting trees



Interpreting trees

These trees are the same! The proximity of the tips does not matter, you have to follow the branches to interpret the relationships.



Interpreting rooted vs unrooted trees



Interpreting rooted vs unrooted trees



Interpreting rooted vs unrooted trees



Tree branch lengths

Cladogram A phylogenetic tree where branches only depict relationships; branch **lengths** have no meaning. Methods that produce cladograms usually estimate **unrooted** trees; the root is assumed or implied via an outgroup.



Tree branch lengths

Phylogram A phylogenetic tree with branch lengths that are proportional to the amount of evolutionary change. Methods that produce phylograms usually estimate **unrooted** trees; the root is assumed or implied via an outgroup.



Tree branch lengths

Chronogram A phylogenetic tree with branch lengths that are proportional to time duration. Methods that produce chronograms estimate **rooted** trees.











Classification: Grouping leaves-the good



Classification: Grouping leaves-the good



Classification: Grouping leaves-the good



Classification: Grouping leaves-the bad

Paraphyletic group A group that consists of an ancestor and some, but not all, of its descendants. Need to add one clade or tip to get monophyly. An "unnatural" group. Bad!



Classification: Grouping leaves-the bad

Paraphyletic group A group that consists of an ancestor and some, but not all, of its descendants. Need to add one clade or tip to get monophyly. An "unnatural" group. Bad!



Classification: Grouping leaves-the ugly

Polyphyletic group A group that consists of unrelated tips. Need to add more than one clade or tip to get monophyly. An "unnatural" group. Ugly!



Other terms



Other terms

Sister group The next most closely related tip or clade; always reciprocal. E.g., in the tree below, B is sister to C (and vice versa), the clade of B and C is sister to A (and vice versa), D is sister to the clade comprised of A, B, and C (and vice versa).



		Characters							
		1	2	3	4	5	6		
Taxa	Homo sapiens	0.13	Α	Α	rounded	1	1610 - 1755		
	Pan paniscus	0.34	A	G	flat	0	621 - 843		
	Gorilla gorilla	0.46	C	G	pointed	0	795 - 1362		

Characters (aka "transformation series") are the columns. The values in the cells are character states (aka "characters").

Charater terminology

Homology A character state that is shared among taxa due to inheritance from a common ancestor (identical by descent).



- Blue character state is homologous.
- Red character state is homologous.

Charater terminology

Homology A character state that is shared among taxa due to inheritance from a common ancestor (identical by descent).



- Blue character state is homologous.
- Red character state is not!

Charater terminology

Homoplasy A character state that is shared because of multiple (convergent) changes. Homo = "same" plasy = "change." Diagnose polyphyletic groups.



- Red character state is homoplasious.
- Blue character state is homologous.

Prefixes:

apo Refers to the new or derived state plesio Refers to the old or primitive state syn or sym Used to indicate shared between taxa aut Used to indicate a state being unique to one taxon Terms: synapomorphy Shared, derived states (homologous). Used to diagnose monophyletic groups.

symplesiomorphy Shared, primitive states (homologous). Diagnose icky, unwanted paraphyletic groups.

autapomorphy Unique derived state.



Blue character state is synapomorphic.



Blue character state is symplesiomorphic.



red character state is autapomorphic.

Outline

What is phylogenetics?

Why is phylogenetics important?

Tree and character terminology and basics

History of phylogenetics

Phylogenetic Methods Molecular Evolution

Phylogeography

A (very) brief history of Systematics

Evolutionary phylogenetics Very subjective and recognized those icky paraphyletic groups. Still remnants of its long legacy in the form of "traditional" paraphyletic groups (e.g., fish, reptiles, lizards).

Numerical phenetics Much more objective and quantitative. Use algorithms to cluster taxa based on similarity of phenotypic characters. Inappropriate use of homoplasious/symplesiomorphic characters often caused paraphyletic/polyphyletic groups.

Phylogenetic systematics (cladistics) Used logical inference to reconstruct relationships based on synapomorphic characters. Won the "war" against phenetics in the 70–80's. Left many cladists suffering from PTSD that are vehemently opposed to statistical inference. Does not deal with homoplastic characters well (they violate the logic).

	Character #											
Taxon	1	2	3	4	5	6	7	8	9	10	11	12
A	0	0	0	0	0	0	0	0	0	0	0	0
В	1	0	0	0	0	1	1	1	1	1	1	1
С	0	1	1	1	0	1	1	1	1	1	1	0
D	0	0	0	0	1	1	1	1	1	0	0	1











Outline

What is phylogenetics?

Why is phylogenetics important?

Tree and character terminology and basics

History of phylogenetics

Phylogenetic Methods Molecular Evolution

Phylogeography

Distance-based methods Group tips based on some measure of evolutionary distance. Very fast. Have to "distill" discrete characters to distances (throwing away information).

Parsimony Search for the tree(s) that require the smallest number of character-state changes (Occam's razor). Not explicitly statistical. Only uses some of the characters. We know it can be positively misleading.

Statistical Inference Use probabilistic models of character evolution. Can use all of the data. More robust and can quantify uncertainty.

> Maximum likelihood Find the tree that maximizes the probability of the observed data. Bayesian inference Find the distribution of trees with the highest probability given the data.

- Earliest approaches attempted to measure changes in DNA indirectly via immunological assays, protein electrophoresis, DNA-DNA hybridization, and restriction enzymes.
- Polymerase chain reaction (PCR) and Sanger sequencing made it possible to observe the character states of the DNA nucleotides directly.

- DNA is a "simple" character; only 4 discrete states (A, C, G &, T) across the entire tree of life!
- For conserved (slowly evolving) regions, you get homologous characters across bacteria to mammals.
- Yet, rapidly evolving regions are variable among closely related organisms.
- Very conducive for modeling due to the small number of states and similar replication machinery (DNA polymerases) across all life.
- We can now sequence entire genomes (billions of characters!) relatively quickly and cheaply—phylogenomics.

Outline

What is phylogenetics?

Why is phylogenetics important?

Tree and character terminology and basics

History of phylogenetics

Phylogenetic Methods Molecular Evolution

Phylogeography

- DNA data made it possible to collect characters and estimate gene trees: genealogical relationships among gene copies carried by individual organisms within and among closely related species.
- Bridges gap between phylogenetics and population genetics.

- We can now model hierarchical processes of evolution.
 - Model evolution of nucleotides along gene trees.
 - Model genealogical processes within populations.
 - Model diversification processes among species.
 - Model spatial dynamics of diversification.

Present



Present



Present



Genealogies within a population



Genealogies within a population





Gene trees within a species tree

Phylogeography—West Nile Virus



Pybus et al. 2012.

Phylogeography—West Nile Virus

video link