April 2, 2014**. Molecular Evolution**

A. Aims: (1) phylogeny reconstruction (2) study of evolution at the molecular level per se.

We will focus on the latter, but there is of course feedback. If we want to <u>study</u> processes of evolution at the molecular level (or any level for that matter), how much do we want to <u>assume</u> about processes of evolution at the molecular level before we get our phylogeny?

Models of phylogeny reconstruction, from simple to complex.

Once we have a model, where do we get the values for the parameters in the model?

1. From the data at hand -- model test
2. From *a priori* knowledge

Should we try to do everything at once (build tree, infer best model of molecular evolution, infer biogeography, etc.)? The Bayesian trend; is it a good idea?

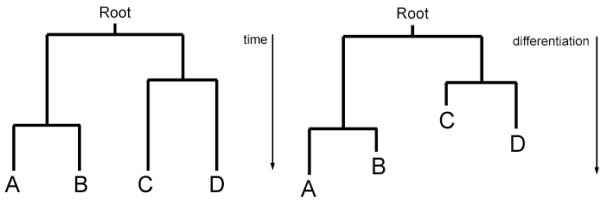
B. Topics for discussion:

Mutation, recombination, and gene conversion

Transposable elements

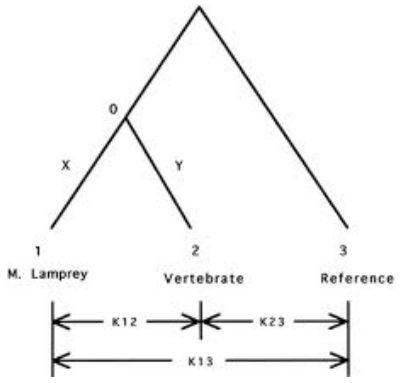
Repetitive elements (microsatellites)

SNPs



Evolutionary rates - can vary at different sites in gene and change between branches on tree.

ultrametric vs. non-ultrametric trees



How to test?

relative rate tests
ML rate tests

Codon usage bias (fig. next page)

G+C content

Natural selection vs neutrality

Detecting selection: D_N/D_S

A ratio greater than one implies positive selection; less than one implies purifying (stabilizing) selection; and a ratio of one indicates neutral (i.e. no) selection.

Transition-transversion bias
purines A G
pyrimidines C T U

RNA, secondary and tertiary structure
structure

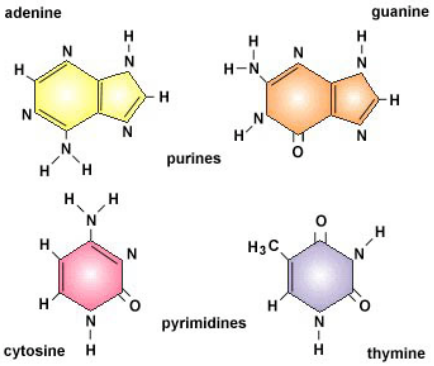
Protein evolution ("proteomics")
secondary and tertiary structure

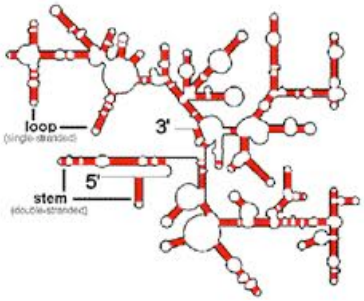
Polyploidy - genome size

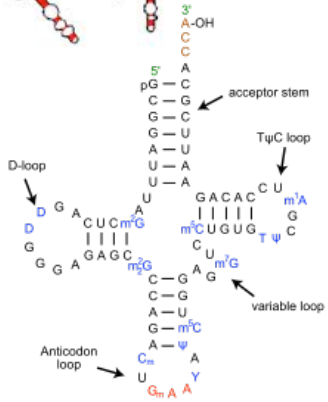
Different genomes
organellar vs. nuclear
gene movement

Comparing genomes (next time)
synteny, rearrangements, indels
Exon shuffling
Introns (different types)

Multigene families
paralogy vs orthology
the fate of duplicated genes: ghost genes, subfunctionalization

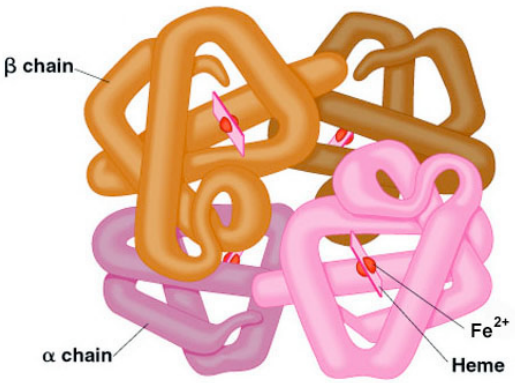






1st	2nd T		2nd C		2nd A		2nd G	
T	TTT 0.43	Phe	TCT 0.18	Ser	TAT 0.42	Tyr	TGT 0.42	Cys
	TTC 0.57		TCC 0.23		TAC 0.58		TGC 0.58	
	TTA 0.06	Leu	TCA 0.15		TAA 0.22	TERM	TGA 0.61	TERM
	TTG 0.12		TCG 0.06		TAG 0.17		TGG 1.00	Trp
C	CTT 0.12		CCT 0.29		CAT 0.41	His	CGT 0.09	
	CTC 0.20	Leu	CCC 0.33	Pro	CAC 0.59		CGC 0.19	Arg
	CTA 0.07		CCA 0.27		CAA 0.27	Gln	CGA 0.10	
	CTG 0.43		CCG 0.11		CAG 0.73		CGG 0.19	
A	ATT 0.35		ACT 0.23		AAT 0.44	Asn	AGT 0.14	Ser
	ATC 0.52	Ile	ACC 0.38	Thr	AAC 0.56		AGC 0.25	
	ATA 0.14		ACA 0.27		AAA 0.40	Lys	AGA 0.21	Arg
	ATG 1.00	Met	ACG 0.12		AAG 0.60		AGG 0.22	
G	GTT 0.17		GCT 0.28		GAT 0.44	Asp	GGT 0.18	
	GTC 0.25	Val	GCC 0.40	Ala	GAC 0.56		GGC 0.33	Gly
	GTA 0.10		GCA 0.22		GAA 0.41	Glu	GGA 0.26	
	GTG 0.48		GCG 0.10		GAG 0.59		GGG 0.23	

Codon Usage in Homo sapiens. The values in red represent the frequency of use for each codon in a group. (From the codon usage database at http://www.kazusa.or.jp/codon/)



(over)