

March 30, 2016. **Molecular Evolution**

A. Distinctly different aims: (1) phylogeny reconstruction; (2) study of evolution at the molecular level per se.

We will focus on the latter today, but there is of course feedback. If we want to study processes of evolution at the molecular level (or any level for that matter), how much do we want to assume 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, e.g., 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?

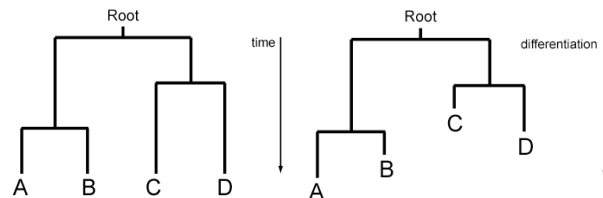
B. Topics for discussion:

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

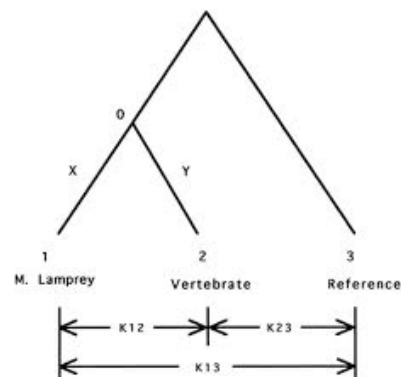
Codon usage bias (fig. next page)

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

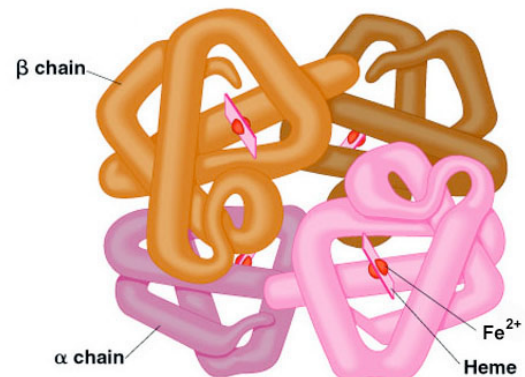
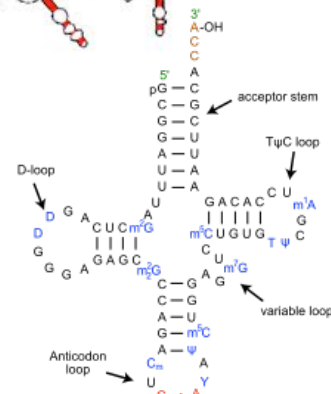
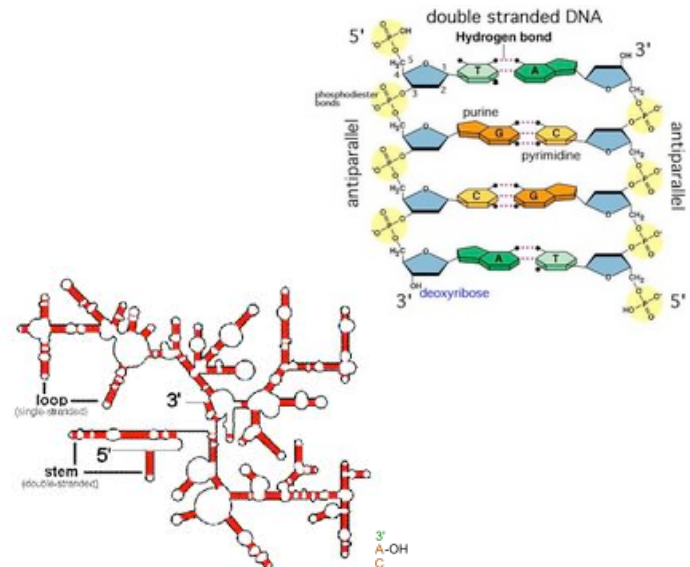
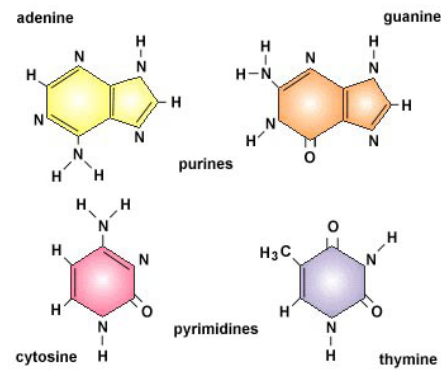
Protein evolution ("proteomics")
 secondary and tertiary structure

Polyploidy - genome size

Different genomes
 organellar vs. nuclear
 gene movement

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

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



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

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/>)