

April 17, 2020. **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?

How do we choose a models for phylogeny reconstruction?

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.), a current Bayesian trend? Is it a good idea or is it better to focus on one problem at a time?

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.

How to test?

- relative rate tests
- ML rate tests

Transition-transversion bias

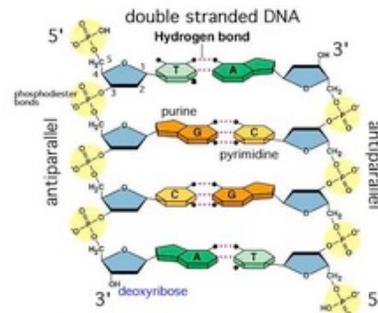
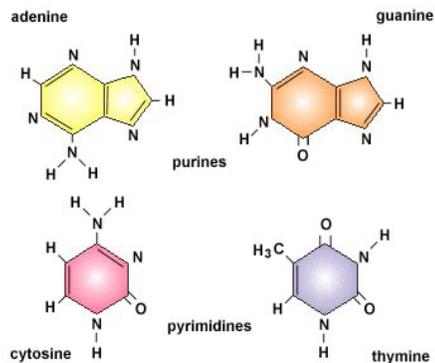
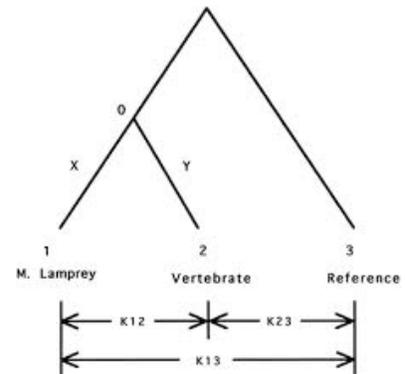
- purines A G
- pyrimidines C T U

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.



RNA, secondary and tertiary structure

Protein evolution ("proteomics")
secondary and tertiary structure

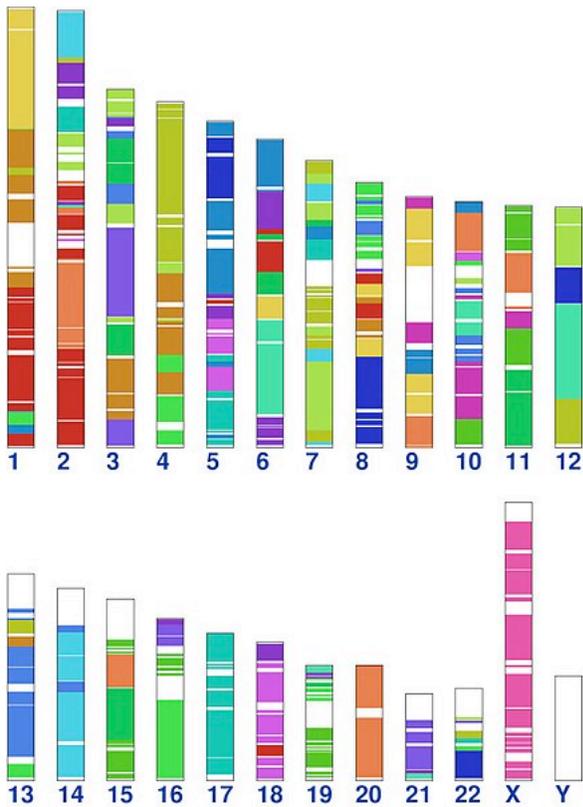
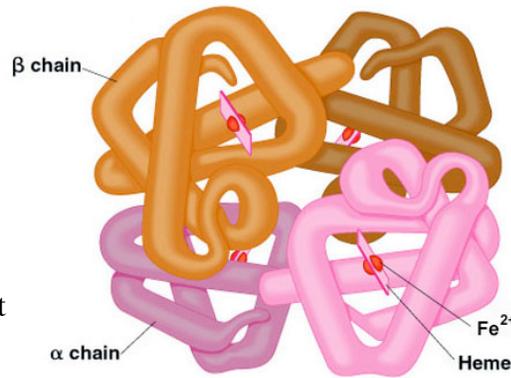
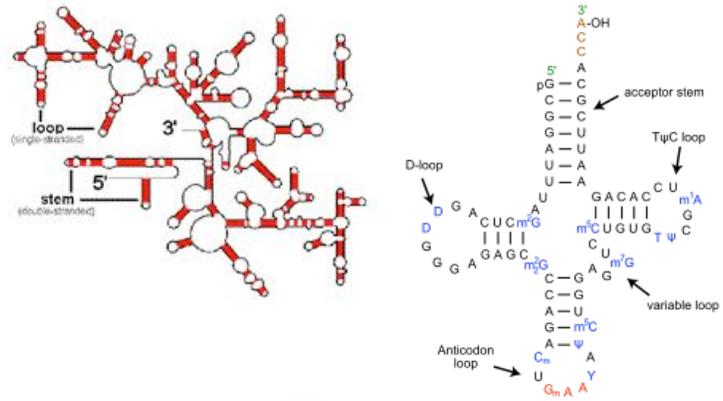
Codon usage bias (see fig. below)

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**



		2nd			
		T	C	A	G
1st					
T	TTT 0.43	TCT 0.18	TAT 0.42	TGT 0.42	
	TTC 0.57 Phe	TCC 0.23	TAC 0.58 Tyr	TGC 0.58 Cys	
	TTA 0.06	TCA 0.15	TAA 0.22	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	CGT 0.09	
	CTC 0.20	CCC 0.33	CAC 0.59 His	CGC 0.19	
	CTA 0.07	CCA 0.27	CAA 0.27	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	AGT 0.14	
	ATC 0.52 Ile	ACC 0.38	AAC 0.56	AGC 0.25	
	ATA 0.14	ACA 0.27	AAA 0.40	AGA 0.21	
	ATG 1.00 Met	ACG 0.12	AAG 0.60	AGG 0.22	
G	GTT 0.17	GCT 0.20	GAT 0.44	GGT 0.18	
	GTC 0.25	GCC 0.40	GAC 0.56	GGC 0.33	
	GTA 0.10	GCA 0.22	GAA 0.41	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/>)

Synteny between human and mouse chromosomes.