

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 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 -- 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.

How to test?

relative rate tests  
ML rate tests

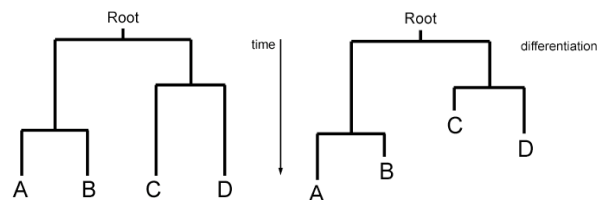
Codon usage bias (fig. next page)

G+C content

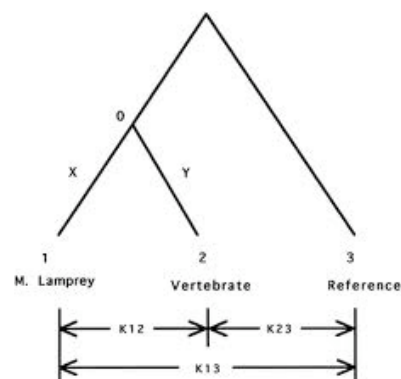
Natural selection vs neutrality

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.



ultrametric vs. non-ultrametric trees



The image displays the chemical structures of four nitrogenous bases, categorized into purines and pyrimidines. Adenine and guanine are purines, while cytosine and thymine are pyrimidines. The structures are color-coded: adenine is yellow, guanine is orange, cytosine is pink, and thymine is purple. Each structure shows the characteristic ring system and the positions of nitrogen atoms and hydrogen atoms.

**adenine**

**guanine**

**purines**

**pyrimidines**

**cytosine**

**thymine**

The diagram illustrates a branched polymer structure. It features a central node from which multiple chains extend. One chain is labeled 'loop (single-stranded)' and another is labeled 'stem (double-stranded)'. The stem is further labeled with '5'' and '3'' ends, indicating its orientation. The structure is composed of red and white spheres connected by lines, representing atoms and bonds respectively.

The diagram illustrates the cloverleaf secondary structure of a tRNA molecule. Key features include:

- 3' end:** The acceptor stem terminates in a 3'-OH group.
- Acceptor stem:** Formed by base pairing between the 5' and 3' ends of the acceptor arm.
- D-loop:** A loop in the D arm containing modified nucleotides like dihydrouridine (D).
- TΨC loop:** A loop in the TΨC arm containing modified nucleotides like pseudouridine (Ψ) and thymine (T).
- Variable loop:** A loop of varying length, often containing modified nucleotides.
- Anticodon loop:** The loop at the 3' end that contains the anticodon, which is complementary to the codon on the mRNA.

A 3D ribbon diagram of a hemoglobin molecule. It consists of two  $\alpha$  chains (colored pink) and two  $\beta$  chains (colored orange). Each chain is associated with a heme group, represented by a red disk-like structure. One heme group is shown in detail, with a label pointing to the central iron atom,  $Fe^{2+}$ , and the heme prosthetic group itself. Labels also identify the  $\alpha$  chain and  $\beta$  chain.

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/>)

2

(over)