

## Assignment 2 / Biometry and bioinformatics II / 2014

---

The goal: To familiarize with genomic differences as regards their phylogenetic signal.

Datafile [data\\_3.txt](#) is a set of complete mitochondrial sequences from a set of mammals.  
Datafile [data\\_4.txt](#) is one gene, cytochrome B, cut from this data.

- Write an interpretation about phylogenetic relationships of the species
  - a) on the basis of the whole mt-genome data
  - b) on the basis of cytochrome B

For a) do only neighbor-joining phylogeny; the data is so big that other analyses will be really slow.

For b) do the same analyses as you did for Assignment 1 (the bacteria data). In addition to writing an interpretation of the results, write about comparison between one-gene-data and whole-mt-genome-data (on the basis of neighbor-joining).

- Inspect visually the whole mt-genome alignment and write about differences you notice between protein coding genes and D-loop: how useful are these two types of sequences in providing information for species comparisons by using phylogenetics (for which sequence alignment, which you are now asked to inspect, is the first step). Next page shows the coordinates for different genes and other pieces in the mt-genome. A verbal story is here enough.
- NOTE: In MEGA, with cytB, you have to define the code: it is NOT "standard", but "vertebrate mitochondrial". When you work with the whole mt-genome data your answer to question "protein coding data" is NO because the data (although it contains protein coding genes, see the table next page) is not a clear 123 123 123.... (=the codons for amino acids). For example the first gene in the data is a transfer-RNA etc.

	Nucleotides in AB499817, the first sequence in datafile	Nucleotides taking into account gaps in aligned file
tRNA-Phe	1-69	1-80
12S ribosomal RNA	70-1023	81-1090
tRNA-Val	1024-1090	1091-1161
16S ribosomal RNA	1091-2670	1162-2840
tRNA-Leu	2671-2745	2841-2917
gene ND1	2748-3704	2919-3882
tRNA-Ile	3704-3722	3882-3901
tRNA-Gln	3769-3843	3948-4025
tRNA-Met	3845-3914	4028-4097
gene ND2	3915-4958	4098-5143
tRNA-Trp	4957-5024	5142-5215
tRNA-Ala	5038-5106	5232-5301
tRNA-Asn	5108-5179	5310-5386
tRNA-Cys	5213-5280	5419-5495
tRNA-Tyr	5281-5348	5496-5572
gene COI	5350-6894	5574-7140
tRNA-Ser	6892-6962	7132-7216
tRNA-Asp	6967-7034	7222-7292
gene COII	7035-7718	7293-7977
tRNA-Lys	7736-7802	7995-8066
gene ATPase subunit 8	7804-8007	8068-8276
gene ATPase subunit 6	7965-8645	8234-8914
gene COIII	8645-9428	8914-9697
tRNA-Gly	9429-9496	9698-9770
gene ND3	9497-9843	9771-10117
tRNA-Arg	9843-9911	10117-10187
gene ND4L	9914-10210	10191-10487
gene ND4	10204-11581	10481-11858
tRNA-His	11580-11650	11857-11930
tRNA-Ser	11651-11710	11931-11995
tRNA-Leu	11711-11780	11996-12067
gene ND5	11781-13601	12068-13895
gene ND6	13585-14112	13879-14406
tRNA-Gln	14111-14181	14405-14476
gene cytB	14186-15325	14482-15625
tRNA-Thr	15326-15395	15626-15703
tRNA-Pro	15395-15460	15703-15772
D-loop	15461-16741	15773-18424