

Bioinformatyka i Biologia Obliczeniowa
Laboratorium 7

Celem ćwiczenia jest zastosowanie algorytmów dopasowywania wielu sekwencji **MSA** (**M**ultiple **S**equences **A**lignment).

Dane: Analizę przeprowadzimy w oparciu o jeden z jego produktów białkowych genomu mitochondrialnego - **cytochrom b** (cytochrome b), którego ewolucyjna szybkość zmienności jest bardzo odpowiednia przy ustalaniu pokrewieństwa pomiędzy gatunkami). Wybierz dane z kilku gatunków, np.: *człowiek (Homo sapiens)*, *neandertalczyk (homo sapiens neanderthalensis)*, *szympan (chimpanzee, pan troglodytes)*, *muszka owocówka (fruit fly, drosophila melanogaster)*, *nosorożec (...)*, *słoń (.....)*, *manut (.....)*

1. Znajdź sekwencje *cytochromu b* do MSA, korzystając ze znanych Ci już serwisów (EBI lub NCBI). Sekwencje zgromadź w jednym pliku tekstowym do późniejszych analiz.
2. Przeprowadź MSA za pomocą:
 - a. narzędzi dostępnych poprzez EBI: <http://www.ebi.ac.uk/Tools/msa/>
 - b. narzędzia COBALT (Constraint-based Multiple Alignment)
<http://www.ncbi.nlm.nih.gov/tools/cobalt/cobalt.cgi>.Wyniki zapisz w pliku o formacie Clustal (Download/Clustal).
3. Narysuj profil swojego zbioru za pomocą **generatora logo WebLogo**:
<http://weblogo.threeplusone.com>
Znajdź, za pomocą logo, obszary najlepiej i najslabiej zakonserwowane.