

# Übungen zum Sequenzanalyse-Praktikum

Universität Bielefeld, SoSe 2015

Dr. Roland Wittler · M.Sc. Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2015summer/SequaPrak>

[praktikum-seqan@CeBiTec.Uni-Bielefeld.DE](mailto:praktikum-seqan@CeBiTec.Uni-Bielefeld.DE)

**Übungsblatt 09 vom 15.06.2015**

**Abgabe bis Donnerstag, 24:00 Uhr.**

## Aufgabe 1 (MSA verwandter Sequenzen mit ClustalW2)

1. Wie lauten die lateinischen Namen des Menschen, der Hausmaus, der Schopffente, der Spitzschwanzente und des Atlantischen Lachses?
2. Suche dir die Proteinsequenzen des Haemoglobin-Alpha-Gens (ggf. Untereinheit A) der Organismen aus Aufgabe 1.1 heraus. Gib im Protokoll nur die *Accession Numbers* der Sequenzen an.
3. Gehe auf <http://www.ebi.ac.uk/Tools/msa/clustalw2/> und füge deine Sequenzen in das Eingabeformular ein. Führe *ClustalW2* mit Standardeinstellungen aus.
4. Als letzte Zeile im multiplen Alignment findest du verschiedene Zeichen. Welche sind es und was bedeuten sie?
5. Lasse dir jetzt die Farben im multiplen Alignment anzeigen. Was zeigen sie an? (Eine kurze Erklärung genügt.)
6. Wie unterscheidet sich ein phylogenetischer Baum von einem Kladogramm?
7. Betrachte nun den phylogenetischen Baum und das Kladogramm. (Bilde diese auch in deinem Protokoll ab.) Zu wem zeigt der Mensch die größte Ähnlichkeit? Ist die Haemoglobin-Sequenz des Menschen ähnlicher der von Vögeln oder der von Fischen?

## Aufgabe 2 (MSA weit verwandter Sequenzen mit ClustalW2)

1. Suche dir nun die Proteinsequenz mit der GI Nummer *313661502* heraus. Um was für eine Sequenz handelt es sich? Füge diese Sequenz zusammen mit den fünf Haemoglobinsequenzen aus Aufgabe 1 in das *ClustalW2*-Eingabeformular ein und führe das Programm mit Standardeinstellungen aus.
2. Vergleiche das Alignment, das du jetzt erhältst, mit dem aus Aufgabe 1. Was fällt auf? Wo sind Unterschiede? Was ist gleich geblieben? Woran liegt das?
3. Betrachte nun den Guide Tree. Unterstützt er deine Beobachtungen?

## Aufgabe 3 (Untersuchung unbekannter Sequenzen mit ClustalW2)

1. Lade dir die Sequenzen *unknown\_query.fas* von der Sequenzanalyse-Seite herunter. Gegeben sind vier bekannte und eine unbekannt Sequenz. Kopiere alle fünf Sequenzen in das Eingabefenster von *ClustalW2* und führe das Programm mit Standardeinstellungen aus.
2. Zu welcher Sequenz hat die Query die höchste Ähnlichkeit? Um welche Sequenz handelt es sich bei dieser? Aus welchem Organismus stammt sie?
3. Was stellen die anderen drei Sequenzen dar und aus welchen Organismen kommen sie?
4. Gehe auf die Seite *Result Summary* des Outputs von *ClustalW2* und starte *JalView*, ein Programm mit grafischer Oberfläche, das dir zusätzlich zum Alignment auch die Konservierung und die Qualität des Alignments anzeigt. Beschreibe kurz, was du siehst (ca. fünf Sätze genügen).
5. Um was für eine Sequenz wird es sich bei der Query wahrscheinlich handeln? Kannst du deine Vermutung mit BLAST bestätigen? Welche Sequenz findest du?