

Übungsblatt 1

Aufgabe 1.1

Das Protein *Myoglobin* besitzt die Fähigkeit Sauerstoff im Blut aufzunehmen und an Muskelzellen abzugeben. Der Anfang der mRNA des *Myoglobin* hat die folgende Nucleotidsequenz¹:

aug ggg cuc agc gac ggg gaa ...

Welcher Aminosäuresequenz entspricht dies? Welche Basensubstitution führt zur gleichen Aminosäuresequenz? Welche führt zu einer anderen?

Aufgabe 1.2

Am *Large Hadron Collider* in CERN werden zu Spitzenzeiten bis zu 10 Gigabyte (1 GB = 2^{30} Byte) an Daten pro Sekunde an Rechenzentren überall auf der Welt verteilt². Insgesamt fallen pro Jahr etwa 15 Petabyte an Daten an (1 PB = 2^{50} Byte). Wie vielen menschlichen Genomäquivalenten entsprechen diese Datenmengen?

Hinweis: Nimm an, dass zur Speicherung der menschlichen DNA-Sequenz 2 Bit (1 Byte = 2^3 Bit) pro Nucleotidpaar erforderlich ist, und die DNA-Sequenz aus insgesamt $3 \cdot 10^9$ Nucleotidpaaren besteht (siehe Vorlesung).

Aufgabe 1.3

Was sind Daten, Informationen und was ist Wissen? Wie können Daten gespeichert werden? Was ist eine Datenbank, ein Datenbank-Managementsystem, ein Datenbanksystem? Welche Probleme können beim Integrieren verschiedener Datenbanken auftreten?

Aufgabe 1.4

Nutze den *SQLite Manager*³ oder ein vergleichbares Programm und erzeuge eine Datenbank mit mehreren Tabellen zum Speichern der cDNA von verschiedenen Spezies.

¹http://www.ncbi.nlm.nih.gov/nuccore/NM_203377

²<http://home.web.cern.ch/about/computing>

³<https://addons.mozilla.org/de/firefox/addon/sqlite-manager/>

- a) Lege die Tabellen so an, dass die folgenden Bedingungen eingehalten werden: In der Datenbank sollen sowohl die cDNA der Spezies, der bearbeitende Wissenschaftler, sowie das Datum des Eintrages gespeichert werden. Für einen Wissenschaftler muss dessen Name und Kontaktadresse abrufbar sein. Für eine Spezies soll sowohl ihr Name als auch eine kurze Beschreibung gespeichert werden können. Beachte, dass für jede Spezies mehrere Sequenz-Einträge möglich sein müssen. Außerdem soll jeder Wissenschaftler in der Lage sein, Einträge für verschiedene Spezies anzulegen.
- b) Schreibe eine SQL-Anfrage, die den Namen und Kontaktadresse der Wissenschaftler anzeigt, welche Einträge zur Spezies *Oryza sativa* oder *Hordeum vulgare* gemacht haben. Nutze dazu die in a) erzeugte Datenbank. Gib auch aus, wann der jeweilige Eintrag gemacht wurde. Füge anschließend einige Werte in die Datenbank ein und teste die Anfrage!

Hinweis: Eine Übersicht der SQL-Befehle mit einigen Beispielen findet man bspw. auf Wikibooks⁴.

⁴http://de.wikibooks.org/wiki/Einf%C3%BChrung_in_SQL