Übungen zum Sequenzanalyse-Praktikum

Universität Bielefeld, SoSe 2013
Prof. Dr. Jens Stoye · M.Sc. Stefan Janssen · B.Sc. Linda Sundermann
`http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2013summer/SequaPrak`
`praktikum-seqan@CeBiTec.Uni-Bielefeld.DE`

Übungsblatt 7 vom 24.05.2013
Abgabe 29.05.2013

Exercise 7.1:

1. Download `http://bibiserv.cebitec.uni-bielefeld.de/sadr2/databasesearch/blast/exercise1/sequence.fas`, a sequence that was just sequenced in a lab of your institution.
2. Run a *blastn* search on the *NCBI* server to identify similar sequences, that are already known.
3. Which gene has been sequenced here?
4. Is this the complete mRNA? How do you know?
5. Do a search against *SwissProt* (using *blastx*). Do you find all sequences you found in the nucleotide database?
6. Have a closer look at the alignments. Why are there small grey characters in the sequence? (Check the *BLAST* FAQs to find the answer.)

Exercise 7.2:

1. Use the sequence from Exercise 7.1 again. Perform a *blastx* search for *Homo sapiens* sequences against *nr* and *SwissProt*.
2. Look for a hit that appears in both outputs and compare the E-values: do they change? Why?

Exercise 7.3:

1. Use the unkown protein sequence from `http://bibiserv.cebitec.uni-bielefeld.de/sadr2/databasesearch/fasta/exercise1/sequence2.fas` and perform a *FASTA* search against *SwissProt*.
2. Run also a blastp search against *SwissProt*.
3. How do the results compare?

Exercise 7.4:

1. `http://bibiserv.cebitec.uni-bielefeld.de/sadr2/databasesearch/fasta/exercise2/sequence3.fas` is a sequence that has a sequencing error. Can you identify the position of this error?
2. Run a *blastx* search against *SwissProt*.
3. Run a *fasty* search against *SwissProt*.
4. How do the results compare?

Exercise 7.5:

1. Perform a "sequence search" with the *Pfam* database, (this is a *HMMER* run) with the same unknown protein sequence as in 7.3.
2. How many different hits do you get?
3. Why are there only so few hits, compared to the *FASTA* results?
4. What do you learn about the protein composition and maybe its function?