

Softwarewerkzeuge der Bioinformatik

Project Report

The reports presents the results and thus constitutes the basis of the assessment. It should be written in a way that the approach, results and discussion is easy to understand without the reader having to consult additional sources. All reports have to adhere to the following guidelines.

1 General Remarks

Language: The project report can be written in **German** or **English**.

Text: The report should be written in full sentences and follow the spelling and grammar rules of the respective language.

Format: There are **no** specific rules regarding font sizes, line spacing, etc. as long as it is readable.

Length: The report should contain the required information, no more, no less. The length of the reports can differ between projects. There are no bonus points for long reports, and unnecessarily long reports mean more work for everyone involved. It is therefore better to be concise and precise. It is possible to fully answer each task in 0.25 to 1.5 pages of pure text (excluding tables and figures).

Figures and Tables: Figures and results should be included in the text and not just linked. Figures and tables must be of a sufficient quality and be readable. All used figures and tables must be referenced and sufficiently explained in the text.

Softwaretools: Softwaretools introduced in the lecture or tutorial should be used if possible or not otherwise stated.

Plagiarism: Solutions and code should be written by you in your own words. Parts that were copied from other groups, the internet or submissions of previous years will be marked with 0 points.

Questions: You can ask questions during the tutorial or contact the responsible tutor via email.

Submission: The reports have to be submitted as a single **PDF** to the specified tutor. Each report should contain the name of the project as well as the names and matriculation numbers of the group members.

2 Structure

Document and justify your approach in your approach and specify which software tools (including chosen parameters), data bases (including accession numbers) or other sources were used. Additionally present and discuss your results.

For each task the following structure should be followed:

2.1 Introduction

Describe the goal of the task in 1–2 sentences.

- In this task, we examine...
- The goal of this task was...

2.2 Approach

Your approach should be described in detail, including parameter choice (e.g. *default settings, an E-value of 0.0001 was chosen,...*) and the employed software tools. You can use e.g. screenshots for this.

2.3 Presentation of the Results

Results have to be presented in a **comprehensible** manner. All results and claims have to be supported (e.g. with figures, screenshots,...). It has to be clear **what the text is referencing**. For example, the referenced sequence positions have to be clearly stated and marked (e.g. in motifs and sequence alignments). If necessary, a sensible selection should be made and described (e.g. select and discuss 5 proteins of 200).

2.4 Discussion and Comparisons

Some tasks ask for a discussion or a comparison of the results. This should only reference results that were presented in the report. Results that were not presented cannot be discussed or used in comparisons.

3 Example

Task: Compare the PRINTS- and PROSITE-motifs of your protein.

Proposed solution:

In this task we examined the motifs of the given protein. Subsequently, we compared the identified motifs. We used PRINTS and PROSITE to identify the motifs by following the corresponding links on the UniProtKB site of our protein.

The search in PRINTS and PROSITE returned several motifs for the given protein. PRINTS contained information about two motifs with the accession XXXX, whereas PROSITE yielded only one motif with the accession YYYYY. Table 1 presents the identified PRINTS- and PROSITE-motifs.

Table 1: Identified PRINTS- and PROSITE-motifs for the given protein.

Motif	PRINTS	PROSITE
Motif 1	GGLIGSIYGV	L-x(2)-S-I-x-[FYGWIV]-K
Motif 2	VTFCIYPLE	-

PRINTS-motif 1 consists of ten amino acids, whereas the PROSITE-motif is two amino acids shorter in comparison. Additionally, PRINTS-motif 2 is nine amino acids long.

Comparing PRINTS-motif 1 with the PROSITE-motif reveals a match between positions 3-9 of PRINTS-motif 1 (GGLIGSIYGV) and positions 1-7 of the PROSITE-motif (L-x(2)-S-I-x-[FYGWIV]-{K}). PRINTS-motif 2 on the other hand differs from the PROSITE-motif.