

projecet theses outline

Sebastian Thieme

April 18, 2010

1 Definition Of Task

Read mapping is one of the key disciplines of sequence analysis today. A lot of scientists are working with a lot of different mapping tools, which have a different efficiency and sensitivity. Hence, there is a need to show the advantages and disadvantages of the used methods for a better understanding and to work in an optimal way with algorithms.

This project thesis deals with a comparison between the BWA - Burrows-Wheeler Alignment Tool and Bowtie. For this comparison I used two not published datasets, which are provided by division Lehrach (vertebrate genomic) of the Max-Planck-Institute of molecular genetic. The problem is that on one hand both methods show similar results in one dataset, but on the other hand there is a big different in the second dataset between both algorithms. The task is to work out the advantages and disadvantages by using both algorithms. Furthermore, I will answer the question why the BWA show lower mapping results as contrasted with the Bowtie and where the differences between both are. Additionally, for a deeper analysis, I will work with a read mapper benchmark to analyse the performance of both. This mapper benchmark, written by Manuel Holtgrewe, is still under construction, but available to use. Last but not least, I will show how the behave is between the gapped-alignment of BWA in comparison to the ungapped-alignment of Bowtie and how many *false-positives/false-negatives* exists.

tasks in overview

1. work out adavantages and disadvantages of BWA and Bowtie.
2. How does they work?
3. Why does BWA show lower mapping results as contrasted with Bowtie?
4. Analysis with the read mapping benchmark written by Manuel Holtgrewe.
5. Where are the differences between the gapped-alignment of BWA in comparison to the ungapped-alignment of Bowtie?
6. How many *false-positives/false-negatives* exists (rate)?