

Advanced Algorithms - Exercise 3

Richter, Bachmann, Kadikowski, Pett, Genov, Mehnert

May 9, 2012

1 Application summary

- **ChiP-Seq**
ChiP stands for chromatin immunoprecipitation. The name gives very much away about the process itself. In short, proteins (i.e. transcription factor binding proteins) are added to chromatin aka DNA, where they bind. These new complex is then substracted by sonification and the small complexes resulting from this are basically fished out by antibody-reactions.
The proteins are then unlinked from the DNA in question, which is finally sequenced. This way one can for example find out about binding site structures for proteins and protein interaction.
- **RNA-Seq**
In RNA-Seq, the whole transcriptome of a cell or sample is sequenced. That is, all of the cDNA contained in a sample. This way, information about the activity of single genes can be derived, which is very interesting in fields like cancer research.
- **Bisulfite mapping**
In Bisulfite mapping, the methylation of the DNA by Bisulfite is used to find exact positions of methylation-based polymorphisms in a single strand DNA. It is based on the fact that Bisulfite preferably deaminates unmethylated Cytosines, which are then converted into Uracil during desulfonation. The sequencing of the prepared DNA in comparison with the unchanged DNA then contains different information that can be used to differentiate SNPs.

2 Quick q-gram construction

There is a q-gram in every position of the text. Therefore, we can go linearly through the text, building the index on the fly.

Input: Text T, size q

Output: q-gram index qgi

```
qgi = {} \\suitable data structure (e.g. dictionary)
```

```
for i = 1, i < length of text - q, i++ \\look at every element of the text
```

```
    qgram = T[i+q] \\take the qgram starting in i
```

```
    if qgram in qgi: \\either add to existing qgram appearance
```

```
        add i to list of appearances
```

```
    else \\or make a new entry in the index
```

```
        add qgram to qgi with position i
```

```
return qgi
```