Practical 2

Visualizing alignment results

Concepts and main aspects of RNA-Seq
WTCHG, 25 April 2016
Irina Pulyakhina

To be able to access a session in the UCSC genome browser that has already been created for you, go to:

```
https://genome-euro.ucsc.edu/cgi-
bin/hgTracks?
hgS_doOtherUser=submit&hgS_otherUserN
ame=pulyakhina&hgS_otherUserSessionNa
me=2016%2DRNAseq_course
```

You have to make sure you copy the whole link not interrupted by any new lines or empty spaces.

In this UCSC session you will find a BAM file and a BigWig file generated for the same sample. Both files contain only the data that was mapped to chromosome 22, therefore the rest of the genome is “empty” in both files.

1) BAM file

Click on the grey area on the left side of the panel containing a BAM file. There you see a popup menu “Display mode”. Try different modes and look at the results. You will have to zoom-in to a smaller window (try chr22:22,642,809-22,664,968).

Try different coloring options. Which coloring schema makes it easier to see mismatches in aligned reads?

Exercise:

- At the moment we do not see the connection between two ends of the same fragment, all reads look as if they were single-end reads. How to draw a line between two ends of the same fragment?
- Look at the edges of the exons. Are all reads always split exactly at the exon-intron boundary?
- Can you change the name displayed for the BAM file? (at the moment “BAM file” on the left on the actual genomic region).

2) BigWig file

Click on the grey are on the left side of the panel containing a BigWig file. Change “data view scaling” and check how it influences the appearance of the data. Try to zoom-in and -out to see whether reads in the BAM file correspond to the coverage peaks in the BigWig file.

Also try changing the type of graph from “bars” to “points” and look at the effect on the data.

Click to “configure” under the genome window. Try to add/remove vertical blue lines, change labela area width, text size and image width and see how it changes the appearance of the browser window.

Exercise:

- Zoom-in to an exonic region. You see that the coverage is quite jumpy. How can you smoothen it?
- How can you save a high-quality PDF with the current browser view?
- Does the coverage of all exons within a randomly taken gene look even?
- Are exons and UTRs covered equally?
- Are 5' UTRs and 3' UTRs covered equally?
- Look at different annotations (UCSC, Ensembl, RefSeq). Looking at the coverage and comparing different annotations, can you find an exon present in your data and annotated in only one annotation?