Genomic Computing

Simone Mosciatti

April 2, 2017

1 Exercise 1

1.1 A

In the region there are two genes, I can see the beginning of SUPT3H in the negative strand and the beginning of RUNX2 on the positive strand.

1.2 B

It is possible to notice a peak of the GC Percent just at the very beginning of SUPT3H exactly in a CpG island.

Similarly we can see a peak of the GC Percent at the beginning of the third intron of RUNX2, which is, as well, inside a CpG island.

It is curius to notice that at the very beginning of RUNX2, roughly 100kb upstream the peak of the GC is way mildier.

The first CpG island following the positive strand, the one associated with SUPT3H is in position: chr6:45345186-45346261, with chromosome band: 6p21.1 and genomic size of 1076.

2 Exercise 2

2.1 A

Definitely no, the first three reads maps in two different regions, the fourth and the fifth maps in three different region while the last map in as much as four different regions.

2.2 B / C

There is an almost perfect match of all the reads with the gene 17 in this coordinates: chr17:41,256,928-41,258,548

The coordinates correspond to the fourth intron of gene BRCA1.

2.3 C

This reads, most likely, come from gene 17.