Part 3 -

Load the attached mouse files to Galaxy. They are ungroomed single-end FASTQ files with Illumina 1.5 phred encoding from a ChIP-seq experiment and downsampled to a part of chromosome 19. In Galaxy, run the **FASTQ Groomer** tool to convert the reads to fastqsanger format. Then, use **Trimmomatic** to require a phred score greater than or equal to 20. Align the trimmed reads to the mm9 reference with **Map with BWA**. Finally, run **MACS2 callpeak** on the experimental ChIP-seq with the control output as the control.

***On Trimmomatic, the sliding window trimming operation was used with 4 bases average across, and 20 average quality minimum.**

- Retrieve the peaks in tabular format. Find the interval chr19:37,340,169-37,340,716. List the value in the fold_enrichment column.

27.13470

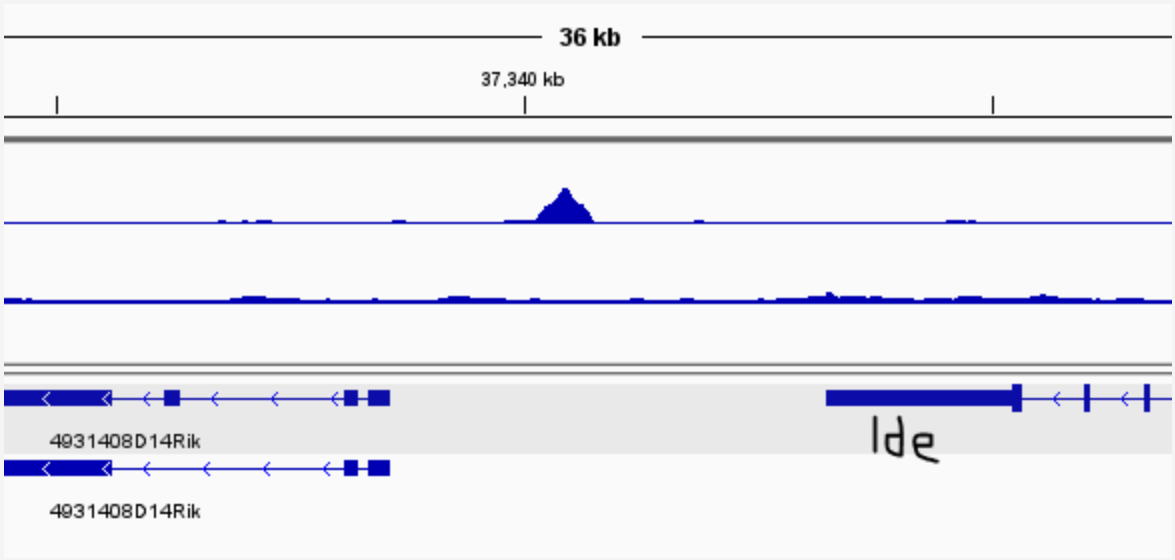
- Load both bedgraph files into IGV, mm9. Go to the interval from Part 3a. What is the nearest transcript?

The nearest transcript from Refseq genes is the gene 4931408D14Rik.

- Relative to the nearest two genes, where (upstream, exon, intron, downstream) is the MACS peak?

The MACS peak is between the two nearest genes of 4931408D14Rik and Ide. It is downstream from 4931408D14Rik and upstream from Ide.

- Submit a screenshot from IGV showing both the MACS peak and a small portion of the nearest two genes.



***Fuller view also attached**