NGS Bioinformatics

Module 9 ChiPSeq Practical assignment

Module topic: ChiPSeq
Contact session title: ChiPSeq
Trainer: Vivek Iyer
Participant: *<write your name here>*
Date: *<write today's date here>*

Introduction

All information is found in ChiPSeq practical material. We will use this document to answer the questions.

Please note

- **Hand-in information** please upload your completed assignment to the Vula 'Assignments' tab. Take note of the final hand-in date for each assignment, which will be indicated on Vula.

**Task 1: S3 Q1 How can you distinguish between the header of the SAM format and the actual align-
ments?**

<start typing your answer here>
_____
_____

Task 2: S3 Q2. What information does the header provide you with?

<start typing your answer here>
_____
_____

Task 3: S3 Q3. Which chromosome are the reads mapped to?

<start typing your answer here>
_____
_____

Task 4: S5 Q1. Look for gene NASP in the search box. Can you see a PAX5 binding site near the NASP gene?

<start typing your answer here>

_____
_____

Task 5: S5 Q2. What is the main difference between the visualisation of BAM and bigWig files?

<start typing your answer here>

_____
_____

Task 6: S8 Q1. The simplest bed file contains just three columns (chromosome, start, end) and is often called BED3 format. What extra columns does BED6 contain?

<start typing your answer here>

_____
_____

Task 7: S8 Q2. In the above examples, what are the lengths of the intervals?

<start typing your answer here>

_____
_____

Task 8: S8 Q3. Can you output a BED6 format with a transcript called "loc1", transcribed on the forward strand and having three exons of length 100 starting at positions 1000, 2000 and 3000?

<start typing your answer here>

_____
_____

Task 9: S8 Q4. What additional information is given in the narrowPeak file, beside the location of the peaks?

<start typing your answer here>

_____
_____

Task 10: S8 Q5. Does the first peak that was called look convincing to you?

<start typing your answer here>

_____

_____

Task 11: S9 Q1. Looking at the output of the bedtools genomecov we ran, what percentage of chromosome 1 do the peaks of PAX5 cover?

<start typing your answer here>

_____

_____

Task 12: S9 Q2. Looking at the output from bedtools intersect, what proportion of PAX5 peaks overlap genes?

<start typing your answer here>

_____

_____

Task 13: S9 Q3. Looking at PAX5_closestTSS.txt, which gene was found to be closest to MACS peak 2?

<start typing your answer here>

_____

_____

Task 14: S10 Q1. Which motif was found to be the most similar to your motif?

<start typing your answer here>

_____

_____