Brad Olinger

Quant Bio

Week 4 Command Notes

10/8/21

Bisulfite mapping with Bismark

Index the reference genome

Command: bismark\_genome\_preparation –parallel 7 .

Map both experiments using Bismark:

Command   
(whole cell mass): Bismark index/ -1 SRR3083926\_1.chr6.fastq -2 SRR3083926\_2.chr6.fastq -B 4.5\_whole

Command (epiblast): Bismark index/ -1 SRR3083929\_1.chr6.fastq -2 SRR3083929\_2.chr6.fastq -B 5.5\_epi

Remove duplicate reads

Command: deduplicate\_bismark 4.5\_whole\_pe.bam 5.5\_epi\_pe.bam

Sort the Bam Files

Command: samtools sort -o 4.sorted.bam 4.5\_whole\_pe.deduplicated.bam

samtools sort -o 5.5.sorted.bam 5.5\_epi\_pe.deduplicated.bam

Index the bam files:

Command: samtools index 4.sorted.bam

Command: samtools index 5.5.sorted.bam

Extract Methylation data

Command: bismark\_methylation\_extractor --bedgraph --comprehensive 4.5\_whole\_pe.deduplicated.bam

Final Image:



Extract promotors and write them to a bed file

Command: awk 'BEGIN{OFS="\t"}{if ($4 == "+") print $3,$5 - 2000,$5,$13,$12,$4; else print $3,$6,$6 + 2000,$13,$12,$4;}' mm10\_refseq\_genes\_chr6\_50M\_60M.bed | grep -v Rik | uniq -f 3 | sort -k2,2n > promoters.bed

Analysis

For each promotor, find the total methylation signal

command: bedtools map -a promoters.bed -b 4.5\_whole\_pe.deduplicated.bedgraph -c 4 -o sum > stage4.txt

bedtools map -a promoters.bed -b 5.5\_epi\_pe.deduplicated.bedgraph -c 4 -o sum > stage5.txt

Note: analysis section is in python script