BIOINFORMATICS COMMANDS

BASH

> receiving input then giving it into a file

grep – “search” function, ( -v is not, “^ ..” is starts with)

zipping files 🡪 bgzip, unzipping files 🡪 gunzip

cut -f #-# 🡪 is cutting fields (columns) from the first num to the second num

wc -l 🡪 is a wordcount but counts lines

indels 🡪 awk '! /\#/' variants.vcf | awk '{if(length($4) > 1 || length($5) > 1) print}'

using bcftools to show all of chromosome one🡪 module load tabix & bcftools, bgzip vcf file, tabix -p vcf filename.vcf.gz, bcftools view -r 1 vcffile.gz > output file.vcf

multiallelic sites 🡪 awk ‘$column# ~/,/’ filename.vcf

Exercise 3 (BAM File Commands)

Showing header lines 🡪 module load samtools, samtools view -H bamfilename.bam

Extract region from file, save in another bam file 🡪 samtools view -b bamfilename.bam “[insert chr#: # - #]” > output.bam

Extracting first # of reads 🡪 samtools view bamfilename.bam | head –[insert #]