Module loading:

module load sra

module load bowtie2

module load samtools

module load picard

Data download:

fasterq-dump ACCESSION #

read merging

/blue/env6932/share/SeqPrep/SeqPrep -f ACCESSION\_1.fq -2 ACCESSION\_2.fq -1 ACCESSION.f.fq.gz -2 ACCESSION.R.fq.gz -s ACCESSION.M.fq.gz

Decompress the files

Gunzip ACCESSION\*fq.gz

Genome indexing

Bowtie2-build –-threads 8 Genomefile.fna Genomefile.fna

Eg. Bowtie2-build –threads 8 GCA\_003718315.1\_ASM371831v1\_genomic.fna GCA\_003718315.1\_ASM371831v1\_genomic.fna

Paired end pipeline:

bowtie2 -x Genomefile.fna -1 ACCESSION.f.fq -2 ACCESSION.r.fq

| samtools view -q 30 -u –

| samtools sort -o ACCESSION.PE.sort.bam

samtools rmdup ACCESSION.PE.sort.bam ACCESSION.PE.rmdup.bam

Single end :

bowtie2 -x Genomefile.fna -U ACCESSION.m.fq

| samtools view -q 30 -u –

| samtools sort -o ACCESSION.SE.sort.bam

samtools rmdup -s ACCESSION.SE.sort.bam ACCESSION.SE.rmdup.bam

samtools merge ACCESSION.merged.unsort.bam ACCESSION.SE.rmdup.bam ACCESSION.PE.rmdup.bam (join files in one)

samtools sort -o ACCESSION.final.bam ACCESSION.merged.unsort.bam (sort the reads)