Commands for Miseq analysis

* To check fastq quality for each file

fastqc \*.fastq

mv \*.txt/.html/.zip new folder location

* **To trim the files**

pip3 install --user --upgrade cutadapt

curl -fsSL [https://github.com/FelixKrueger/TrimGalore/archive/0.6.0.tar.gz -o trim\_galore.tar.gz](https://github.com/FelixKrueger/TrimGalore/archive/0.6.0.tar.gz%20-o%20trim_galore.tar.gz)

tar xvzf trim\_galore.tar.gz

for f in \*\_R1\_001.fastq; do r1=$f ; r2=${f/\_R1\_001.fastq/}\_R2\_001.fastq; TrimGalore-0.6.0/trim\_galore --paired $r1 $r2 -o /Volumes/archive/macknightlab/RYEGRASS\_sequencing/Miseq\_Samarth/ryegrass\_miseq/untrimmed/trimmed/ --phred33 --length 36 --fastqc --fastqc\_args "--outdir /Volumes/archive/macknightlab/RYEGRASS\_sequencing/Miseq\_Samarth/ryegrass\_miseq/untrimmed/trimmed/QC" --path\_to\_cutadapt /Volumes/archive/macknightlab/RYEGRASS\_sequencing/Miseq\_Samarth/ryegrass\_miseq/untrimmed/cutadapt; done

* **To index reference file**

bwa index lope\_V1.0.fasta samtools faidx lope\_V1.0.fasta

* **To align files and generate .sam file**

for f in /Volumes/archive/trimmed/\*R1\_001\_val\_1.fq; do r1=$f; r2=${f/R1\_001\_val\_1.fq}R2\_001\_val\_2.fq; out1=${f##\*/}; out2=${out1/\_fwd\*/}; bwa mem -M lope\_V1.0.fasta $r1 $r2 > /Volumes/archive/trimmed/Scaffold\_aligned\_files/$out2.sam; done

* **To get a .bam file**

for f in /Volumes/archive/trimmed/scaffold\_aligned\_files/\*.sam; do r1=$f; r2=${f##\*/}; r3=${r2/.sam/}.bam; samtools sort $r1 > /Volumes/archive/trimmed/scaffold\_aligned\_files/$r3; done

These files can be viewed at Geneious

**Variant calling analysis**

* **To index reference sequence for variant calling**

java -jar /Volumes/archive/trimmed/scaffold\_aligned\_files/picard/build/libs/picard.jar CreateSequenceDictionary -R norm\_refgenome.fasta

* To index bam files for variant calling

for f in \*RG.bam; do java -jar /Volumes/archive/trimmed/scaffold\_aligned\_files/picard/build/libs/picard.jar BuildBamIndex -I $f; done

* **To Addorreplace read groups**

for f in /Volumes/archive/trimmed/scaffold\_aligned\_files/\*.bam; do r1=$f; r2=${f##\*/}; r3=${r2/.bam/}; r4=${r2/.bam/}\_RG.bam; java -jar /Volumes/archive/trimmed/scaffold\_aligned\_files/picard/build/libs/picard.jar AddOrReplaceReadGroups I=$r1 O=$r4 RGID=4 RGLB=lib1 RGPL=illumina RGPU=unit1 RGSM=20; done

* **To run haplotype caller**

for f in \*RG.bam; do r1=$f; r2=${f/\_RG.bam/}.vcf.gz; /Volumes/archive/trimmed/gatk-4.1.9.0/gatk --java-options "-Xmx8g" HaplotypeCaller -R /Volumes/archive/trimmed/norm\_refgenome.fasta -I $f -O $r2; done