

```
# Pre Alignment QC
```

```
# 1.0 fastqc on raw data
```

```
fastqc SRR4104638_R1.fastq SRR4104638_R2.fastq
```

```
# 2.0 trimomatic
```

```
Trimmomatic \
```

```
PE \
```

```
-phred33 \
```

```
SRR4104638_R1.fastq SRR4104638_R2.fastq \
```

```
SRR4104638_R1_paired.fq.gz SRR4104638_R1_unpaired.fq.gz \
```

```
SRR4104638_R2_paired.fq.gz SRR4104638_R2_unpaired.fq.gz \
```

```
ILLUMINACLIP:all_adapter.fa:2:30:10 \
```

```
LEADING:3 \
```

```
TRAILING:3 \
```

```
SLIDINGWINDOW:4:15 \
```

```
MINLEN:50
```

```
# 3.0 fastqc on paired trimmed fastq data
```

```
# you can also run this on the unpaired data
```

```
fastqc SRR4104638_R1_paired.fq.gz SRR4104638_R2_paired.fq.gz
```

```
# Alignment
```

```
# 0.0 Index the reference genome
```

```
bwa index -a bwtsv Human.b38.fasta
```

```
# 1.0 alignment with BWA on trimmed data
```

```
bwa mem \
```

```
-t 12 \
```

```
-R '@RG ID:MiSeq SM:SRR4104638 PL:illumina LB:XT40
```

```
PU:H0164ALXX140820.2' \
```

```
Human.b38 \
```

```
SRR4104638_R1_paired.fq.gz \
```

```
SRR4104638_R2_paired.fq.gz > SRR4104638.b38.bwa.raw.sam
```

```
# 2.0 sort sam, convert to bam
```

```
samtools sort \
```

```
-O bam \
```

```
-o SRR4104638.b38.bwa.sorted.bam \
```

```
SRR4104638.b38.bwa.raw.sam
```

```
# 2.1 index bam
```

```
samtools index SRR4104638.b38.bwa.sorted.bam
```

```
# Basic Alignment post processing
```

```
# 1.0 Mark Duplicates
```

```
java -jar picard.jar MarkDuplicates \  
    --INPUT=SRR4104638.b38.bwa.sorted.bam \  
    --OUTPUT=SRR4104638.b38.bwa.marked_dup.bam \  
    --CREATE_INDEX=true \  
    --METRICS_FILE=SRR04638.b38.bwa.marked_dup_metrics.txt
```

```
# 2.0 Filter BAM File
```

```
# https://broadinstitute.github.io/picard/explain-flags.html
```

```
samtools view \  
-h \  
-b \  
-q 20 \  
-F 1796 \  
SRR4104638.b38.bwa.marked_dup.bam > SRR4104638.b38.bwa.filtered.bam
```

```
# 2.1 index filtered bam
```

```
samtools index SRR4104638.b38.bwa.filtered.bam
```

```
# Basic Variant calling
```

```
# 1.0 Call Variants with Freebayes
```

```
freebayes \
```

```
-f Human.b38.fasta \
```

```
SRR4104638.b38.bwa.filtered.bam > SRR4104638.b38.bwa.freebayes.raw.vcf
```

```
# 2.0 Hard Filter Variants
```

```
vcffilter \
```

```
-f "QUAL > 1 & QUAL / AO > 10 & SAF > 0 & SAR > 0 & RPR > 1 & RPL > 1" \
```

```
SRR4104638.b38.bwa.freebayes.raw.vcf >
```

```
SRR4104638.b38.bwa.freebayes.filtered.vcf
```