# Convert SAM/BAM file to R1 & R2 FASTQ.

samtools fastq -1 KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.R1.fastq -2 KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.R2.fastq KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bam

# Map Illumina FASTQ sequences to the PacBio reference assembly using BWA-MEM.

bwa mem K\_pneumoniae\_PB.fasta KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.R1.fastq KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.R2.fastq > KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.sam

# No longer need intermediate files, remove them.

\rm KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.R1.fastq

\rm KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.R2.fastq

# Convert SAM to BAM using SAMtools.

samtools view -Sb KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.sam > KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.bam

# No longer need intermediate files, remove them.

\rm KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.sam

# Positional sorting of the alignments using SAMtools.

samtools sort KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.bam -o KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.srt.bam

# No longer need intermediate files, remove them.

\rm KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.bam

# Removal of duplicate sequences using SAMtools.

samtools rmdup -S KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.srt.bam KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.dedup.bam

# No longer need intermediate files, remove them.

\rm KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.srt.bam

# Index the sorted BAM file using SAMtools.

samtools index KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.dedup.bam

# Generate a mapping stats file using SAMtools.

samtools flagstat KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.dedup.bam > KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.srt.bam.flagstat

# SNV calling using VarScan2. Requirements: 10X coverage; Q30

samtools mpileup -A -f K\_pneumoniae\_PB.fasta KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.dedup.bam | java -jar VarScan.v2.4.3.jar mpileup2snp --min-coverage 10 --min-avg-qual 30 --output-vcf 1 > KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.srt.bam.varscan

# Coverage metrics using Varscan2. Requirements: 10X coverage; Q30

samtools mpileup -A -f K\_pneumoniae\_PB.fasta KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.dedup.bam | awk '($4>0)' | java -jar VarScan.v2.4.3.jar readcounts --min-coverage 10 --min-base-qual 30 --output-vcf 1 --output-file KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.srt.bam.varscan.counts

# INDEL calling using VarScan2. Requirements: 10X coverage; Q30

samtools mpileup -A -f K\_pneumoniae\_PB.fasta KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.dedup.bam | java -jar VarScan.v2.4.3.jar mpileup2indel --min-coverage 10 --min-avg-qual 30 --output-vcf 1 > KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.srt.bam.varscan.indels

# Coverage metrics using SAMtools. (no depth requirements)

samtools depth -a KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.dedup.bam > KPAA-TOP52-U1\_gerald\_HCK5HBBXX\_8\_AGGCAGAA-CTCTCTAT.bwa.srt.bam.samtools.depth