

# CORSO DI METODI MOLECOLARI E BIOINFORMATICA

LM Biologia Evoluzionistica, Università di Padova

Docente: Dr. Andrea Binatti, Dr. Enrico Gaffo,

Prof. Stefania Bortoluzzi

## Esercitazione 6

Padova, 26 novembre 2021

### GUIDA COMANDI

***Nei campi preceduti dal simbolo “\$” significa che dovete inserire voi il file***

#### Allineamento sul genoma di riferimento

```
bwa mem -R "@RG\tID:sample\tLB:exome\tSM:sample\tPL:ILLUMINA" $genoma  
$primo_fastq $secondo_fastq > $reads_allineate_SAM
```

#### Conversione SAM in BAM

```
/opt/samtools/bin/samtools view -b $reads_allineate_SAM >  
$reads_allineate_BAM
```

#### Ordinamento delle reads allineate

```
/opt/samtools/bin/samtools sort $reads_allineate_BAM >  
$reads_ordinate_BAM
```

#### Indicizzazione del BAM

```
/opt/samtools/bin/samtools index $reads_ordinate_BAM
```

## Rimozione duplicati di PCR

```
/opt/gatk/gatk MarkDuplicates -I $reads_ordinate_BAM -O  
$no_duplicati_BAM -M metrics.txt -REMOVE_DUPLICATES true -CREATE_INDEX  
true
```

## Ricalibrazione della qualità delle basi

```
/opt/gatk/gatk BaseRecalibrator -I $no_duplicati_BAM -R $genoma  
--known-sites $db SNP -O model.grp
```

```
/opt/gatk/gatk ApplyBQSR -R $genoma -I $no_duplicati_BAM -bqsr  
model.grp -O $allineamento_ricalibrato_BAM
```

## Chiamata delle varianti germline

```
/opt/gatk/gatk HaplotypeCaller -R $genoma -I  
$allineamento_ricalibrato_BAM -O $germline_varianti_VCF
```

```
/opt/gatk/gatk VariantFiltration -V $germline_varianti_VCF -filter  
"QUAL < 30.0" --filter-name "QUAL30" -filter "MQ < 40.0" --filter-name  
"MQ40" -filter "DP < 30" --filter-name "DP30" -O  
$varianti_filtrate_VCF
```

```
/opt/gatk/gatk SelectVariants -R ref.fa -V $varianti_filtrate_VCF  
--exclude-filtered -O $varianti_selezionate_VCF
```

## Annotazione delle varianti germline

```
/opt/gatk/gatk VariantAnnotator -R $genoma -V $varianti_filtrate_VCF  
-O $varianti_annotate_VCF --db SNP $db SNP
```