

# GATK Test Protocol

Zhibo Huang

## Abstract

Analysis scripts of GATK pipeline

**Citation:** Zhibo Huang GATK Test Protocol. [protocols.io](https://doi.org/10.17504/protocols.io.mhdc326)

[dx.doi.org/10.17504/protocols.io.mhdc326](https://doi.org/10.17504/protocols.io.mhdc326)

**Published:** 10 Jan 2018

## Protocol

### GATK3 scripts

#### Step 1.

alignment

#### SOFTWARE PACKAGE (CentOS - )

**bwa, 0.7.15-r1140**

cmd [COMMAND \(CentOS - 7.3.1611\)](#)

```
bwa mem -R "@RG\tID:NA12878-test\tSM:NA12878-test\tLB:NA12878-test\tPL:illumina\tCN:BGI" -  
M -  
t 48 ${refpath}/human_g1k_v37.fasta ${datapath}/ERR194147_1.fastq.gz ${datapath}/ERR194147_  
2.fastq.gz | samtools view -S -b -o ${alignment}/na12878.raw.bam -  
alignment
```

#### EXPECTED RESULTS

`${alignment}/na12878.sort.bam`

### GATK3 scripts

#### Step 2.

rmDup

#### SOFTWARE PACKAGE

**Picard, 2.8** [\[link\]](#)

cmd [COMMAND](#)

```
samtools index ${alignment}/na12878.sort.dup.bam  
bamindex
```

#### EXPECTED RESULTS

`${alignment}/na12878.sort.dup.bam`

## GATK3 scripts

### Step 3.

BQSR

 SOFTWARE PACKAGE (CentOS - )

**GATK, 3.7**

cmd **COMMAND**

```
java -Xmx250g -jar ${software}/GenomeAnalysisTK.jar -R ${refpath}/human_g1k_v37.fasta -  
T BaseRecalibrator -nct 48 -I ${alignment}/na12878.sort.dup.bam -  
knownSites ${refpath}/dbSNP_138.b37.vcf -o ${alignment}/na12878.recal_data.grp  
BQSR
```

## GATK3 scripts

### Step 4.

Printreads

 SOFTWARE PACKAGE (CentOS - )

**GATK, 3.7**

cmd **COMMAND**

```
java -Xmx250g -jar ${software}/GenomeAnalysisTK.jar -R ${refpath}/human_g1k_v37.fasta -  
T PrintReads -nct 7 -filterNoBases -I ${alignment}/na12878.sort.dup.bam -  
BQSR ${alignment}/na12878.recal_data.grp -o ${alignment}/na12878.sort.recal.bam  
Printreads
```

## GATK3 scripts

### Step 5.

Variation

 SOFTWARE PACKAGE (CentOS - )

**GATK, 3.7**

cmd **COMMAND**

```
java -Xmx250g -jar ${GATK}/GenomeAnalysisTK.jar -l INFO -  
R ${refpath}/human_g1k_v37.fasta -T HaplotypeCaller -rf BadCigar -  
I ${alignment}/na12878.final.bam -o ${variation}/na12878.vcf.gz -nct 48 -  
D ${refpath}/dbSNP_138.b37.vcf  
Variation
```