

# Qualimap Analysis Results

*BAM QC analysis*

*Generated by Qualimap v.2.2.1*

*2016/10/23 11:33:53*

# 1. Input data & parameters

## 1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/BB017.sorted.bam.ba
m -nw 400 -hm 3
```

## 1.2. Alignment

Command line:	bwa mem -M /home/vdp5/data/salvador_vivax_asia_2016/first-SAMEA2376790/pvivax_sal1_SAMEA2376790.fasta /home/vdp5/data/cambodia_samples/sequences_gz/BB017-BiooBarcode_34_CATGGC_R1.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/BB017-BiooBarcode_34_CATGGC_R2.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 11:33:53 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/BB017.sorted.bam.b

am

## 2. Summary

### 2.1. Globals

Reference size	23,958,997
Number of reads	16,955,752
Mapped reads	8,570,672 / 50.55%
Unmapped reads	8,385,080 / 49.45%
Mapped paired reads	8,570,672 / 50.55%
Mapped reads, first in pair	4,295,399 / 25.33%
Mapped reads, second in pair	4,275,273 / 25.21%
Mapped reads, both in pair	8,195,794 / 48.34%
Mapped reads, singletons	374,878 / 2.21%
Read min/max/mean length	30 / 100 / 99.9
Duplicated reads (estimated)	5,158,213 / 30.42%
Duplication rate	60.63%
Clipped reads	811,423 / 4.79%

### 2.2. ACGT Content

Number/percentage of A's	238,258,936 / 28.69%
Number/percentage of C's	176,671,129 / 21.28%
Number/percentage of T's	239,197,216 / 28.81%
Number/percentage of G's	176,253,063 / 21.23%
Number/percentage of N's	69,743 / 0.01%
GC Percentage	42.5%

## 2.3. Coverage

Mean	34.6952
Standard Deviation	27.5148

## 2.4. Mapping Quality

Mean Mapping Quality	58.71
----------------------	-------

## 2.5. Insert size

Mean	1,039.17
Standard Deviation	29,063.04
P25/Median/P75	315 / 357 / 400

## 2.6. Mismatches and indels

General error rate	1.03%
Mismatches	7,958,082
Insertions	209,096
Mapped reads with at least one insertion	2.32%
Deletions	273,751
Mapped reads with at least one deletion	3.03%
Homopolymer indels	65.7%

## 2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

gi 1074120478 emb LT615256.1	977217	32681231	33.4432	13.2732
gi 1074120682 emb LT615257.1	860454	29498162	34.2821	20.7547
gi 1074120865 emb LT615258.1	989719	38284788	38.6825	38.1606
gi 1074121086 emb LT615259.1	935450	34726867	37.1232	37.5731
gi 1074121301 emb LT615260.1	1432239	52386399	36.5766	28.0941
gi 1074121615 emb LT615261.1	1080962	37367913	34.5691	17.9564
gi 1074121871 emb LT615262.1	1545099	52415309	33.9236	13.1465
gi 1074122235 emb LT615263.1	1585108	54591932	34.4405	21.3476
gi 1074122590 emb LT615264.1	2122358	72056519	33.9512	16.7787
gi 1074123050 emb LT615265.1	1754192	60918988	34.7277	31.9099
gi 1074123421 emb LT615	2150147	79294770	36.8788	59.4064

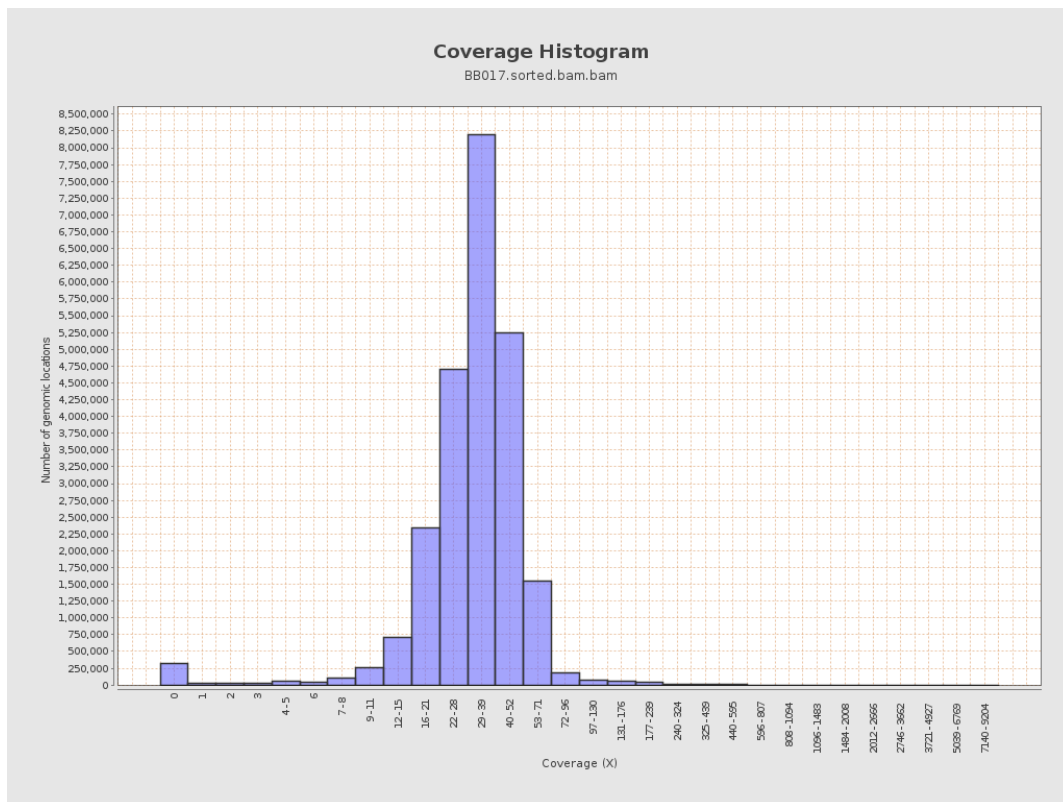
266.1				
gi 107412389 8 emb LT615 267.1	3031036	103597858	34.179	15.4087
gi 107412458 8 emb LT615 268.1	2359348	74918134	31.7537	21.7443
gi 107412506 5 emb LT615 269.1	3135668	108522881	34.6092	12.5932

### 3. Results : Coverage across reference

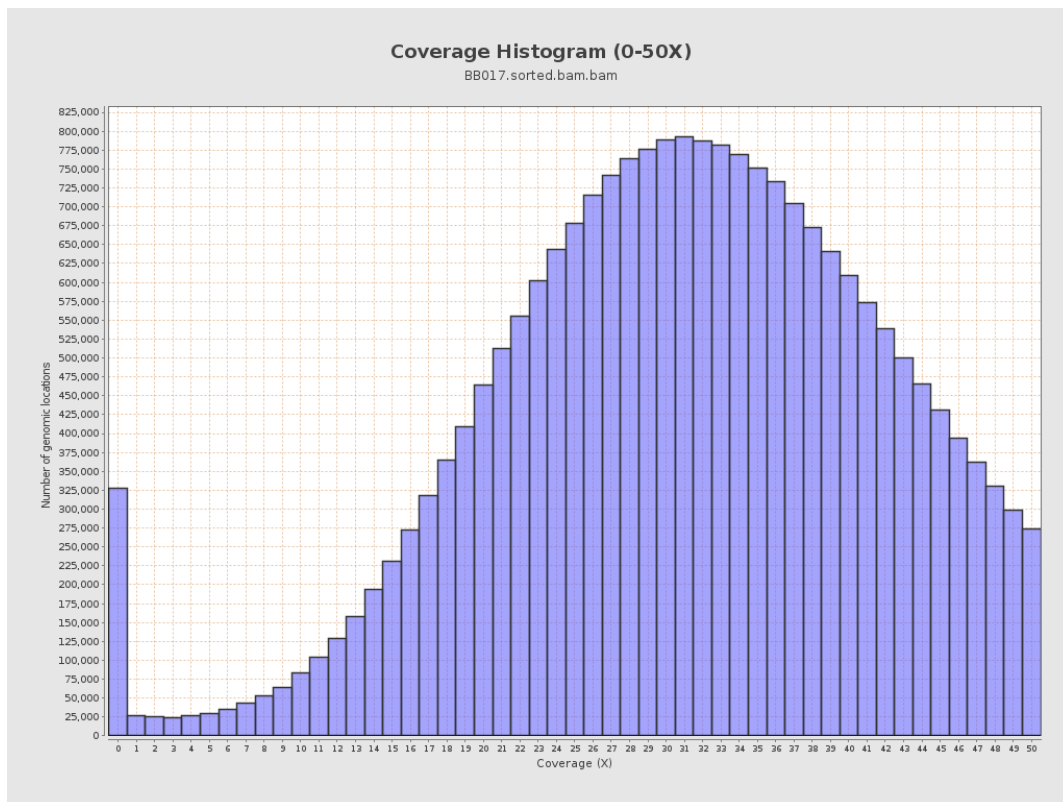




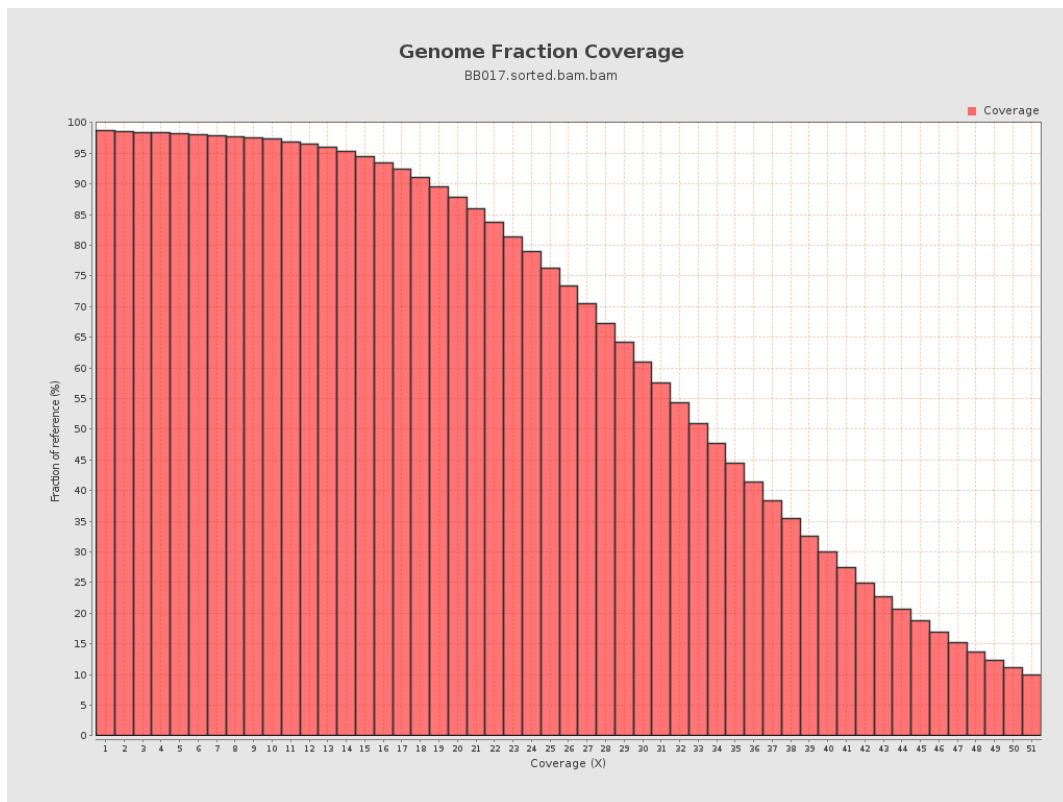
## 4. Results : Coverage Histogram



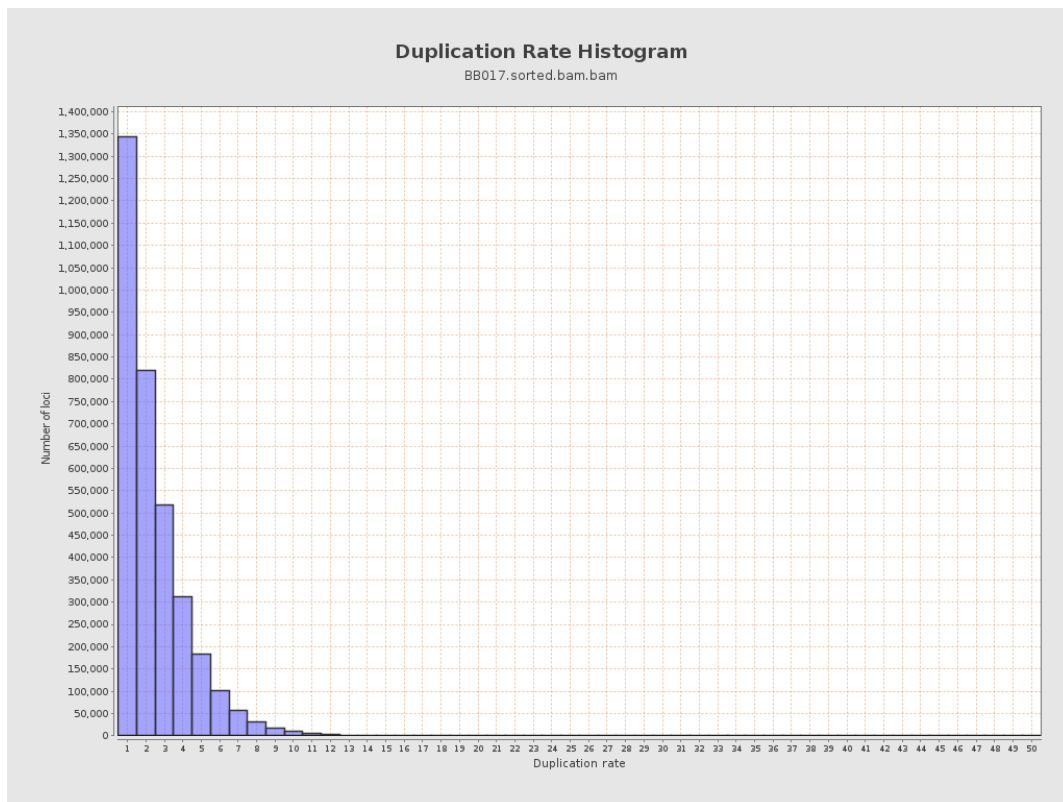
## 5. Results : Coverage Histogram (0-50X)



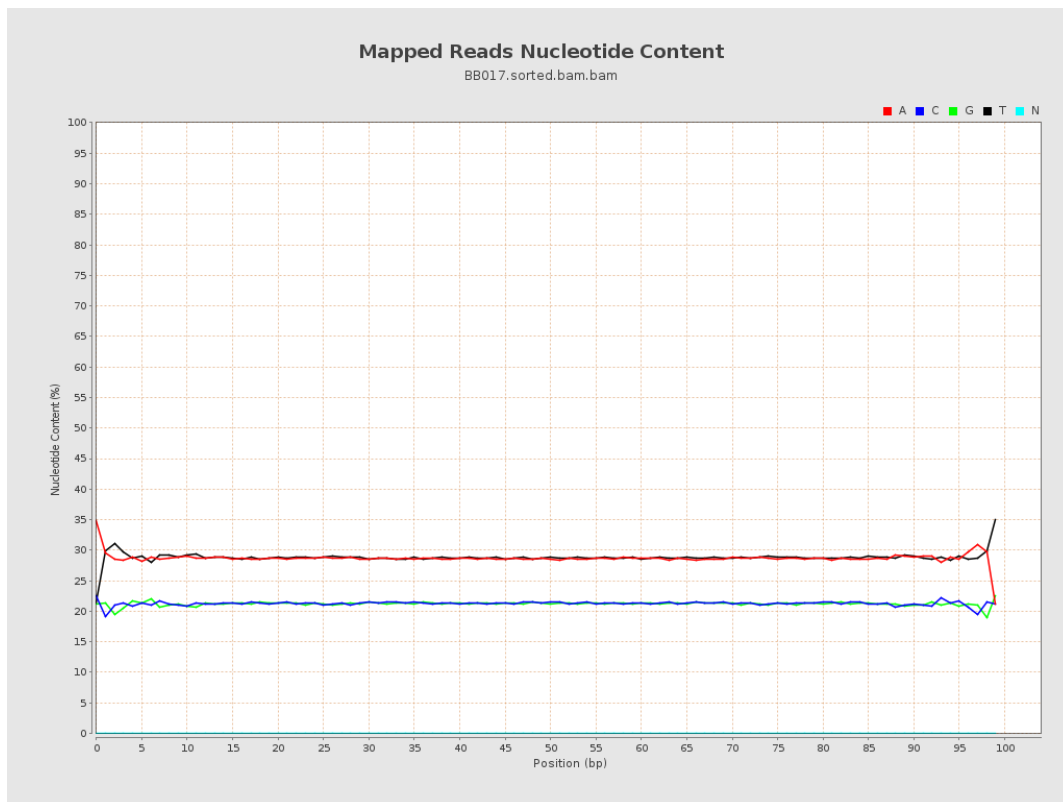
## 6. Results : Genome Fraction Coverage



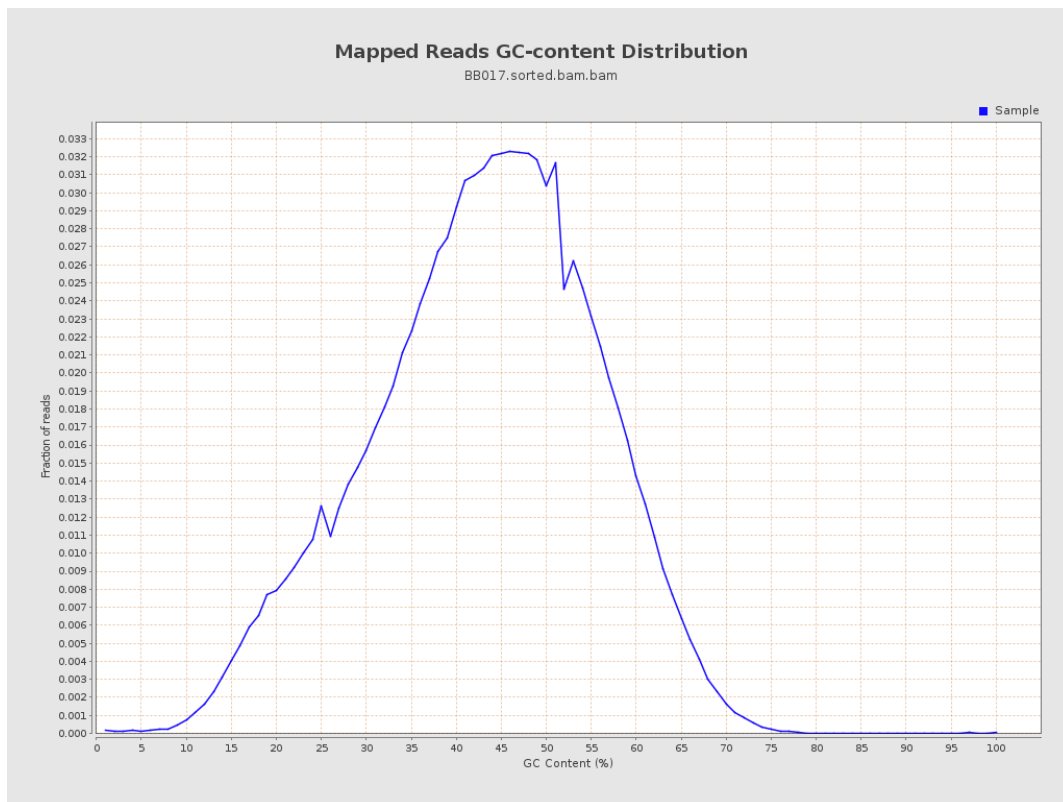
## 7. Results : Duplication Rate Histogram



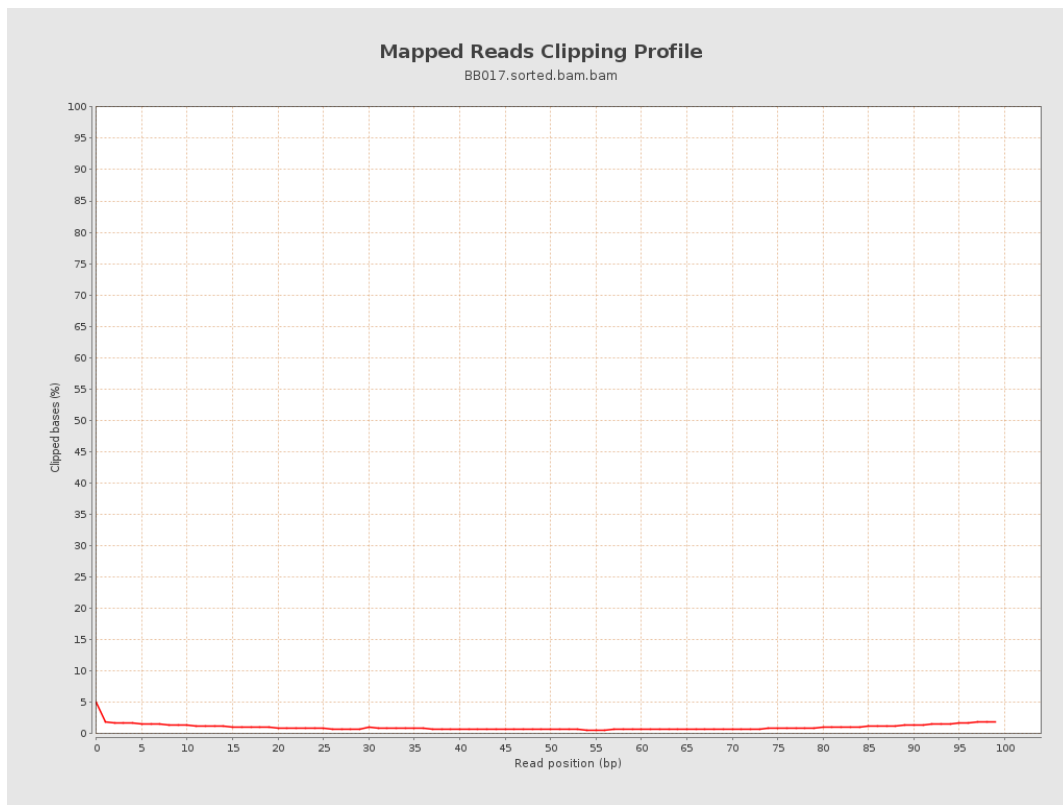
## 8. Results : Mapped Reads Nucleotide Content



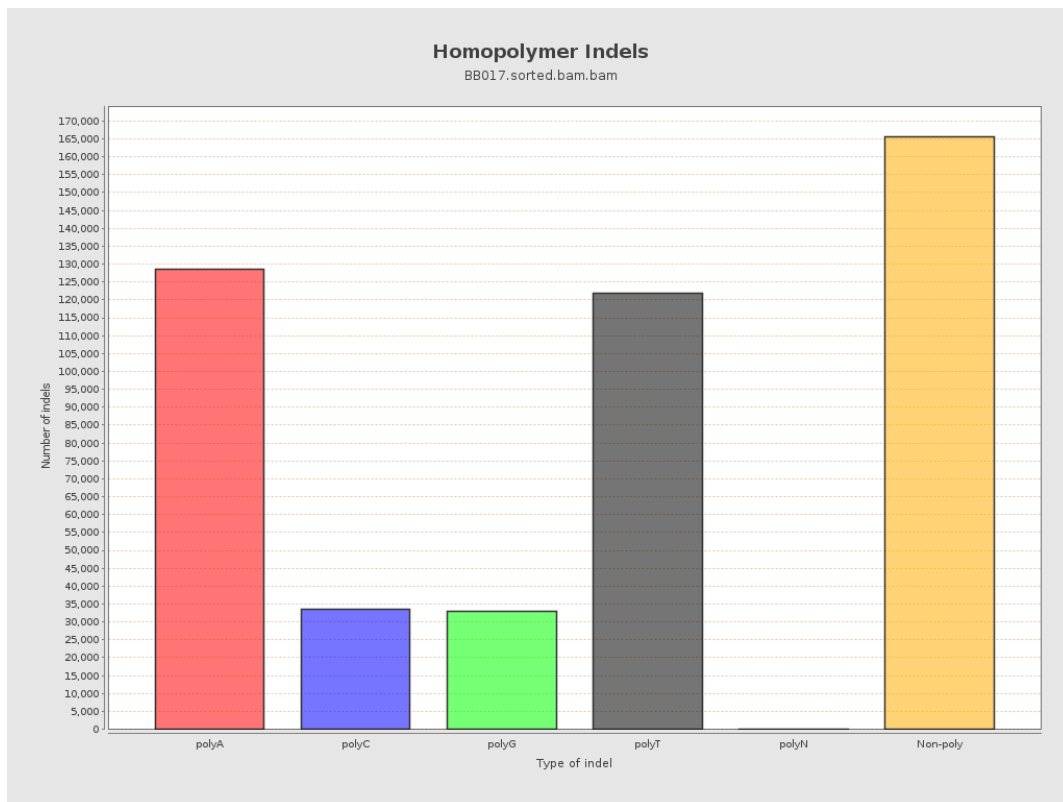
## 9. Results : Mapped Reads GC-content Distribution



## 10. Results : Mapped Reads Clipping Profile

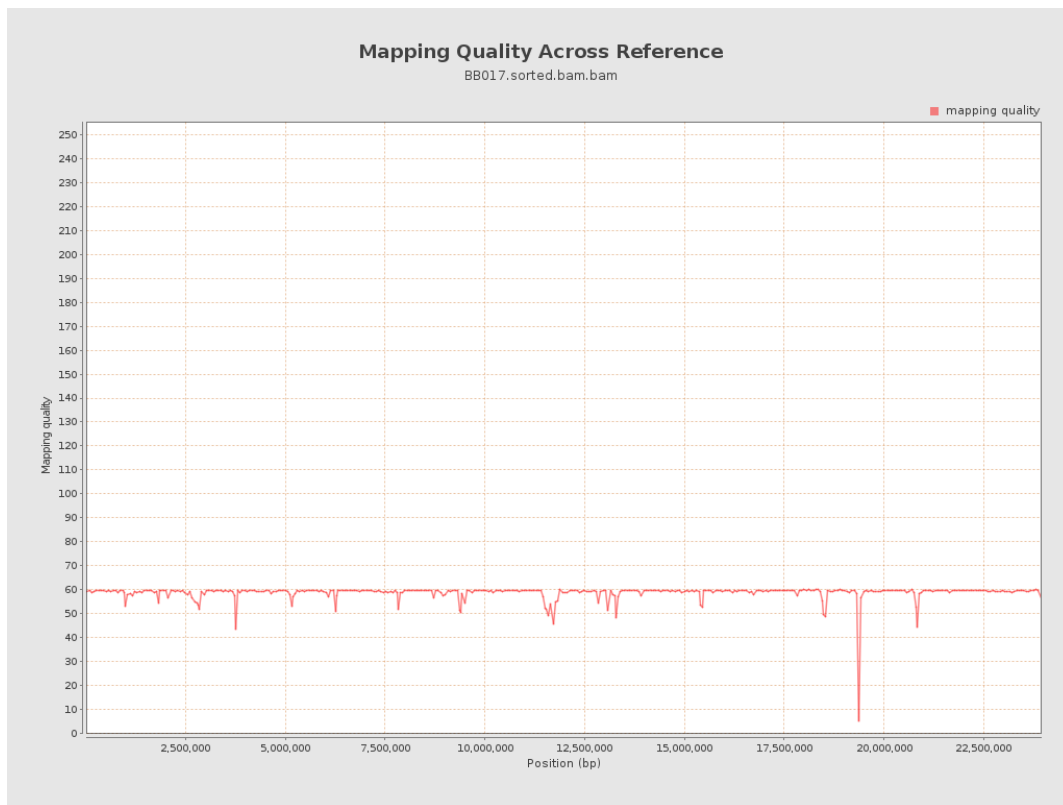


## 11. Results : Homopolymer Indels

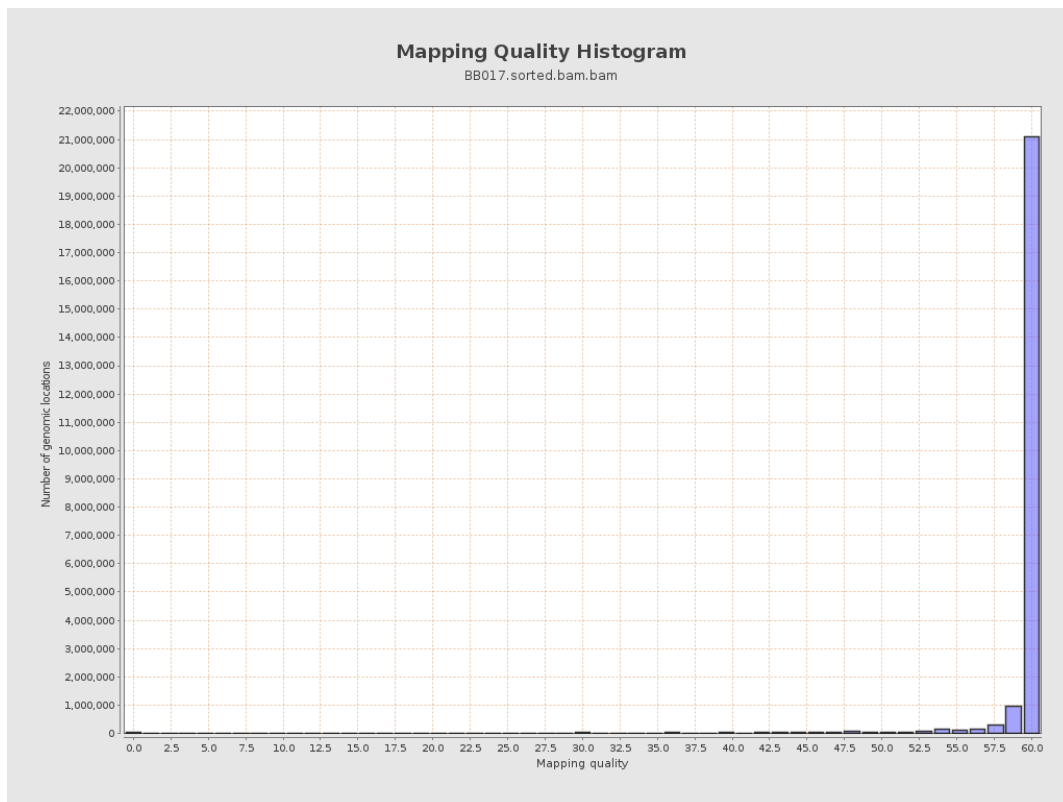




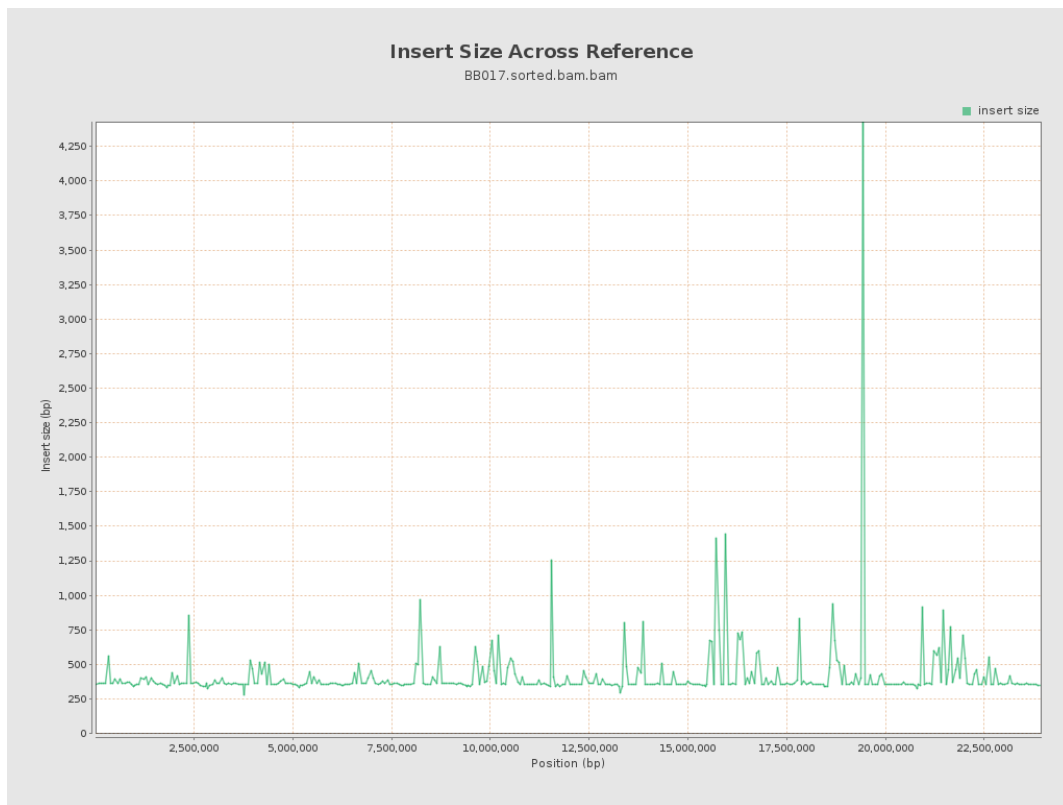
## 12. Results : Mapping Quality Across Reference



## 13. Results : Mapping Quality Histogram



## 14. Results : Insert Size Across Reference



## 15. Results : Insert Size Histogram

