

# Qualimap Analysis Results

*BAM QC analysis*

*Generated by Qualimap v.2.2.1*

*2016/10/23 11:40:38*

# 1. Input data & parameters

## 1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/BB035.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/BB035-BiooBarcode_39_CTATAC_R2.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/BB035-BiooBarcode_39_CTATAC_R1.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 11:40:38 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/BB035.sorted.bam.b

am

## 2. Summary

### 2.1. Globals

Reference size	23,958,997
Number of reads	13,122,212
Mapped reads	9,421,208 / 71.8%
Unmapped reads	3,701,004 / 28.2%
Mapped paired reads	9,421,208 / 71.8%
Mapped reads, first in pair	4,692,941 / 35.76%
Mapped reads, second in pair	4,728,267 / 36.03%
Mapped reads, both in pair	9,221,382 / 70.27%
Mapped reads, singletons	199,826 / 1.52%
Read min/max/mean length	30 / 100 / 99.86
Duplicated reads (estimated)	4,850,991 / 36.97%
Duplication rate	51.28%
Clipped reads	756,546 / 5.77%

### 2.2. ACGT Content

Number/percentage of A's	257,844,427 / 28.08%
Number/percentage of C's	200,969,175 / 21.89%
Number/percentage of T's	258,923,005 / 28.2%
Number/percentage of G's	200,536,322 / 21.84%
Number/percentage of N's	76,683 / 0.01%
GC Percentage	43.72%

## 2.3. Coverage

Mean	38.3663
Standard Deviation	21.8984

## 2.4. Mapping Quality

Mean Mapping Quality	58.81
----------------------	-------

## 2.5. Insert size

Mean	1,126.85
Standard Deviation	31,632.88
P25/Median/P75	318 / 365 / 414

## 2.6. Mismatches and indels

General error rate	0.9%
Mismatches	7,643,387
Insertions	206,624
Mapped reads with at least one insertion	2.1%
Deletions	271,462
Mapped reads with at least one deletion	2.76%
Homopolymer indels	67.62%

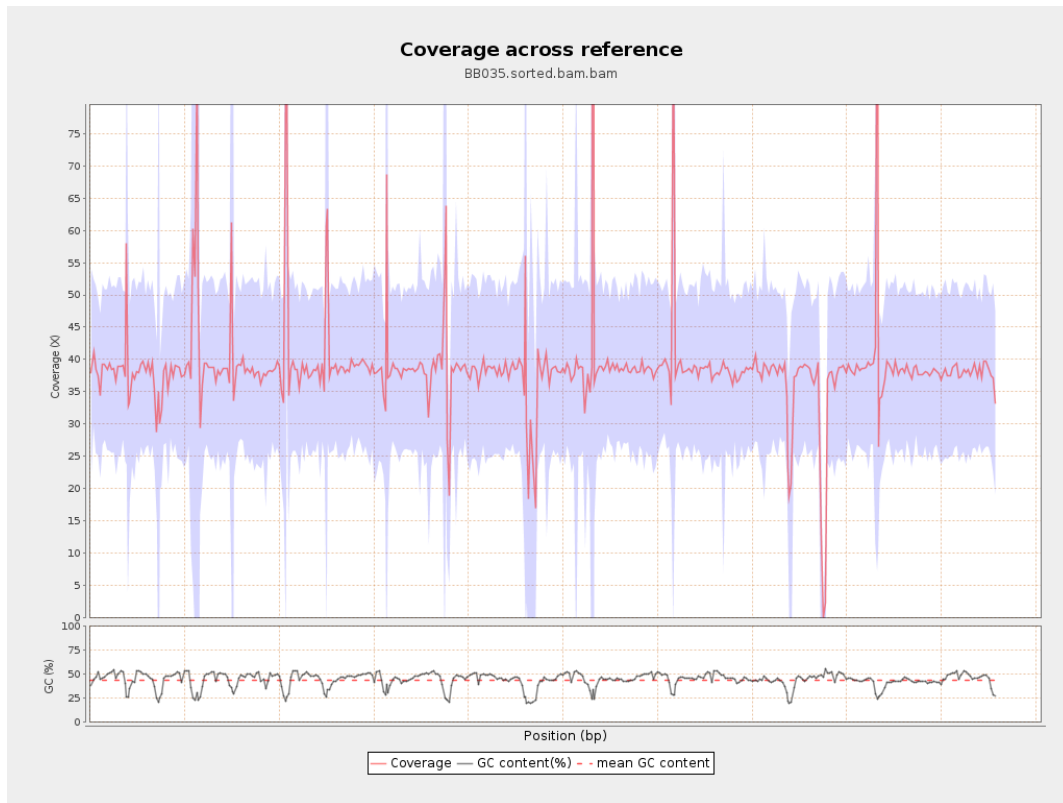
## 2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

gi 1074120478 emb LT615256.1	977217	37843982	38.7263	12.87
gi 1074120682 emb LT615257.1	860454	31643820	36.7757	18.0435
gi 1074120865 emb LT615258.1	989719	40070633	40.4869	27.0788
gi 1074121086 emb LT615259.1	935450	36886475	39.4318	28.4384
gi 1074121301 emb LT615260.1	1432239	56094897	39.1659	21.8913
gi 1074121615 emb LT615261.1	1080962	42711222	39.5122	19.8032
gi 1074121871 emb LT615262.1	1545099	59654597	38.6089	12.3818
gi 1074122235 emb LT615263.1	1585108	62553242	39.4631	20.336
gi 1074122590 emb LT615264.1	2122358	80147028	37.7632	14.9342
gi 1074123050 emb LT615265.1	1754192	63009247	35.9192	20.6653
gi 1074123421 emb LT615	2150147	87651596	40.7654	42.8036

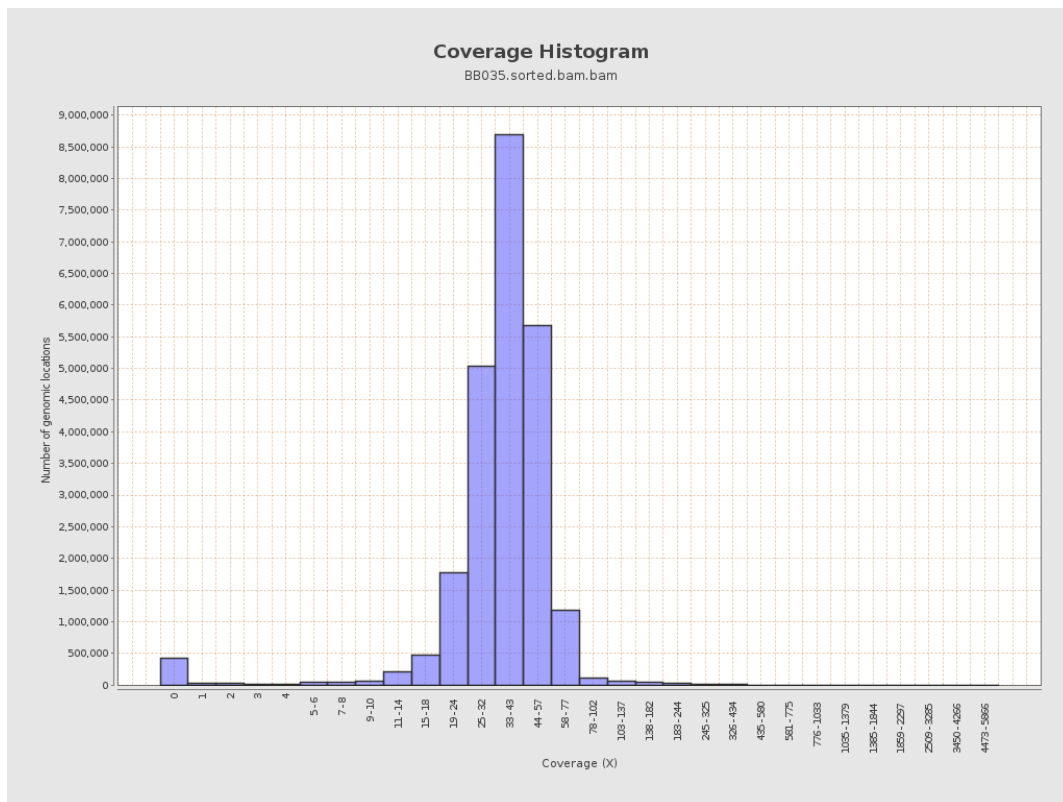
266.1				
gi 107412389 8 emb LT615 267.1	3031036	116350675	38.3864	13.4101
gi 107412458 8 emb LT615 268.1	2359348	85876156	36.3983	24.3523
gi 107412506 5 emb LT615 269.1	3135668	118725074	37.8628	11.4433

### 3. Results : Coverage across reference

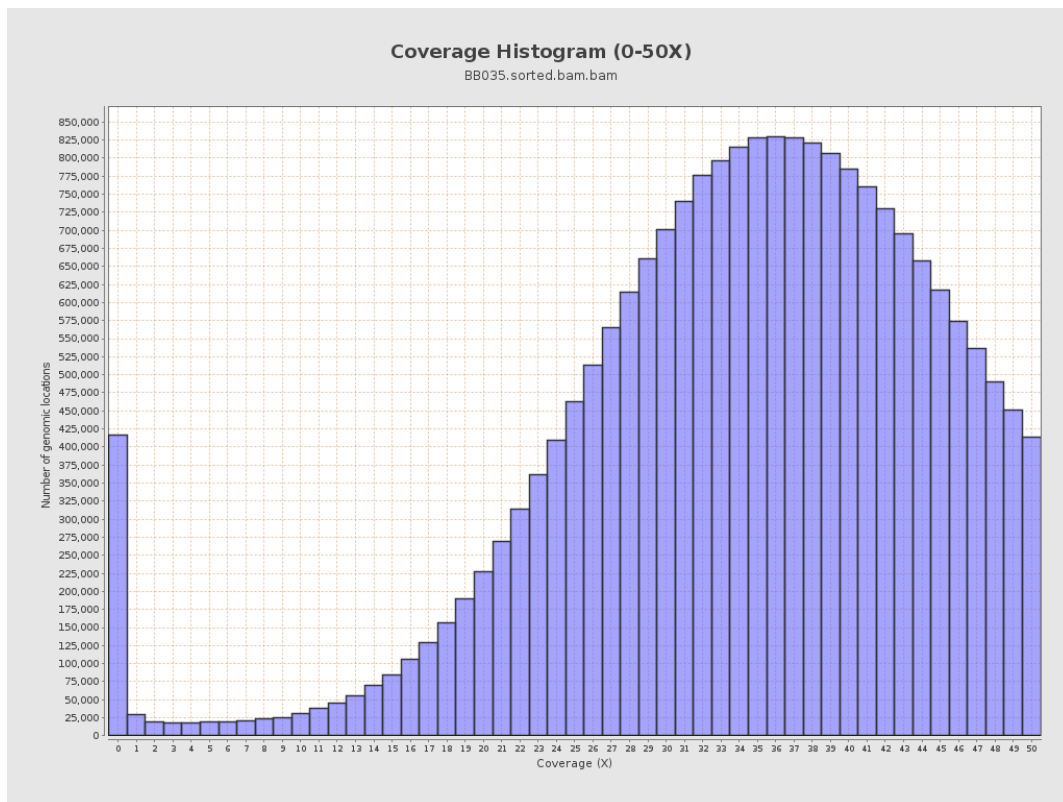




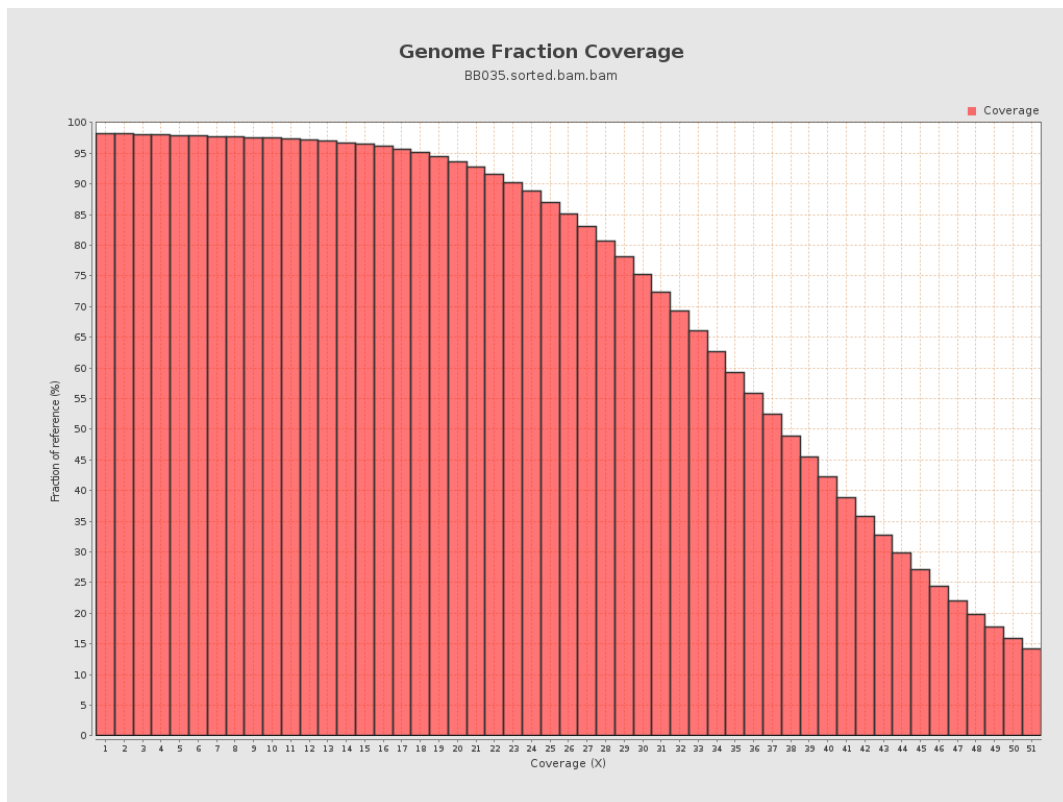
## 4. Results : Coverage Histogram



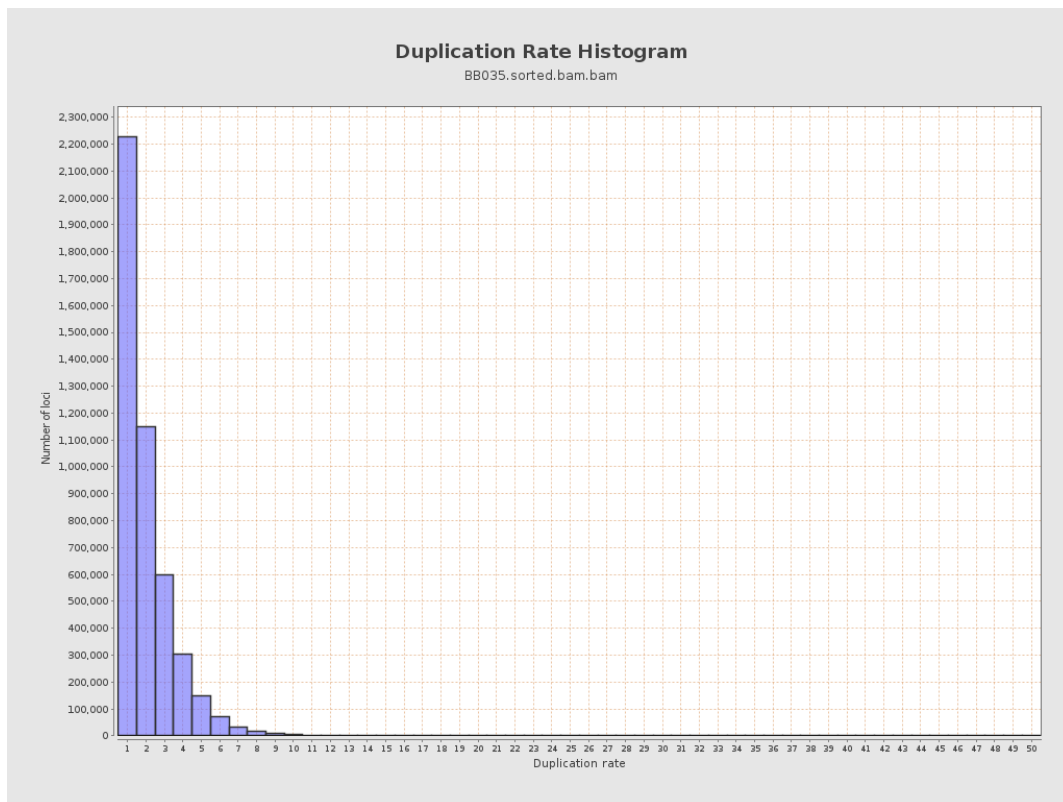
## 5. Results : Coverage Histogram (0-50X)



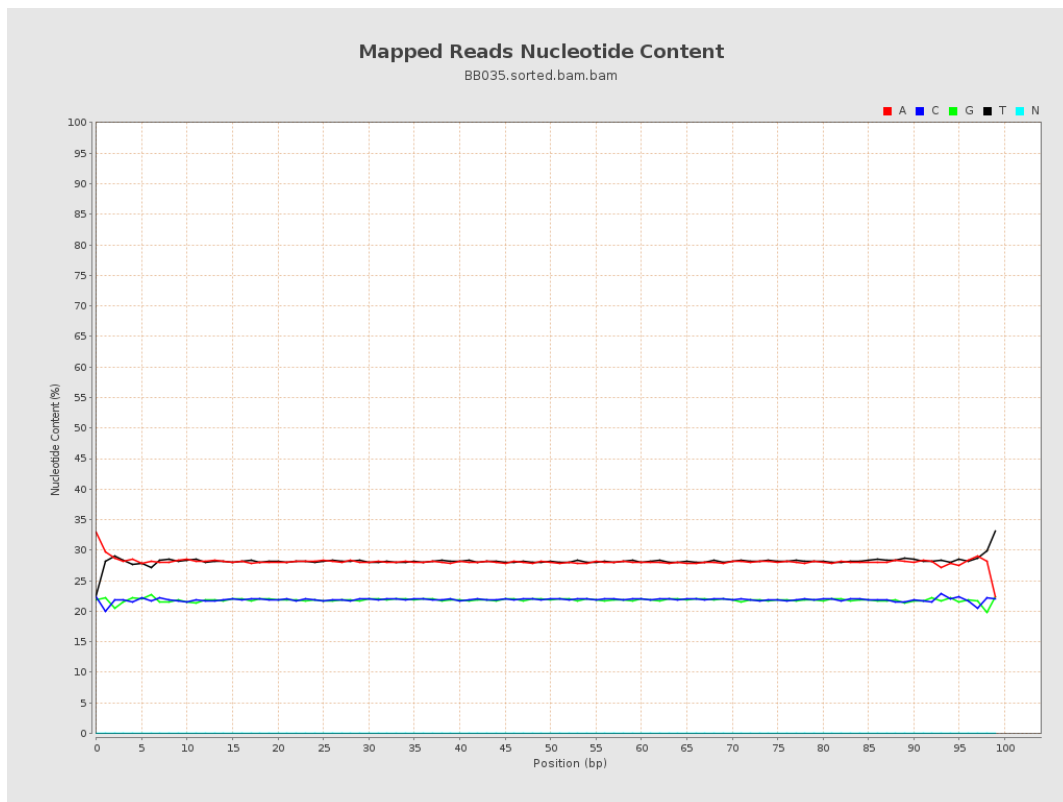
## 6. Results : Genome Fraction Coverage



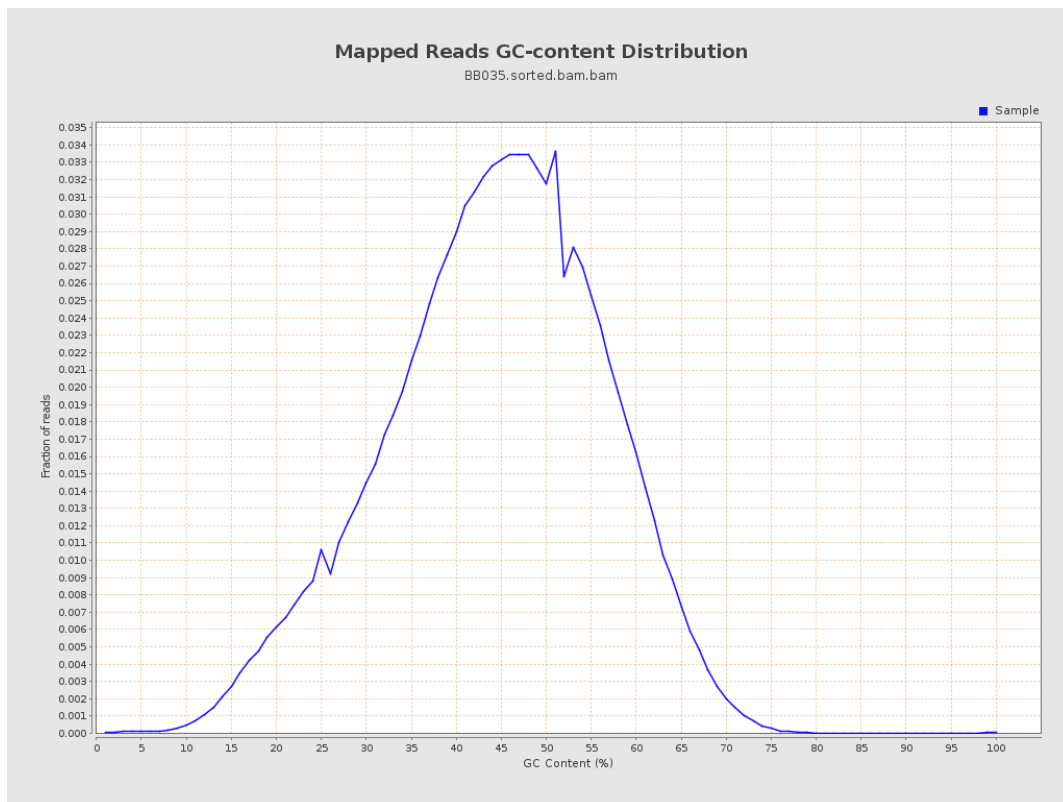
## 7. Results : Duplication Rate Histogram



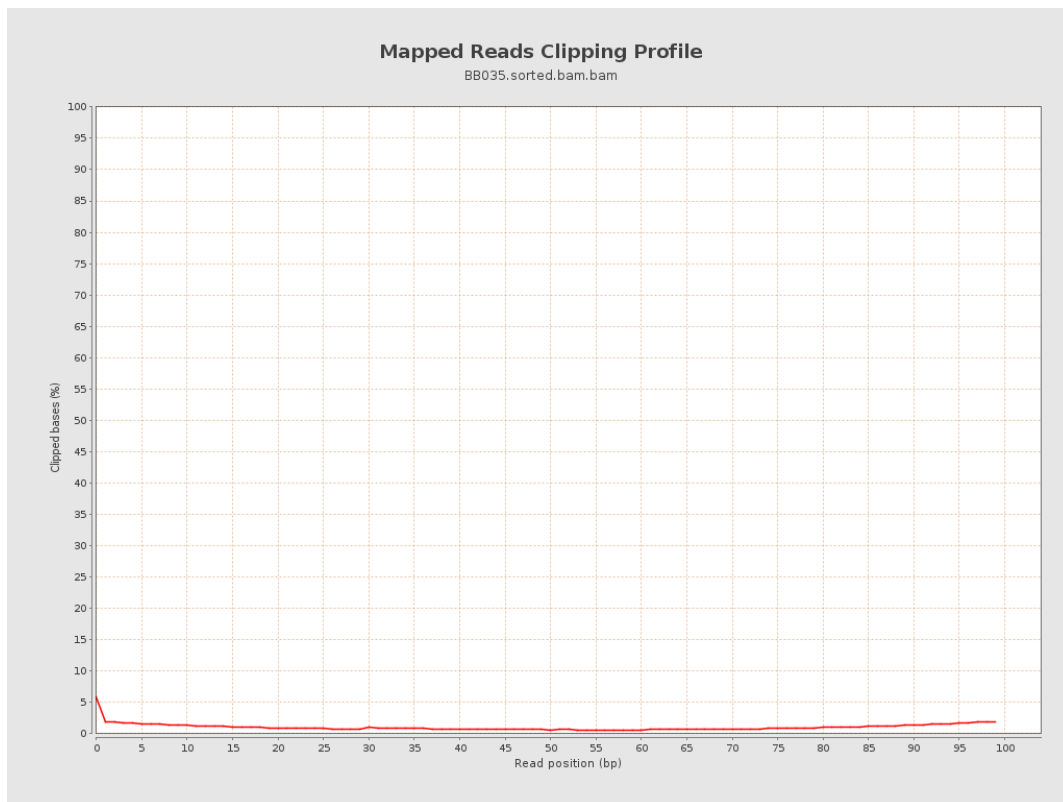
## 8. Results : Mapped Reads Nucleotide Content



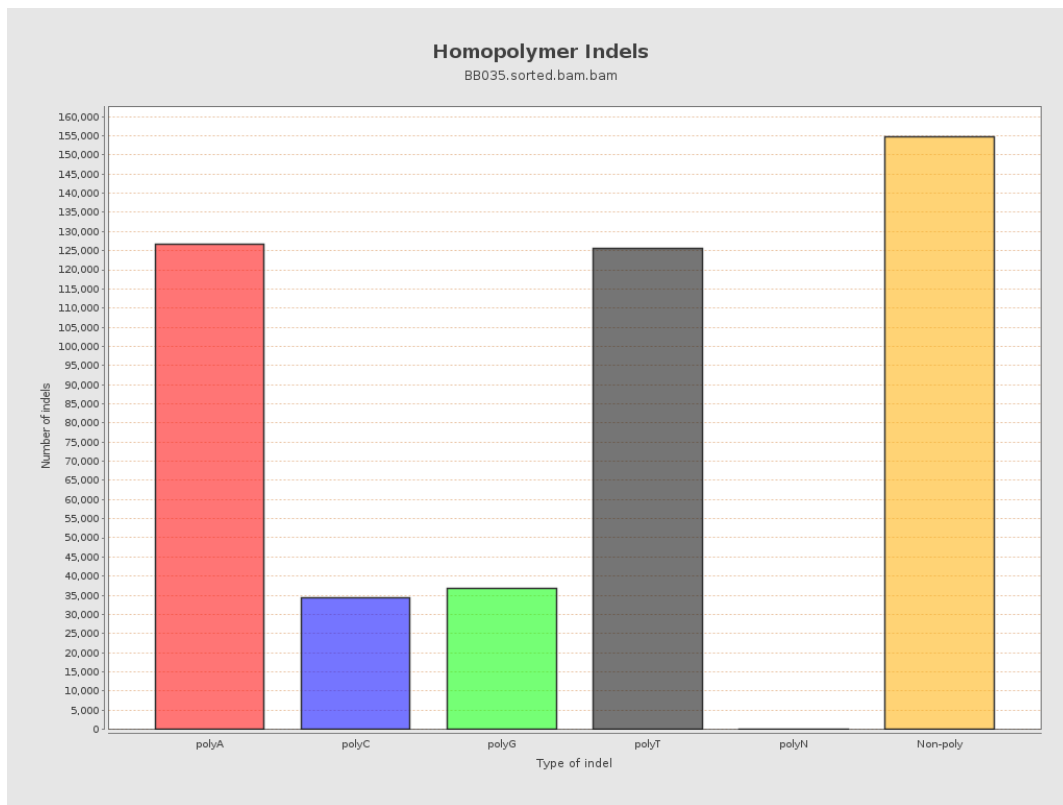
## 9. Results : Mapped Reads GC-content Distribution



## 10. Results : Mapped Reads Clipping Profile

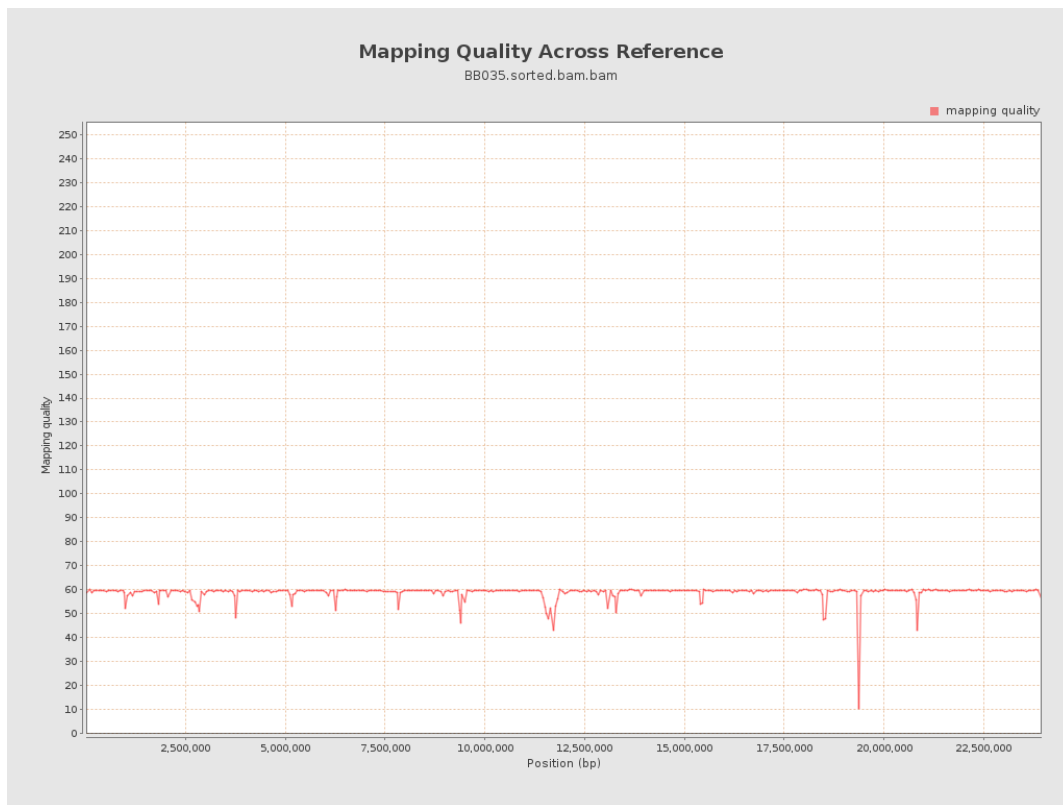


## 11. Results : Homopolymer Indels

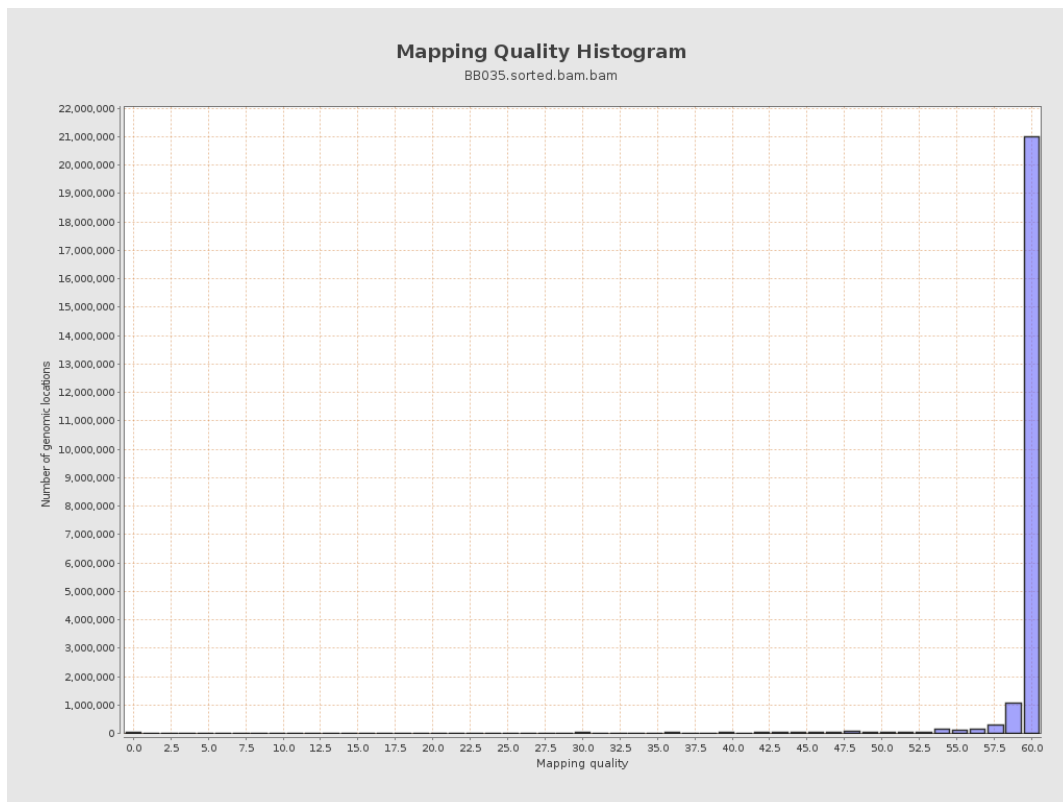




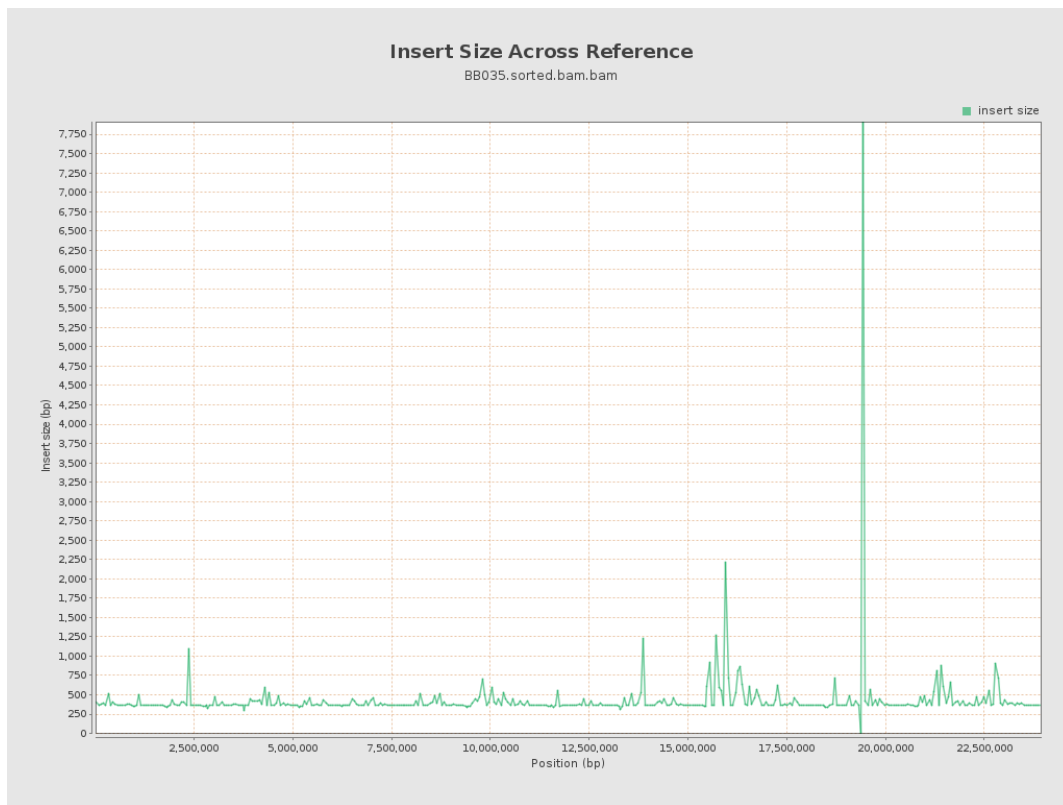
## 12. Results : Mapping Quality Across Reference



## 13. Results : Mapping Quality Histogram



## 14. Results : Insert Size Across Reference



## 15. Results : Insert Size Histogram

