

# Qualimap Analysis Results

*BAM QC analysis*

*Generated by Qualimap v.2.2.1*

*2016/10/23 12:10:51*

# 1. Input data & parameters

## 1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/OM044.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/OM044-BiooBarcode8_TTAGGC_R1.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/OM044-BiooBarcode8_TTAGGC_R2.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 12:10:50 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/OM044.sorted.bam.bam

## 2. Summary

### 2.1. Globals

Reference size	23,958,997
Number of reads	44,628,889
Mapped reads	8,478,062 / 19%
Unmapped reads	36,150,827 / 81%
Mapped paired reads	8,478,062 / 19%
Mapped reads, first in pair	4,249,750 / 9.52%
Mapped reads, second in pair	4,228,312 / 9.47%
Mapped reads, both in pair	8,143,513 / 18.25%
Mapped reads, singletons	334,549 / 0.75%
Read min/max/mean length	30 / 100 / 99.96
Duplicated reads (estimated)	2,199,151 / 4.93%
Duplication rate	21.08%
Clipped reads	1,060,086 / 2.38%

### 2.2. ACGT Content

Number/percentage of A's	240,871,951 / 29.77%
Number/percentage of C's	163,332,272 / 20.18%
Number/percentage of T's	241,954,318 / 29.9%
Number/percentage of G's	163,084,297 / 20.15%
Number/percentage of N's	66,025 / 0.01%
GC Percentage	40.34%

## 2.3. Coverage

Mean	33.812
Standard Deviation	44.4381

## 2.4. Mapping Quality

Mean Mapping Quality	58.18
----------------------	-------

## 2.5. Insert size

Mean	727.96
Standard Deviation	24,494.53
P25/Median/P75	276 / 290 / 300

## 2.6. Mismatches and indels

General error rate	1.44%
Mismatches	11,065,115
Insertions	258,355
Mapped reads with at least one insertion	2.88%
Deletions	296,278
Mapped reads with at least one deletion	3.27%
Homopolymer indels	63.2%

## 2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

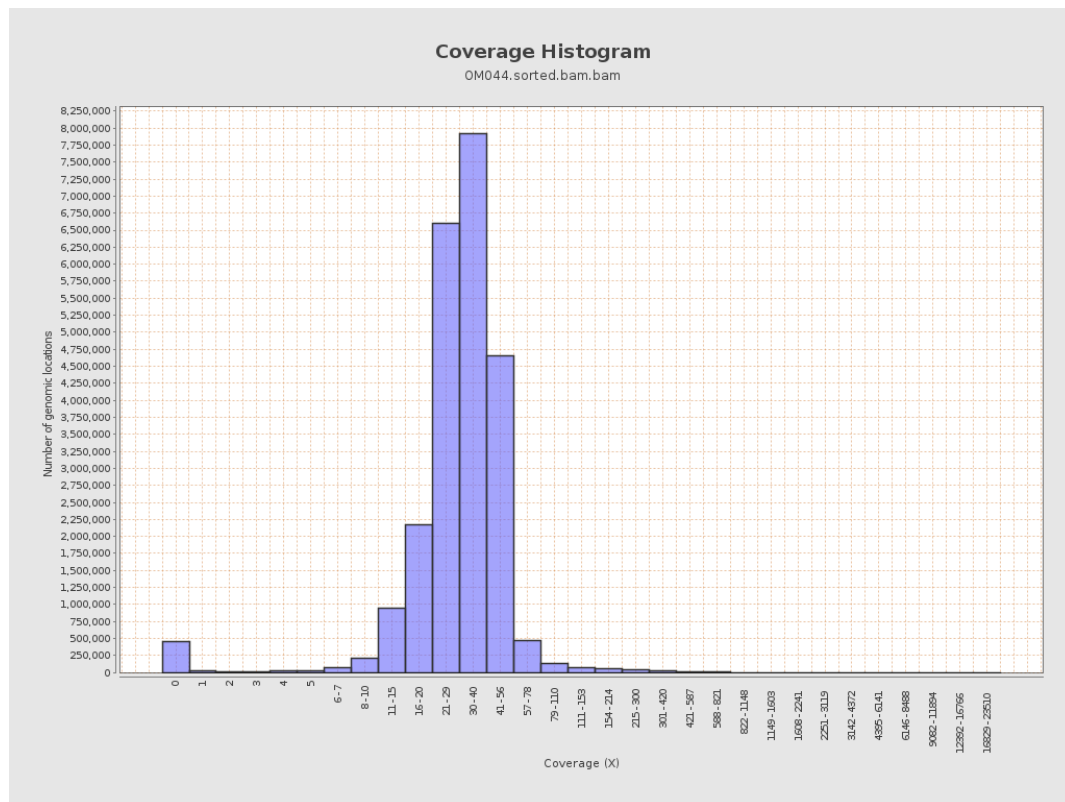
gi 1074120478 emb LT615256.1	977217	28947874	29.6228	14.9066
gi 1074120682 emb LT615257.1	860454	27763348	32.2659	25.8346
gi 1074120865 emb LT615258.1	989719	36748604	37.1303	51.0361
gi 1074121086 emb LT615259.1	935450	35397092	37.8396	55.0018
gi 1074121301 emb LT615260.1	1432239	52407721	36.5915	41.181
gi 1074121615 emb LT615261.1	1080962	36749646	33.9972	34.1273
gi 1074121871 emb LT615262.1	1545099	49656804	32.1383	15.4575
gi 1074122235 emb LT615263.1	1585108	53621268	33.8281	56.2496
gi 1074122590 emb LT615264.1	2122358	69107659	32.5617	17.6584
gi 1074123050 emb LT615265.1	1754192	58225718	33.1923	101.5739
gi 1074123421 emb LT615	2150147	76561401	35.6075	54.9191

266.1				
gi 107412389 8 emb LT615 267.1	3031036	101075974	33.347	32.8384
gi 107412458 8 emb LT615 268.1	2359348	76521594	32.4334	34.5459
gi 107412506 5 emb LT615 269.1	3135668	107315854	34.2242	14.2298

### 3. Results : Coverage across reference

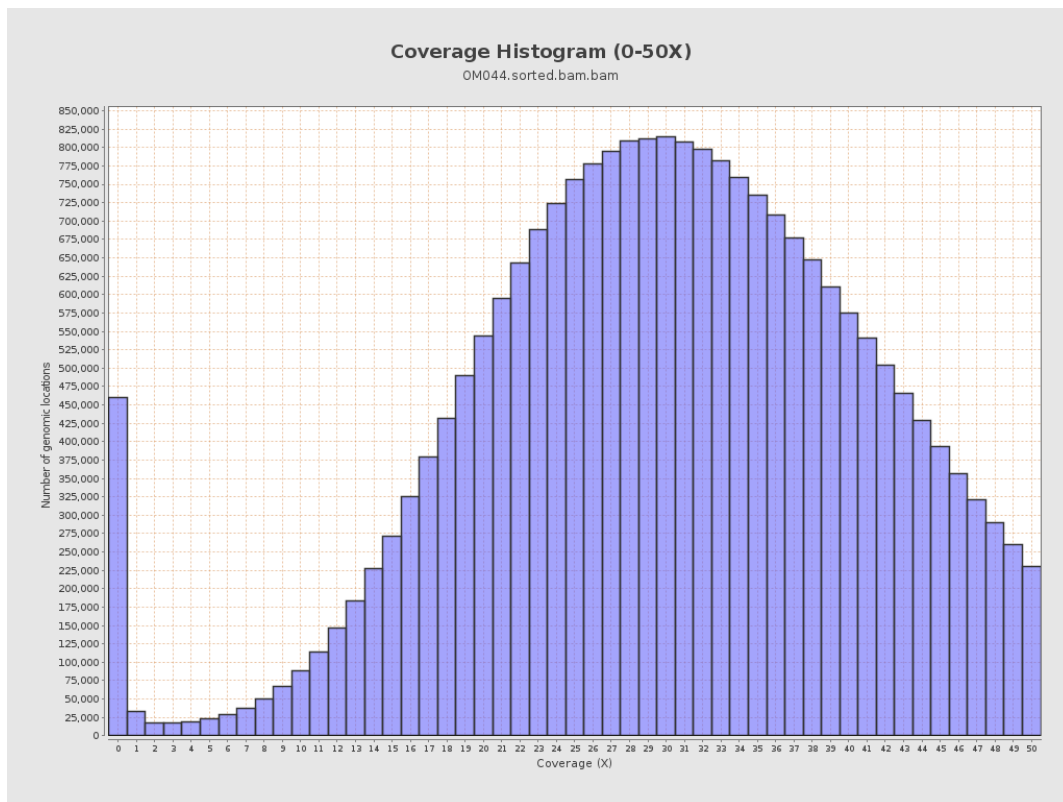


## 4. Results : Coverage Histogram

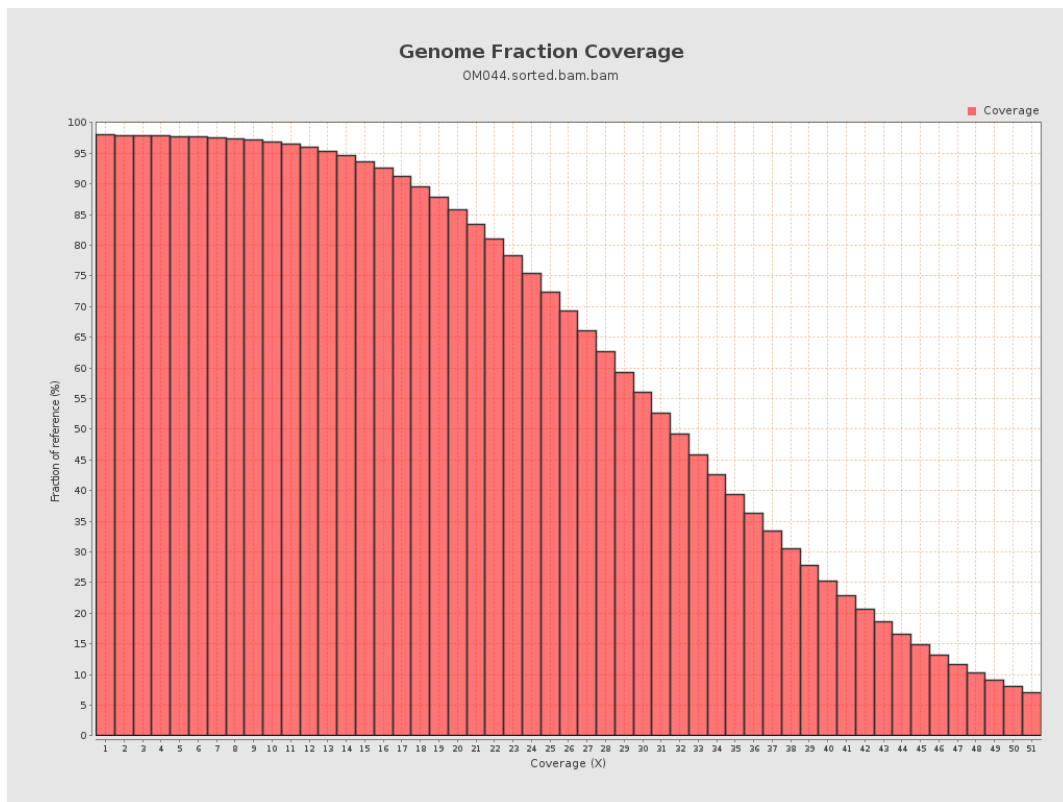




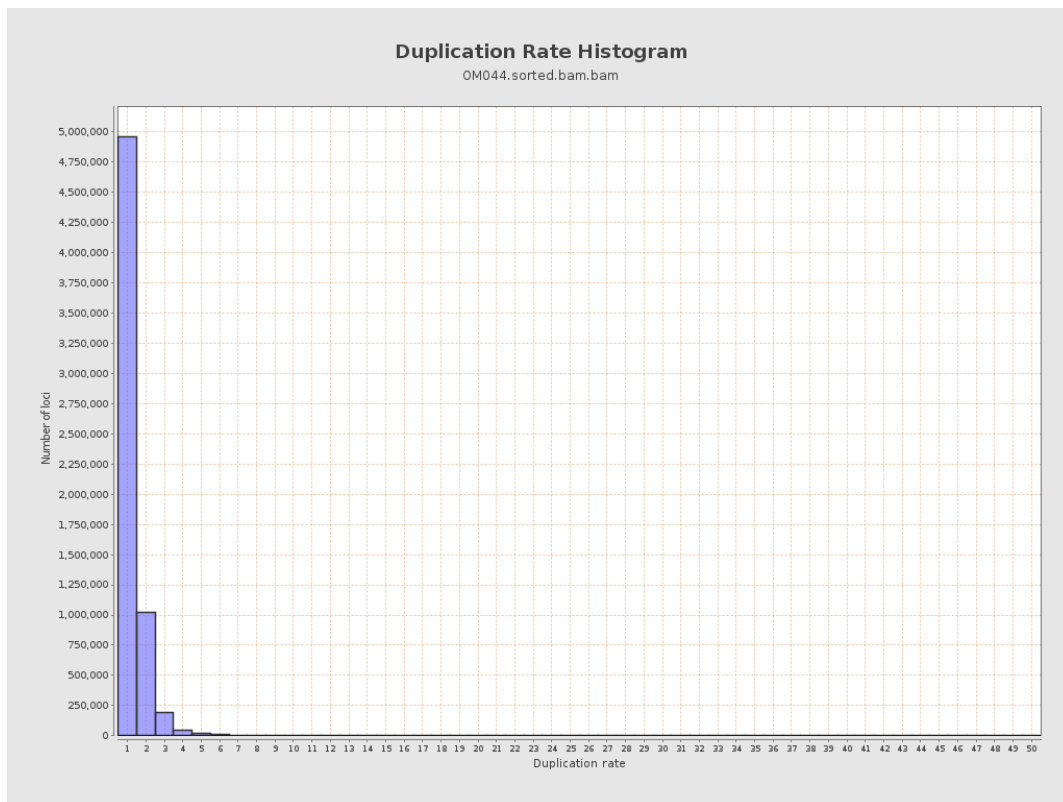
## 5. Results : Coverage Histogram (0-50X)



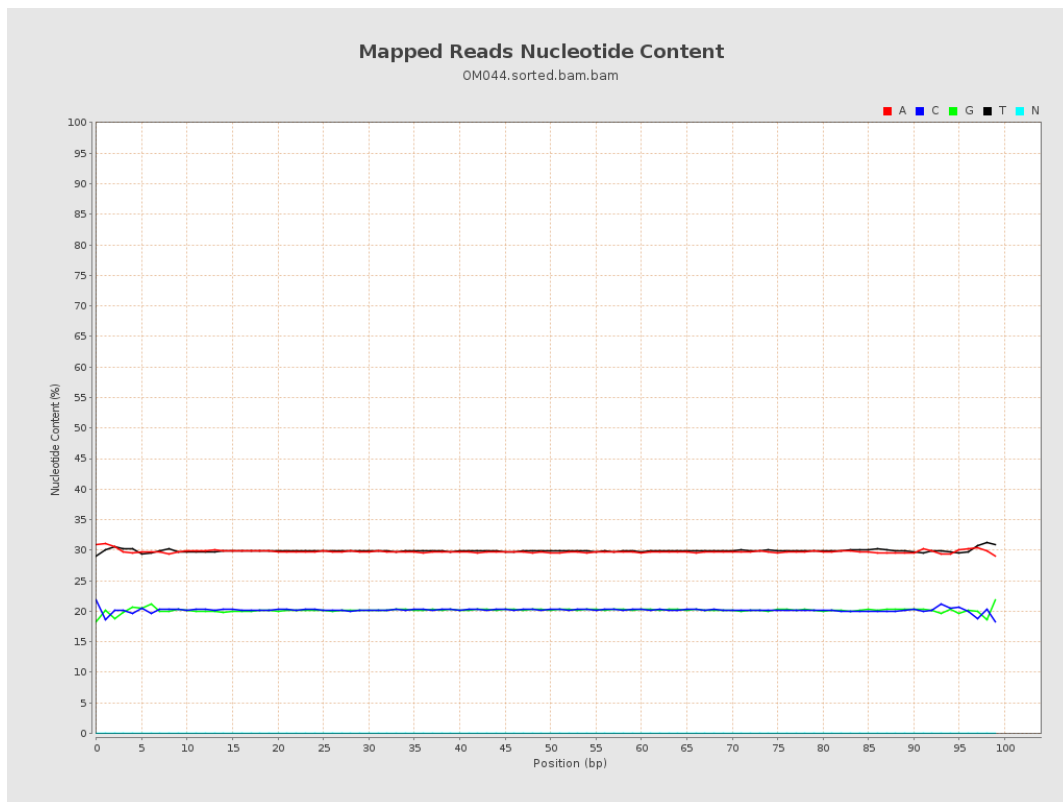
## 6. Results : Genome Fraction Coverage



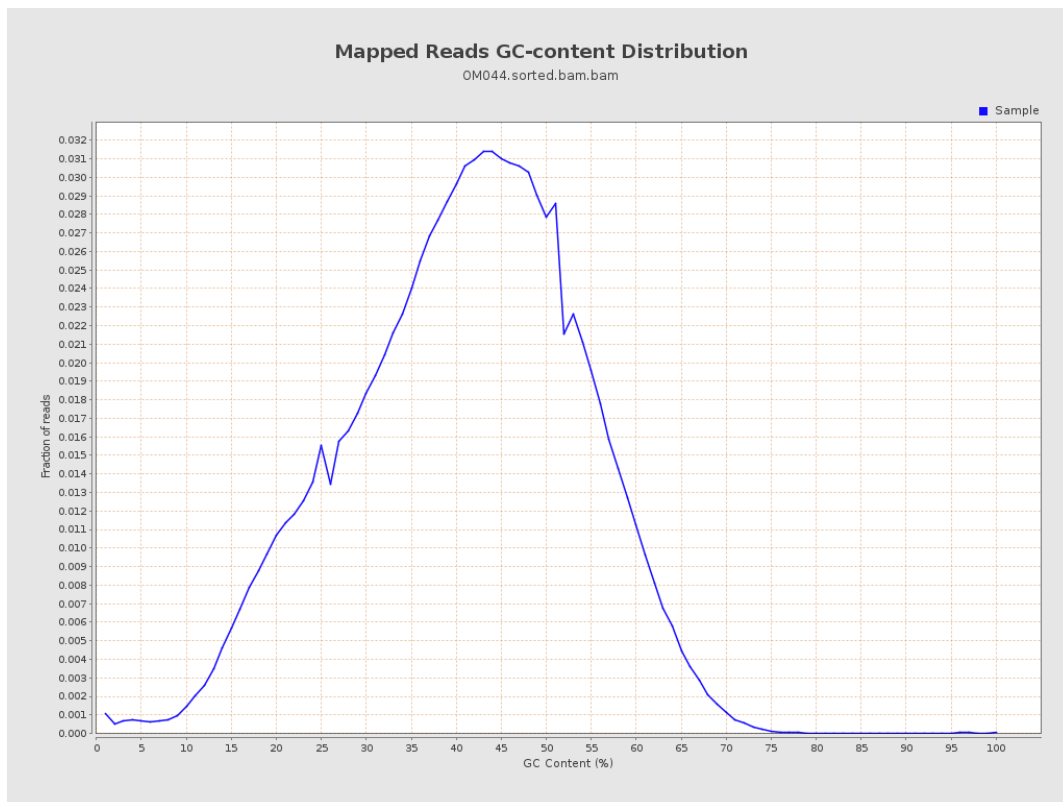
## 7. Results : Duplication Rate Histogram



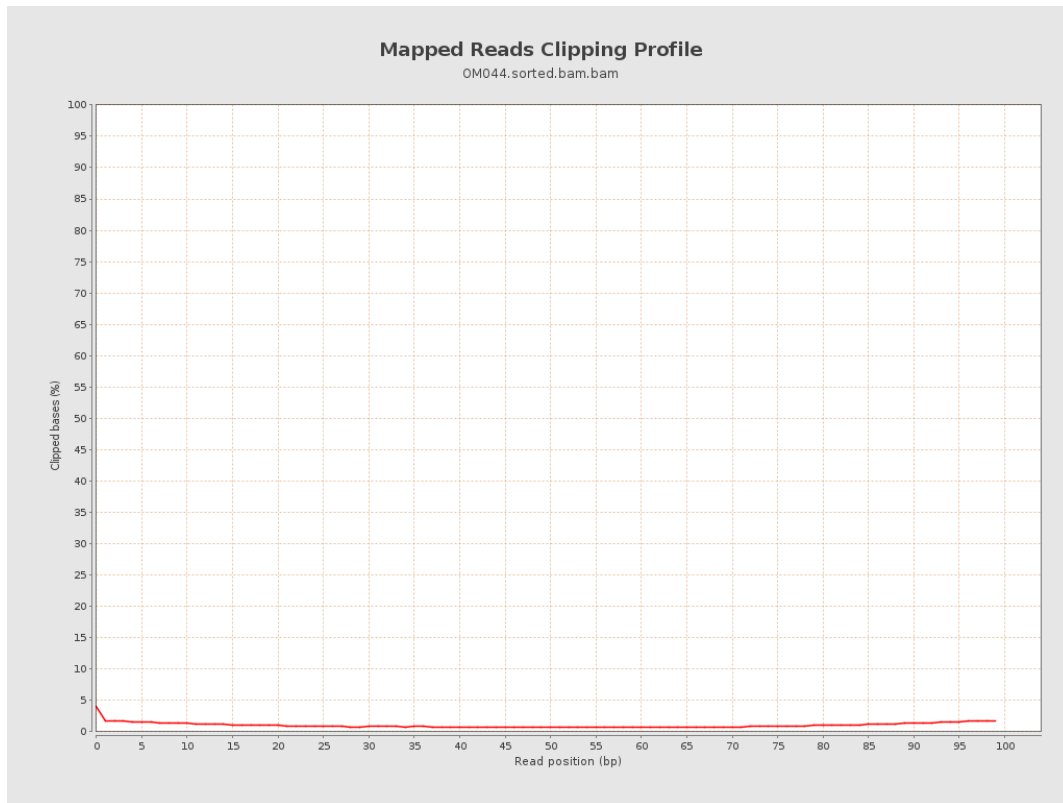
## 8. Results : Mapped Reads Nucleotide Content



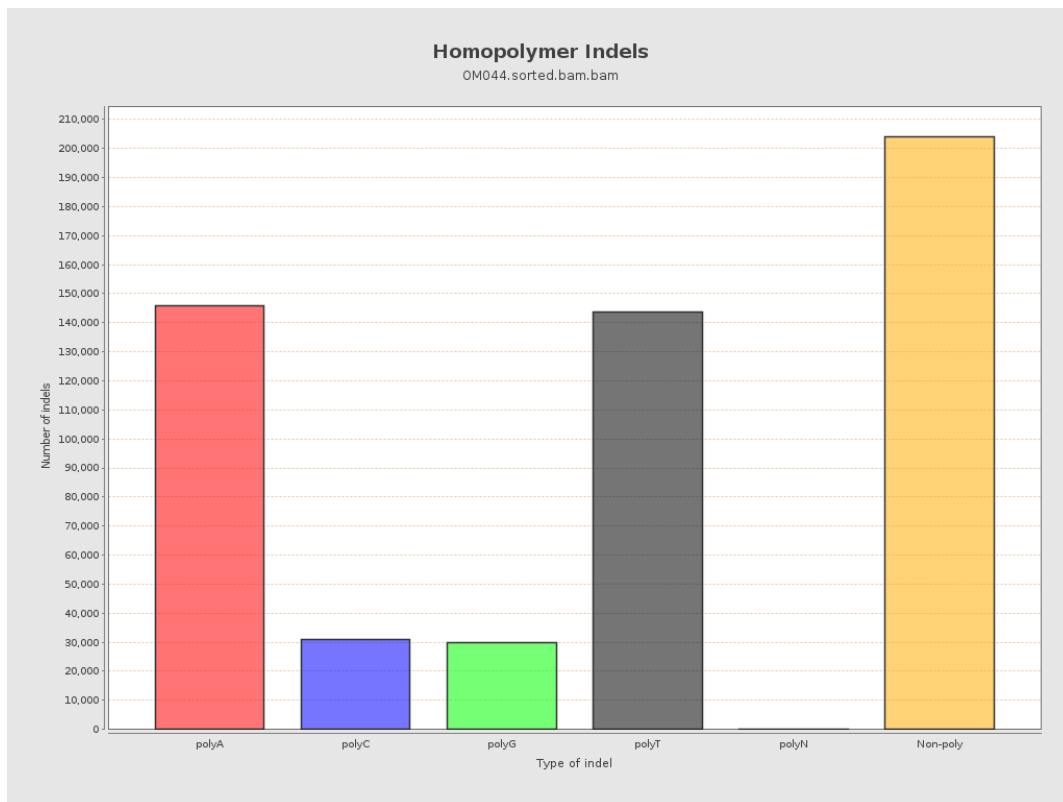
## 9. Results : Mapped Reads GC-content Distribution



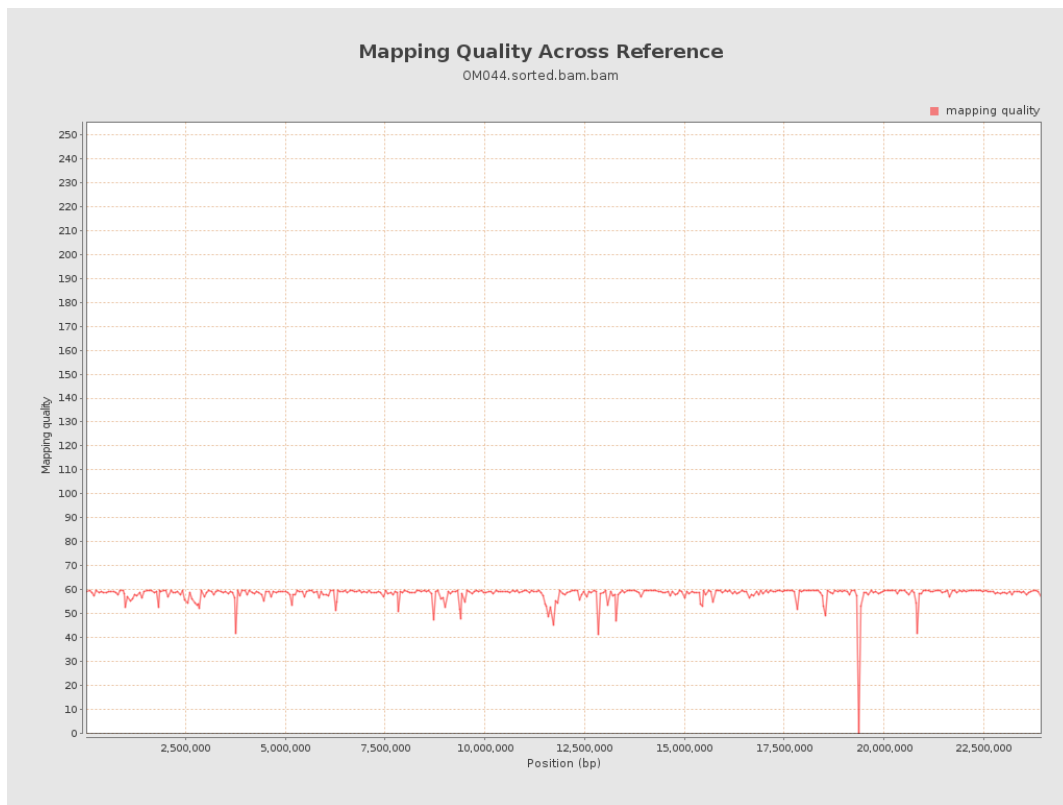
## 10. Results : Mapped Reads Clipping Profile



## 11. Results : Homopolymer Indels

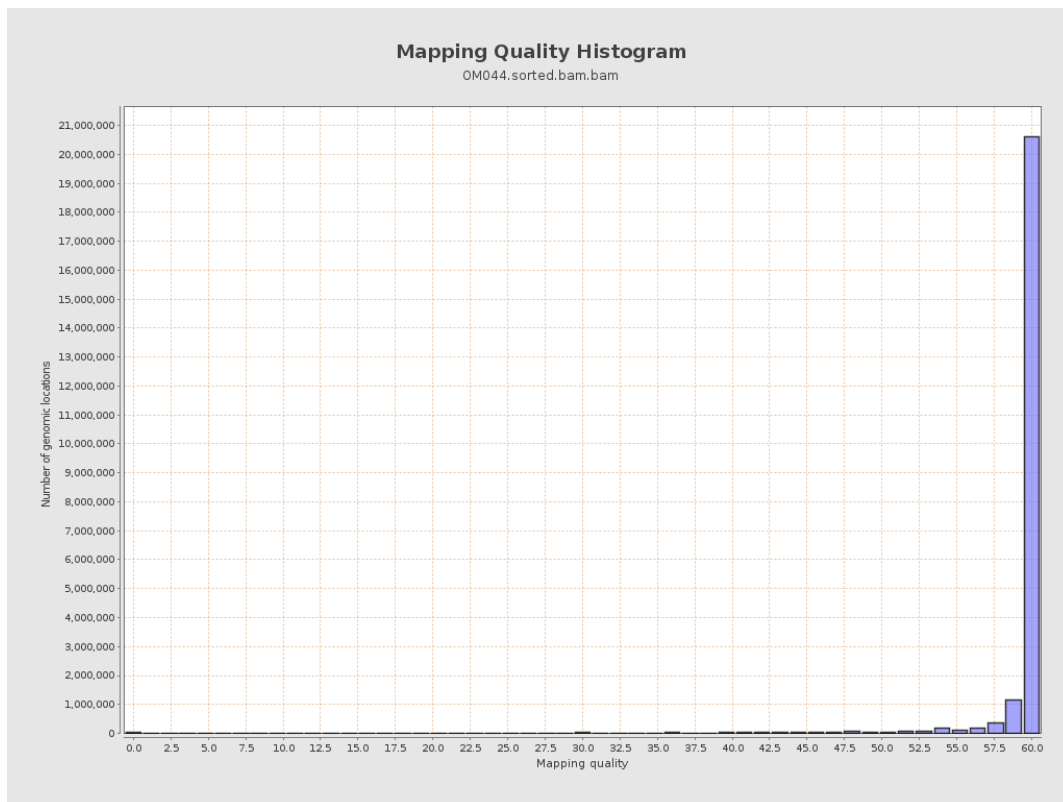


## 12. Results : Mapping Quality Across Reference

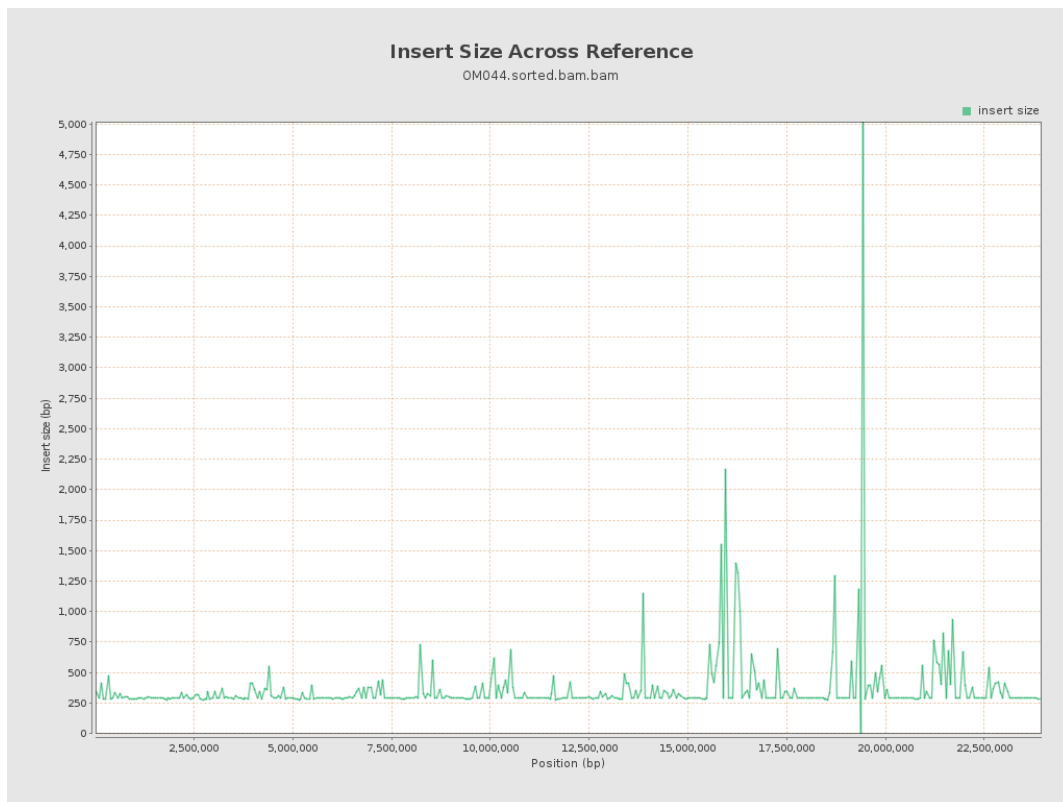




## 13. Results : Mapping Quality Histogram



## 14. Results : Insert Size Across Reference



## 15. Results : Insert Size Histogram

