

# Qualimap Analysis Results

*BAM QC analysis*

*Generated by Qualimap v.2.2.1*

*2016/10/23 11:47:51*

# 1. Input data & parameters

## 1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/KP053.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/KP053-BiooBarcode_25_ACTGAT_R1.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/KP053-BiooBarcode_25_ACTGAT_R2.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 11:47:50 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/KP053.sorted.bam.b

am

## 2. Summary

### 2.1. Globals

Reference size	23,958,997
Number of reads	16,131,587
Mapped reads	10,013,989 / 62.08%
Unmapped reads	6,117,598 / 37.92%
Mapped paired reads	10,013,989 / 62.08%
Mapped reads, first in pair	5,024,841 / 31.15%
Mapped reads, second in pair	4,989,148 / 30.93%
Mapped reads, both in pair	9,824,843 / 60.9%
Mapped reads, singletons	189,146 / 1.17%
Read min/max/mean length	30 / 100 / 99.86
Duplicated reads (estimated)	3,799,676 / 23.55%
Duplication rate	36.95%
Clipped reads	967,863 / 6%

### 2.2. ACGT Content

Number/percentage of A's	276,660,914 / 28.54%
Number/percentage of C's	207,623,133 / 21.42%
Number/percentage of T's	277,686,548 / 28.65%
Number/percentage of G's	207,255,597 / 21.38%
Number/percentage of N's	80,009 / 0.01%
GC Percentage	42.81%

## 2.3. Coverage

Mean	40.4959
Standard Deviation	26.9235

## 2.4. Mapping Quality

Mean Mapping Quality	58.8
----------------------	------

## 2.5. Insert size

Mean	1,100.09
Standard Deviation	31,023.04
P25/Median/P75	276 / 350 / 424

## 2.6. Mismatches and indels

General error rate	1.01%
Mismatches	9,096,541
Insertions	236,547
Mapped reads with at least one insertion	2.25%
Deletions	311,809
Mapped reads with at least one deletion	2.97%
Homopolymer indels	65.44%

## 2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

gi 1074120478 emb LT615256.1	977217	38171475	39.0614	13.5789
gi 1074120682 emb LT615257.1	860454	32449402	37.712	18.9853
gi 1074120865 emb LT615258.1	989719	43969060	44.4258	38.7693
gi 1074121086 emb LT615259.1	935450	40291952	43.0723	39.2335
gi 1074121301 emb LT615260.1	1432239	60475277	42.2243	28.1685
gi 1074121615 emb LT615261.1	1080962	45941455	42.5005	23.1374
gi 1074121871 emb LT615262.1	1545099	61575698	39.8523	11.8454
gi 1074122235 emb LT615263.1	1585108	65036423	41.0296	21.3826
gi 1074122590 emb LT615264.1	2122358	83778746	39.4744	19.159
gi 1074123050 emb LT615265.1	1754192	68426680	39.0075	26.0227
gi 1074123421 emb LT615	2150147	92622063	43.0771	55.681

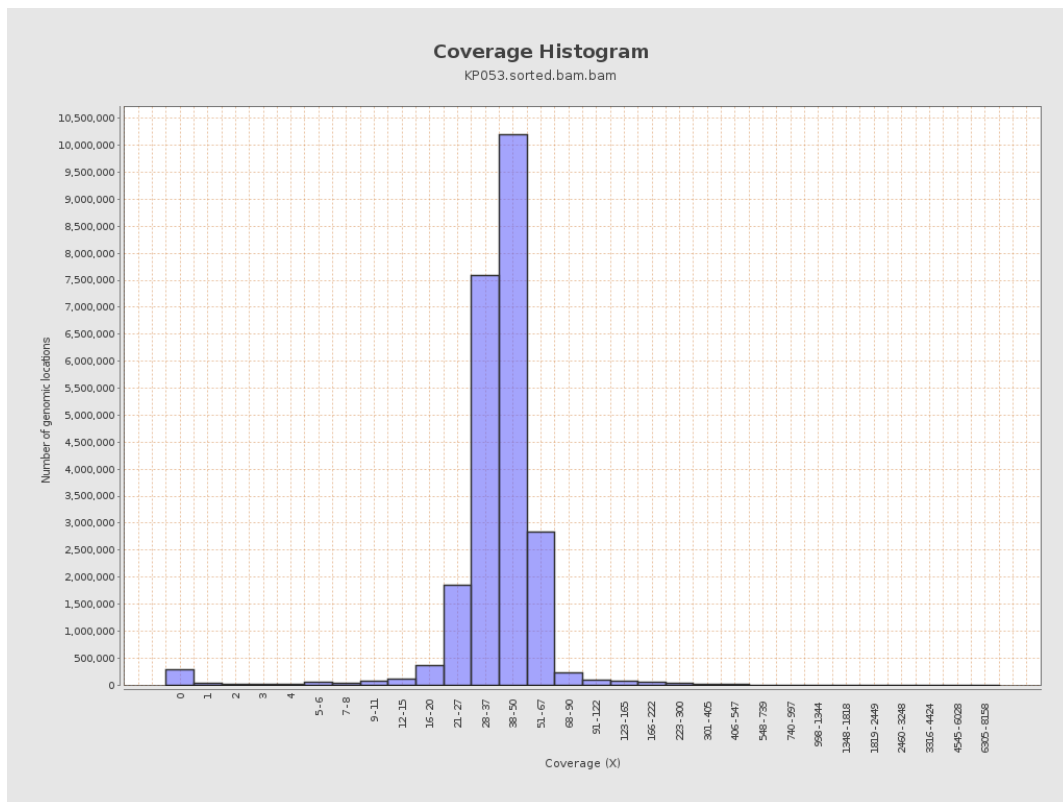
266.1				
gi 107412389 8 emb LT615 267.1	3031036	121197900	39.9856	13.9579
gi 107412458 8 emb LT615 268.1	2359348	90077106	38.1788	26.6536
gi 107412506 5 emb LT615 269.1	3135668	126227680	40.2554	10.0919

### 3. Results : Coverage across reference

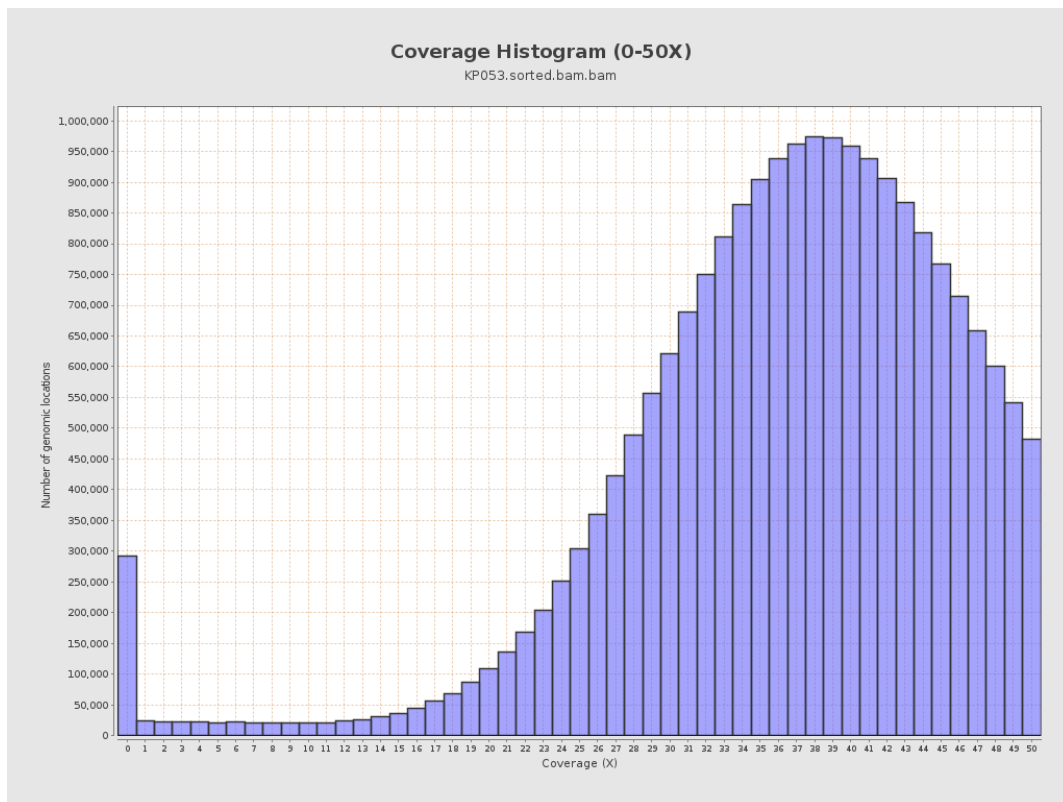




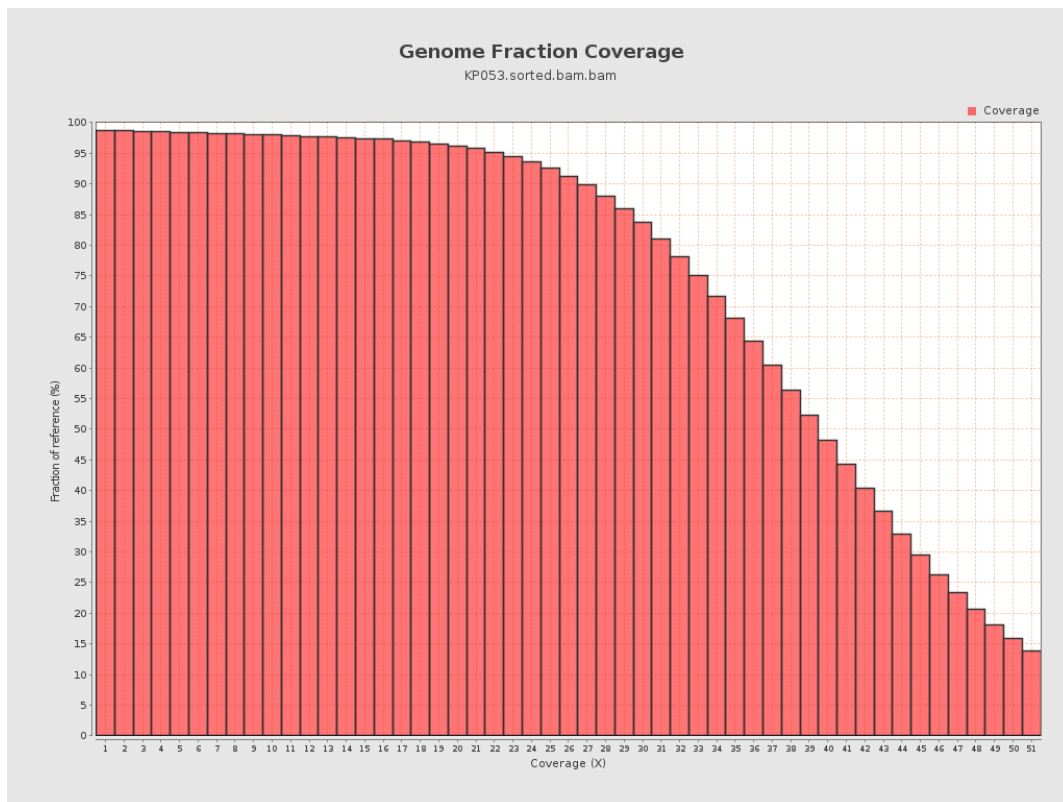
## 4. Results : Coverage Histogram



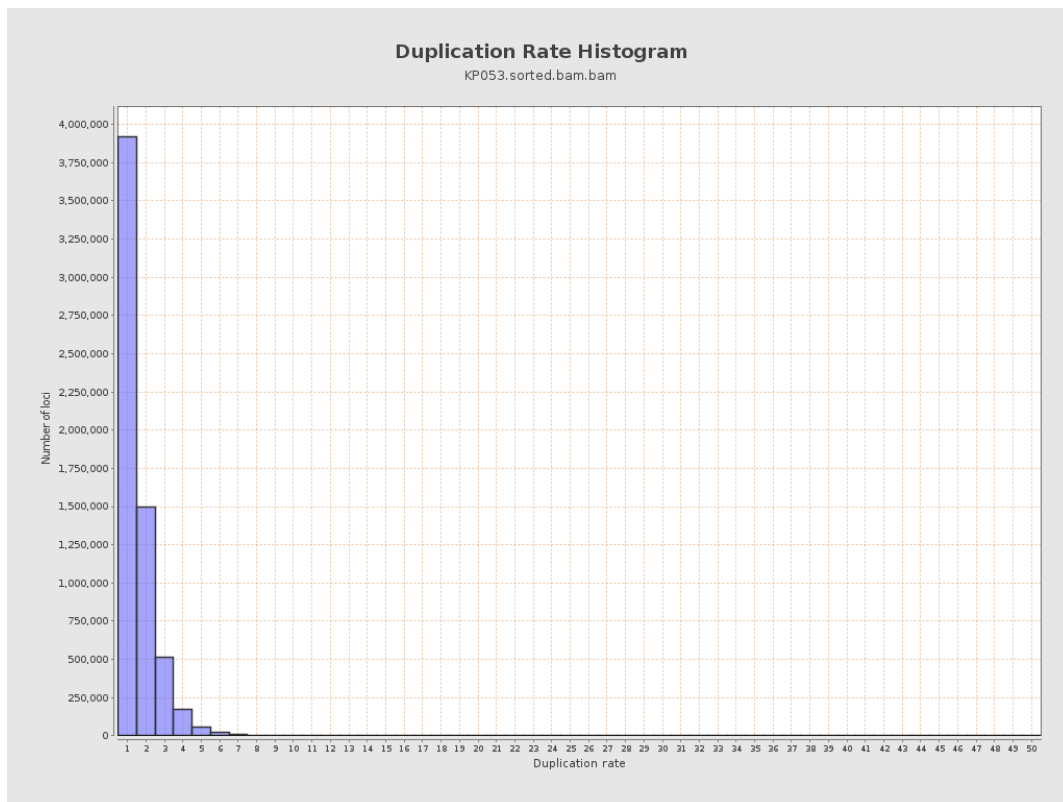
## 5. Results : Coverage Histogram (0-50X)



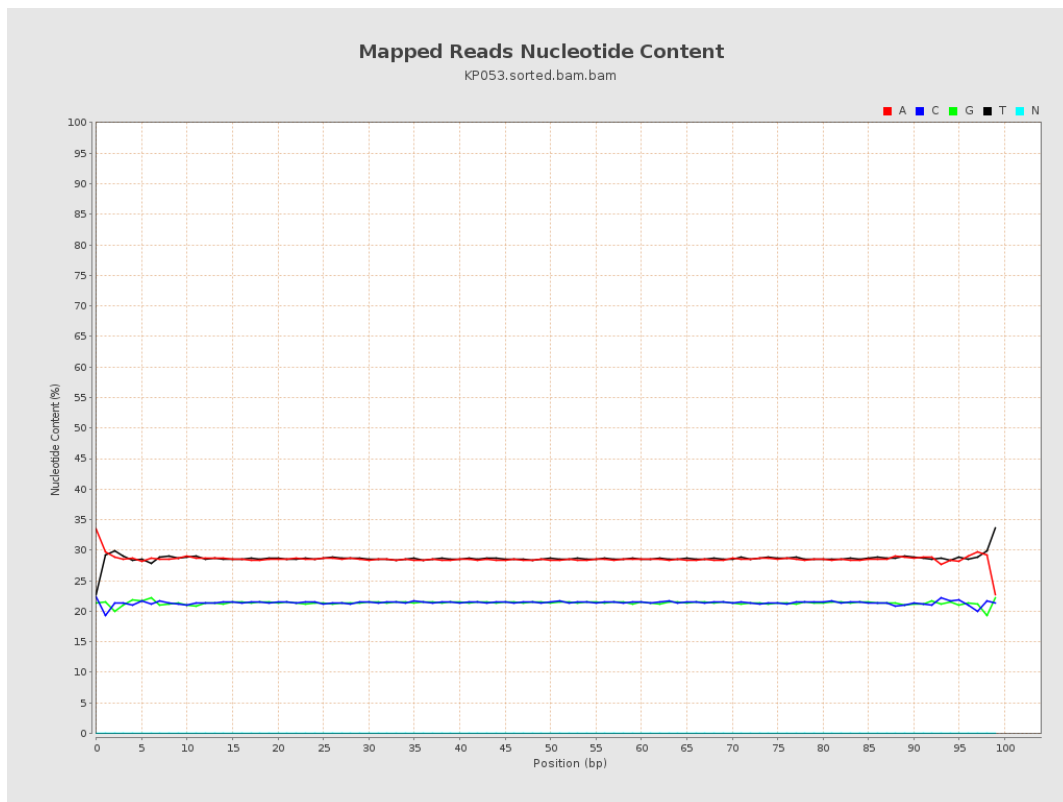
## 6. Results : Genome Fraction Coverage



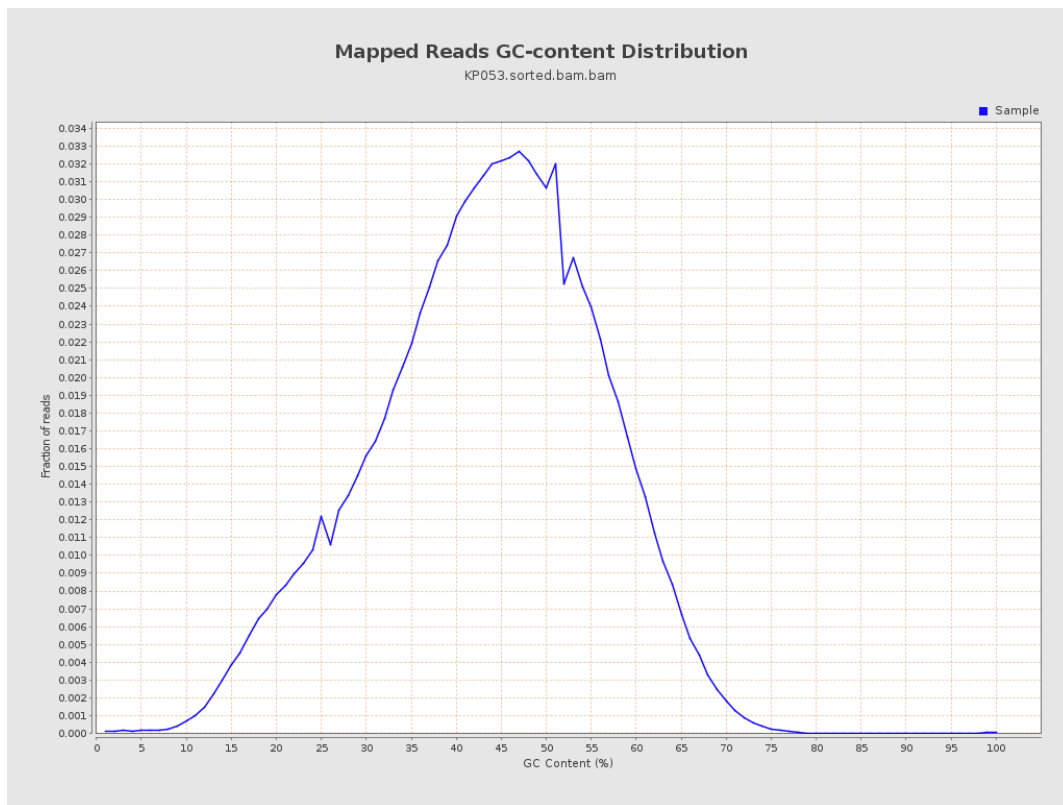
## 7. Results : Duplication Rate Histogram



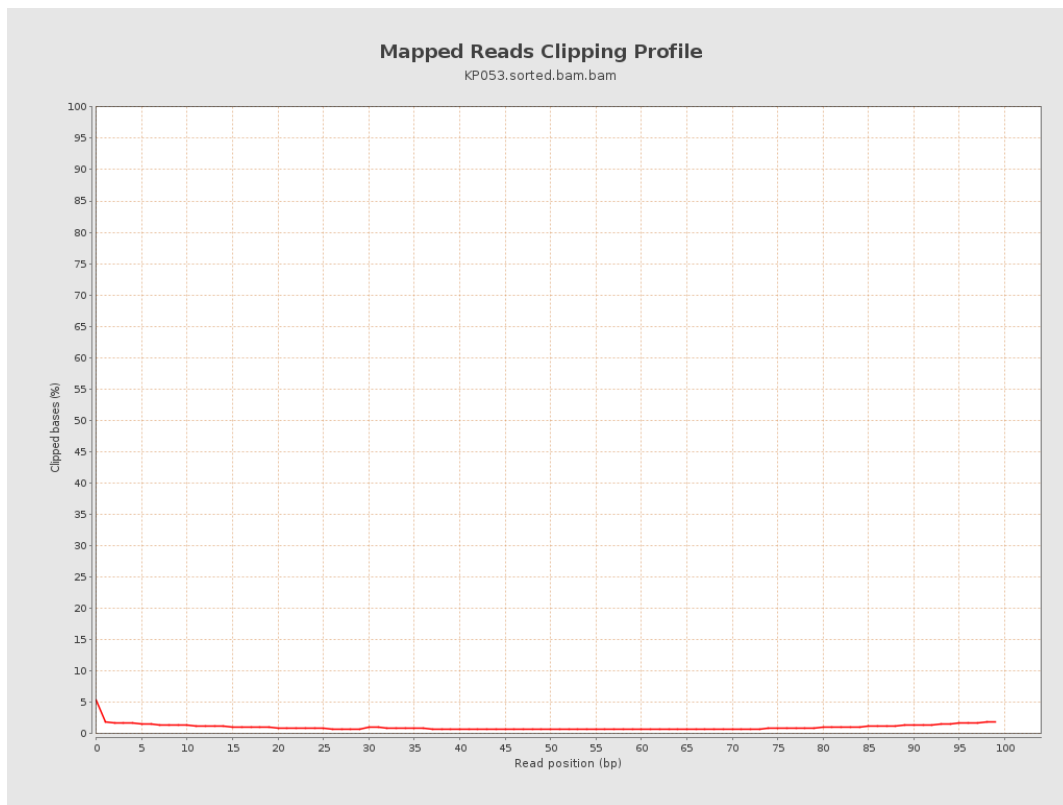
## 8. Results : Mapped Reads Nucleotide Content



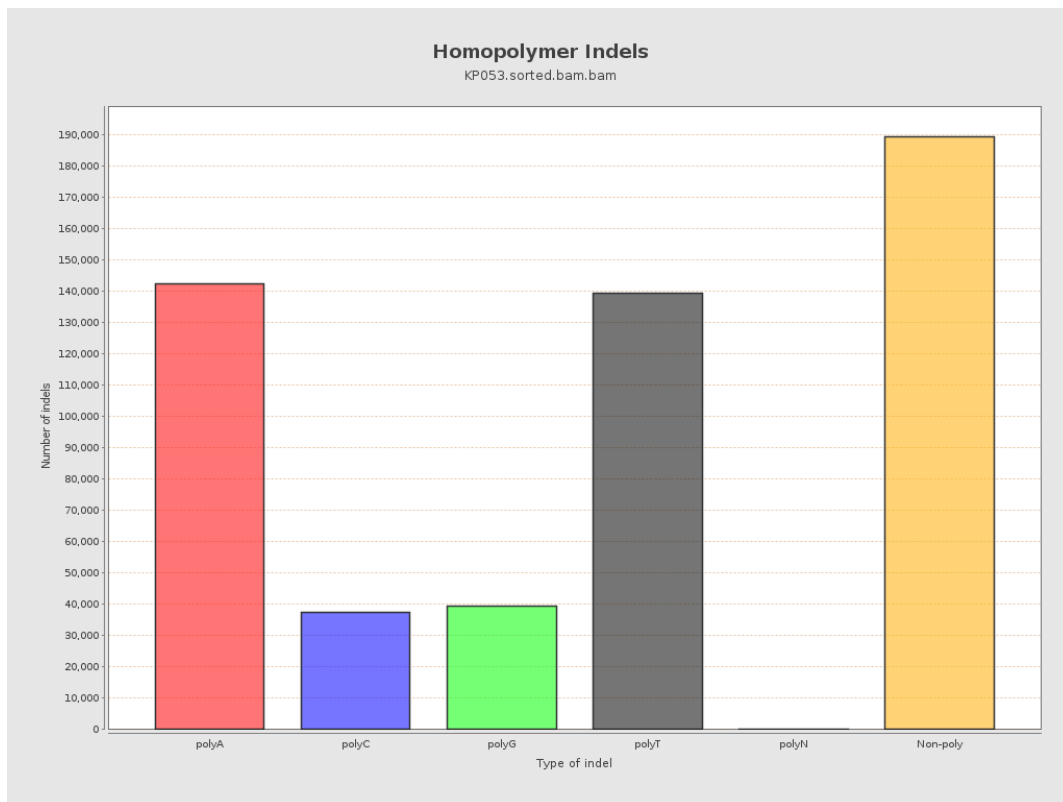
## 9. Results : Mapped Reads GC-content Distribution



## 10. Results : Mapped Reads Clipping Profile

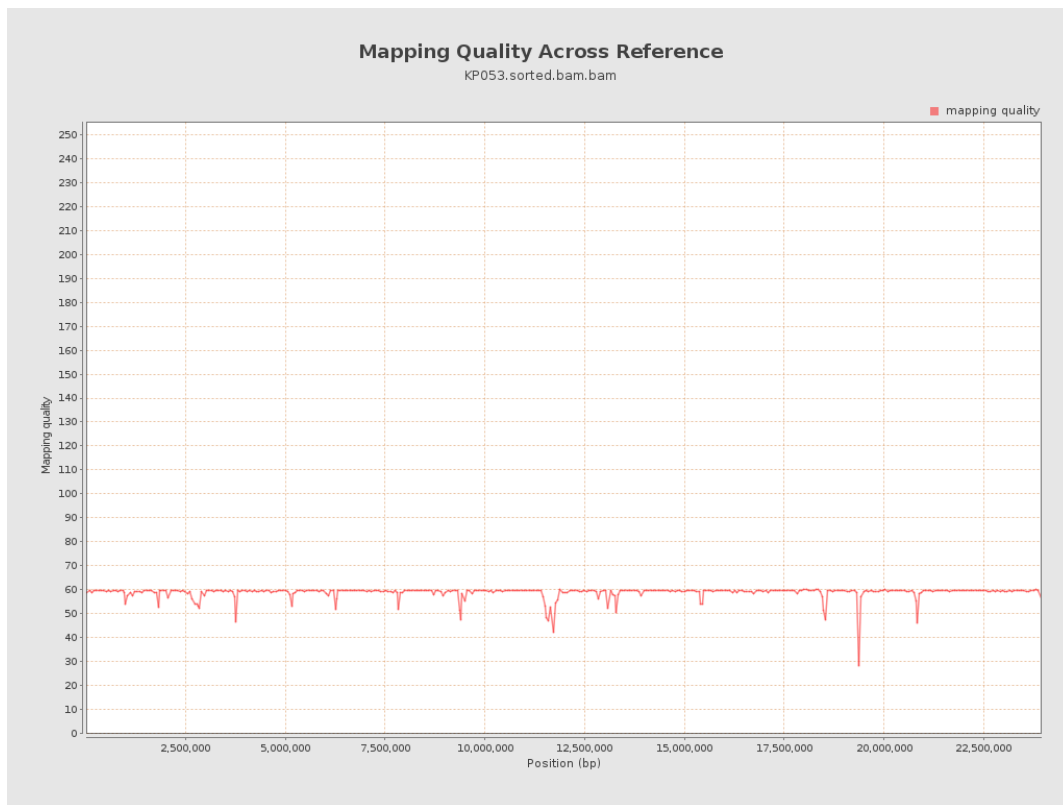


## 11. Results : Homopolymer Indels

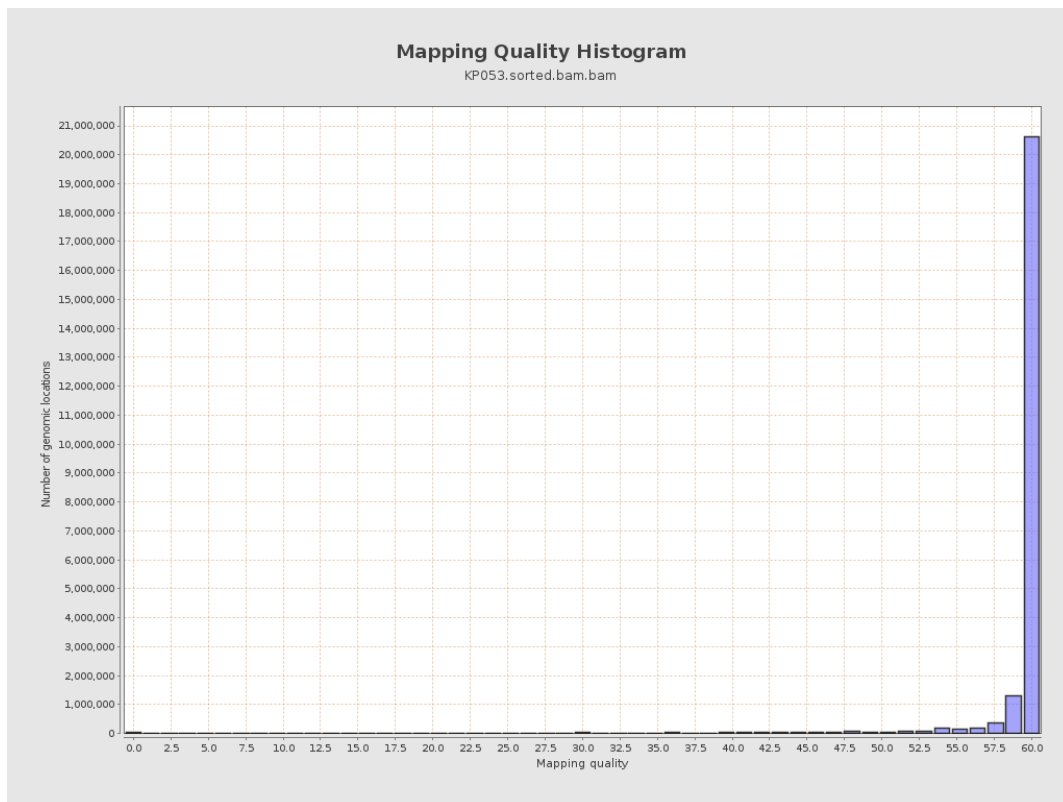




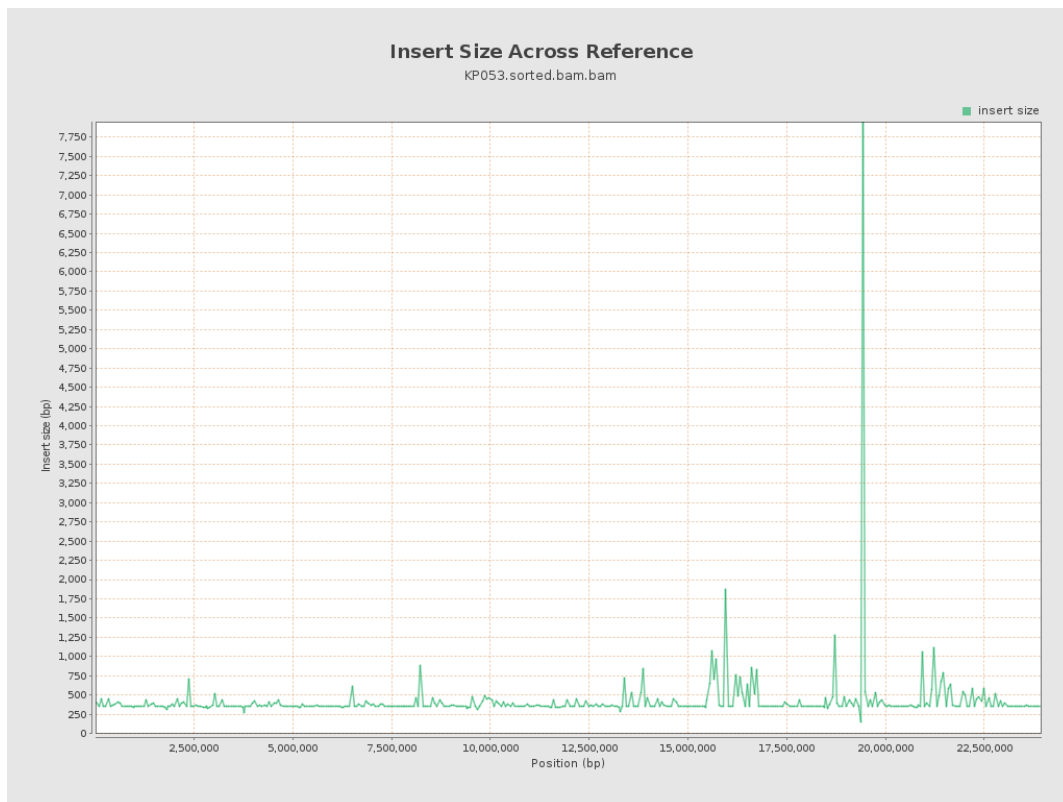
## 12. Results : Mapping Quality Across Reference



## 13. Results : Mapping Quality Histogram



## 14. Results : Insert Size Across Reference



## 15. Results : Insert Size Histogram

