

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 13:53:05

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/OM270.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/OM270-BiooBarcode16_CCGTCC_R1.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/OM270-BiooBarcode16_CCGTCC_R2.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 13:53:04 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/OM270.sorted.bam.

	bam
--	-----

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	36,844,160
Mapped reads	5,317,856 / 14.43%
Unmapped reads	31,526,304 / 85.57%
Mapped paired reads	5,317,856 / 14.43%
Mapped reads, first in pair	2,669,263 / 7.24%
Mapped reads, second in pair	2,648,593 / 7.19%
Mapped reads, both in pair	5,037,444 / 13.67%
Mapped reads, singletons	280,412 / 0.76%
Read min/max/mean length	30 / 100 / 99.97
Duplicated reads (estimated)	1,179,554 / 3.2%
Duplication rate	16.94%
Clipped reads	747,897 / 2.03%

2.2. ACGT Content

Number/percentage of A's	150,505,414 / 29.92%
Number/percentage of C's	100,628,472 / 20%
Number/percentage of T's	151,481,489 / 30.11%
Number/percentage of G's	100,408,181 / 19.96%
Number/percentage of N's	41,768 / 0.01%
GC Percentage	39.97%

2.3. Coverage

Mean	21.0182
Standard Deviation	35.3145

2.4. Mapping Quality

Mean Mapping Quality	57.84
----------------------	-------

2.5. Insert size

Mean	834.82
Standard Deviation	25,659.84
P25/Median/P75	315 / 333 / 343

2.6. Mismatches and indels

General error rate	1.52%
Mismatches	7,257,251
Insertions	169,377
Mapped reads with at least one insertion	3%
Deletions	192,139
Mapped reads with at least one deletion	3.35%
Homopolymer indels	62.13%

2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

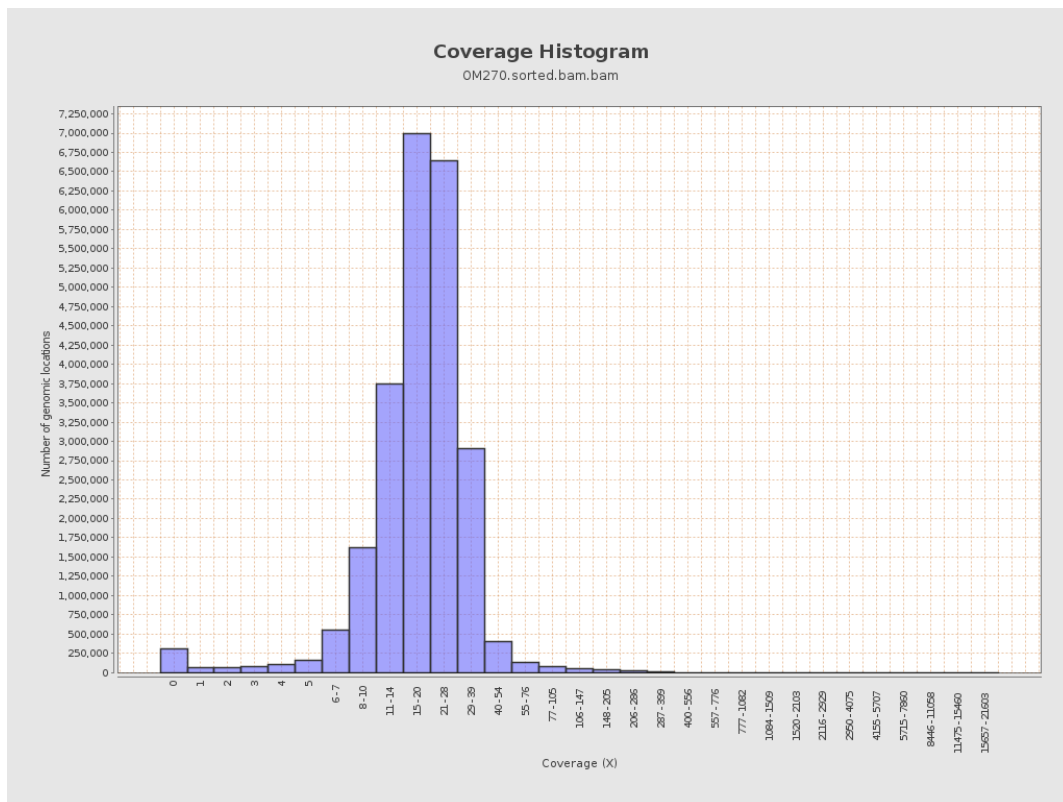
gi 1074120478 emb LT615256.1	977217	18478172	18.909	12.7178
gi 1074120682 emb LT615257.1	860454	17390178	20.2105	18.2456
gi 1074120865 emb LT615258.1	989719	22053853	22.2829	29.119
gi 1074121086 emb LT615259.1	935450	20427137	21.8367	29.7996
gi 1074121301 emb LT615260.1	1432239	31933900	22.2965	25.8375
gi 1074121615 emb LT615261.1	1080962	22794880	21.0876	21.5221
gi 1074121871 emb LT615262.1	1545099	30210762	19.5526	11.08
gi 1074122235 emb LT615263.1	1585108	33427567	21.0885	46.7497
gi 1074122590 emb LT615264.1	2122358	43134486	20.3239	12.1967
gi 1074123050 emb LT615265.1	1754192	38663849	22.0408	93.1938
gi 1074123421 emb LT615	2150147	47121639	21.9155	40.6897

266.1				
gi 107412389 8 emb LT615 267.1	3031036	62174795	20.5127	25.9729
gi 107412458 8 emb LT615 268.1	2359348	48290402	20.4677	21.5385
gi 107412506 5 emb LT615 269.1	3135668	67473851	21.5182	11.0762

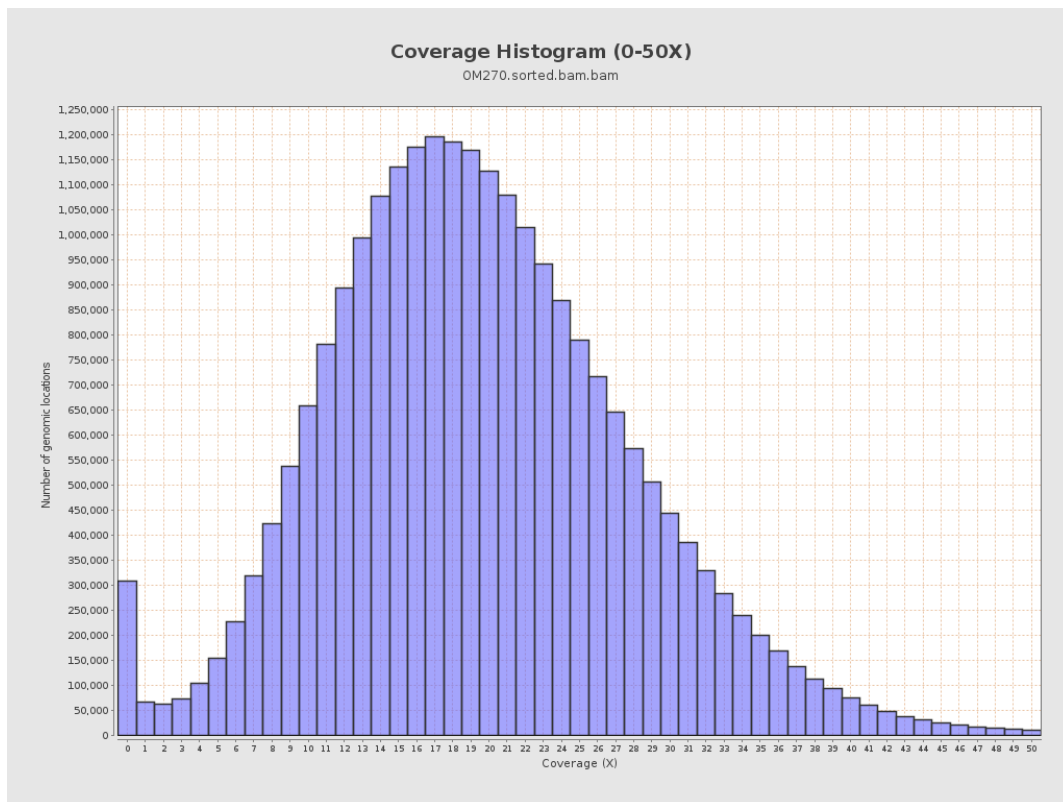
3. Results : Coverage across reference



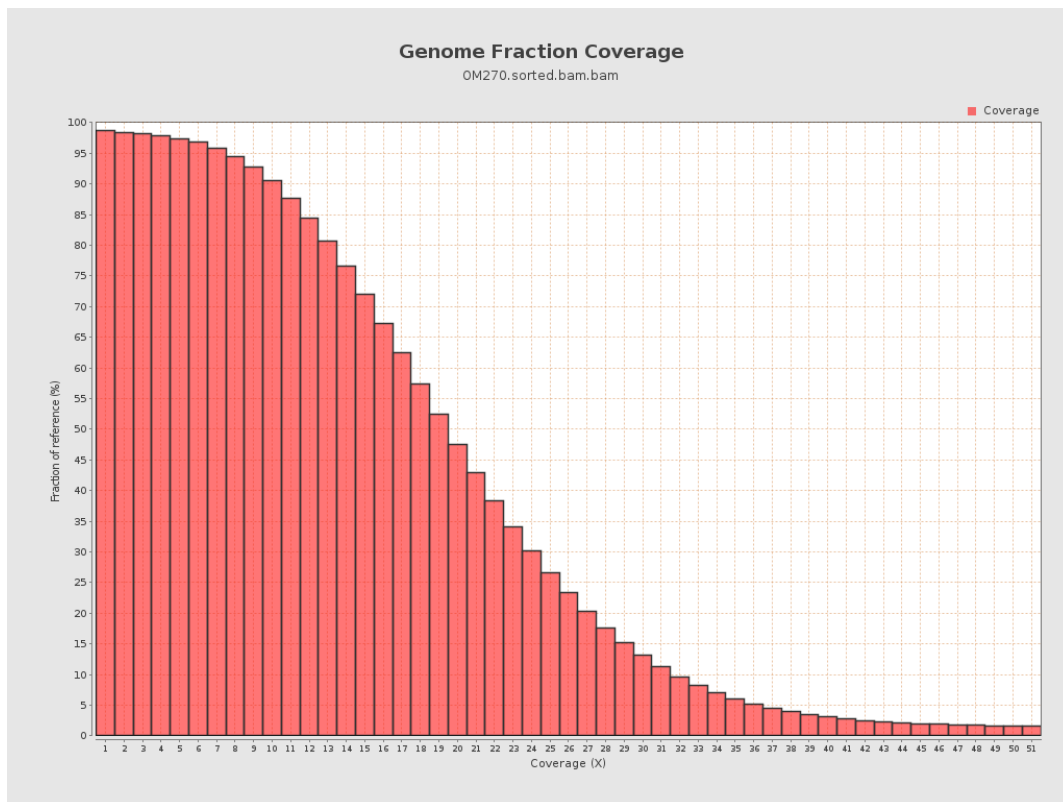
4. Results : Coverage Histogram



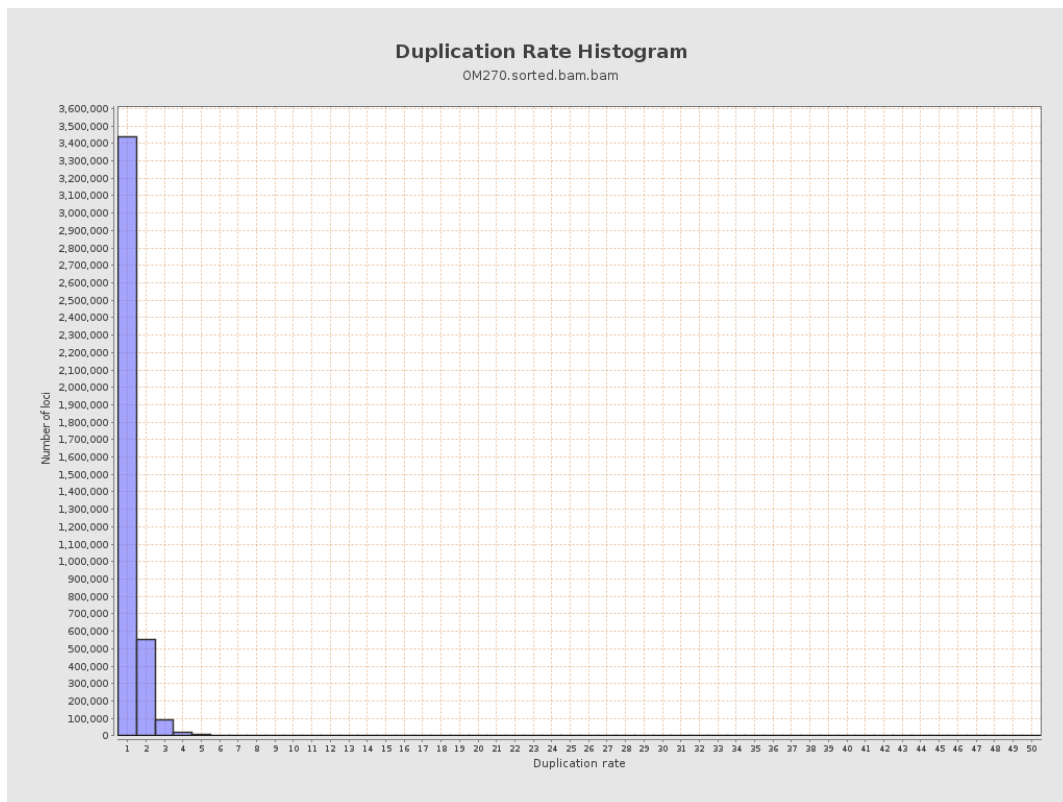
5. Results : Coverage Histogram (0-50X)



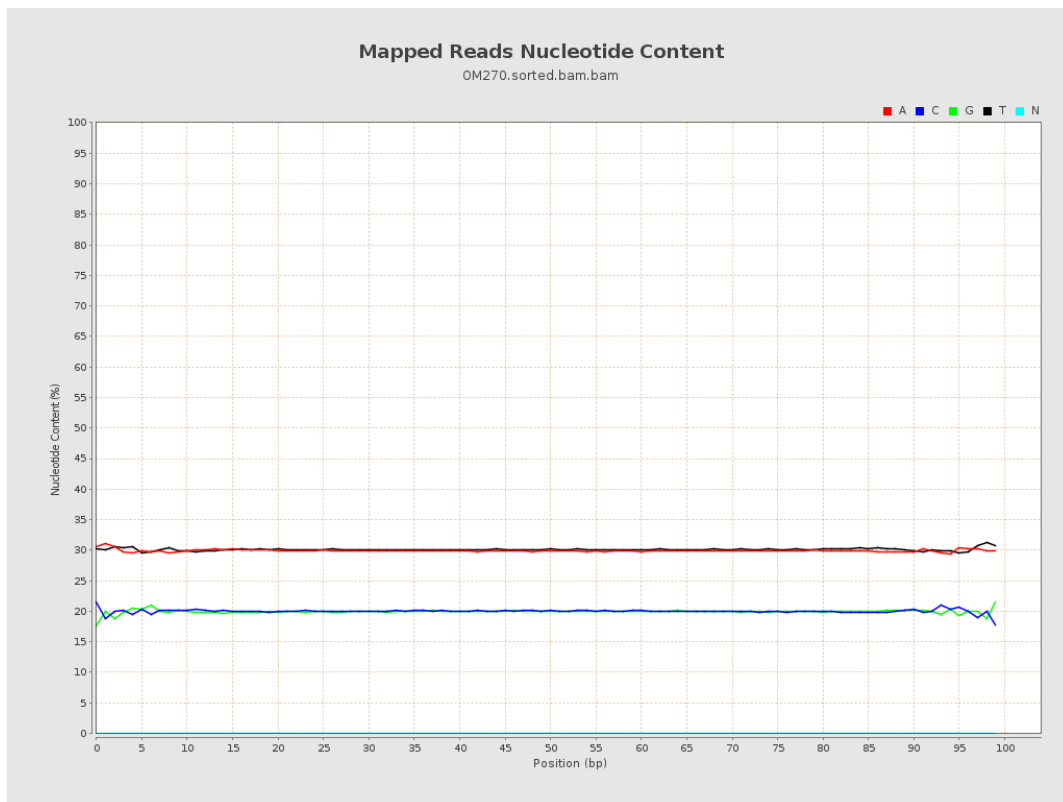
6. Results : Genome Fraction Coverage



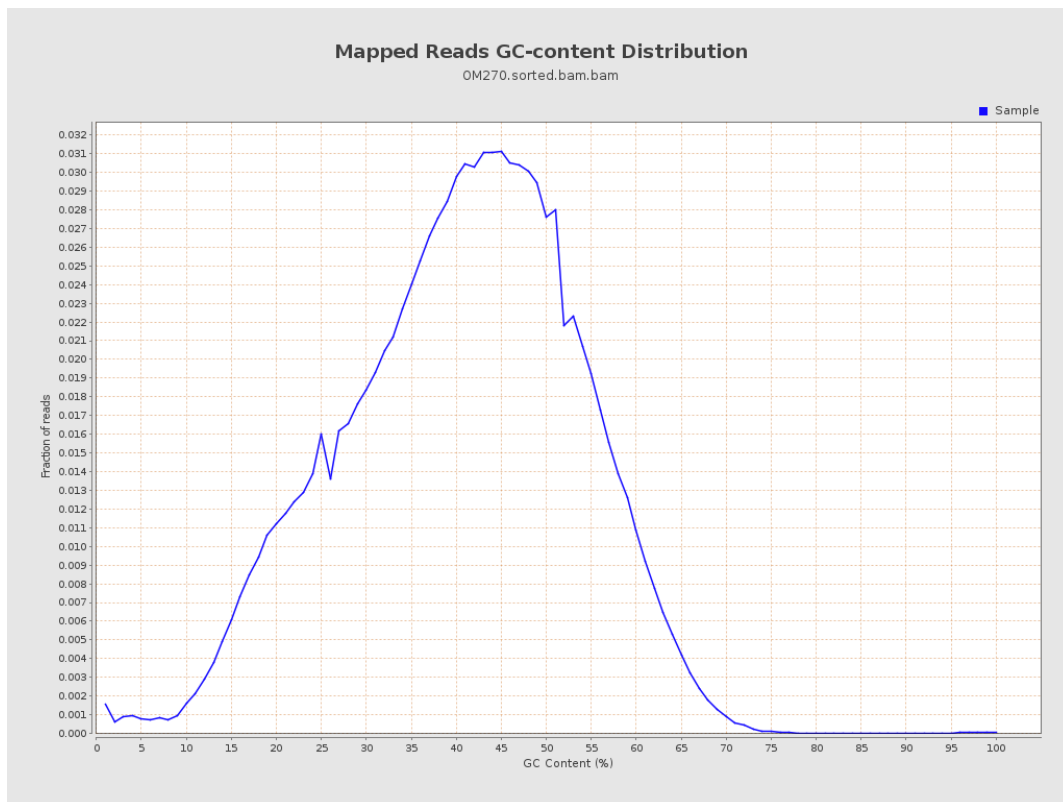
7. Results : Duplication Rate Histogram



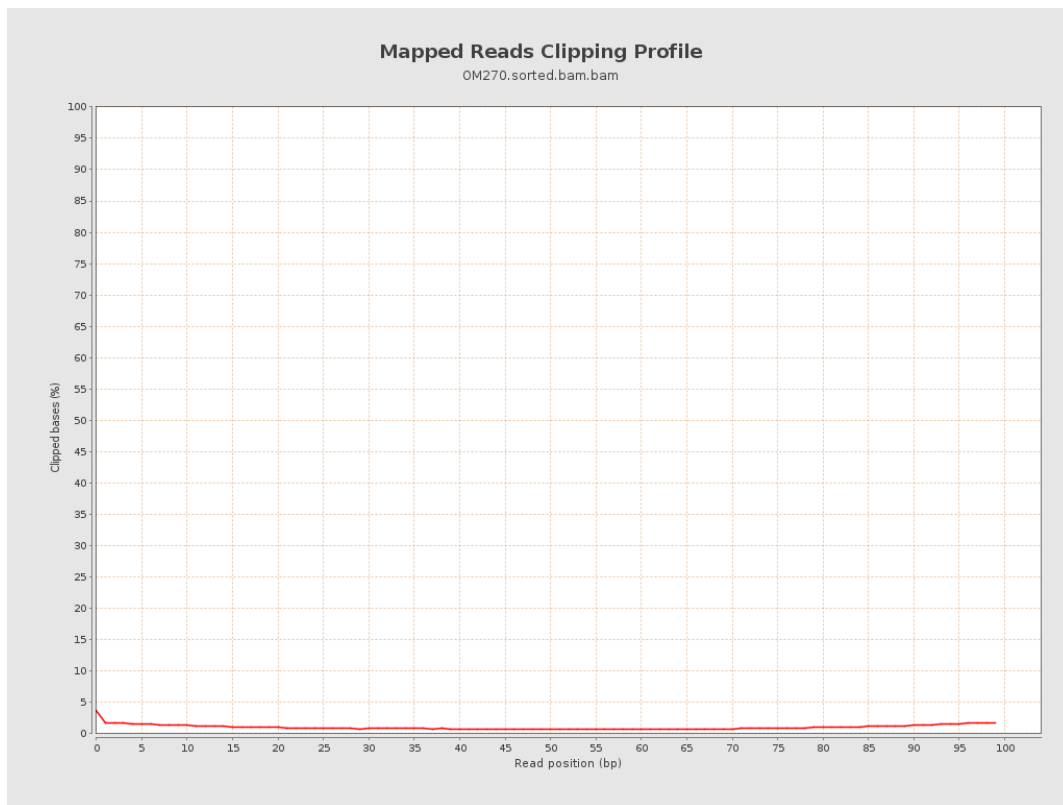
8. Results : Mapped Reads Nucleotide Content



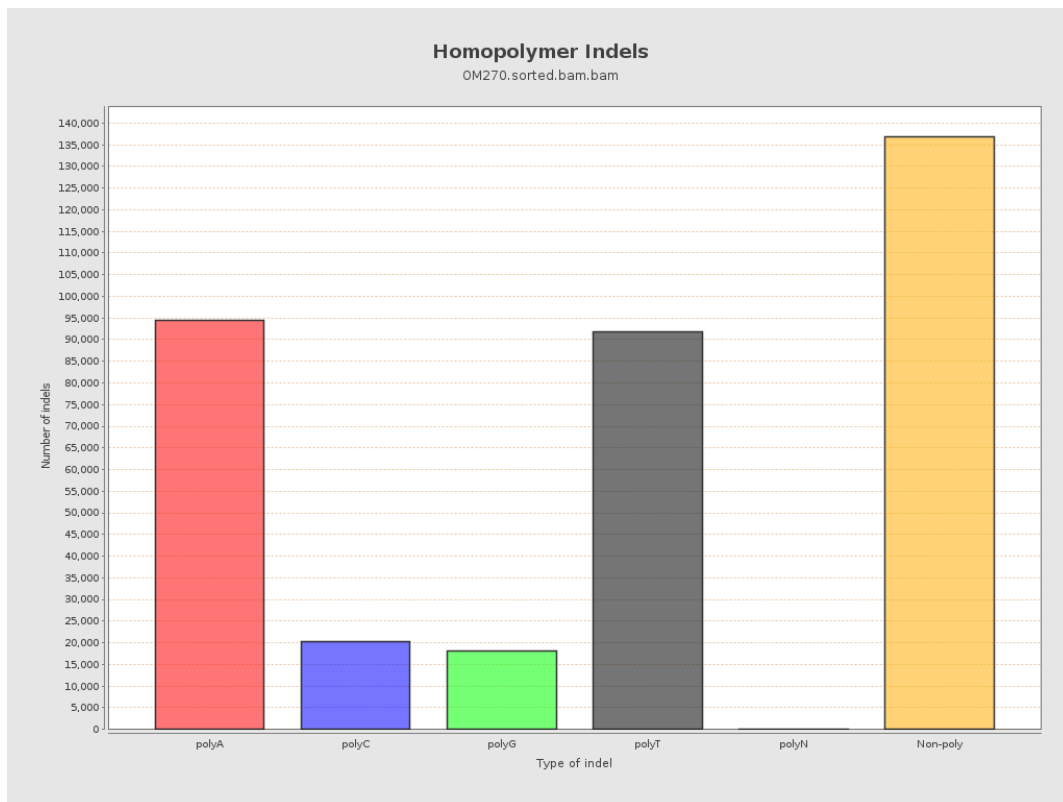
9. Results : Mapped Reads GC-content Distribution



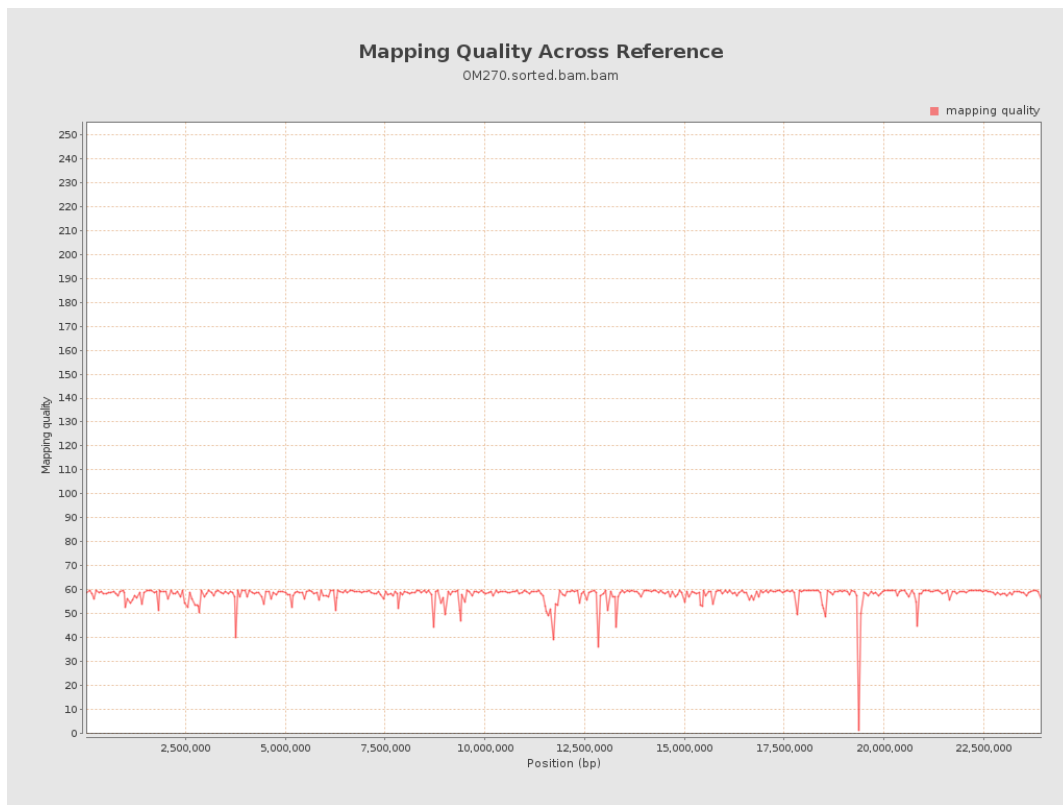
10. Results : Mapped Reads Clipping Profile



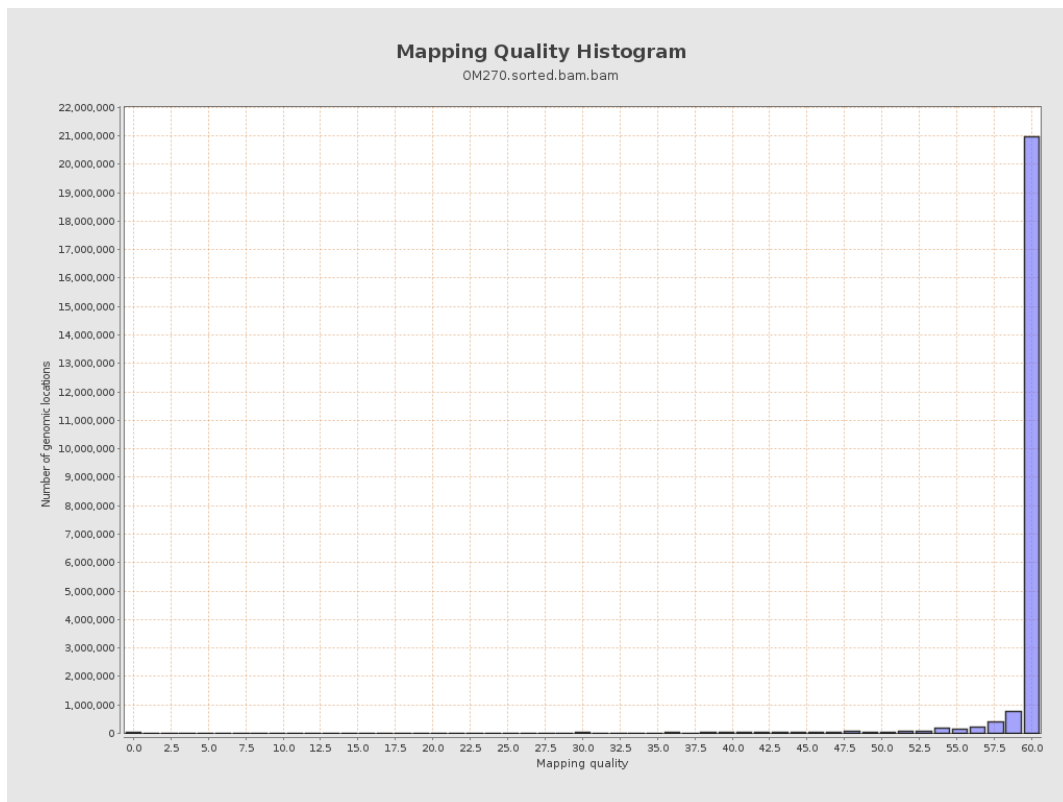
11. Results : Homopolymer Indels



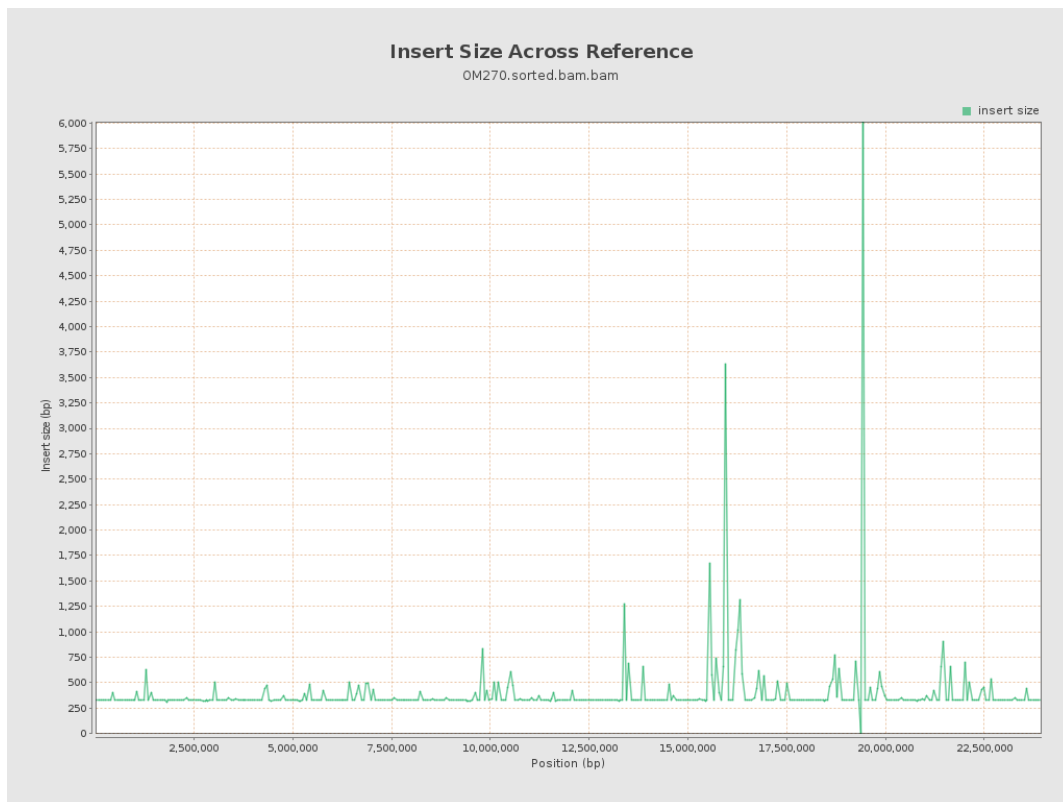
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

