

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 13:40:25

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/OM164.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/OM164-BiooBarcode18_GTCCGC_R2.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/OM164-BiooBarcode18_GTCCGC_R1.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 13:40:25 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/OM164.sorted.bam.

	bam
--	-----

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	10,370,392
Mapped reads	6,819,282 / 65.76%
Unmapped reads	3,551,110 / 34.24%
Mapped paired reads	6,819,282 / 65.76%
Mapped reads, first in pair	3,400,575 / 32.79%
Mapped reads, second in pair	3,418,707 / 32.97%
Mapped reads, both in pair	6,688,421 / 64.5%
Mapped reads, singletons	130,861 / 1.26%
Read min/max/mean length	30 / 100 / 99.9
Duplicated reads (estimated)	1,467,773 / 14.15%
Duplication rate	17.58%
Clipped reads	749,924 / 7.23%

2.2. ACGT Content

Number/percentage of A's	197,628,004 / 29.97%
Number/percentage of C's	131,946,721 / 20.01%
Number/percentage of T's	198,207,002 / 30.05%
Number/percentage of G's	131,722,035 / 19.97%
Number/percentage of N's	54,150 / 0.01%
GC Percentage	39.98%

2.3. Coverage

Mean	27.5553
Standard Deviation	25.6765

2.4. Mapping Quality

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

2.5. Insert size

Mean	772.11
Standard Deviation	21,334.27
P25/Median/P75	323 / 335 / 344

2.6. Mismatches and indels

General error rate	1.57%
Mismatches	9,844,968
Insertions	210,091
Mapped reads with at least one insertion	2.93%
Deletions	241,538
Mapped reads with at least one deletion	3.35%
Homopolymer indels	63.37%

2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

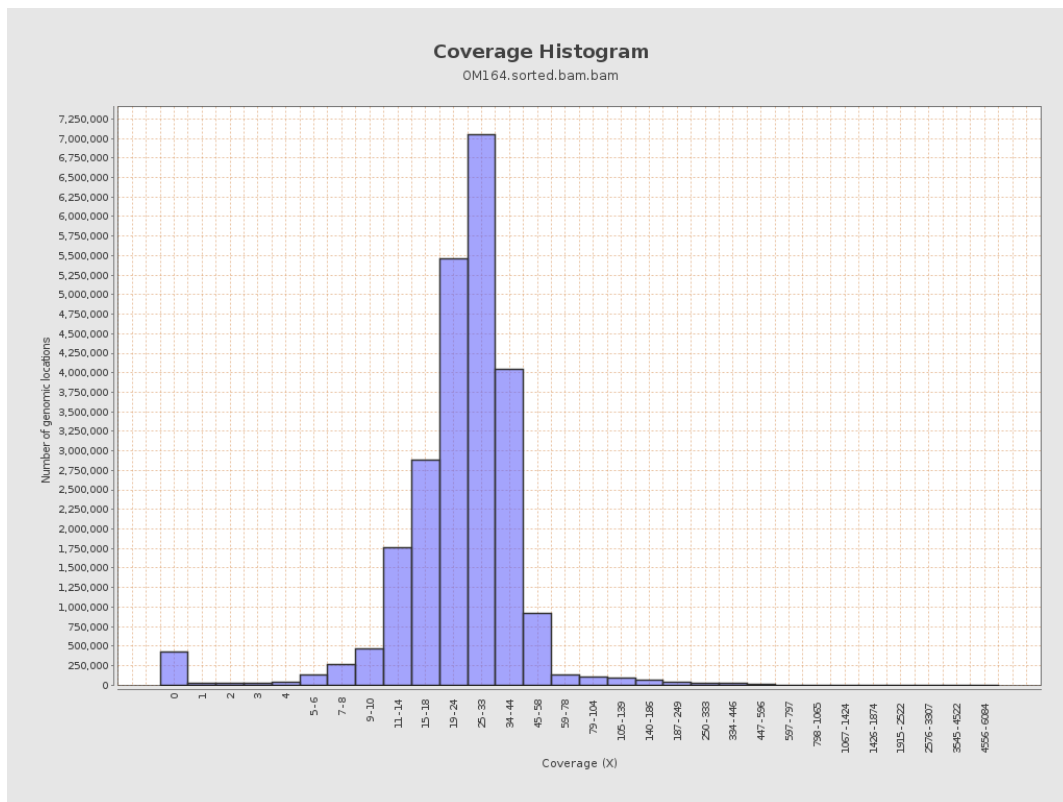
gi 1074120478 emb LT615256.1	977217	23409563	23.9553	12.0662
gi 1074120682 emb LT615257.1	860454	22803906	26.5022	19.0077
gi 1074120865 emb LT615258.1	989719	32452962	32.7901	47.14
gi 1074121086 emb LT615259.1	935450	27737678	29.6517	41.5548
gi 1074121301 emb LT615260.1	1432239	40822135	28.5023	26.7822
gi 1074121615 emb LT615261.1	1080962	31652494	29.2818	34.0386
gi 1074121871 emb LT615262.1	1545099	39643908	25.6578	11.1842
gi 1074122235 emb LT615263.1	1585108	43146079	27.2196	19.2316
gi 1074122590 emb LT615264.1	2122358	55795480	26.2894	14.4409
gi 1074123050 emb LT615265.1	1754192	48529867	27.6651	22.1184
gi 1074123421 emb LT615	2150147	62932194	29.2688	44.5329

266.1				
gi 107412389 8 emb LT615 267.1	3031036	80512799	26.5628	13.3549
gi 107412458 8 emb LT615 268.1	2359348	63722292	27.0084	28.6108
gi 107412506 5 emb LT615 269.1	3135668	87036836	27.757	10.7681

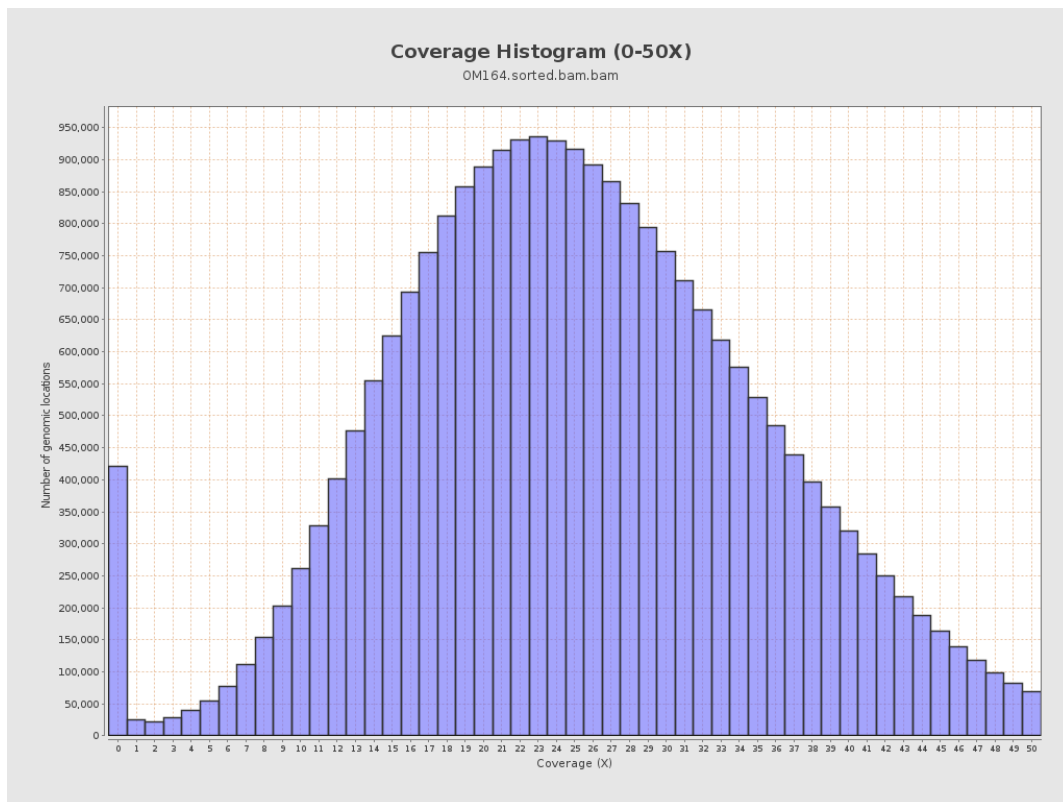
3. Results : Coverage across reference



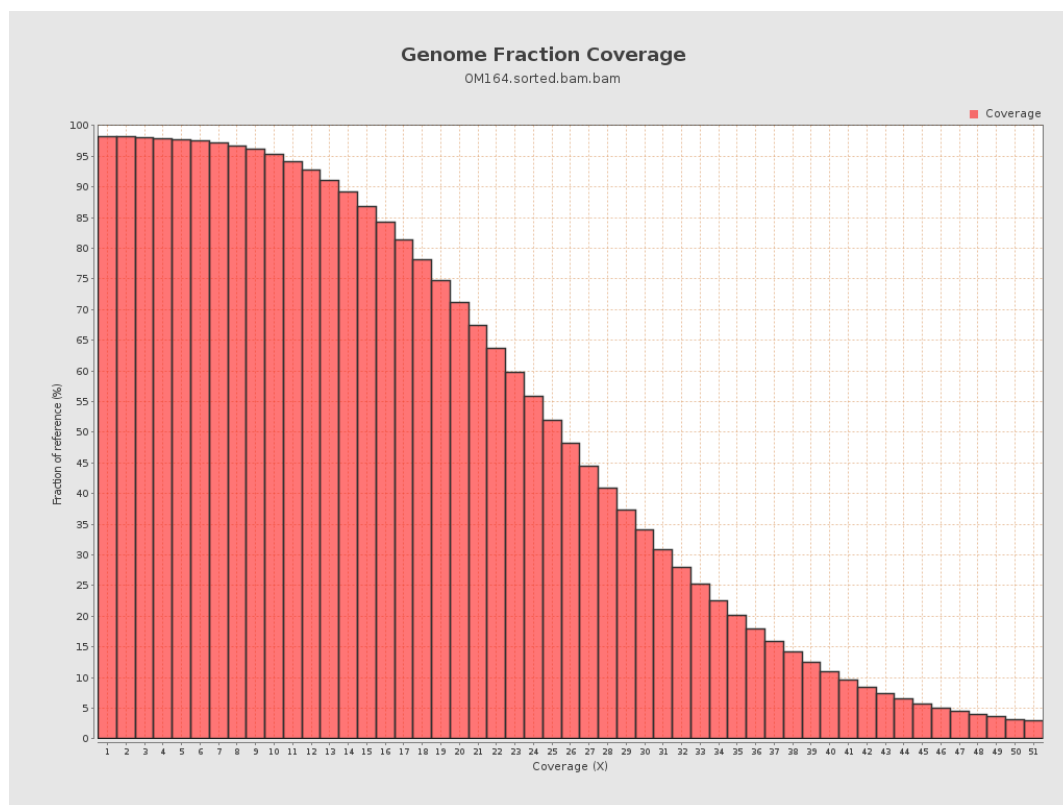
4. Results : Coverage Histogram



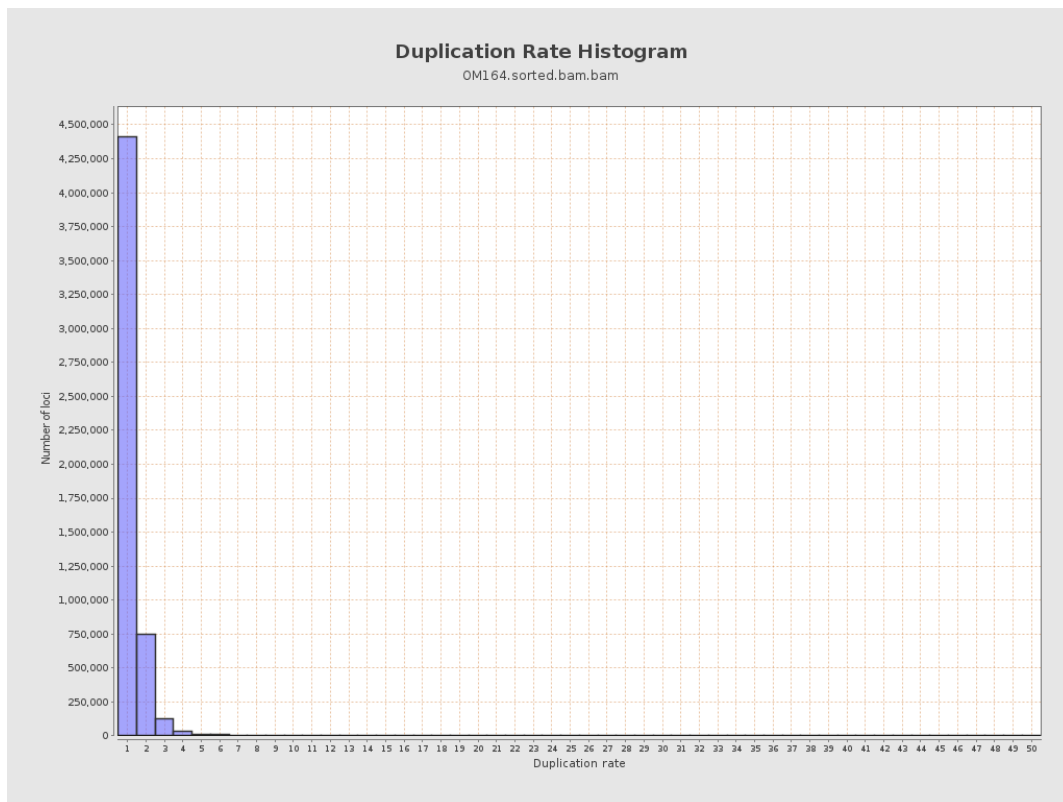
5. Results : Coverage Histogram (0-50X)



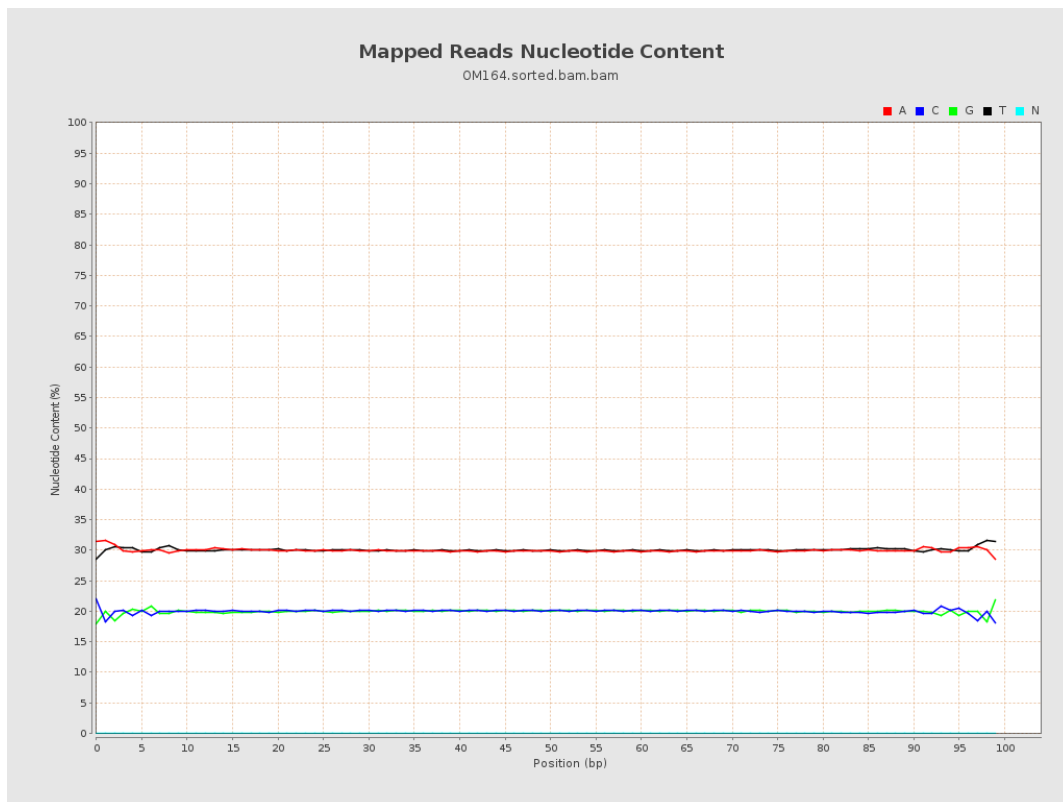
6. Results : Genome Fraction Coverage



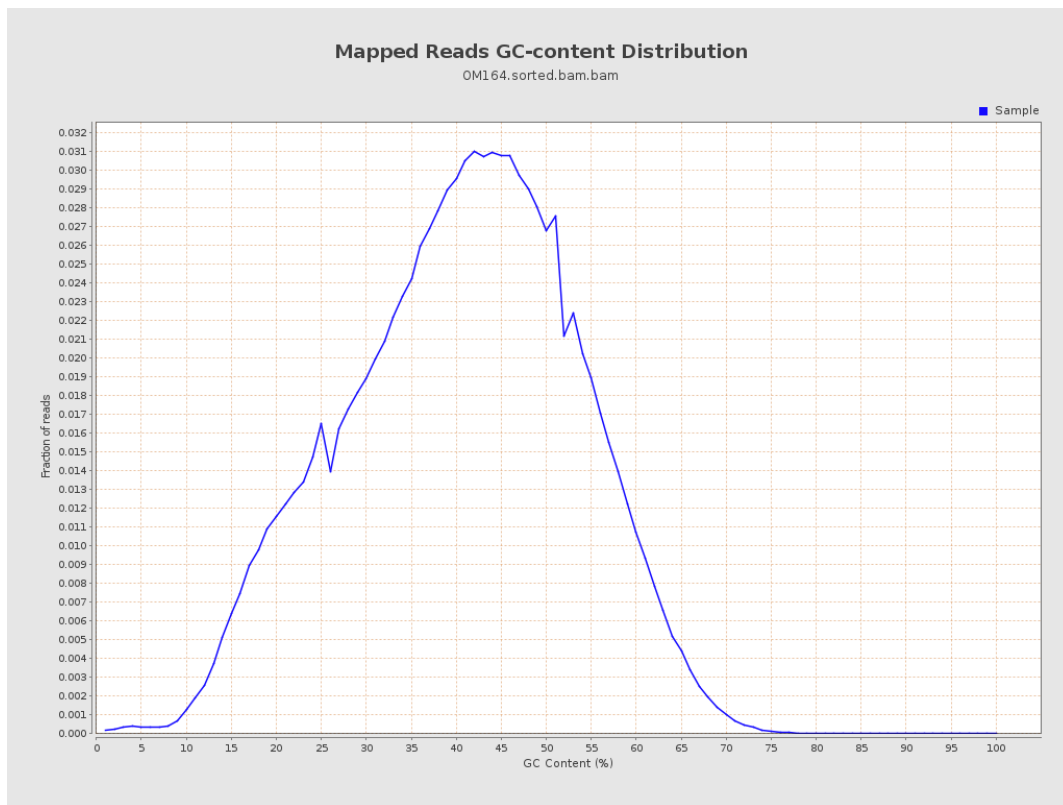
7. Results : Duplication Rate Histogram



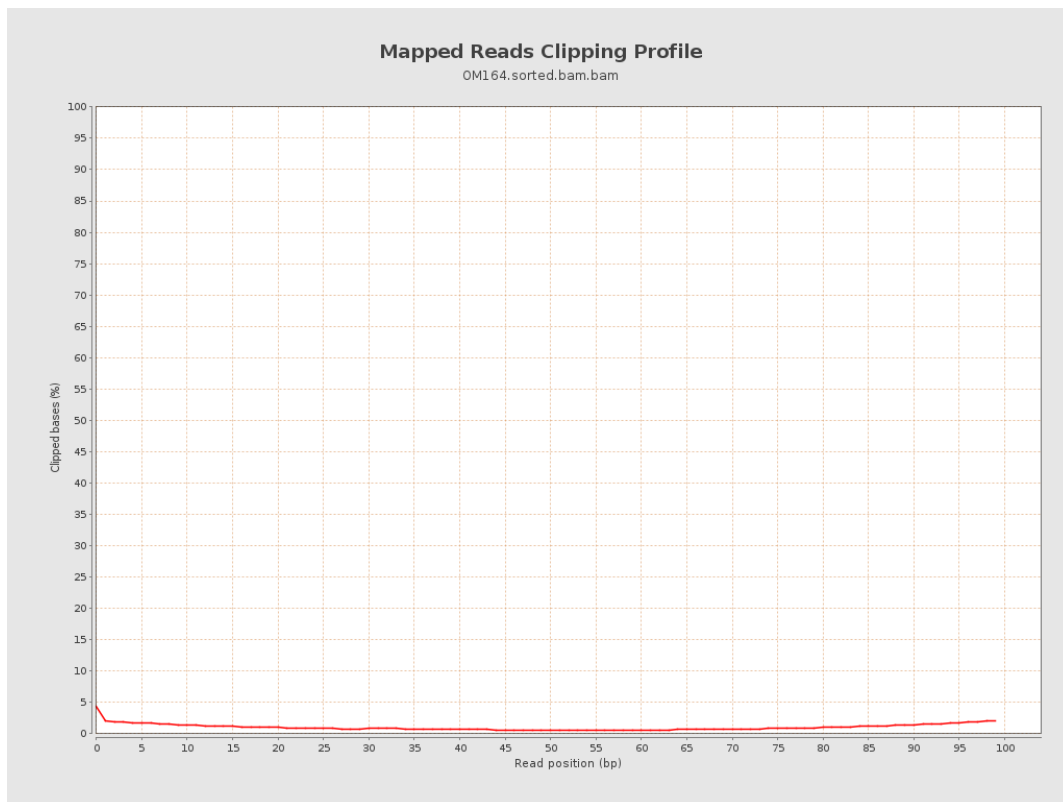
8. Results : Mapped Reads Nucleotide Content



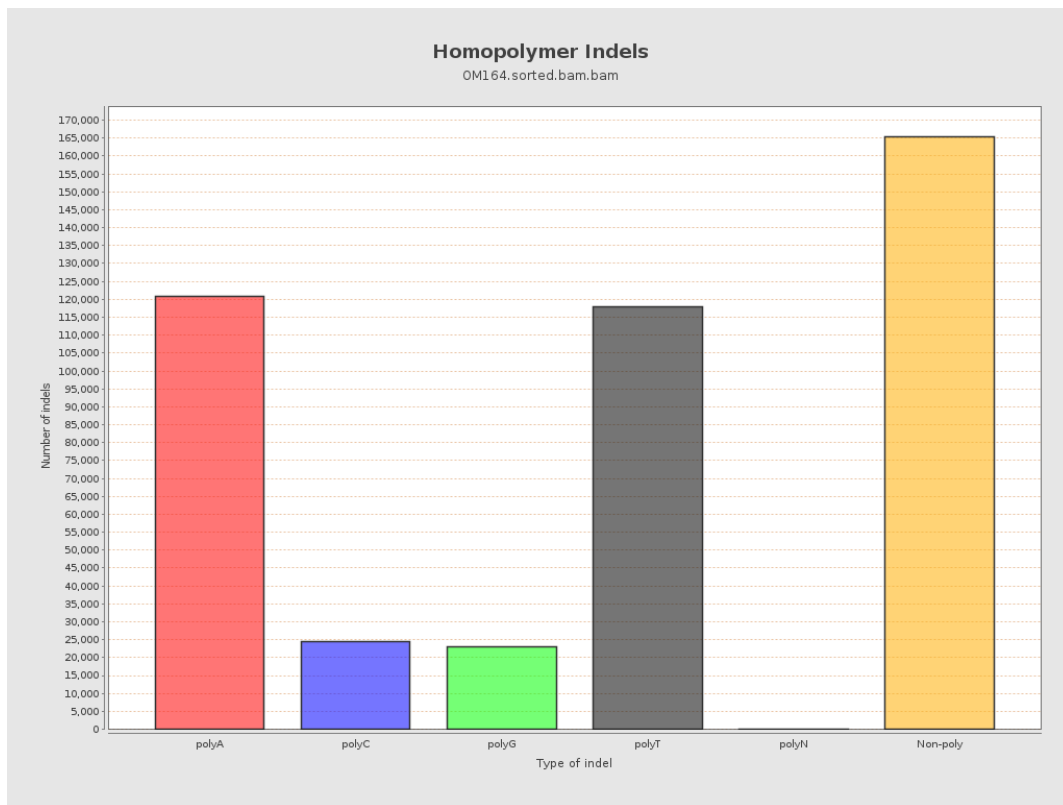
9. Results : Mapped Reads GC-content Distribution



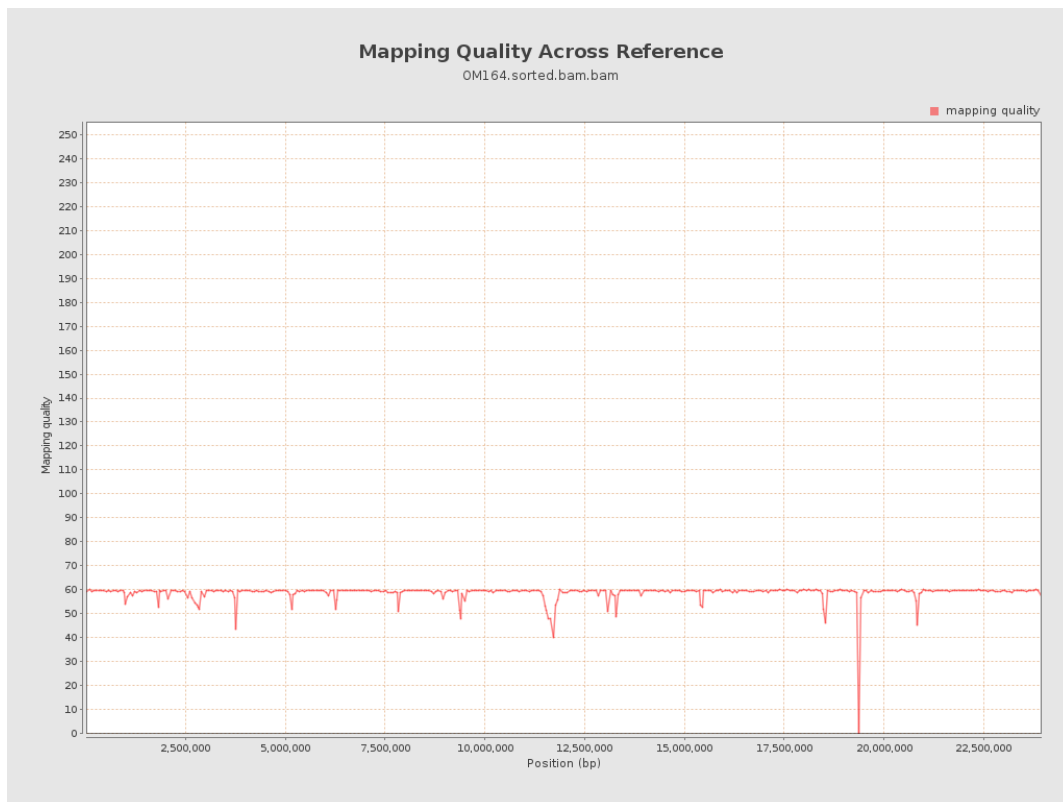
10. Results : Mapped Reads Clipping Profile



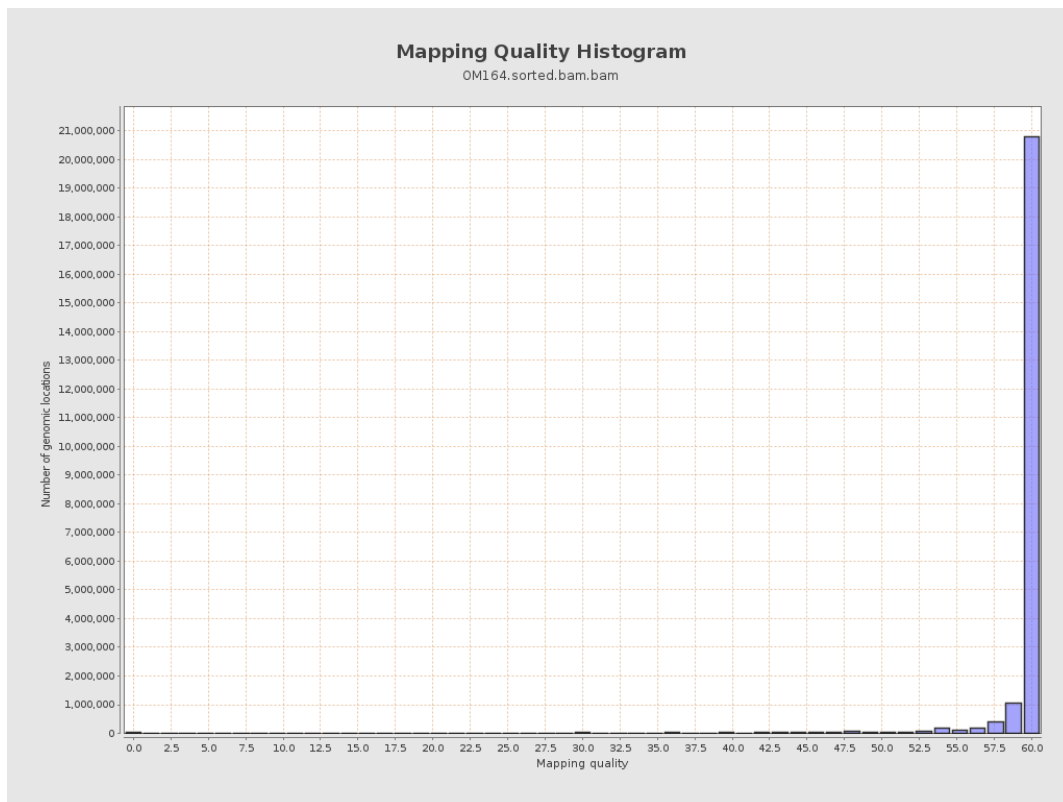
11. Results : Homopolymer Indels



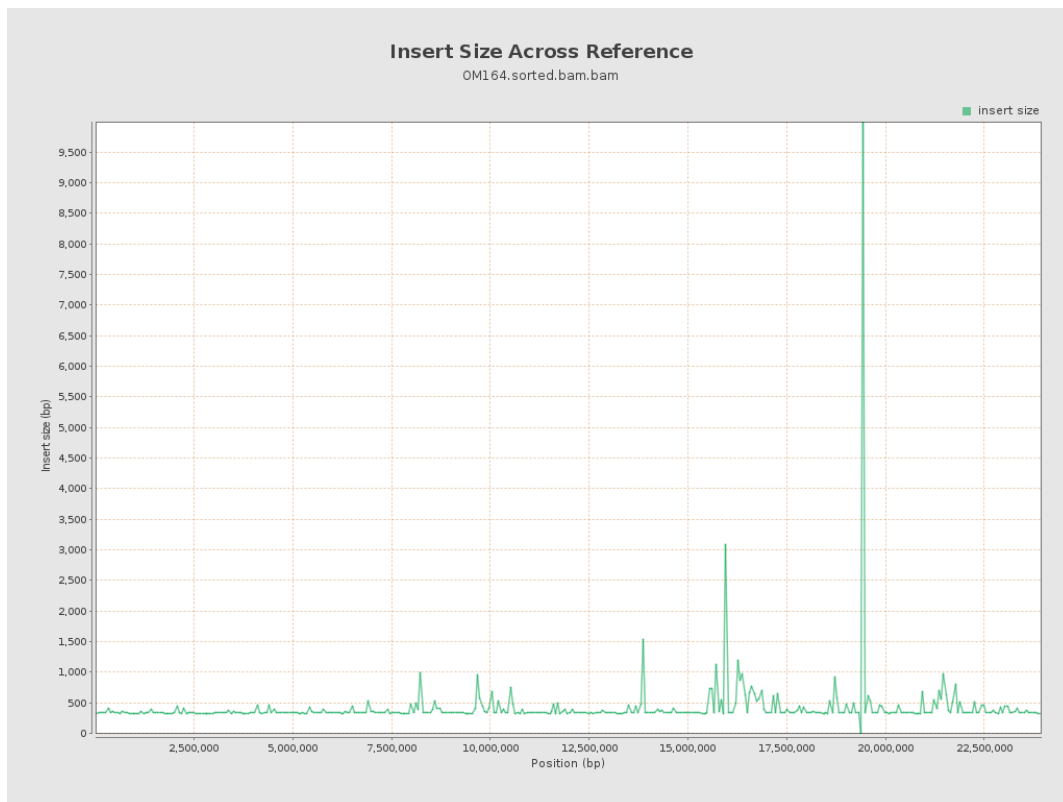
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

