

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 12:24:24

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/OM078.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/OM078-BiooBarcode12_GGCTAC_R1.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/OM078-BiooBarcode12_GGCTAC_R2.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 12:24:23 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/OM078.sorted.bam.

	bam
--	-----

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	12,888,063
Mapped reads	3,744,684 / 29.06%
Unmapped reads	9,143,379 / 70.94%
Mapped paired reads	3,744,684 / 29.06%
Mapped reads, first in pair	1,877,681 / 14.57%
Mapped reads, second in pair	1,867,003 / 14.49%
Mapped reads, both in pair	3,629,890 / 28.16%
Mapped reads, singletons	114,794 / 0.89%
Read min/max/mean length	30 / 100 / 99.95
Duplicated reads (estimated)	591,676 / 4.59%
Duplication rate	11.94%
Clipped reads	457,290 / 3.55%

2.2. ACGT Content

Number/percentage of A's	107,394,509 / 29.91%
Number/percentage of C's	71,886,830 / 20.02%
Number/percentage of T's	108,004,381 / 30.08%
Number/percentage of G's	71,758,074 / 19.99%
Number/percentage of N's	28,695 / 0.01%
GC Percentage	40.01%

2.3. Coverage

Mean	15.0021
Standard Deviation	15.2659

2.4. Mapping Quality

Mean Mapping Quality	58.37
----------------------	-------

2.5. Insert size

Mean	801.02
Standard Deviation	23,346.61
P25/Median/P75	294 / 308 / 317

2.6. Mismatches and indels

General error rate	1.54%
Mismatches	5,249,166
Insertions	118,422
Mapped reads with at least one insertion	3%
Deletions	136,476
Mapped reads with at least one deletion	3.42%
Homopolymer indels	62.34%

2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

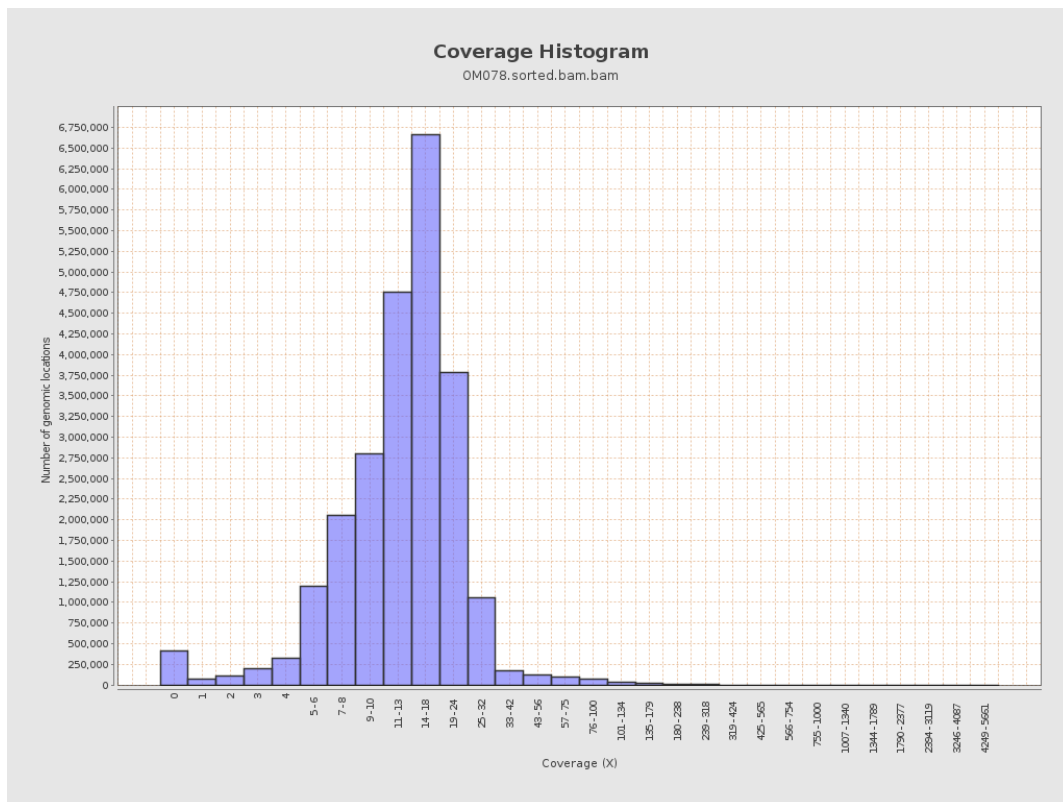
gi 1074120478 emb LT615256.1	977217	13523454	13.8387	11.0973
gi 1074120682 emb LT615257.1	860454	13795516	16.0328	12.6885
gi 1074120865 emb LT615258.1	989719	17557302	17.7397	22.2361
gi 1074121086 emb LT615259.1	935450	14776498	15.7961	20.9916
gi 1074121301 emb LT615260.1	1432239	21944995	15.3222	14.4025
gi 1074121615 emb LT615261.1	1080962	15336143	14.1875	13.151
gi 1074121871 emb LT615262.1	1545099	22344799	14.4617	7.7456
gi 1074122235 emb LT615263.1	1585108	23492180	14.8206	16.2322
gi 1074122590 emb LT615264.1	2122358	30340665	14.2957	8.1354
gi 1074123050 emb LT615265.1	1754192	28046628	15.9883	28.2302
gi 1074123421 emb LT615	2150147	33499519	15.5801	21.218

266.1				
gi 107412389 8 emb LT615 267.1	3031036	44033394	14.5275	10.1746
gi 107412458 8 emb LT615 268.1	2359348	33968210	14.3973	14.0654
gi 107412506 5 emb LT615 269.1	3135668	46775820	14.9173	6.3191

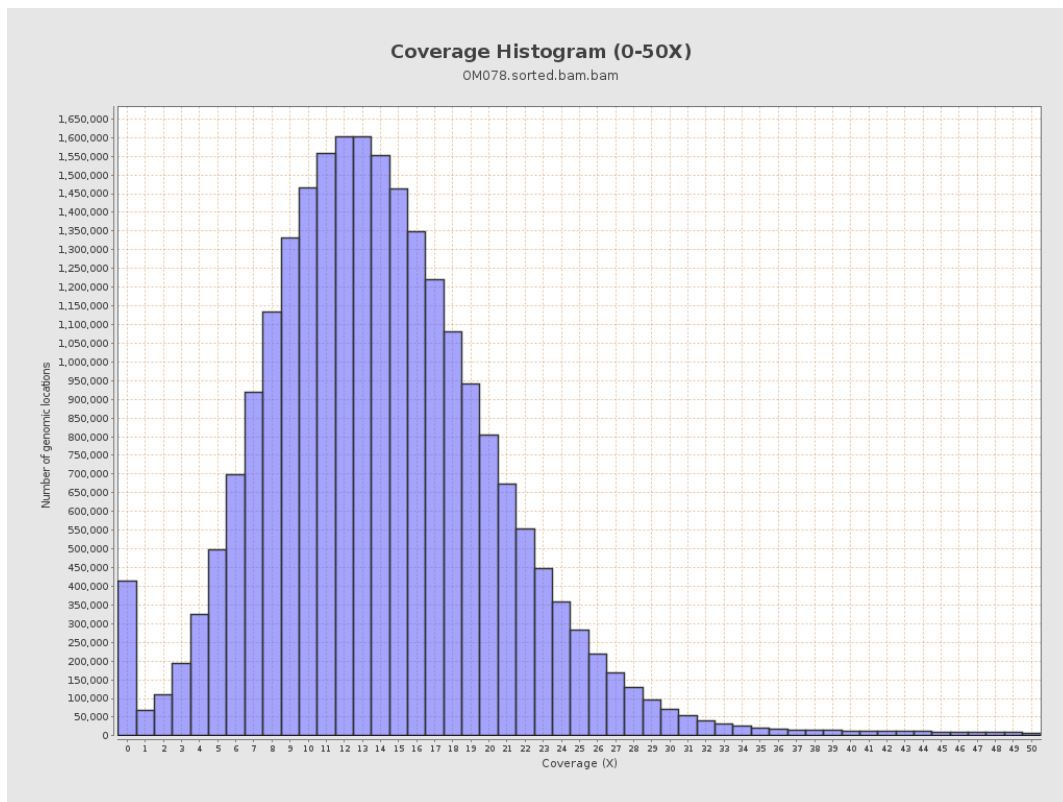
3. Results : Coverage across reference



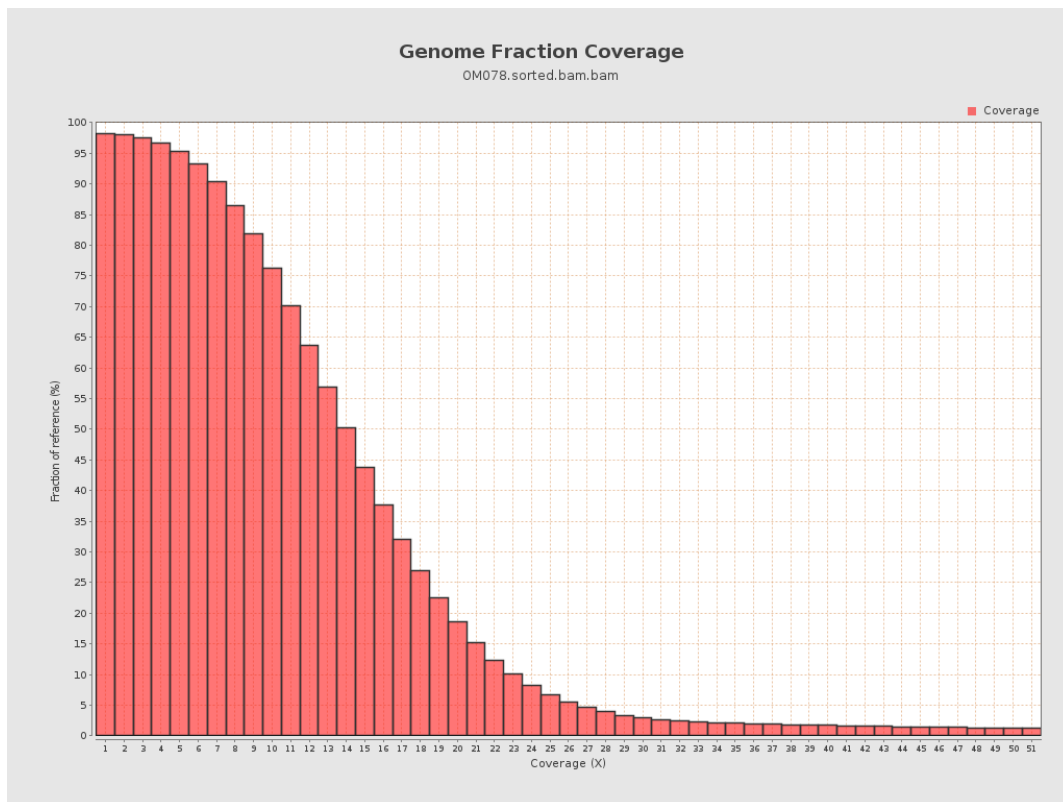
4. Results : Coverage Histogram



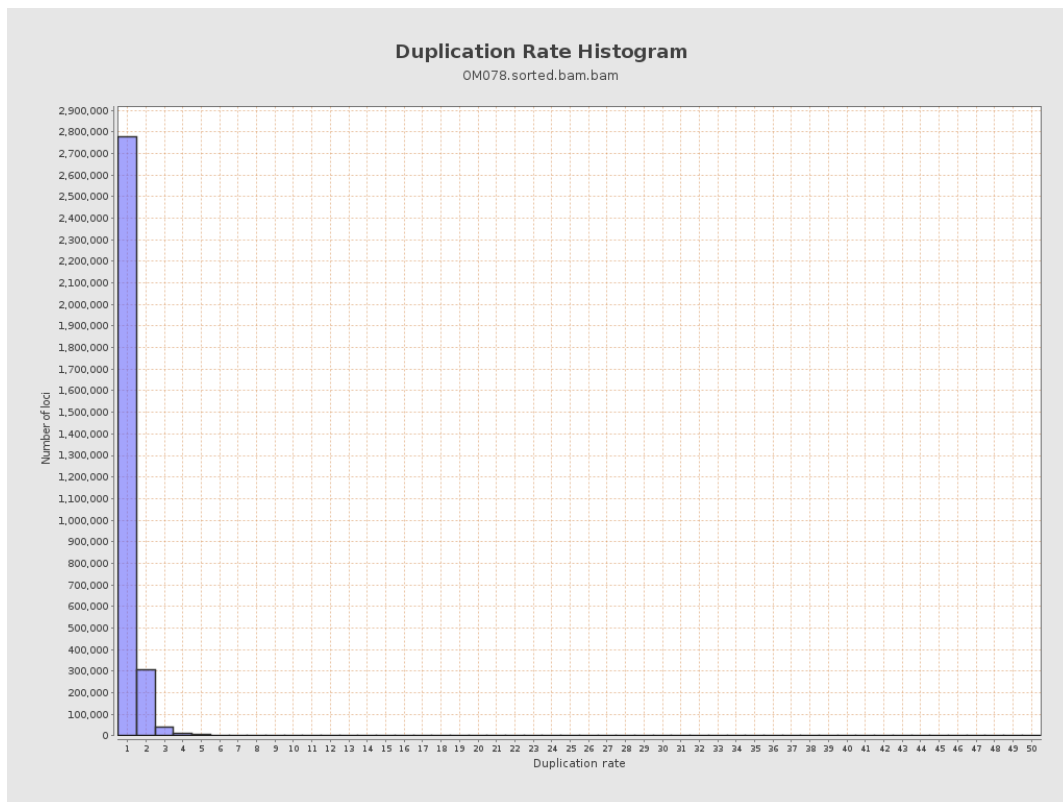
5. Results : Coverage Histogram (0-50X)



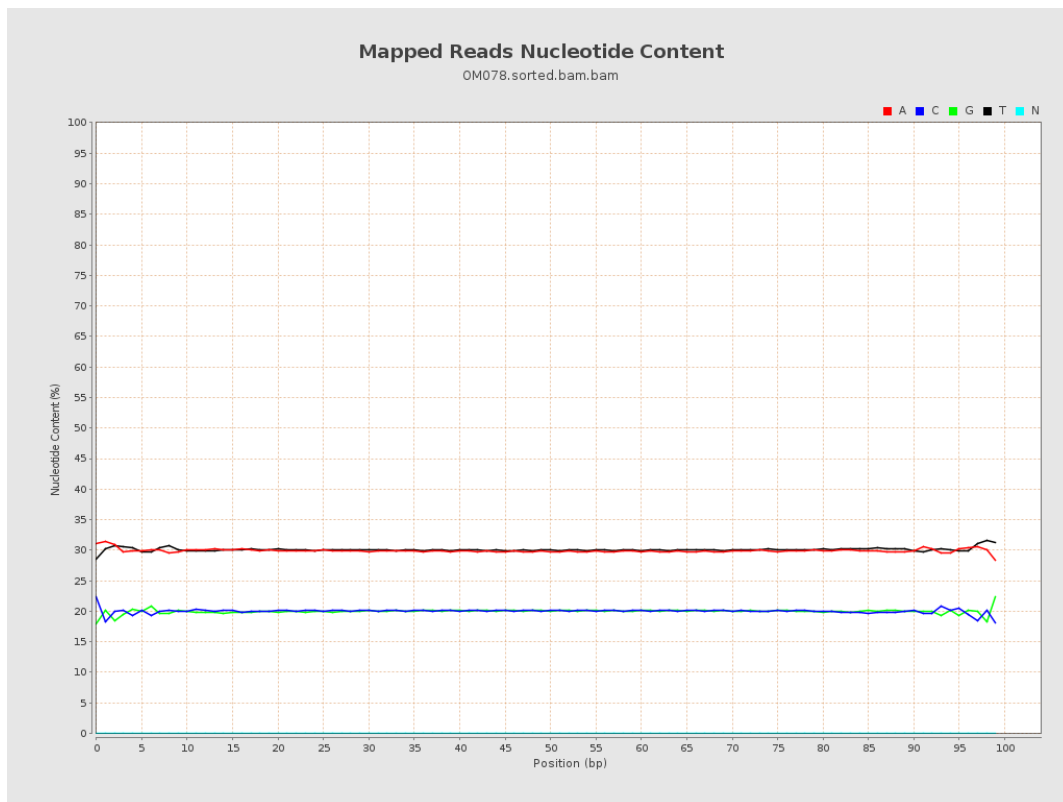
6. Results : Genome Fraction Coverage



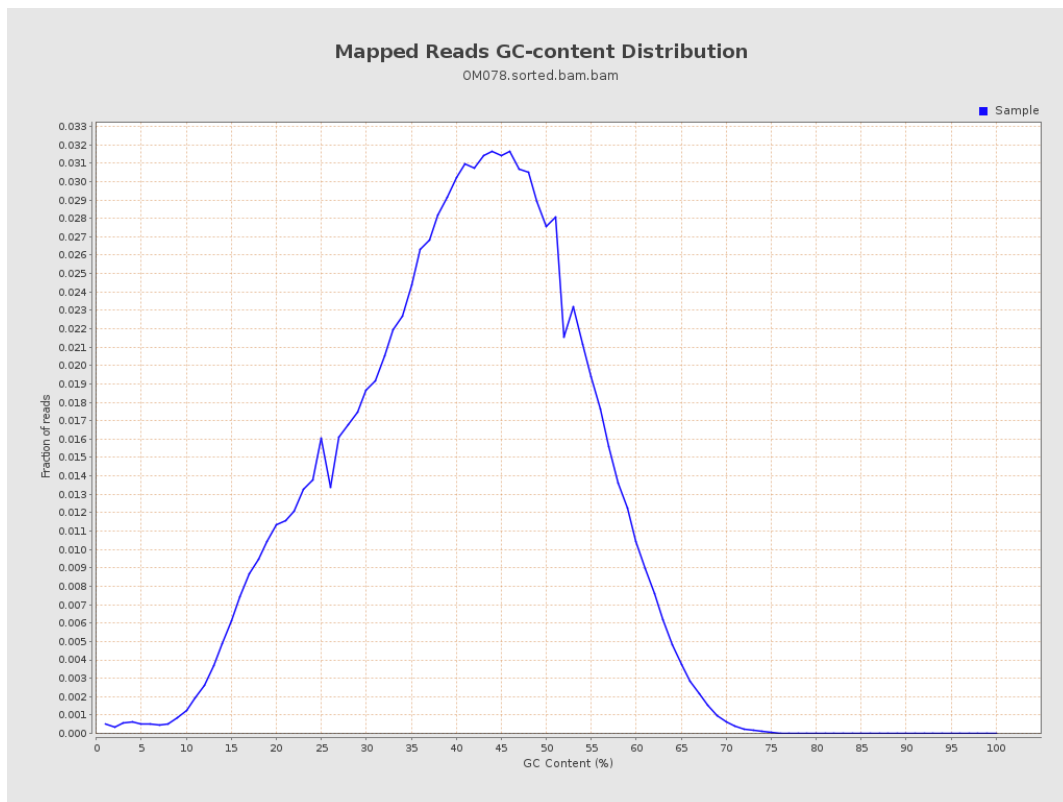
7. Results : Duplication Rate Histogram



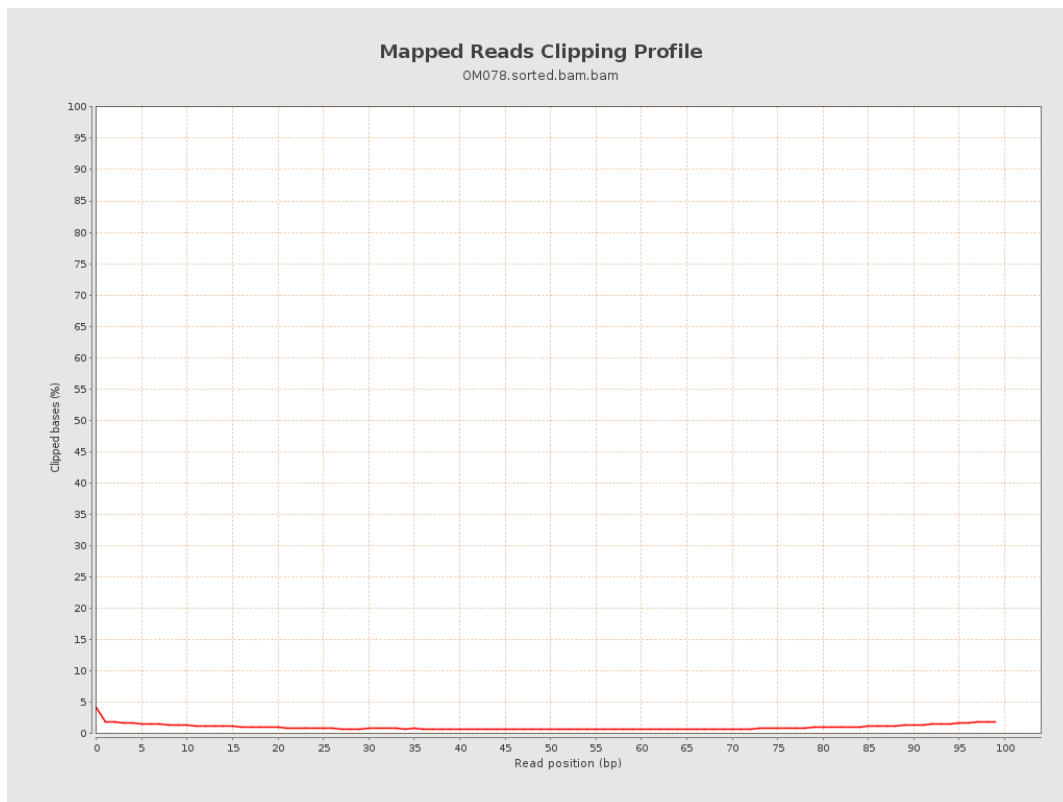
8. Results : Mapped Reads Nucleotide Content



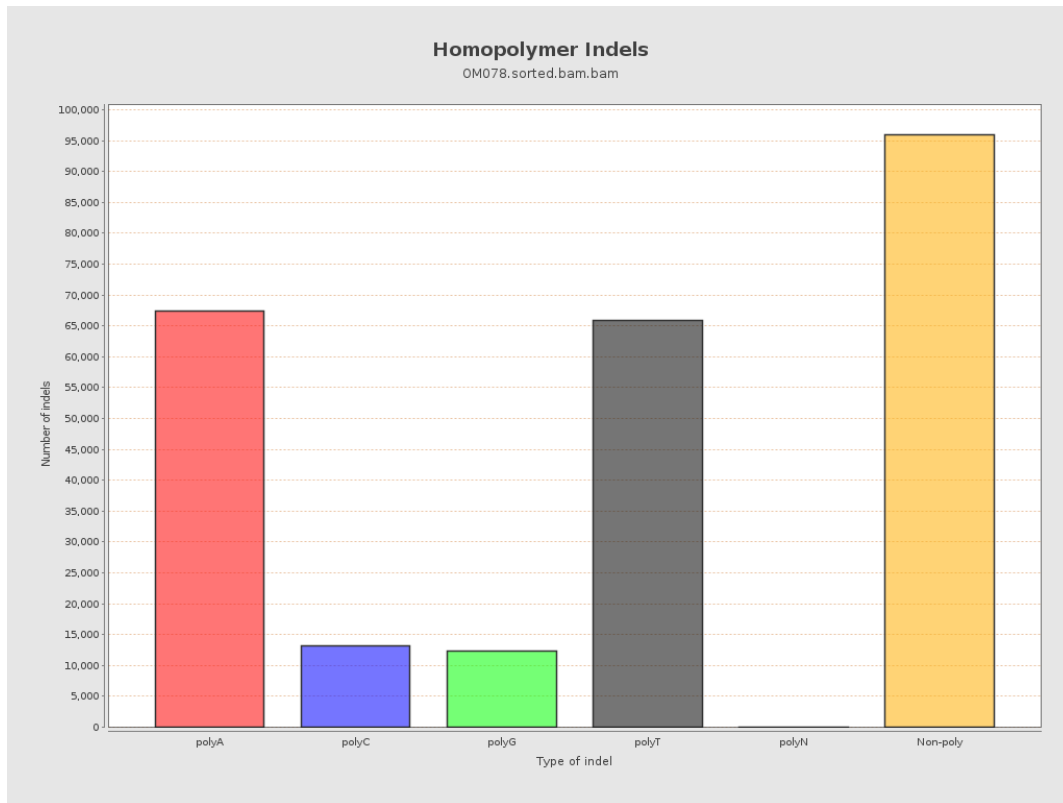
9. Results : Mapped Reads GC-content Distribution



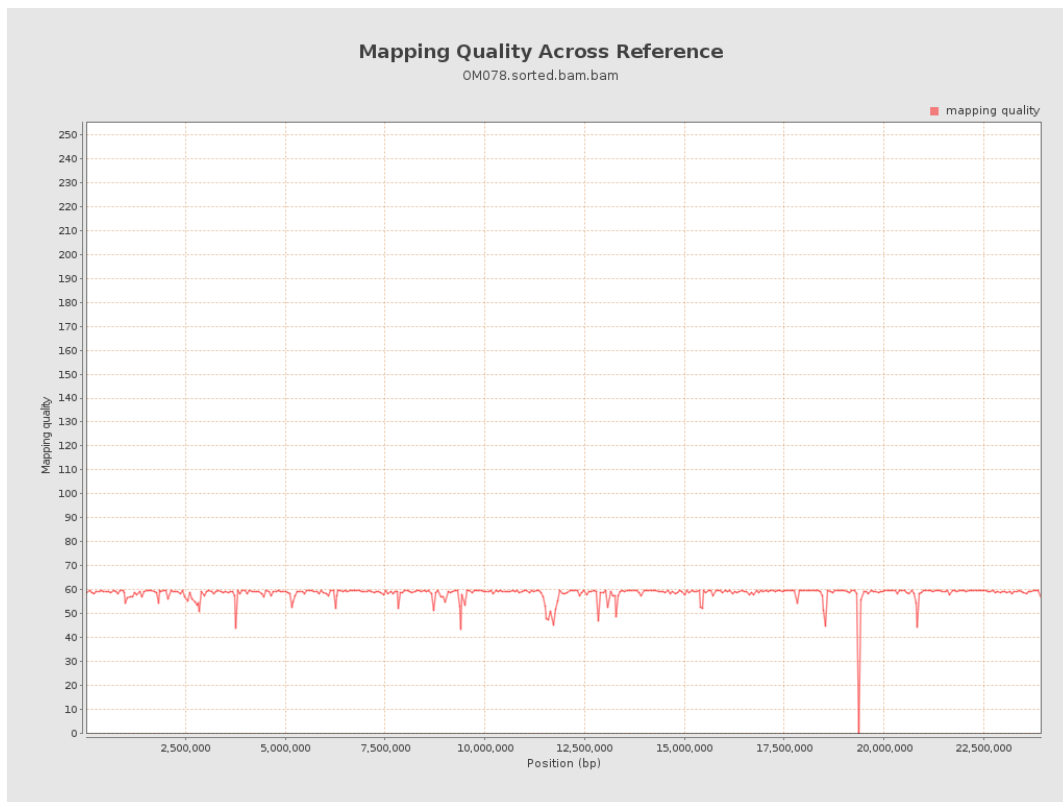
10. Results : Mapped Reads Clipping Profile



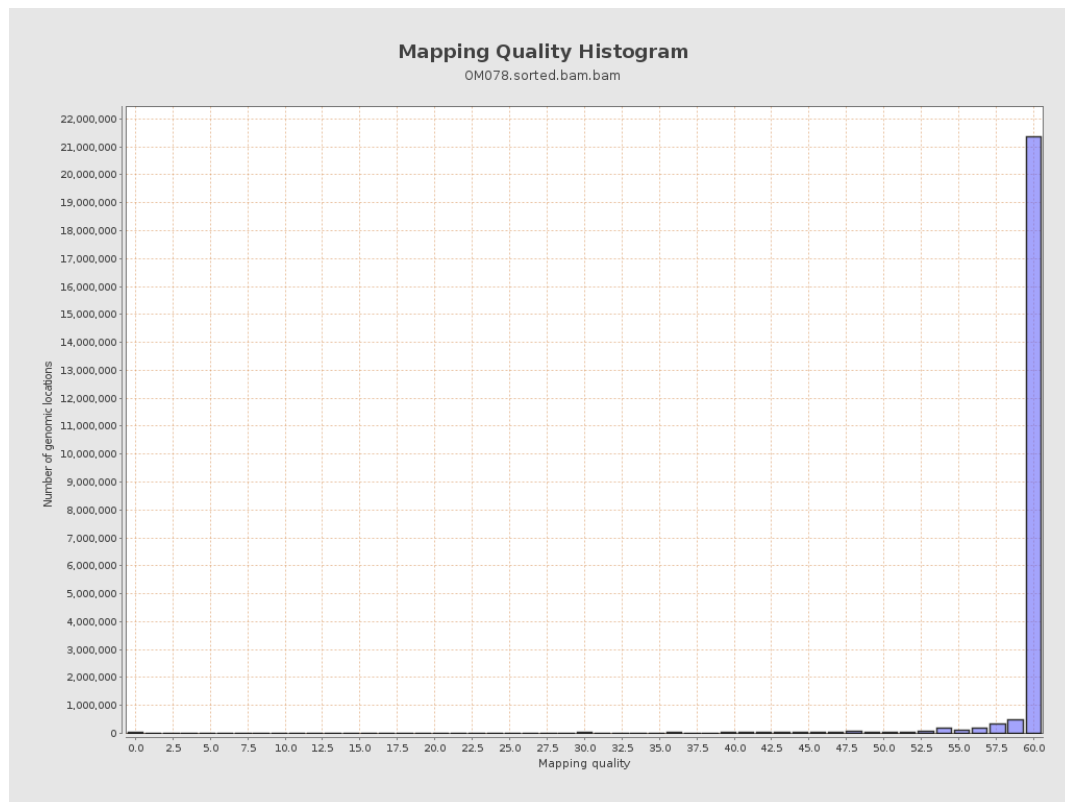
11. Results : Homopolymer Indels



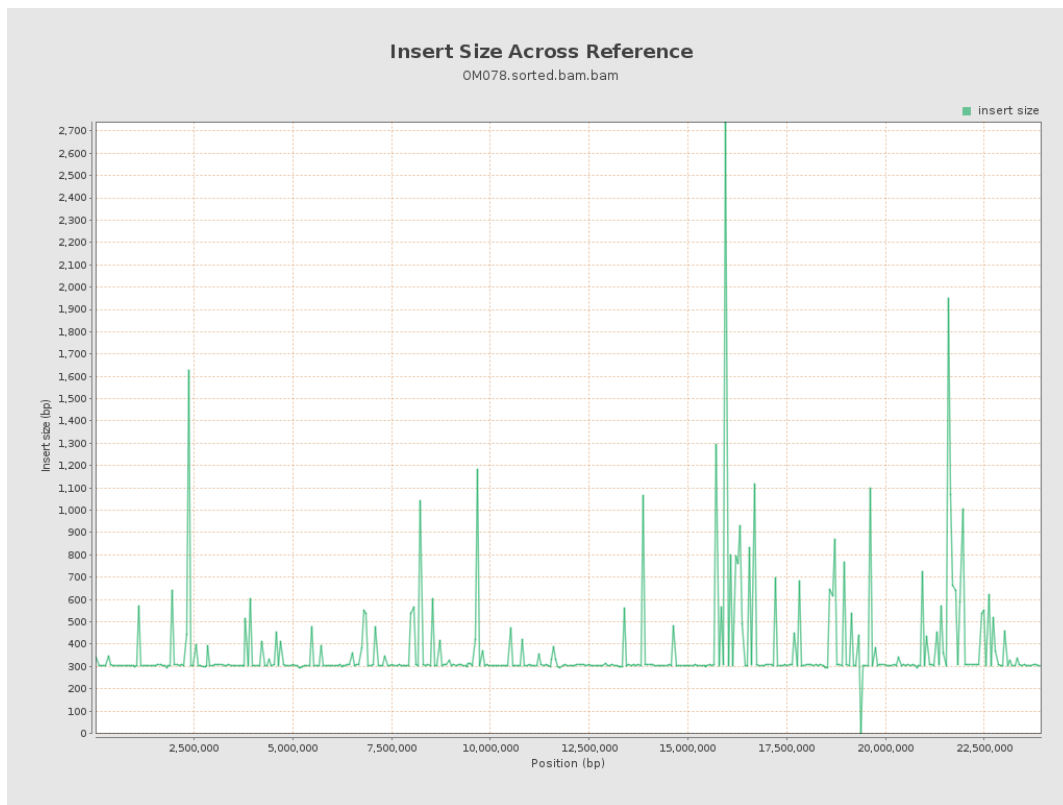
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

