

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 12:26:37

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/OM092.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/OM092-BiooBarcode13_AGTCAA_R1.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/OM092-BiooBarcode13_AGTCAA_R2.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 12:26:36 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/OM092.sorted.bam.

	bam
--	-----

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	16,054,406
Mapped reads	5,794,941 / 36.1%
Unmapped reads	10,259,465 / 63.9%
Mapped paired reads	5,794,941 / 36.1%
Mapped reads, first in pair	2,902,422 / 18.08%
Mapped reads, second in pair	2,892,519 / 18.02%
Mapped reads, both in pair	5,660,643 / 35.26%
Mapped reads, singletons	134,298 / 0.84%
Read min/max/mean length	30 / 100 / 99.93
Duplicated reads (estimated)	1,139,420 / 7.1%
Duplication rate	15.46%
Clipped reads	664,621 / 4.14%

2.2. ACGT Content

Number/percentage of A's	166,744,255 / 29.89%
Number/percentage of C's	112,009,954 / 20.08%
Number/percentage of T's	167,412,340 / 30.01%
Number/percentage of G's	111,721,450 / 20.03%
Number/percentage of N's	49,038 / 0.01%
GC Percentage	40.1%

2.3. Coverage

Mean	23.3101
Standard Deviation	22.6887

2.4. Mapping Quality

Mean Mapping Quality	58.51
----------------------	-------

2.5. Insert size

Mean	659.85
Standard Deviation	22,485.94
P25/Median/P75	194 / 202 / 212

2.6. Mismatches and indels

General error rate	1.45%
Mismatches	7,688,916
Insertions	183,566
Mapped reads with at least one insertion	2.99%
Deletions	208,810
Mapped reads with at least one deletion	3.39%
Homopolymer indels	62.71%

2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

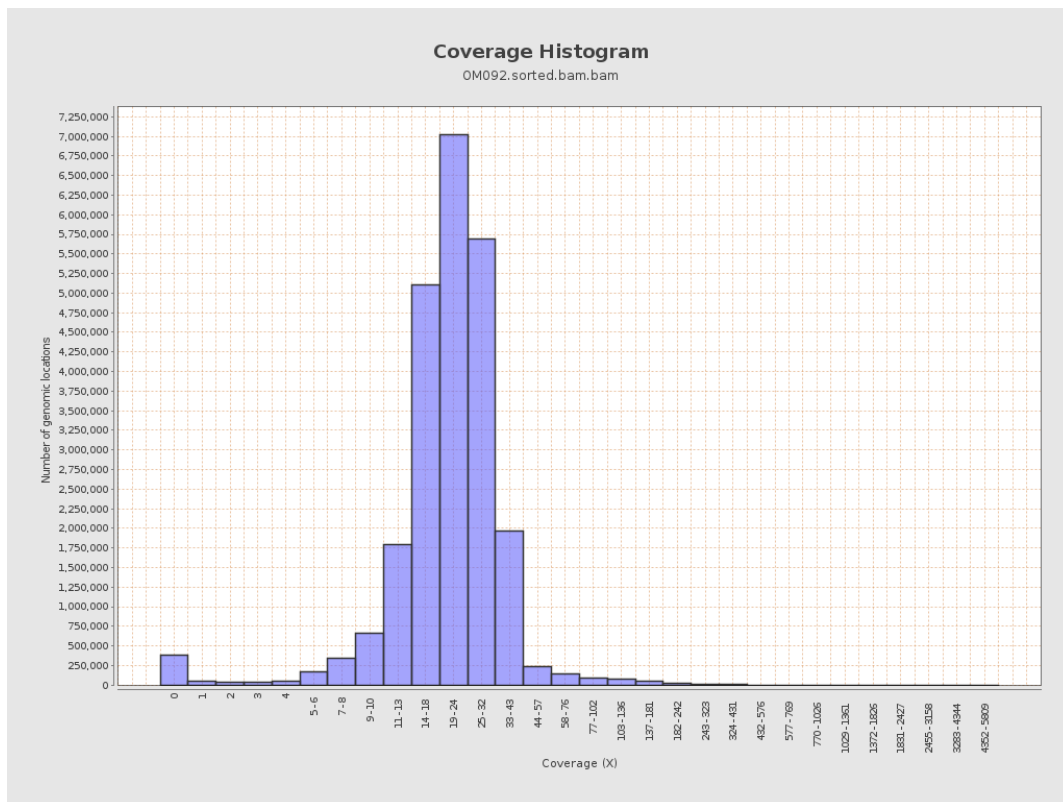
gi 1074120478 emb LT615256.1	977217	20327107	20.801	9.3274
gi 1074120682 emb LT615257.1	860454	18687084	21.7177	15.9132
gi 1074120865 emb LT615258.1	989719	27741112	28.0293	39.1273
gi 1074121086 emb LT615259.1	935450	23411962	25.0275	35.1736
gi 1074121301 emb LT615260.1	1432239	34109174	23.8153	21.4934
gi 1074121615 emb LT615261.1	1080962	26364847	24.3902	21.5726
gi 1074121871 emb LT615262.1	1545099	33609301	21.7522	9.2911
gi 1074122235 emb LT615263.1	1585108	36650879	23.122	20.9267
gi 1074122590 emb LT615264.1	2122358	47478961	22.3709	12.4366
gi 1074123050 emb LT615265.1	1754192	43670666	24.895	33.7922
gi 1074123421 emb LT615	2150147	51632325	24.0134	35.7421

266.1				
gi 107412389 8 emb LT615 267.1	3031036	67555659	22.288	12.135
gi 107412458 8 emb LT615 268.1	2359348	54863674	23.2537	25.8169
gi 107412506 5 emb LT615 269.1	3135668	72384038	23.0841	8.3739

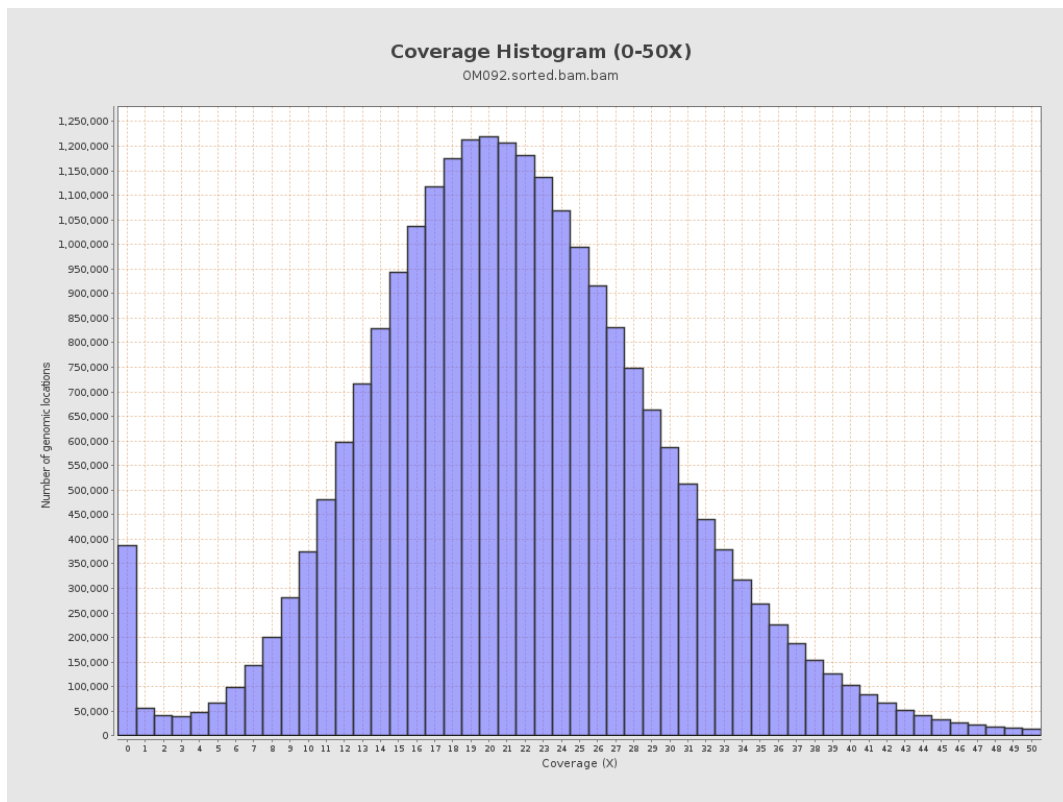
3. Results : Coverage across reference



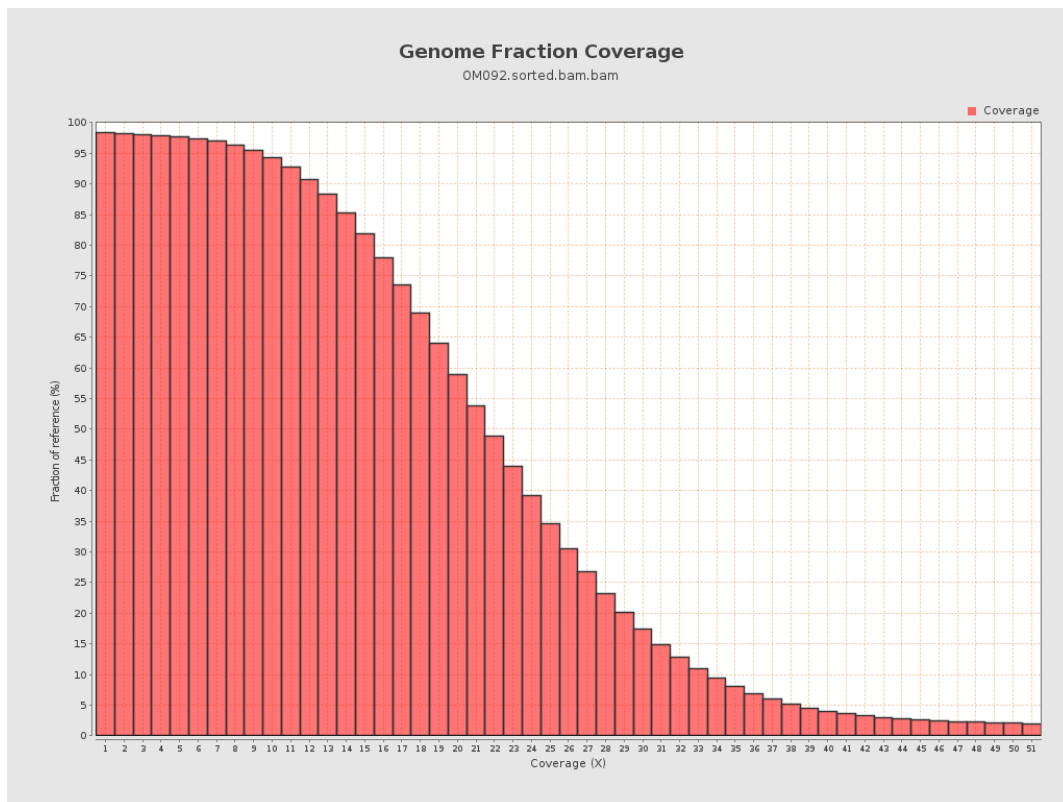
4. Results : Coverage Histogram



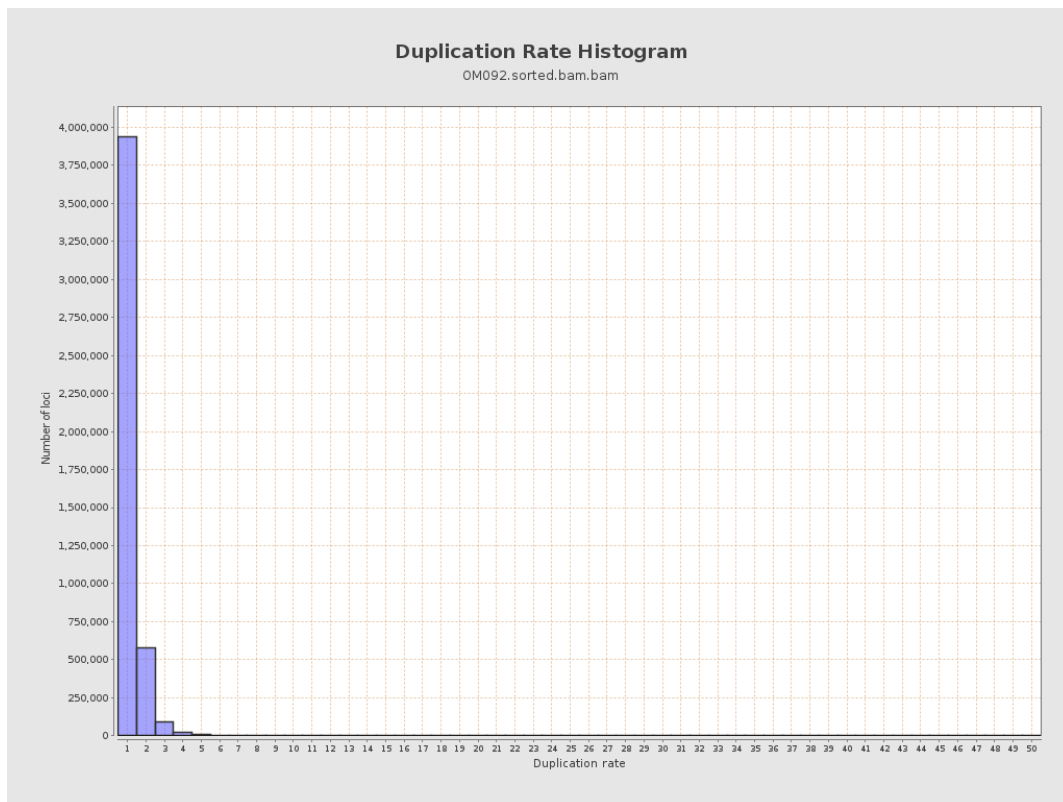
5. Results : Coverage Histogram (0-50X)



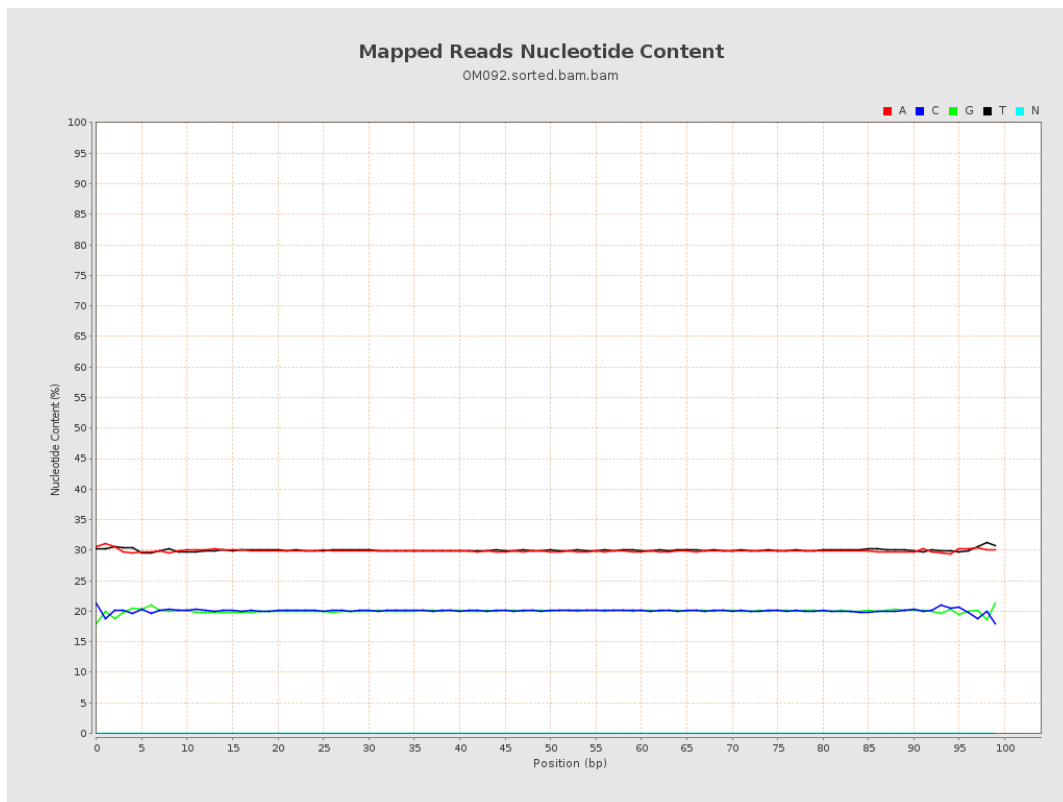
6. Results : Genome Fraction Coverage



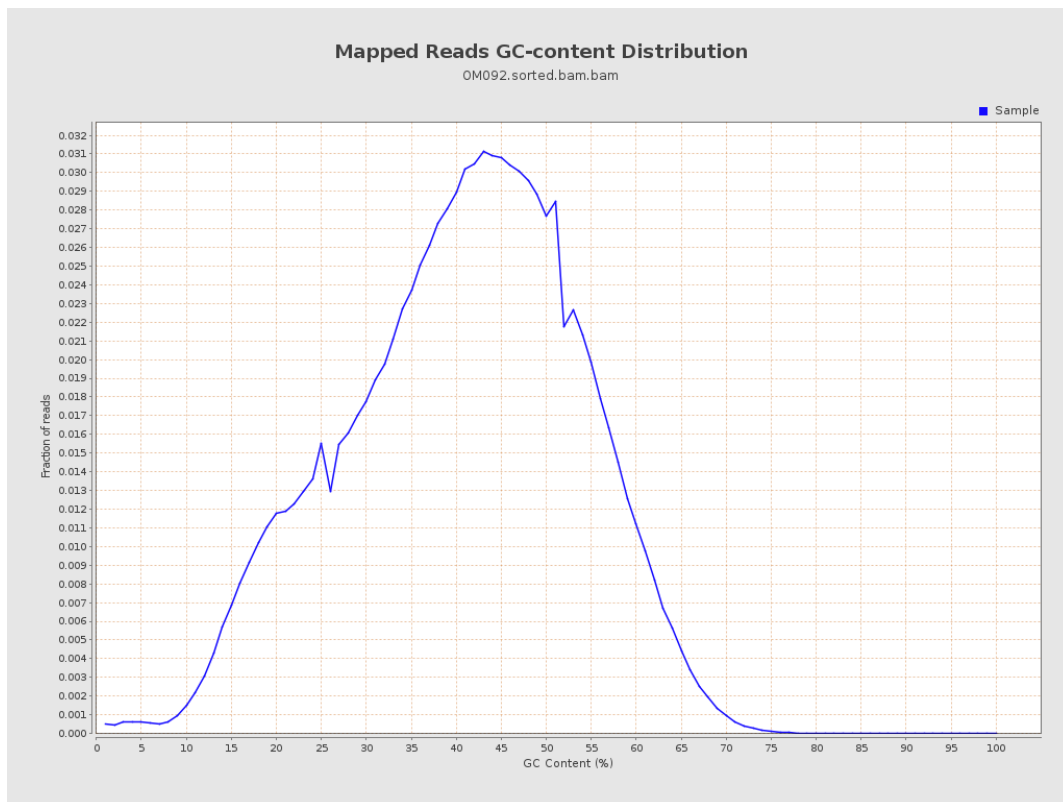
7. Results : Duplication Rate Histogram



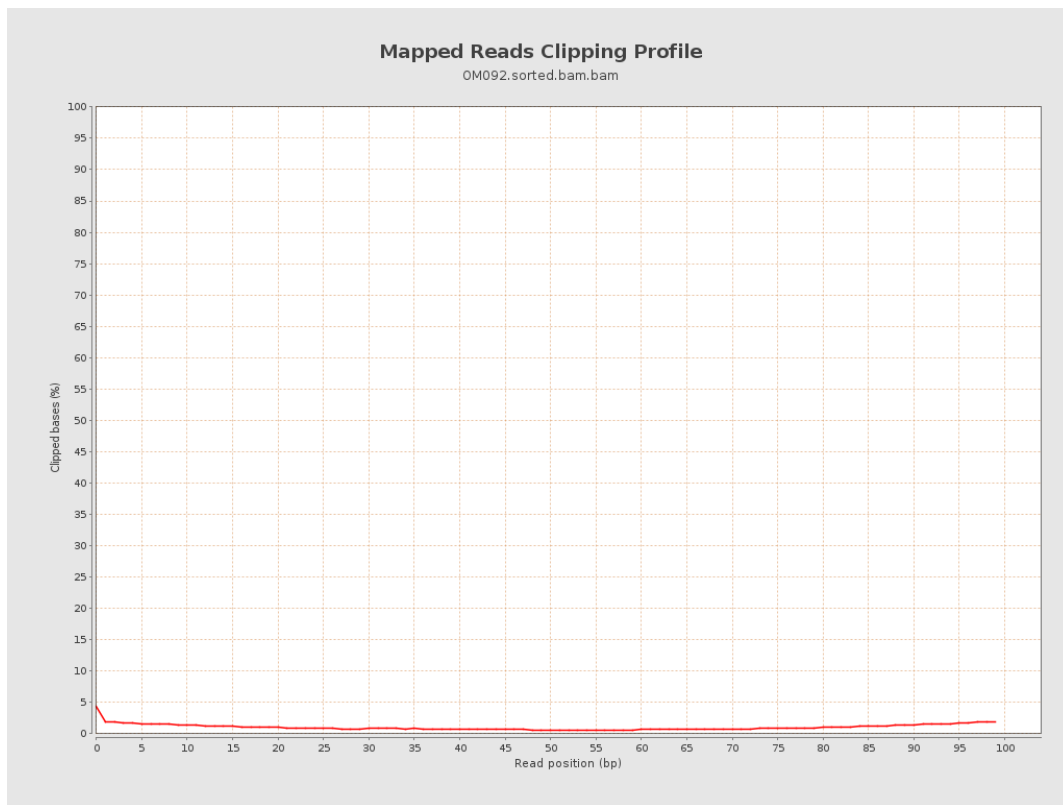
8. Results : Mapped Reads Nucleotide Content



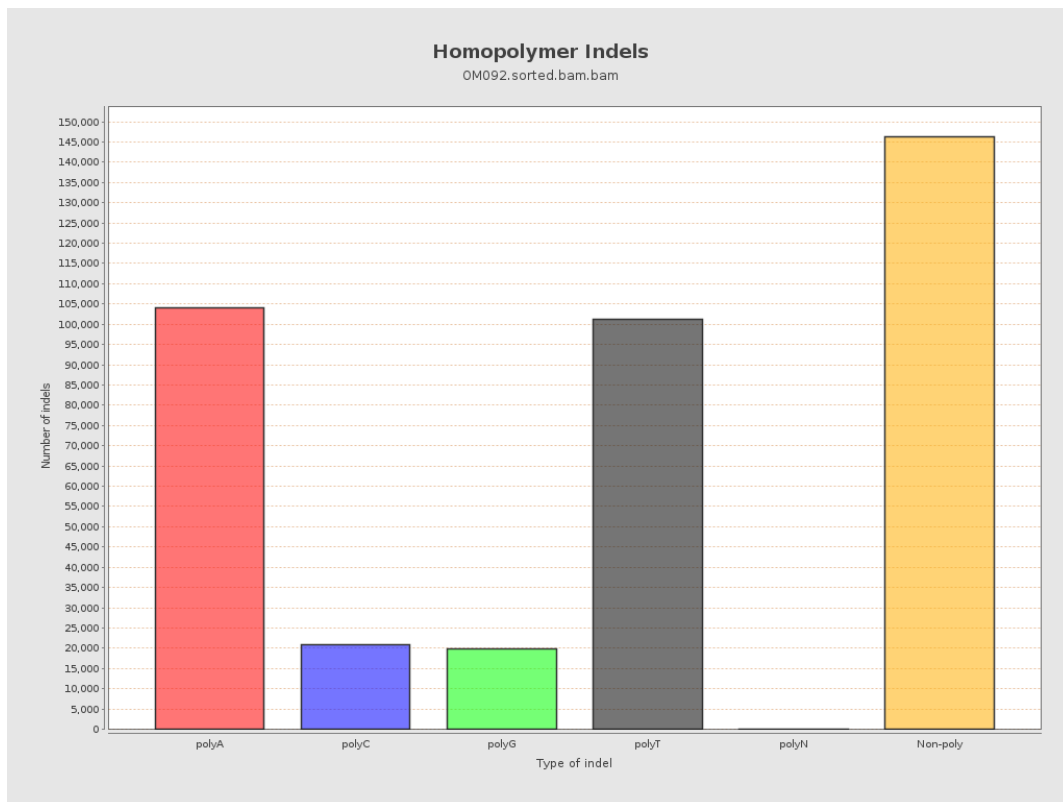
9. Results : Mapped Reads GC-content Distribution



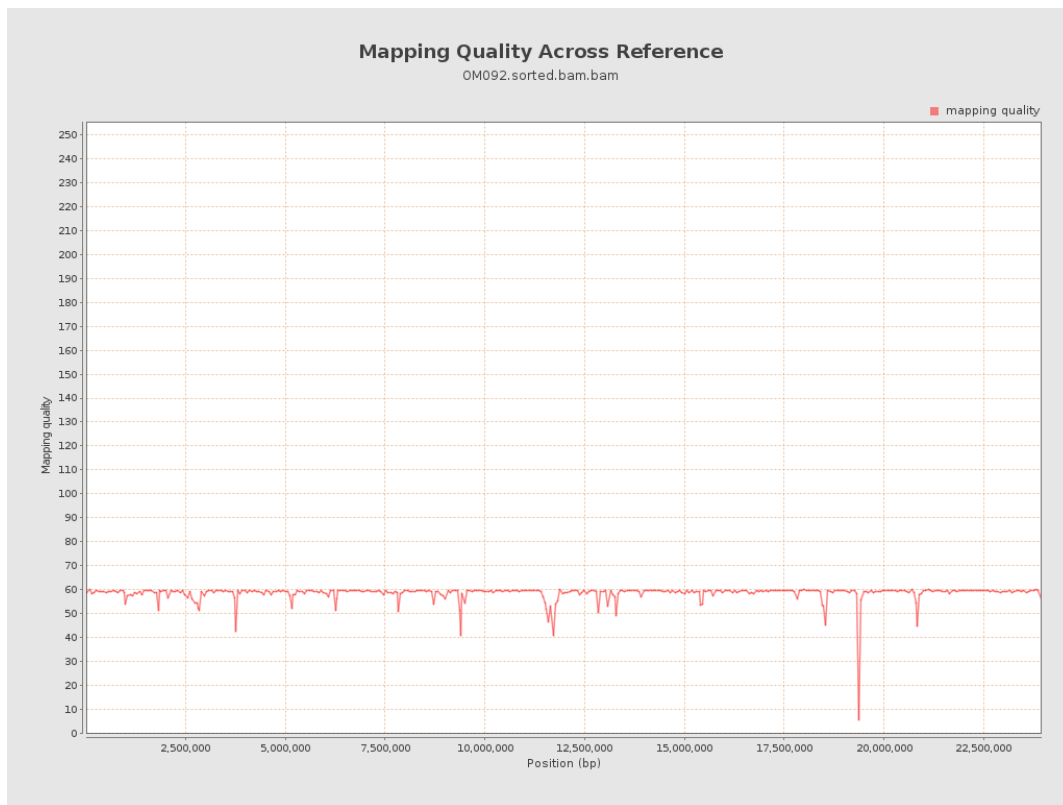
10. Results : Mapped Reads Clipping Profile



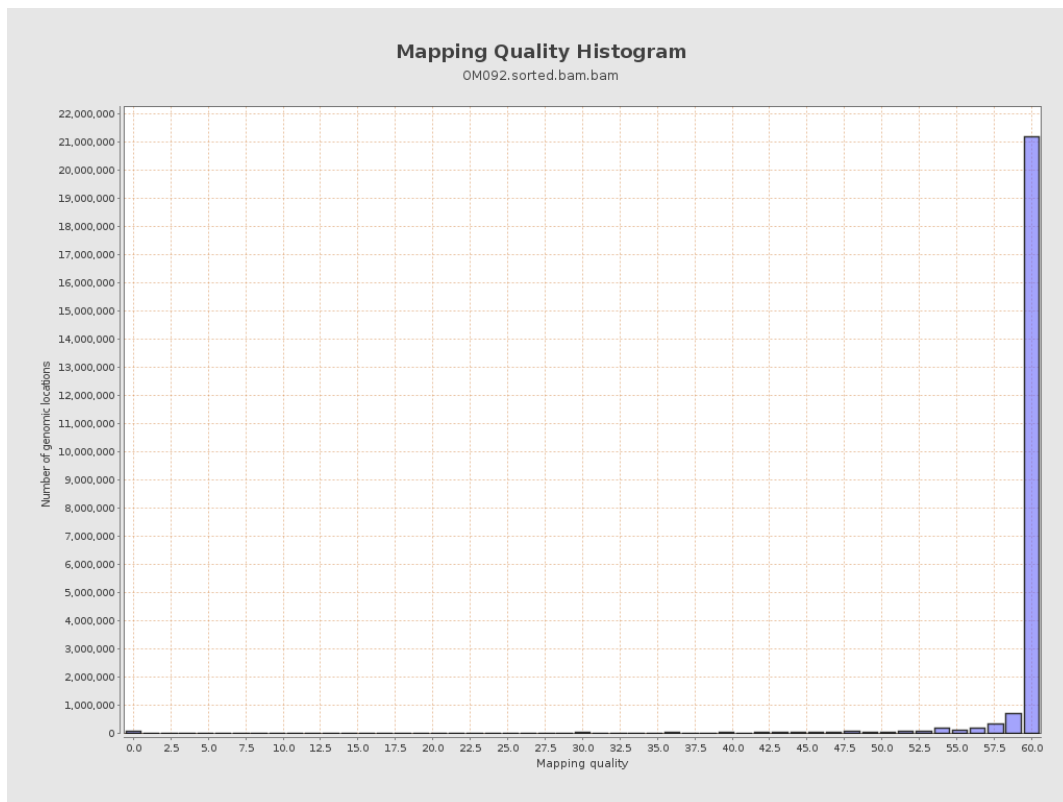
11. Results : Homopolymer Indels



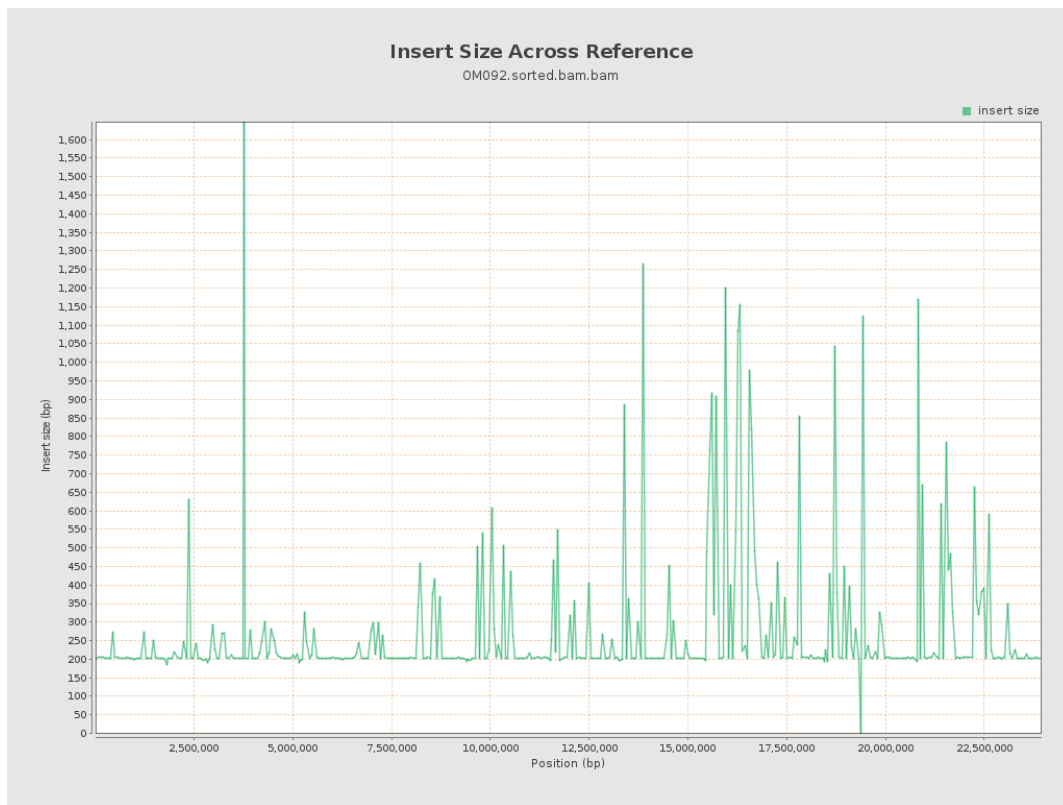
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

