

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 14:15:44

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/OM314.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/OM314-BiooBarcode40_CTCAGA_R1.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/OM314-BiooBarcode40_CTCAGA_R2.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 14:15:44 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/OM314.sorted.bam.

	bam
--	-----

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	1,302,907
Mapped reads	650,350 / 49.92%
Unmapped reads	652,557 / 50.08%
Mapped paired reads	650,350 / 49.92%
Mapped reads, first in pair	326,440 / 25.05%
Mapped reads, second in pair	323,910 / 24.86%
Mapped reads, both in pair	635,691 / 48.79%
Mapped reads, singletons	14,659 / 1.13%
Read min/max/mean length	30 / 100 / 99.92
Duplicated reads (estimated)	38,388 / 2.95%
Duplication rate	4.56%
Clipped reads	73,568 / 5.65%

2.2. ACGT Content

Number/percentage of A's	18,758,244 / 29.89%
Number/percentage of C's	12,626,275 / 20.12%
Number/percentage of T's	18,819,403 / 29.98%
Number/percentage of G's	12,563,989 / 20.02%
Number/percentage of N's	5,260 / 0.01%
GC Percentage	40.13%

2.3. Coverage

Mean	2.6226
Standard Deviation	3.2606

2.4. Mapping Quality

Mean Mapping Quality	58.57
----------------------	-------

2.5. Insert size

Mean	983.05
Standard Deviation	30,964.05
P25/Median/P75	330 / 342 / 351

2.6. Mismatches and indels

General error rate	1.58%
Mismatches	942,205
Insertions	20,566
Mapped reads with at least one insertion	2.99%
Deletions	23,030
Mapped reads with at least one deletion	3.34%
Homopolymer indels	63.8%

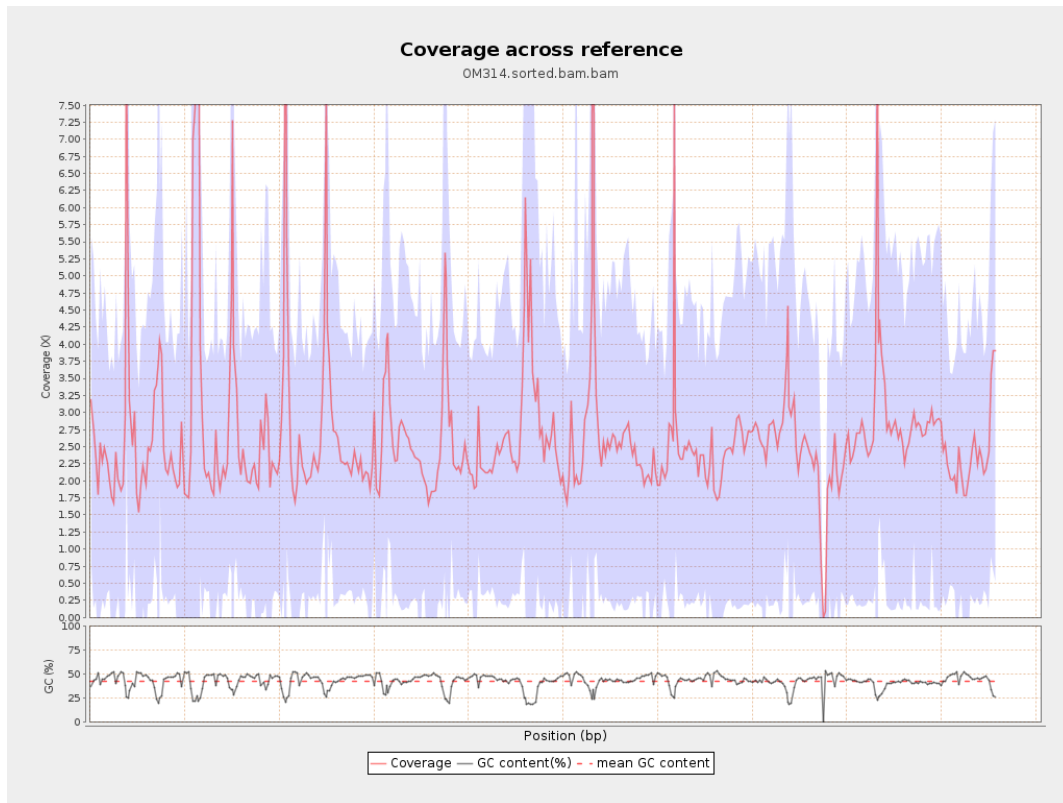
2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

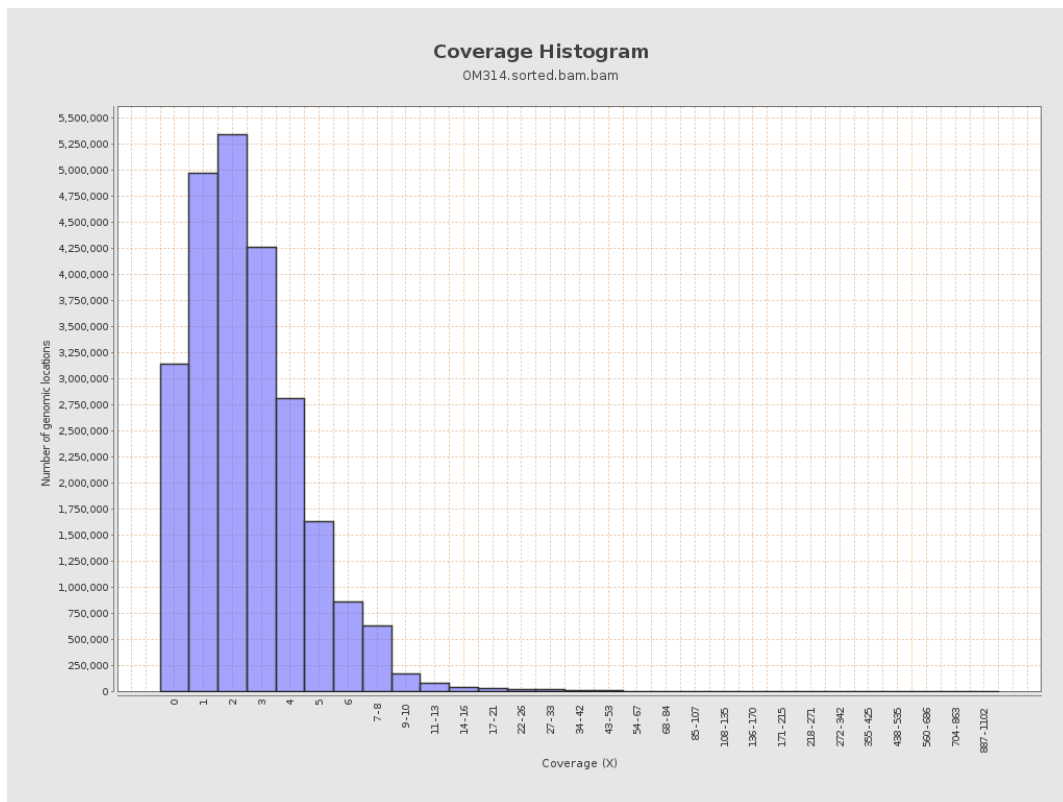
gi 1074120478 emb LT615256.1	977217	2334161	2.3886	2.0774
gi 1074120682 emb LT615257.1	860454	2283655	2.654	2.3837
gi 1074120865 emb LT615258.1	989719	2953087	2.9838	4.4147
gi 1074121086 emb LT615259.1	935450	2763748	2.9545	5.1586
gi 1074121301 emb LT615260.1	1432239	3984719	2.7822	3.8849
gi 1074121615 emb LT615261.1	1080962	2902771	2.6854	3.0639
gi 1074121871 emb LT615262.1	1545099	3792767	2.4547	1.8104
gi 1074122235 emb LT615263.1	1585108	4069119	2.5671	2.2092
gi 1074122590 emb LT615264.1	2122358	5336555	2.5144	2.0114
gi 1074123050 emb LT615265.1	1754192	4895908	2.791	3.0555
gi 1074123421 emb LT615	2150147	5658336	2.6316	6.568

266.1				
gi 107412389 8 emb LT615 267.1	3031036	7803757	2.5746	1.9415
gi 107412458 8 emb LT615 268.1	2359348	5673851	2.4048	2.6206
gi 107412506 5 emb LT615 269.1	3135668	8383008	2.6734	1.869

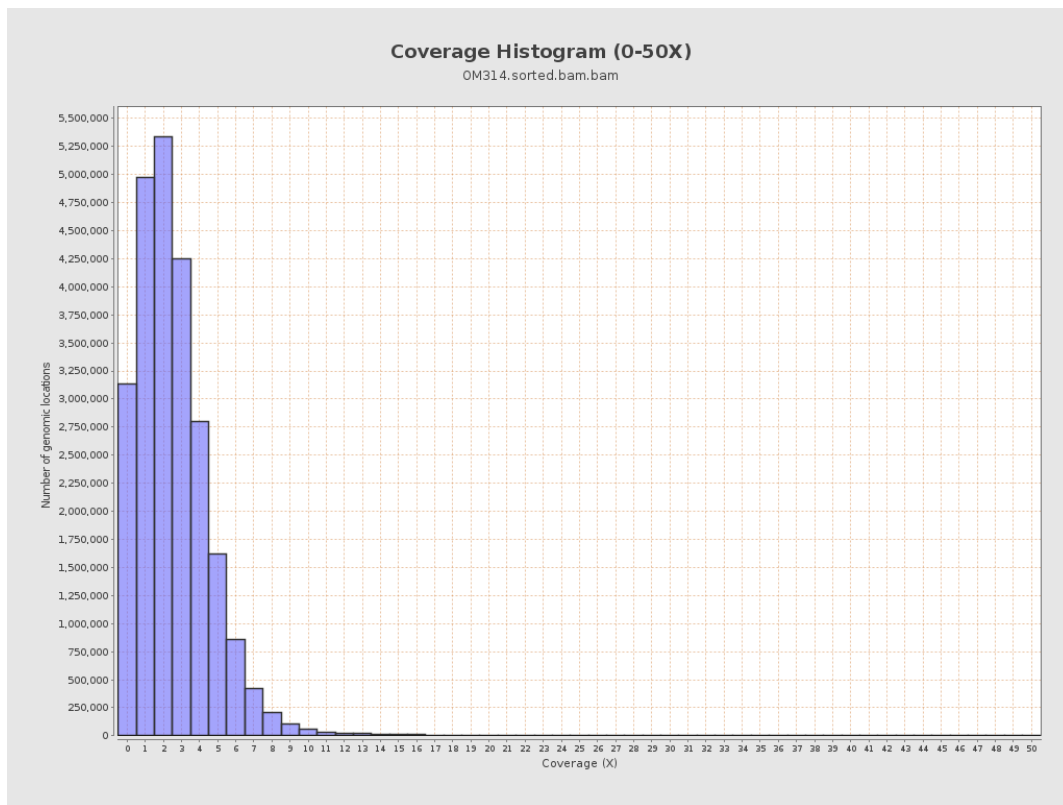
3. Results : Coverage across reference



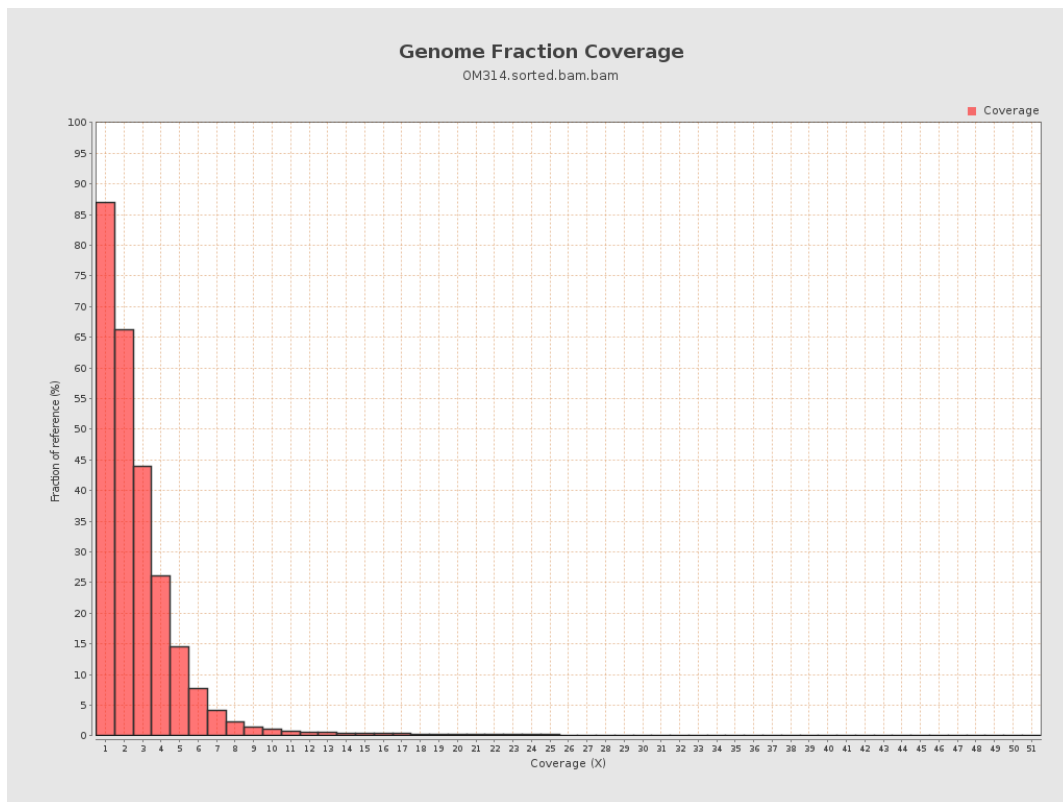
4. Results : Coverage Histogram



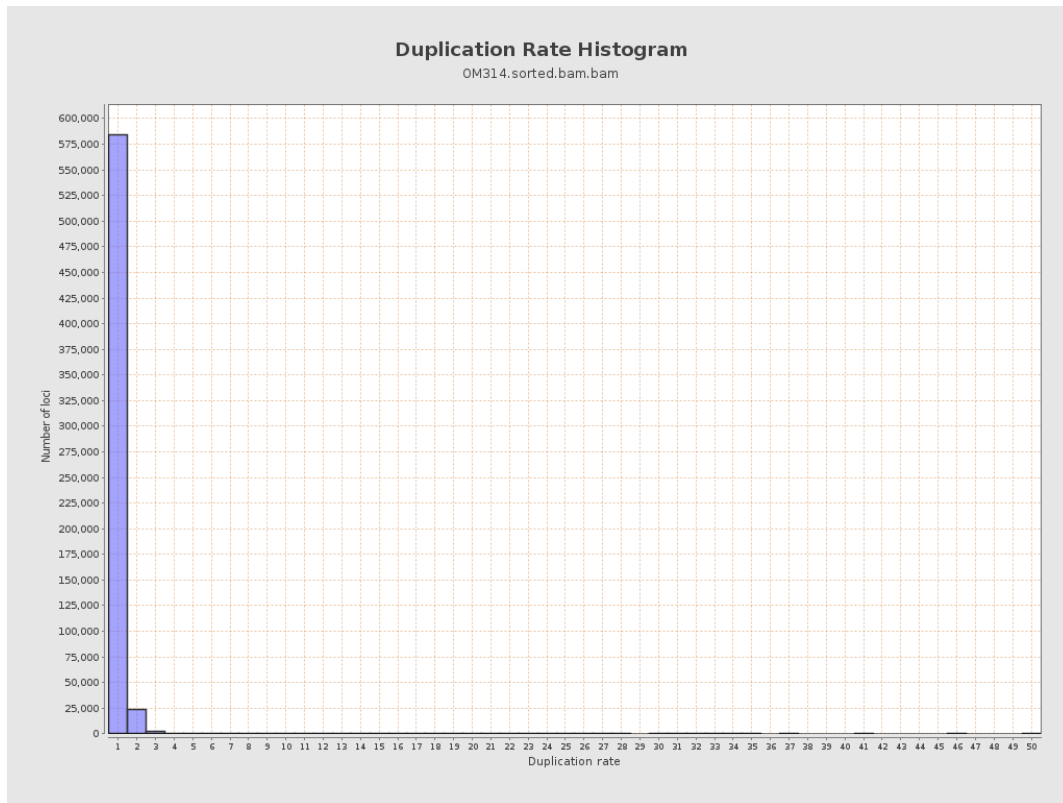
5. Results : Coverage Histogram (0-50X)



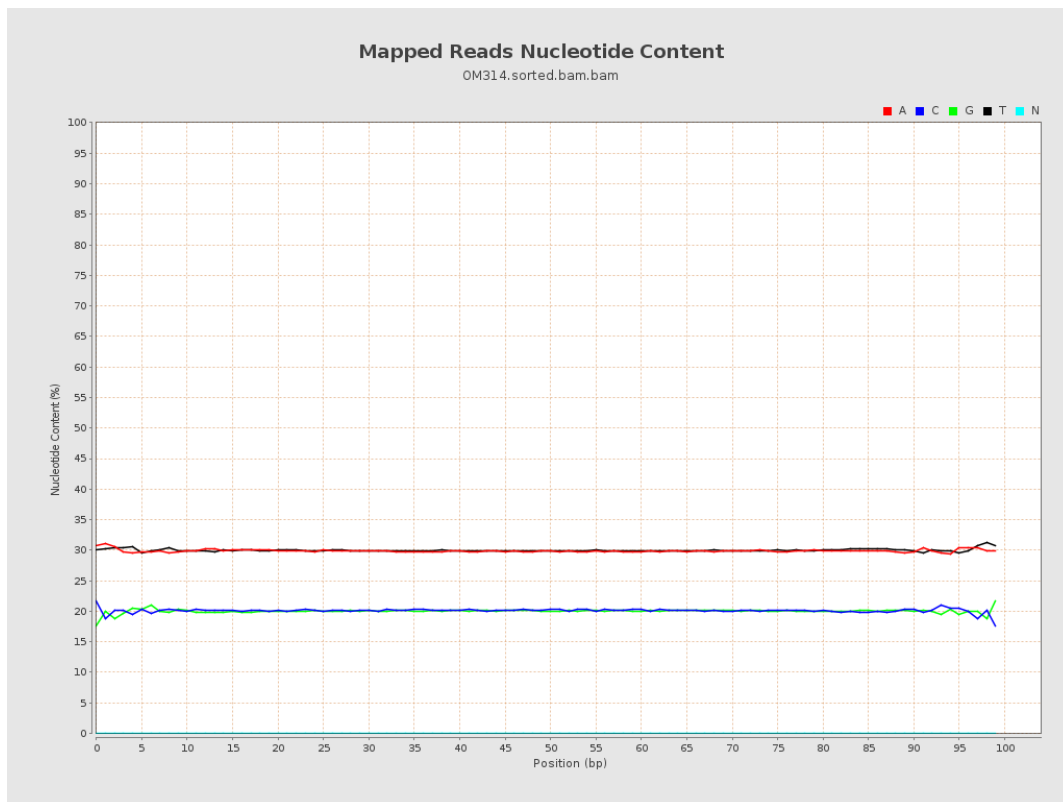
6. Results : Genome Fraction Coverage



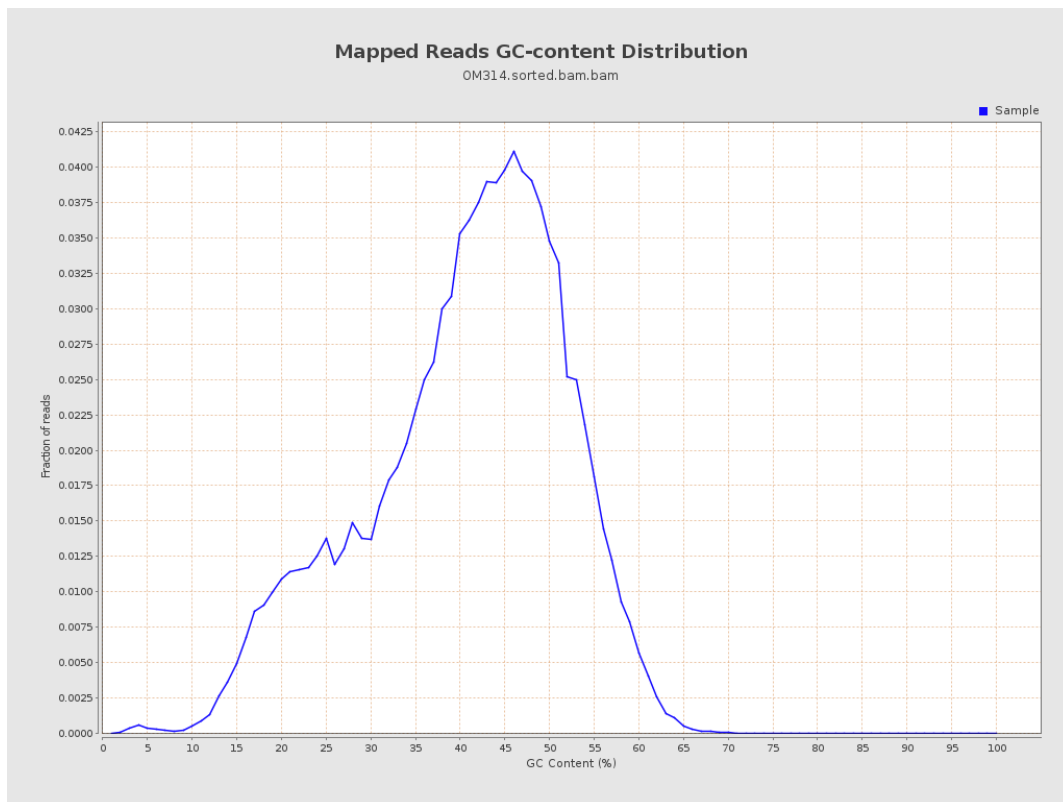
7. Results : Duplication Rate Histogram



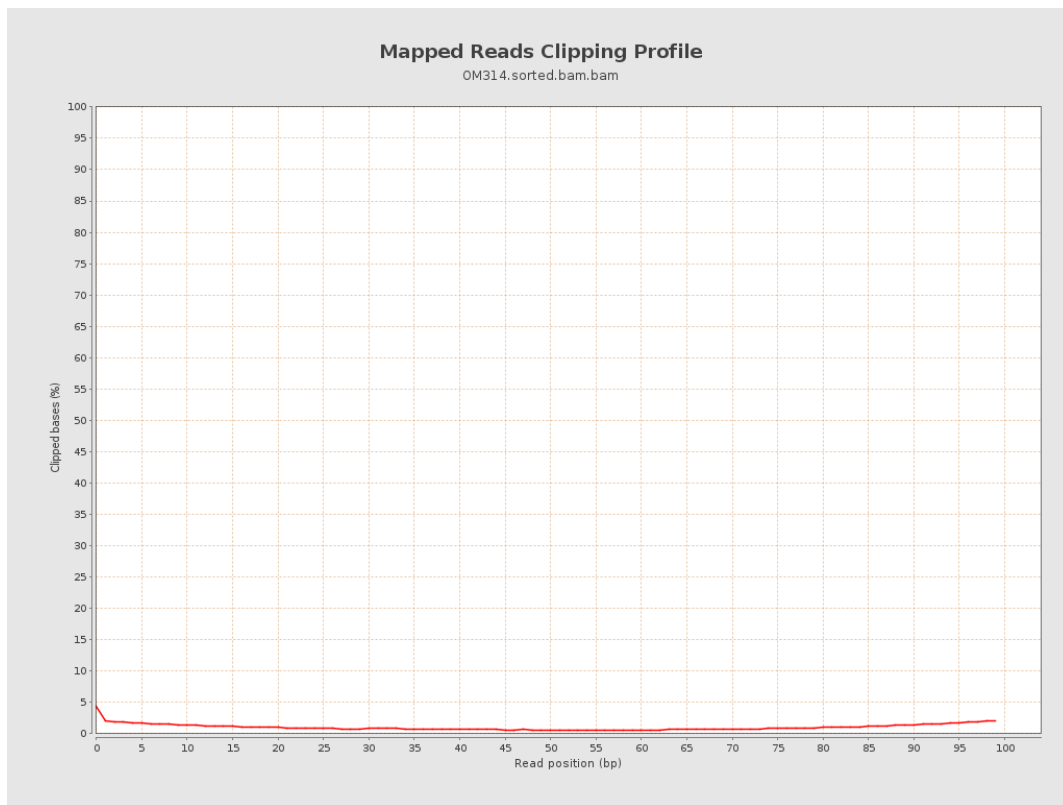
8. Results : Mapped Reads Nucleotide Content



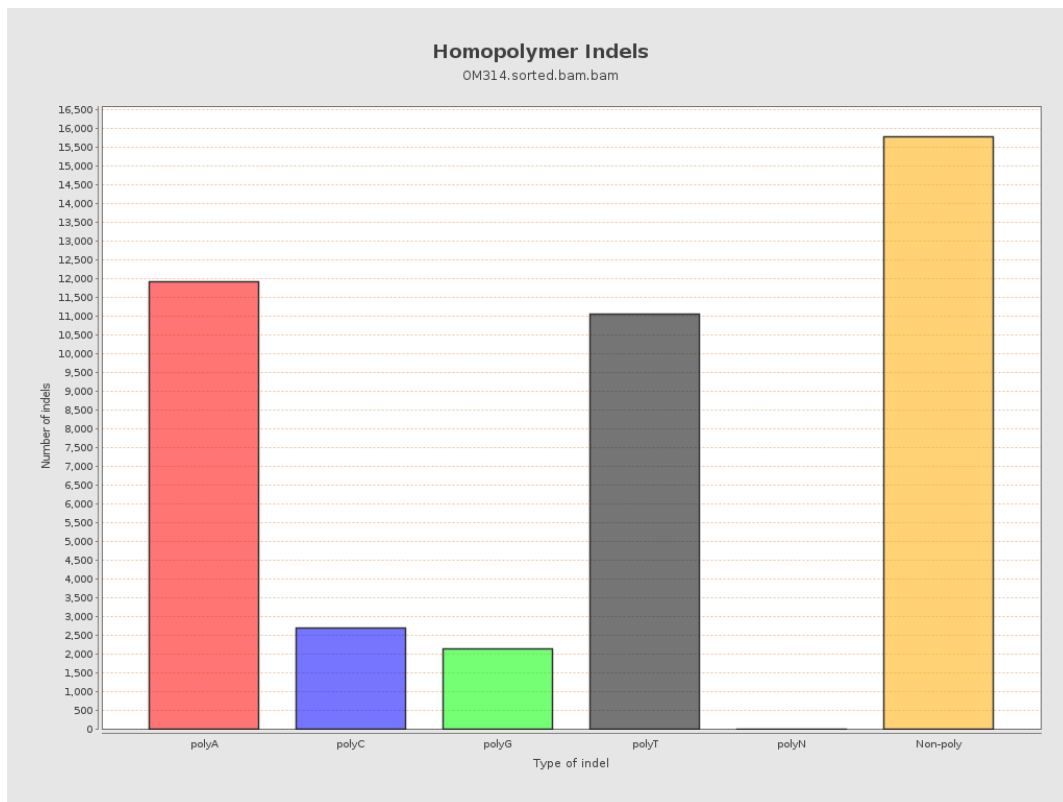
9. Results : Mapped Reads GC-content Distribution



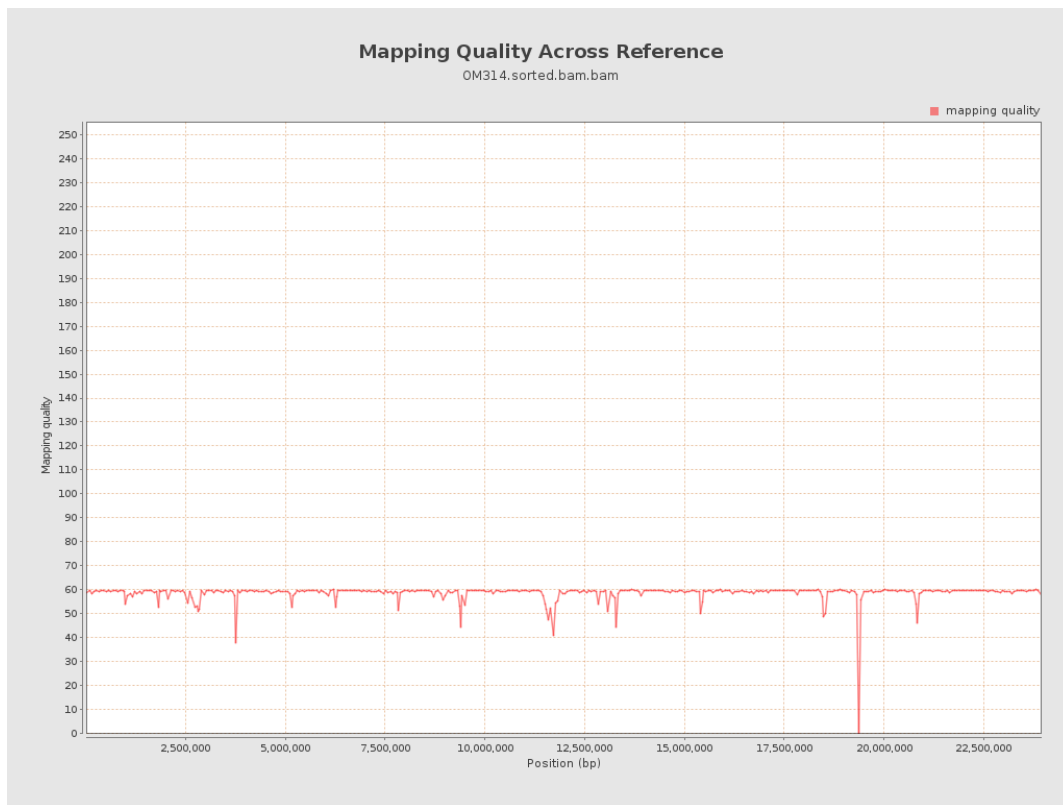
10. Results : Mapped Reads Clipping Profile



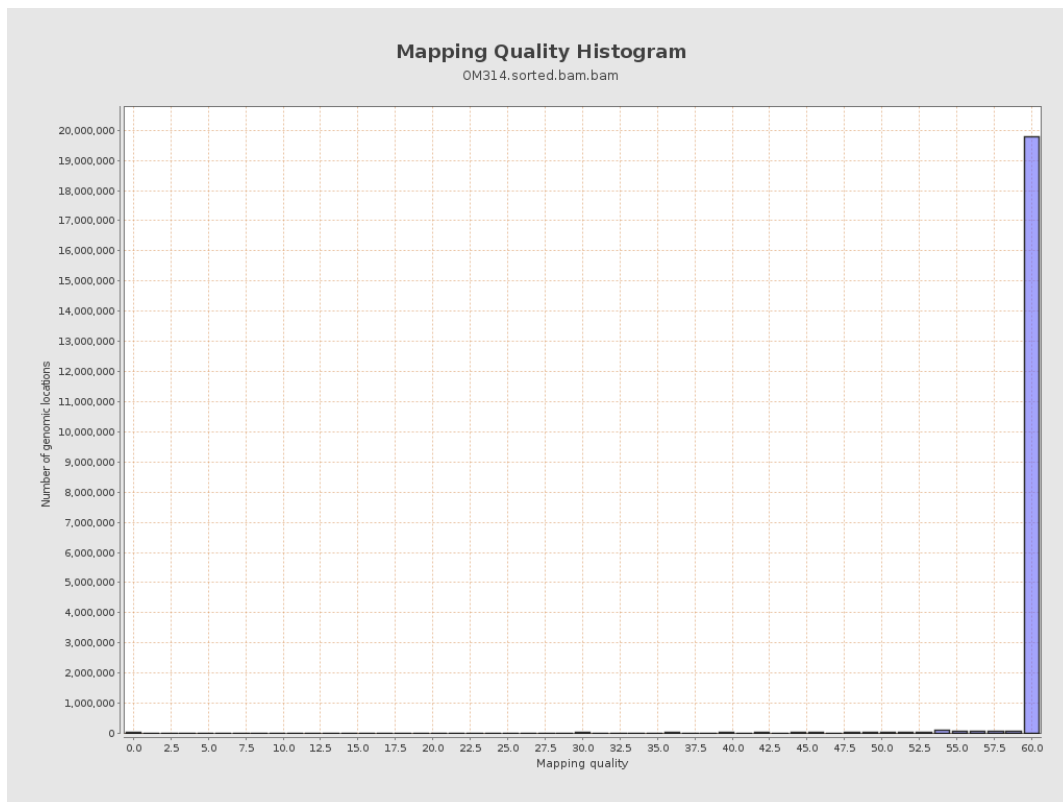
11. Results : Homopolymer Indels



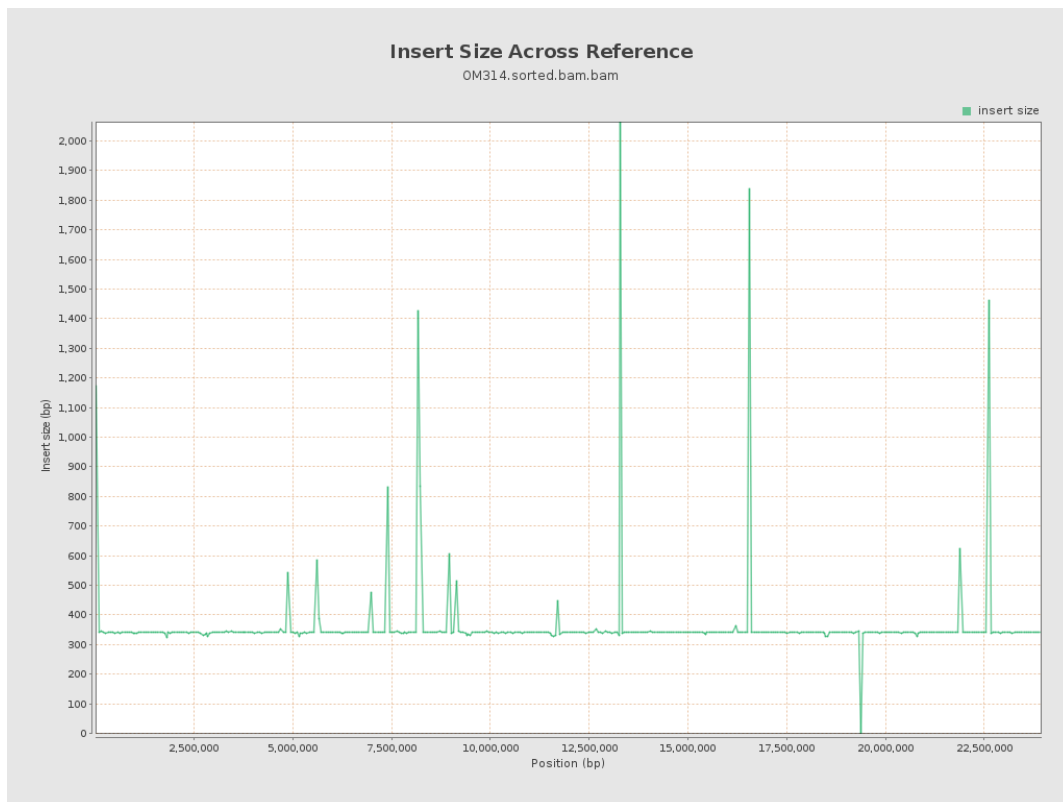
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

