

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 14:16:43

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/OM317.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/OM317-BiooBarcode41_GACGAC_R2.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/OM317-BiooBarcode41_GACGAC_R1.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 14:16:43 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/OM317.sorted.bam.

	bam
--	-----

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	7,543,203
Mapped reads	4,879,784 / 64.69%
Unmapped reads	2,663,419 / 35.31%
Mapped paired reads	4,879,784 / 64.69%
Mapped reads, first in pair	2,434,895 / 32.28%
Mapped reads, second in pair	2,444,889 / 32.41%
Mapped reads, both in pair	4,803,860 / 63.68%
Mapped reads, singletons	75,924 / 1.01%
Read min/max/mean length	30 / 100 / 99.89
Duplicated reads (estimated)	826,712 / 10.96%
Duplication rate	13.81%
Clipped reads	488,739 / 6.48%

2.2. ACGT Content

Number/percentage of A's	140,967,048 / 29.78%
Number/percentage of C's	95,575,072 / 20.19%
Number/percentage of T's	141,292,474 / 29.85%
Number/percentage of G's	95,468,902 / 20.17%
Number/percentage of N's	41,452 / 0.01%
GC Percentage	40.36%

2.3. Coverage

Mean	19.7752
Standard Deviation	17.0088

2.4. Mapping Quality

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

2.5. Insert size

Mean	706.64
Standard Deviation	25,566.16
P25/Median/P75	219 / 228 / 238

2.6. Mismatches and indels

General error rate	1.37%
Mismatches	6,163,630
Insertions	149,548
Mapped reads with at least one insertion	2.9%
Deletions	166,448
Mapped reads with at least one deletion	3.23%
Homopolymer indels	64.6%

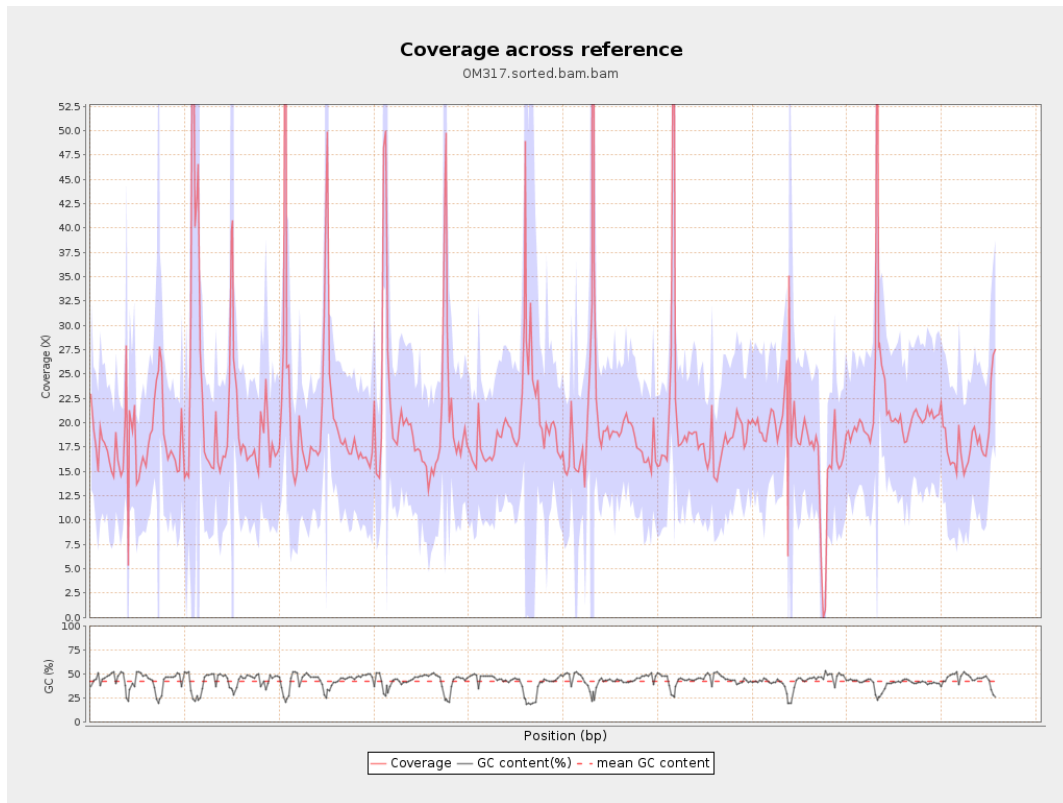
2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

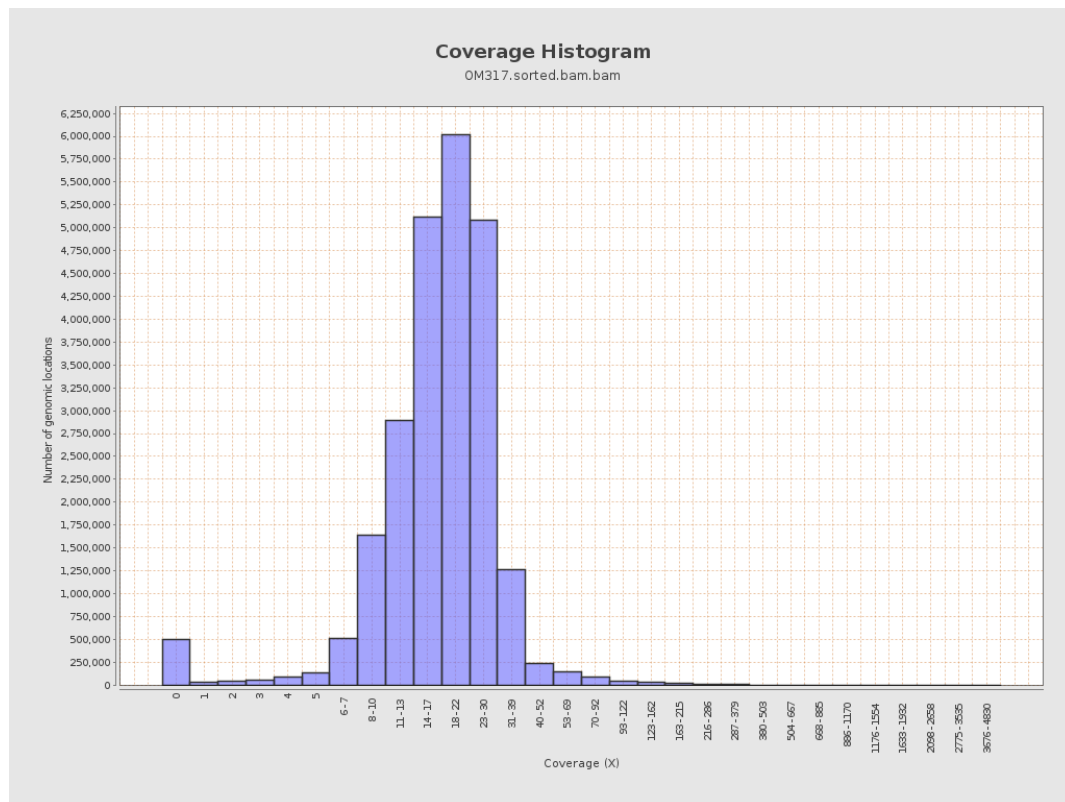
gi 1074120478 emb LT615256.1	977217	17199504	17.6005	7.3909
gi 1074120682 emb LT615257.1	860454	15625070	18.1591	12.6194
gi 1074120865 emb LT615258.1	989719	22956350	23.1948	27.4867
gi 1074121086 emb LT615259.1	935450	19576166	20.927	24.7973
gi 1074121301 emb LT615260.1	1432239	30284368	21.1448	20.5411
gi 1074121615 emb LT615261.1	1080962	21124527	19.5423	13.7342
gi 1074121871 emb LT615262.1	1545099	31154830	20.1636	11.1125
gi 1074122235 emb LT615263.1	1585108	31231134	19.7028	13.3172
gi 1074122590 emb LT615264.1	2122358	40882830	19.2629	10.5348
gi 1074123050 emb LT615265.1	1754192	34739001	19.8034	17.6385
gi 1074123421 emb LT615	2150147	45138493	20.9932	32.9173

266.1				
gi 107412389 8 emb LT615 267.1	3031036	57844560	19.0841	8.4036
gi 107412458 8 emb LT615 268.1	2359348	43737964	18.5382	16.1961
gi 107412506 5 emb LT615 269.1	3135668	62299781	19.8681	7.1975

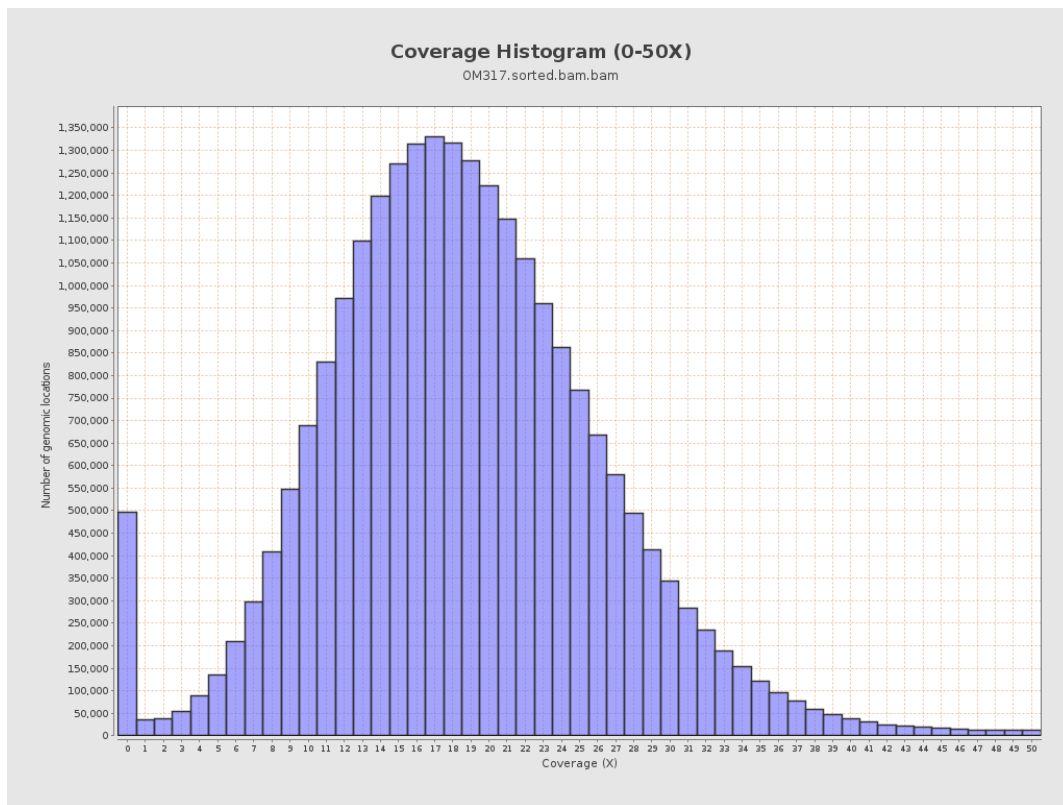
3. Results : Coverage across reference



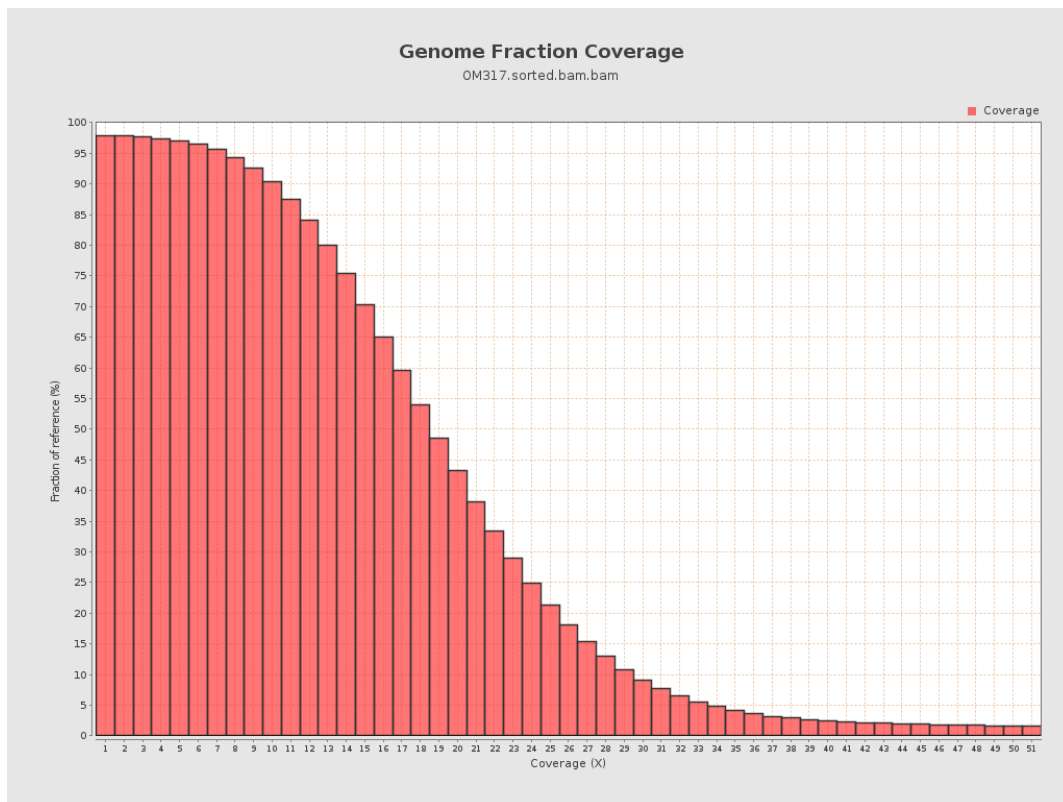
4. Results : Coverage Histogram



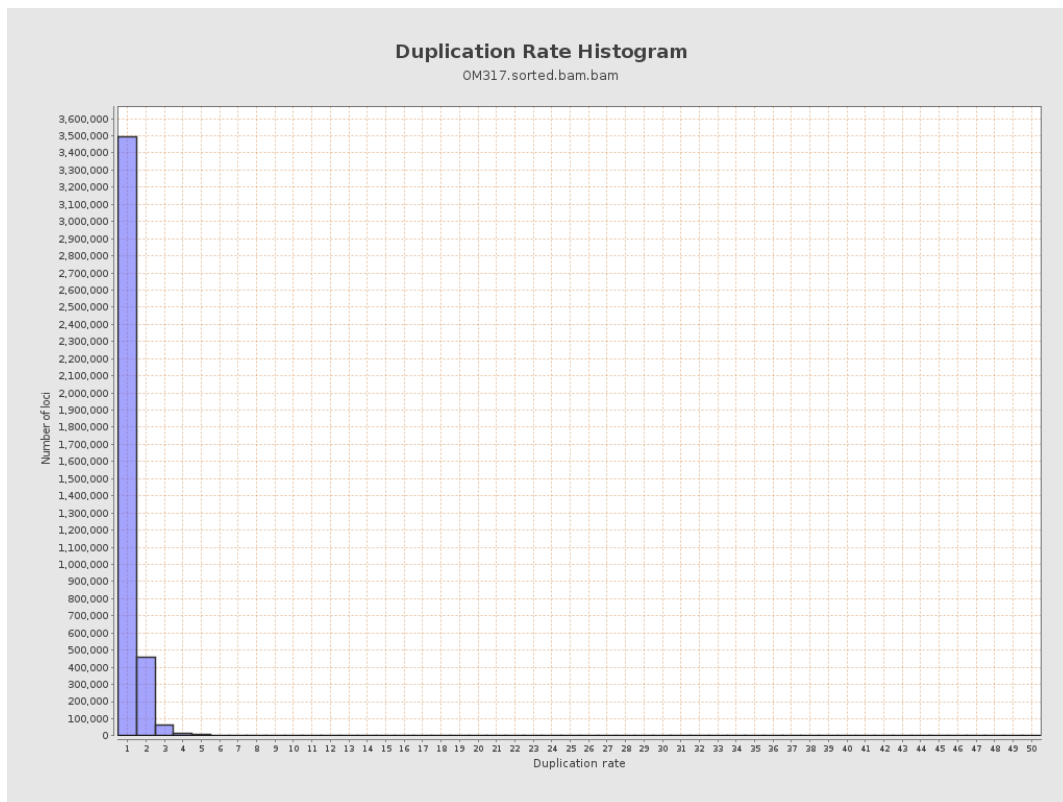
5. Results : Coverage Histogram (0-50X)



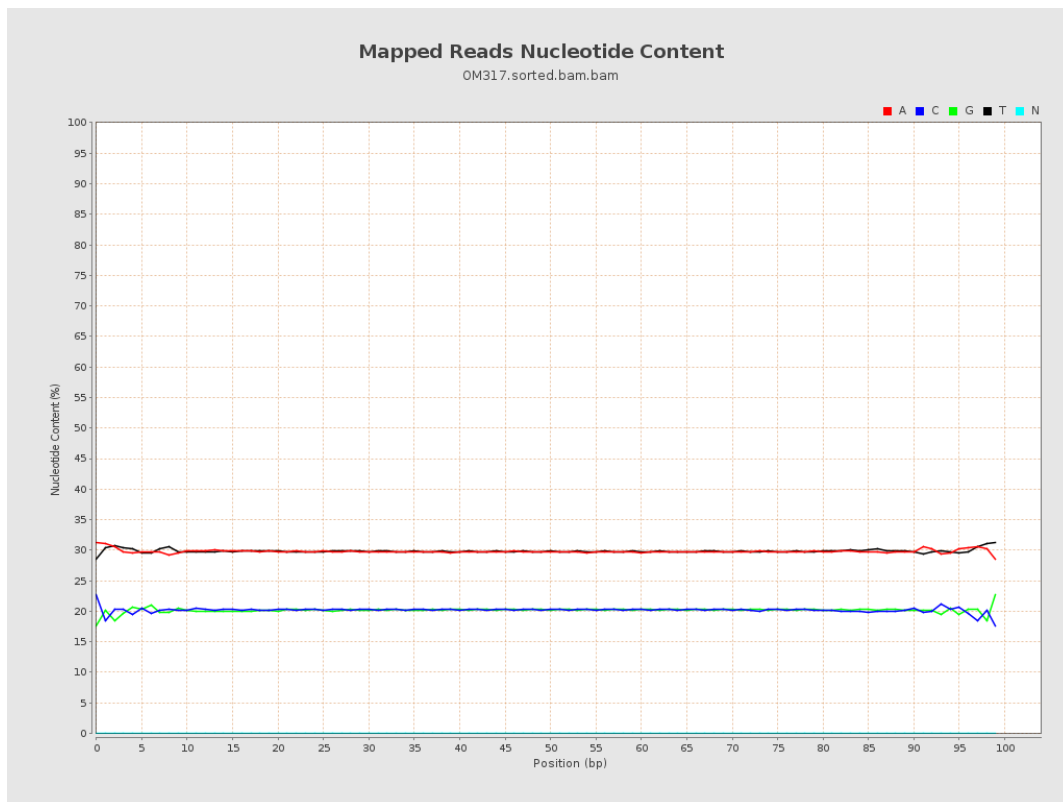
6. Results : Genome Fraction Coverage



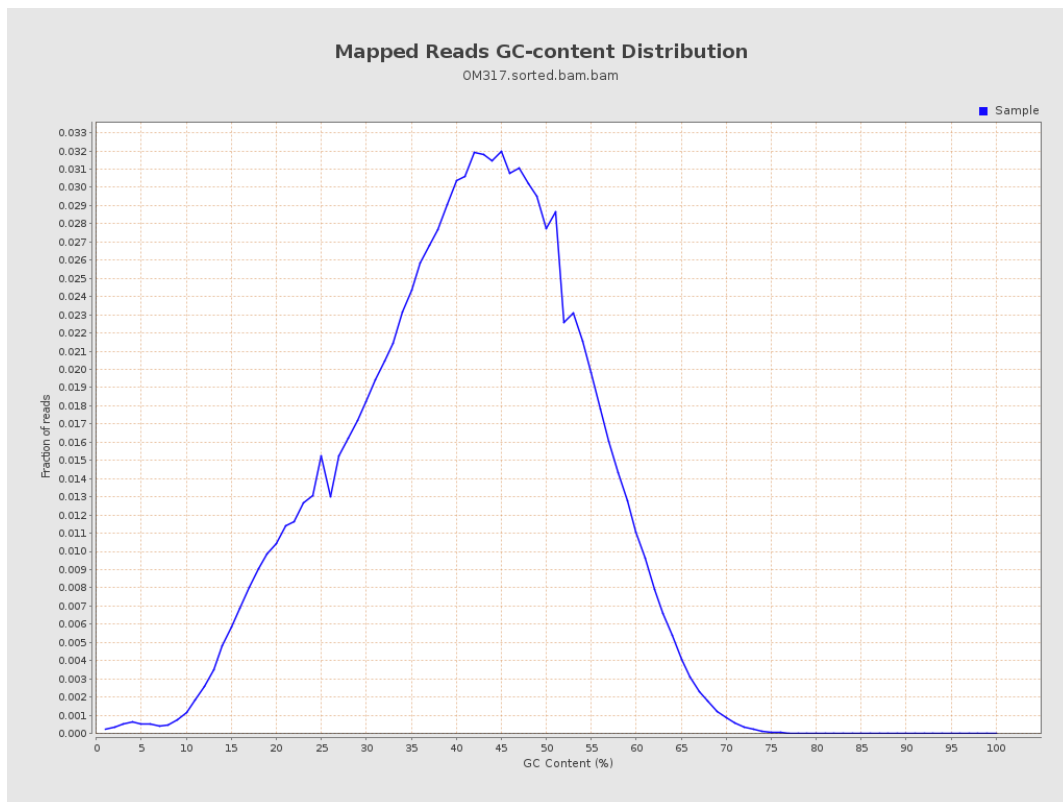
7. Results : Duplication Rate Histogram



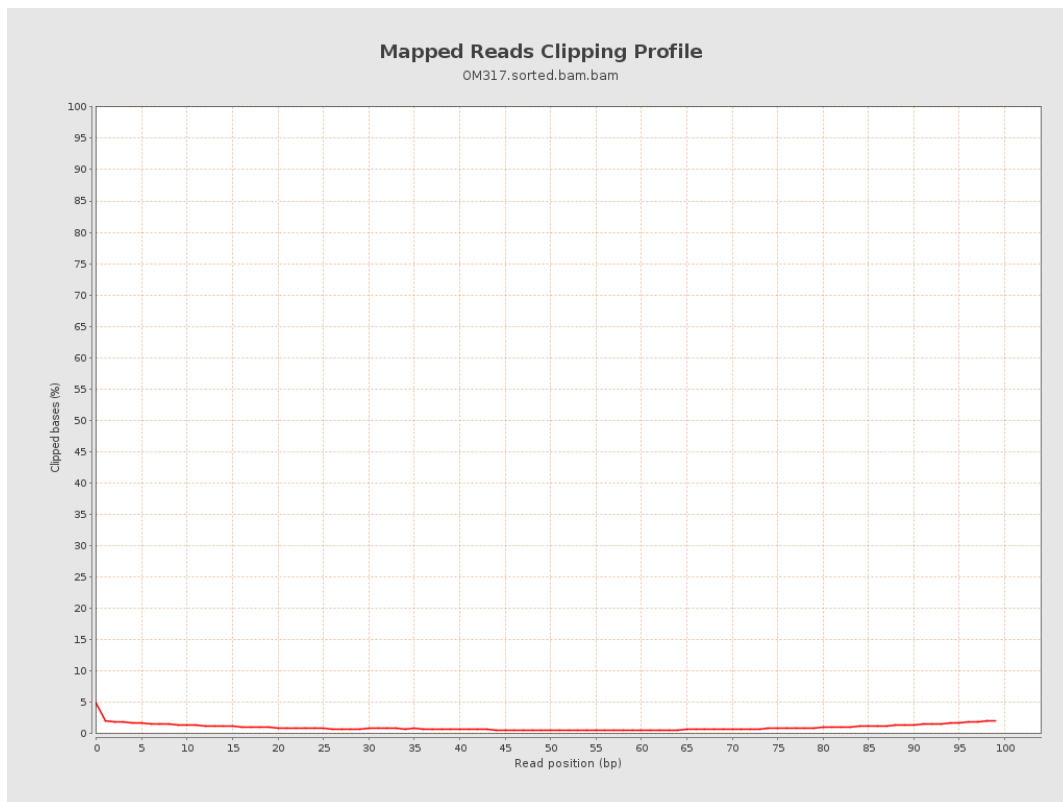
8. Results : Mapped Reads Nucleotide Content



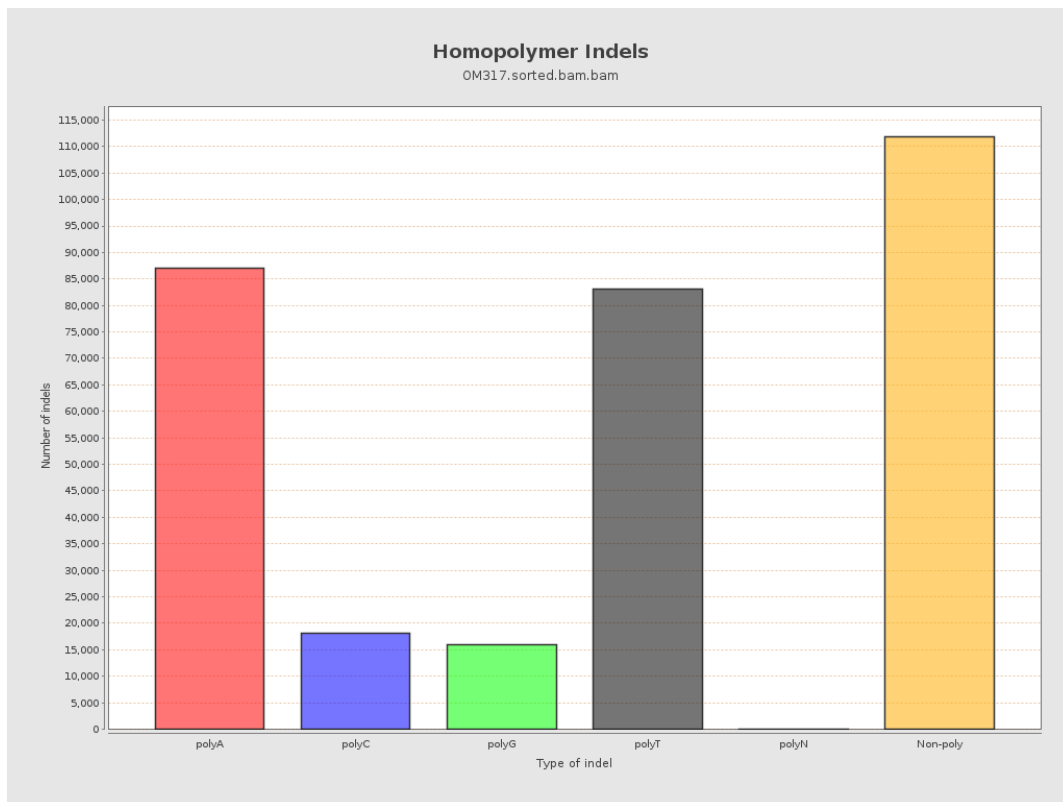
9. Results : Mapped Reads GC-content Distribution



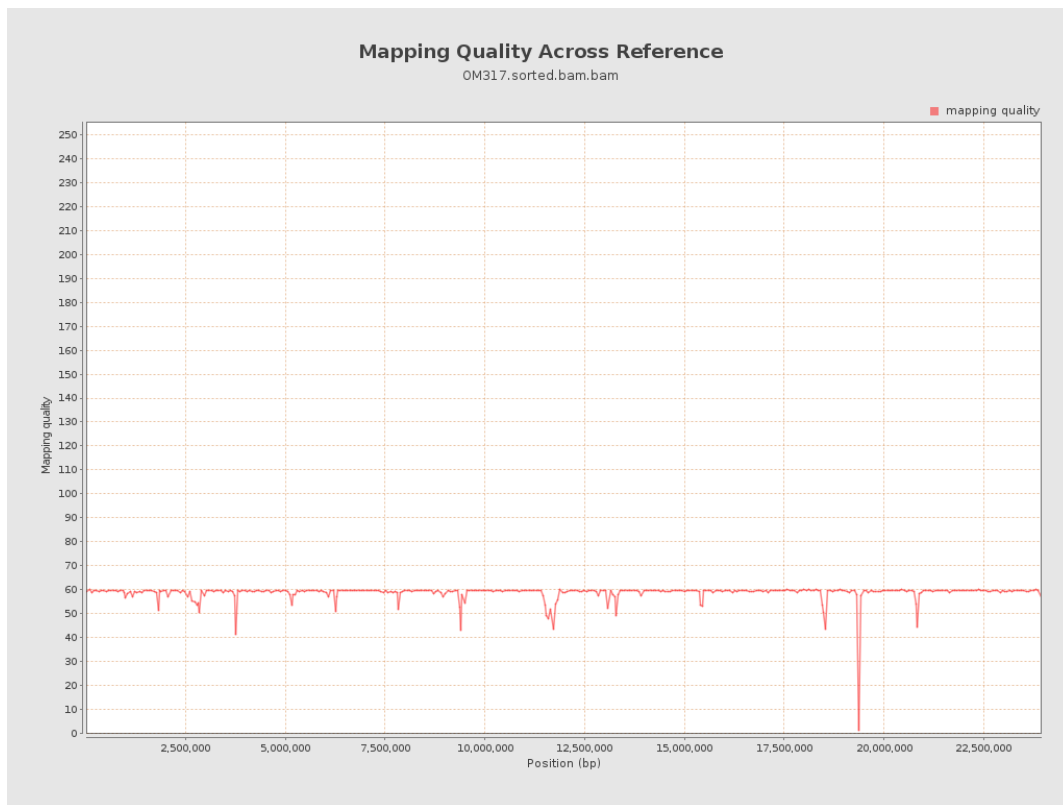
10. Results : Mapped Reads Clipping Profile



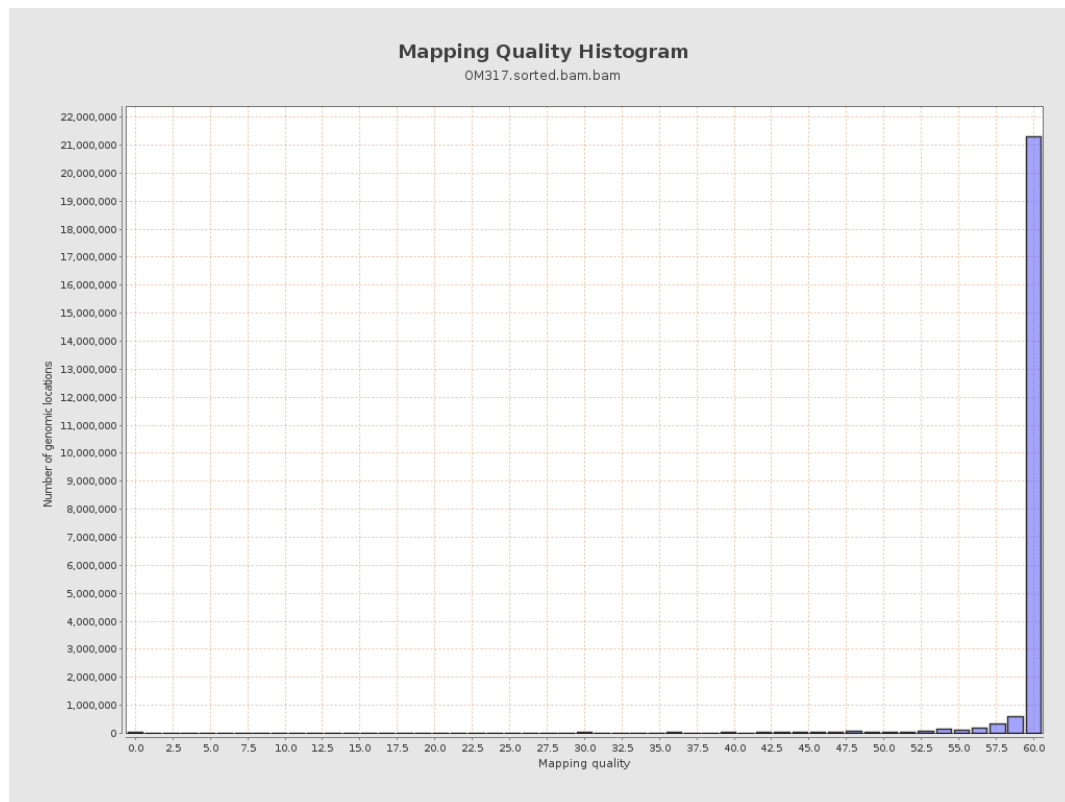
11. Results : Homopolymer Indels



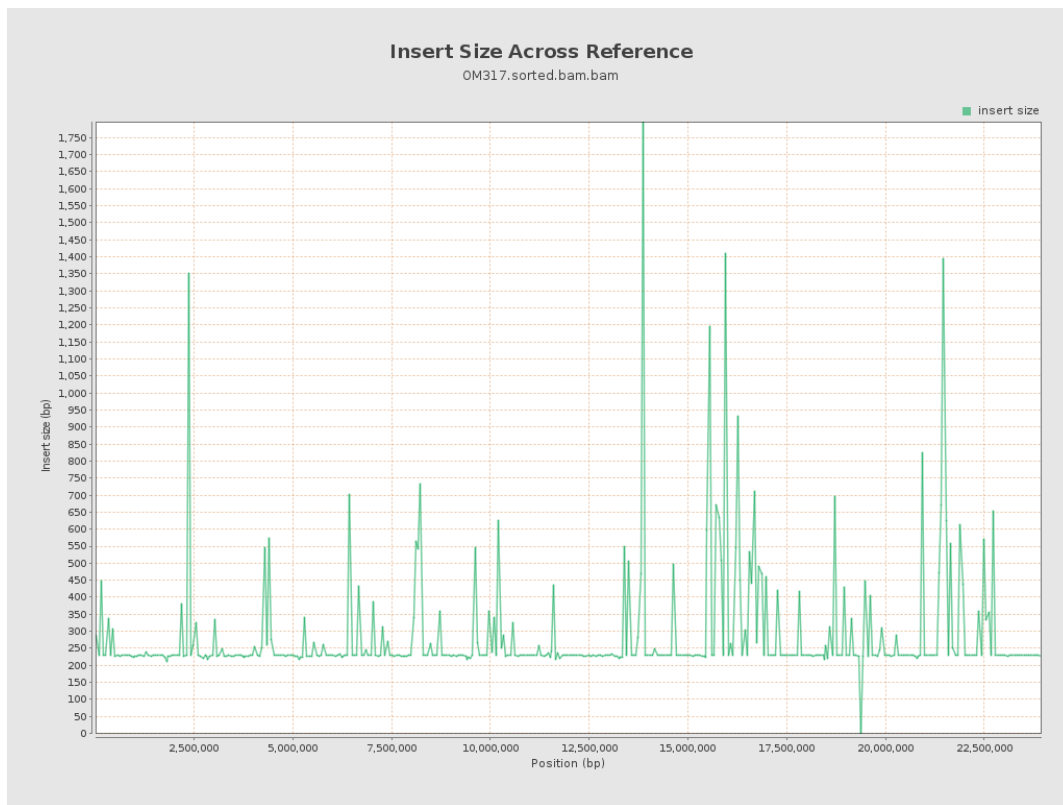
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

