

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 14:22:42

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/OM322.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/OM322-BiooBarcode42_TAATCG_R1.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/OM322-BiooBarcode42_TAATCG_R2.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 14:22:41 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/OM322.sorted.bam.

	bam
--	-----

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	53,547,333
Mapped reads	18,790,181 / 35.09%
Unmapped reads	34,757,152 / 64.91%
Mapped paired reads	18,790,181 / 35.09%
Mapped reads, first in pair	9,430,476 / 17.61%
Mapped reads, second in pair	9,359,705 / 17.48%
Mapped reads, both in pair	18,215,161 / 34.02%
Mapped reads, singletons	575,020 / 1.07%
Read min/max/mean length	30 / 100 / 99.94
Duplicated reads (estimated)	7,791,172 / 14.55%
Duplication rate	37.57%
Clipped reads	2,452,072 / 4.58%

2.2. ACGT Content

Number/percentage of A's	546,798,146 / 30.38%
Number/percentage of C's	352,527,949 / 19.59%
Number/percentage of T's	549,022,589 / 30.5%
Number/percentage of G's	351,619,414 / 19.53%
Number/percentage of N's	147,808 / 0.01%
GC Percentage	39.12%

2.3. Coverage

Mean	75.2112
Standard Deviation	86.0414

2.4. Mapping Quality

Mean Mapping Quality	58.47
----------------------	-------

2.5. Insert size

Mean	953.81
Standard Deviation	29,710.24
P25/Median/P75	340 / 354 / 364

2.6. Mismatches and indels

General error rate	1.74%
Mismatches	29,938,954
Insertions	646,919
Mapped reads with at least one insertion	3.25%
Deletions	725,998
Mapped reads with at least one deletion	3.62%
Homopolymer indels	61.52%

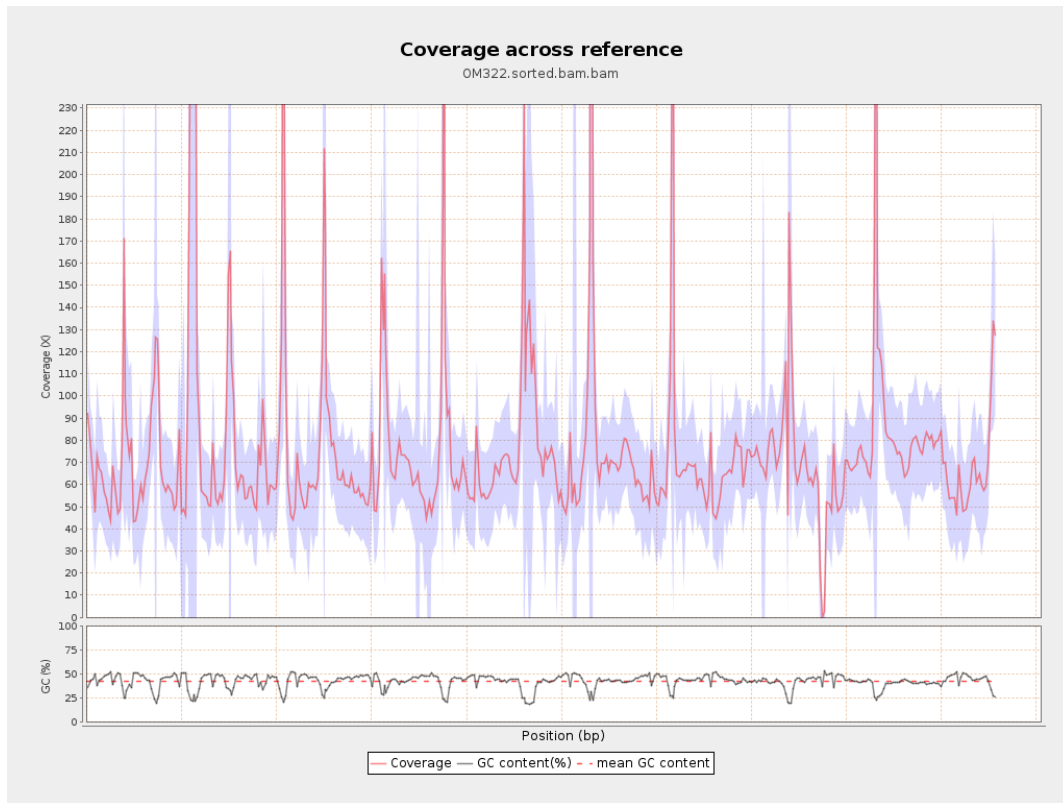
2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

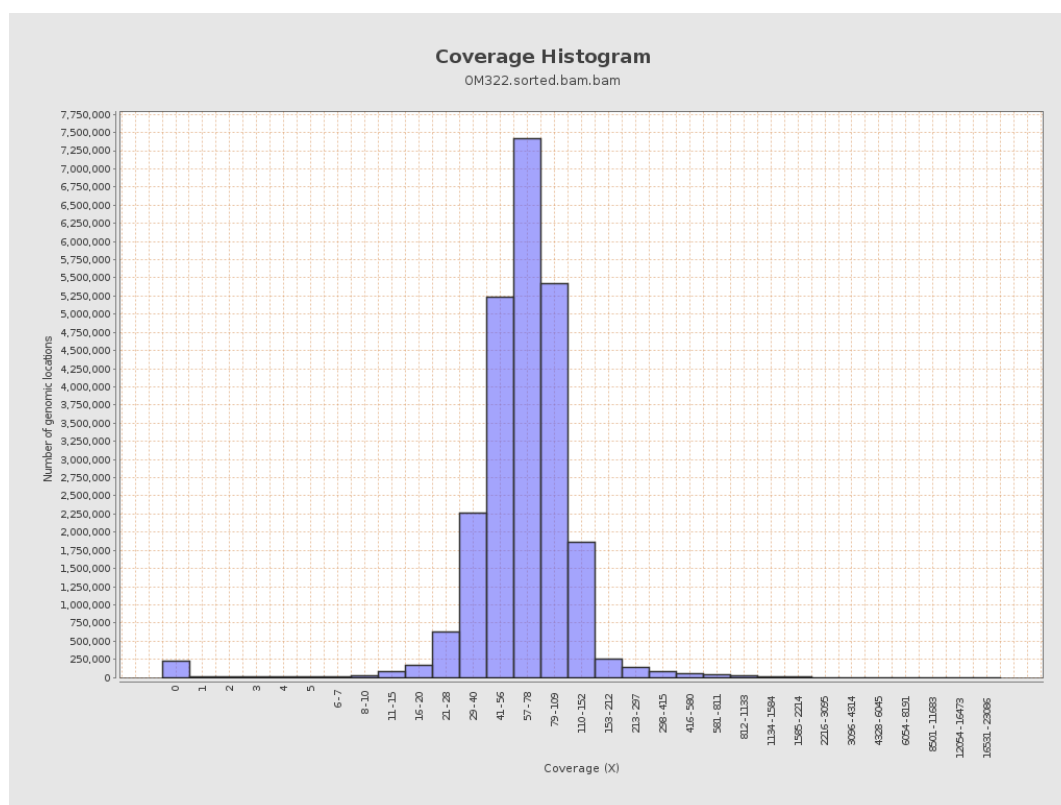
gi 1074120478 emb LT615256.1	977217	62259938	63.7115	34.9679
gi 1074120682 emb LT615257.1	860454	62886589	73.0854	57.1542
gi 1074120865 emb LT615258.1	989719	93701200	94.6745	152.3681
gi 1074121086 emb LT615259.1	935450	79854863	85.3652	146.034
gi 1074121301 emb LT615260.1	1432239	114152545	79.7022	91.5154
gi 1074121615 emb LT615261.1	1080962	79243283	73.3081	69.0863
gi 1074121871 emb LT615262.1	1545099	108921962	70.4951	36.1984
gi 1074122235 emb LT615263.1	1585108	116463065	73.4733	70.2569
gi 1074122590 emb LT615264.1	2122358	150788561	71.0477	45.348
gi 1074123050 emb LT615265.1	1754192	139163799	79.3321	105.2954
gi 1074123421 emb LT615	2150147	171095049	79.5737	157.2388

266.1				
gi 107412389 8 emb LT615 267.1	3031036	215080202	70.9593	44.8934
gi 107412458 8 emb LT615 268.1	2359348	172752783	73.2206	84.1873
gi 107412506 5 emb LT615 269.1	3135668	235620399	75.142	29.8257

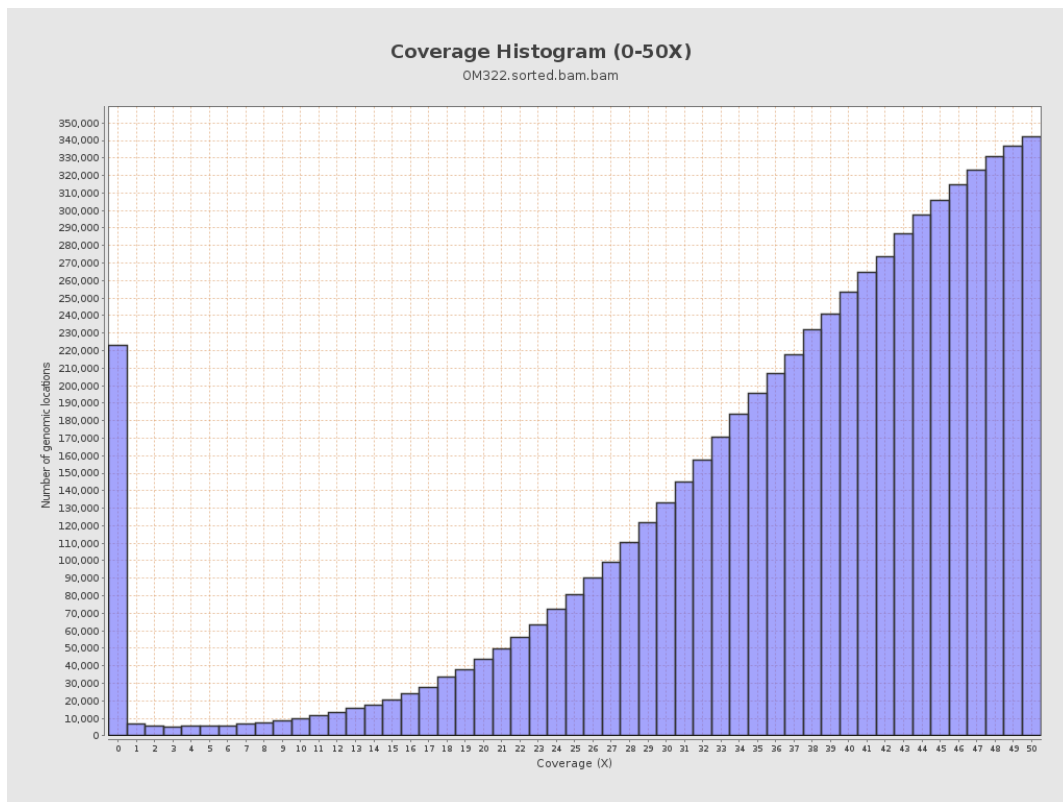
3. Results : Coverage across reference



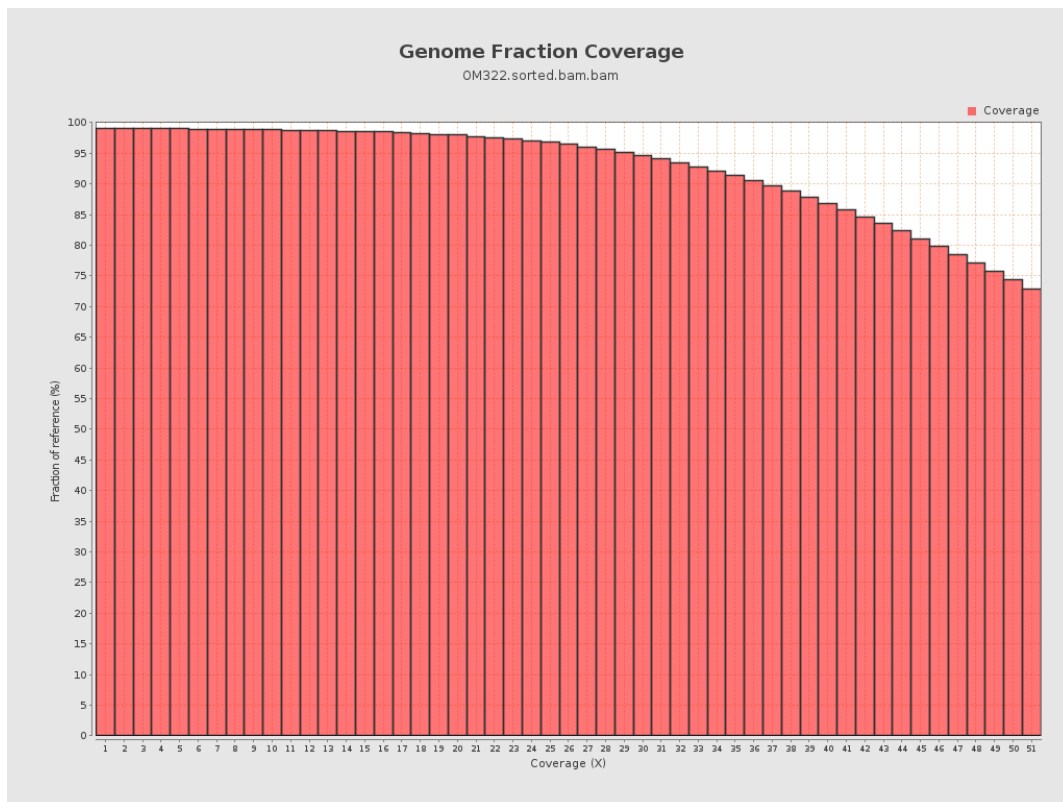
4. Results : Coverage Histogram



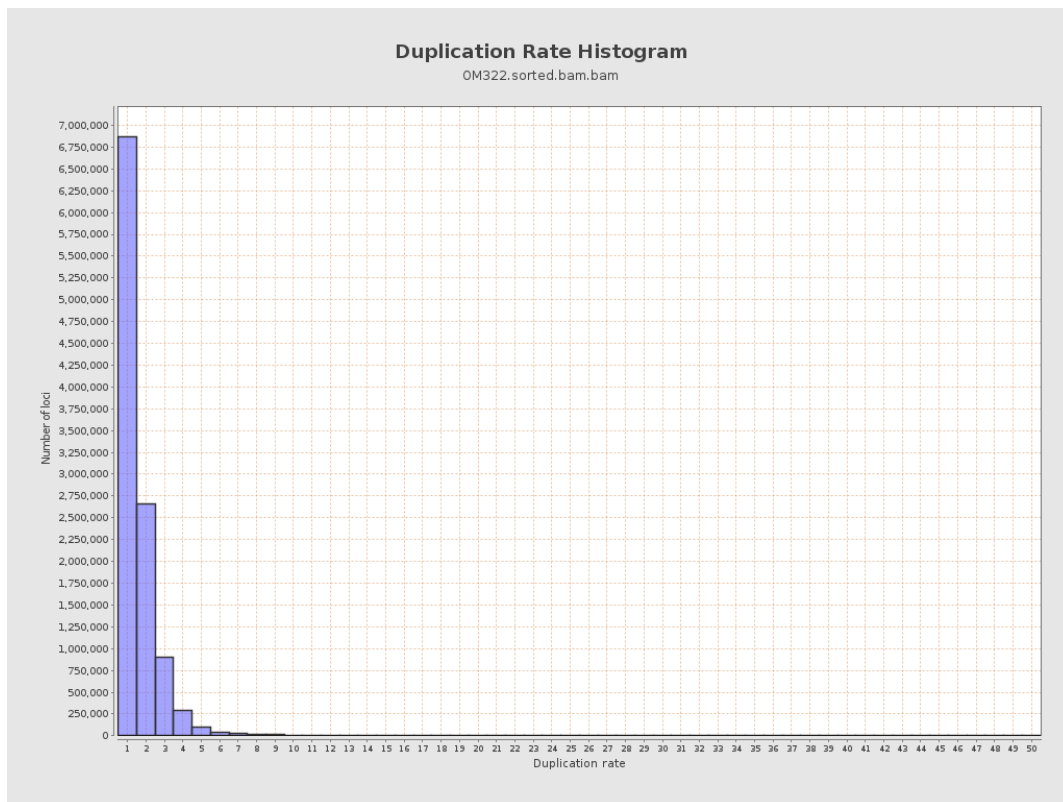
5. Results : Coverage Histogram (0-50X)



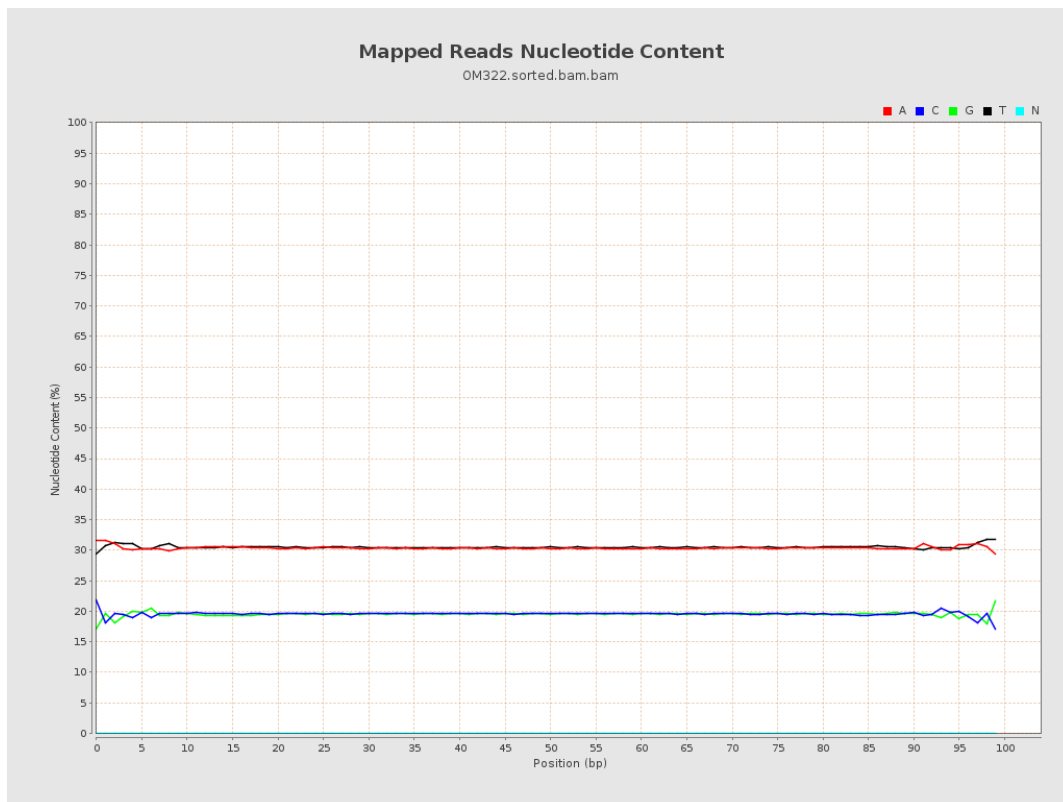
6. Results : Genome Fraction Coverage



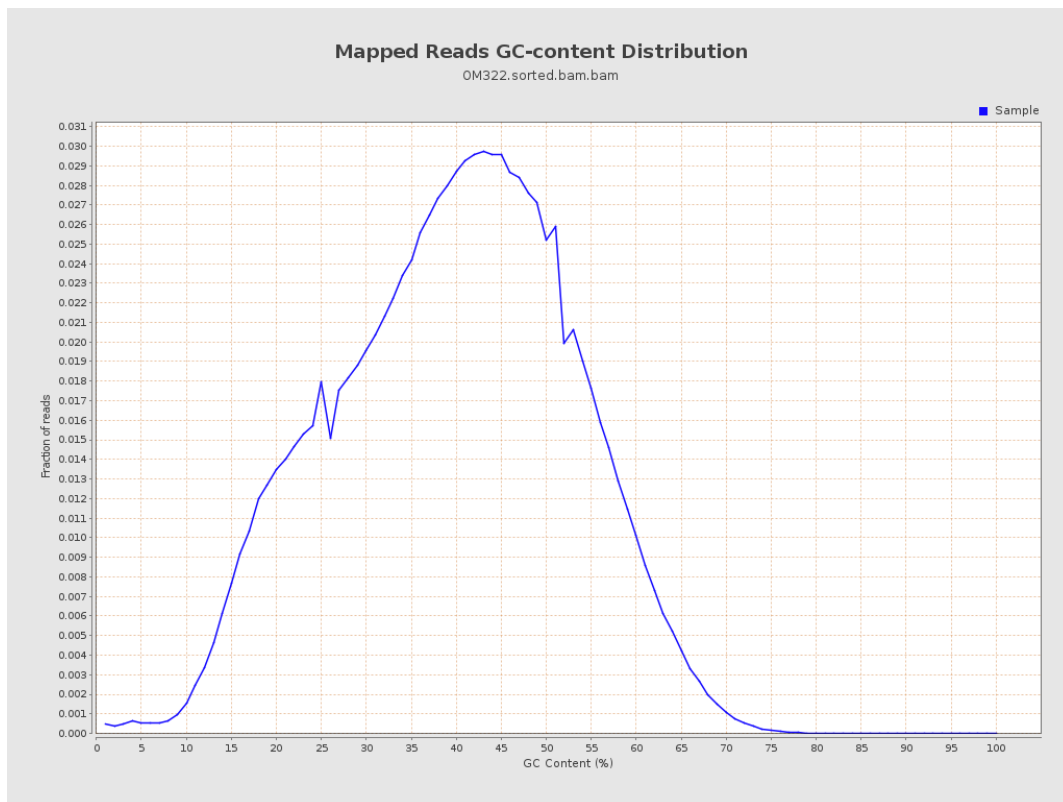
7. Results : Duplication Rate Histogram



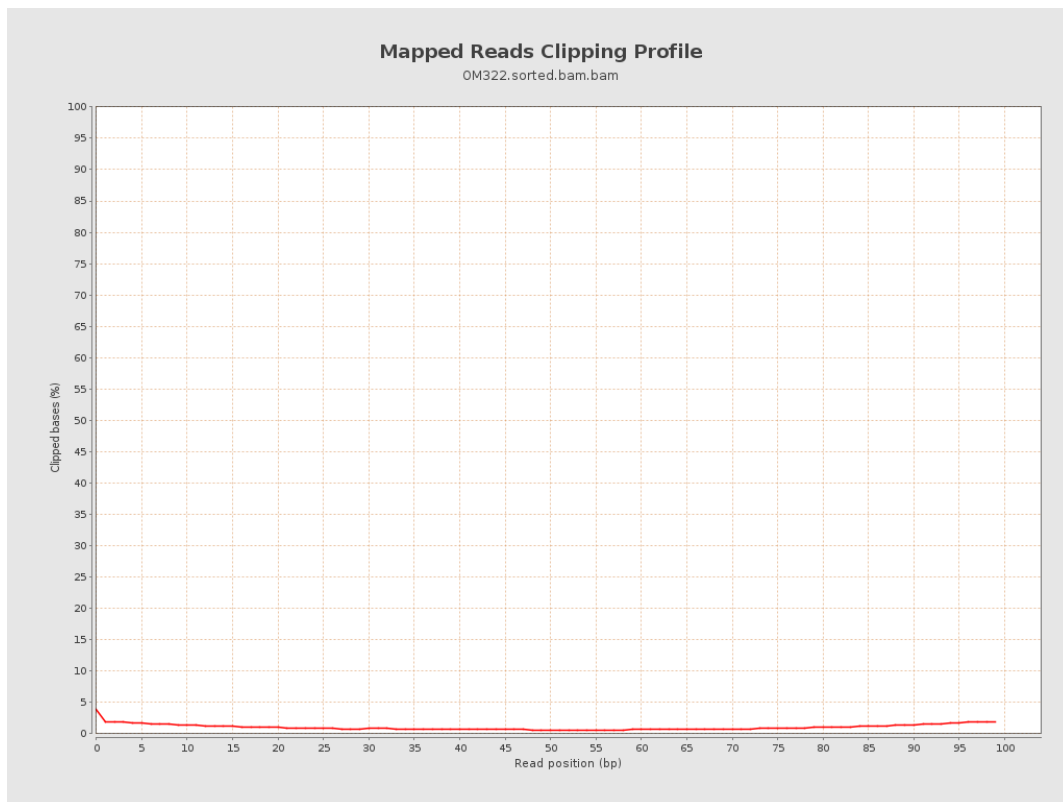
8. Results : Mapped Reads Nucleotide Content



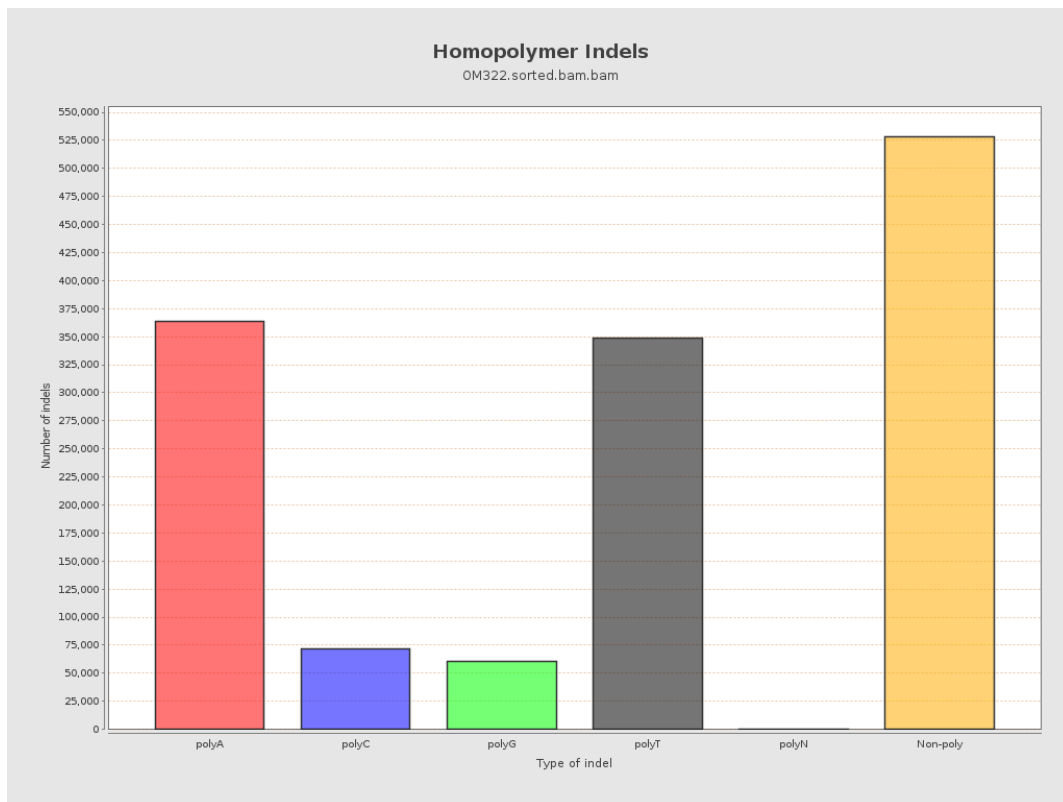
9. Results : Mapped Reads GC-content Distribution



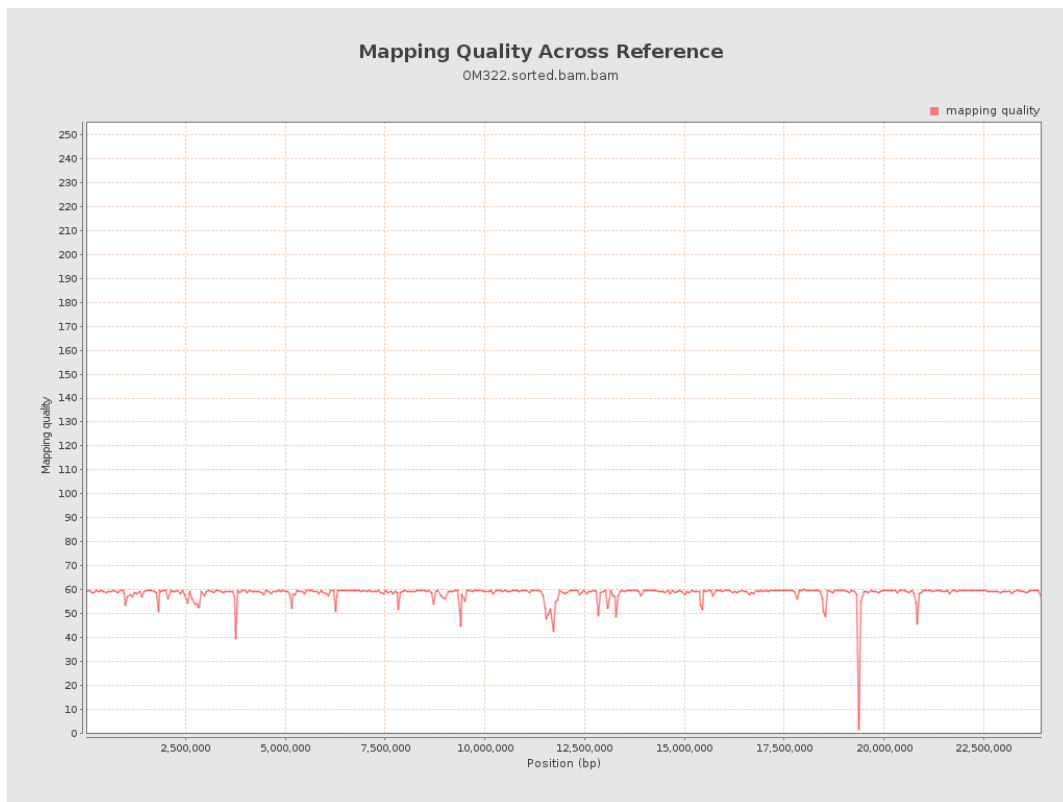
10. Results : Mapped Reads Clipping Profile



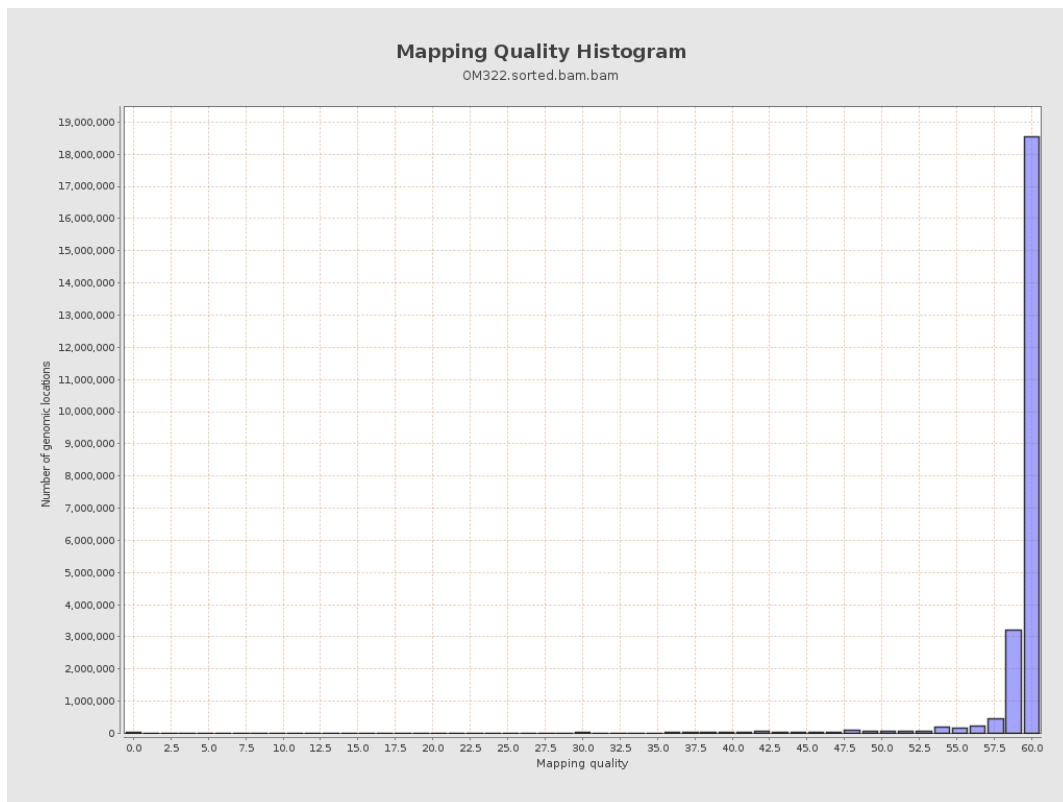
11. Results : Homopolymer Indels



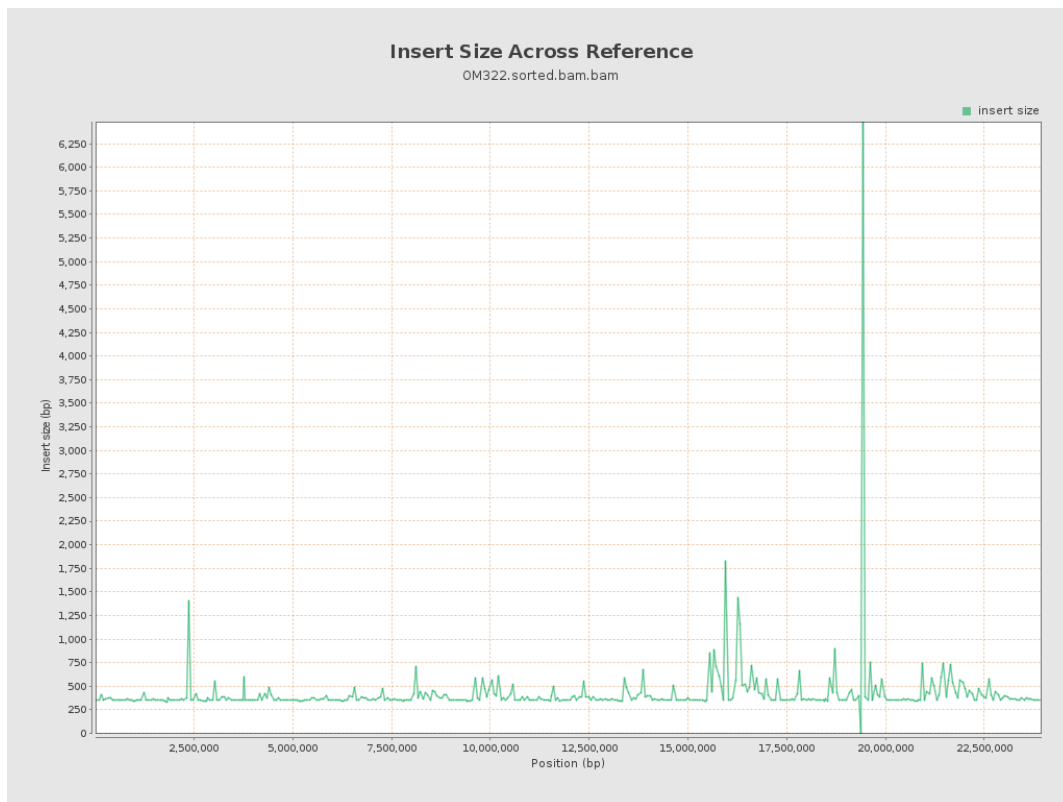
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

