

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 13:09:26

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/OM115.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/OM115-BiooBarcode14_AGTTCC_R1.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/OM115-BiooBarcode14_AGTTCC_R2.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 13:09:26 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/OM115.sorted.bam.

	bam
--	-----

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	22,954,073
Mapped reads	4,719,017 / 20.56%
Unmapped reads	18,235,056 / 79.44%
Mapped paired reads	4,719,017 / 20.56%
Mapped reads, first in pair	2,364,223 / 10.3%
Mapped reads, second in pair	2,354,794 / 10.26%
Mapped reads, both in pair	4,540,396 / 19.78%
Mapped reads, singletons	178,621 / 0.78%
Read min/max/mean length	30 / 100 / 99.96
Duplicated reads (estimated)	912,414 / 3.97%
Duplication rate	15.16%
Clipped reads	584,901 / 2.55%

2.2. ACGT Content

Number/percentage of A's	134,884,346 / 29.9%
Number/percentage of C's	90,396,007 / 20.04%
Number/percentage of T's	135,609,763 / 30.06%
Number/percentage of G's	90,242,803 / 20%
Number/percentage of N's	38,739 / 0.01%
GC Percentage	40.04%

2.3. Coverage

Mean	18.8497
Standard Deviation	23.5856

2.4. Mapping Quality

Mean Mapping Quality	58.2
----------------------	------

2.5. Insert size

Mean	846.42
Standard Deviation	28,702.56
P25/Median/P75	278 / 289 / 297

2.6. Mismatches and indels

General error rate	1.46%
Mismatches	6,225,641
Insertions	147,112
Mapped reads with at least one insertion	2.94%
Deletions	169,217
Mapped reads with at least one deletion	3.36%
Homopolymer indels	62.82%

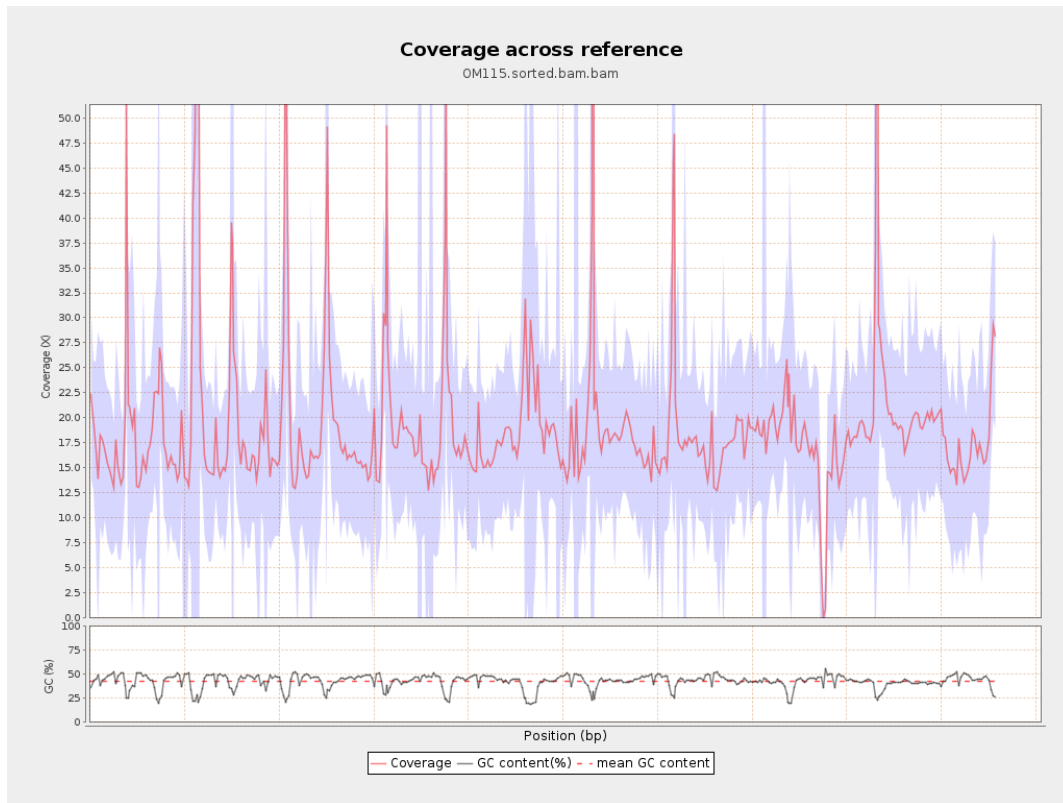
2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

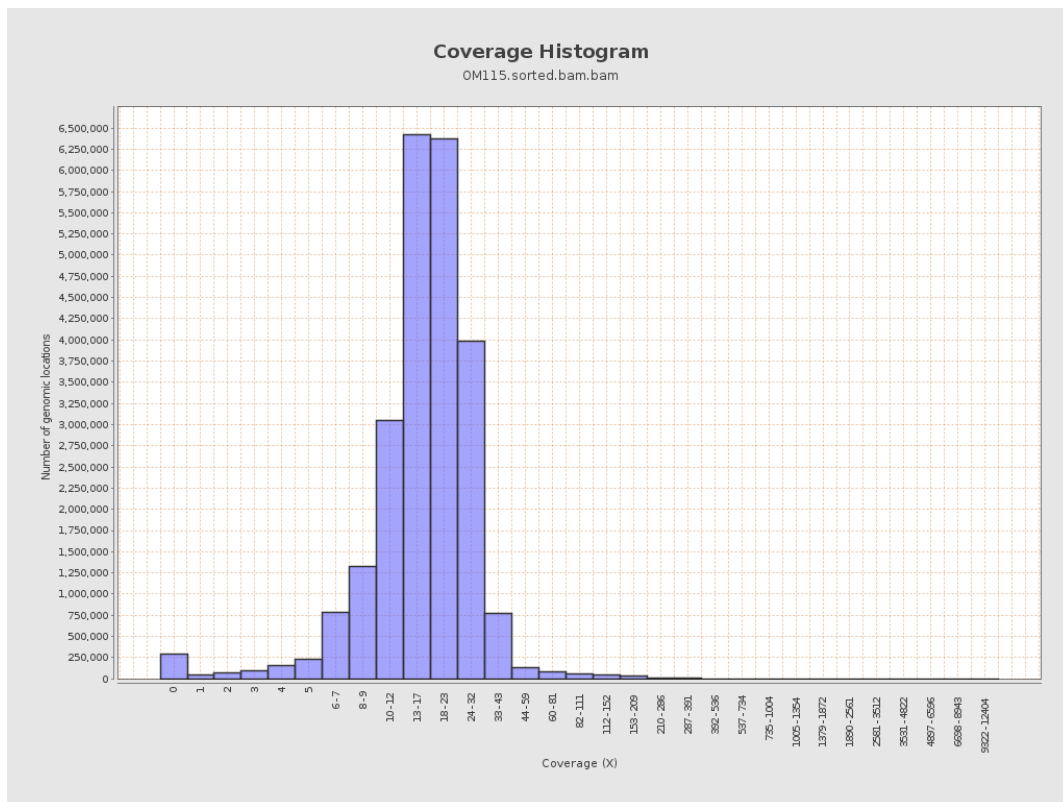
gi 1074120478 emb LT615256.1	977217	16750022	17.1405	10.3095
gi 1074120682 emb LT615257.1	860454	15576472	18.1026	13.661
gi 1074120865 emb LT615258.1	989719	20472536	20.6852	26.7716
gi 1074121086 emb LT615259.1	935450	19815430	21.1828	29.564
gi 1074121301 emb LT615260.1	1432239	28059370	19.5913	18.4361
gi 1074121615 emb LT615261.1	1080962	20035258	18.5347	14.8352
gi 1074121871 emb LT615262.1	1545099	28044222	18.1504	9.5311
gi 1074122235 emb LT615263.1	1585108	30218991	19.0643	29.9045
gi 1074122590 emb LT615264.1	2122358	38299449	18.0457	10.3535
gi 1074123050 emb LT615265.1	1754192	34324920	19.5674	54.8873
gi 1074123421 emb LT615	2150147	42103444	19.5817	29.7783

266.1				
gi 107412389 8 emb LT615 267.1	3031036	55411457	18.2814	15.6642
gi 107412458 8 emb LT615 268.1	2359348	42609262	18.0598	18.9125
gi 107412506 5 emb LT615 269.1	3135668	59899321	19.1026	8.5291

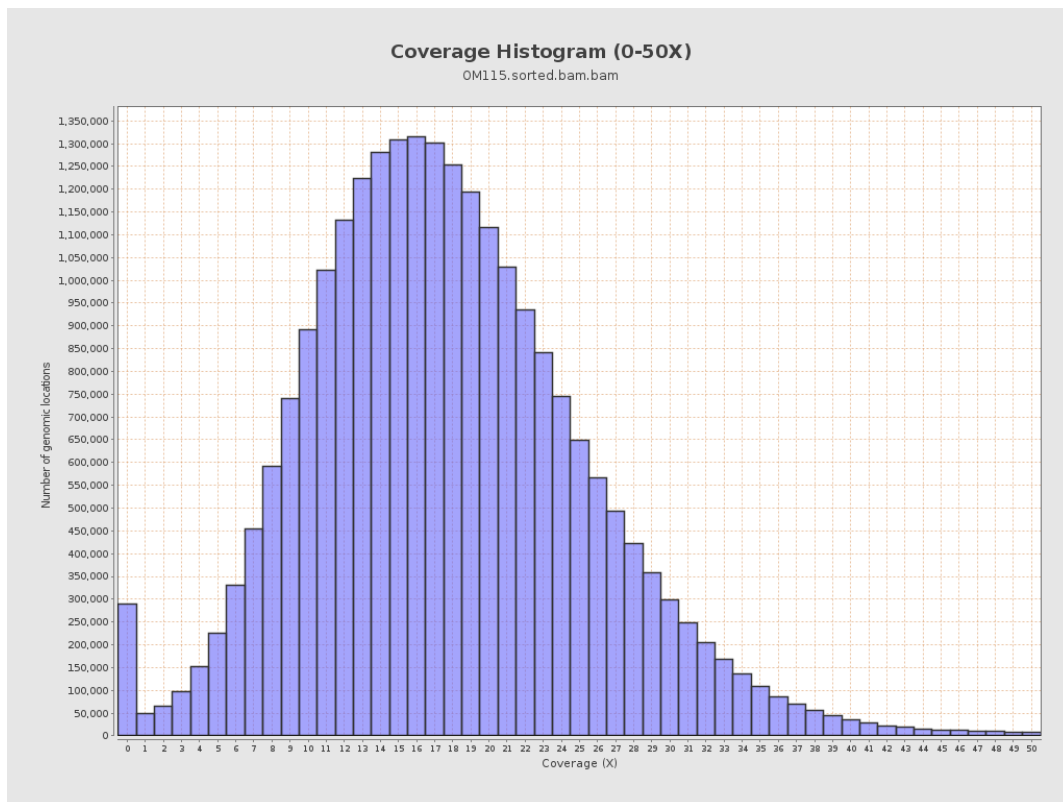
3. Results : Coverage across reference



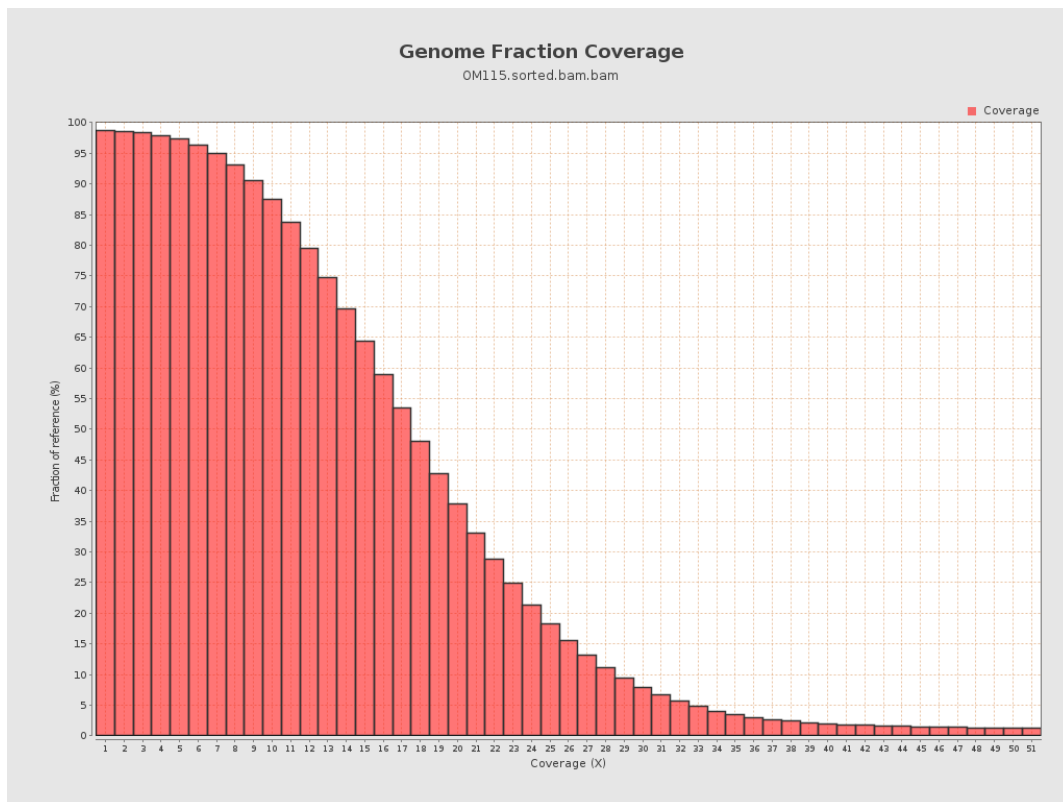
4. Results : Coverage Histogram



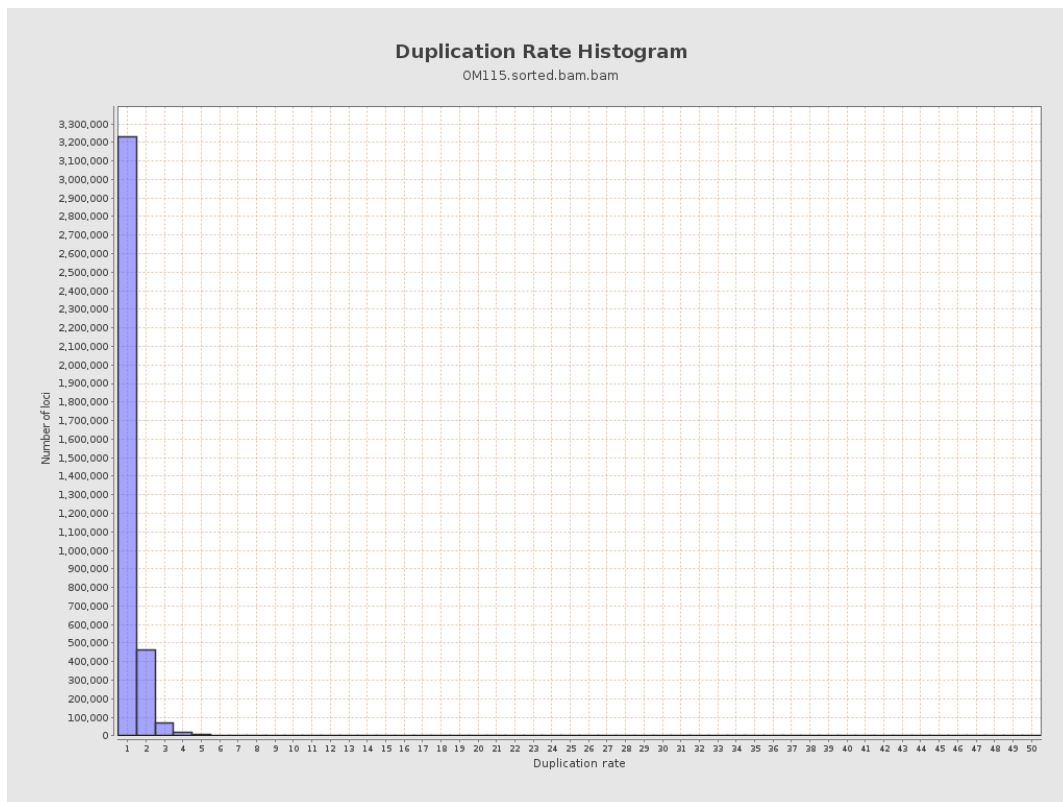
5. Results : Coverage Histogram (0-50X)



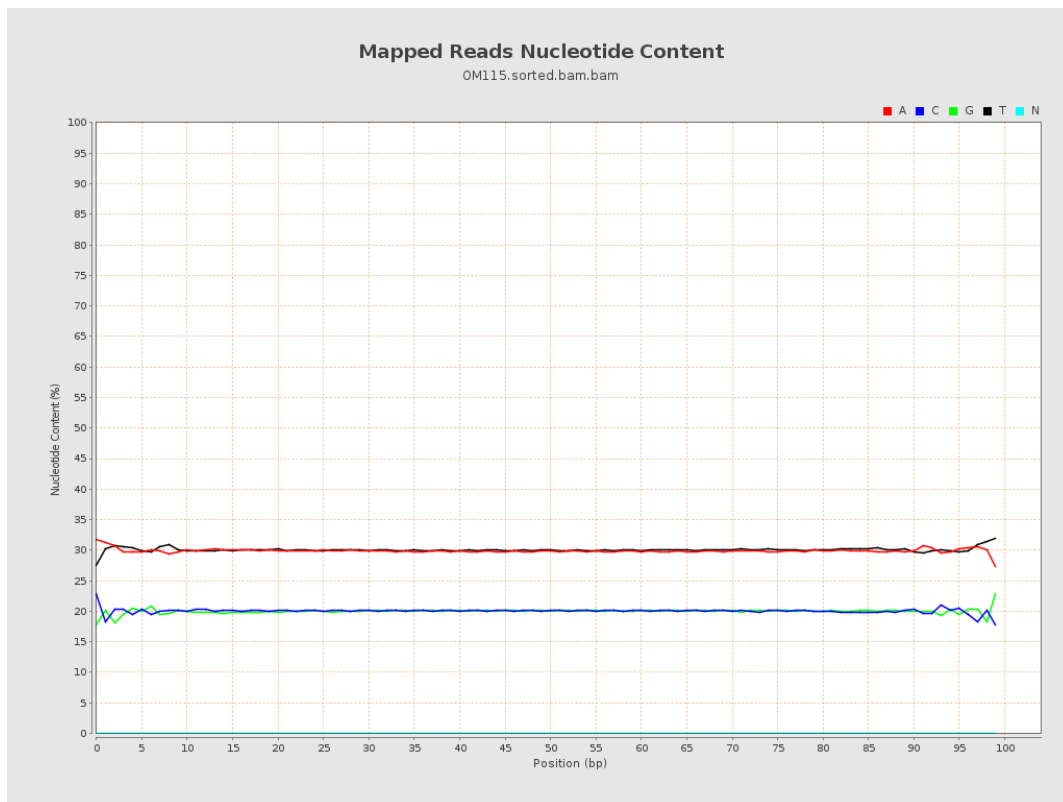
6. Results : Genome Fraction Coverage



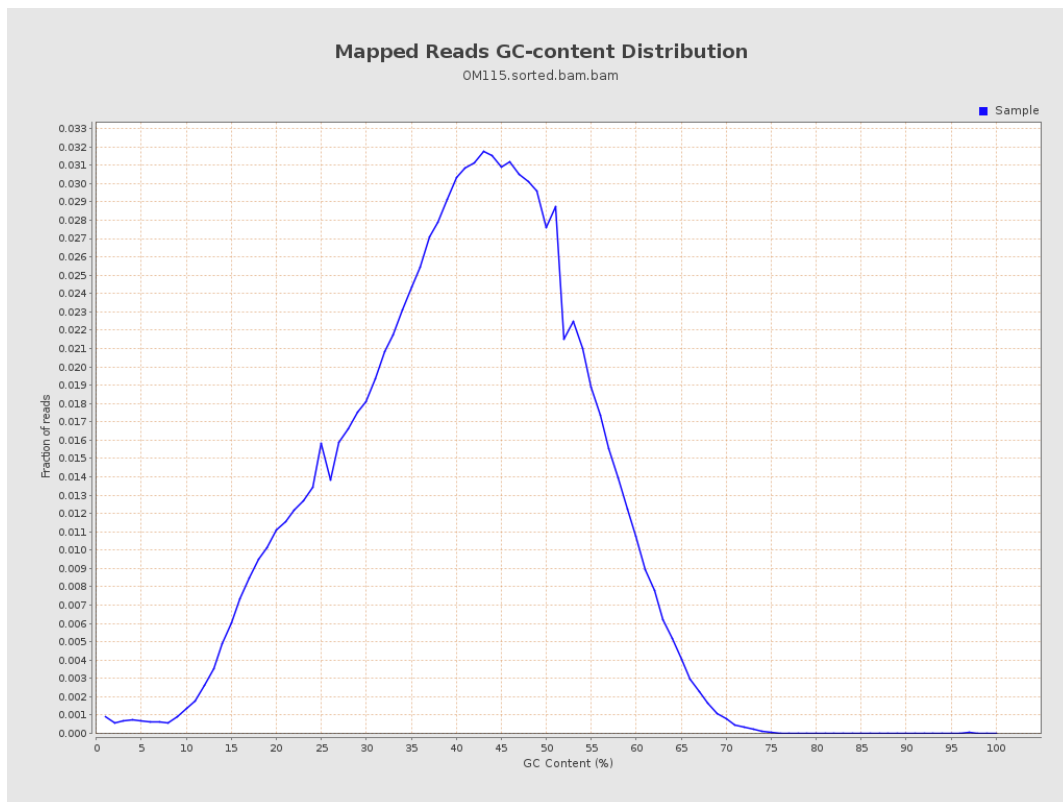
7. Results : Duplication Rate Histogram



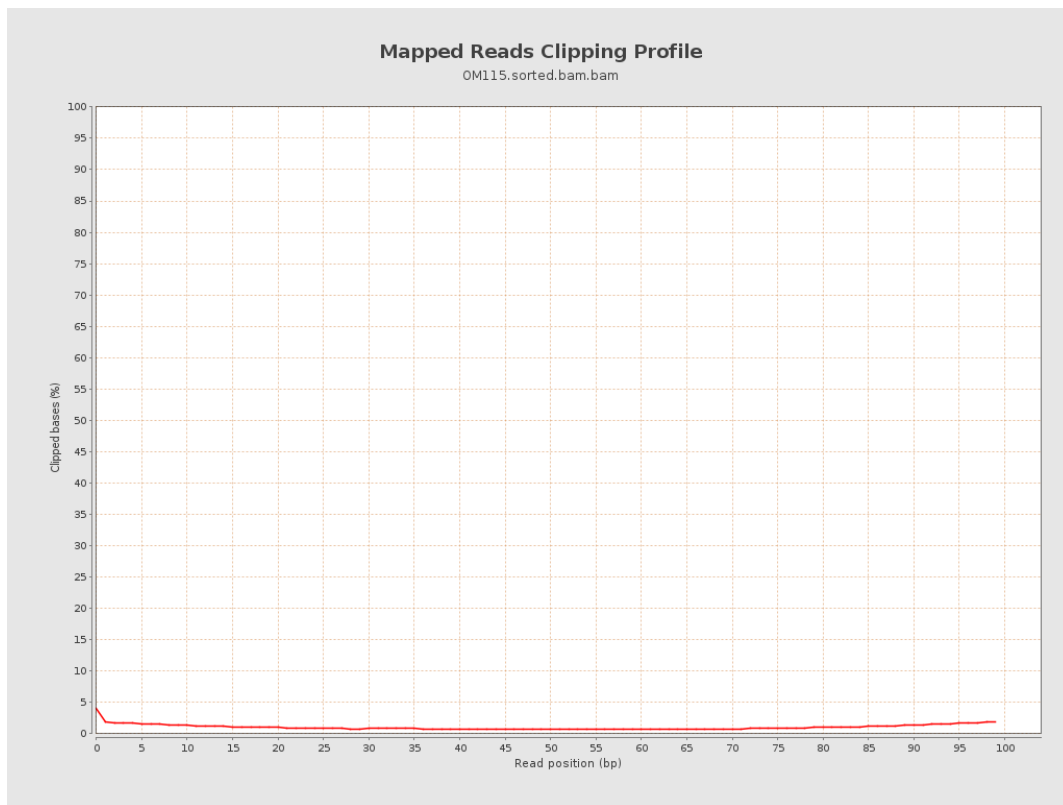
8. Results : Mapped Reads Nucleotide Content



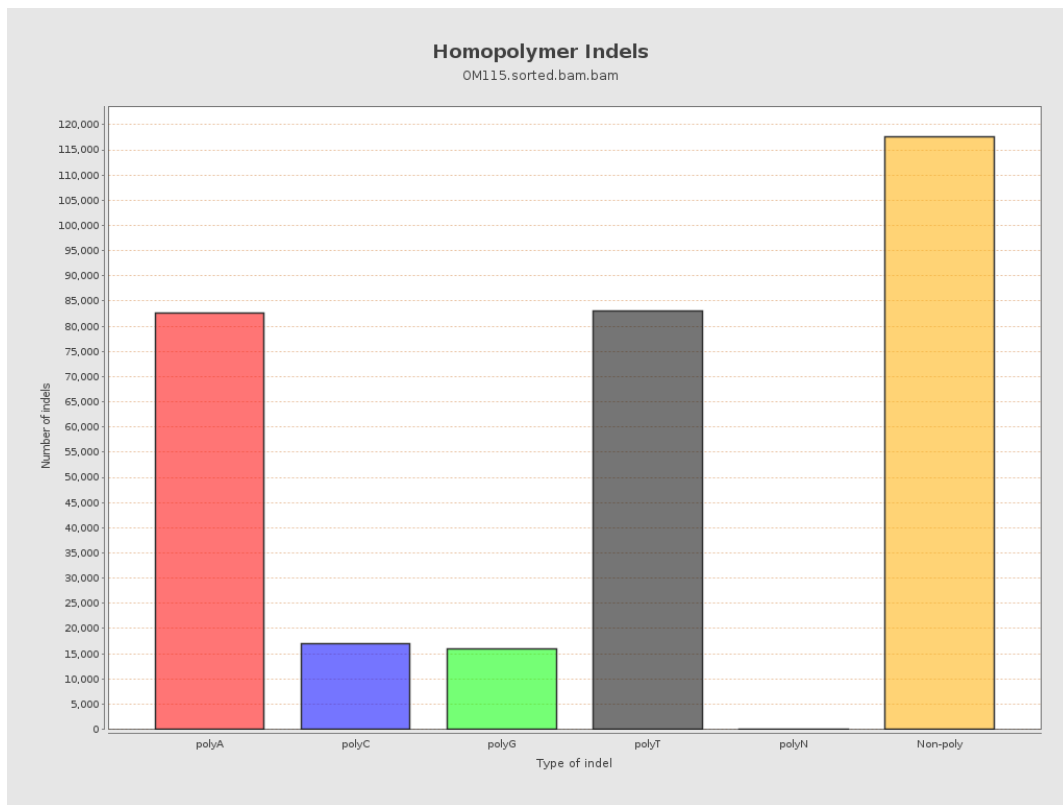
9. Results : Mapped Reads GC-content Distribution



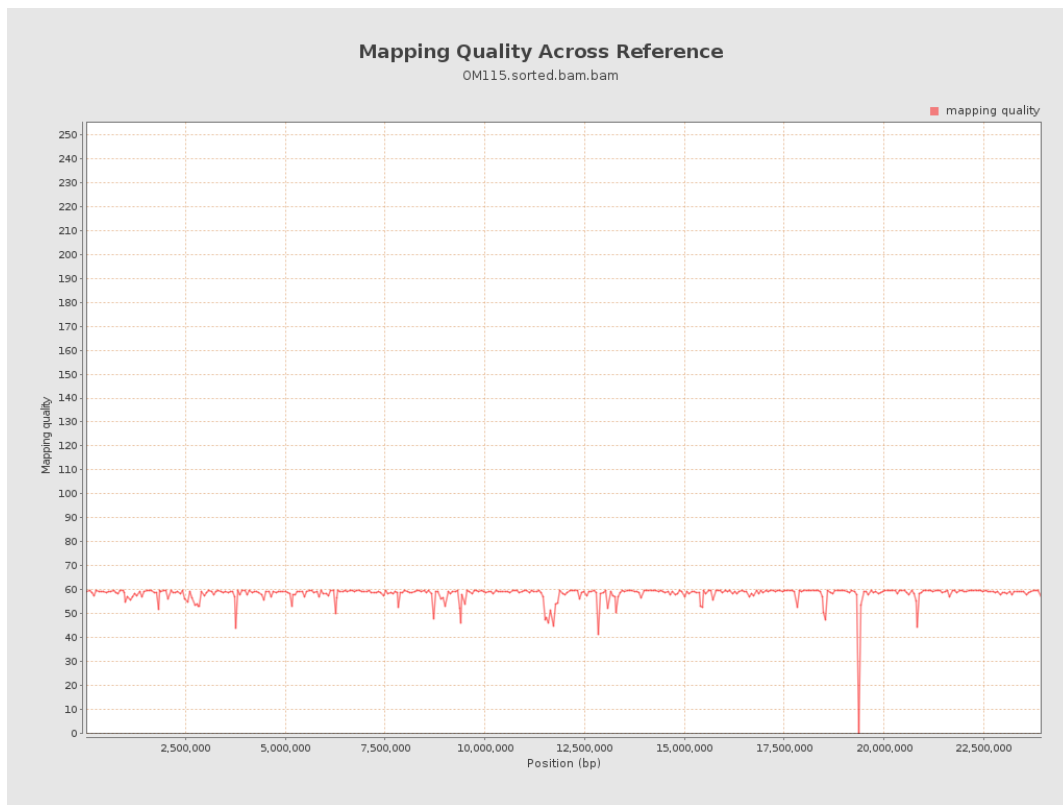
10. Results : Mapped Reads Clipping Profile



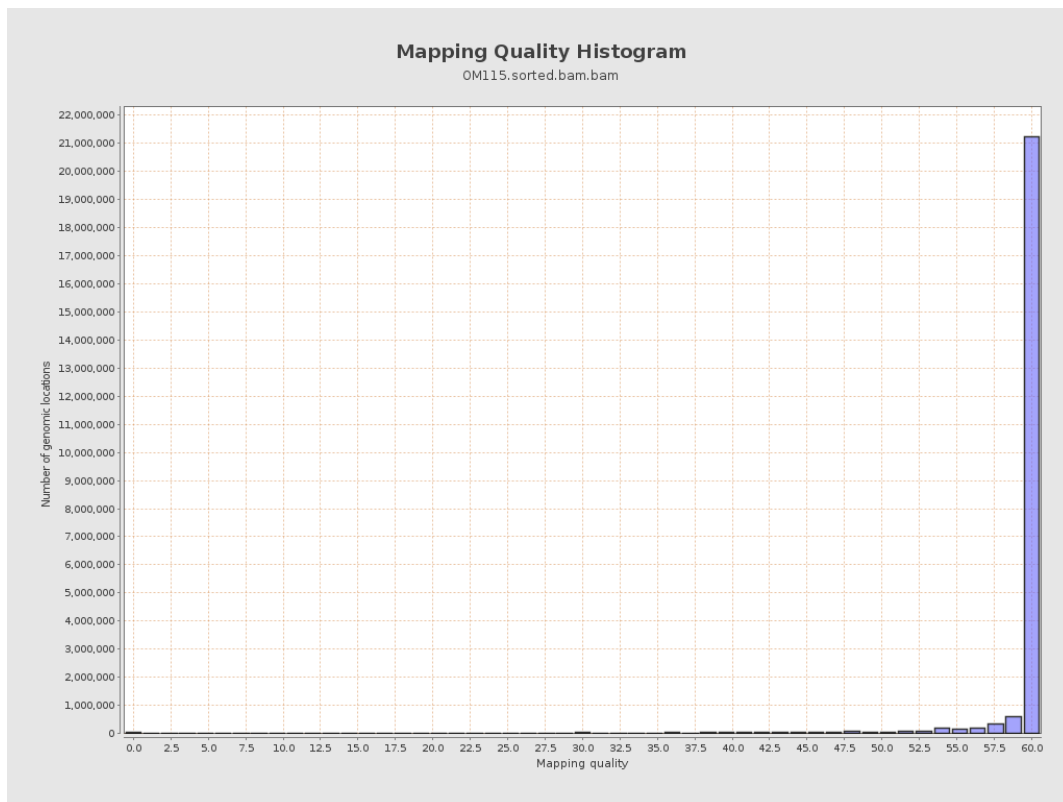
11. Results : Homopolymer Indels



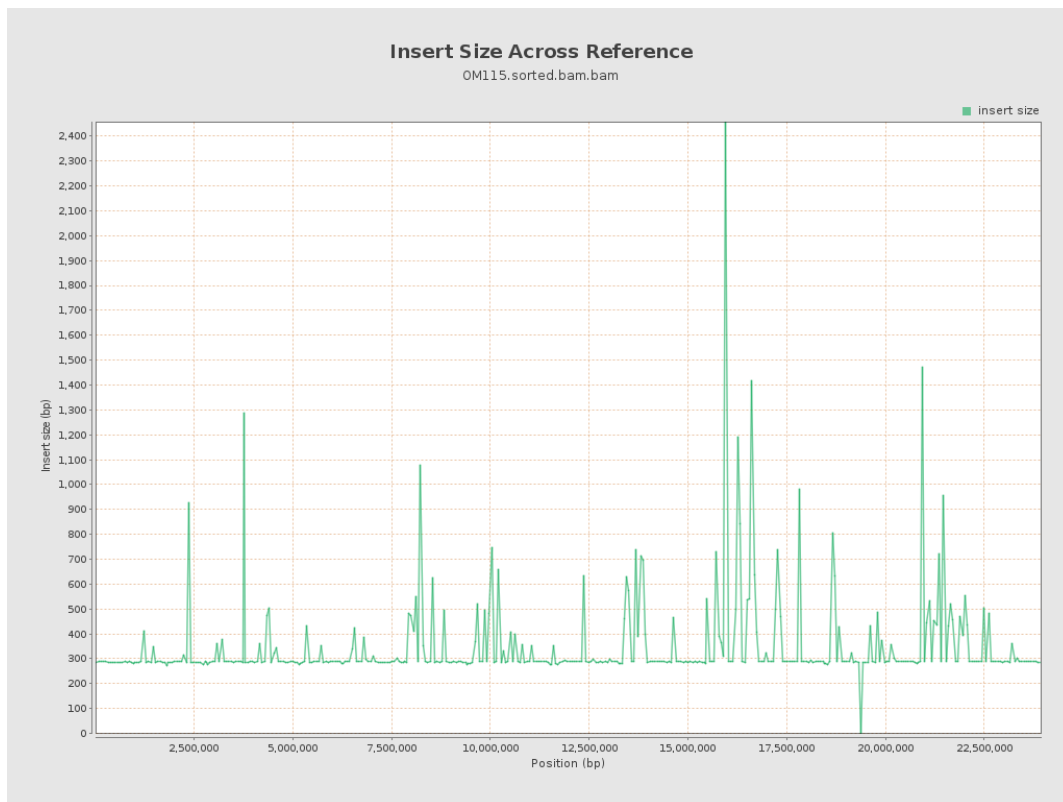
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

