

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 11:31:47

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/BB014.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/BB014-BiooBarcode_32_CACTCA_R2.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/BB014-BiooBarcode_32_CACTCA_R1.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 11:31:46 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/BB014.sorted.bam.b

am

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	21,362,097
Mapped reads	12,173,139 / 56.98%
Unmapped reads	9,188,958 / 43.02%
Mapped paired reads	12,173,139 / 56.98%
Mapped reads, first in pair	6,063,028 / 28.38%
Mapped reads, second in pair	6,110,111 / 28.6%
Mapped reads, both in pair	11,956,747 / 55.97%
Mapped reads, singletons	216,392 / 1.01%
Read min/max/mean length	30 / 100 / 99.89
Duplicated reads (estimated)	4,393,795 / 20.57%
Duplication rate	34.45%
Clipped reads	1,155,220 / 5.41%

2.2. ACGT Content

Number/percentage of A's	332,229,585 / 28.23%
Number/percentage of C's	256,082,815 / 21.76%
Number/percentage of T's	333,503,635 / 28.34%
Number/percentage of G's	255,154,803 / 21.68%
Number/percentage of N's	99,226 / 0.01%
GC Percentage	43.44%

2.3. Coverage

Mean	49.1775
Standard Deviation	41.8243

2.4. Mapping Quality

Mean Mapping Quality	58.67
----------------------	-------

2.5. Insert size

Mean	768.75
Standard Deviation	24,702.37
P25/Median/P75	238 / 291 / 349

2.6. Mismatches and indels

General error rate	0.95%
Mismatches	10,353,890
Insertions	281,843
Mapped reads with at least one insertion	2.2%
Deletions	373,750
Mapped reads with at least one deletion	2.92%
Homopolymer indels	66.71%

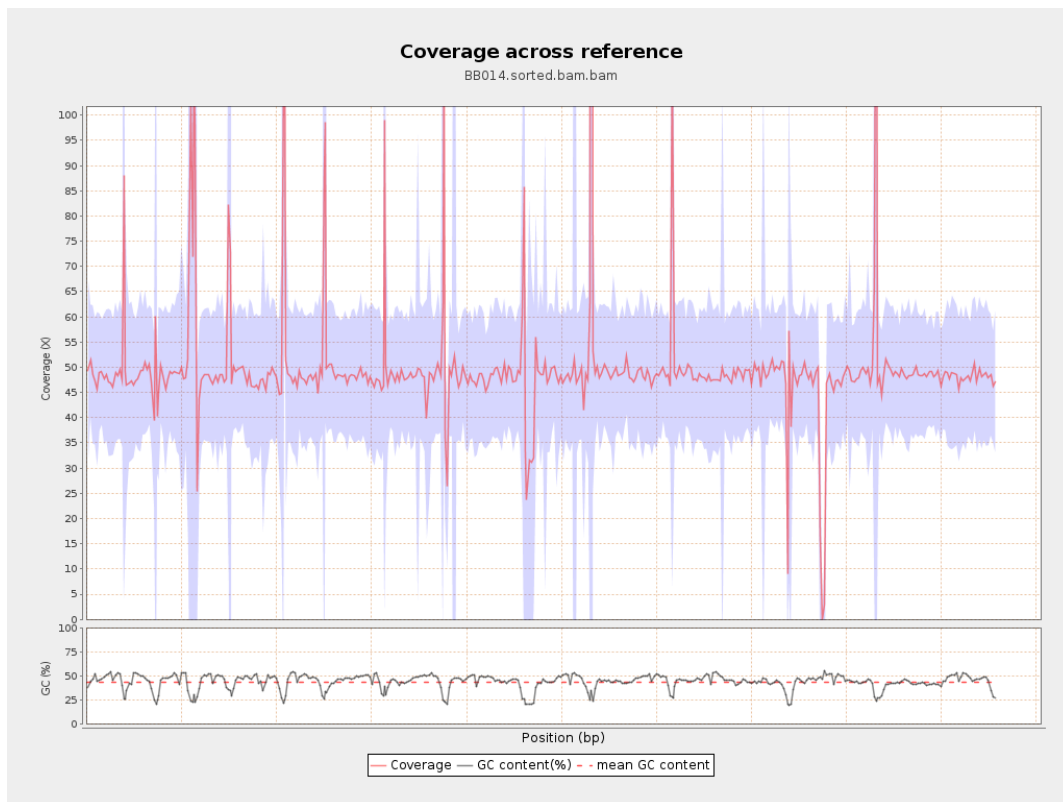
2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

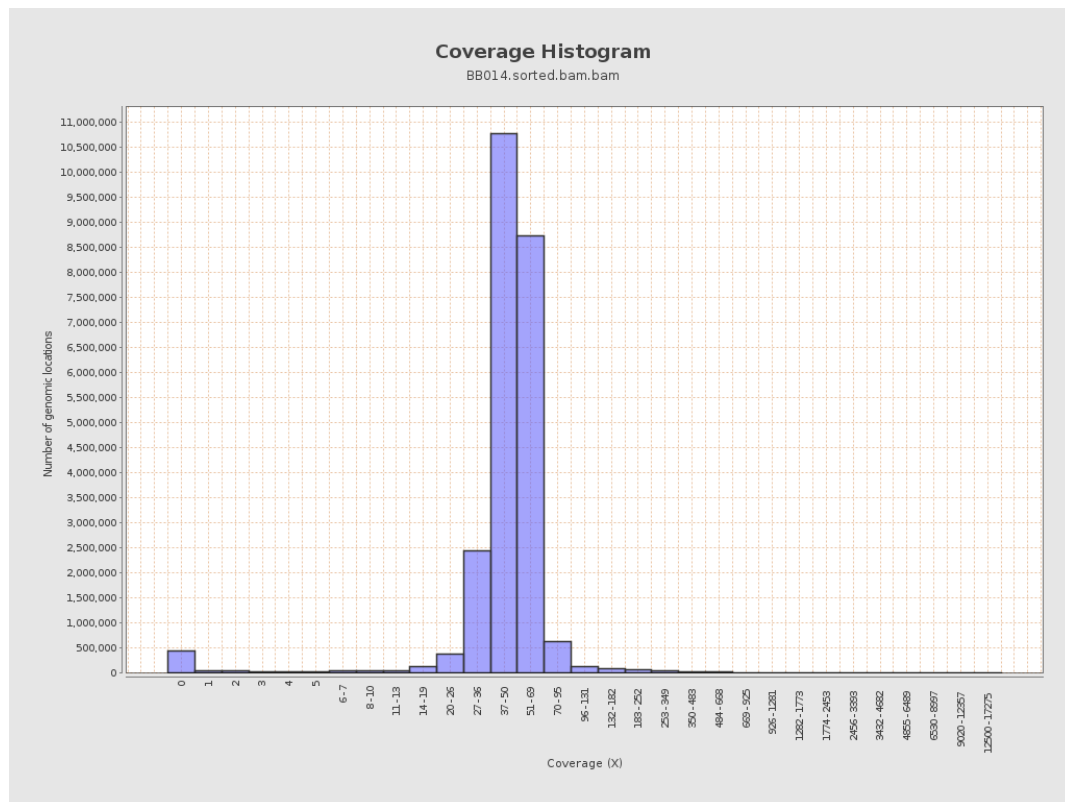
gi 1074120478 emb LT615256.1	977217	47833449	48.9486	17.4368
gi 1074120682 emb LT615257.1	860454	41513375	48.2459	23.0289
gi 1074120865 emb LT615258.1	989719	53181840	53.7343	45.8527
gi 1074121086 emb LT615259.1	935450	47309129	50.5737	54.9557
gi 1074121301 emb LT615260.1	1432239	71312656	49.791	33.8321
gi 1074121615 emb LT615261.1	1080962	55093216	50.9668	28.4907
gi 1074121871 emb LT615262.1	1545099	75056143	48.5769	13.5572
gi 1074122235 emb LT615263.1	1585108	78673036	49.6326	24.8477
gi 1074122590 emb LT615264.1	2122358	102332478	48.2164	29.1333
gi 1074123050 emb LT615265.1	1754192	80806967	46.0651	36.0382
gi 1074123421 emb LT615	2150147	113162161	52.63	105.2727

266.1				
gi 107412389 8 emb LT615 267.1	3031036	147882463	48.7894	17.0001
gi 107412458 8 emb LT615 268.1	2359348	112244066	47.5742	36.939
gi 107412506 5 emb LT615 269.1	3135668	151842942	48.4244	11.1251

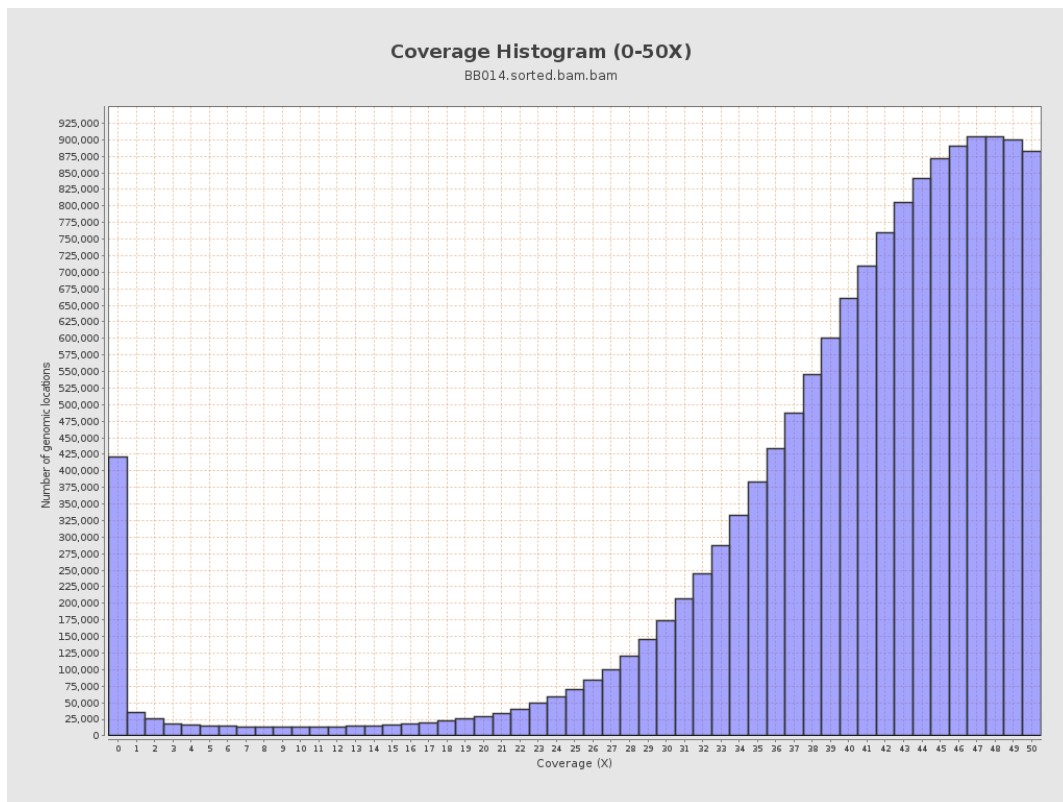
3. Results : Coverage across reference



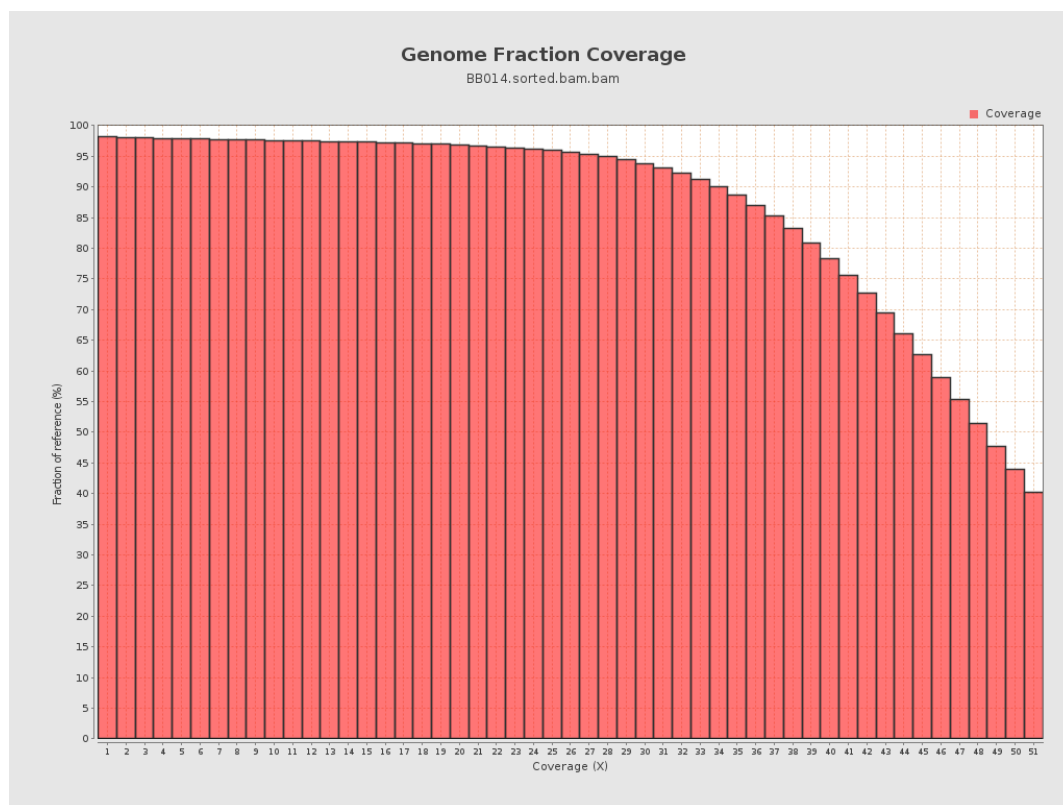
4. Results : Coverage Histogram



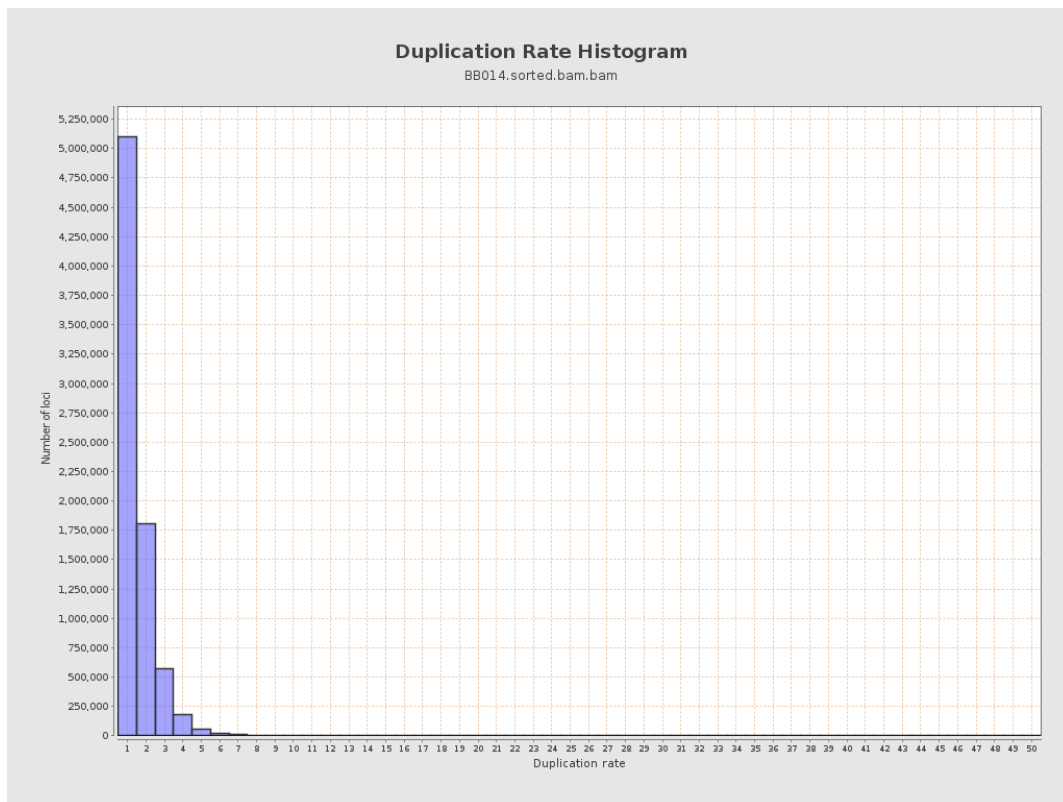
5. Results : Coverage Histogram (0-50X)



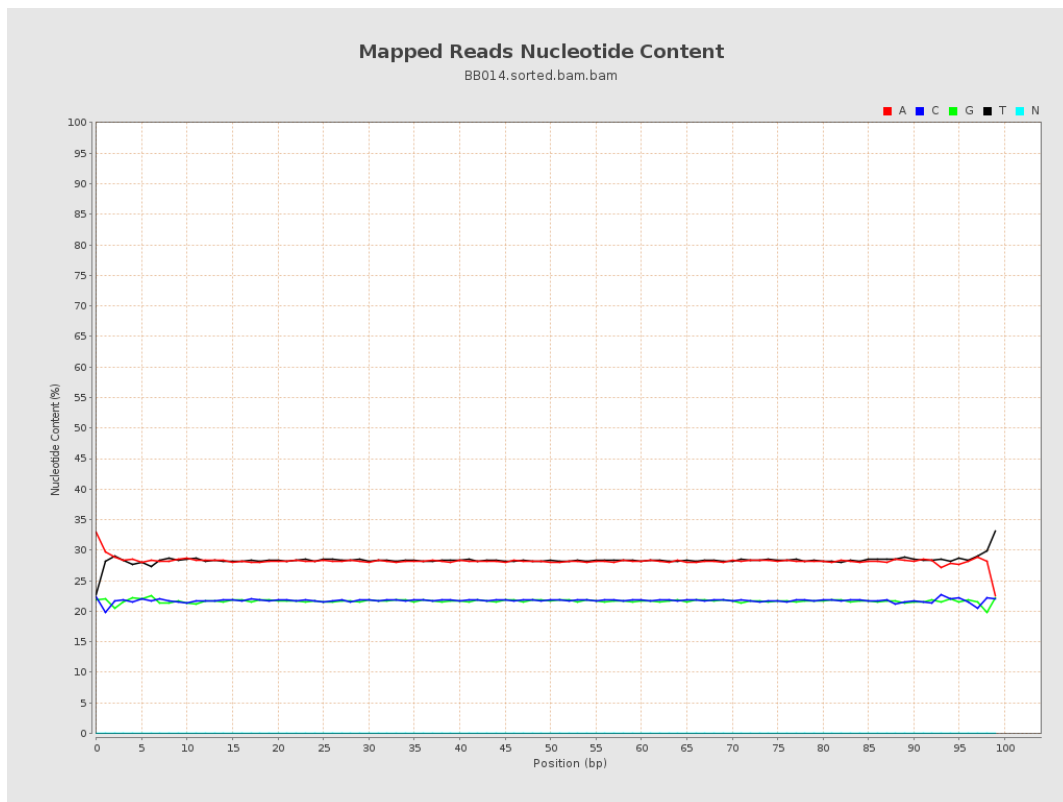
6. Results : Genome Fraction Coverage



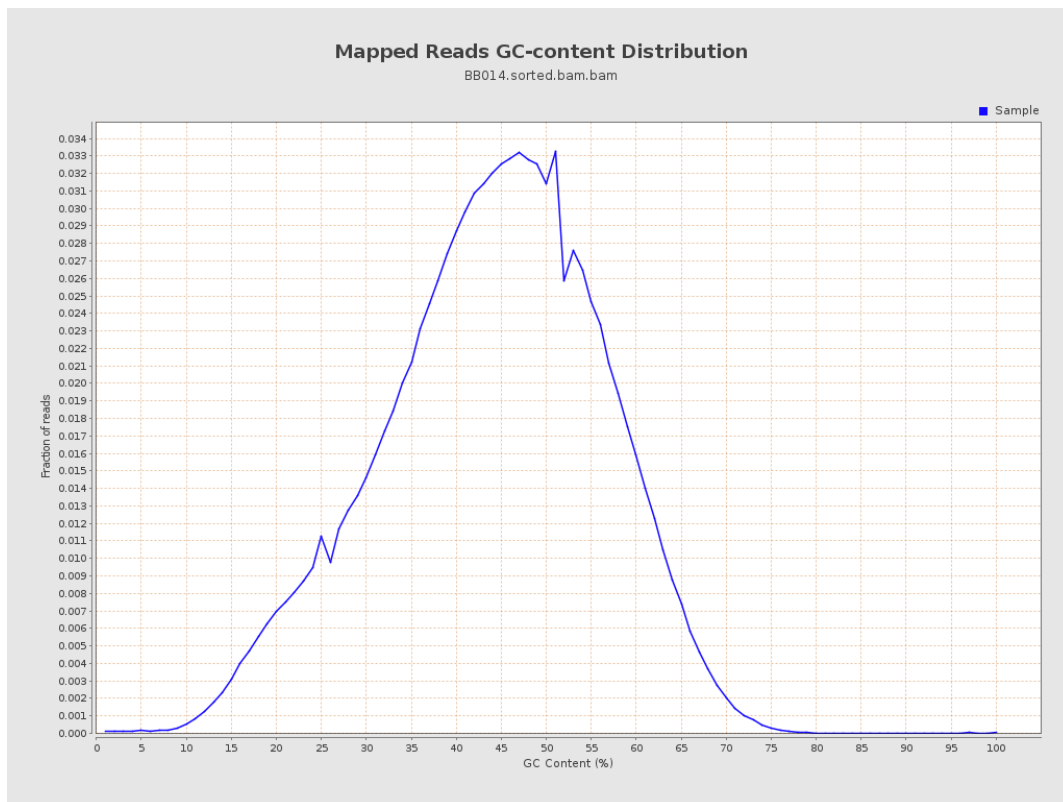
7. Results : Duplication Rate Histogram



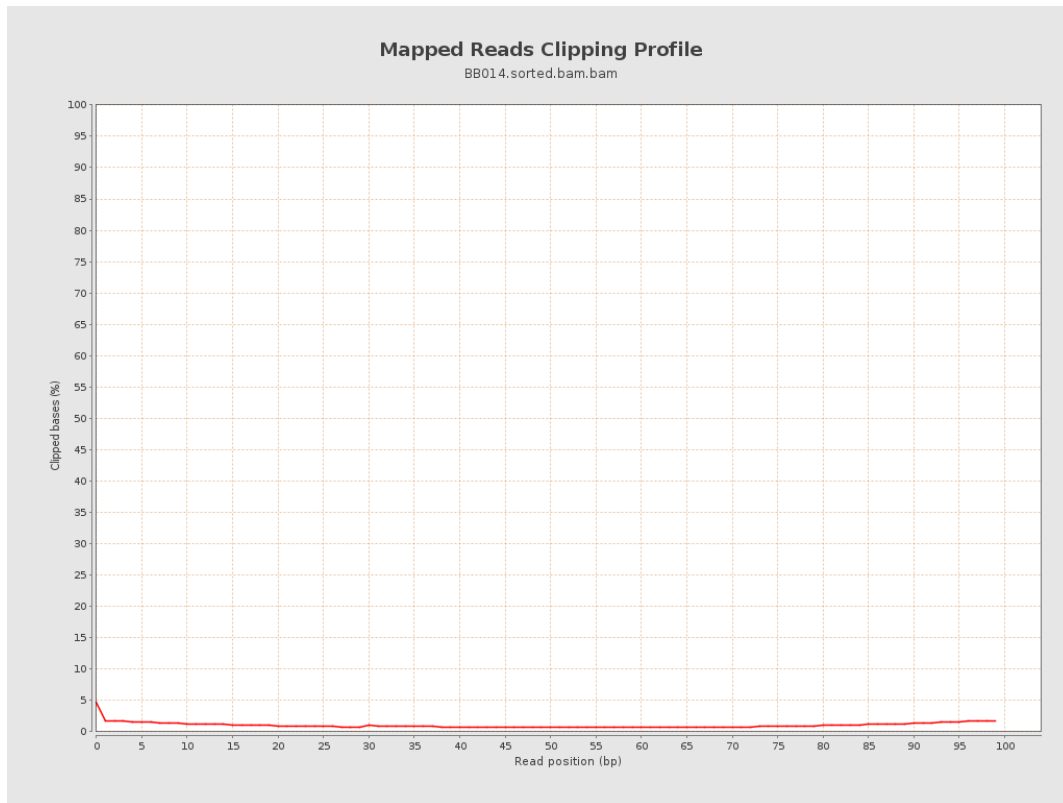
8. Results : Mapped Reads Nucleotide Content



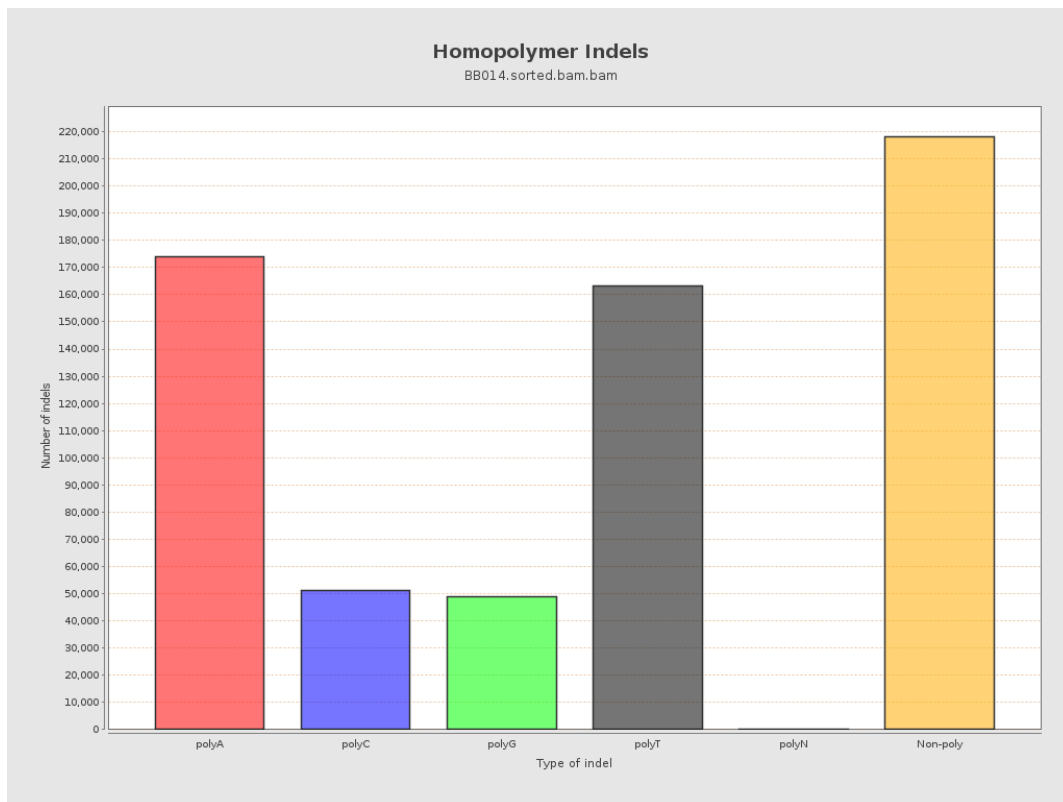
9. Results : Mapped Reads GC-content Distribution



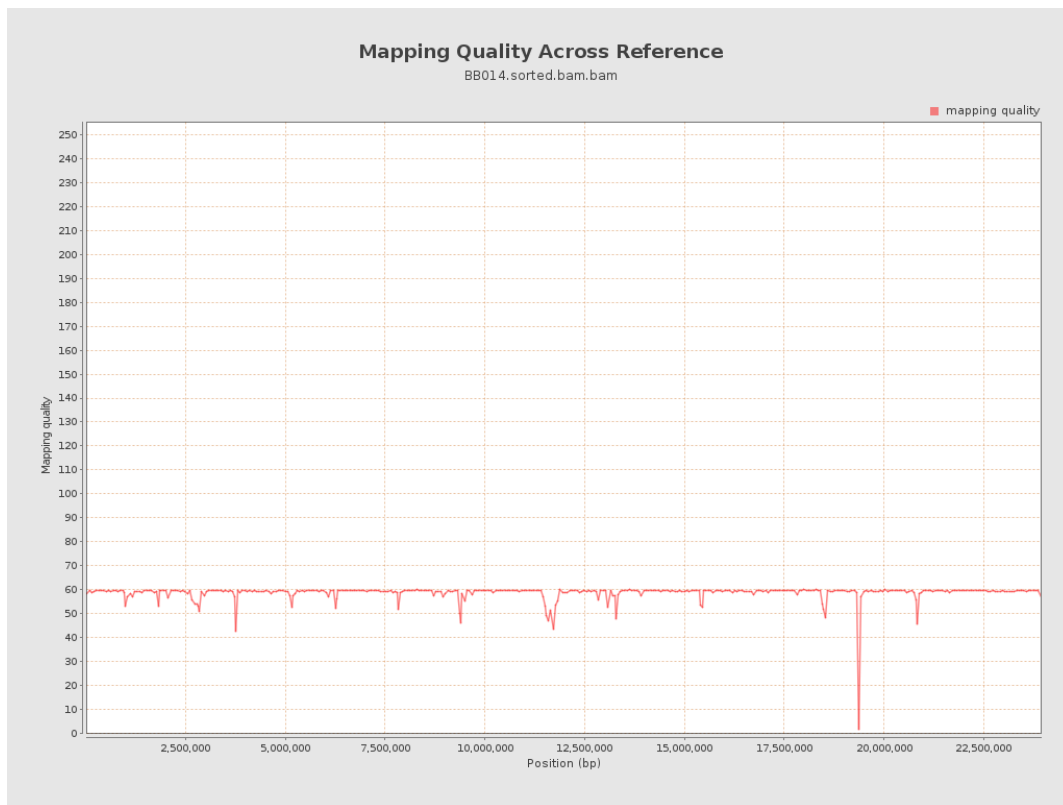
10. Results : Mapped Reads Clipping Profile



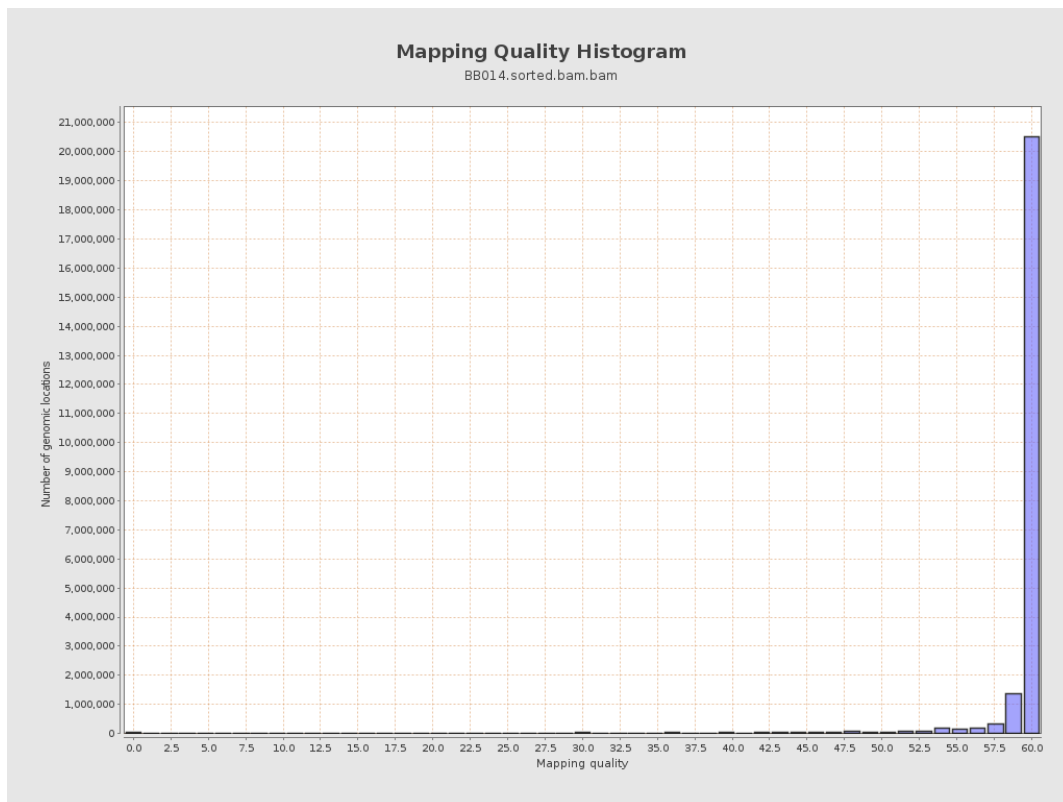
11. Results : Homopolymer Indels



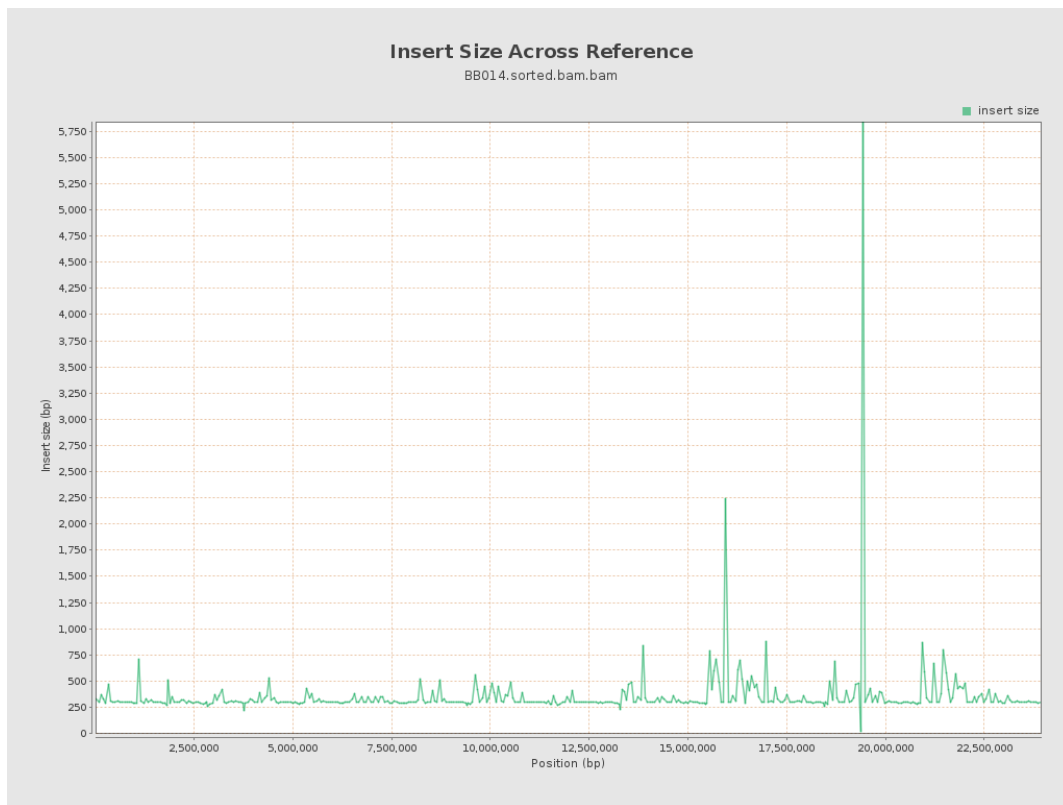
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

