

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 11:30:32

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/BB012.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/BB012-BiooBarcode_31_CACGAT_R2.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/BB012-BiooBarcode_31_CACGAT_R1.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 11:30:31 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/BB012.sorted.bam.b

am

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	24,404,928
Mapped reads	15,952,952 / 65.37%
Unmapped reads	8,451,976 / 34.63%
Mapped paired reads	15,952,952 / 65.37%
Mapped reads, first in pair	7,931,571 / 32.5%
Mapped reads, second in pair	8,021,381 / 32.87%
Mapped reads, both in pair	15,433,717 / 63.24%
Mapped reads, singletons	519,235 / 2.13%
Read min/max/mean length	30 / 100 / 99.86
Duplicated reads (estimated)	7,289,761 / 29.87%
Duplication rate	45.48%
Clipped reads	1,215,262 / 4.98%

2.2. ACGT Content

Number/percentage of A's	433,294,642 / 27.85%
Number/percentage of C's	344,823,911 / 22.16%
Number/percentage of T's	434,874,396 / 27.95%
Number/percentage of G's	343,042,985 / 22.05%
Number/percentage of N's	131,164 / 0.01%
GC Percentage	44.21%

2.3. Coverage

Mean	65.0118
Standard Deviation	58.5047

2.4. Mapping Quality

Mean Mapping Quality	58.69
----------------------	-------

2.5. Insert size

Mean	1,113.08
Standard Deviation	30,353.78
P25/Median/P75	318 / 362 / 409

2.6. Mismatches and indels

General error rate	0.8%
Mismatches	11,417,482
Insertions	317,525
Mapped reads with at least one insertion	1.9%
Deletions	440,186
Mapped reads with at least one deletion	2.64%
Homopolymer indels	71.28%

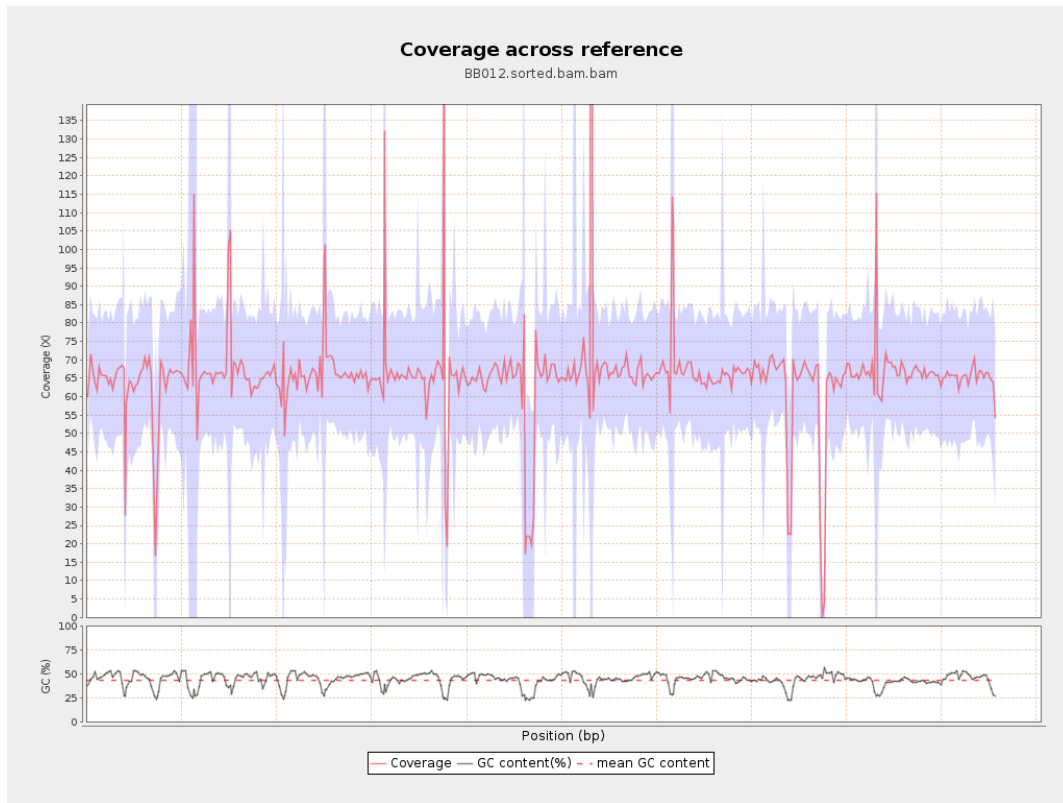
2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

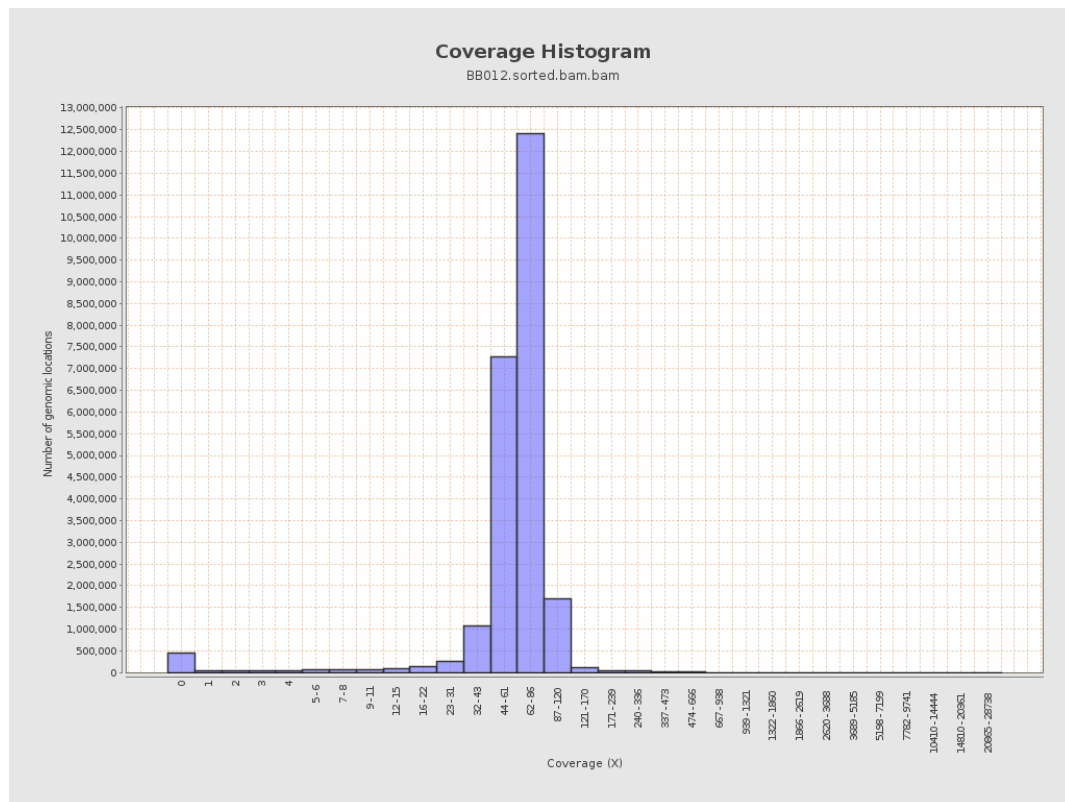
gi 1074120478 emb LT615256.1	977217	64196230	65.6929	16.8334
gi 1074120682 emb LT615257.1	860454	50559260	58.7588	27.2456
gi 1074120865 emb LT615258.1	989719	64955951	65.6307	46.9726
gi 1074121086 emb LT615259.1	935450	63240098	67.6039	62.1049
gi 1074121301 emb LT615260.1	1432239	93836717	65.5175	46.8276
gi 1074121615 emb LT615261.1	1080962	71138000	65.8099	29.5517
gi 1074121871 emb LT615262.1	1545099	102195615	66.1418	16.828
gi 1074122235 emb LT615263.1	1585108	105924918	66.825	30.4441
gi 1074122590 emb LT615264.1	2122358	135115846	63.6631	23.4083
gi 1074123050 emb LT615265.1	1754192	105220186	59.9821	40.6575
gi 1074123421 emb LT615	2150147	150069197	69.7949	166.1736

266.1				
gi 107412389 8 emb LT615 267.1	3031036	201742140	66.5588	21.4956
gi 107412458 8 emb LT615 268.1	2359348	143220225	60.7033	38.4848
gi 107412506 5 emb LT615 269.1	3135668	206204130	65.7608	15.7569

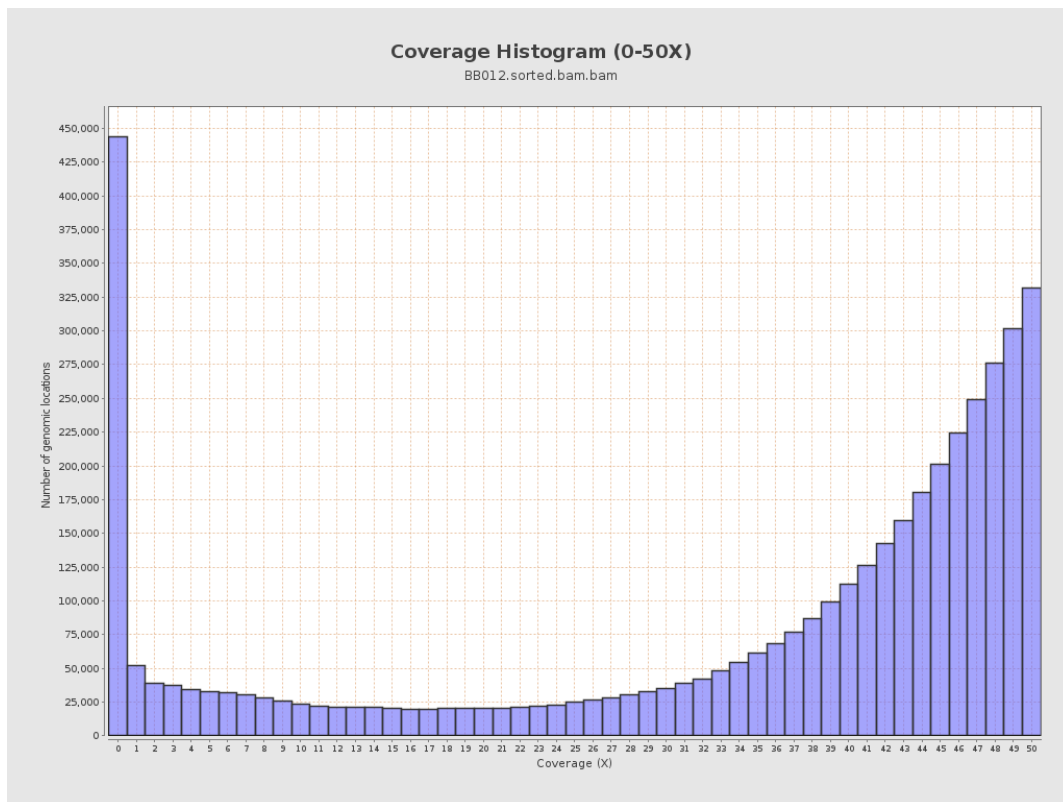
3. Results : Coverage across reference



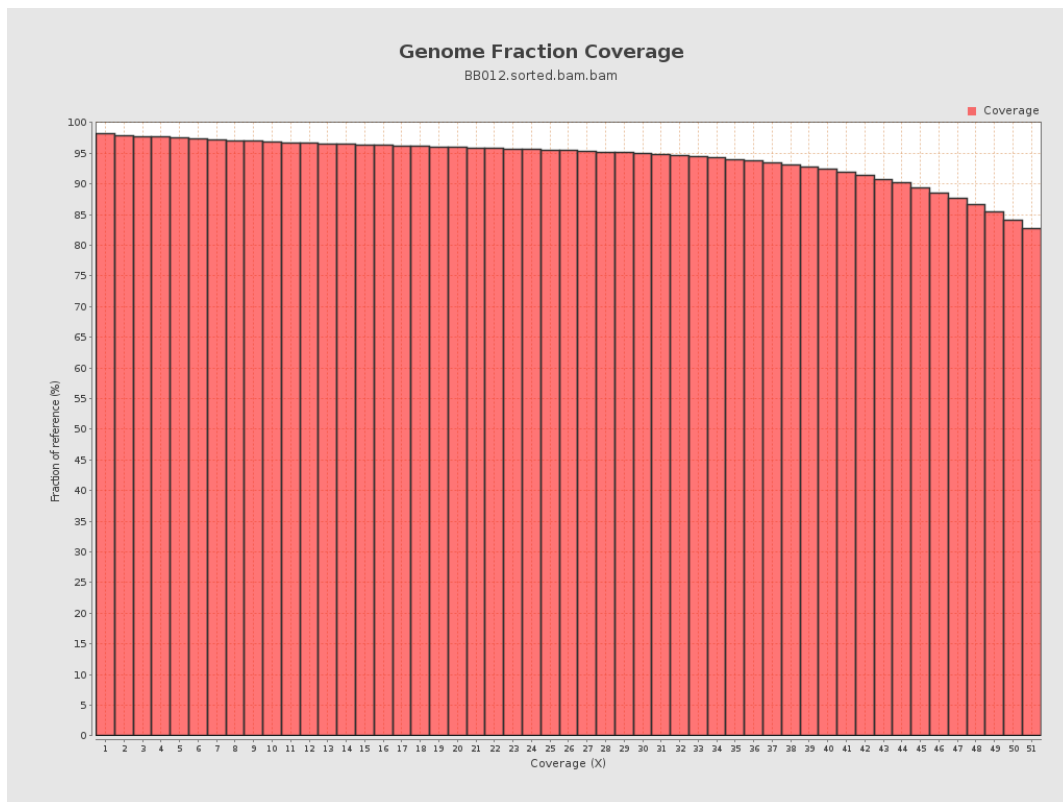
4. Results : Coverage Histogram



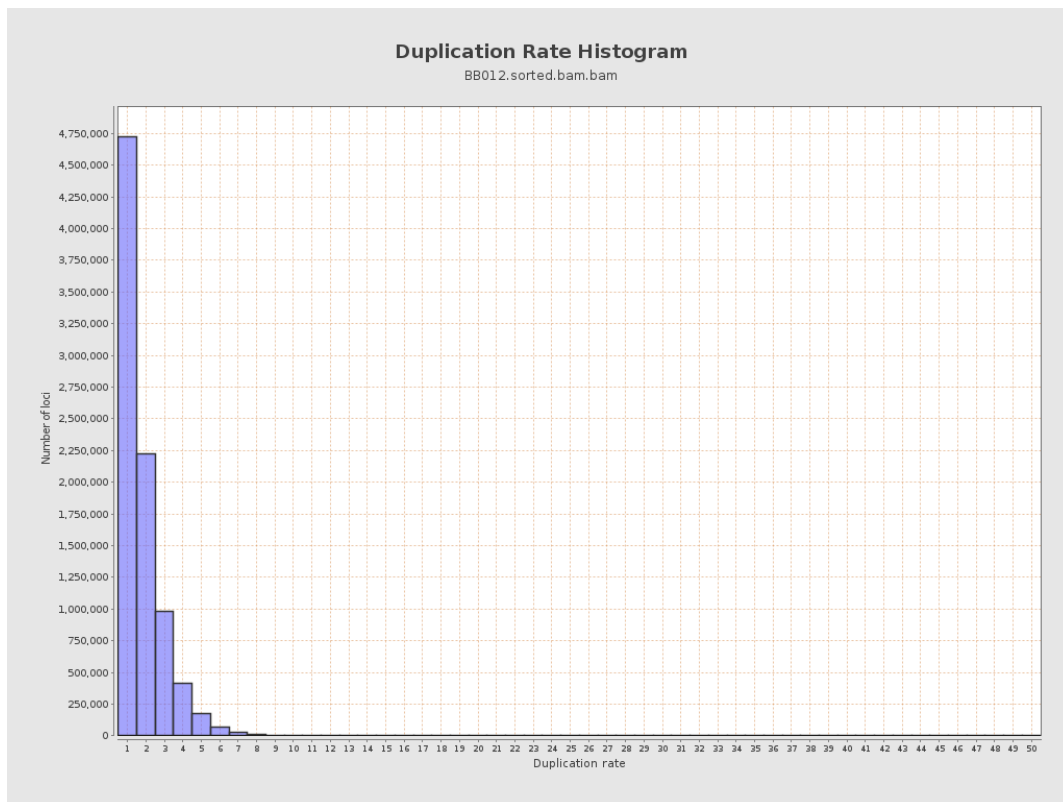
5. Results : Coverage Histogram (0-50X)



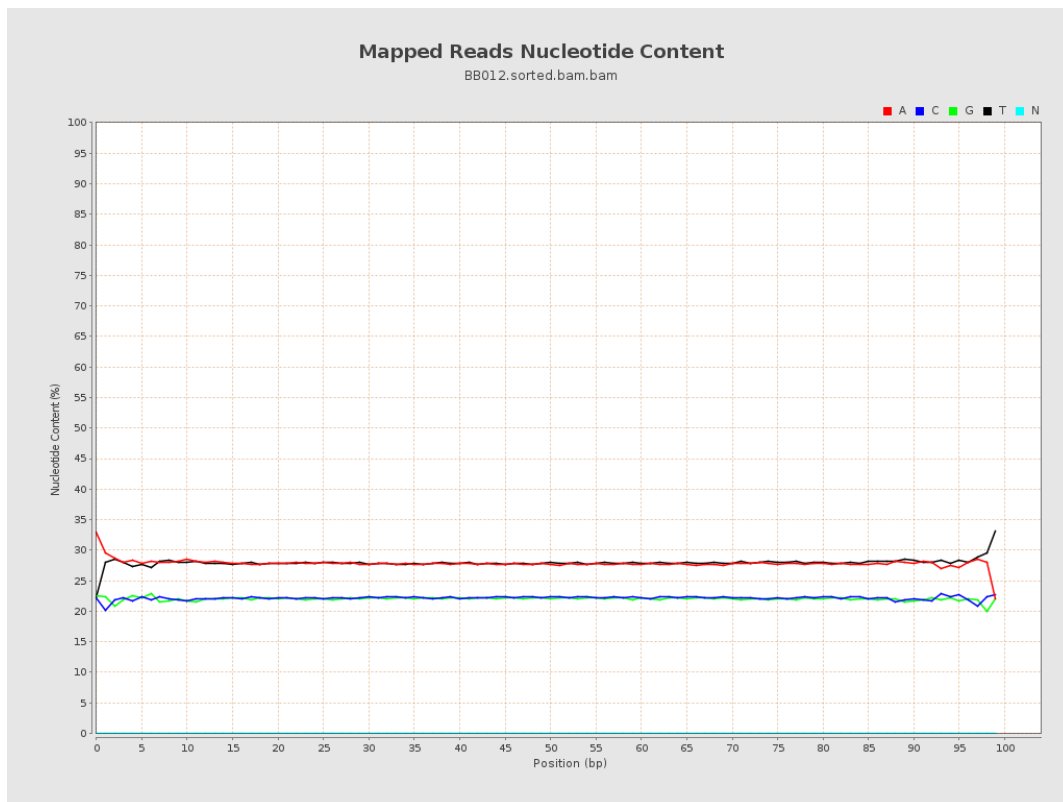
6. Results : Genome Fraction Coverage



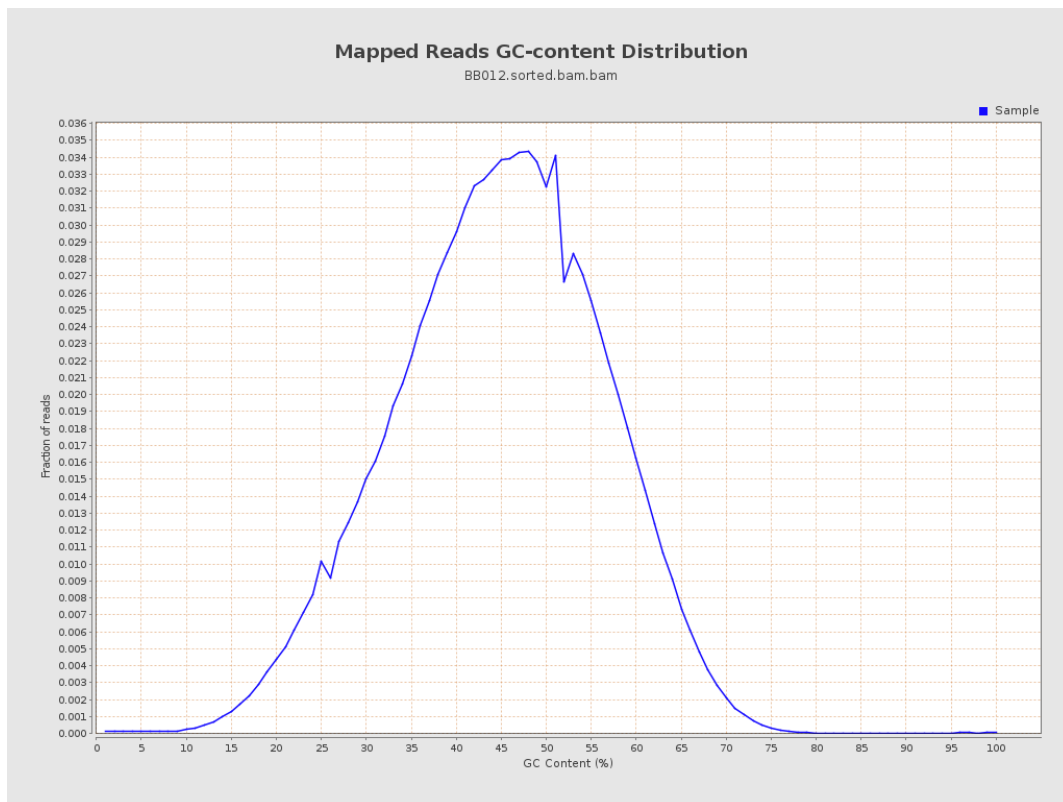
7. Results : Duplication Rate Histogram



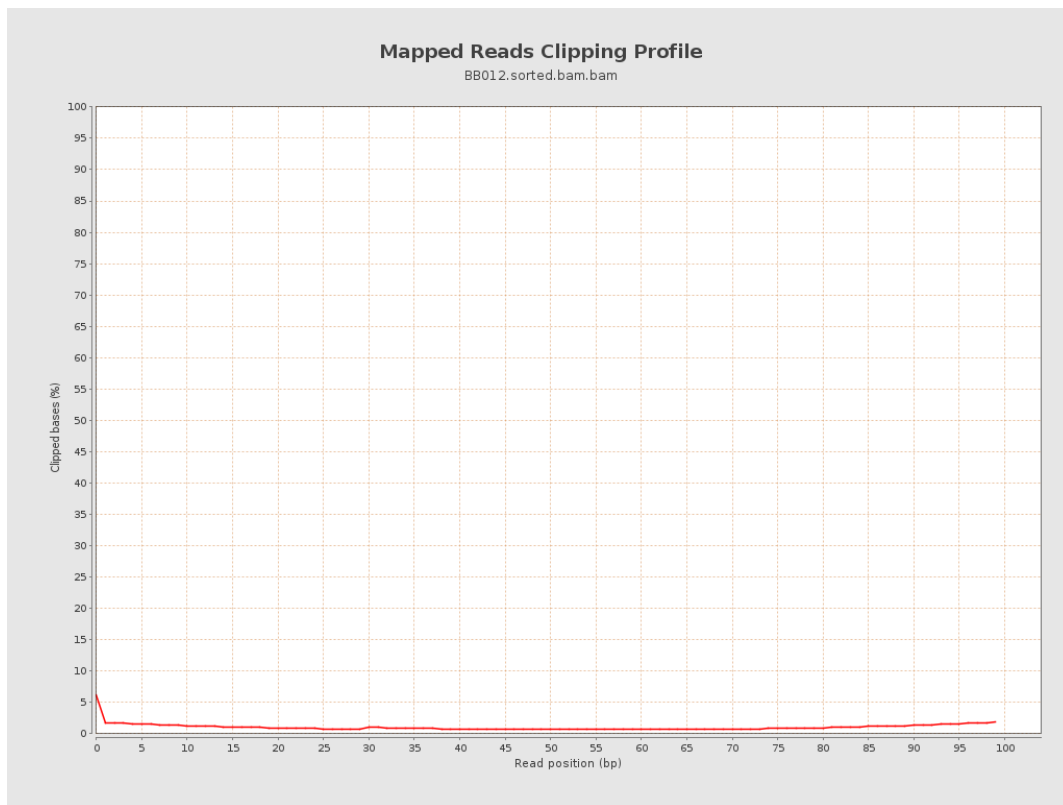
8. Results : Mapped Reads Nucleotide Content



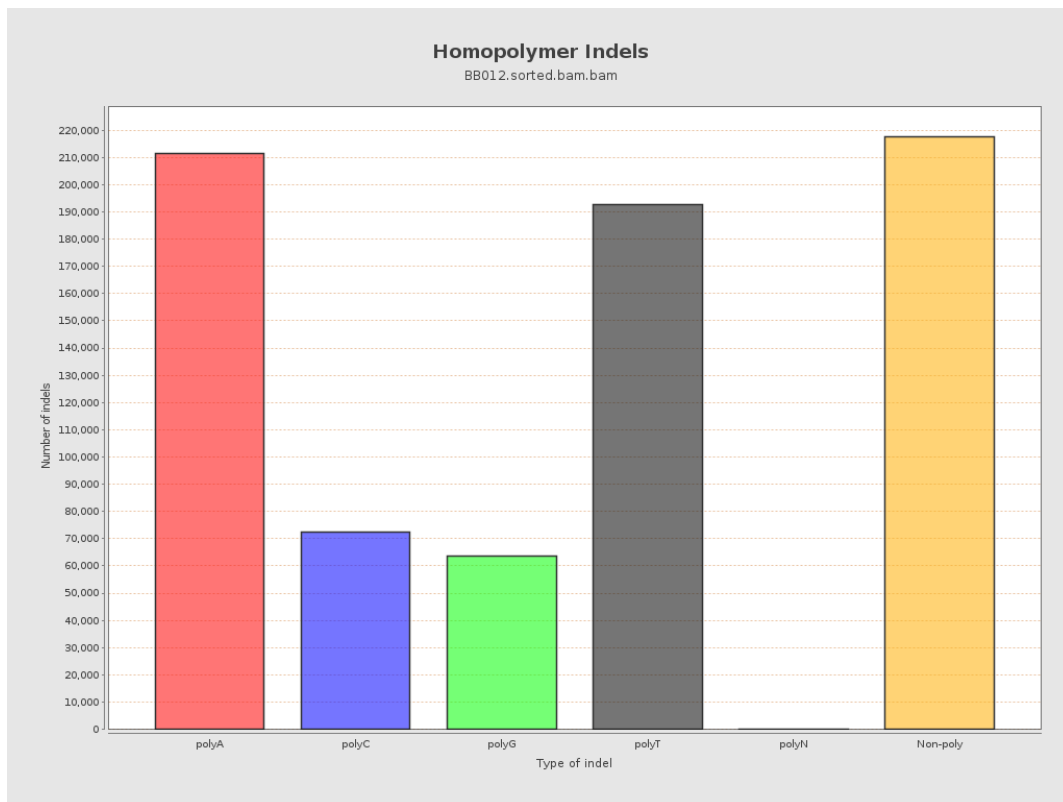
9. Results : Mapped Reads GC-content Distribution



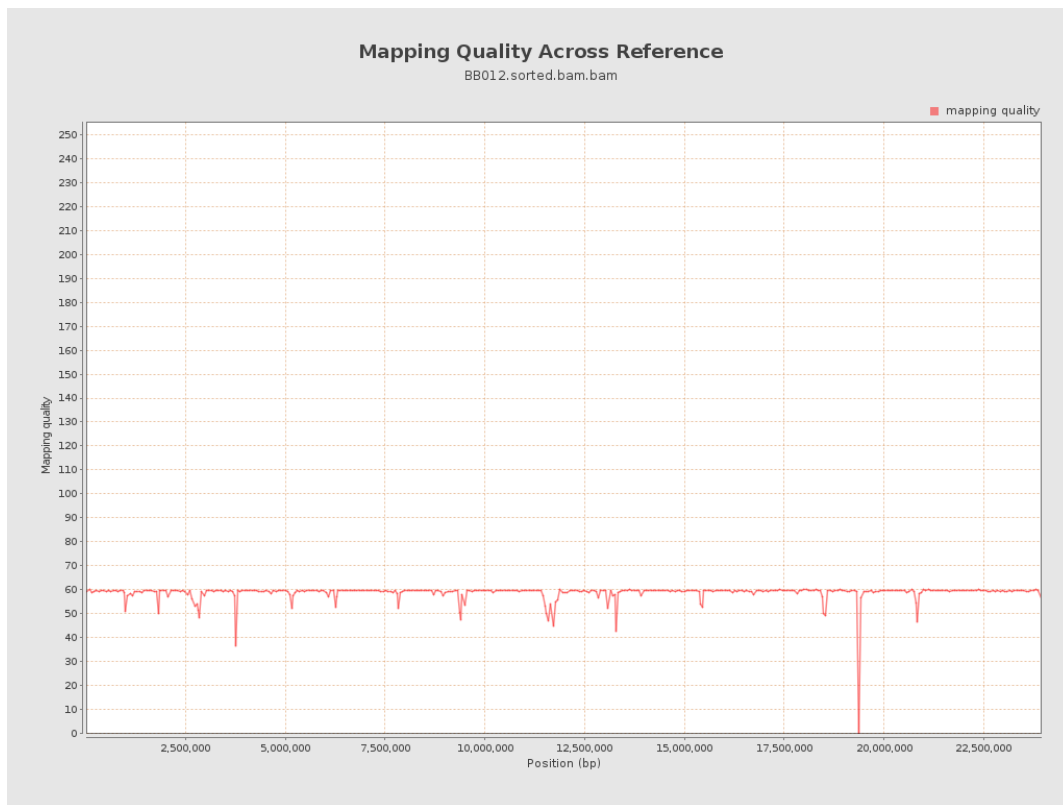
10. Results : Mapped Reads Clipping Profile



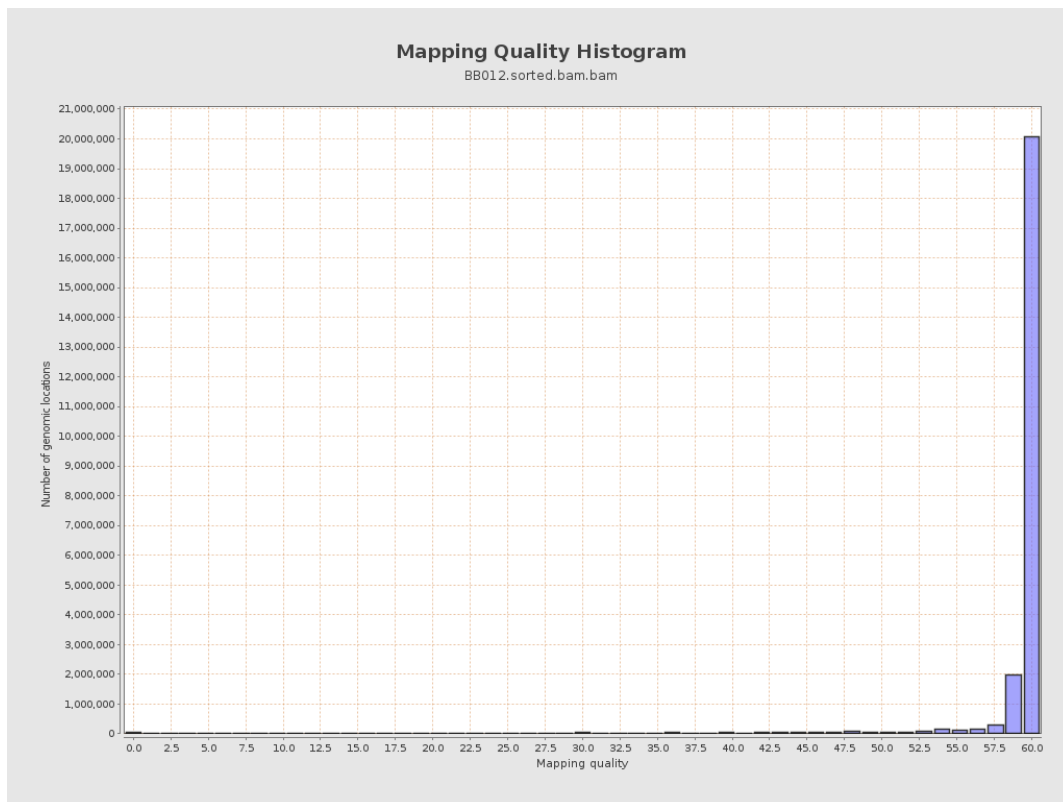
11. Results : Homopolymer Indels



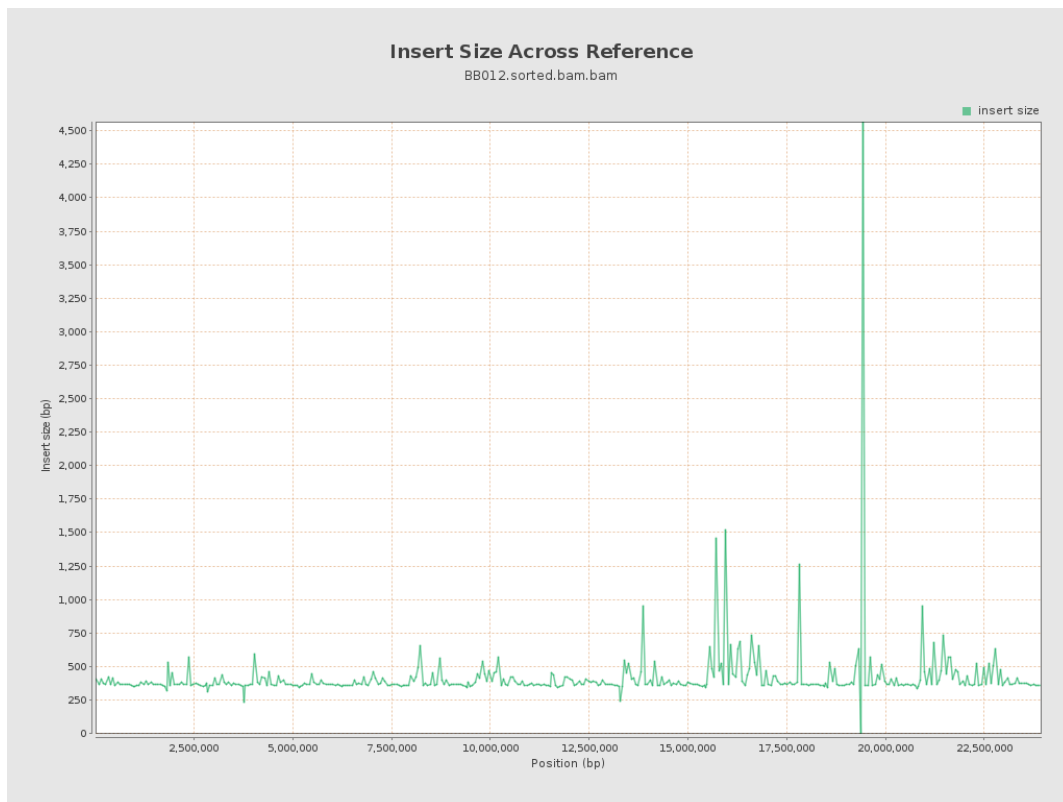
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

