

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 11:34:55

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/BB019.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/BB019-BiooBarcode_35_CATTTT_R1.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/BB019-BiooBarcode_35_CATTTT_R2.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 11:34:55 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/BB019.sorted.bam.b

am

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	17,792,305
Mapped reads	8,716,423 / 48.99%
Unmapped reads	9,075,882 / 51.01%
Mapped paired reads	8,716,423 / 48.99%
Mapped reads, first in pair	4,385,057 / 24.65%
Mapped reads, second in pair	4,331,366 / 24.34%
Mapped reads, both in pair	8,505,558 / 47.8%
Mapped reads, singletons	210,865 / 1.19%
Read min/max/mean length	30 / 100 / 99.89
Duplicated reads (estimated)	4,538,334 / 25.51%
Duplication rate	52.8%
Clipped reads	653,115 / 3.67%

2.2. ACGT Content

Number/percentage of A's	234,560,750 / 27.61%
Number/percentage of C's	190,046,644 / 22.37%
Number/percentage of T's	235,506,763 / 27.73%
Number/percentage of G's	189,296,582 / 22.29%
Number/percentage of N's	73,258 / 0.01%
GC Percentage	44.66%

2.3. Coverage

Mean	35.4874
Standard Deviation	25.9598

2.4. Mapping Quality

Mean Mapping Quality	58.63
----------------------	-------

2.5. Insert size

Mean	819.87
Standard Deviation	25,494.14
P25/Median/P75	313 / 354 / 397

2.6. Mismatches and indels

General error rate	0.74%
Mismatches	5,725,157
Insertions	165,347
Mapped reads with at least one insertion	1.82%
Deletions	226,861
Mapped reads with at least one deletion	2.5%
Homopolymer indels	71.71%

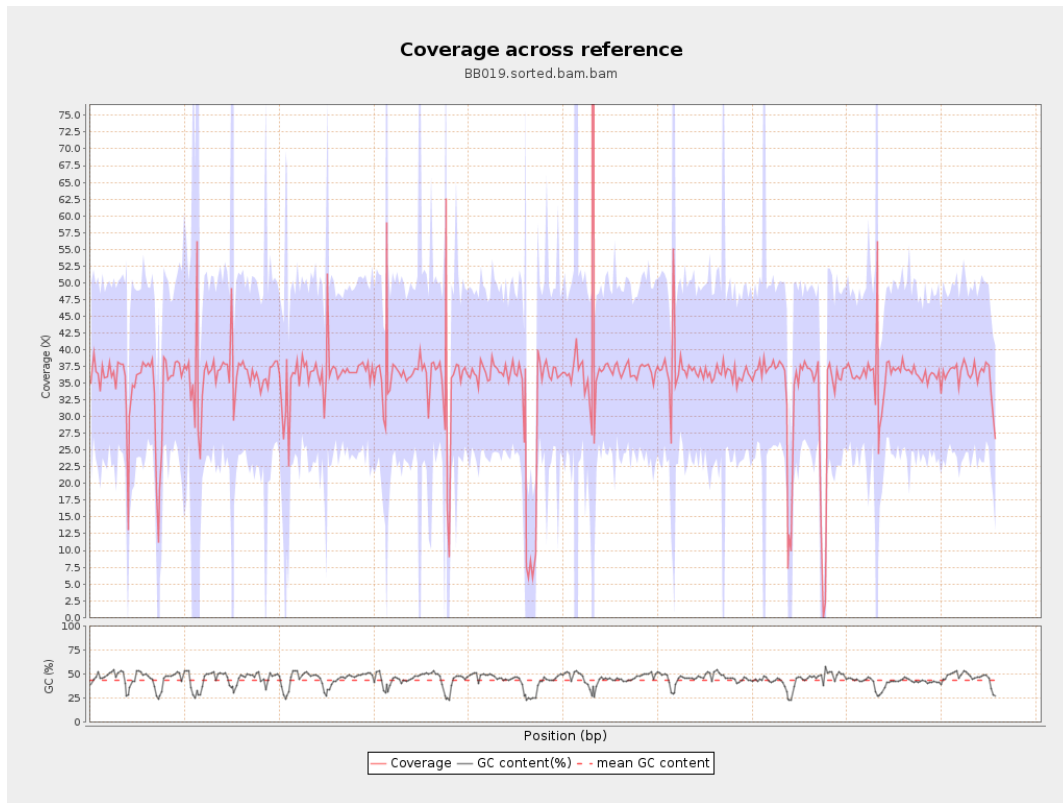
2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

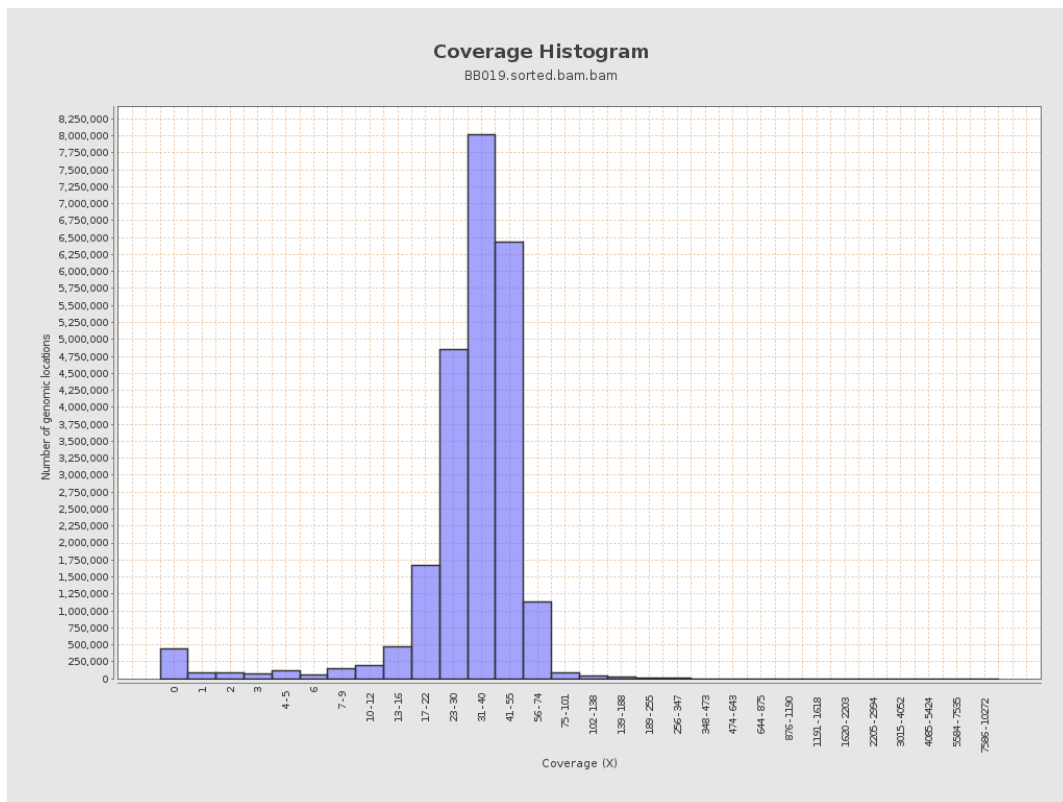
gi 1074120478 emb LT615256.1	977217	35647997	36.4791	12.0651
gi 1074120682 emb LT615257.1	860454	28009261	32.5517	16.7814
gi 1074120865 emb LT615258.1	989719	35085523	35.45	23.1048
gi 1074121086 emb LT615259.1	935450	33746513	36.0752	25.2381
gi 1074121301 emb LT615260.1	1432239	50820307	35.4831	21.4585
gi 1074121615 emb LT615261.1	1080962	38463238	35.5824	14.3523
gi 1074121871 emb LT615262.1	1545099	56371643	36.4842	12.1266
gi 1074122235 emb LT615263.1	1585108	57948913	36.5583	19.2915
gi 1074122590 emb LT615264.1	2122358	74432852	35.0708	14.4192
gi 1074123050 emb LT615265.1	1754192	56090809	31.9753	34.8491
gi 1074123421 emb LT615	2150147	81568478	37.9362	61.4225

266.1				
gi 107412389 8 emb LT615 267.1	3031036	110976822	36.6135	16.4767
gi 107412458 8 emb LT615 268.1	2359348	78448454	33.2501	20.032
gi 107412506 5 emb LT615 269.1	3135668	112631677	35.9195	11.5993

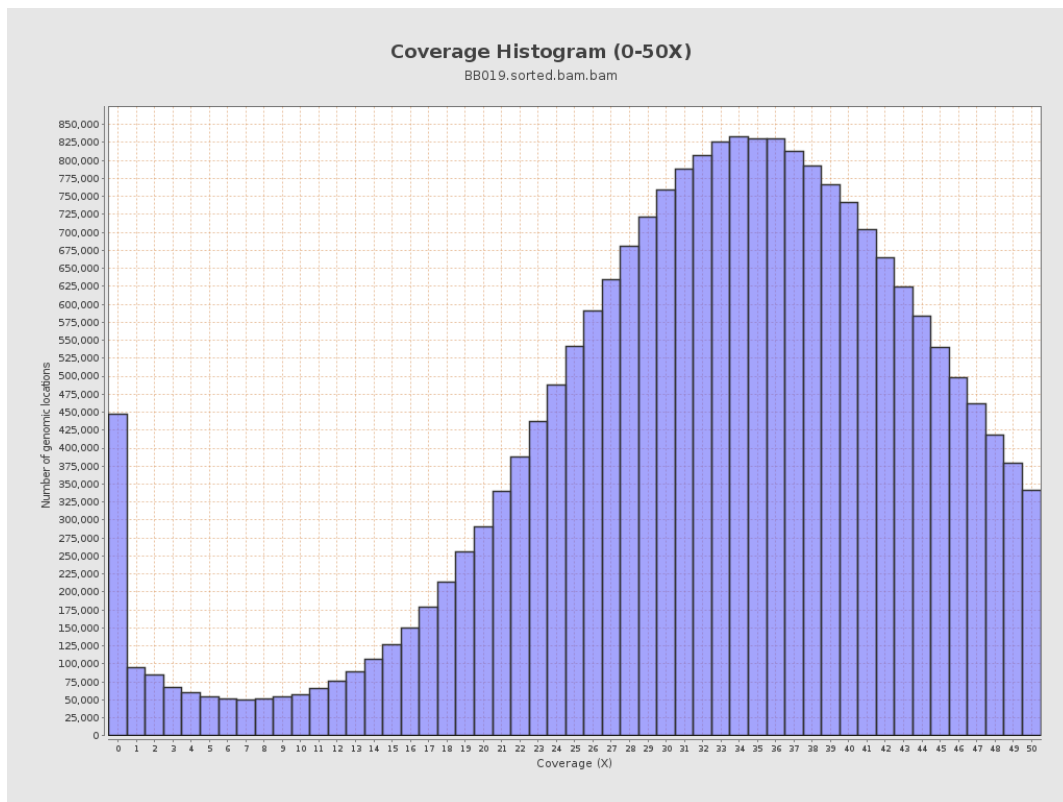
3. Results : Coverage across reference



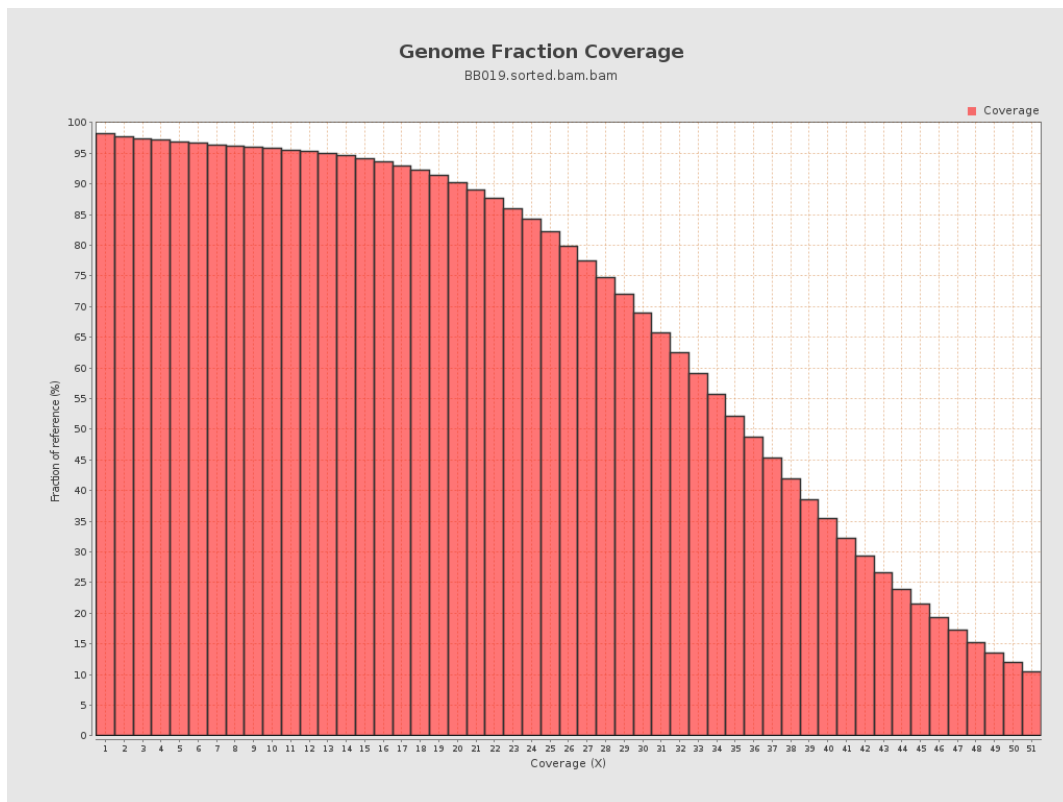
4. Results : Coverage Histogram



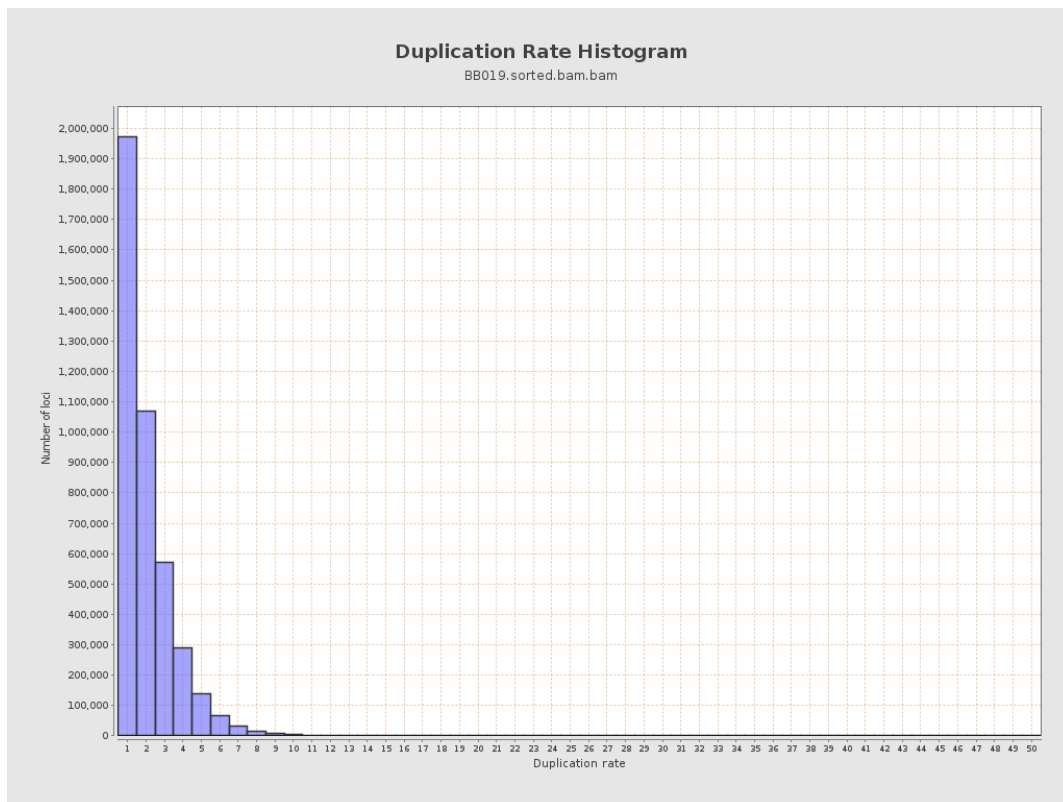
5. Results : Coverage Histogram (0-50X)



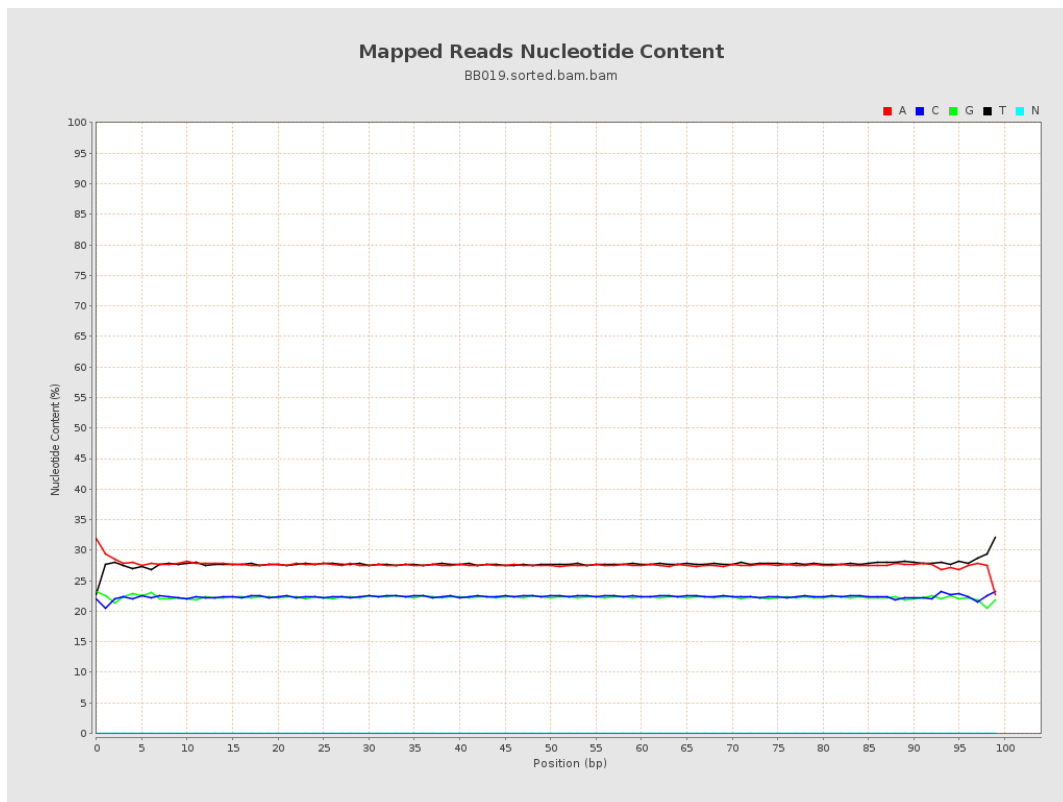
6. Results : Genome Fraction Coverage



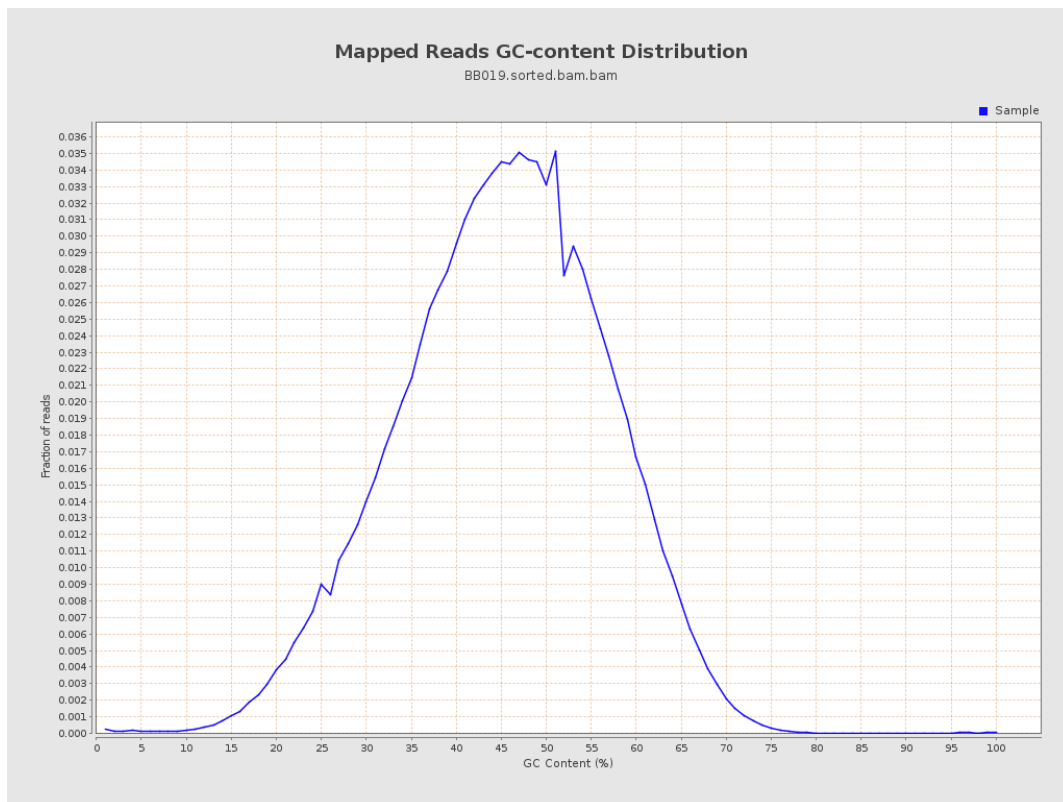
7. Results : Duplication Rate Histogram



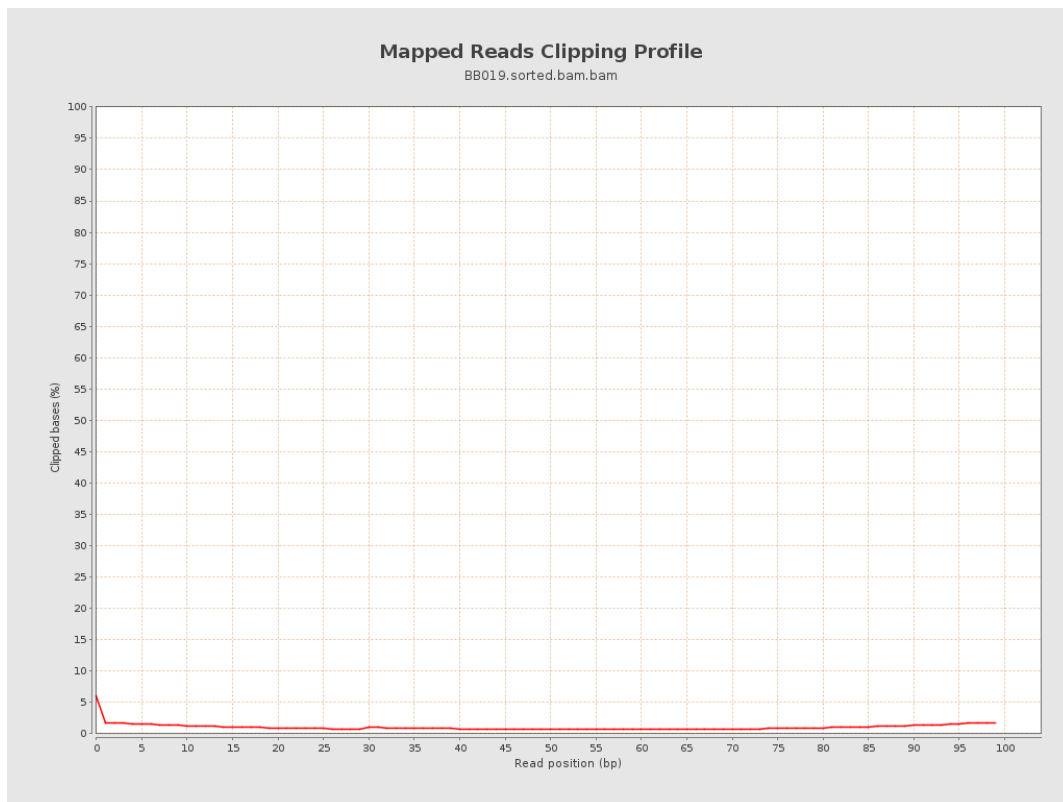
8. Results : Mapped Reads Nucleotide Content



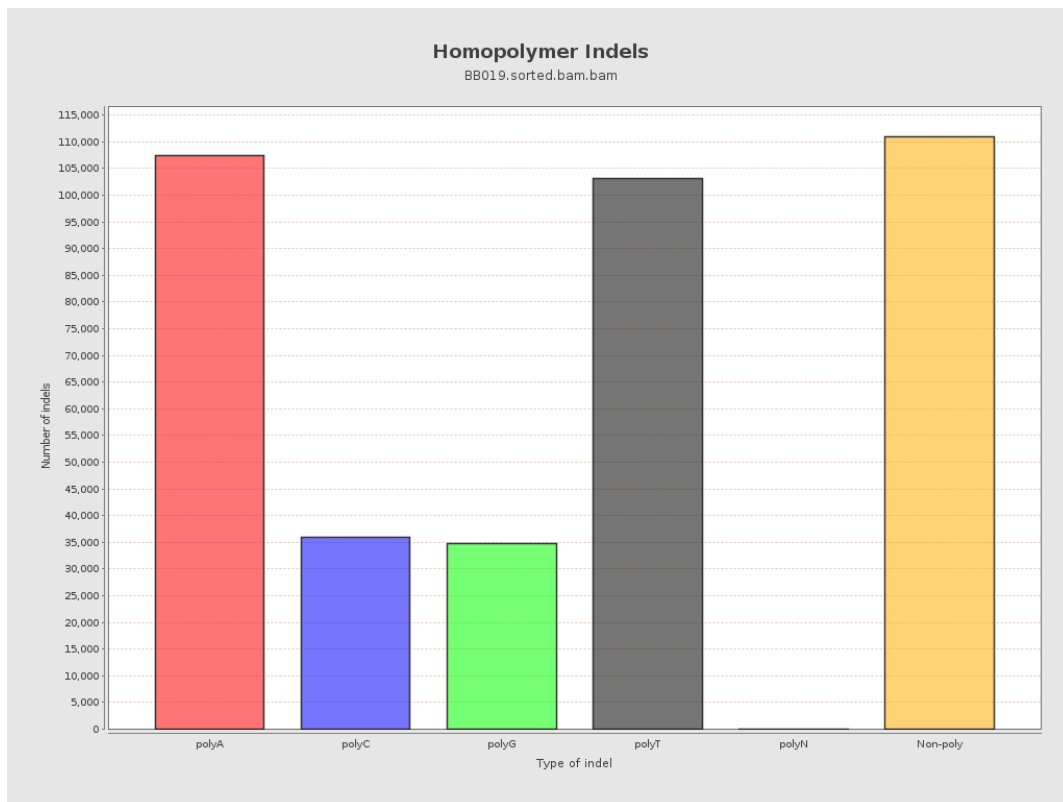
9. Results : Mapped Reads GC-content Distribution



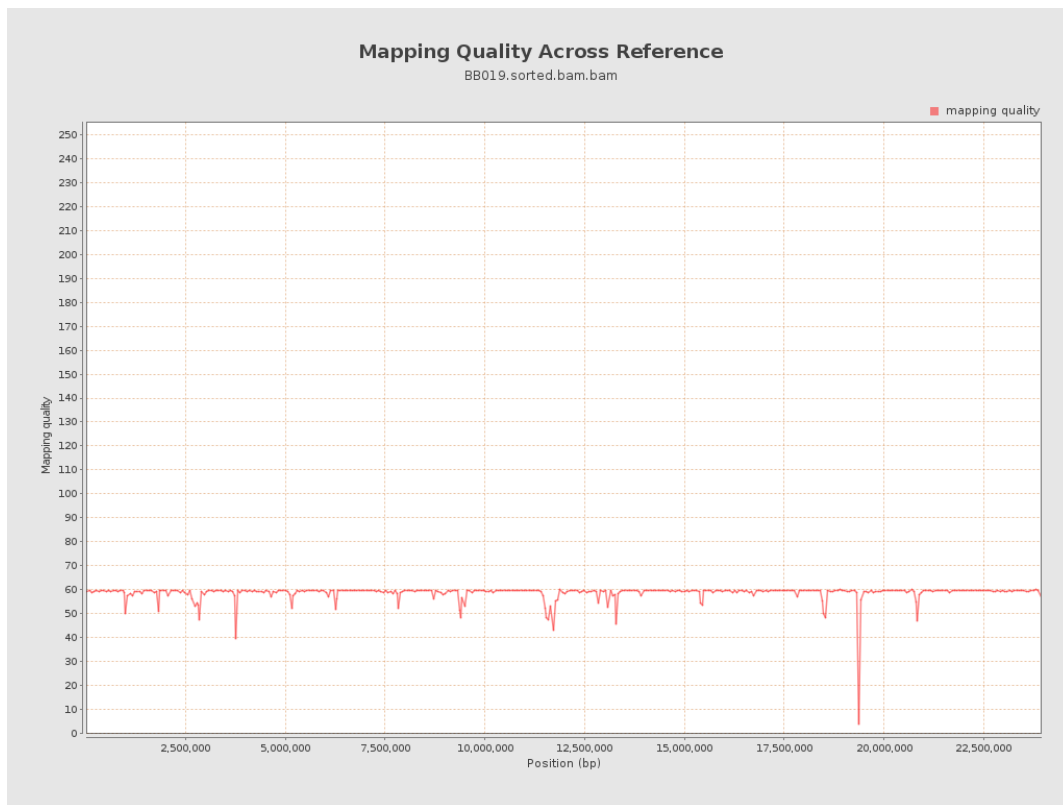
10. Results : Mapped Reads Clipping Profile



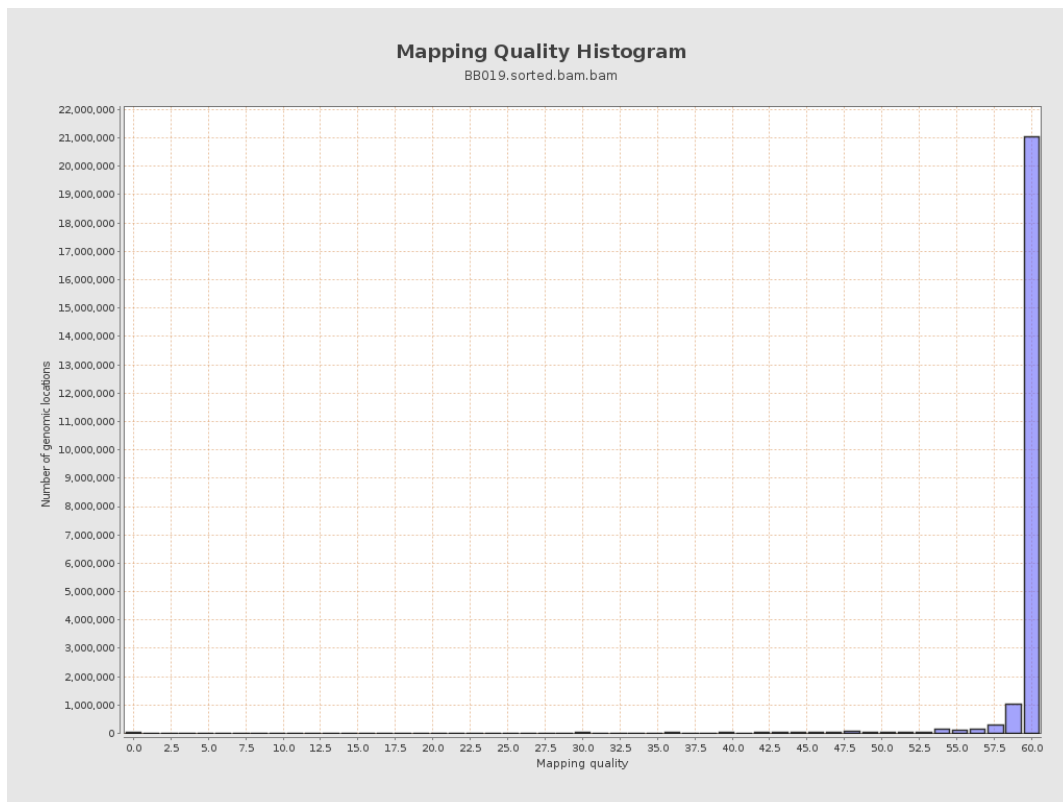
11. Results : Homopolymer Indels



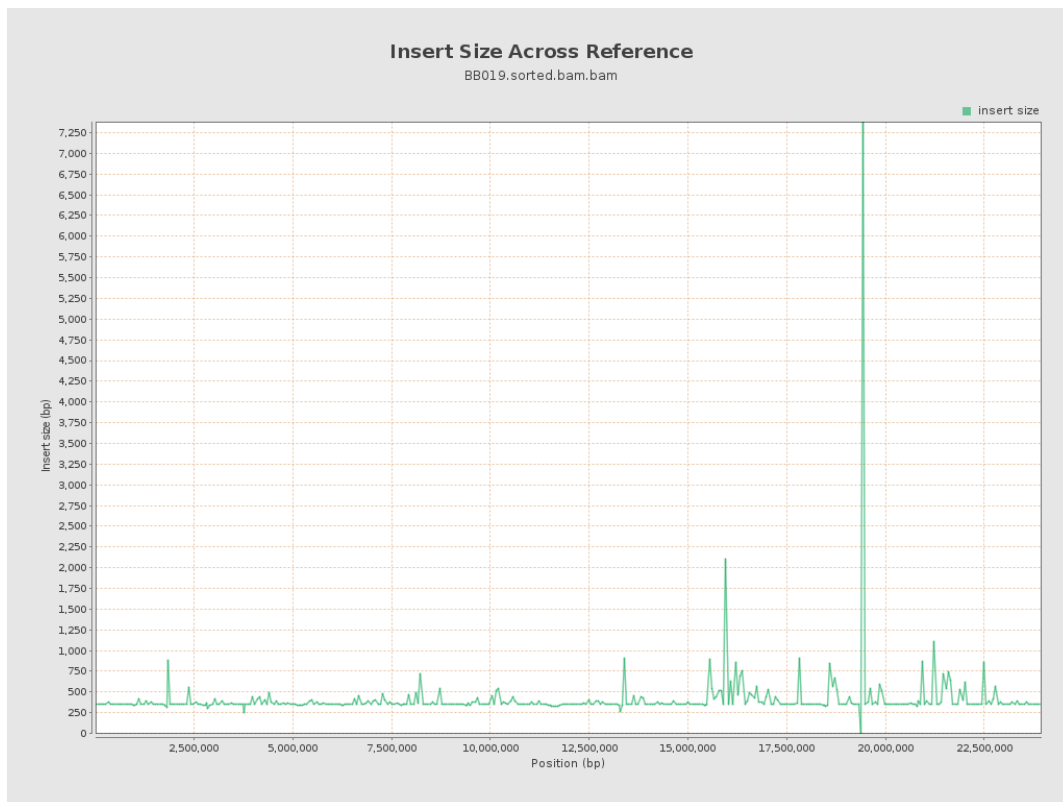
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

