

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 11:52:24

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/KP068.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/KP068-BiooBarcode_28_CAAAAG_R2.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/KP068-BiooBarcode_28_CAAAAG_R1.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 11:52:24 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/KP068.sorted.bam.b

am

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	15,361,644
Mapped reads	6,810,757 / 44.34%
Unmapped reads	8,550,887 / 55.66%
Mapped paired reads	6,810,757 / 44.34%
Mapped reads, first in pair	3,383,584 / 22.03%
Mapped reads, second in pair	3,427,173 / 22.31%
Mapped reads, both in pair	6,622,633 / 43.11%
Mapped reads, singletons	188,124 / 1.22%
Read min/max/mean length	30 / 100 / 99.91
Duplicated reads (estimated)	3,122,784 / 20.33%
Duplication rate	46.34%
Clipped reads	730,023 / 4.75%

2.2. ACGT Content

Number/percentage of A's	188,346,052 / 28.75%
Number/percentage of C's	138,965,658 / 21.21%
Number/percentage of T's	189,092,397 / 28.86%
Number/percentage of G's	138,790,878 / 21.18%
Number/percentage of N's	53,280 / 0.01%
GC Percentage	42.39%

2.3. Coverage

Mean	27.3757
Standard Deviation	22.0061

2.4. Mapping Quality

Mean Mapping Quality	58.61
----------------------	-------

2.5. Insert size

Mean	1,043.4
Standard Deviation	30,044.58
P25/Median/P75	284 / 377 / 461

2.6. Mismatches and indels

General error rate	1.06%
Mismatches	6,506,877
Insertions	164,237
Mapped reads with at least one insertion	2.29%
Deletions	217,037
Mapped reads with at least one deletion	3.03%
Homopolymer indels	65.68%

2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

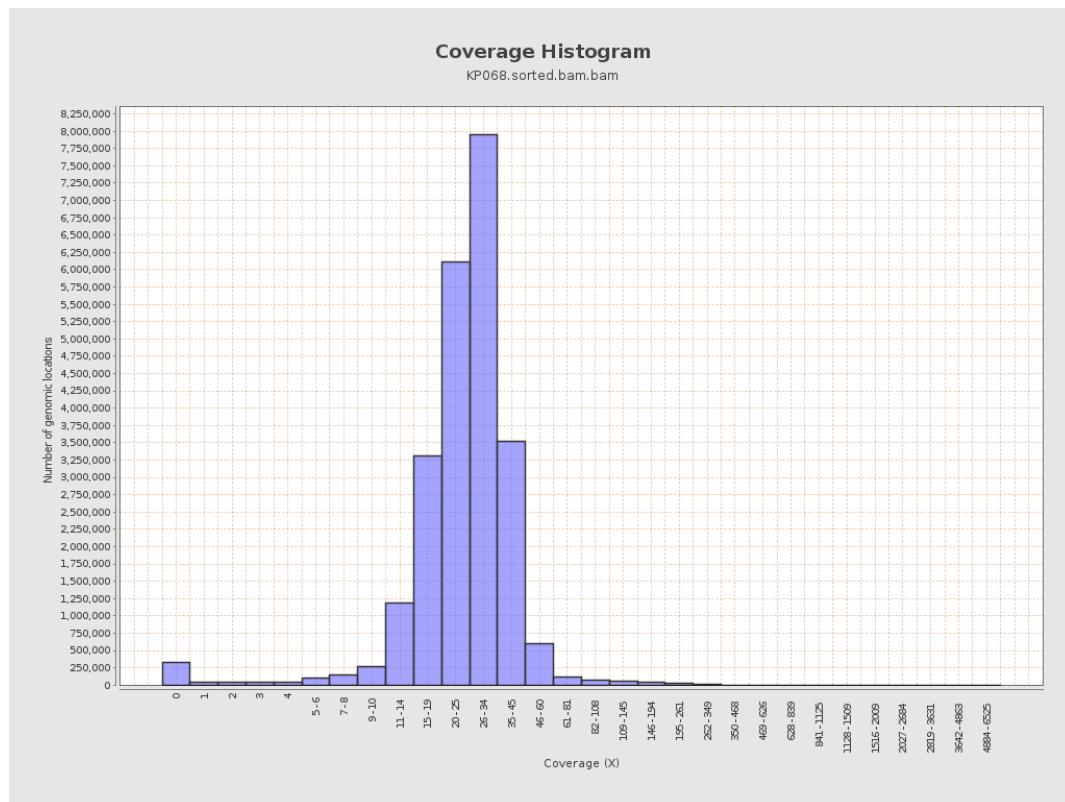
gi 1074120478 emb LT615256.1	977217	25355945	25.9471	10.7048
gi 1074120682 emb LT615257.1	860454	23571657	27.3944	16.4101
gi 1074120865 emb LT615258.1	989719	29819218	30.129	27.2278
gi 1074121086 emb LT615259.1	935450	28140009	30.0818	31.8379
gi 1074121301 emb LT615260.1	1432239	41045414	28.6582	23.9411
gi 1074121615 emb LT615261.1	1080962	28964945	26.7955	16.431
gi 1074121871 emb LT615262.1	1545099	42201618	27.3132	10.2952
gi 1074122235 emb LT615263.1	1585108	43631557	27.5259	18.9383
gi 1074122590 emb LT615264.1	2122358	56968176	26.8419	22.8569
gi 1074123050 emb LT615265.1	1754192	46414794	26.4594	31.646
gi 1074123421 emb LT615	2150147	61160893	28.445	37.6892

266.1				
gi 107412389 8 emb LT615 267.1	3031036	83126665	27.4252	15.5896
gi 107412458 8 emb LT615 268.1	2359348	59278255	25.1248	17.68
gi 107412506 5 emb LT615 269.1	3135668	86214739	27.4949	9.8221

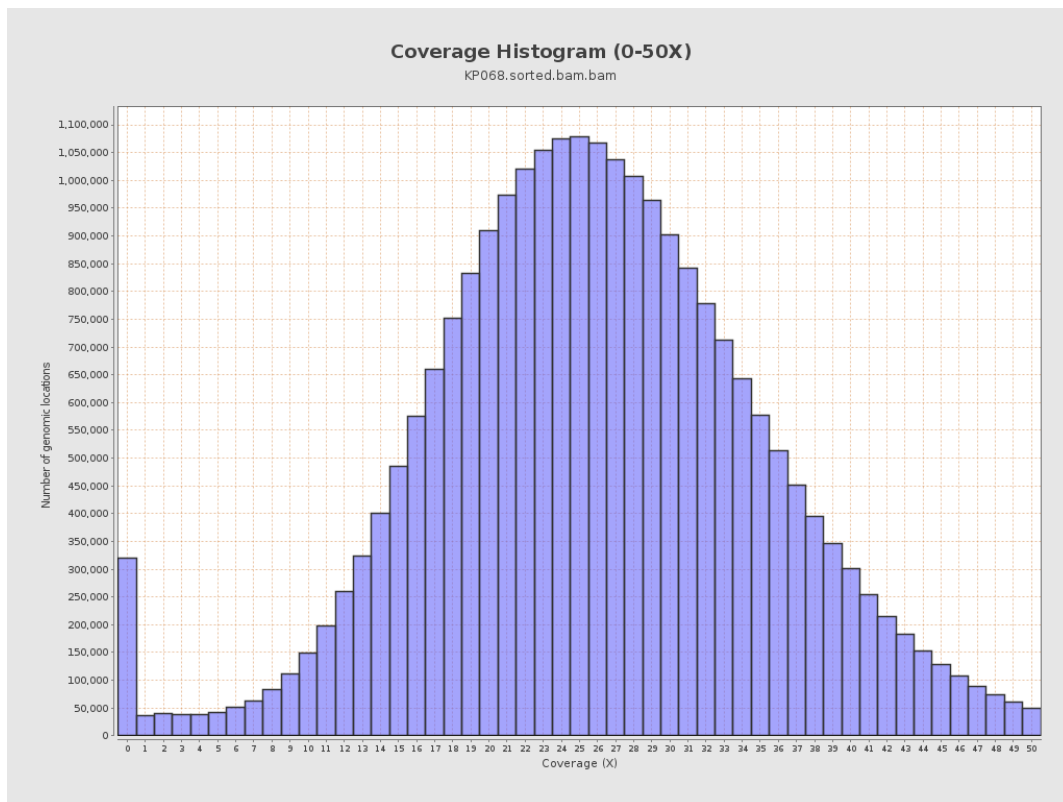
3. Results : Coverage across reference



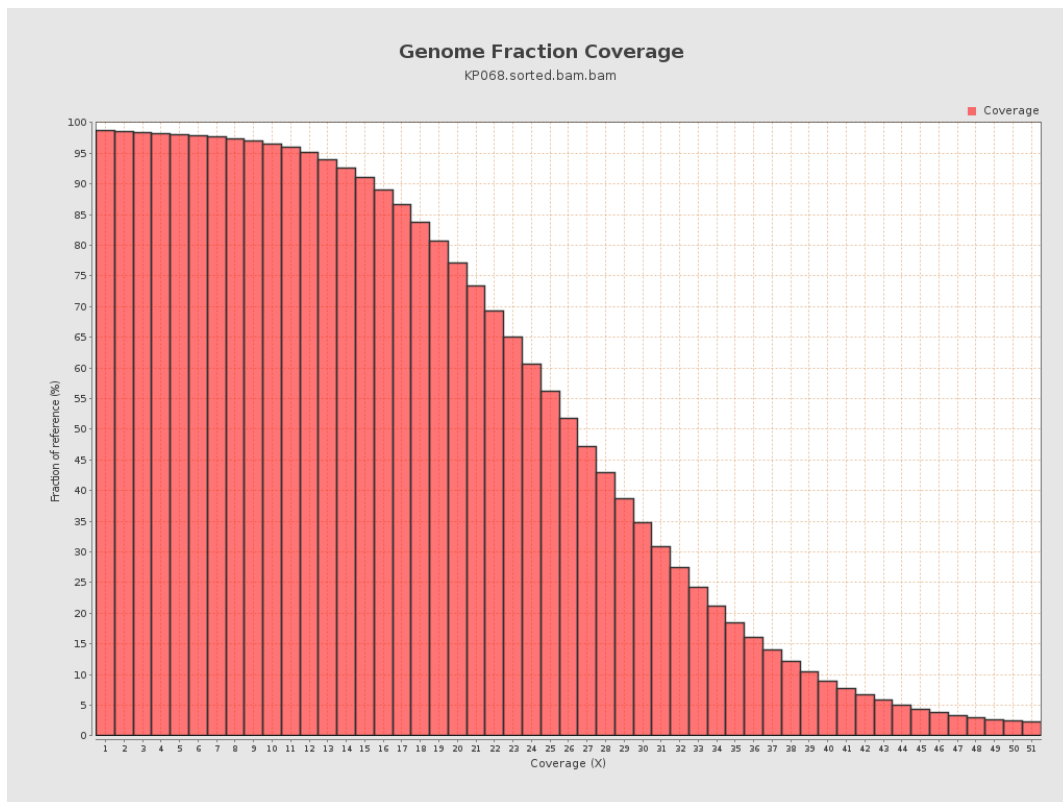
4. Results : Coverage Histogram



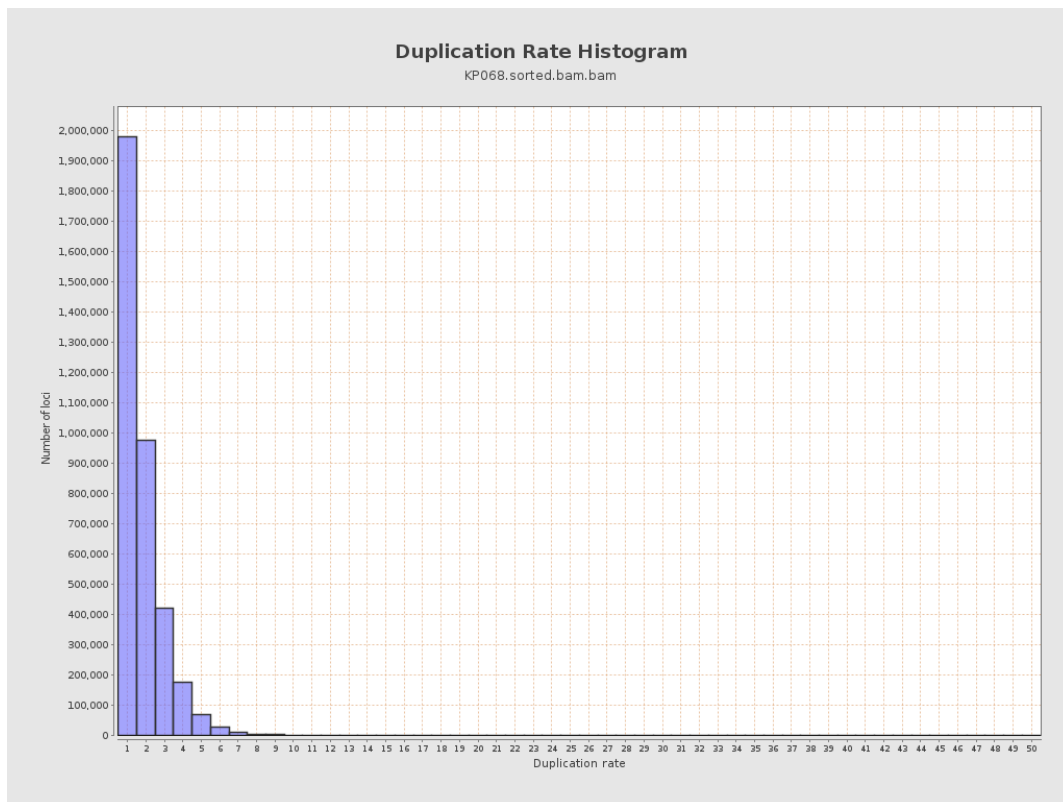
5. Results : Coverage Histogram (0-50X)



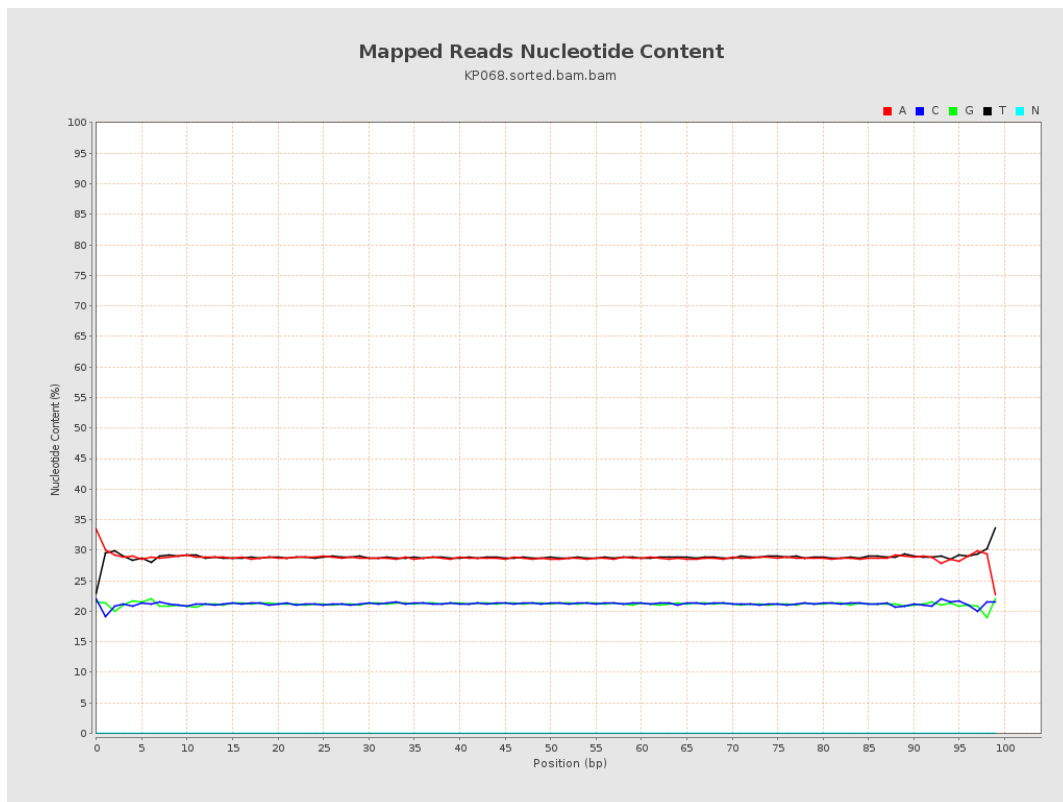
6. Results : Genome Fraction Coverage



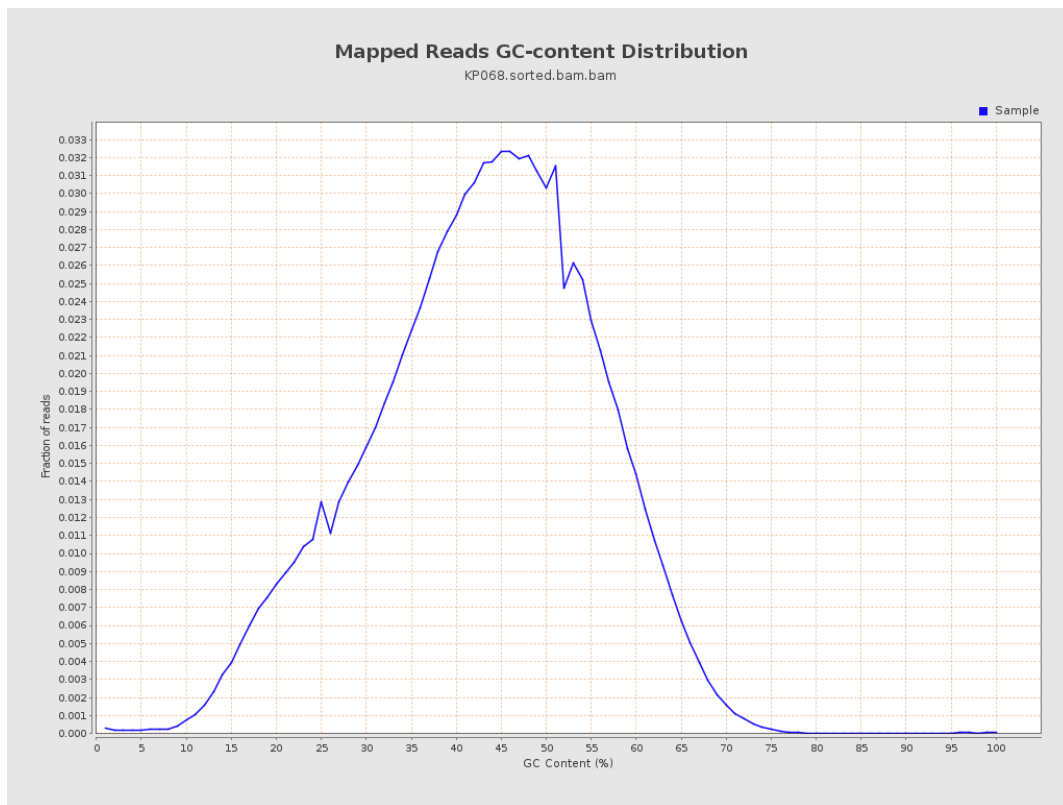
7. Results : Duplication Rate Histogram



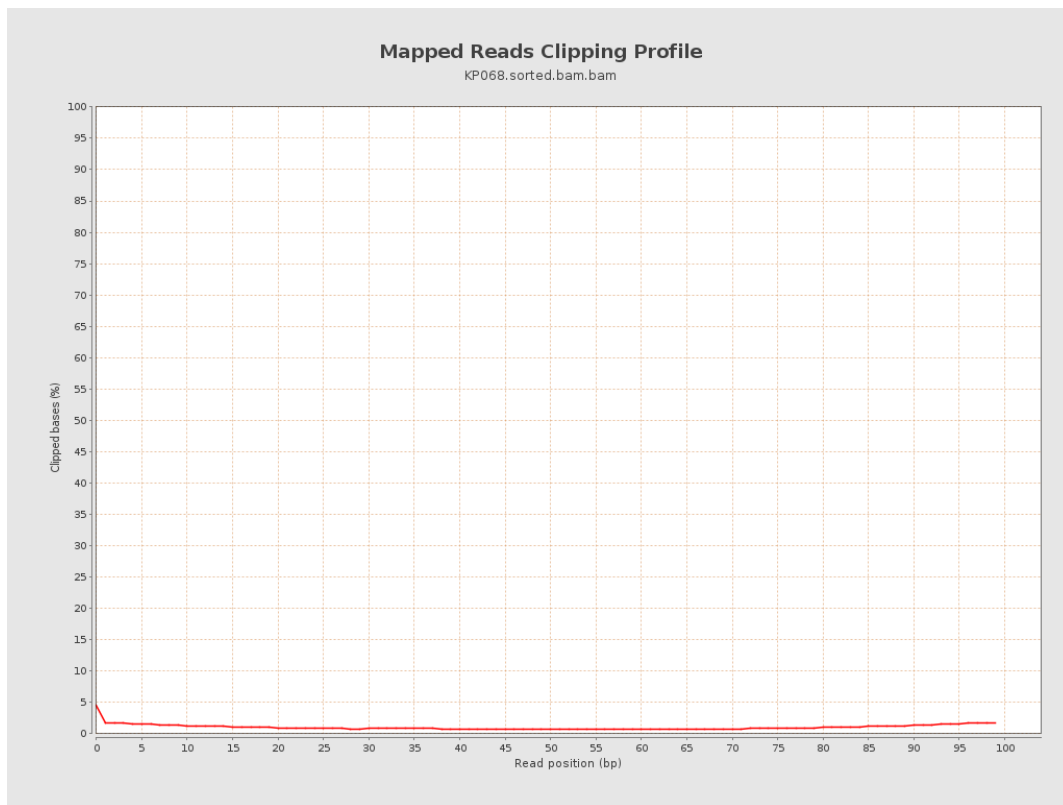
8. Results : Mapped Reads Nucleotide Content



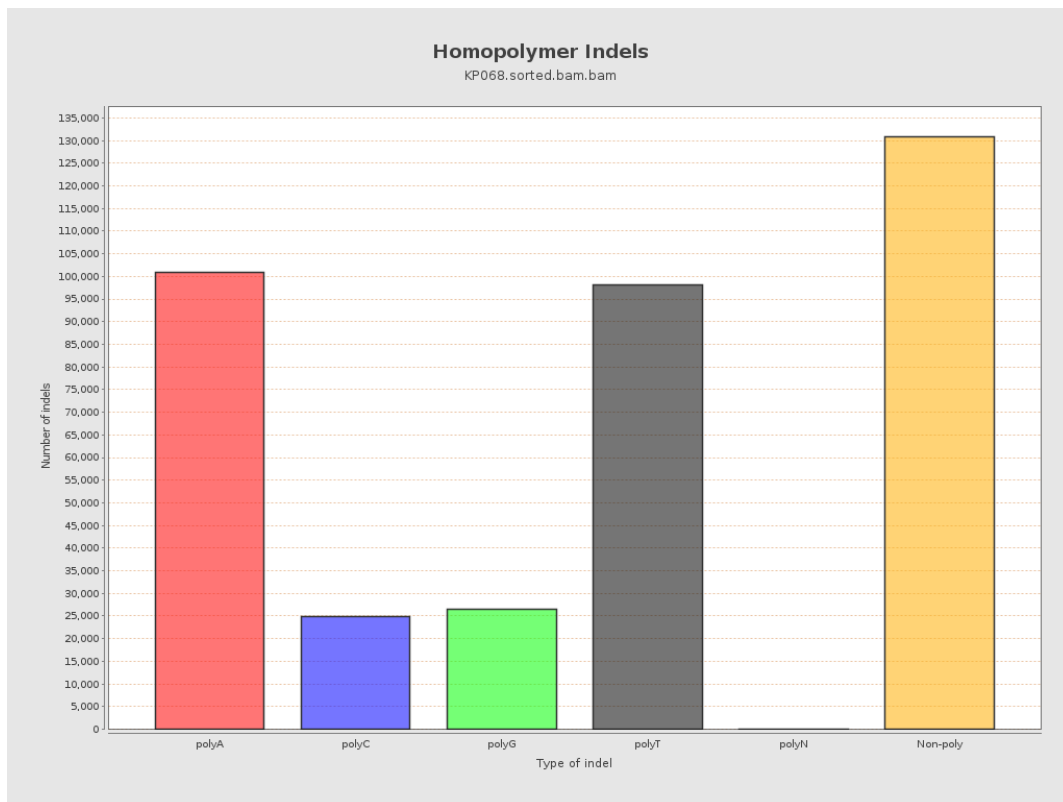
9. Results : Mapped Reads GC-content Distribution



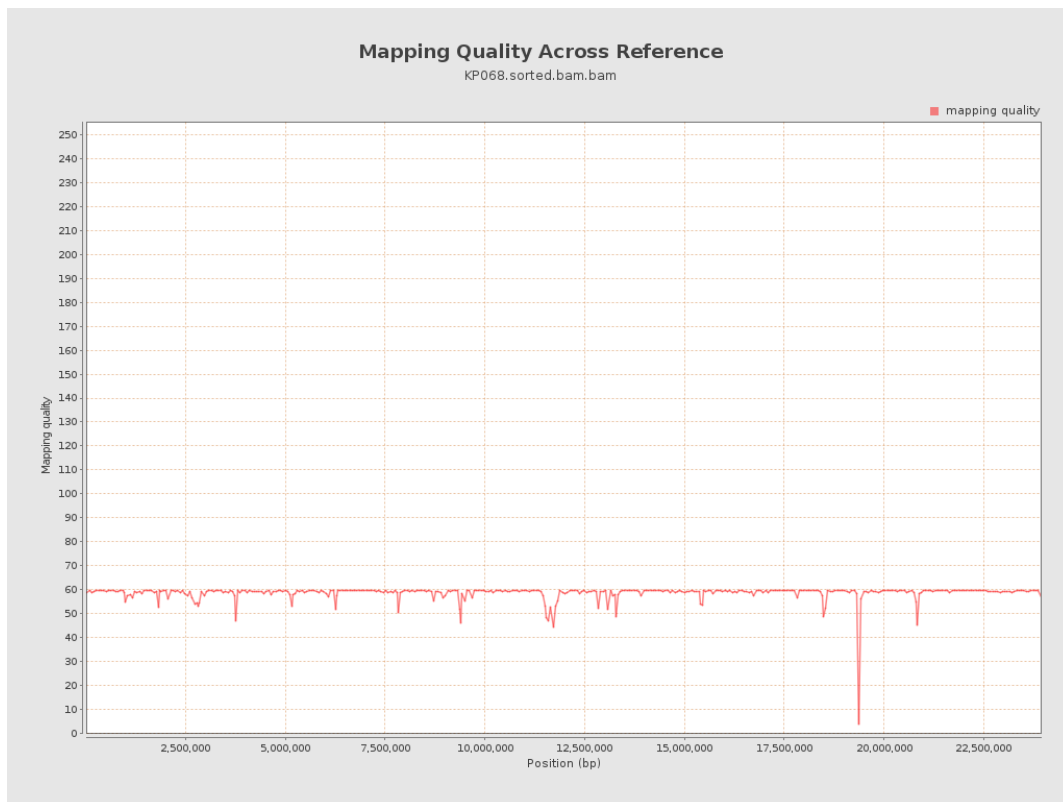
10. Results : Mapped Reads Clipping Profile



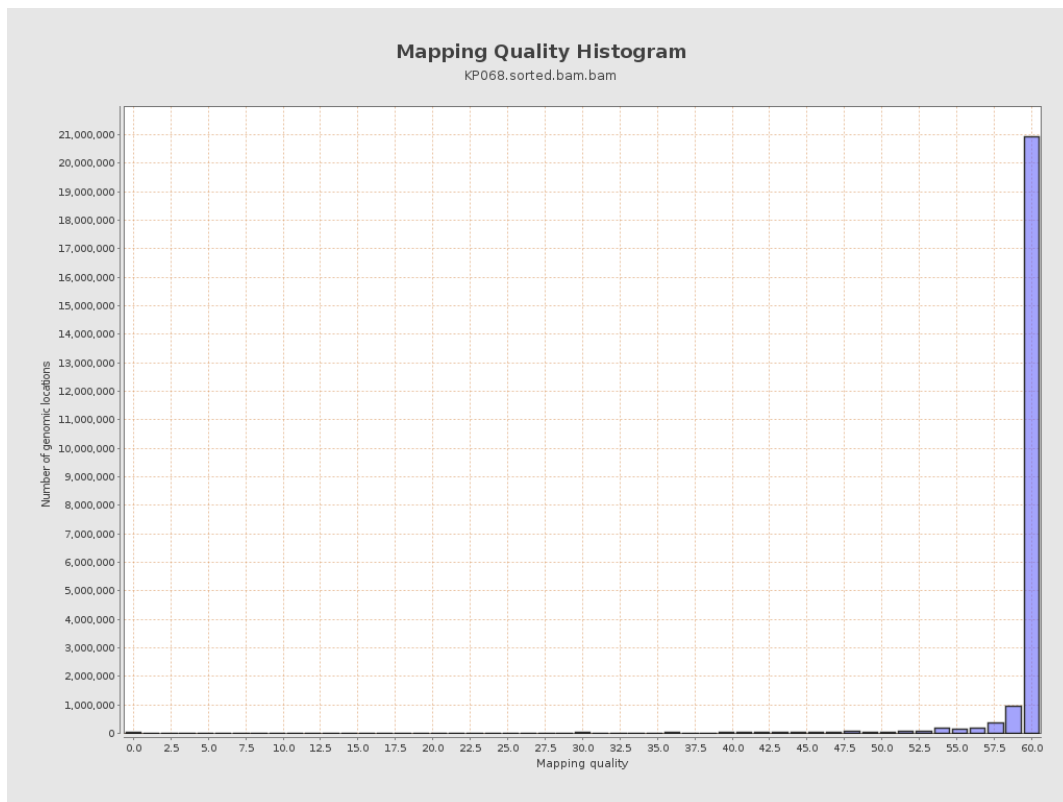
11. Results : Homopolymer Indels



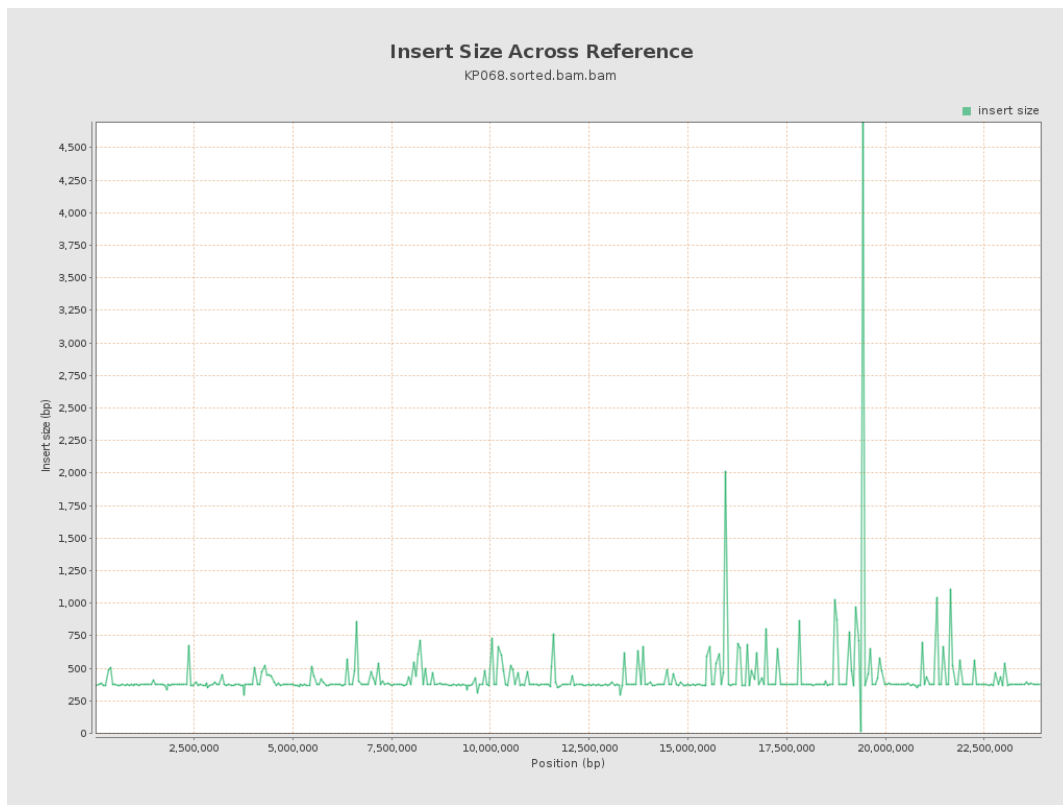
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

