

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 11:51:02

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/KP067.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/KP067-BiooBarcode_27_ATTCCT_R2.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/KP067-BiooBarcode_27_ATTCCT_R1.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 11:51:02 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/KP067.sorted.bam.b

am

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	19,065,137
Mapped reads	14,751,581 / 77.37%
Unmapped reads	4,313,556 / 22.63%
Mapped paired reads	14,751,581 / 77.37%
Mapped reads, first in pair	7,351,486 / 38.56%
Mapped reads, second in pair	7,400,095 / 38.81%
Mapped reads, both in pair	14,536,063 / 76.24%
Mapped reads, singletons	215,518 / 1.13%
Read min/max/mean length	30 / 100 / 99.85
Duplicated reads (estimated)	6,484,696 / 34.01%
Duplication rate	43.52%
Clipped reads	1,331,637 / 6.98%

2.2. ACGT Content

Number/percentage of A's	410,477,186 / 28.64%
Number/percentage of C's	305,897,577 / 21.35%
Number/percentage of T's	411,640,187 / 28.73%
Number/percentage of G's	305,007,323 / 21.28%
Number/percentage of N's	118,968 / 0.01%
GC Percentage	42.63%

2.3. Coverage

Mean	59.8749
Standard Deviation	40.7093

2.4. Mapping Quality

Mean Mapping Quality	58.86
----------------------	-------

2.5. Insert size

Mean	1,039.93
Standard Deviation	29,712.94
P25/Median/P75	271 / 339 / 412

2.6. Mismatches and indels

General error rate	1.01%
Mismatches	13,476,354
Insertions	355,328
Mapped reads with at least one insertion	2.29%
Deletions	463,026
Mapped reads with at least one deletion	2.99%
Homopolymer indels	65.74%

2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

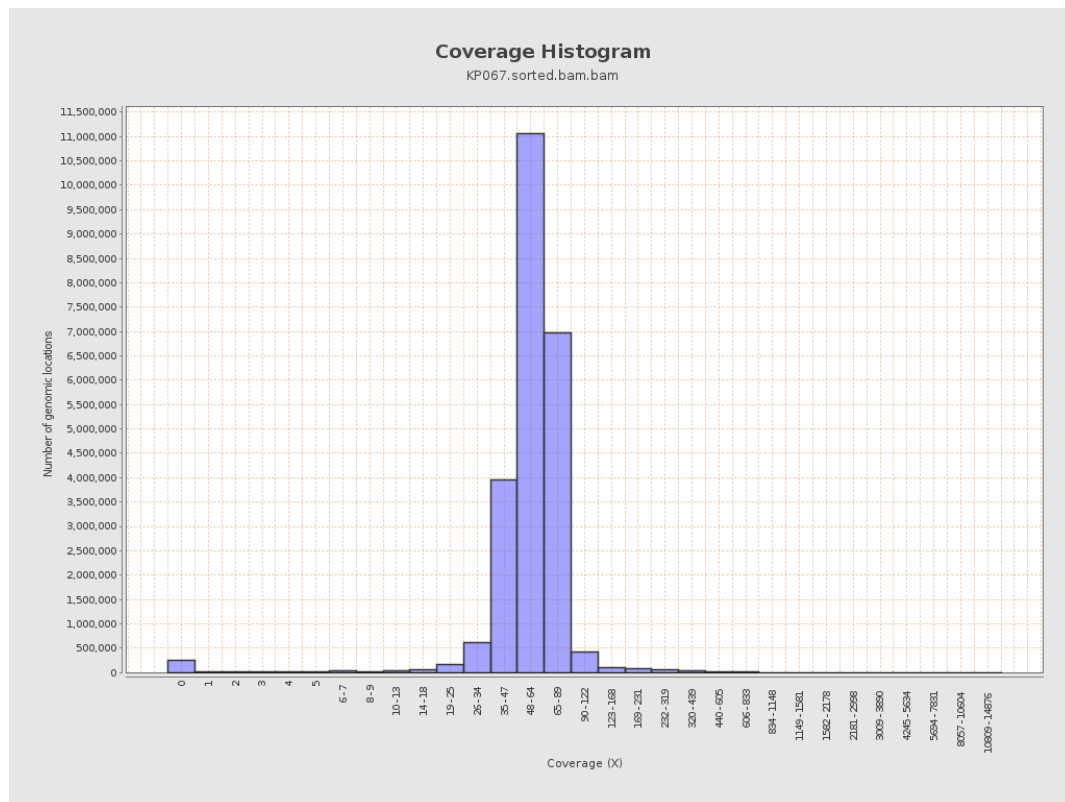
gi 1074120478 emb LT615256.1	977217	55805644	57.1067	22.7918
gi 1074120682 emb LT615257.1	860454	49167440	57.1413	26.5738
gi 1074120865 emb LT615258.1	989719	61108117	61.7429	47.9017
gi 1074121086 emb LT615259.1	935450	59640123	63.7555	65.8021
gi 1074121301 emb LT615260.1	1432239	87098781	60.813	38.9402
gi 1074121615 emb LT615261.1	1080962	65356628	60.4615	32.224
gi 1074121871 emb LT615262.1	1545099	92111215	59.6151	16.5954
gi 1074122235 emb LT615263.1	1585108	95504442	60.2511	26.8745
gi 1074122590 emb LT615264.1	2122358	124420987	58.6239	20.704
gi 1074123050 emb LT615265.1	1754192	102806500	58.6062	31.6147
gi 1074123421 emb LT615	2150147	136363053	63.4203	94.9494

266.1				
gi 107412389 8 emb LT615 267.1	3031036	181418804	59.8537	19.365
gi 107412458 8 emb LT615 268.1	2359348	134390462	56.9608	38.469
gi 107412506 5 emb LT615 269.1	3135668	189349694	60.3858	14.3092

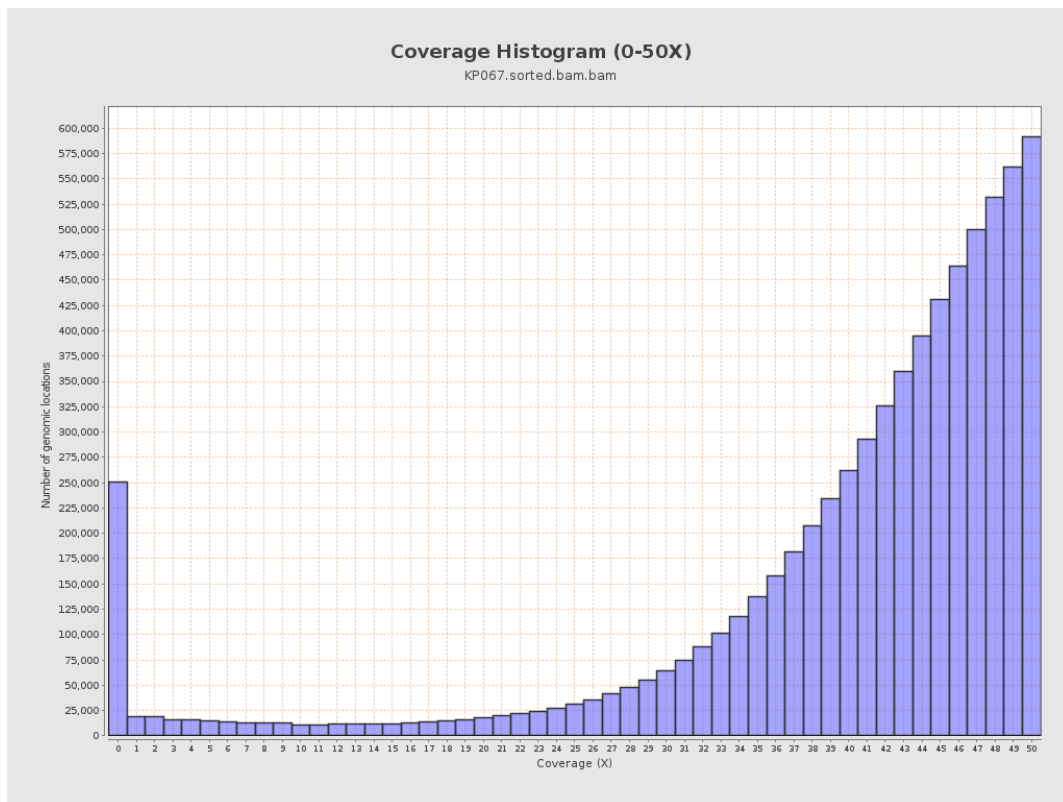
3. Results : Coverage across reference



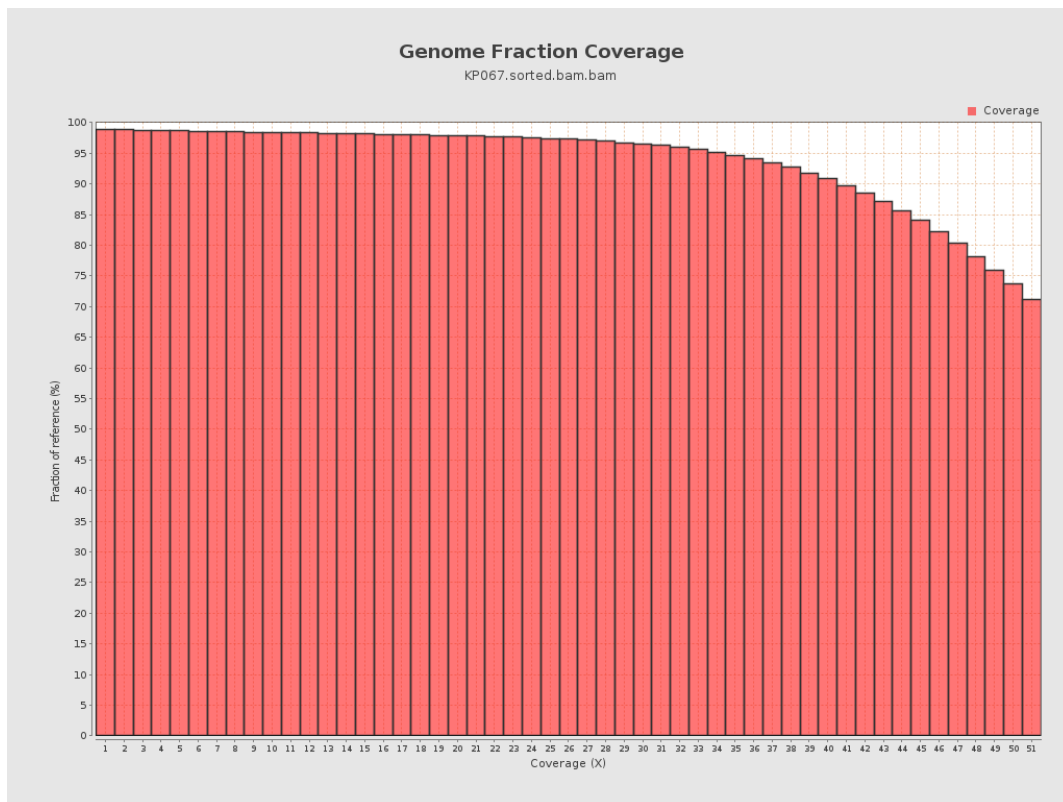
4. Results : Coverage Histogram



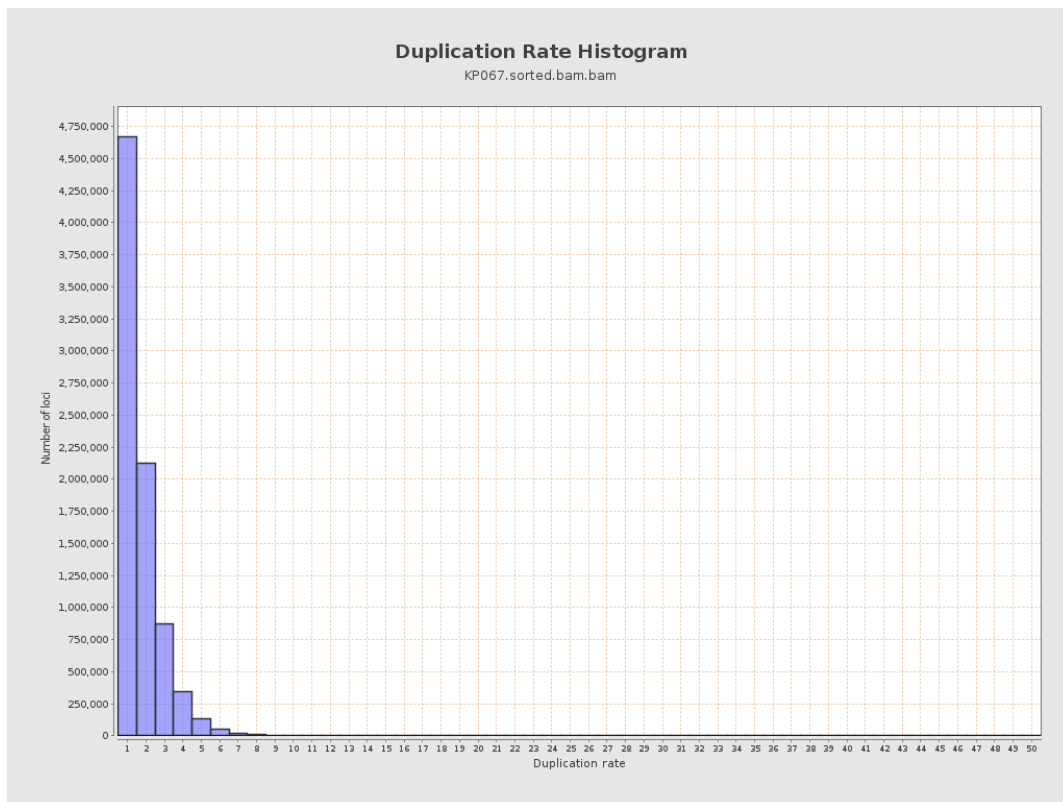
5. Results : Coverage Histogram (0-50X)



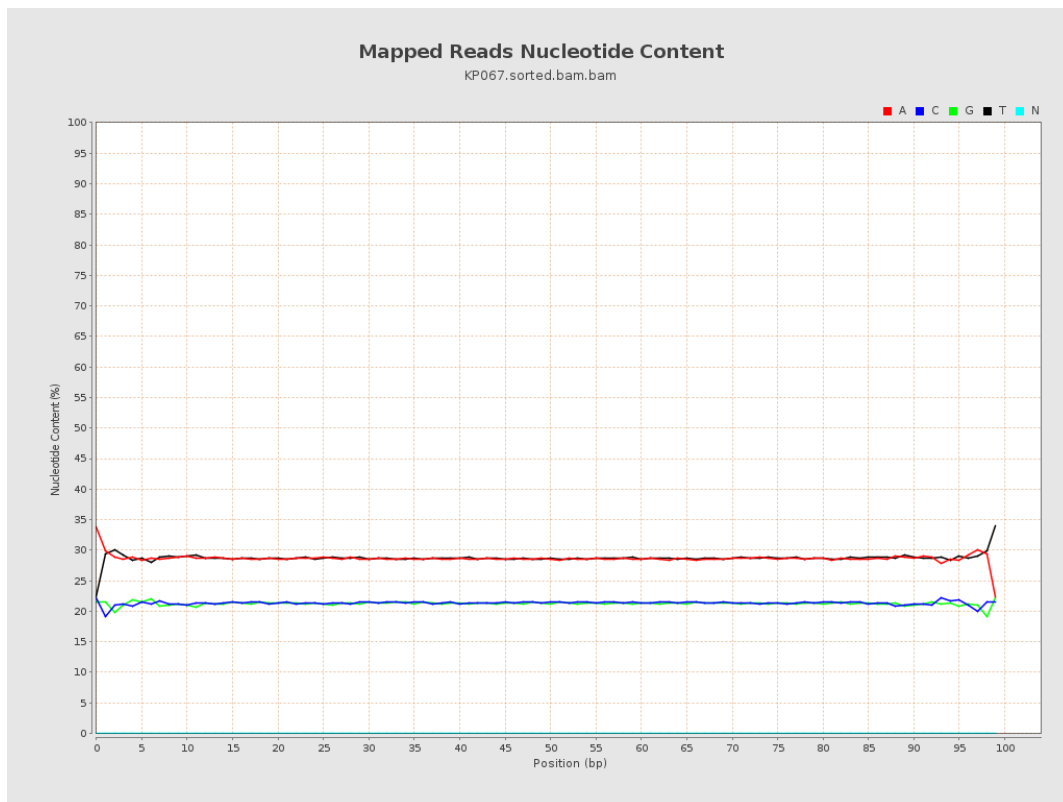
6. Results : Genome Fraction Coverage



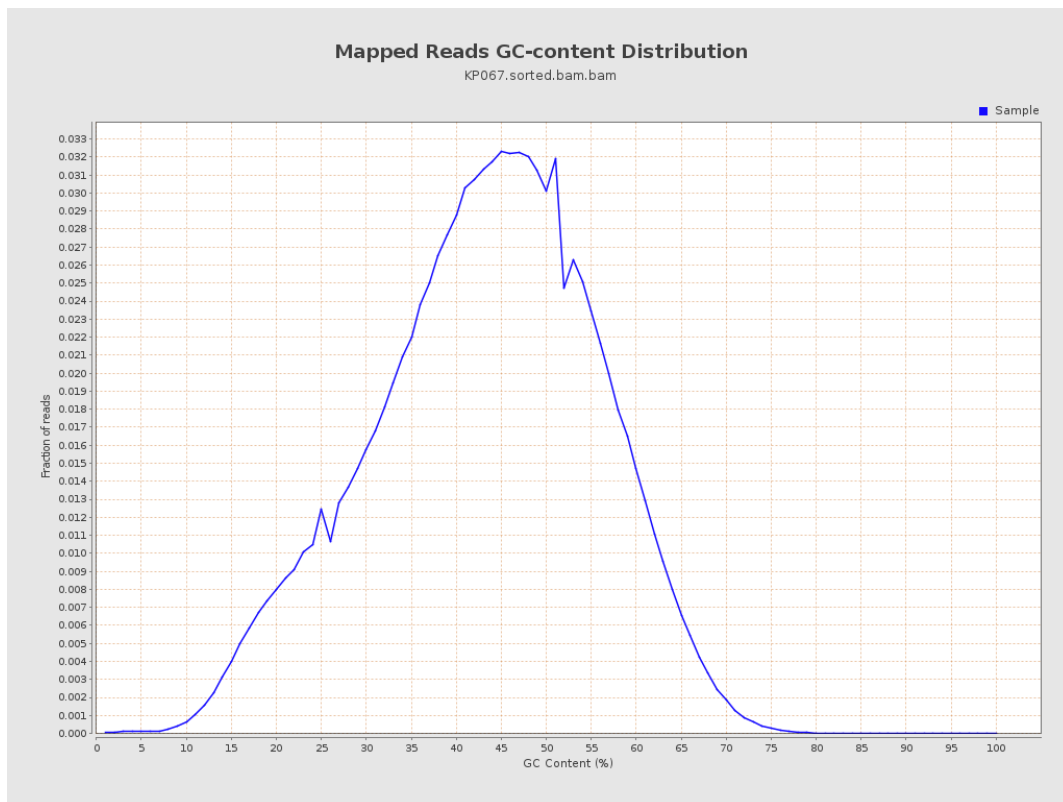
7. Results : Duplication Rate Histogram



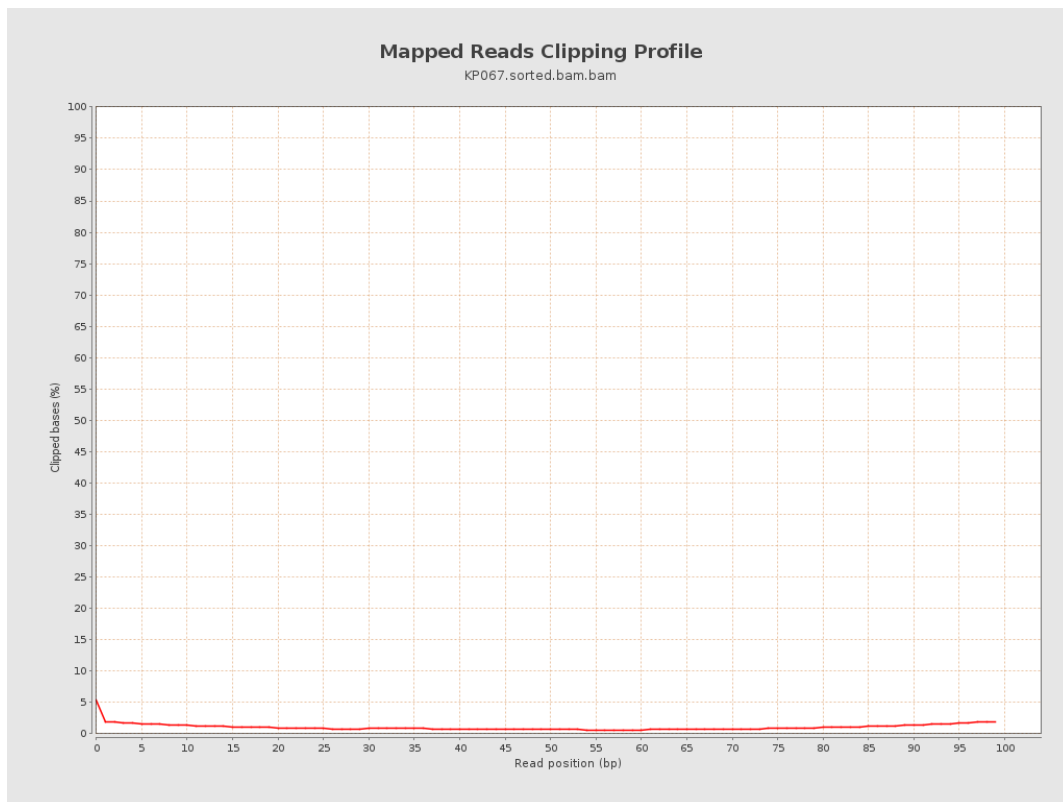
8. Results : Mapped Reads Nucleotide Content



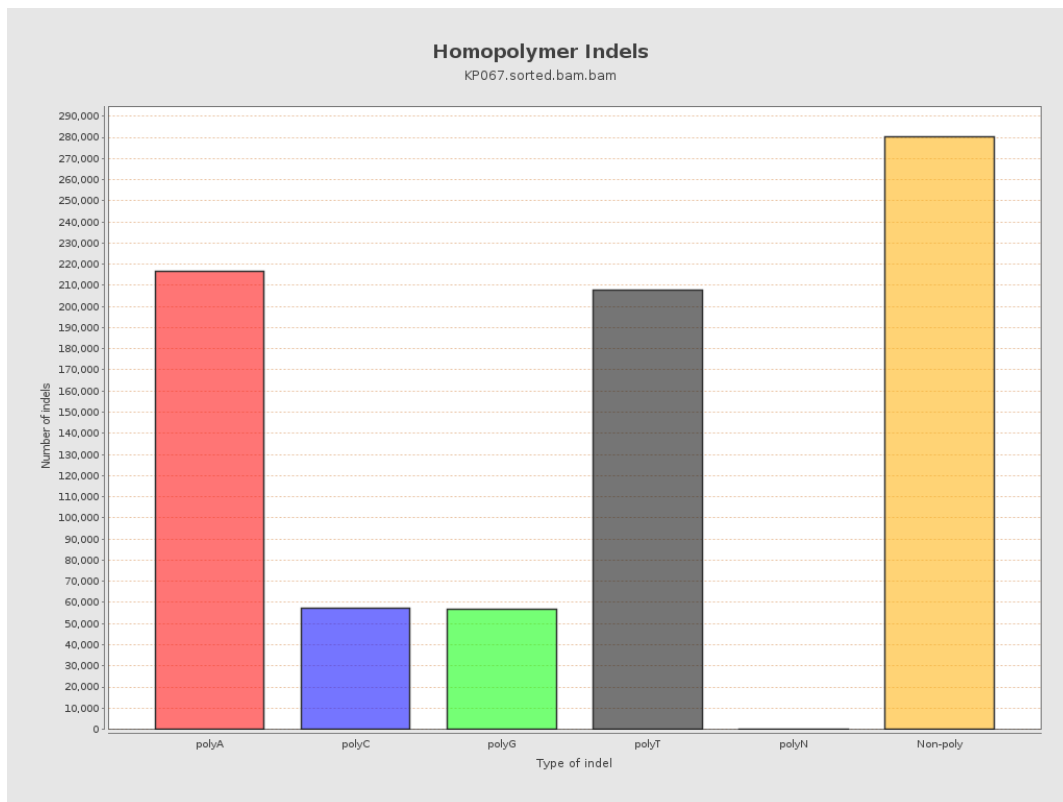
9. Results : Mapped Reads GC-content Distribution



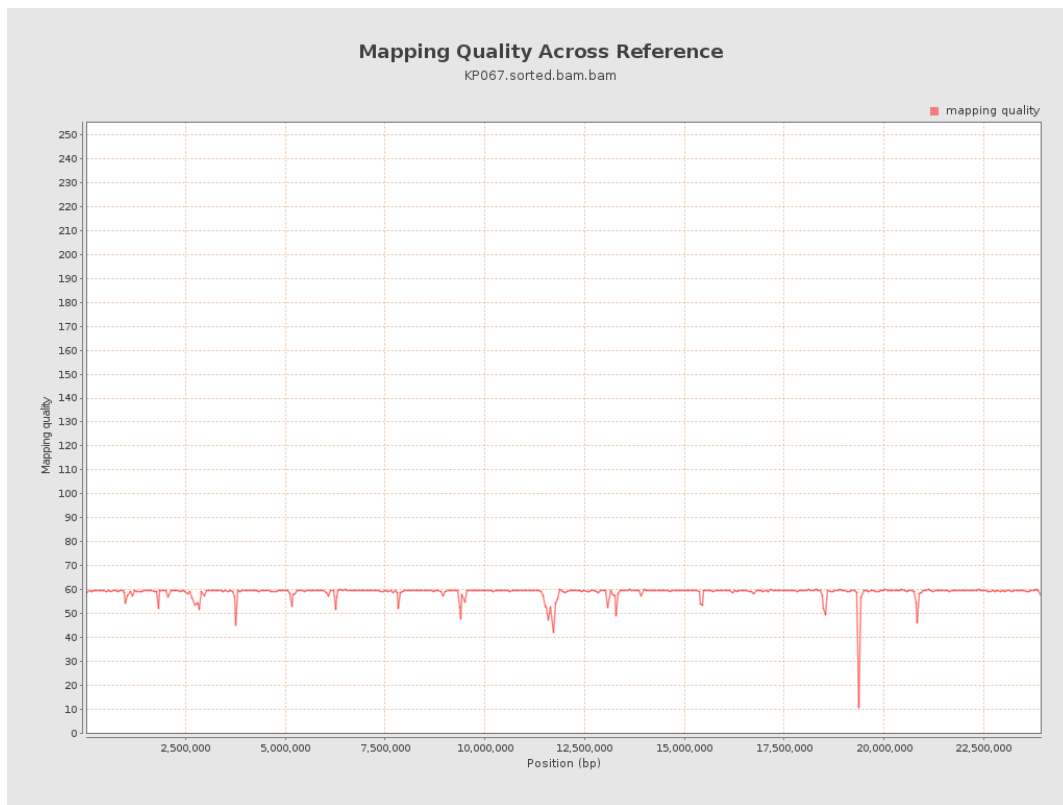
10. Results : Mapped Reads Clipping Profile



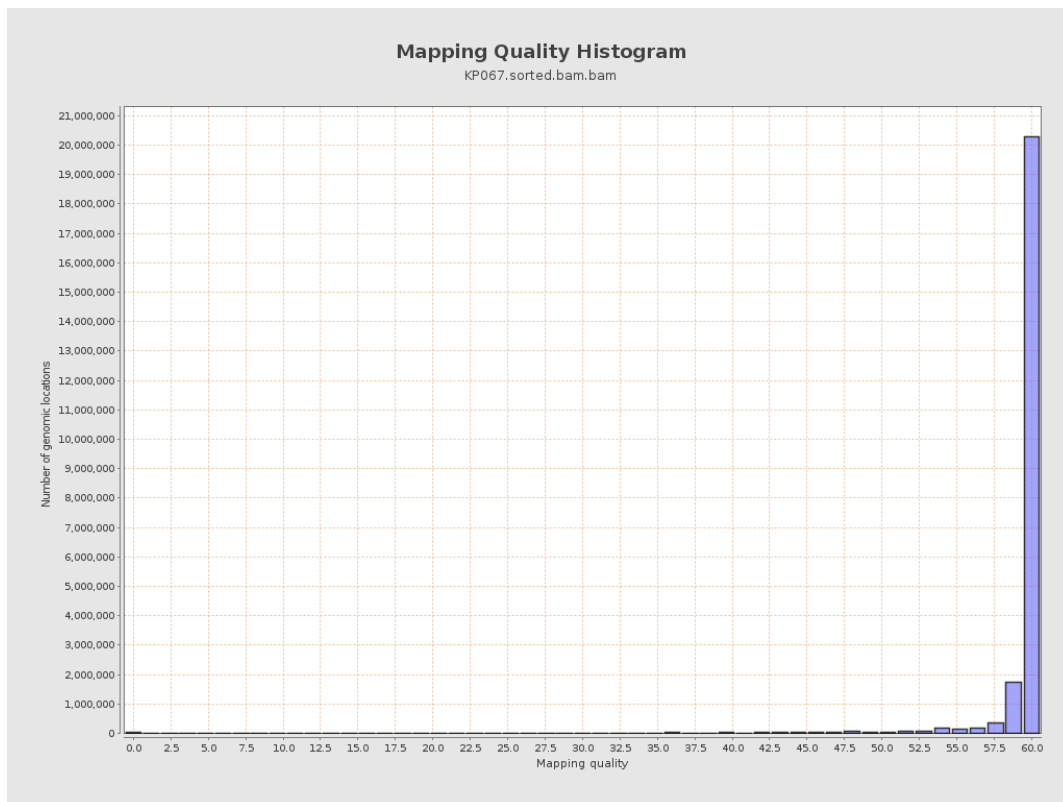
11. Results : Homopolymer Indels



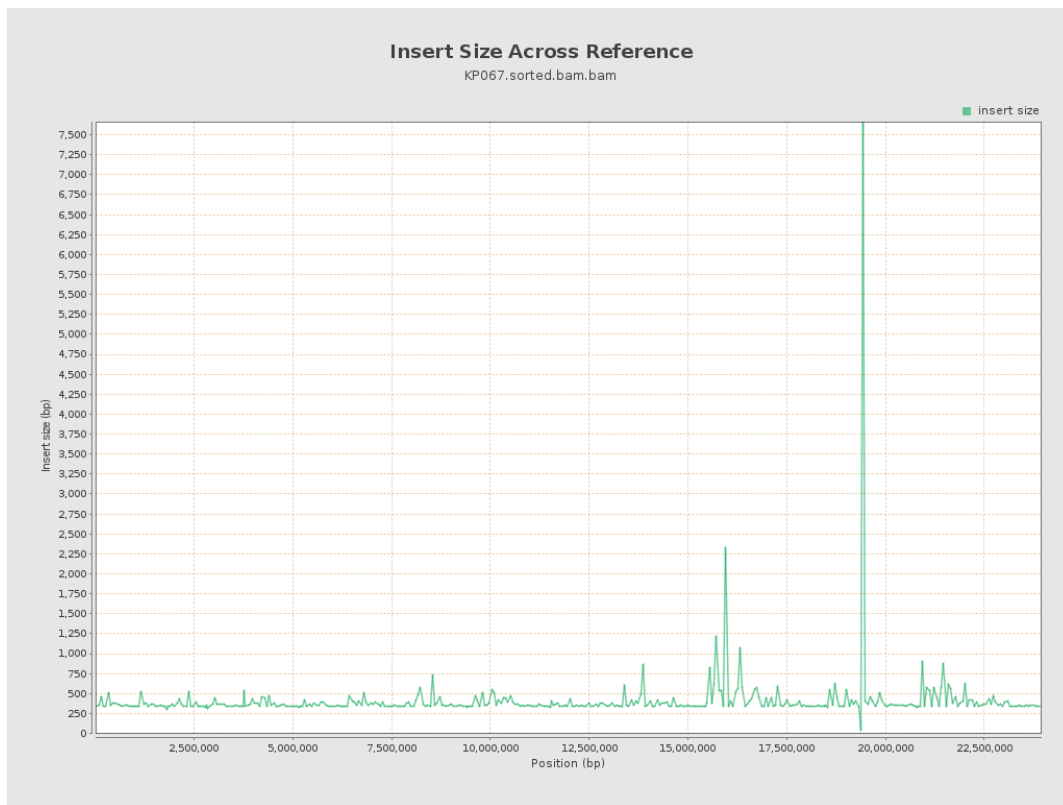
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

