

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 11:46:12

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/KP034.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/KP034-BiooBarcode_24_GGTAGC_R1.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/KP034-BiooBarcode_24_GGTAGC_R2.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 11:46:11 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/KP034.sorted.bam.b

am

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	18,755,497
Mapped reads	12,722,653 / 67.83%
Unmapped reads	6,032,844 / 32.17%
Mapped paired reads	12,722,653 / 67.83%
Mapped reads, first in pair	6,385,504 / 34.05%
Mapped reads, second in pair	6,337,149 / 33.79%
Mapped reads, both in pair	12,505,988 / 66.68%
Mapped reads, singletons	216,665 / 1.16%
Read min/max/mean length	30 / 100 / 99.87
Duplicated reads (estimated)	7,012,447 / 37.39%
Duplication rate	56.71%
Clipped reads	1,232,361 / 6.57%

2.2. ACGT Content

Number/percentage of A's	352,491,440 / 28.61%
Number/percentage of C's	263,259,130 / 21.37%
Number/percentage of T's	353,613,660 / 28.7%
Number/percentage of G's	262,563,950 / 21.31%
Number/percentage of N's	103,050 / 0.01%
GC Percentage	42.68%

2.3. Coverage

Mean	51.4733
Standard Deviation	36.2434

2.4. Mapping Quality

Mean Mapping Quality	58.76
----------------------	-------

2.5. Insert size

Mean	914.35
Standard Deviation	26,796.19
P25/Median/P75	265 / 337 / 411

2.6. Mismatches and indels

General error rate	1.02%
Mismatches	11,771,515
Insertions	303,292
Mapped reads with at least one insertion	2.27%
Deletions	403,022
Mapped reads with at least one deletion	3.02%
Homopolymer indels	65.35%

2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

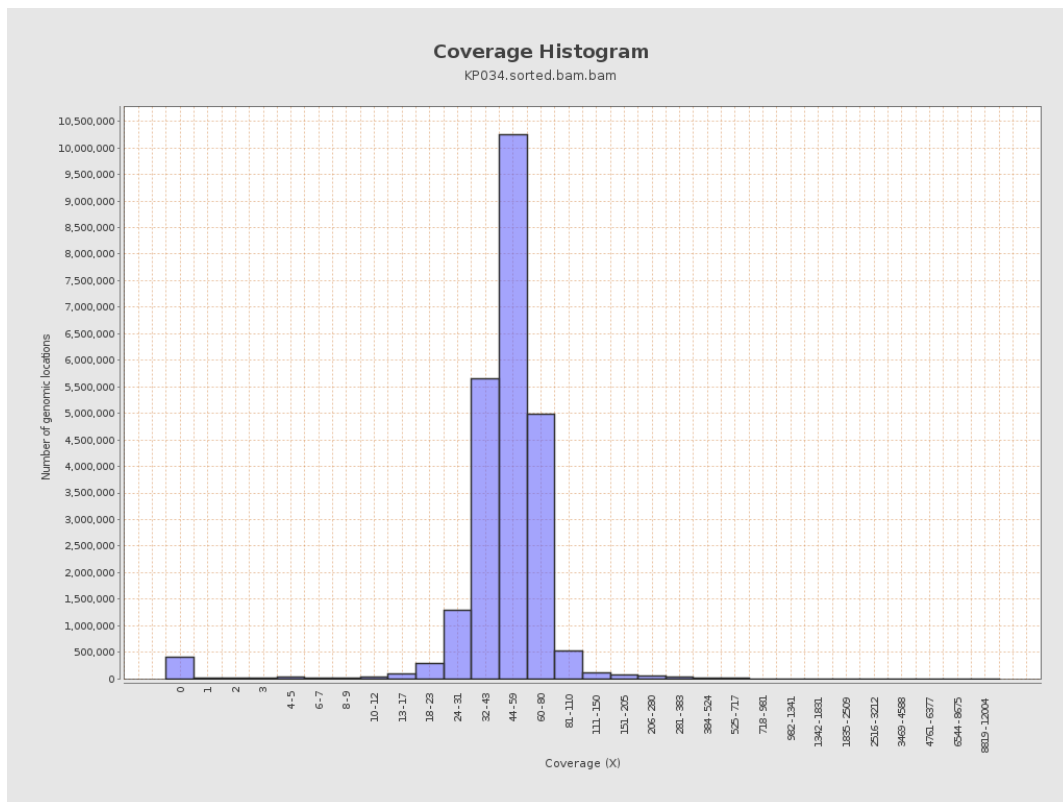
gi 1074120478 emb LT615256.1	977217	49611532	50.7682	22.37
gi 1074120682 emb LT615257.1	860454	43435981	50.4803	26.0089
gi 1074120865 emb LT615258.1	989719	54492632	55.0587	47.7456
gi 1074121086 emb LT615259.1	935450	50936577	54.4514	59.6316
gi 1074121301 emb LT615260.1	1432239	75059315	52.407	33.9537
gi 1074121615 emb LT615261.1	1080962	56175613	51.9682	31.3477
gi 1074121871 emb LT615262.1	1545099	78102560	50.5486	16.6249
gi 1074122235 emb LT615263.1	1585108	82455981	52.0192	26.615
gi 1074122590 emb LT615264.1	2122358	106623462	50.2382	21.9675
gi 1074123050 emb LT615265.1	1754192	89115427	50.8014	34.9593
gi 1074123421 emb LT615	2150147	115056343	53.5109	76.0172

266.1				
gi 107412389 8 emb LT615 267.1	3031036	154238689	50.8865	17.0993
gi 107412458 8 emb LT615 268.1	2359348	117097624	49.6313	37.2452
gi 107412506 5 emb LT615 269.1	3135668	160847603	51.2961	15.5241

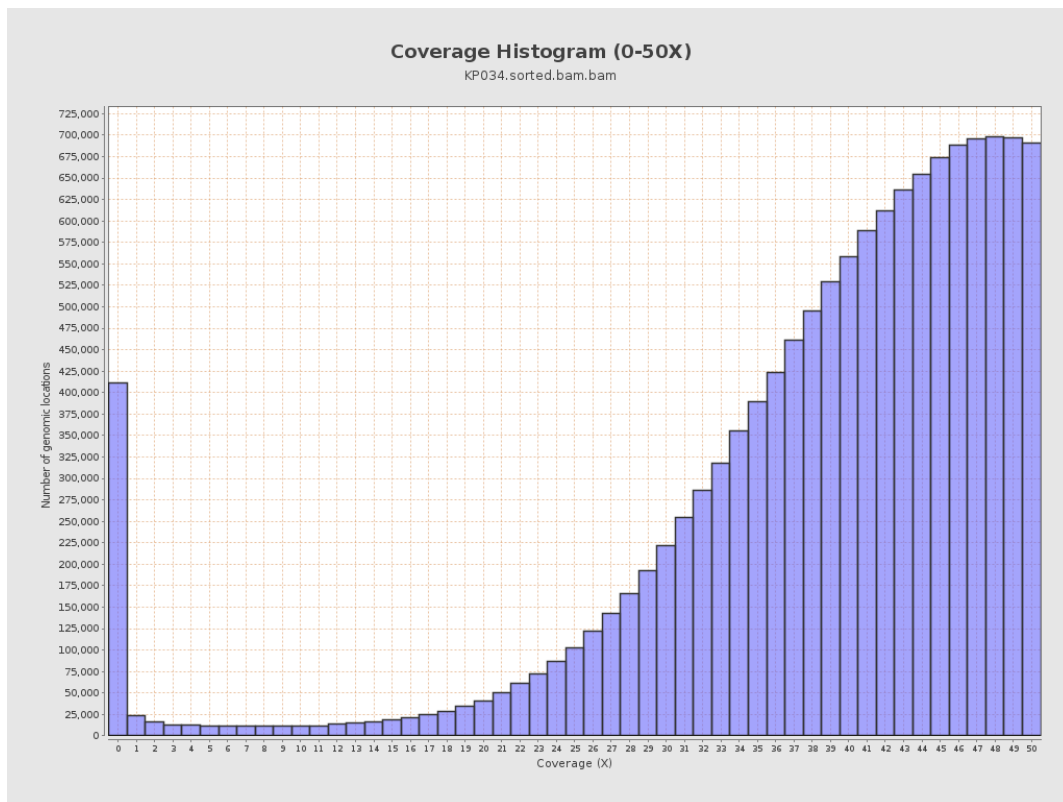
3. Results : Coverage across reference



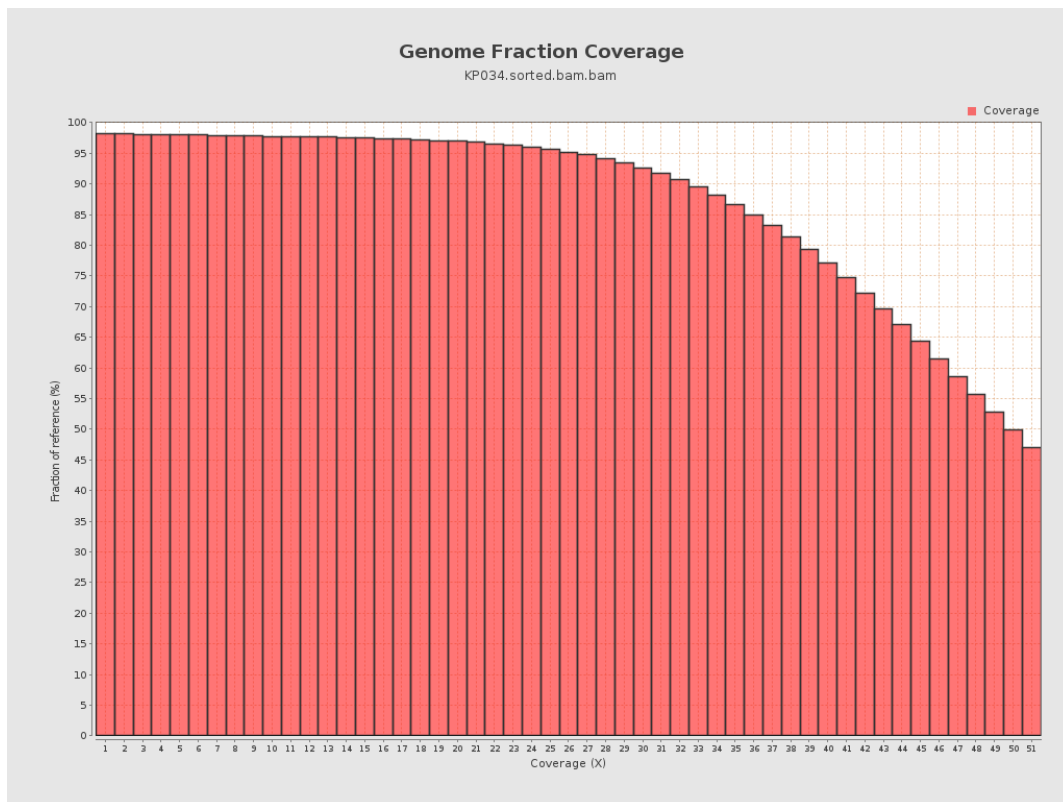
4. Results : Coverage Histogram



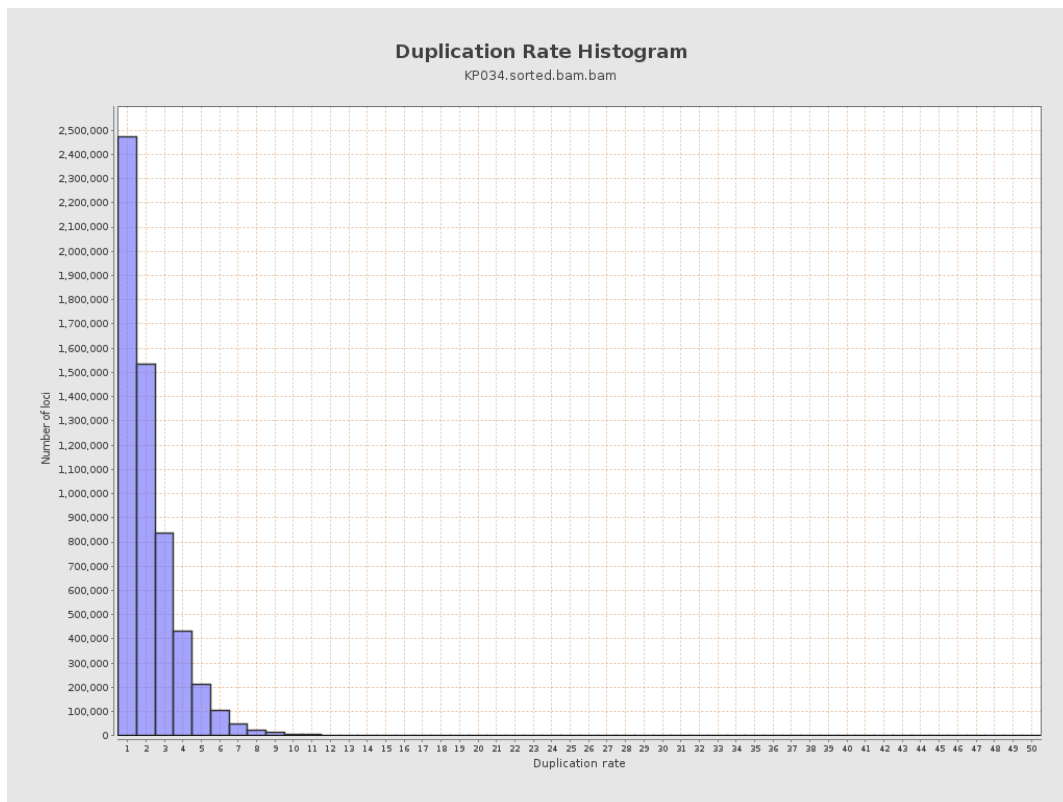
5. Results : Coverage Histogram (0-50X)



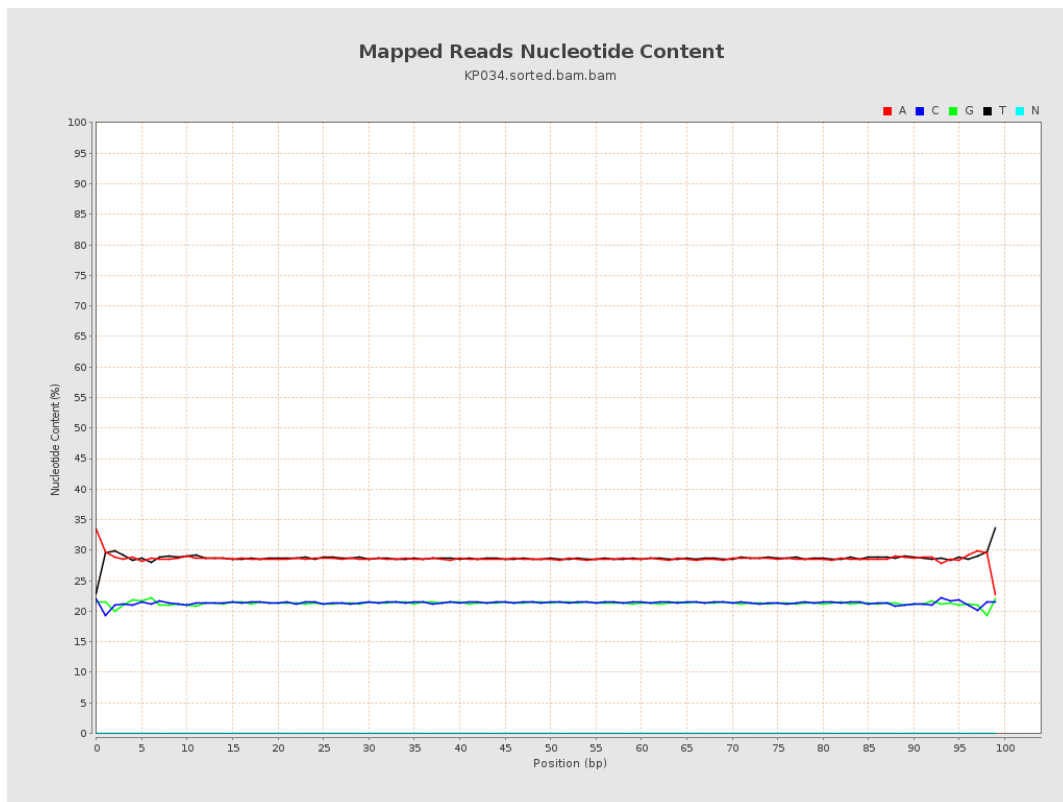
6. Results : Genome Fraction Coverage



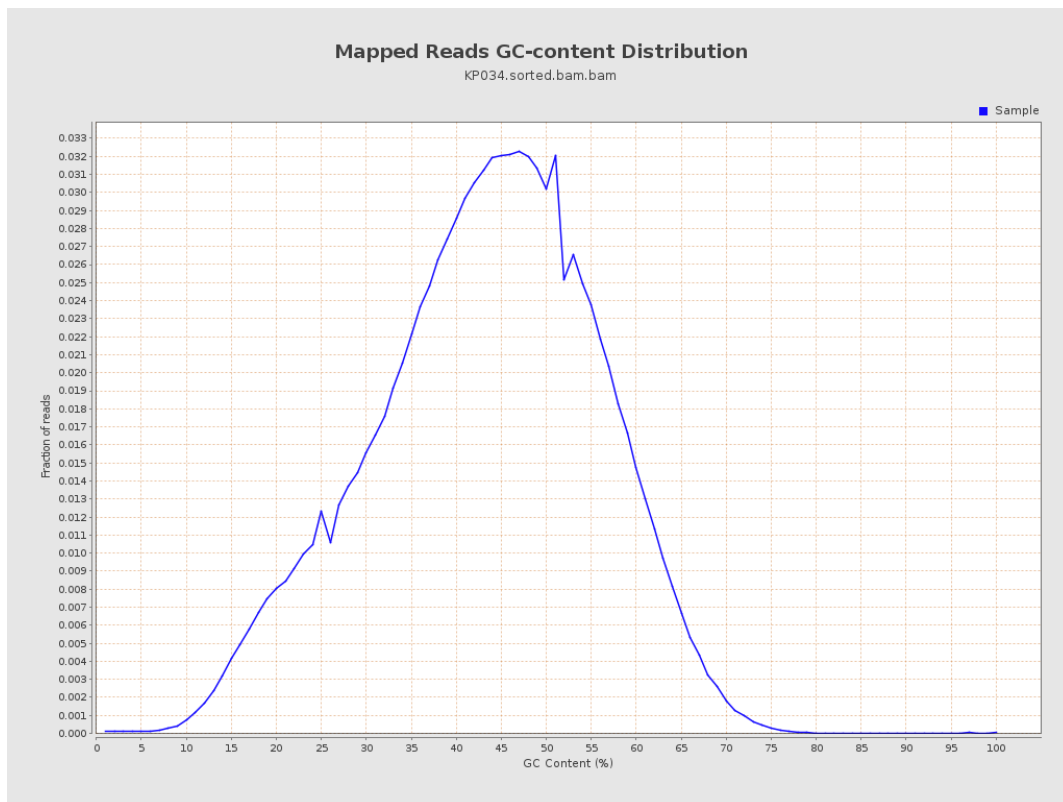
7. Results : Duplication Rate Histogram



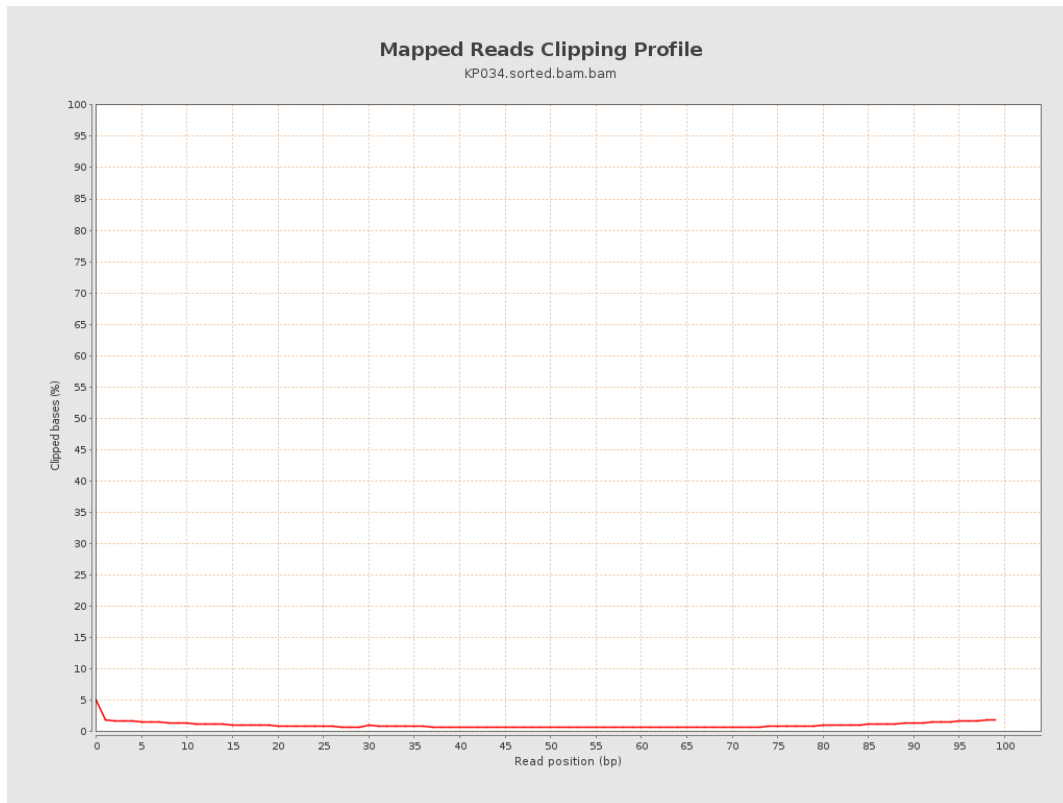
8. Results : Mapped Reads Nucleotide Content



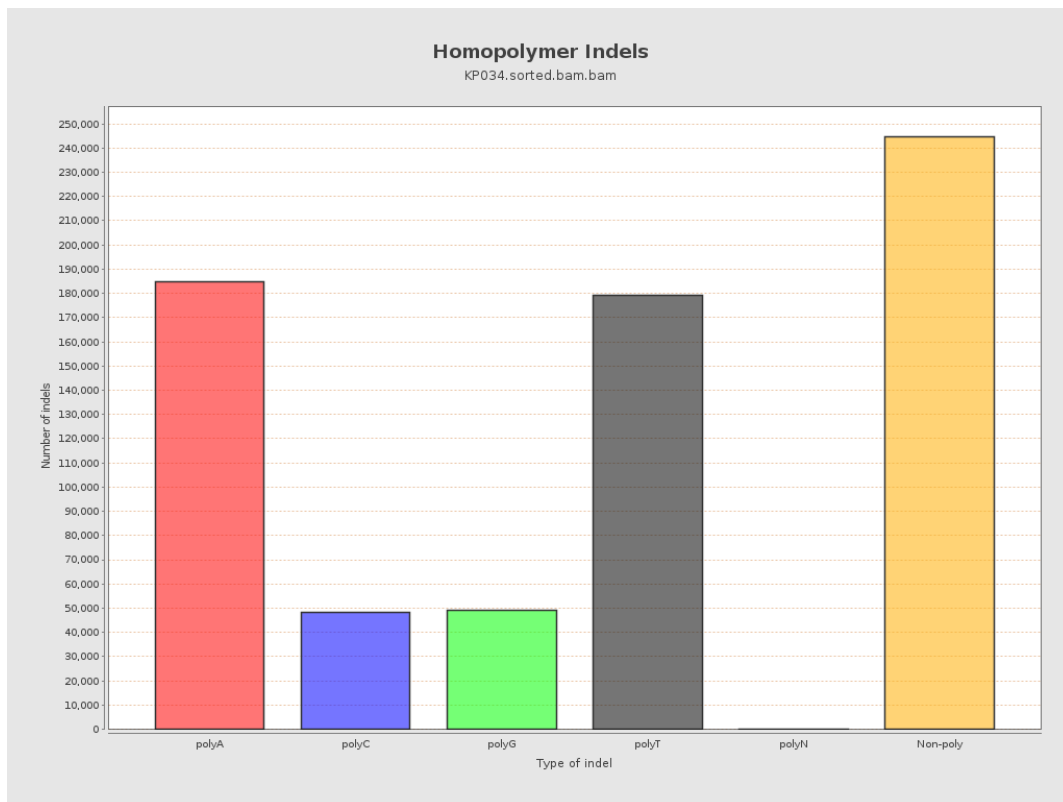
9. Results : Mapped Reads GC-content Distribution



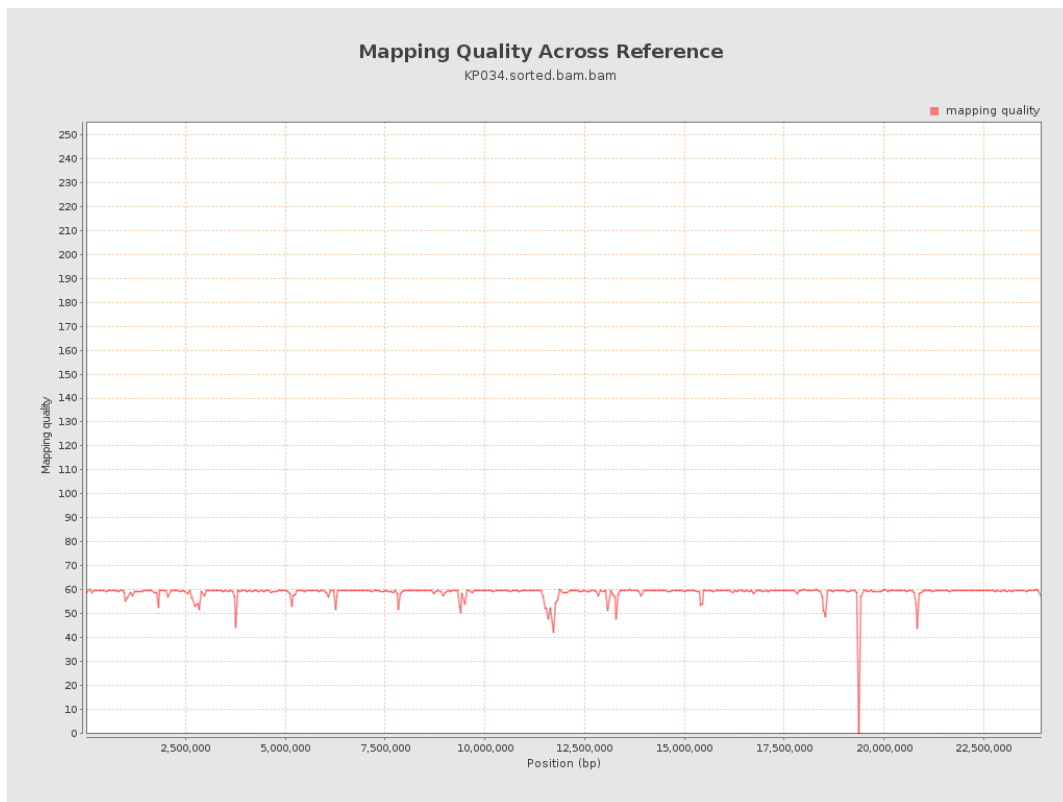
10. Results : Mapped Reads Clipping Profile



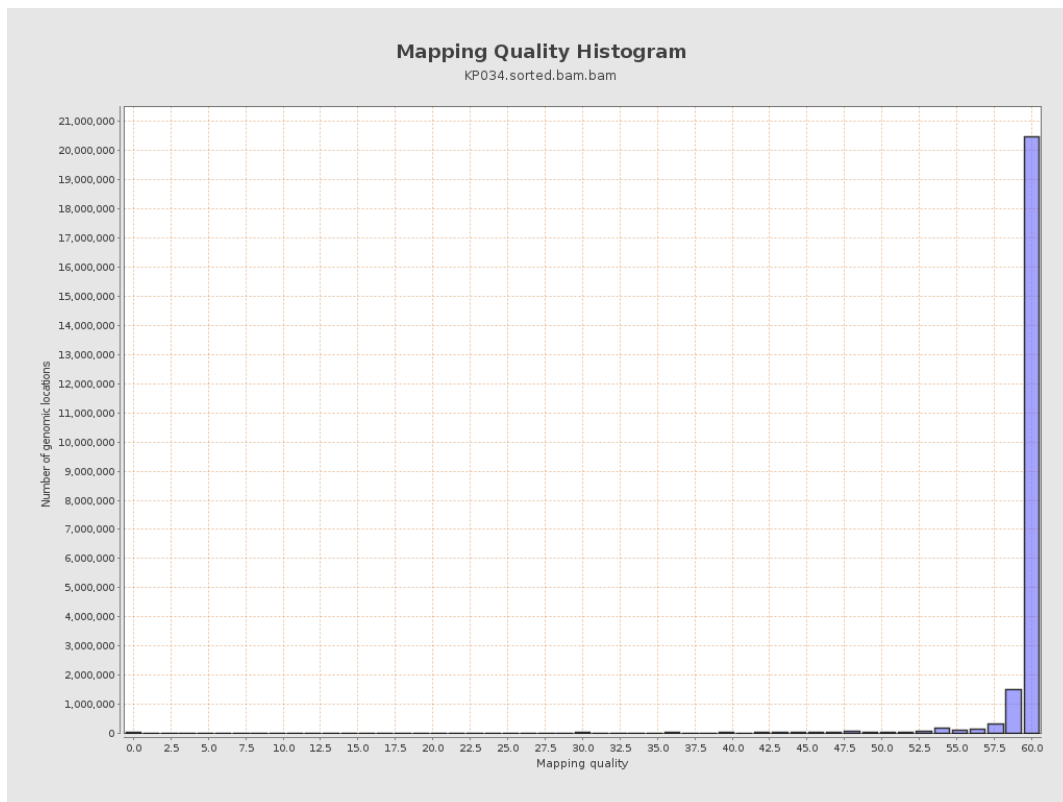
11. Results : Homopolymer Indels



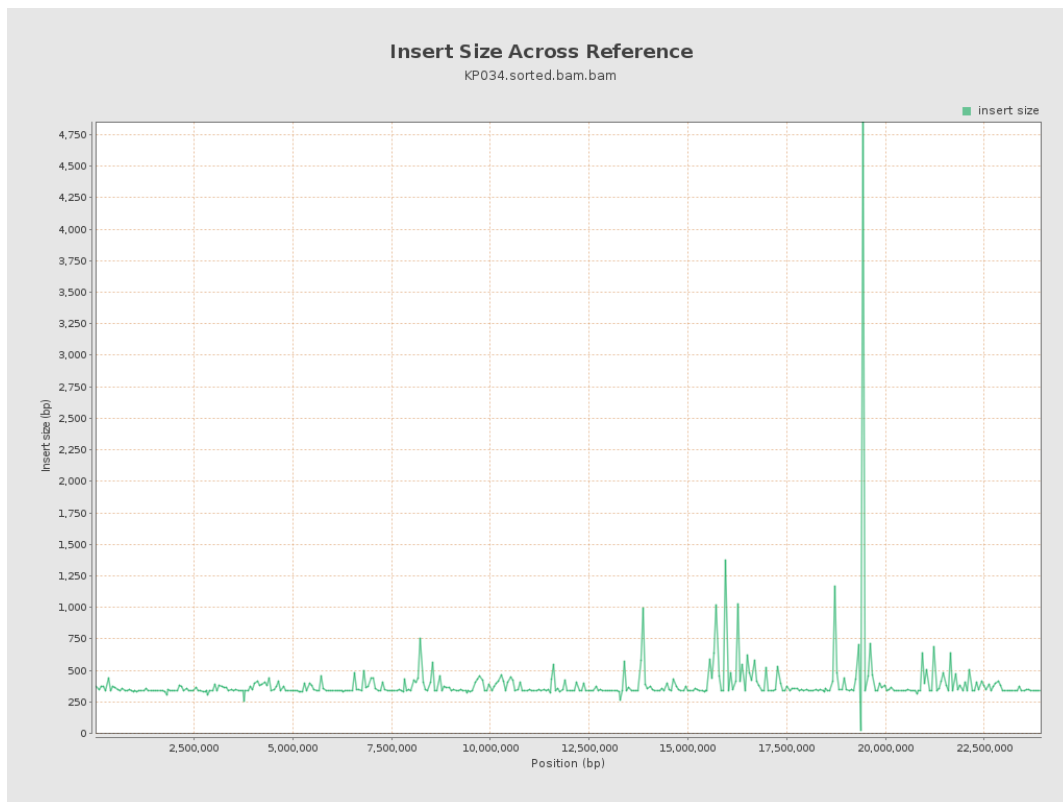
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

