

Qualimap Analysis Results

BAM QC analysis

Generated by Qualimap v.2.2.1

2016/10/23 11:44:23

1. Input data & parameters

1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/KP017.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/KP017-BiooBarcode_23_GAGTGG_R2.fastq.gz /home/vdp5/data/cambodia_samples/sequences_gz/KP017-BiooBarcode_23_GAGTGG_R1.fastq.gz
Draw chromosome limits:	no
Analyze overlapping paired-end reads:	no
Program:	bwa (0.7.15-r1140)
Analysis date:	Sun Oct 23 11:44:22 EDT 2016
Size of a homopolymer:	3
Skip duplicate alignments:	no
Number of windows:	400
BAM file:	/home/vdp5/data/cambodia_samples/sequences_bam/KP017.sorted.bam.b

am

2. Summary

2.1. Globals

Reference size	23,958,997
Number of reads	22,797,277
Mapped reads	18,190,531 / 79.79%
Unmapped reads	4,606,746 / 20.21%
Mapped paired reads	18,190,531 / 79.79%
Mapped reads, first in pair	9,063,542 / 39.76%
Mapped reads, second in pair	9,126,989 / 40.04%
Mapped reads, both in pair	17,966,020 / 78.81%
Mapped reads, singletons	224,511 / 0.98%
Read min/max/mean length	30 / 100 / 99.82
Duplicated reads (estimated)	7,032,848 / 30.85%
Duplication rate	39.31%
Clipped reads	1,435,946 / 6.3%

2.2. ACGT Content

Number/percentage of A's	490,929,424 / 27.7%
Number/percentage of C's	395,019,825 / 22.29%
Number/percentage of T's	492,420,023 / 27.79%
Number/percentage of G's	393,822,032 / 22.22%
Number/percentage of N's	151,125 / 0.01%
GC Percentage	44.51%

2.3. Coverage

Mean	74.0412
Standard Deviation	37.539

2.4. Mapping Quality

Mean Mapping Quality	58.72
----------------------	-------

2.5. Insert size

Mean	1,035.58
Standard Deviation	30,854.5
P25/Median/P75	246 / 307 / 366

2.6. Mismatches and indels

General error rate	0.73%
Mismatches	11,900,084
Insertions	338,124
Mapped reads with at least one insertion	1.8%
Deletions	475,448
Mapped reads with at least one deletion	2.52%
Homopolymer indels	71.21%

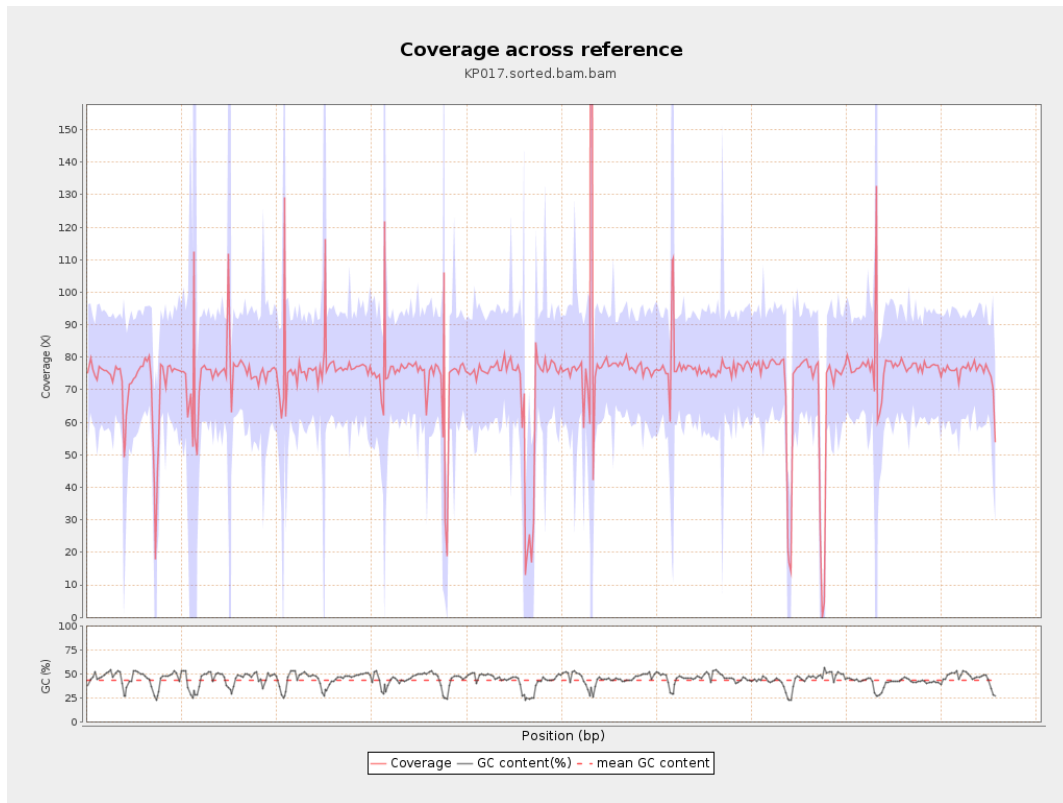
2.7. Chromosome stats

Name	Length	Mapped bases	Mean coverage	Standard deviation

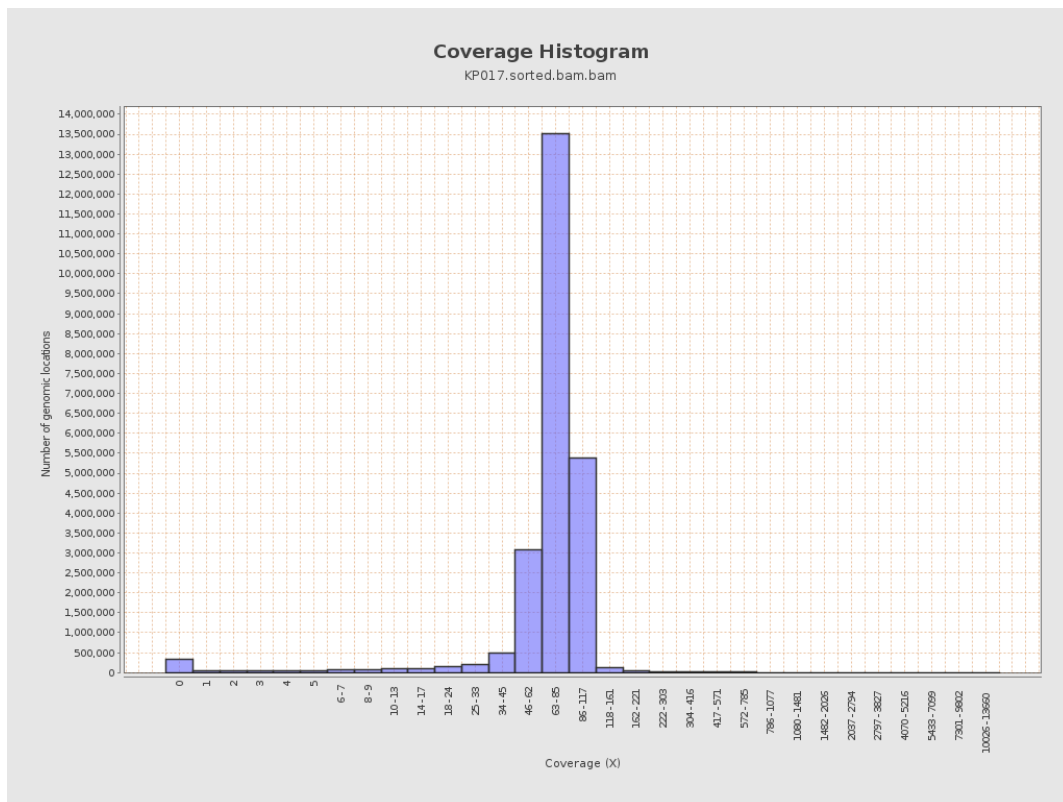
gi 1074120478 emb LT615256.1	977217	73366105	75.0766	17.7179
gi 1074120682 emb LT615257.1	860454	58954658	68.5158	27.8127
gi 1074120865 emb LT615258.1	989719	70941940	71.6789	38.1194
gi 1074121086 emb LT615259.1	935450	69703696	74.5135	47.2251
gi 1074121301 emb LT615260.1	1432239	106076268	74.0632	30.3812
gi 1074121615 emb LT615261.1	1080962	81720440	75.5997	30.2748
gi 1074121871 emb LT615262.1	1545099	117675030	76.1602	18.6371
gi 1074122235 emb LT615263.1	1585108	120216927	75.8415	26.6443
gi 1074122590 emb LT615264.1	2122358	154878865	72.9749	24.2046
gi 1074123050 emb LT615265.1	1754192	117756741	67.1288	32.7788
gi 1074123421 emb LT615	2150147	168264929	78.2574	85.5185

266.1				
gi 107412389 8 emb LT615 267.1	3031036	232372577	76.6644	20.5345
gi 107412458 8 emb LT615 268.1	2359348	164607338	69.7681	41.6361
gi 107412506 5 emb LT615 269.1	3135668	237416488	75.7148	16.2199

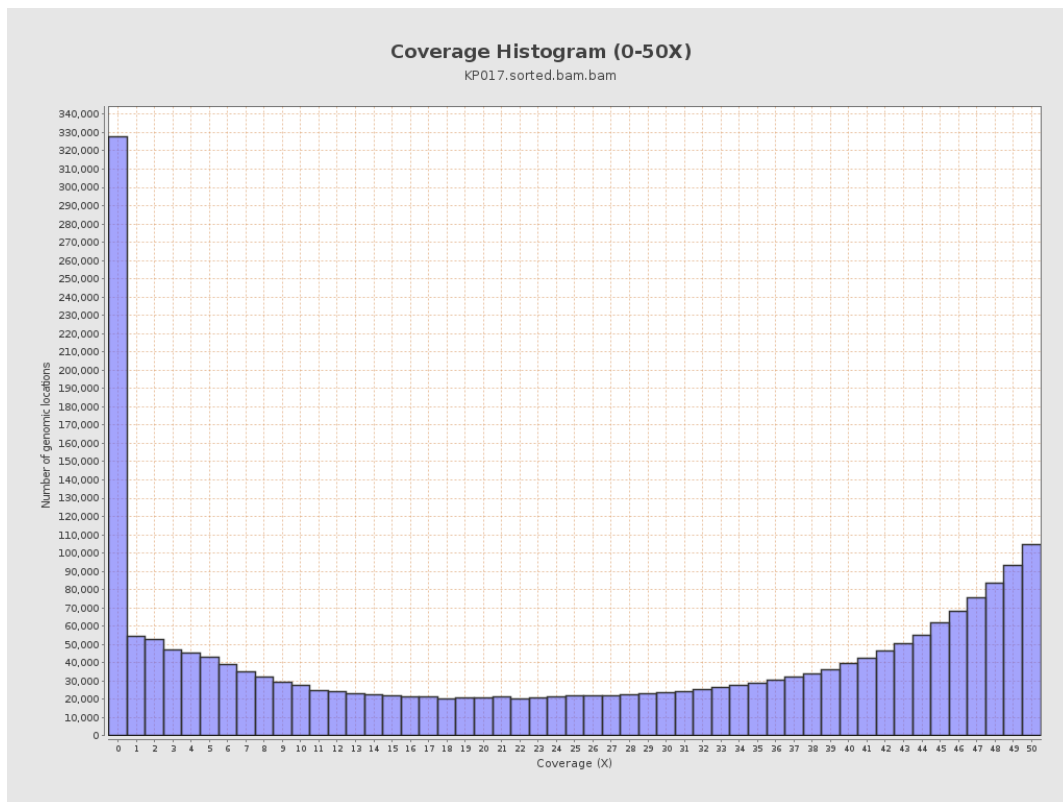
3. Results : Coverage across reference



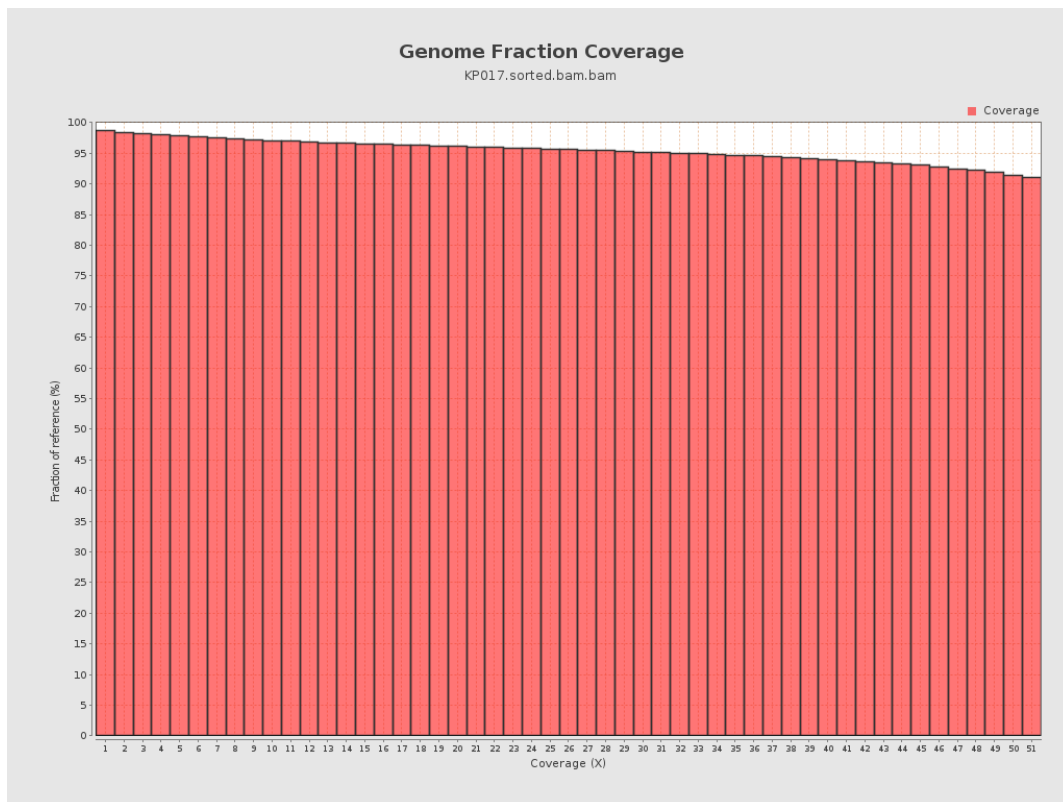
4. Results : Coverage Histogram



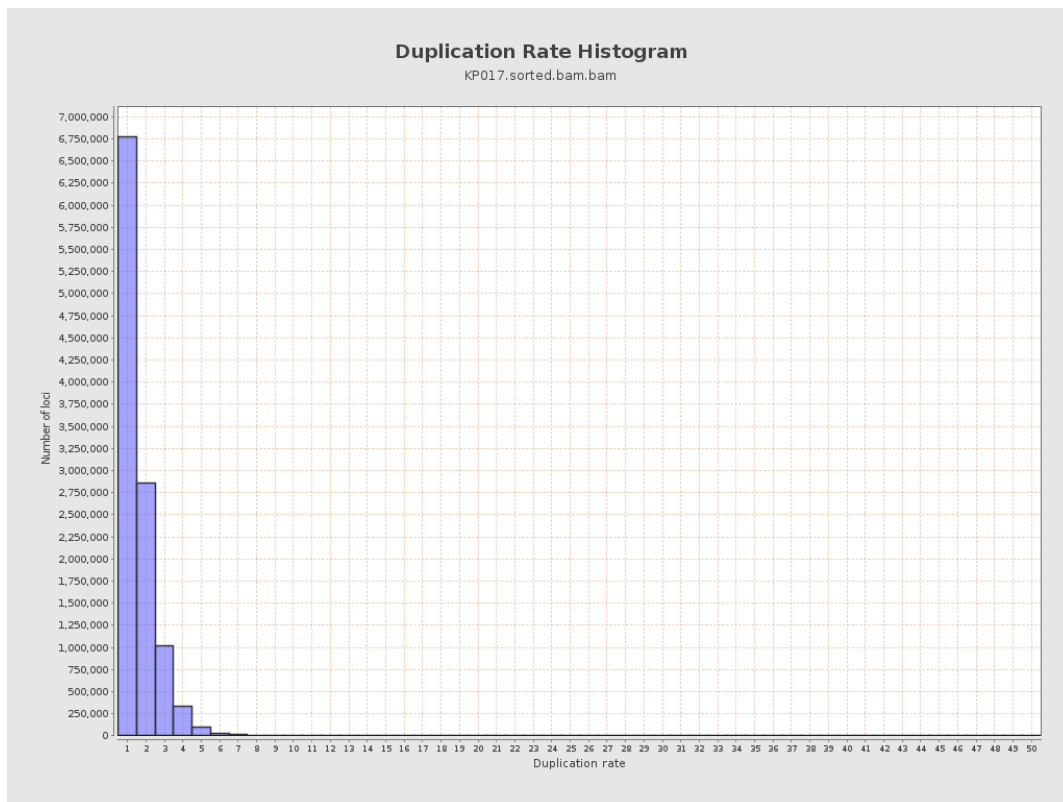
5. Results : Coverage Histogram (0-50X)



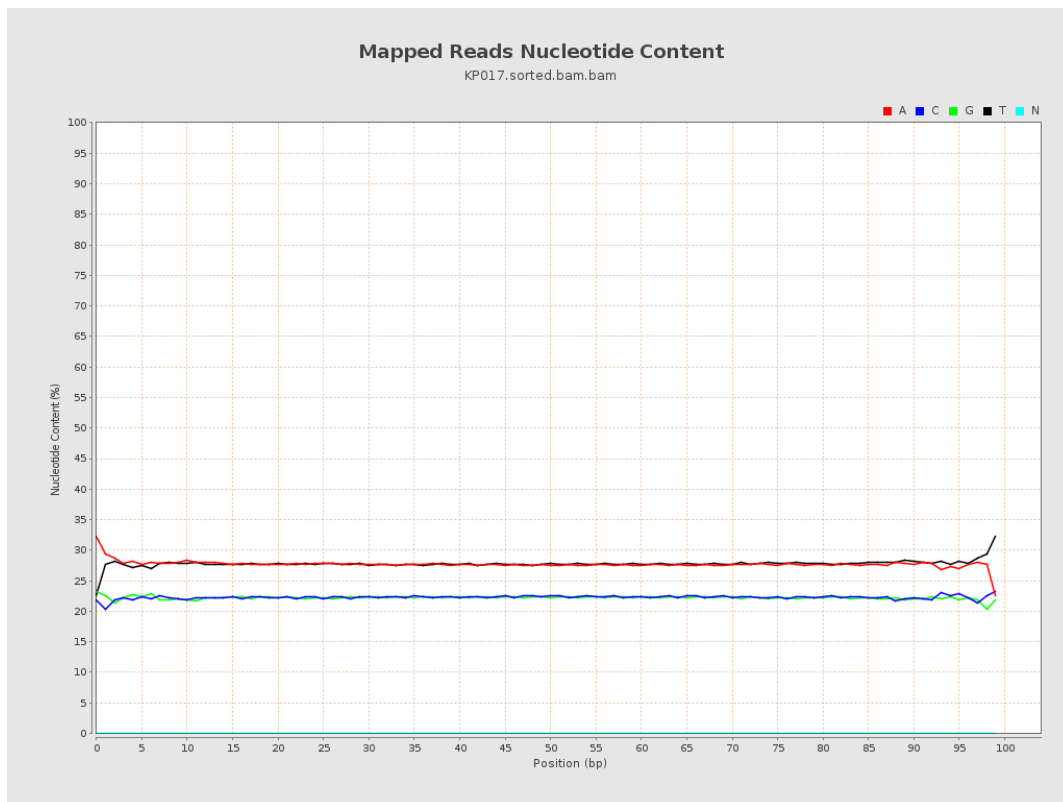
6. Results : Genome Fraction Coverage



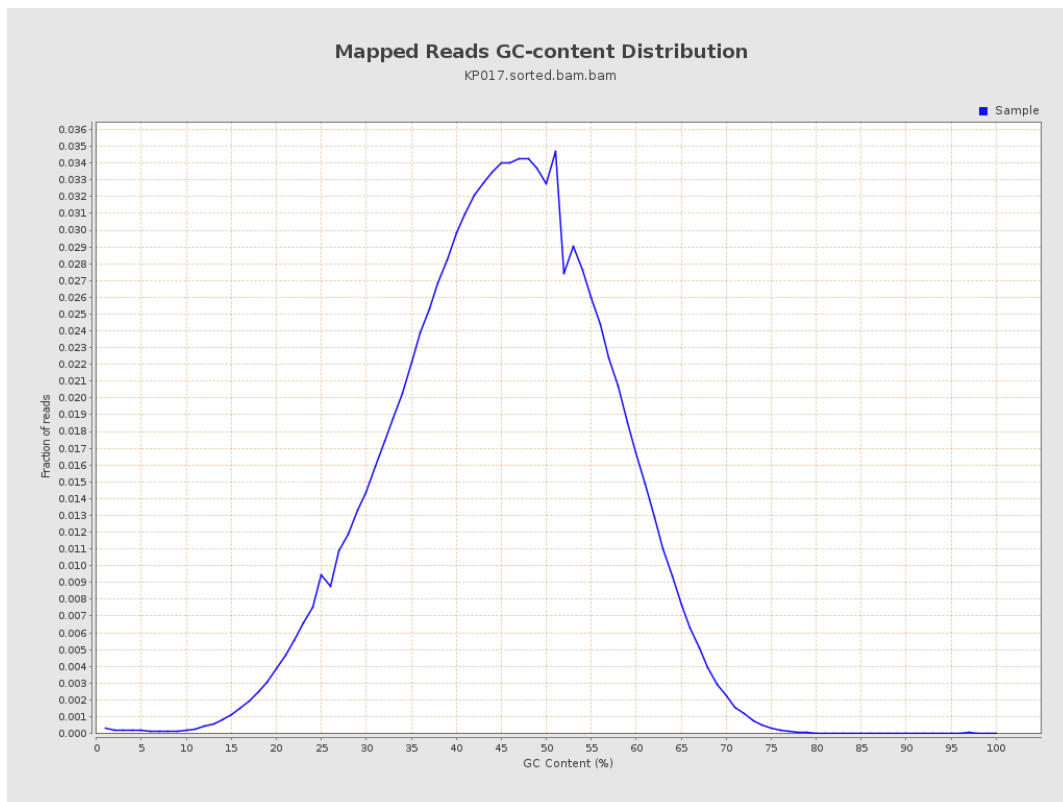
7. Results : Duplication Rate Histogram



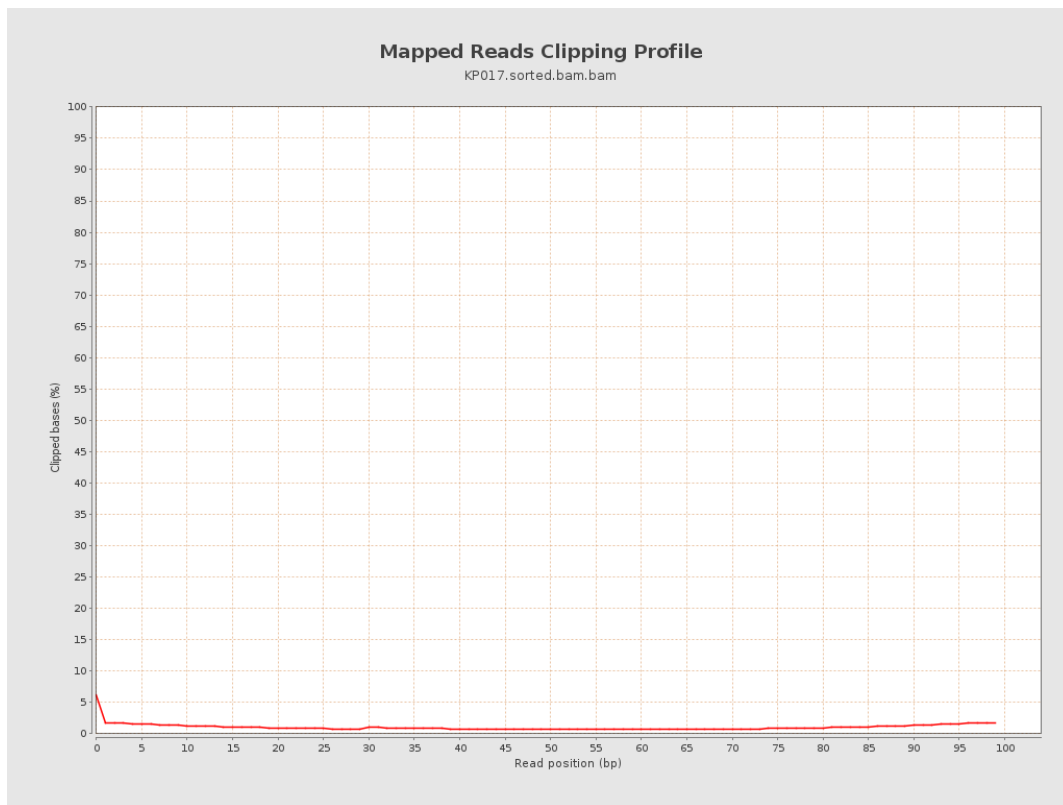
8. Results : Mapped Reads Nucleotide Content



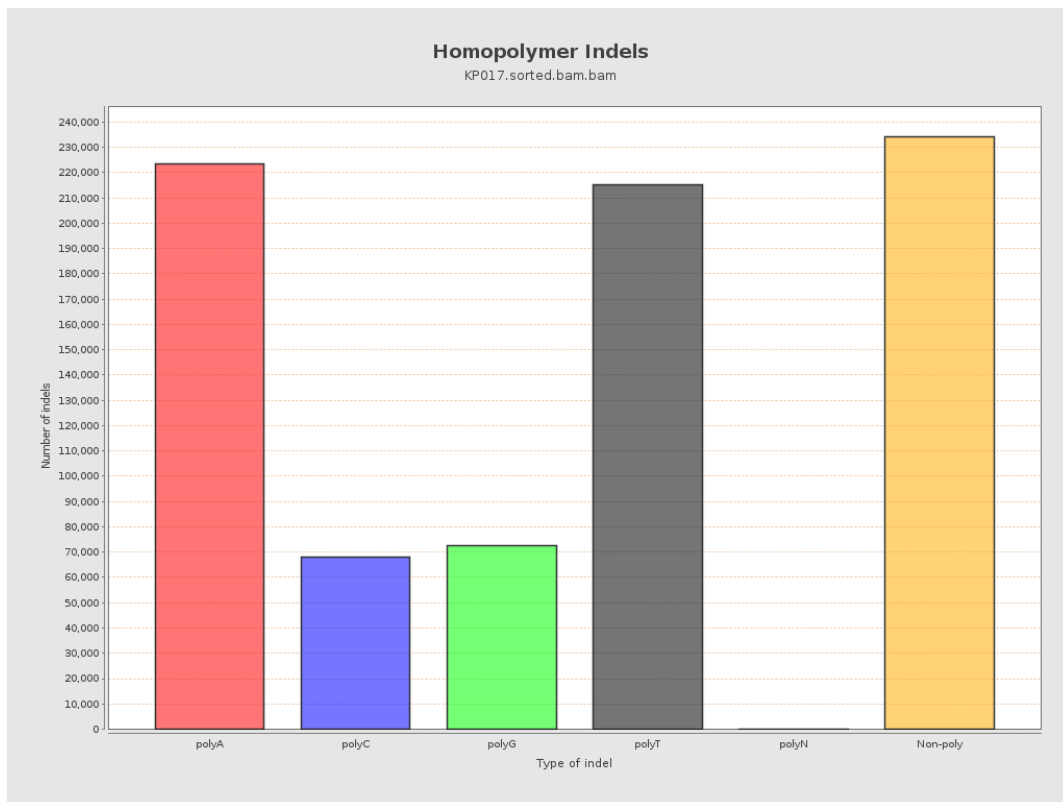
9. Results : Mapped Reads GC-content Distribution



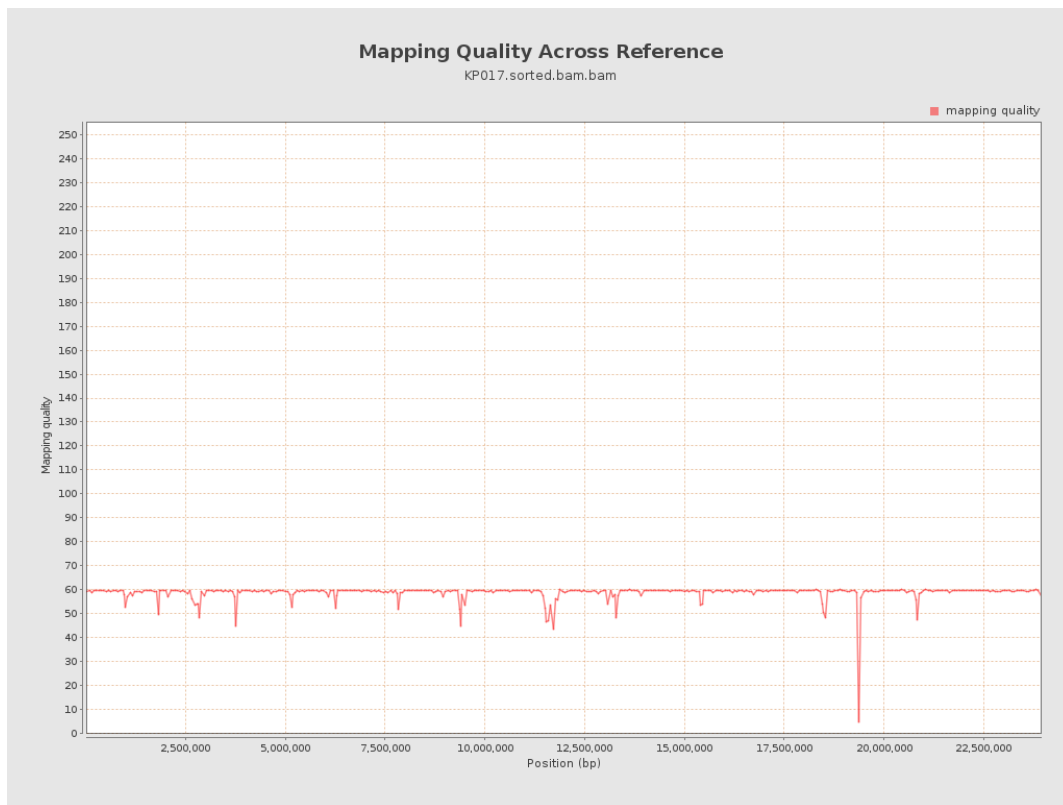
10. Results : Mapped Reads Clipping Profile



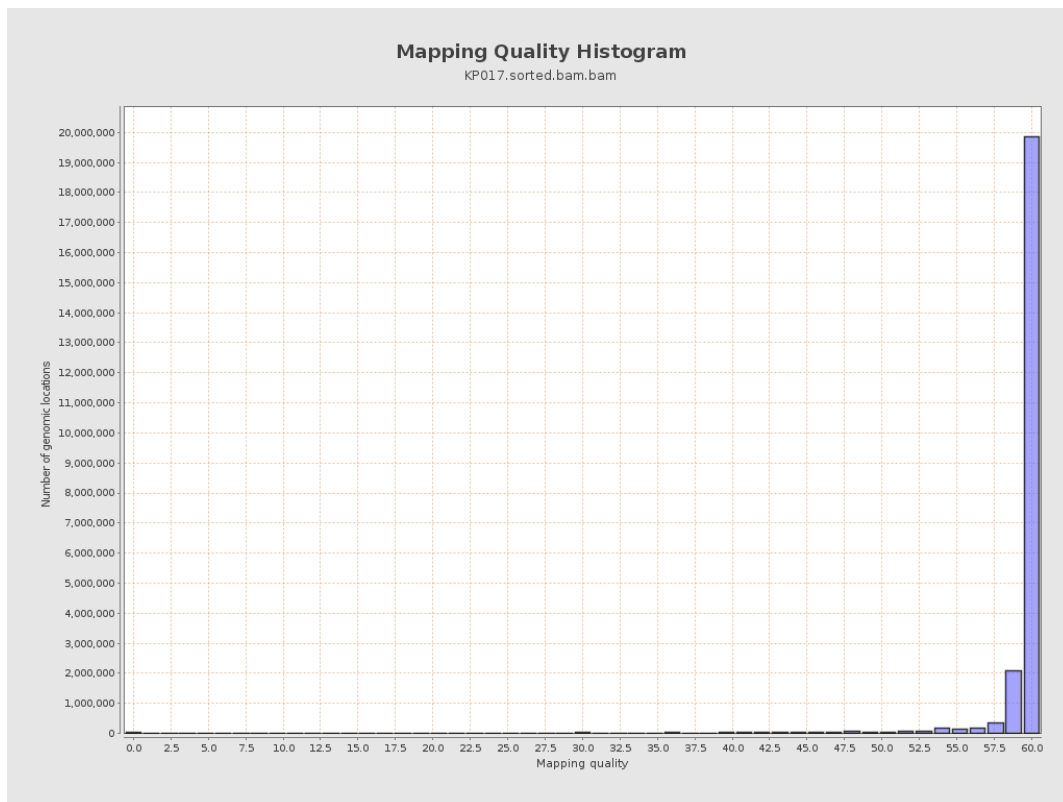
11. Results : Homopolymer Indels



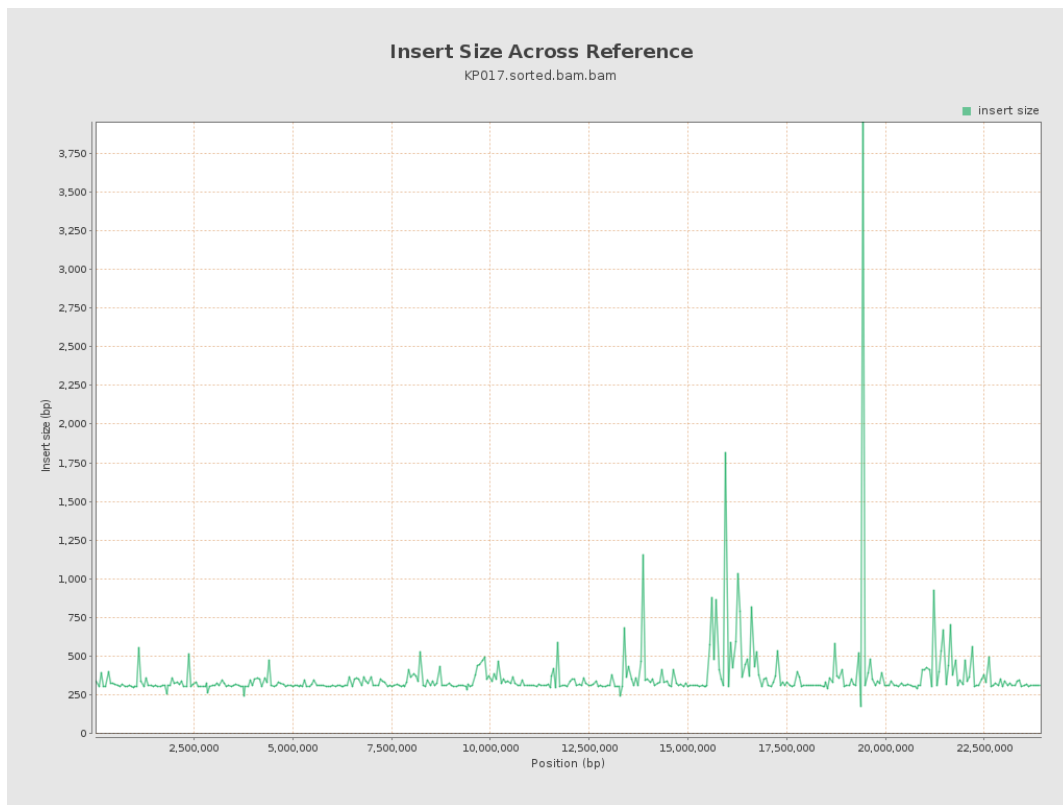
12. Results : Mapping Quality Across Reference



13. Results : Mapping Quality Histogram



14. Results : Insert Size Across Reference



15. Results : Insert Size Histogram

