

# Qualimap Analysis Results

*BAM QC analysis*

*Generated by Qualimap v.2.2.1*

*2016/10/23 13:33:59*

# 1. Input data & parameters

## 1.1. QualiMap command line

```
qualimap bamqc -bam
/home/vdp5/data/cambodia_samples/sequences_bam/OM143.sorted.bam.ba
m -nw 400 -hm 3
```

## 1.2. Alignment

|                                       |  |
|---------------------------------------|--|
| Command line:                         | bwa mem -M<br>/home/vdp5/data/salvador_vivax_asia_2016/first-SAMEA2376790/pvivax_sal1_SAMEA2376790.fasta<br>/home/vdp5/data/cambodia_samples/sequences_gz/OM143-BiooBarcode17_GTAGAG_R2.fastq.gz<br>/home/vdp5/data/cambodia_samples/sequences_gz/OM143-BiooBarcode17_GTAGAG_R1.fastq.gz |
| Draw chromosome limits:               | no   |
| Analyze overlapping paired-end reads: | no   |
| Program:                              | bwa (0.7.15-r1140)   |
| Analysis date:                        | Sun Oct 23 13:33:58 EDT 2016   |
| Size of a homopolymer:                | 3  |
| Skip duplicate alignments:            | no   |
| Number of windows:                    | 400  |
| BAM file:                             | /home/vdp5/data/cambodia_samples/sequences_bam/OM143.sorted.bam.   |

|  |     |
|--|-----|
|  | bam |
|--|-----|

## 2. Summary

### 2.1. Globals

|                              |                    |
|------------------------------|--------------------|
| Reference size               | 23,958,997         |
| Number of reads              | 6,807,050          |
| Mapped reads                 | 5,012,018 / 73.63% |
| Unmapped reads               | 1,795,032 / 26.37% |
| Mapped paired reads          | 5,012,018 / 73.63% |
| Mapped reads, first in pair  | 2,500,820 / 36.74% |
| Mapped reads, second in pair | 2,511,198 / 36.89% |
| Mapped reads, both in pair   | 4,936,099 / 72.51% |
| Mapped reads, singletons     | 75,919 / 1.12%     |
| Read min/max/mean length     | 30 / 100 / 99.86   |
| Duplicated reads (estimated) | 832,858 / 12.24%   |
| Duplication rate             | 14.11%             |
| Clipped reads                | 476,201 / 7%       |

### 2.2. ACGT Content

|                          |                      |
|--------------------------|----------------------|
| Number/percentage of A's | 143,087,867 / 29.4%  |
| Number/percentage of C's | 100,245,739 / 20.6%  |
| Number/percentage of T's | 143,499,186 / 29.49% |
| Number/percentage of G's | 99,808,569 / 20.51%  |
| Number/percentage of N's | 25,127 / 0.01%       |
| GC Percentage            | 41.11%               |

## 2.3. Coverage

|                    |         |
|--------------------|---------|
| Mean               | 20.3313 |
| Standard Deviation | 17.5595 |

## 2.4. Mapping Quality

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

## 2.5. Insert size

|                    |                 |
|--------------------|-----------------|
| Mean               | 838.41          |
| Standard Deviation | 25,640.92       |
| P25/Median/P75     | 301 / 312 / 321 |

## 2.6. Mismatches and indels

|  |           |
|--|-----------|
| General error rate                       | 1.21%     |
| Mismatches                               | 5,563,721 |
| Insertions                               | 140,285   |
| Mapped reads with at least one insertion | 2.66%     |
| Deletions                                | 162,084   |
| Mapped reads with at least one deletion  | 3.07%     |
| Homopolymer indels                       | 64.8%     |

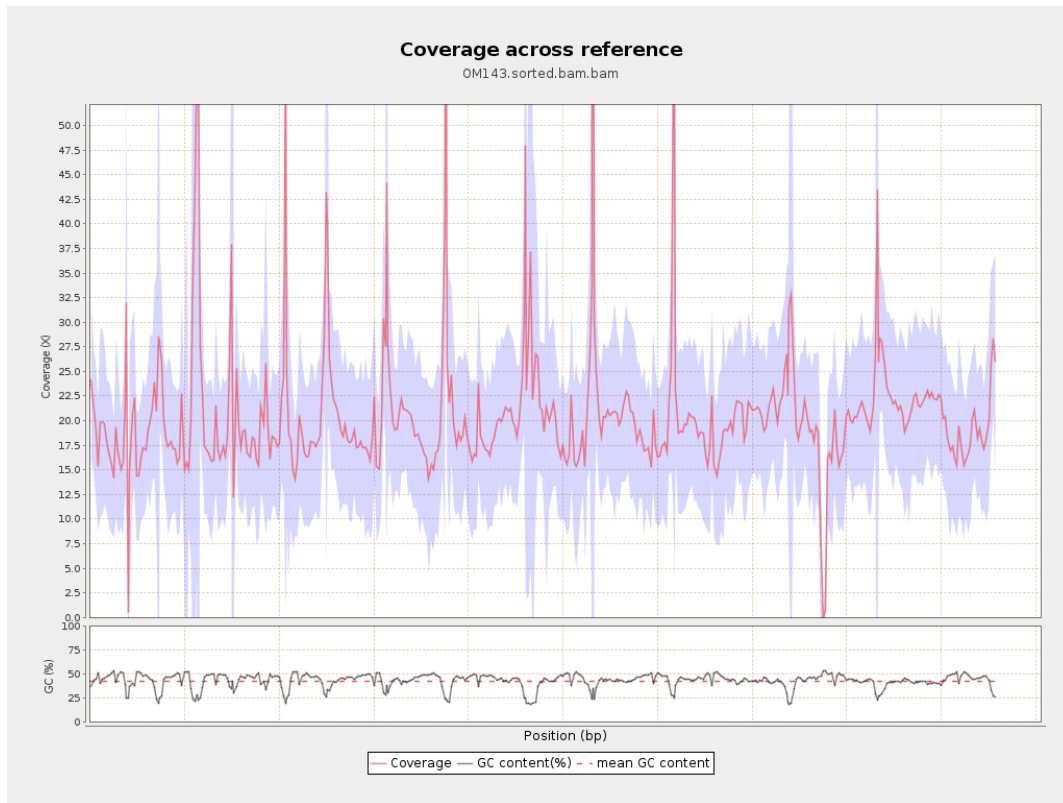
## 2.7. Chromosome stats

| Name | Length | Mapped bases | Mean coverage | Standard deviation |
|------|--------|--------------|---------------|--------------------|
|      |        |              |               |                    |

|                              |         |          |         |         |
|------------------------------|---------|----------|---------|---------|
| gi 1074120478 emb LT615256.1 | 977217  | 18059186 | 18.4802 | 7.913   |
| gi 1074120682 emb LT615257.1 | 860454  | 15697627 | 18.2434 | 12.4583 |
| gi 1074120865 emb LT615258.1 | 989719  | 21160397 | 21.3802 | 20.3772 |
| gi 1074121086 emb LT615259.1 | 935450  | 20625008 | 22.0482 | 28.3674 |
| gi 1074121301 emb LT615260.1 | 1432239 | 28754369 | 20.0765 | 19.5961 |
| gi 1074121615 emb LT615261.1 | 1080962 | 21033818 | 19.4584 | 12.0462 |
| gi 1074121871 emb LT615262.1 | 1545099 | 30539187 | 19.7652 | 7.9481  |
| gi 1074122235 emb LT615263.1 | 1585108 | 32608225 | 20.5716 | 12.1292 |
| gi 1074122590 emb LT615264.1 | 2122358 | 42751588 | 20.1434 | 9.5007  |
| gi 1074123050 emb LT615265.1 | 1754192 | 36847116 | 21.0052 | 14.1088 |
| gi 1074123421 emb LT615      | 2150147 | 46399708 | 21.5798 | 40.6671 |

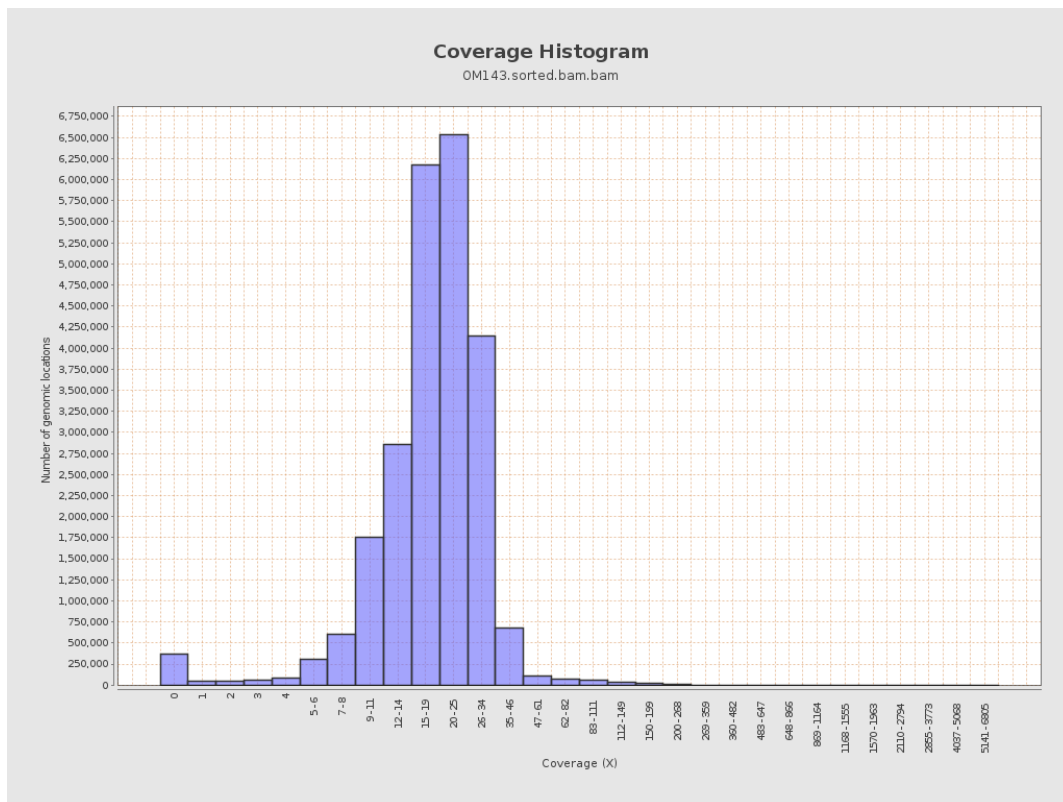
|                                      |         |          |         |         |
|--------------------------------------|---------|----------|---------|---------|
| 266.1                                |         |          |         |         |
| gi 107412389<br>8 emb LT615<br>267.1 | 3031036 | 61514097 | 20.2947 | 8.7891  |
| gi 107412458<br>8 emb LT615<br>268.1 | 2359348 | 45474006 | 19.274  | 14.1254 |
| gi 107412506<br>5 emb LT615<br>269.1 | 3135668 | 65652197 | 20.9372 | 7.349   |

### 3. Results : Coverage across reference

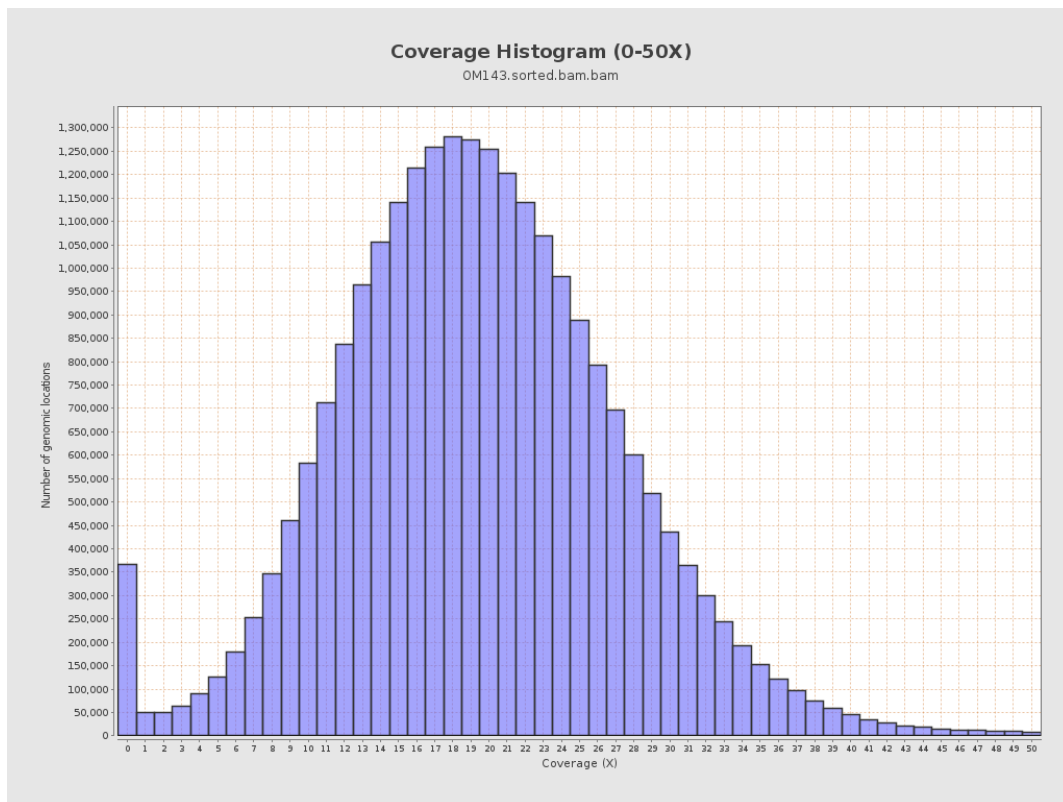




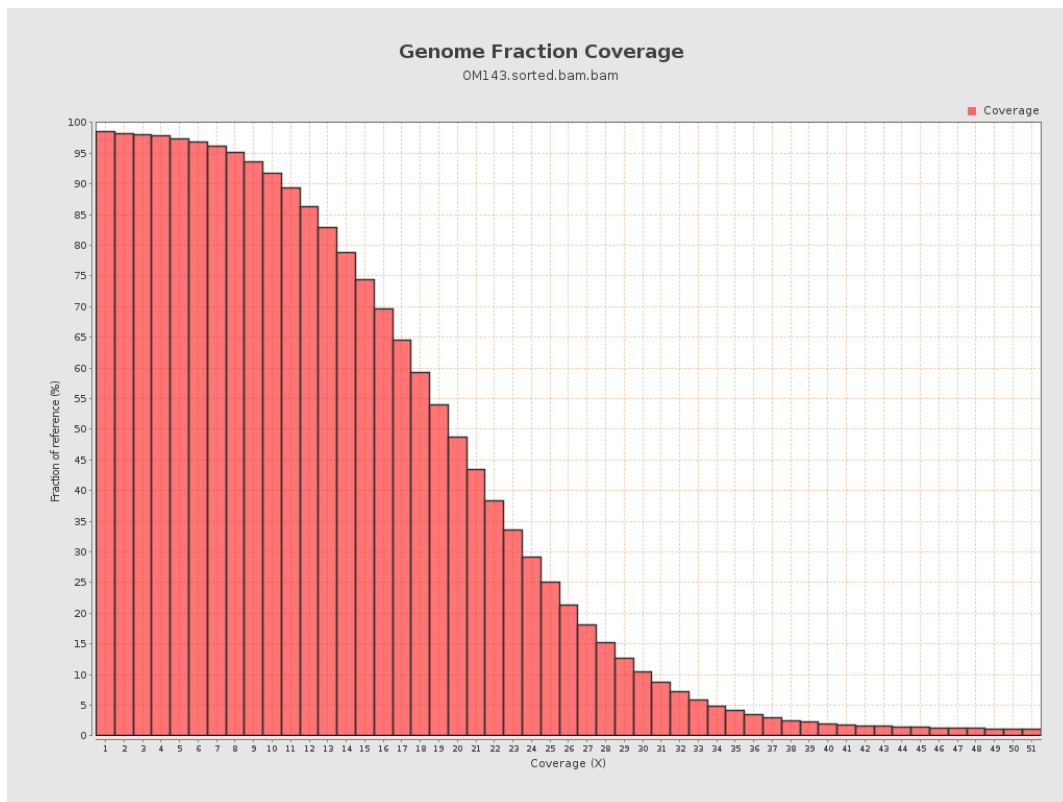
## 4. Results : Coverage Histogram



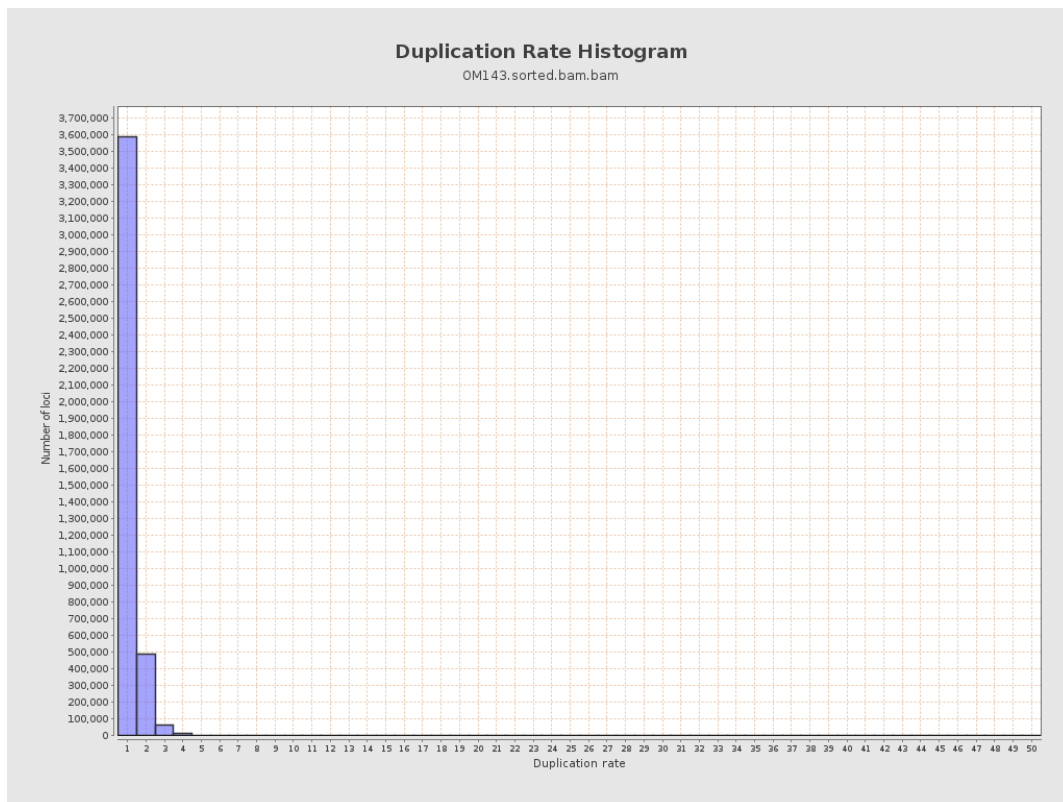
## 5. Results : Coverage Histogram (0-50X)



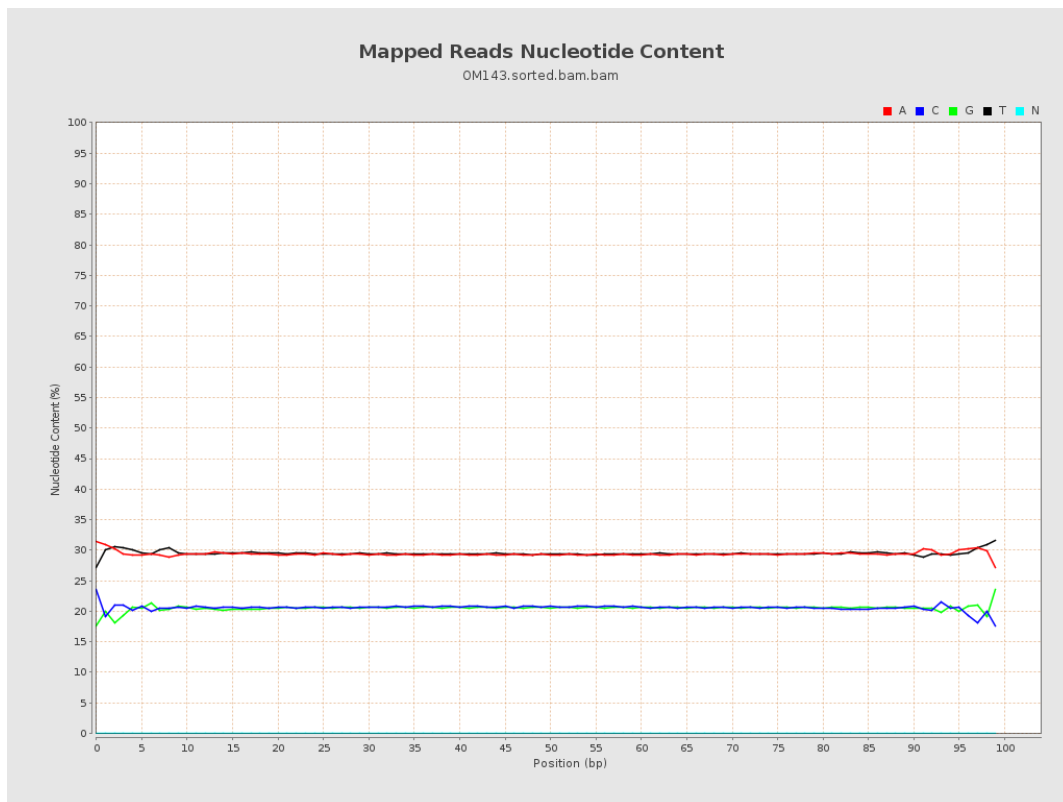
## 6. Results : Genome Fraction Coverage



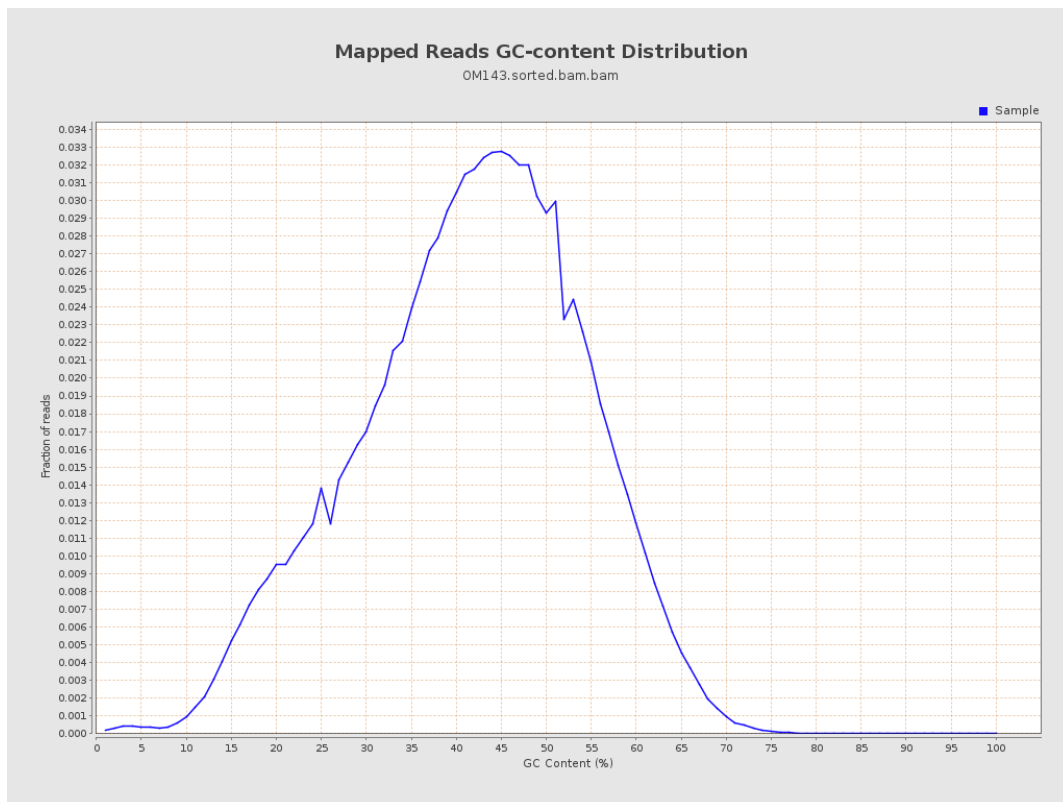
## 7. Results : Duplication Rate Histogram



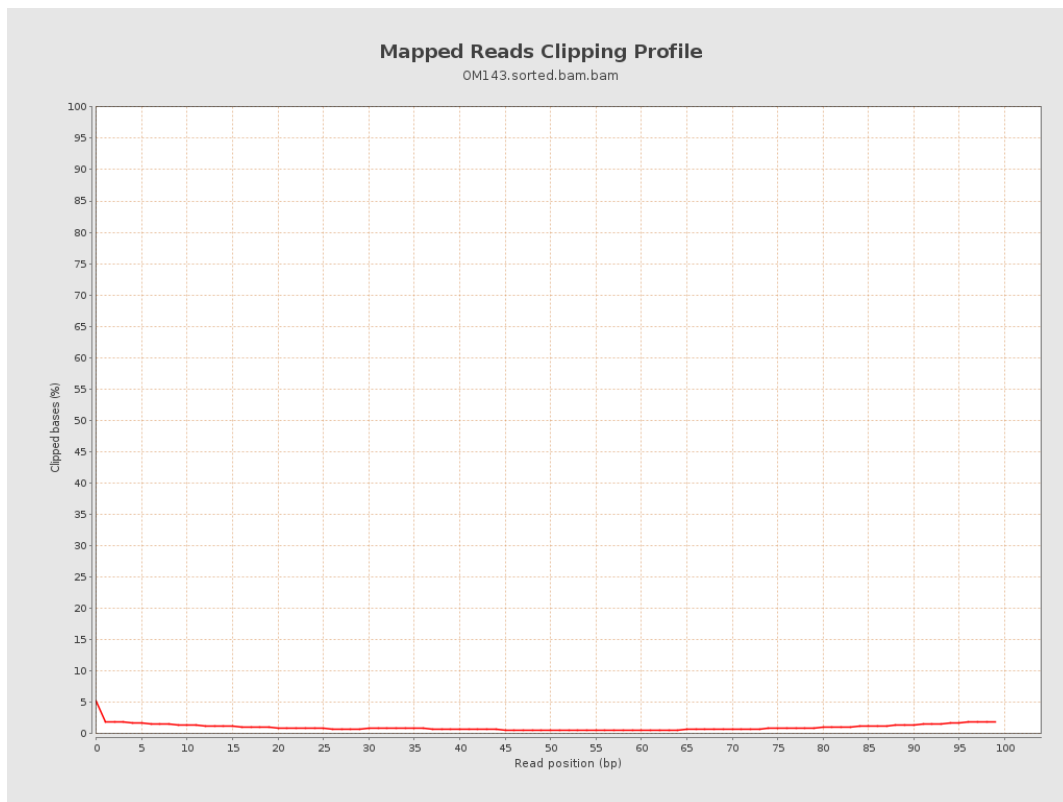
## 8. Results : Mapped Reads Nucleotide Content



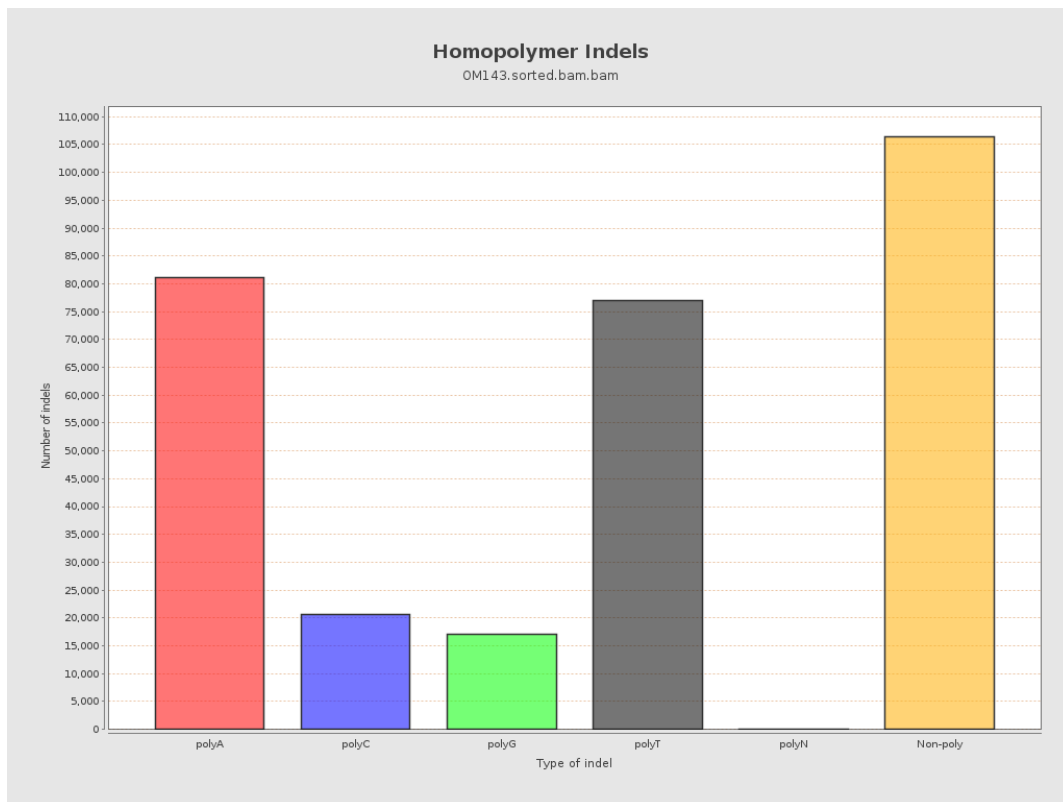
## 9. Results : Mapped Reads GC-content Distribution



## 10. Results : Mapped Reads Clipping Profile

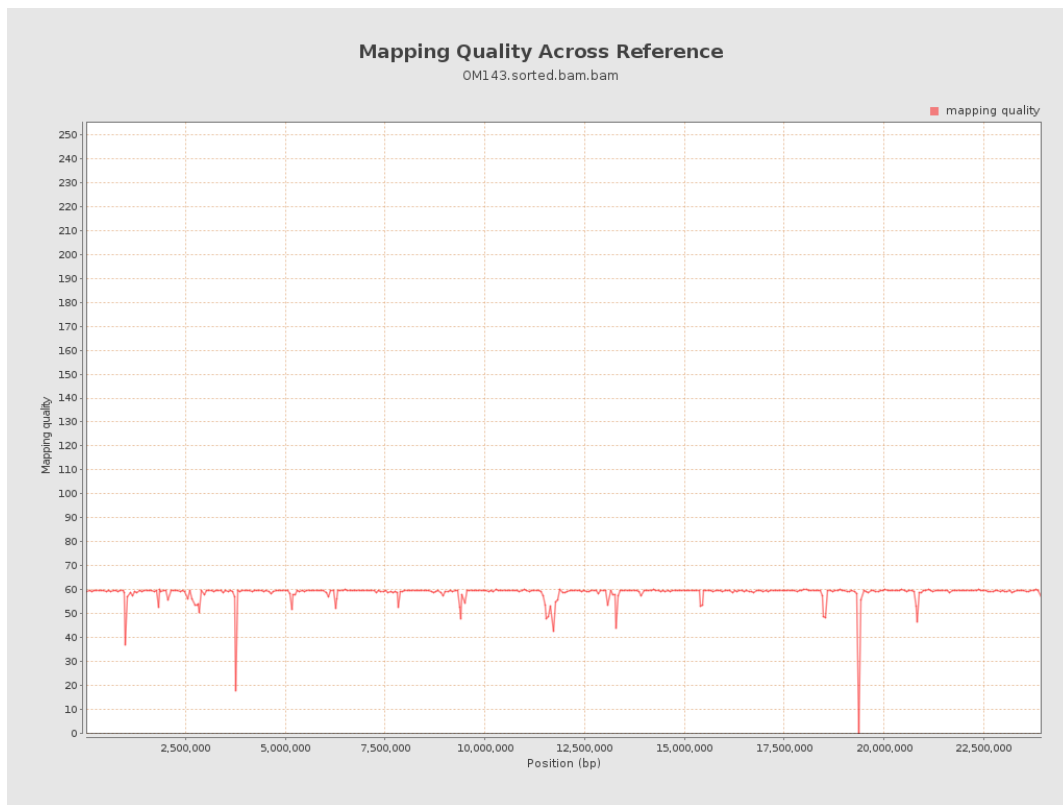


## 11. Results : Homopolymer Indels

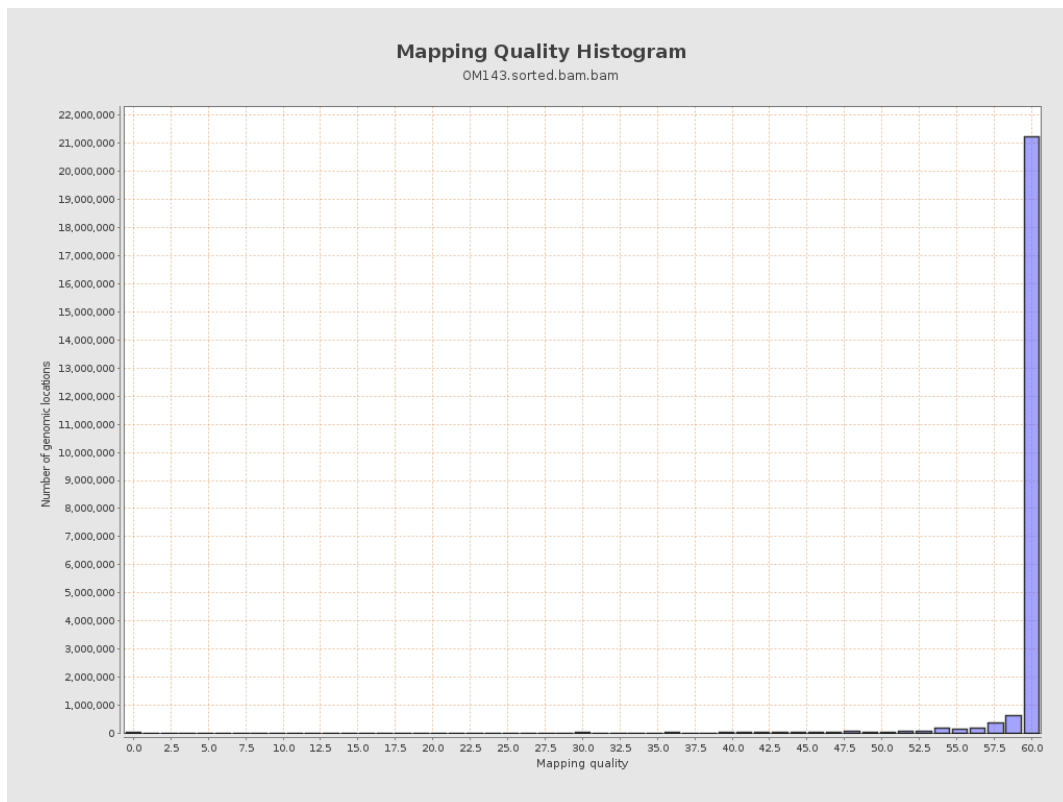




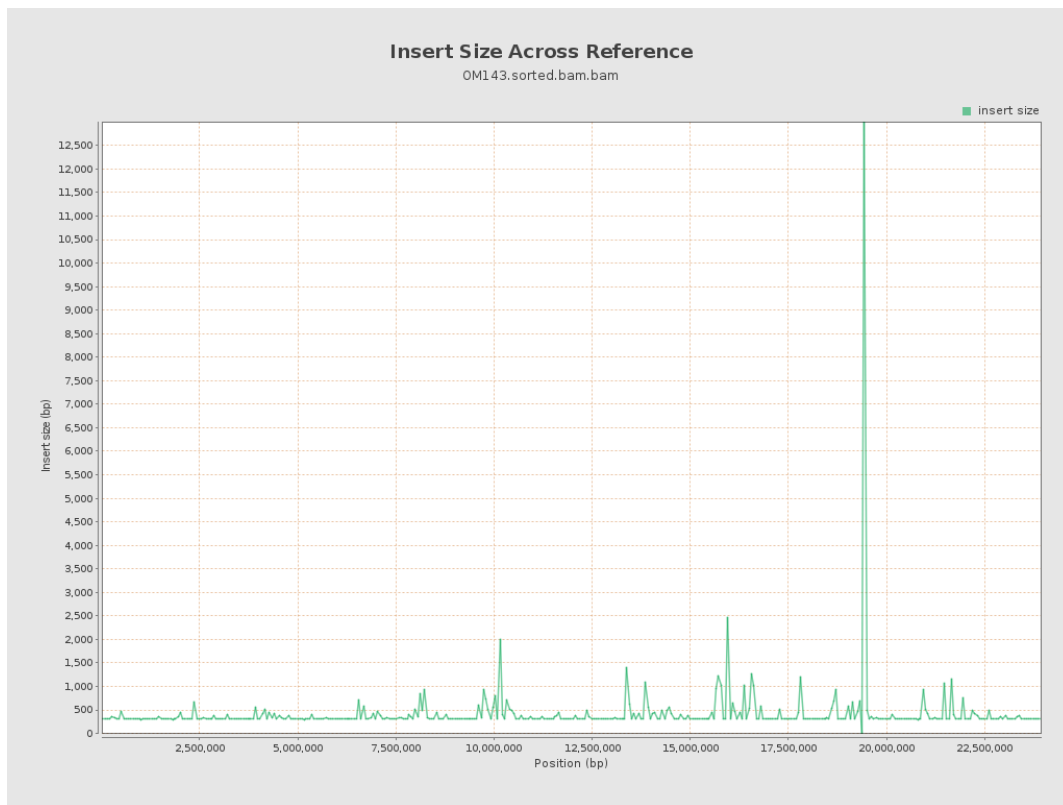
## 12. Results : Mapping Quality Across Reference



## 13. Results : Mapping Quality Histogram



## 14. Results : Insert Size Across Reference



## 15. Results : Insert Size Histogram

