# Supplementary Material: Quality control of alignment sequencing data obtained with Qualimap 2 (Okonechnikov et al., 2016).

# Qualimap Analysis Results

*BAM QC analysis*

*Generated by Qualimap v.2.2*

UY/11/CA/18

# 1. Input data & parameters

## 1.1. QualiMap command line

**UY/11/CA/18.sort.bam**

## 1.2. Alignment

| | |
|---|---|
| Draw chromosome limits: | no |
| Analyze overlapping paired-end reads: | no |
| Program: | bowtie2 (2.1.0) |
| Analysis date: | Thu May 19 16:29:48 GMT-03:00 2016 |
| Size of a homopolymer: | 3 |
| Skip duplicate alignments: | no |
| Number of windows: | 400 |
| BAM file: | UY/11/CA/18.sort.bam |

# 2. Summary

## 2.1. Globals

| | |
|---|---|
| Reference size | 27.636 |
| Number of reads | 2.278.516 |
| Mapped reads | 1.967.637 / 86,36% |
| Unmapped reads | 310.879 / 13,64% |
| Mapped paired reads | 1.967.637 / 86,36% |
| Mapped reads, first in pair | 998.910 / 43,84% |
| Mapped reads,  second in pair | 968.727 / 42,52% |
| Mapped reads, both in pair | 1.886.066 / 82,78% |
| Mapped reads, singletons | 81.571 / 3,58% |
| Read min/max/mean length | 20 / 300 / 169,85 |
| Duplicated reads (estimated) | 1.942.330 / 85,25% |
| Duplication rate | 98,78% |
| Clipped reads | 0 / 0% |

## 2.2. ACGT Content

| | |
|---|---|
| Number/percentage of A's | 96.958.906 / 28,78% |
| Number/percentage of C's | 54.668.304 / 16,23% |
| Number/percentage of T's | 110.782.702 / 32,88% |
| Number/percentage of G's | 74.491.179 / 22,11% |
| Number/percentage of N's | 0 / 0% |
| GC Percentage | 38,34% |

## 2.3. Coverage

| | |
|---|---|
| Mean | 12.196,4277 |
| Standard Deviation | 5.743,8155 |

## 2.4. Mapping Quality

| | |
|---|---|
| Mean Mapping Quality | 23,16 |

## 2.5. Insert size

| | |
|---|---|
| Mean | 525,24 |
| Standard Deviation | 163,05 |
| P25/Median/P75 | 409 / 513 / 630 |

## 2.6. Mismatches and indels

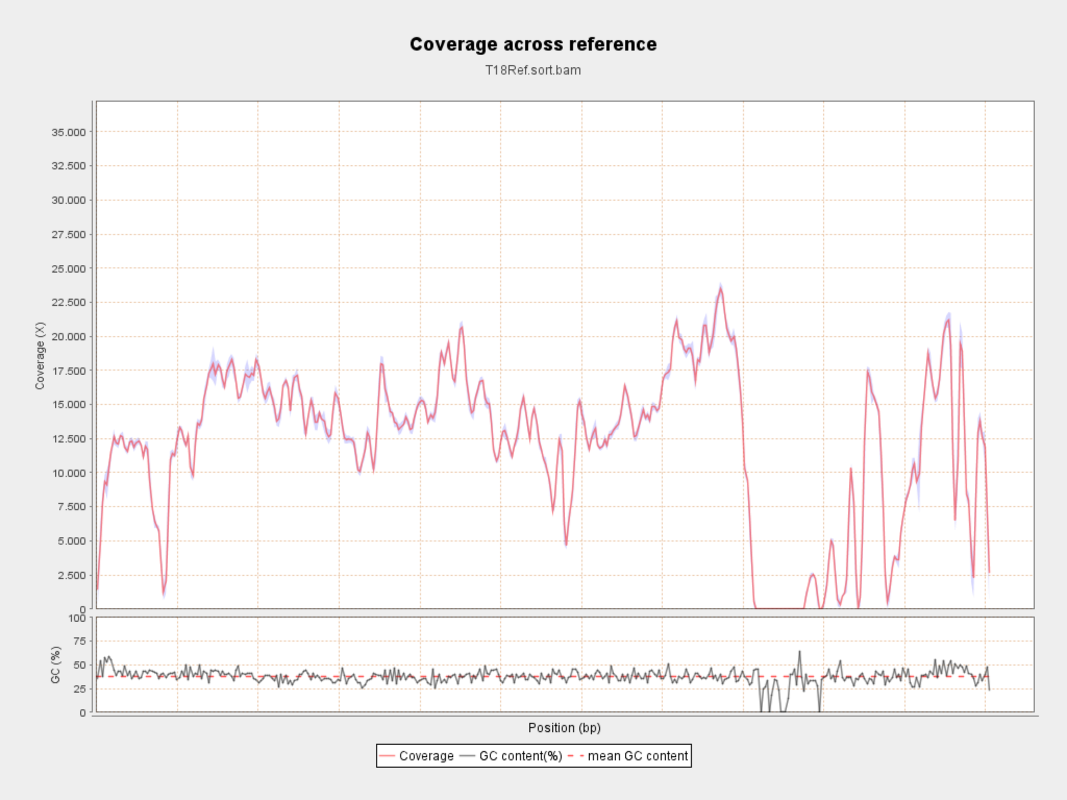
| | |
|---|---|
| General error rate | 4,07% |
| Mismatches | 13.491.862 |
| Insertions | 96.914 |
| Mapped reads with at least one insertion | 4,09% |
| Deletions | 45.195 |
| Mapped reads with at least one deletion | 2,1% |
| Homopolymer indels | 36,92% |

## 2.7. Chromosome stats

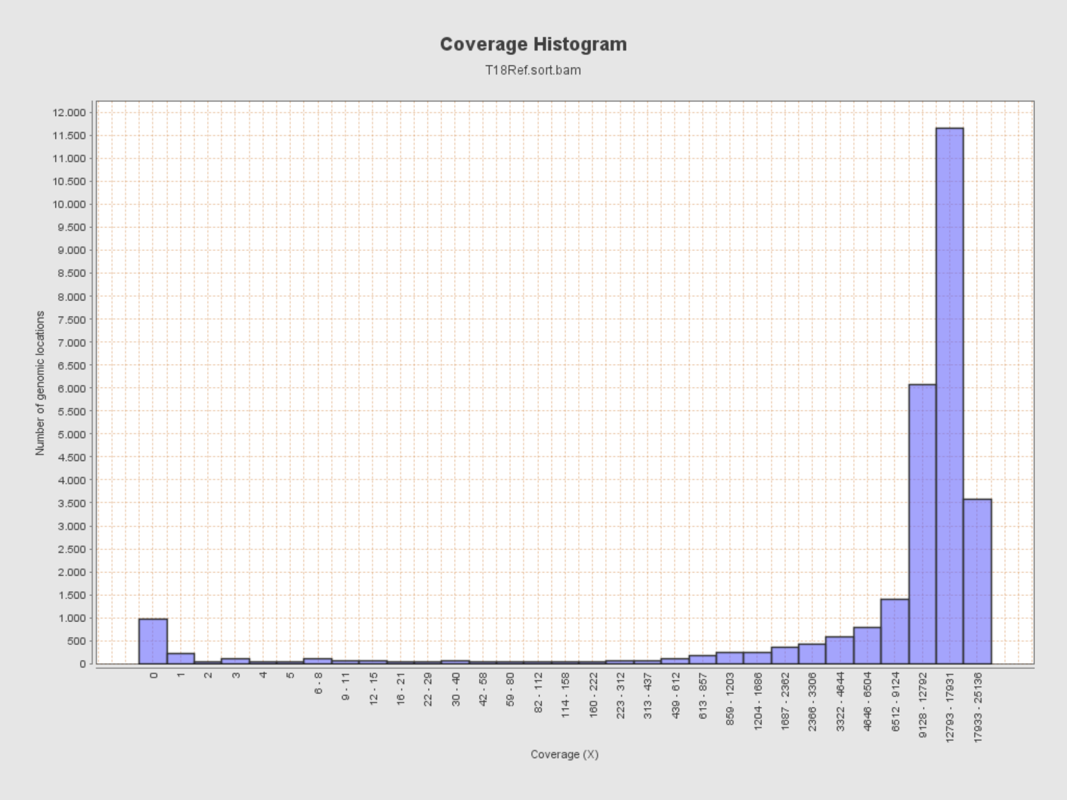
| **Name** | **Length** | **Mapped bases** | **Mean coverage** | **Standard deviation** |
|---|---|---|---|---|
| | | | | |

| | | | | |
|---|---|---|---|---|
| gi\|171988306\|gb\|EU418976.1\|InfectiousbronchitisvirusstrainArkDPI11,completegenome | 27636 | 337060477 | 12.196,4277 | 5.743,8155 |

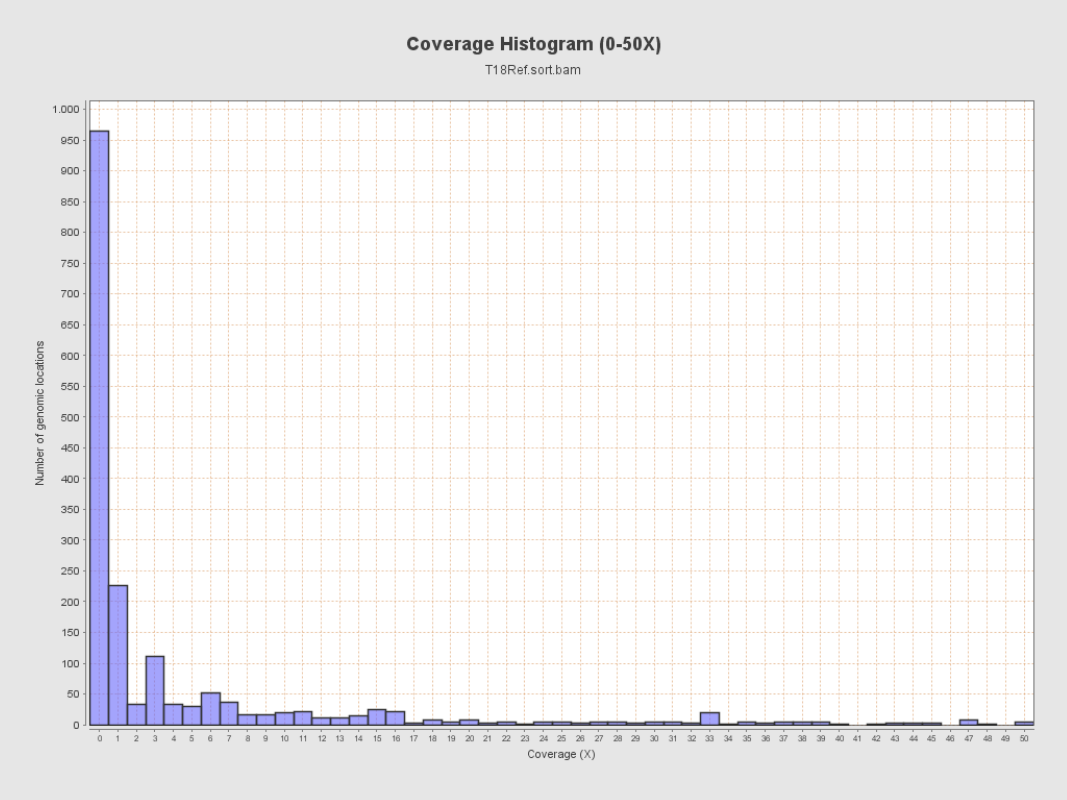
# 3. Results : Coverage across reference



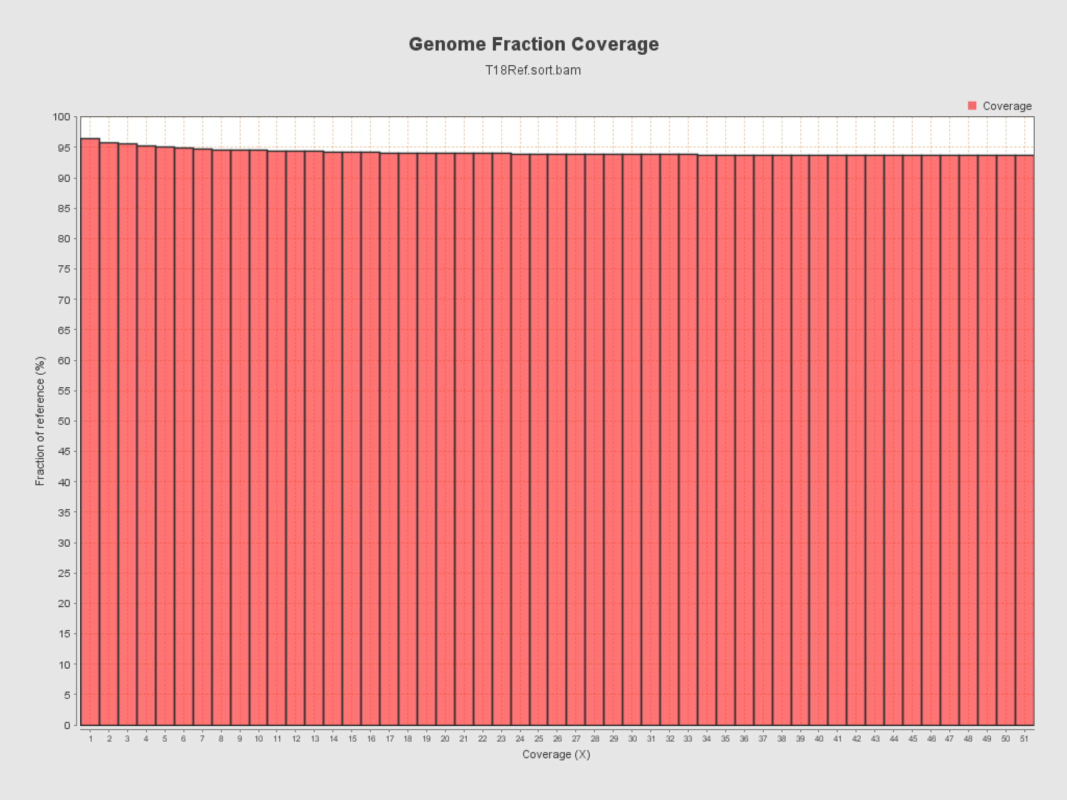
# 4. Results : Coverage Histogram



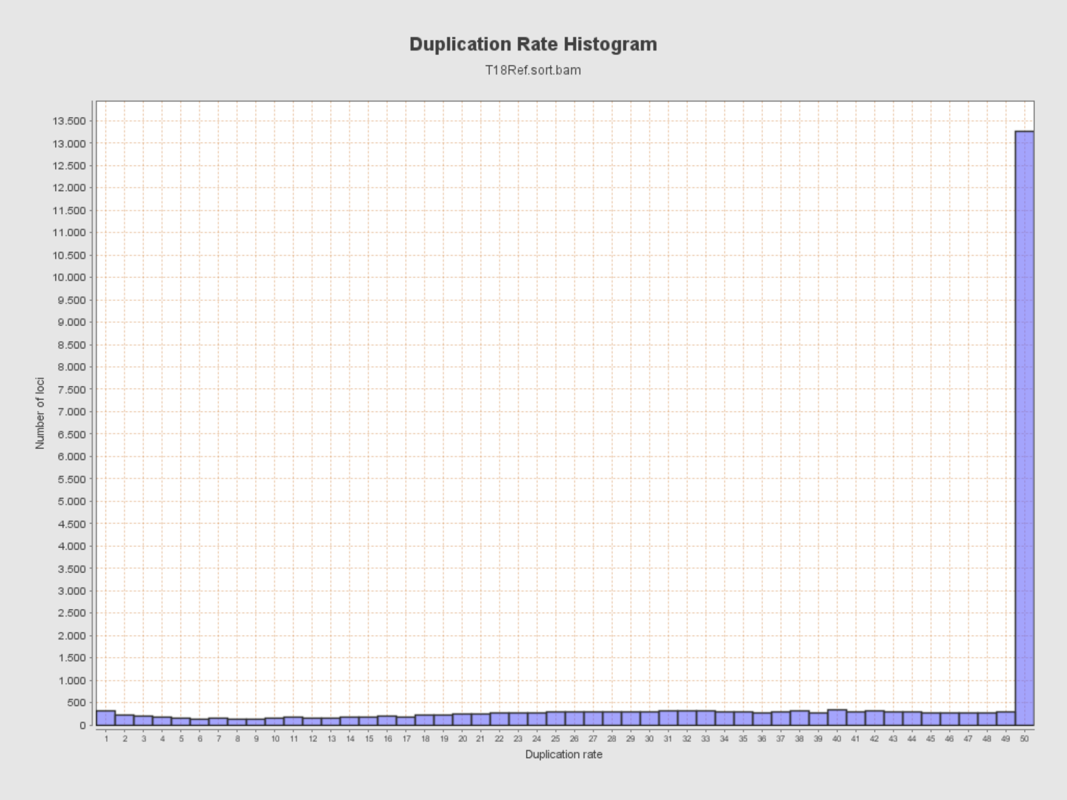
# 5. Results : Coverage Histogram (0-50X)



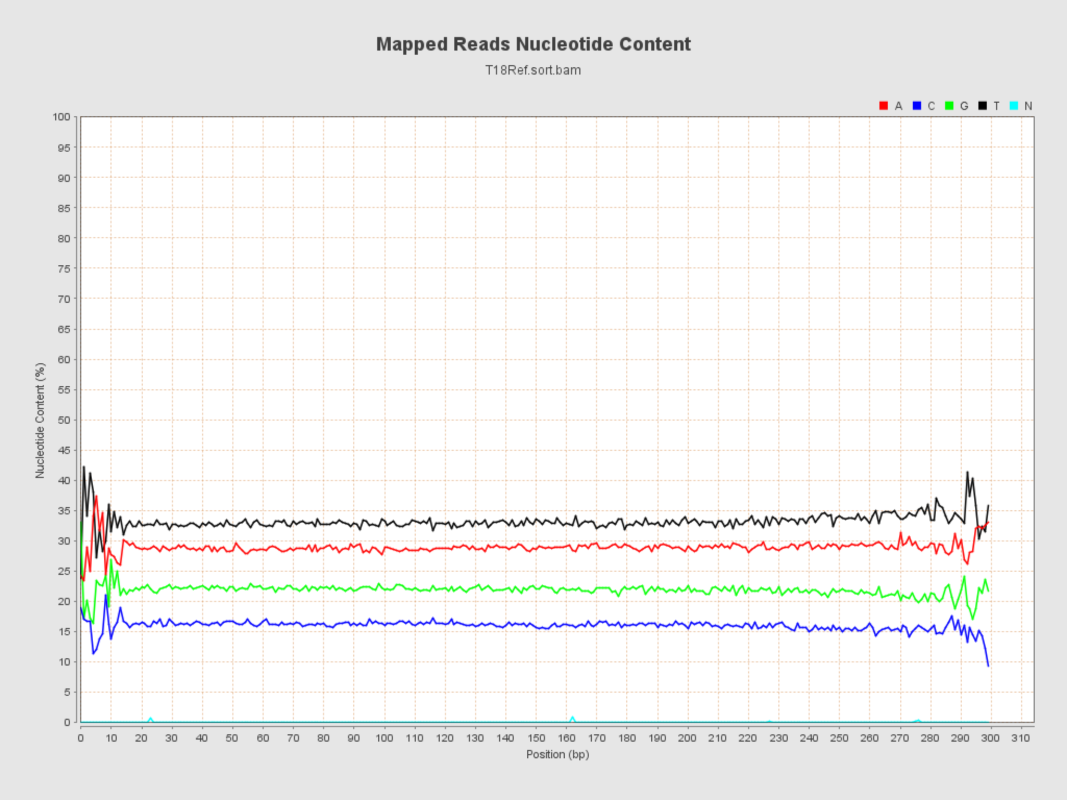
# 6. Results : Genome Fraction Coverage



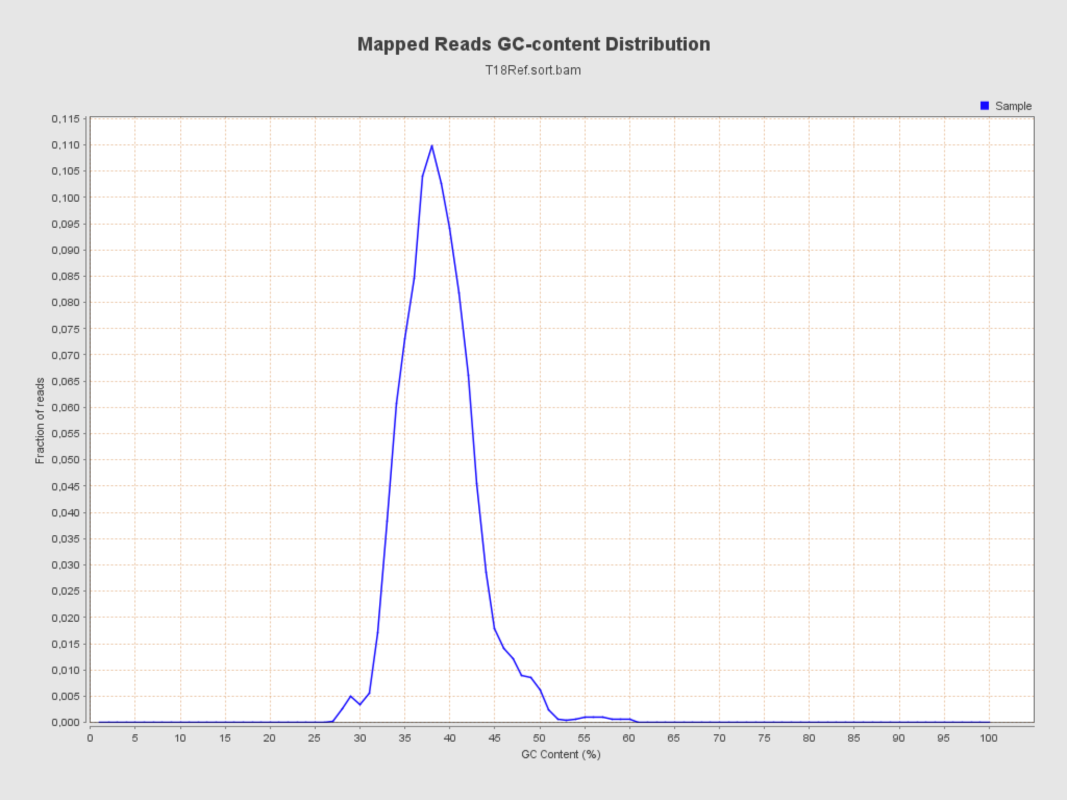
# 7. Results : Duplication Rate Histogram



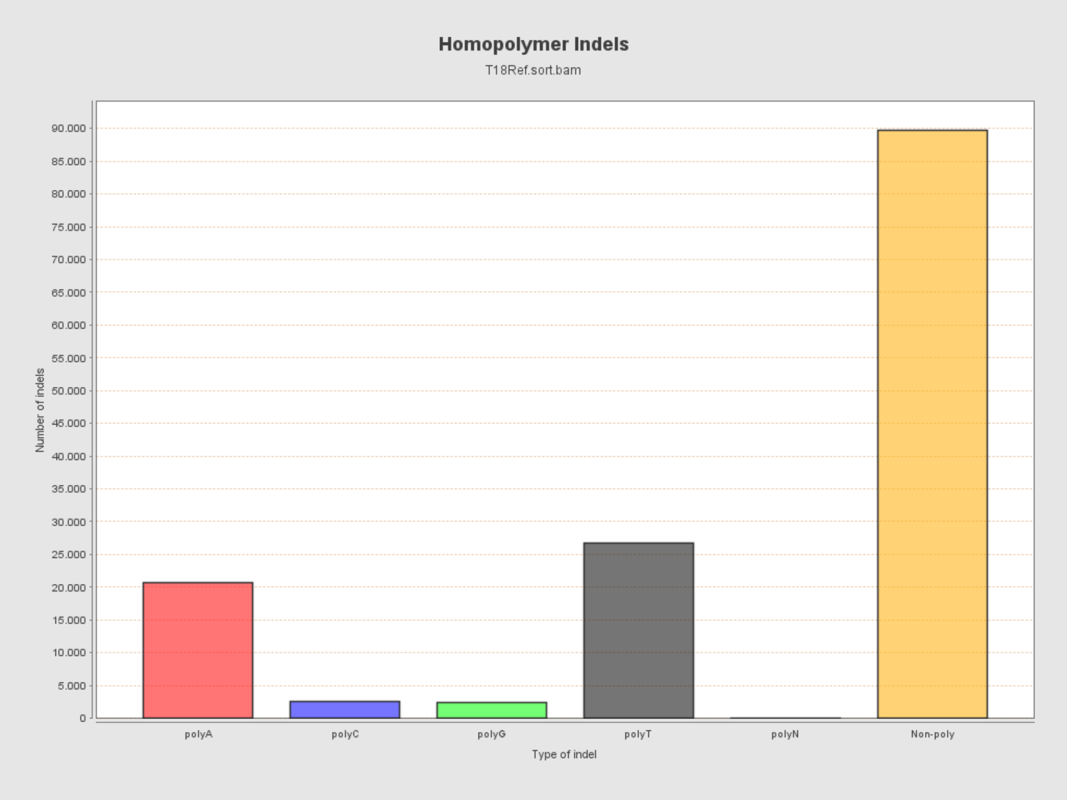
# 8. Results : Mapped Reads Nucleotide Content



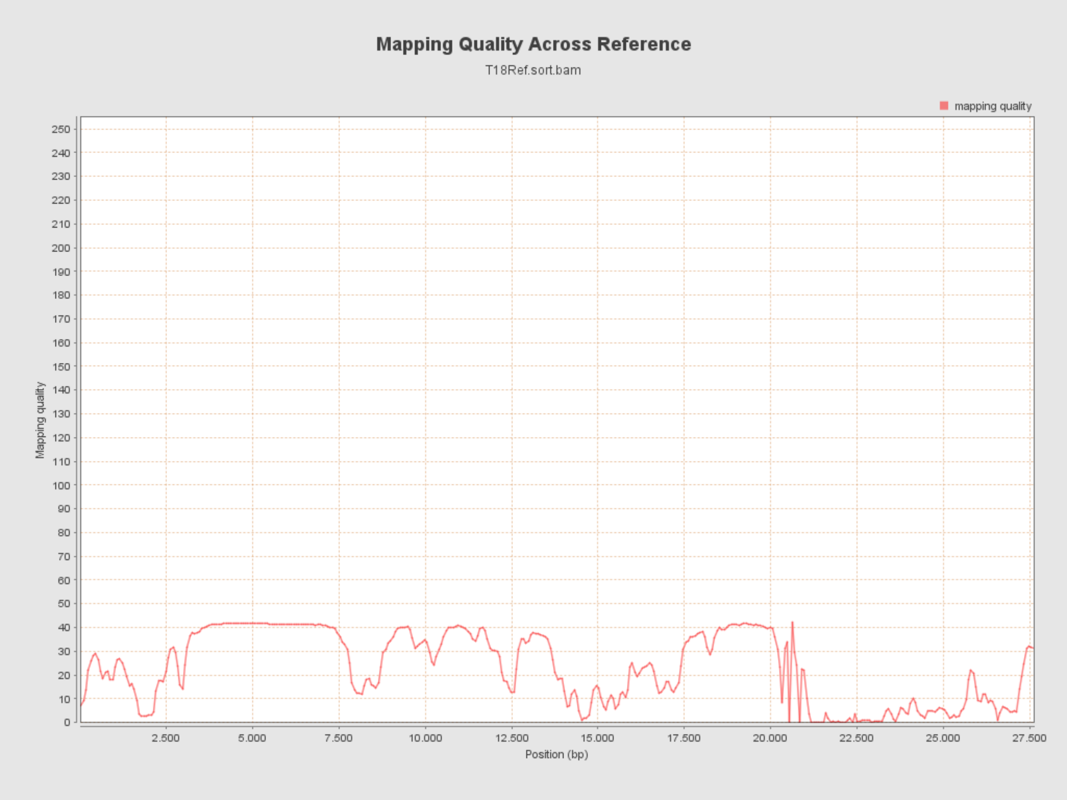
# 9. Results : Mapped Reads GC-content Distribution



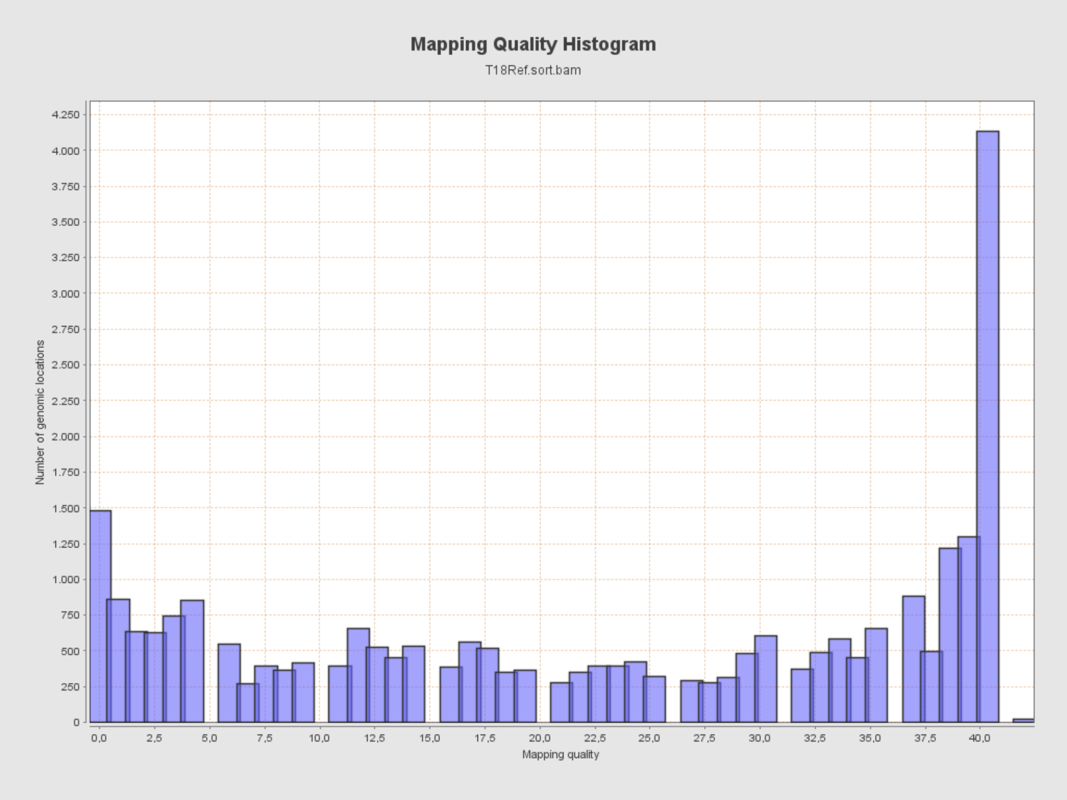
# 10. Results : Homopolymer Indels



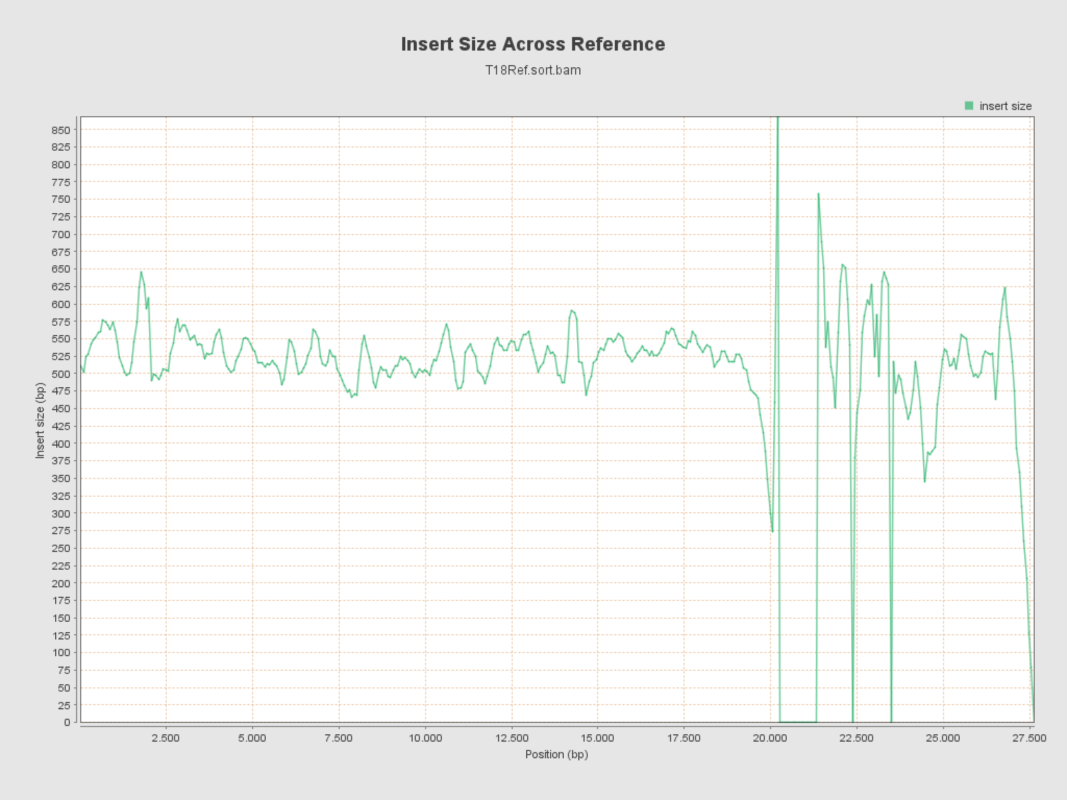
# 11. Results : Mapping Quality Across Reference



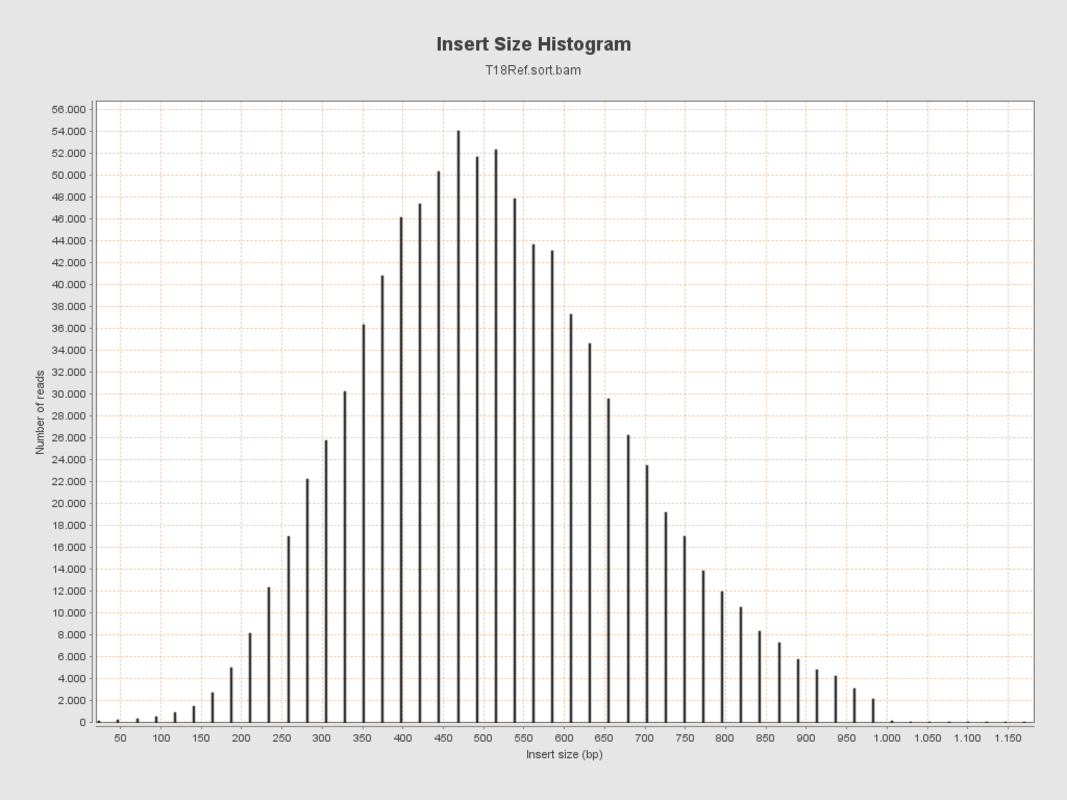
# 12. Results : Mapping Quality Histogram



# 13. Results : Insert Size Across Reference



# 14. Results : Insert Size Histogram



# Qualimap Analysis Results

*BAM QC analysis*

*Generated by Qualimap v.2.2*

UY/09/CA/01

# 1. Input data & parameters

## 1.1. QualiMap command line

**UY/ 09/CA/01.sort.bam**

## 1.2. Alignment

| | |
|---|---|
| Draw chromosome limits: | no |
| Analyze overlapping paired-end reads: | yes |
| Program: | bowtie2 (2.1.0) |
| Analysis date: | Thu May 19 16:34:27 GMT-03:00 2016 |
| Size of a homopolymer: | 3 |
| Skip duplicate alignments: | no |
| Number of windows: | 400 |
| BAM file: | UY/09/CA/01.sort.bam |

# 2. Summary

## 2.1. Globals

| | |
|---|---|
| Reference size | 27.370 |
| Number of reads | 3.723.022 |
| Mapped reads | 2.932.135 / 78,76% |
| Unmapped reads | 790.887 / 21,24% |
| Mapped paired reads | 2.932.135 / 78,76% |
| Mapped reads, first in pair | 1.493.733 / 40,12% |
| Mapped reads,  second in pair | 1.438.402 / 38,64% |
| Mapped reads, both in pair | 2.809.452 / 75,46% |
| Mapped reads, singletons | 122.683 / 3,3% |
| Read min/max/mean length | 20 / 300 / 176,31 |
| Overlapping read pairs | 897.273 / 48,2% |
| Duplicated reads (estimated) | 2.906.485 / 78,07% |
| Duplication rate | 98,62% |
| Clipped reads | 0 / 0% |

## 2.2. ACGT Content

| | |
|---|---|
| Number/percentage of A's | 151.104.114 / 28,91% |
| Number/percentage of C's | 85.627.002 / 16,38% |
| Number/percentage of T's | 173.307.368 / 33,15% |
| Number/percentage of G's | 112.683.228 / 21,56% |
| Number/percentage of N's | 0 / 0% |
| GC Percentage | 37,94% |

## 2.3. Coverage

| | |
|---|---|
| Mean | 19.104,9693 |
| Standard Deviation | 8.467,8684 |
| Mean (paired-end reads overlap ignored) | 15.332,25 |

## 2.4. Mapping Quality

| | |
|---|---|
| Mean Mapping Quality | 22,61 |

## 2.5. Insert size

| | |
|---|---|
| Mean | 387,96 |
| Standard Deviation | 145,67 |
| P25/Median/P75 | 285 / 359 / 463 |

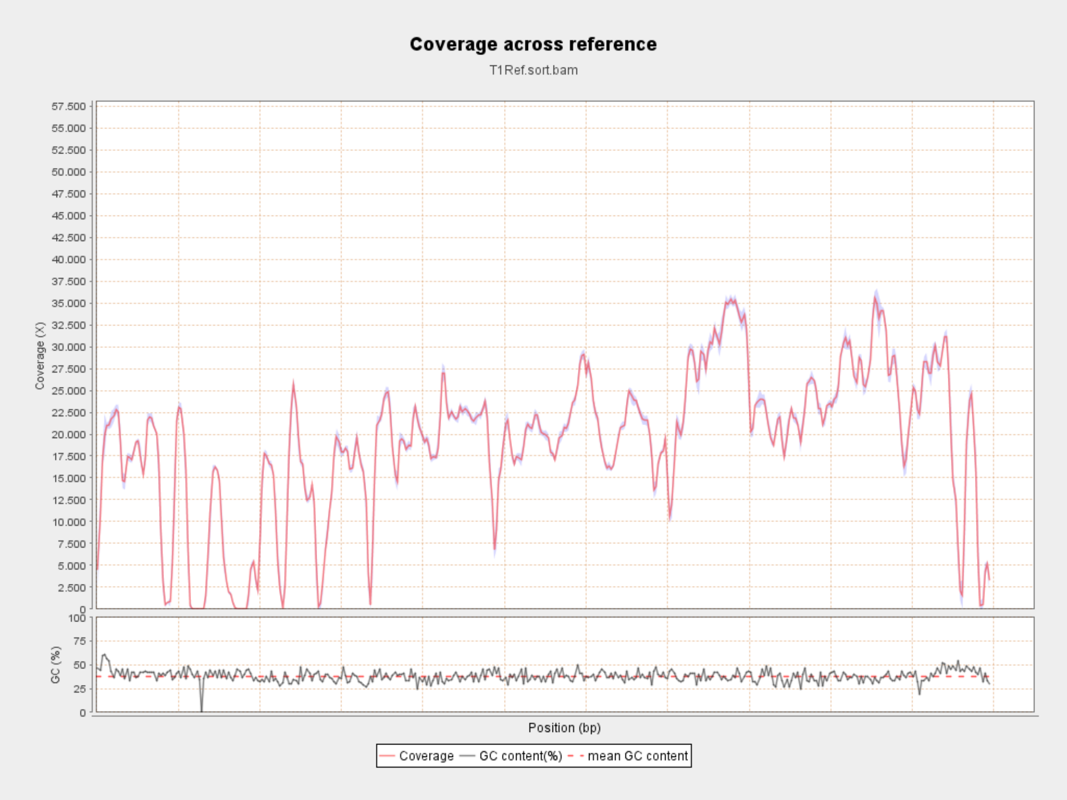
## 2.6. Mismatches and indels

| | |
|---|---|
| General error rate | 4,22% |
| Mismatches | 21.609.966 |
| Insertions | 163.918 |
| Mapped reads with at least one insertion | 4,39% |
| Deletions | 79.815 |
| Mapped reads with at least one deletion | 2,59% |
| Homopolymer indels | 52,49% |

## 2.7. Chromosome stats

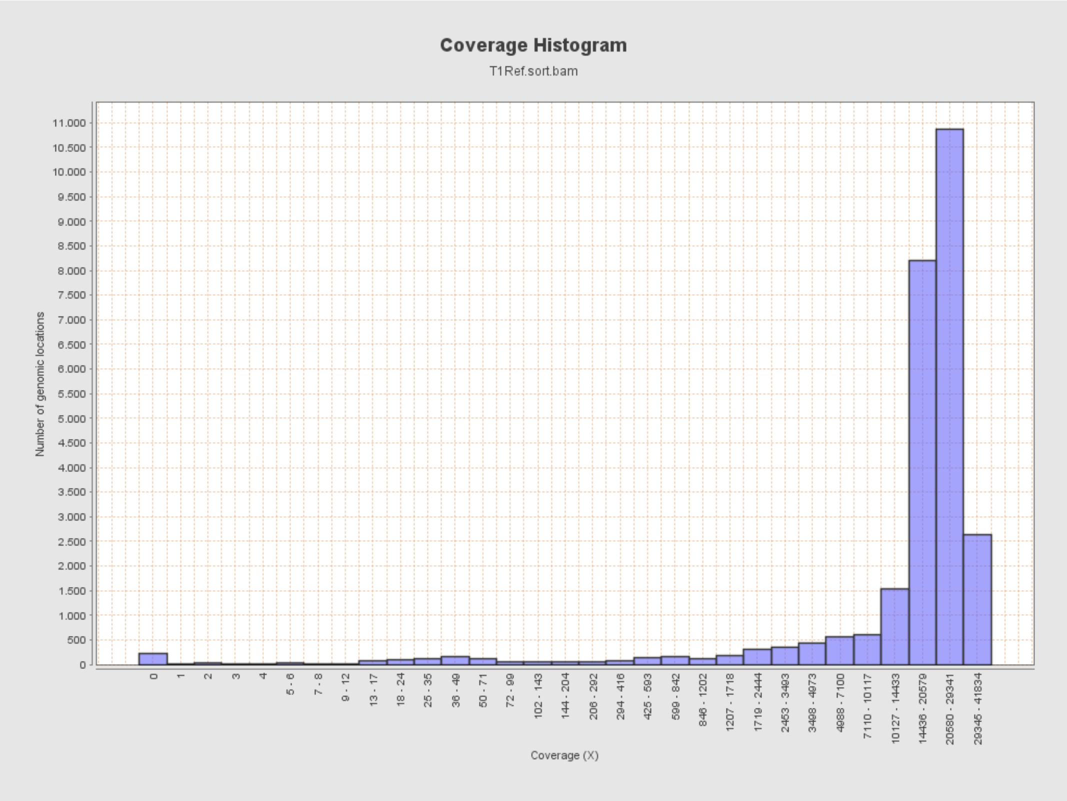
| **Name** | **Length** | **Mapped bases** | **Mean coverage** | **Standard deviation** |
|---|---|---|---|---|
| gi\|914339339\|gb\|KP780179.1\|Infectiousbronchitisvirusstraingamma CoV/Ck/Italy/I2022/13,completegenome | 27370 | 522903011 | 19.104,9693 | 8.467,8684 |

# 3. Results : Coverage across reference

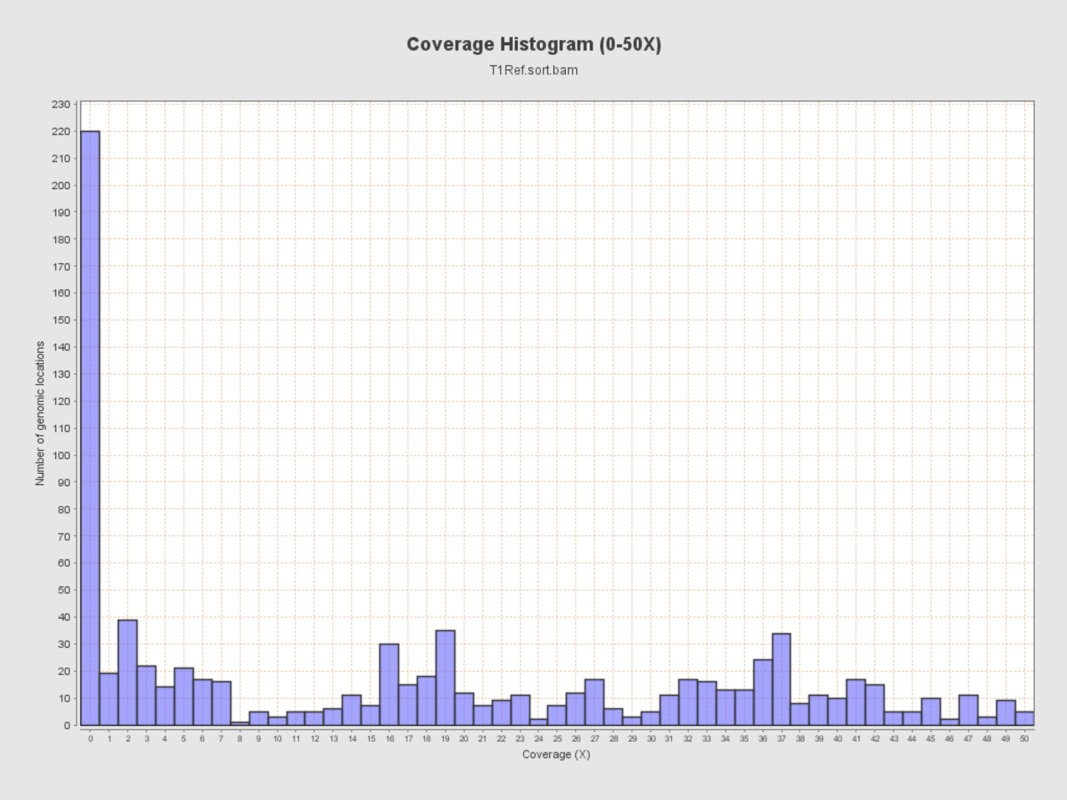


]

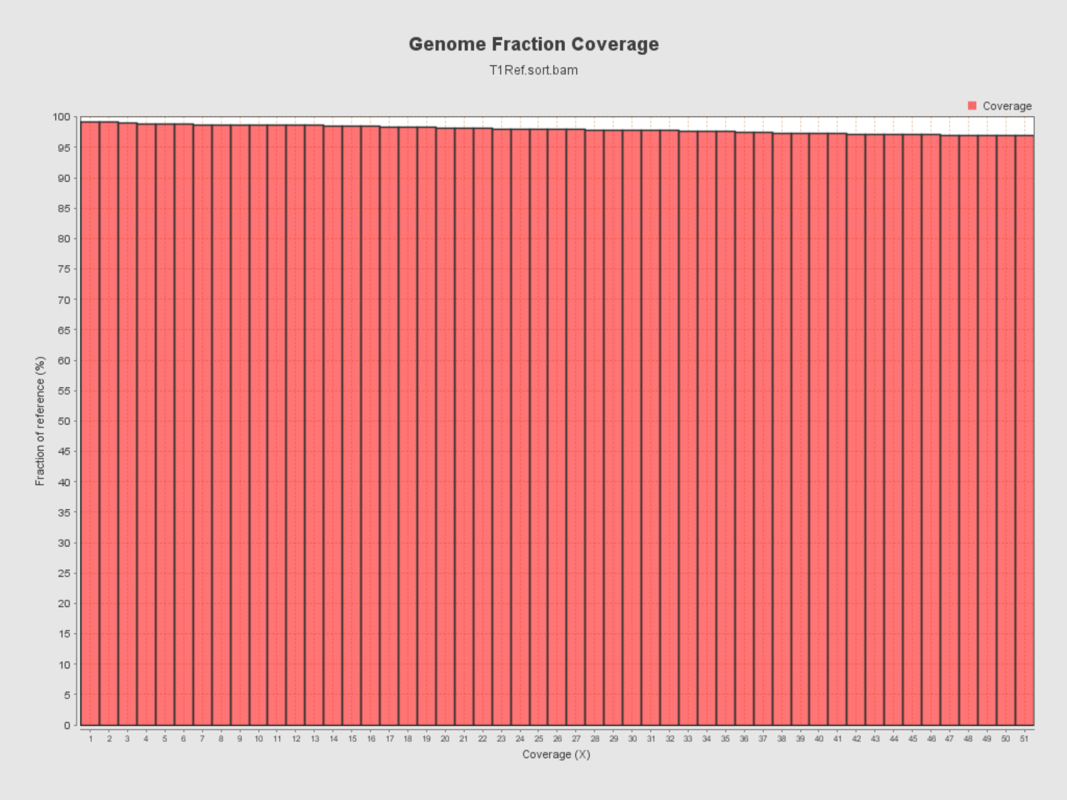
# 4. Results : Coverage Histogram



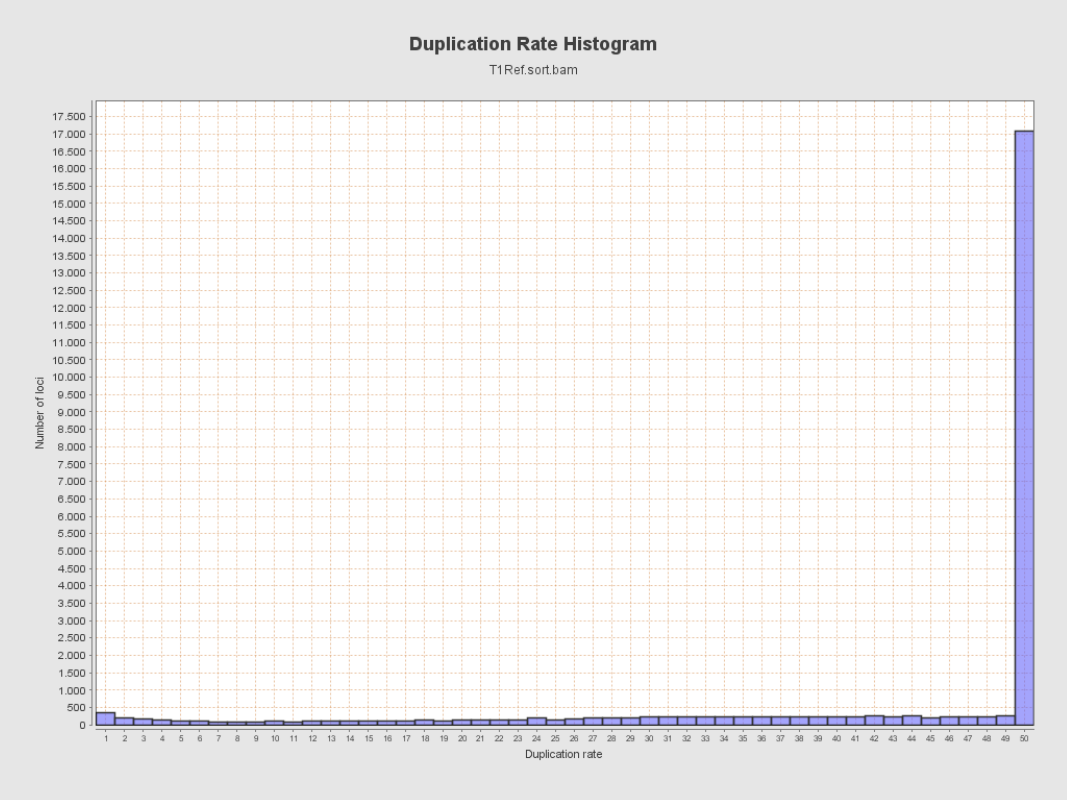
# 5. Results : Coverage Histogram (0-50X)



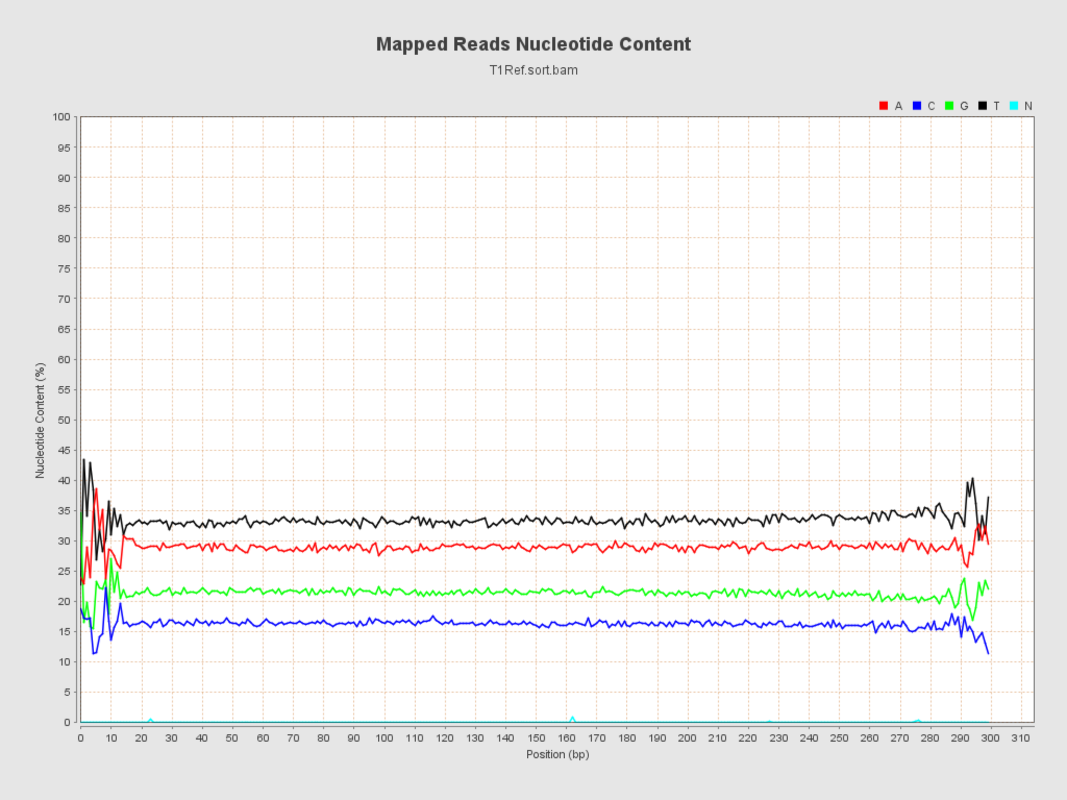
# 6. Results : Genome Fraction Coverage



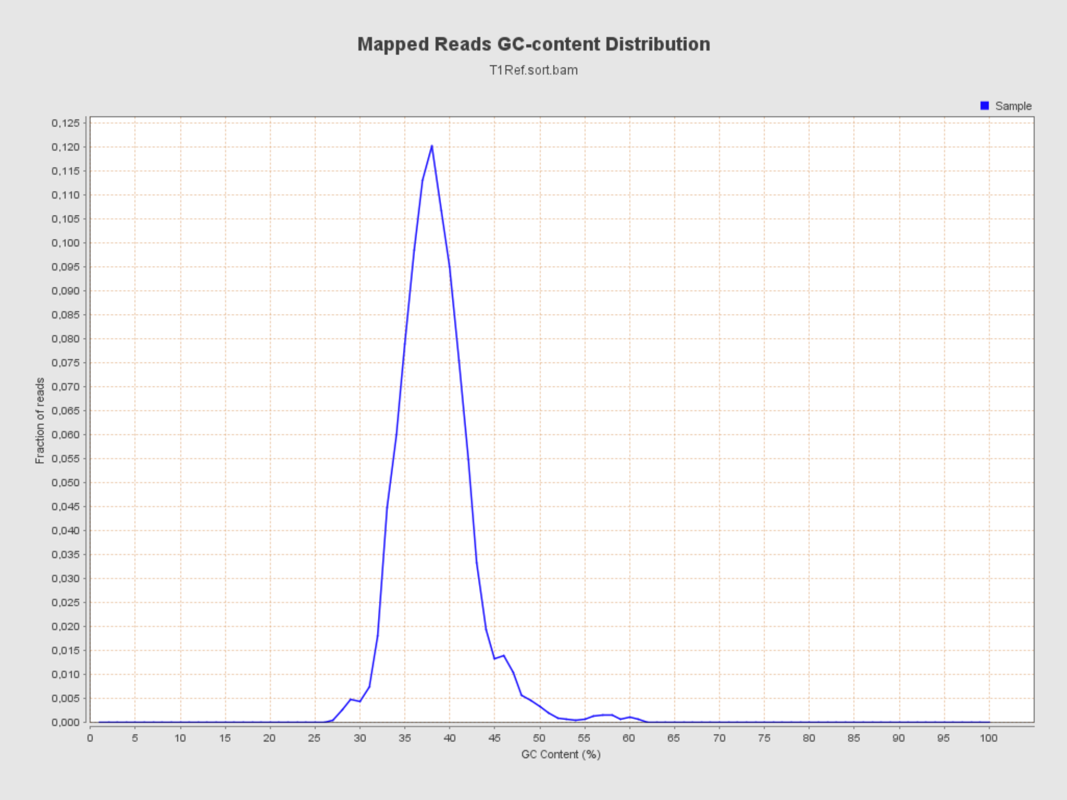
# 7. Results : Duplication Rate Histogram



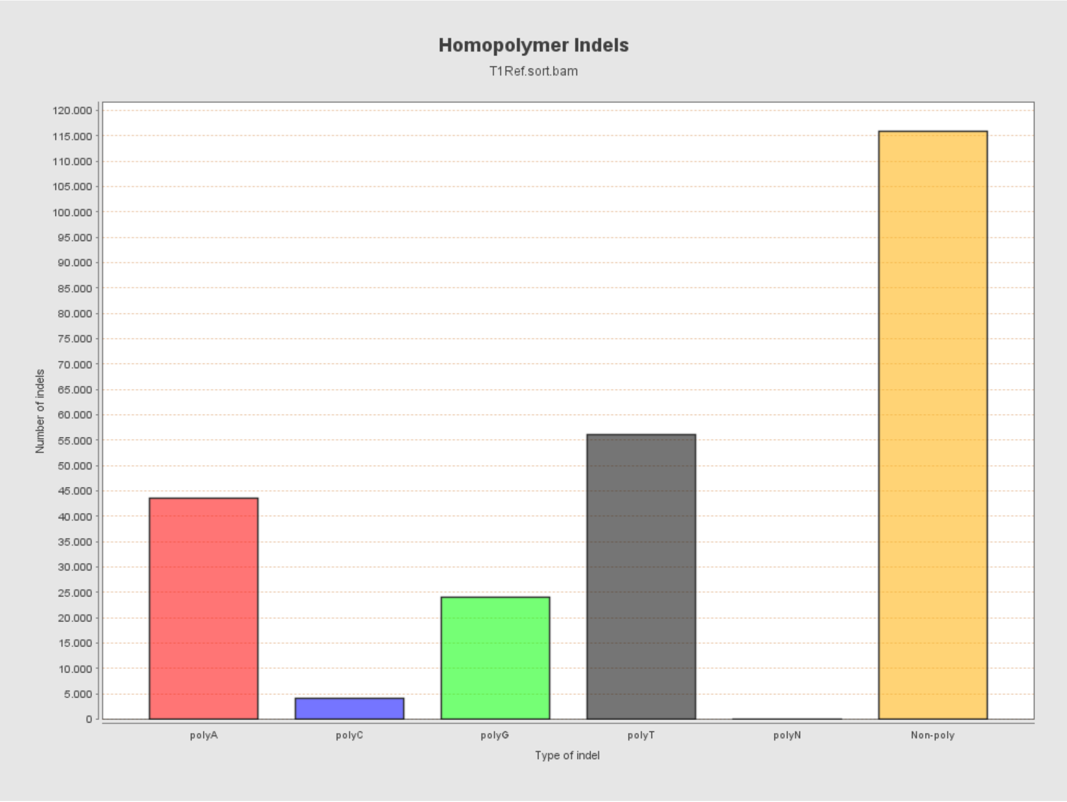
# 8. Results : Mapped Reads Nucleotide Content



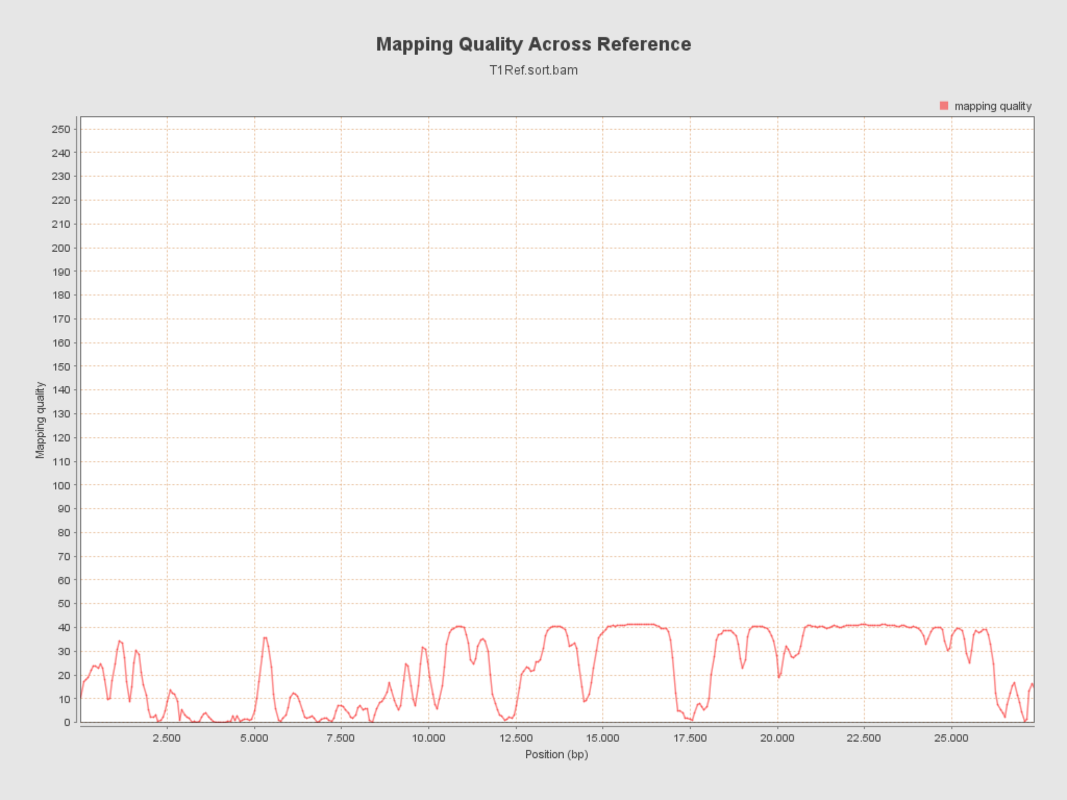
# 9. Results : Mapped Reads GC-content Distribution



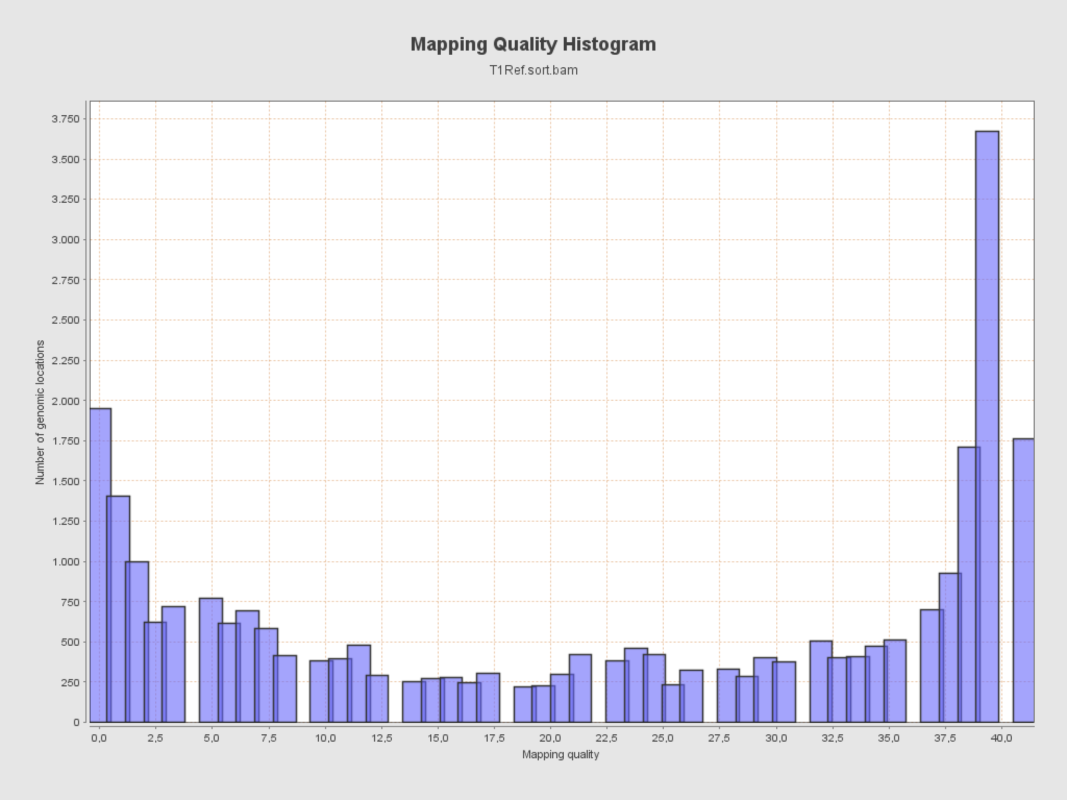
# 10. Results : Homopolymer Indels



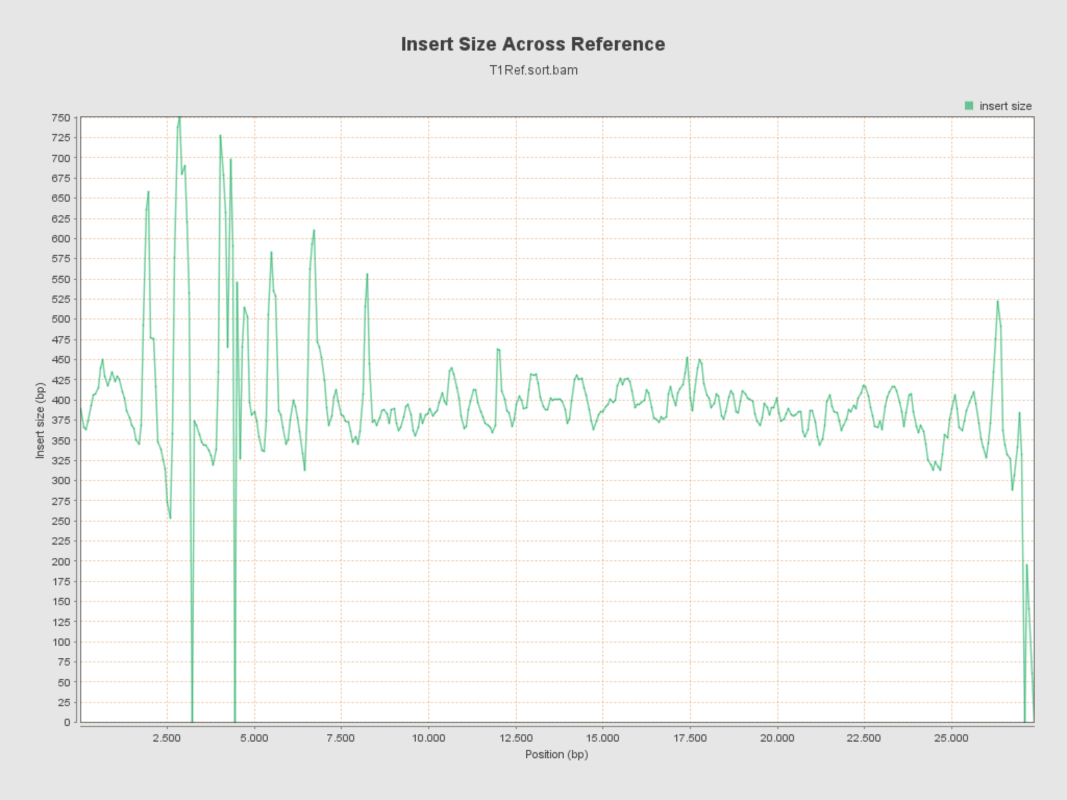
# 11. Results : Mapping Quality Across Reference



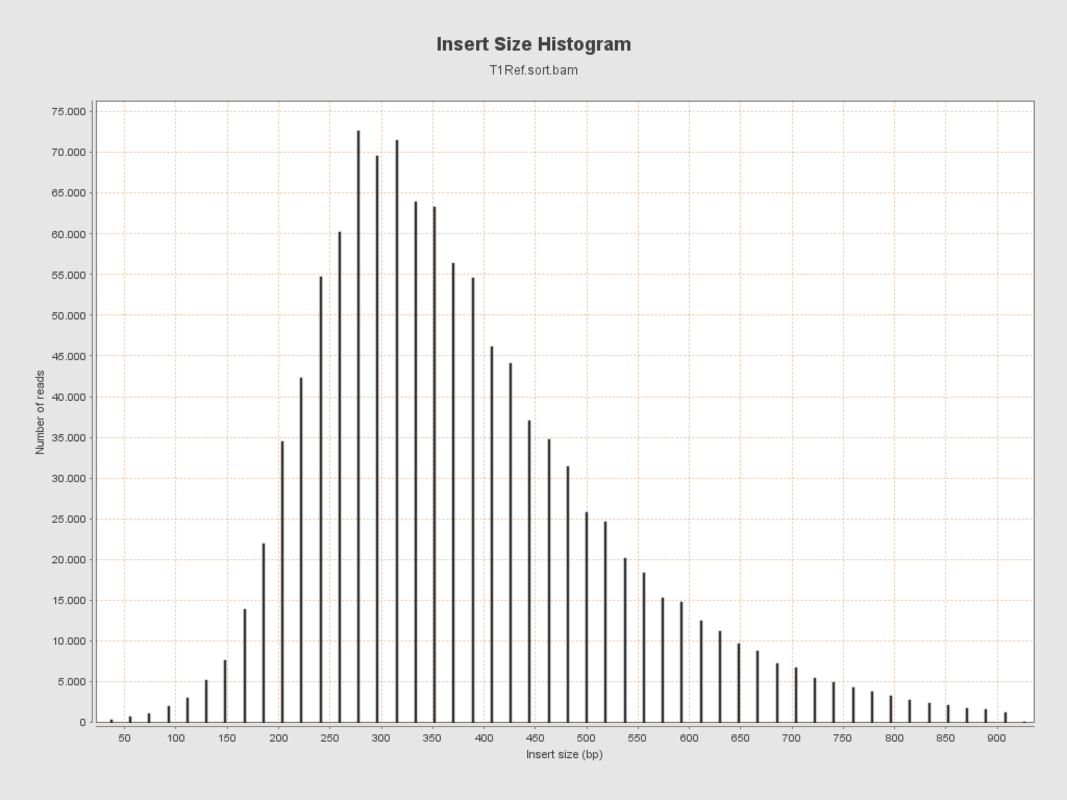
# 12. Results : Mapping Quality Histogram



# 13. Results : Insert Size Across Reference



# 14. Results : Insert Size Histogram



## Reference

Okonechnikov, K., Conesa, A., García-Alcalde, F., 2016 Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics 32(2), 292–294.