

Quality Filtering:

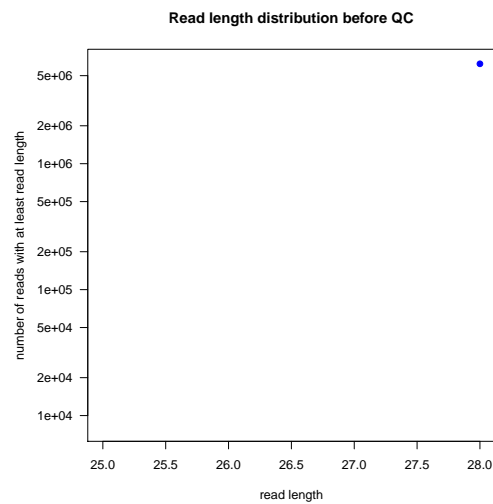


Figure 1: Read length reverse cumulative distribution before filtering

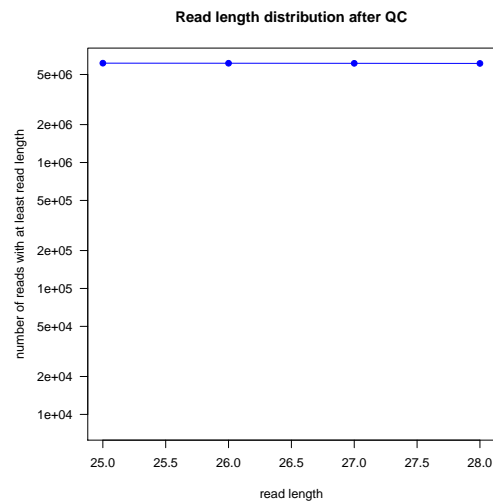


Figure 2: Read length reverse cumulative distribution after filtering

Adaptor Finding:

0.1

Chosen adaptor: GATCGGAAGAGCTCGTATGCGTCTCTCTGCTTG

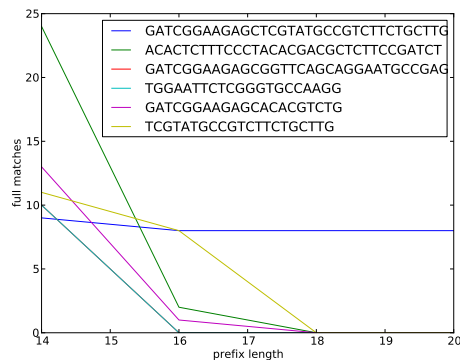


Figure 3: Plot for Adaptor Finding: Length of Adaptor versus Full Matches

Adaptor Removal and Quality Filtering:

0.2

```
# 5'/3' adaptors      : / GATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG
# 5'/3' min aln. len. : 0 / 0
# min len (1 mm)      : 7
# min len (2 mm)      : 10
# max mm (full)       : 2
# sequence filter     : on
# Mon Feb  2 02:27:31 2015: READING SEQUENCES
# Mon Feb  2 02:28:45 2015: done reading sequences (n=5765821)
# Mon Feb  2 02:28:45 2015: REMOVING ADAPTORS
# Mon Feb  2 02:34:08 2015: done removing adaptors (n=5765656)
# Mon Feb  2 02:34:08 2015: FILTERING AND OUTPUT
# Mon Feb  2 02:35:51 2015: done FILTERING AND OUTPUT
# SUMMARY:
# Number of sequences:
#
#           input sequences : 6149250 (100.0%)
#           unique sequences : 5765821 ( 93.8%)
#
# Adaptor removal:
#           full matches to 5'-adaptor : 0 ( 0.0%)
#   inexact suffix matches to 5'-adaptor : 0 ( 0.0%)
#           exact suffix matches to 5'-adaptor : 0 ( 0.0%)
#           full matches to 3'-adaptor : 0 ( 0.0%)
#   inexact prefix matches to 3'-adaptor : 6887 ( 0.1%)
#           exact prefix matches to 3'-adaptor : 1782069 ( 29.0%)
#
#           total matches to 5'-adaptor : 0 ( 0.0%)
#           total matches to 3'-adaptor : 1788956 ( 29.1%)
#
#           unique sequences (w/o adaptors) : 5765656 ( 93.8%)
#
# Sequence filtering:
#           5'-adaptor alignment too short : 0 ( 0.0%)
```

```

#           3'-adaptor alignment too short :           0 (  0.0%)
#                                           too short :           183 (  0.0%)
#           too many non-base characters :           0 (  0.0%)
#                                           low entropy :          31442 (  0.5%)
#
# Final sequences:
#                                           passed :   6117625 ( 99.5%)

```

Top 10 tags before adaptor removal:

```

>seq0    2210
AAAAAAAAAAAAAAAAAAAAAAAAAAAA
>seq1    345
AATGGAATGGAATGGAATGGAATGGAAT
>seq2    343
GAATGGAATGGAATGGAATGGAATGGAA
>seq3    320
TGGAATGGAATGGAATGGAATGGAATGG
>seq4    267
ACACACACACACACACACACACACACAC
>seq5    237
CACACACACACACACACACACACACACA
>seq6    229
GGAATGGAATGGAATGGAATGGAATGGA
>seq7    192
CTTCCTTGTTGTTGTTGTTATTCAACTCA
>seq8    178
CTGCTCTATGAAAAGAAAGGTTAAACTC
>seq9    173
AAAATGTGGCACATATACACCATGGAAT

```

Mapping:

Number of mapped reads

```

Total reads mapped:           5587889
Total unmapped reads:        203456
Total overmapped reads (>100 hits):   326280

```

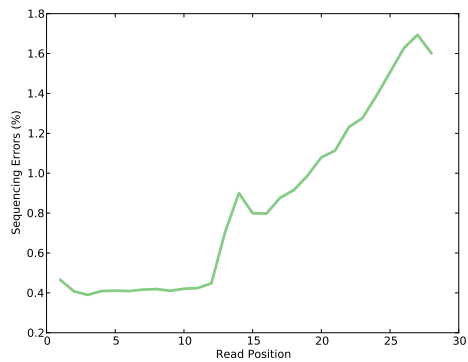


Figure 4: Fraction of all reads that had an error at read position

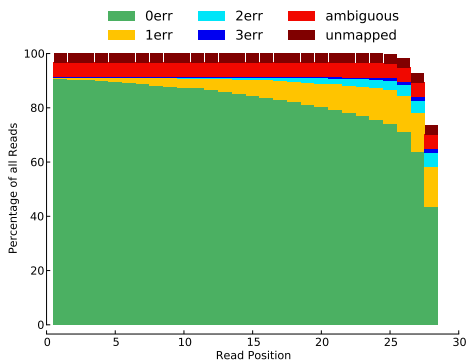


Figure 5: Fraction of all reads that have a certain mapping error up to read position

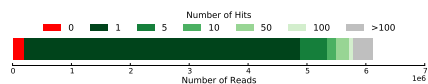


Figure 6: Number of reads that have a certain number of hits

Fragment Size Estimation:

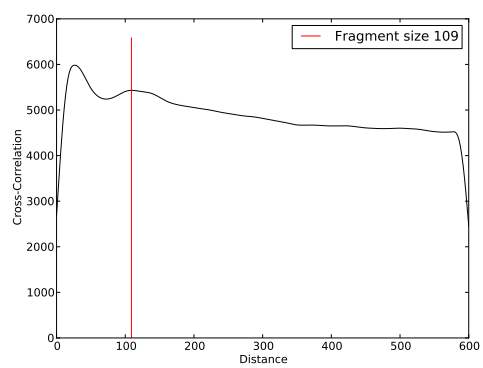


Figure 7: Fragment size correlation plot