

Quality Filtering:

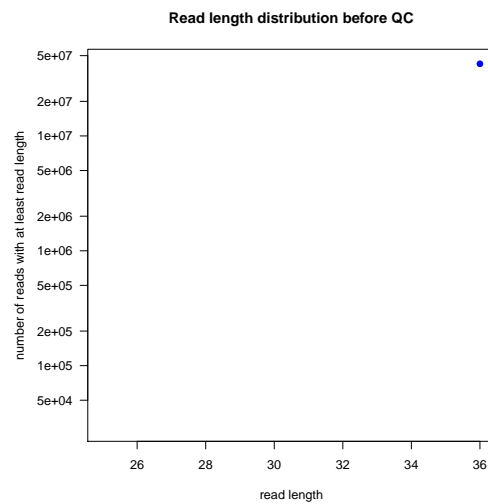


Figure 1: Read length reverse cumulative distribution before filtering

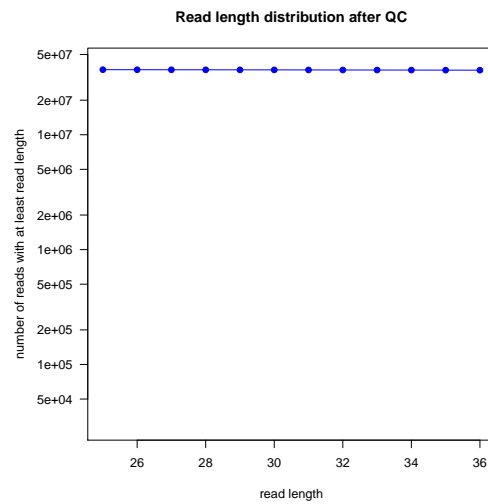


Figure 2: Read length reverse cumulative distribution after filtering

Adaptor Finding:

0.1

Chosen adaptor: ACACTCTTTCCCTACACGACGCTCTTCCGATCT

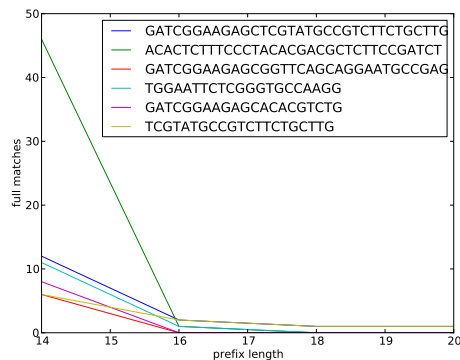


Figure 3: Plot for Adaptor Finding: Length of Adaptor versus Full Matches

Adaptor Removal and Quality Filtering:

0.2

```
# 5'/3' adaptors      : / AACTCTTTCCCTACACGACGCTCTTCCGATCT
# 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 02:48:16 2015: READING SEQUENCES
# Mon Feb  2 02:55:48 2015: done reading sequences (n=33761292)
# Mon Feb  2 02:55:48 2015: REMOVING ADAPTORS
# Mon Feb  2 03:35:02 2015: done removing adaptors (n=33759255)
# Mon Feb  2 03:35:02 2015: FILTERING AND OUTPUT
# Mon Feb  2 03:46:59 2015: done FILTERING AND OUTPUT
# SUMMARY:
# Number of sequences:
#
#           input sequences : 36861300 (100.0%)
#           unique sequences : 33761292 ( 91.6%)
#
# 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 : 124279 ( 0.3%)
#     exact prefix matches to 3'-adaptor : 12477972 ( 33.9%)
#
#           total matches to 5'-adaptor : 0 ( 0.0%)
#           total matches to 3'-adaptor : 12602251 ( 34.2%)
#
#           unique sequences (w/o adaptors) : 33759255 ( 91.6%)
#
# Sequence filtering:
#           5'-adaptor alignment too short : 0 ( 0.0%)
```

```

#           3'-adaptor alignment too short :           0 (  0.0%)
#                                           too short :           3 (  0.0%)
#           too many non-base characters :           0 (  0.0%)
#                                           low entropy :      88213 (  0.2%)
#
# Final sequences :
#                                           passed :  36773084 ( 99.8%)

```

```

Top 10 tags before adaptor removal:
>seq0    678
CTAACCCTAACCCTAACCCTAACCCTAACCCTAACC
>seq1    613
ATTCCATTCCATTCCATTCCATTCCATTCCATTCCA
>seq2    587
TTCCATTCCATTCCATTCCATTCCATTCCATTCCAT
>seq3    586
TGGAATGGAATGGAATGGAATGGAATGGAATGGAAT
>seq4    580
GAATGGAATGGAATGGAATGGAATGGAATGGAATGG
>seq5    557
TCCATTCCATTCCATTCCATTCCATTCCATTCCATT
>seq6    556
GGAATGGAATGGAATGGAATGGAATGGAATGGAATG
>seq7    548
AATGGAATGGAATGGAATGGAATGGAATGGAATGGA
>seq8    547
ACACACACACACACACACACACACACACACACACAC
>seq9    537
GTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGG

```

Mapping:

Number of mapped reads

```

Total reads mapped:      34788972
Total unmapped reads:    664310
Total overmapped reads (>100 hits):  1319801

```

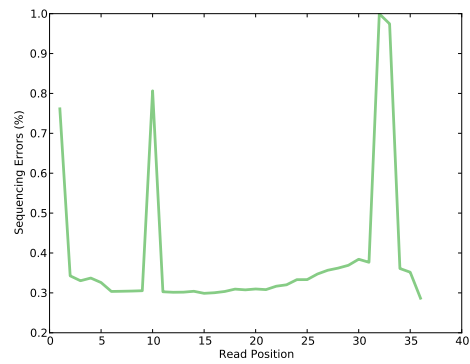


Figure 4: Fraction of all reads that had an error at read position

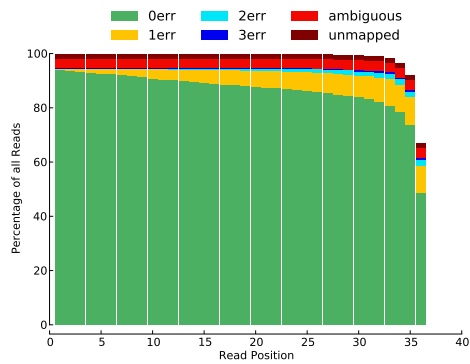


Figure 5: Fraction of all reads that have a certain mapping error up to read position

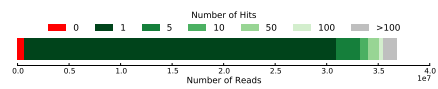


Figure 6: Number of reads that have a certain number of hits

Fragment Size Estimation:

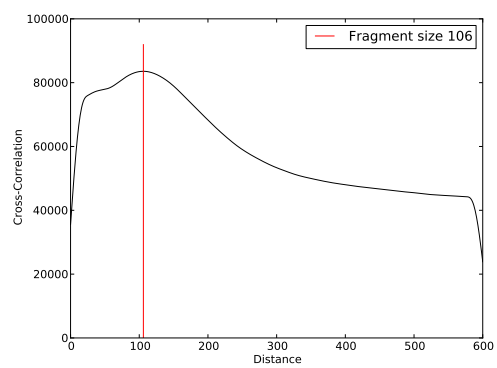


Figure 7: Fragment size correlation plot