

Quality Filtering:

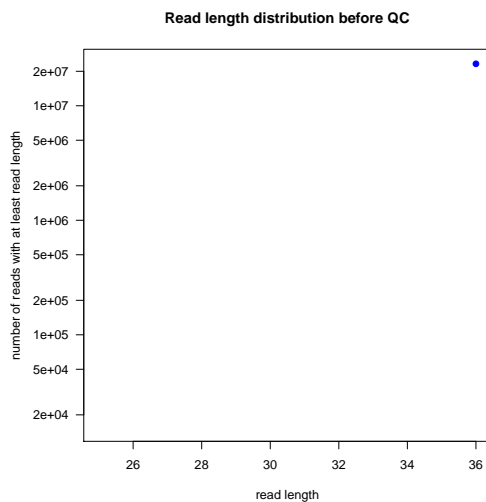


Figure 1: Read length reverse cumulative distribution before filtering

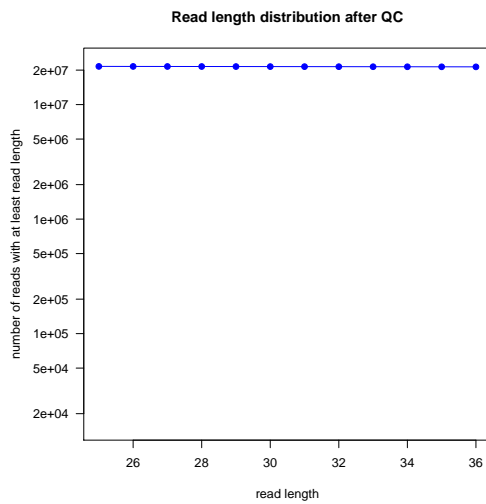


Figure 2: Read length reverse cumulative distribution after filtering

Adaptor Finding:

0.1

Chosen adaptor: GATCGGAAGAGCTCGTATGCCGTCCTCTGCTTG

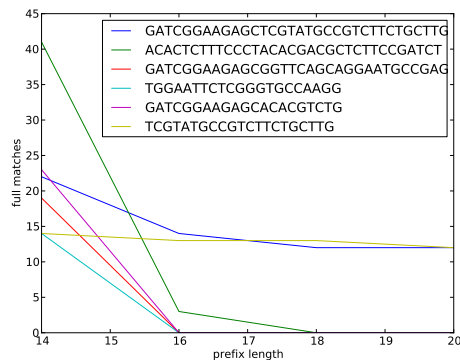


Figure 3: Plot for Adaptor Finding: Length of Adaptor versus Full Matches

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
# Sun Feb 1 20:20:36 2015: READING SEQUENCES
# Sun Feb 1 20:24:59 2015: done reading sequences (n=19978415)
# Sun Feb 1 20:24:59 2015: REMOVING ADAPTORS
# Sun Feb 1 20:47:21 2015: done removing adaptors (n=19977798)
# Sun Feb 1 20:47:21 2015: FILTERING AND OUTPUT
# Sun Feb 1 20:54:32 2015: done FILTERING AND OUTPUT
# SUMMARY:
# Number of sequences:
#                   input sequences : 21555006 (100.0%)
#                   unique sequences : 19978415 ( 92.7%)
#
# 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 :        1071 ( 0.0%)
#   inexact prefix matches to 3'-adaptor :       22933 ( 0.1%)
#       exact prefix matches to 3'-adaptor :    6328018 ( 29.4%)
#
#       total matches to 5'-adaptor :          0 ( 0.0%)
#       total matches to 3'-adaptor :    6352022 ( 29.5%)
#
#       unique sequences (w/o adaptors) : 19977798 ( 92.7%)
#
# Sequence filtering:
#       5'-adaptor alignment too short :          0 ( 0.0%)

```

```

#           3'-adaptor alignment too short :           0 (  0.0%)
#                                           too short :       1098 (  0.0%)
#           too many non-base characters :           0 (  0.0%)
#                                           low entropy :      47945 (  0.2%)
#
# Final sequences :
#                                           passed :  21505963 ( 99.8%)

```

```

Top 10 tags before adaptor removal:
>seq0   935
GATCGGAAGAGCTCGTATGCGCTCTCTGCTTGAAA
>seq1   559
CTAACCCCTAACCCCTAACCCCTAACCCCTAACCC
>seq2   492
CTGCTCTATGAAAAGAAAGGTTAAACTCTGTGAGTT
>seq3   455
ACACACACACACACACACACACACACACACACAC
>seq4   418
CCTAACCCCTAACCCCTAACCCCTAACCCCTAAC
>seq5   406
CACACACACACACACACACACACACACACACACA
>seq6   365
GGFTAGGGFTAGGGFTAGGGFTAGGGFTAGGGFTAG
>seq7   357
CATTCATTCATTCATTCATTCATTCATTCATTC
>seq8   345
TCTGCTCTGTGTAAAGGATCGTTCAACTCTGTGAGT
>seq9   343
TGGAATGGAATGGAATGGAATGGAATGGAATGGAAT

```

Mapping:

Number of mapped reads

```

Total reads mapped:           20127066
Total unmapped reads:        619702
Total overmapped reads (>100 hits):  759195

```

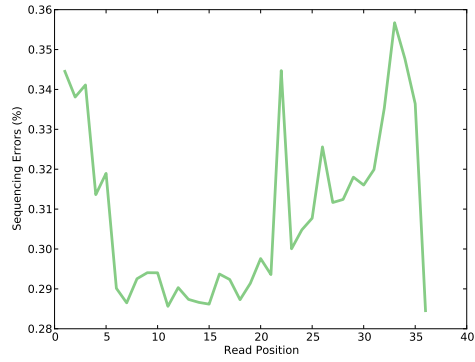


Figure 4: Fraction of all reads that had an error at read position

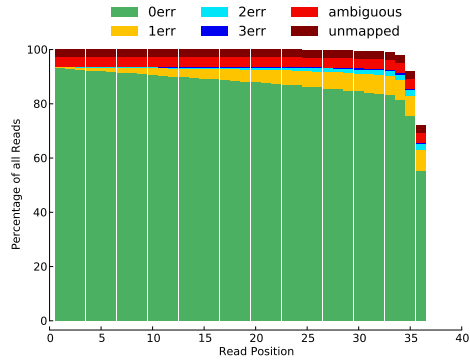


Figure 5: Fraction of all reads that have a certain mapping error up to read position

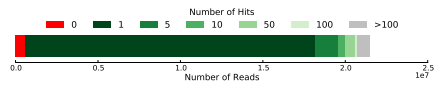


Figure 6: Number of reads that have a certain number of hits

Fragment Size Estimation:

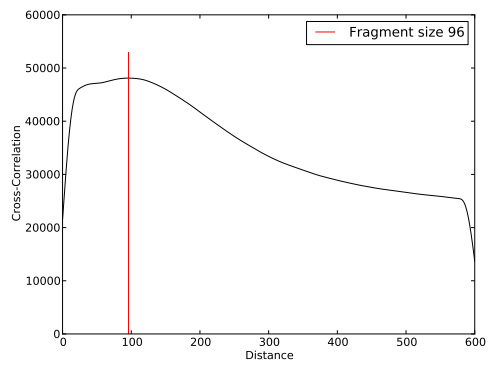


Figure 7: Fragment size correlation plot