Combined Summary Report

March 29, 2016

1 Pre-Processing Data

This section contains the input read data before processing.

1.1 Base Quality Graph

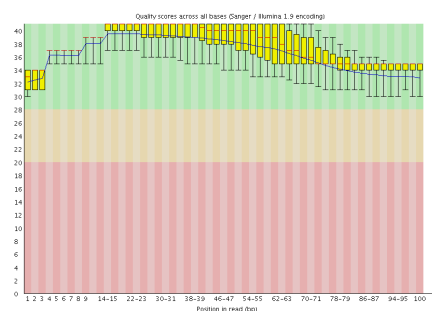*This graph shows an overview of the range of quality values across all bases at each position in the FastQ file.*
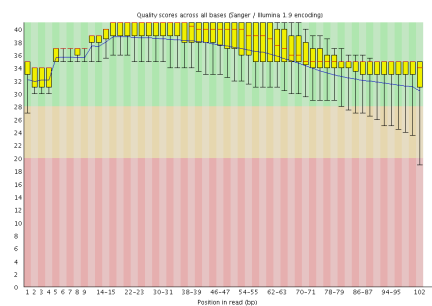

Figure 1: DRR016127_1

Figure 2: DRR016127_2

1.2 Data Summary

- DRR016127_1 Basic Statistics:

  Total Sequences 12516488
  Sequences flagged as poor quality 0
  Sequence length 100
  %GC 46

Other Statistics Summary:

Metrics ResultBasic Statistics PASS
Per base sequence quality PASS
Per tile sequence quality PASS
Per sequence quality scores PASS
Per base sequence content FAIL
This could've failed due to fragmentation bias in the first 13bp. Checking...
Only fragment bias in the first 13bp found.Per sequence GC content WARN
Per base N content PASS
Sequence Length Distribution PASS
Sequence Duplication Levels FAIL
Overrepresented sequences WARN
Adapter Content PASS
Kmer Content FAIL

- DRR016127_2 Basic Statistics:

  Total Sequences 12516488
  Sequences flagged as poor quality 0
  Sequence length 102
  %GC 47

  Other Statistics Summary:

  Metrics ResultBasic Statistics PASS
  Per base sequence quality PASS
  Per tile sequence quality PASS
  Per sequence quality scores PASS
  Per base sequence content FAIL
  This could've failed due to fragmentation bias in the first 13bp. Checking...
  Only fragment bias in the first 13bp found.Per sequence GC content PASS
  Per base N content PASS
  Sequence Length Distribution PASS
  Sequence Duplication Levels FAIL
  Overrepresented sequences PASS
  Adapter Content PASS
  Kmer Content FAIL

2 Post-Processing Data

2.1 Base Quality Graph

This graph shows an overview of the range of quality values across all bases at each position in the FastQ file.

2.2 Data Summary

- DRR016127_1 Basic Statistics:

  Total Sequences 12516488
  Sequences flagged as poor quality 0
  Sequence length 100
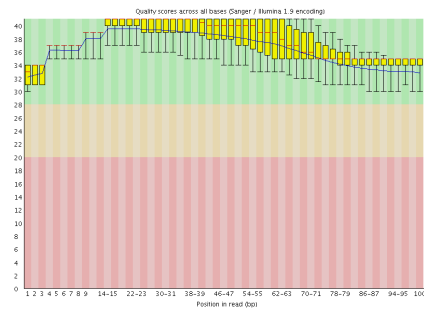
Figure 3: DRR016127_1
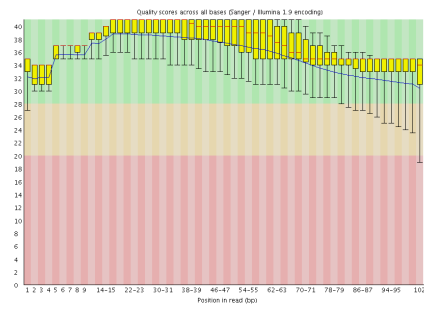
Figure 4: DRR016127_2

%GC 46

Other Statistics Summary:

Metrics ResultBasic Statistics PASS
Per base sequence quality PASS
Per tile sequence quality PASS
Per sequence quality scores PASS
Per base sequence content FAIL
This could've failed due to fragmentation bias in the first 13bp. Checking...
Only fragment bias in the first 13bp found.Per sequence GC content WARN
Per base N content PASS
Sequence Length Distribution PASS
Sequence Duplication Levels FAIL
Overrepresented sequences WARN
Adapter Content PASS
Kmer Content FAIL

- DRR016127_2 Basic Statistics:

Total Sequences 12516488
Sequences flagged as poor quality 0
Sequence length 102
%GC 47

Other Statistics Summary:

Metrics ResultBasic Statistics PASS

Per base sequence quality PASS
Per tile sequence quality PASS
Per sequence quality scores PASS
Per base sequence content FAIL
This could've failed due to fragmentation bias in the first 13bp. Checking...
Only fragment bias in the first 13bp found.Per sequence GC content PASS
Per base N content PASS
Sequence Length Distribution PASS
Sequence Duplication Levels FAIL
Overrepresented sequences PASS
Adapter Content PASS
Kmer Content FAIL

3 Post Mapping Data

3.1 Data Summary

Post Mapping Metrics

25075522 + 0 in total (QC-passed reads + QC-failed reads)
0 + 0 secondary
42546 + 0 supplementary
0 + 0 duplicates
21270168 + 0 mapped (84.82% : N/A)
25032976 + 0 paired in sequencing
12516488 + 0 read1
12516488 + 0 read2
20945984 + 0 properly paired (83.67% : N/A)
21077176 + 0 with itself and mate mapped
150446 + 0 singletons (0.60% : N/A)
83326 + 0 with mate mapped to a different chr
45380 + 0 with mate mapped to a different chr (mapQ>=5)

3.2 Visual Summary

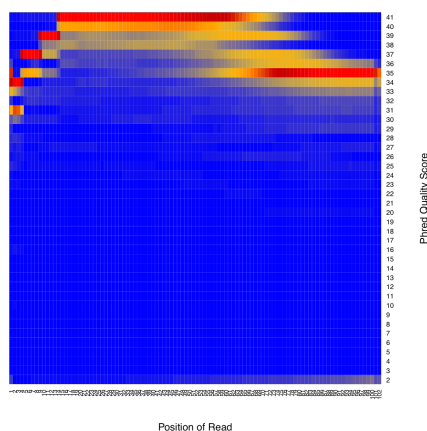*This graph shows an overview of the range of quality values across all bases at each position in the alignment file.*

Figure 5: Alignment Quality

4

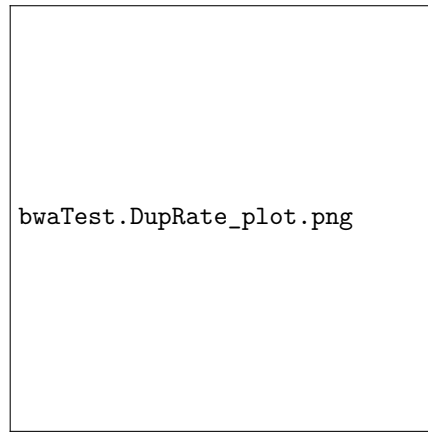*This graph shows read duplication rate.*

Figure 6: Duplication Rate

- *Sequence based: reads with identical sequence are regarded as duplicated reads.*

- *Mapping based: reads mapped to the exactly same genomic location are regarded as duplicated reads.*