

# Quality Assessment: FastQC

Ben Tully

## Abstract

Quality impact results is the universal first step for all sequencing methods. Sequencing results distrputed in FASTQ format.

**Citation:** Ben Tully Quality Assessment: FastQC. **protocols.io**

dx.doi.org/10.17504/protocols.io.fa3bign

**Published:** 25 Jul 2016

## Protocol

### Step 1.

Move to quality control directory:

```
cmd COMMAND  
$ cd /home/c-debi/ecogeo/qualitycontrol
```

### Step 2.

View sample FASTQ file:

```
cmd COMMAND  
$ less igm1000002065_R1.fastq
```

 **EXPECTED RESULTS**

```
@SRR492065.1453522 HWI-EAS385_0095_FC:2:33:6836:12013  
length=100  
ATTTTGTCTCAATCAATTTTGACATAGAAATGCCATTCGGACACAAAATCACCGCCTTAAT  
TATTTCTCTTCCAAATCATTGTGATCATTACTAATCAA  
+  
IIIIIIIIIIIIIIIGIIHIIIIHIIIIIIIIIGIIIGIHHIIIIIIIIIBIIFHHIG  
IIHFHIIHBIIIIIIIIGIFGDHIIIIIIHIGIEI
```

### Step 3.

There are dozens of trimming programs and quality assessment tools - many utilize similar criteria, a sliding window that analyzes quality scores and trims when the score falls below an average.

### Step 4.

FastQC - allows for visualization of quality scores:

```
cmd COMMAND
```

```
$ fastqc
```

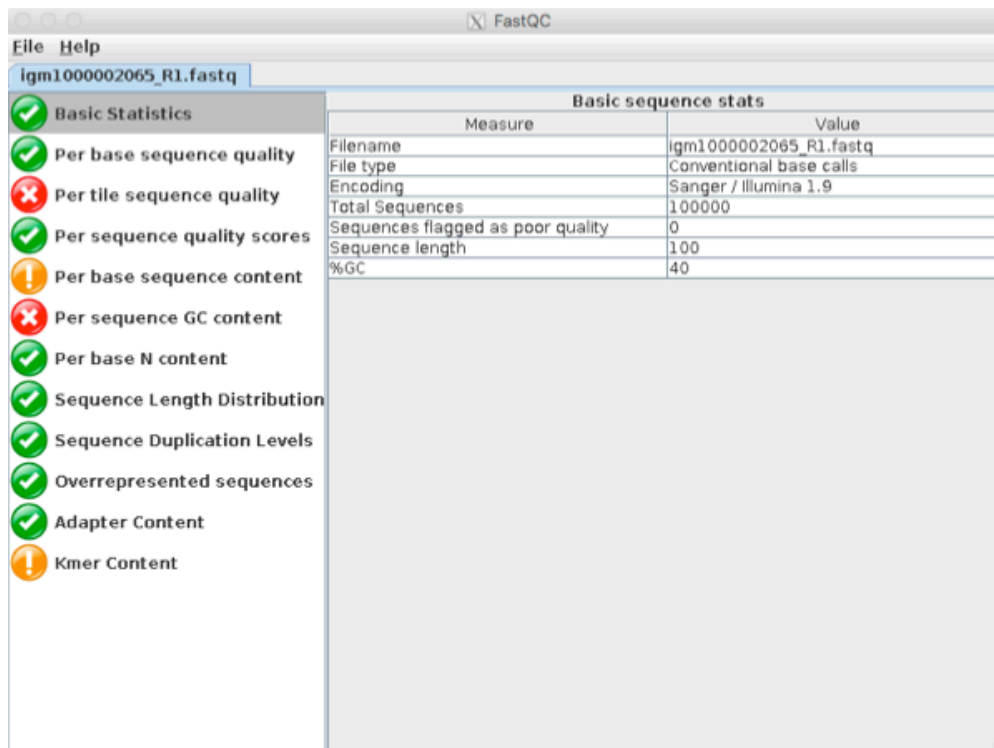
### Step 5.

File > Open ... img1000002065\_R1.fastq



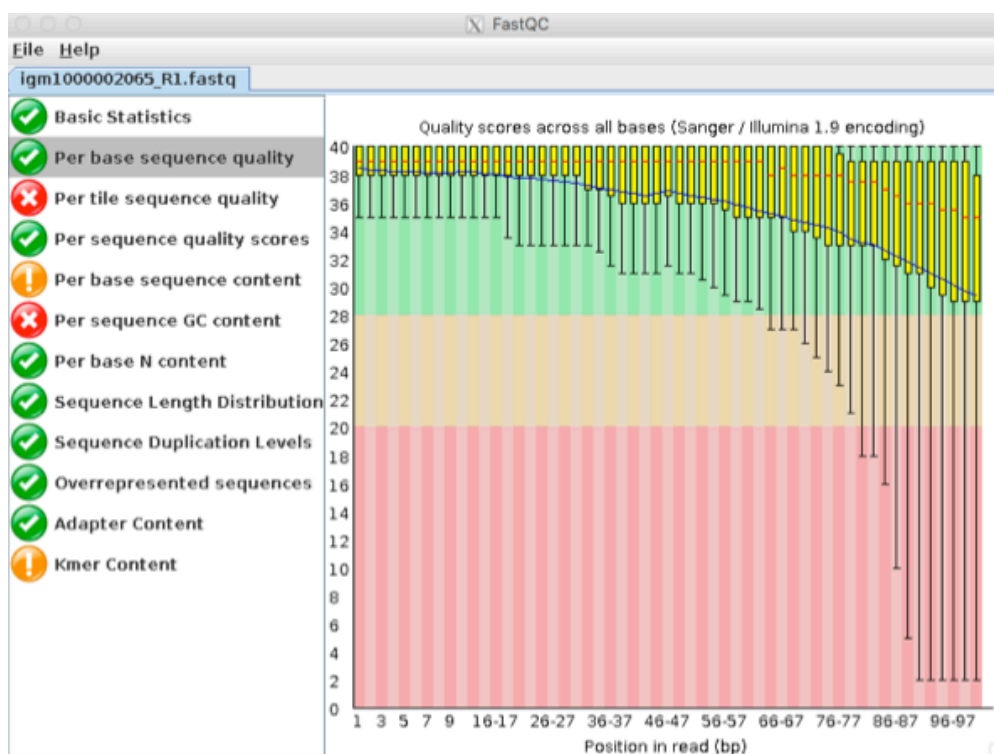
### Step 6.

img1000002065\_R1.fastq



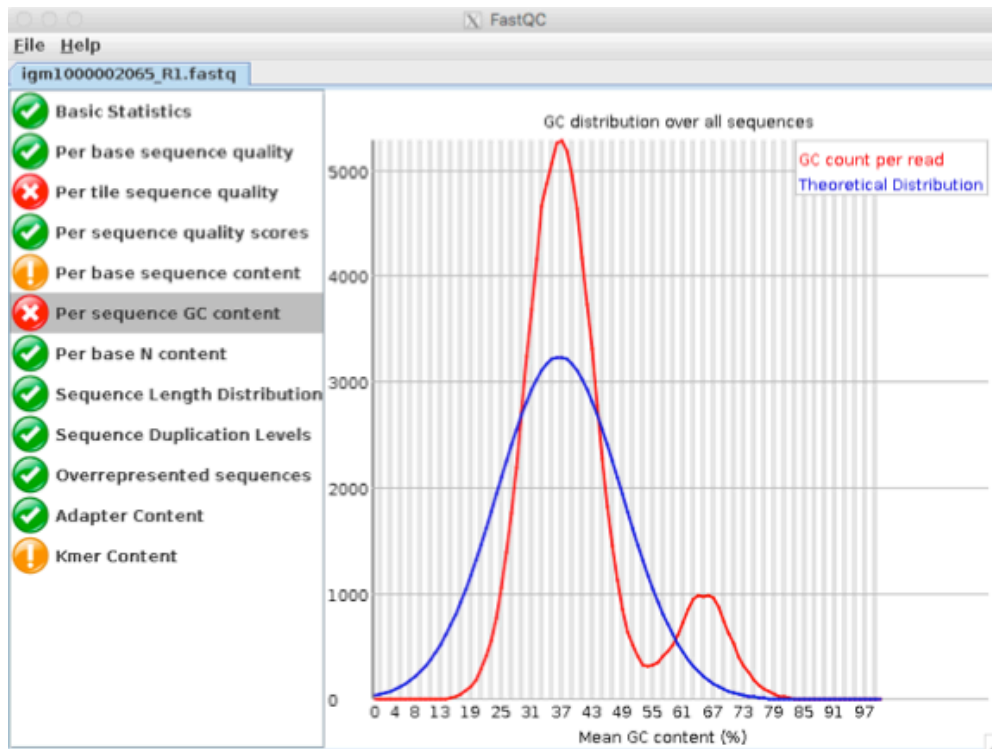
## Step 7.

img1000002065\_R1.fastq Quality Scores



## Step 8.

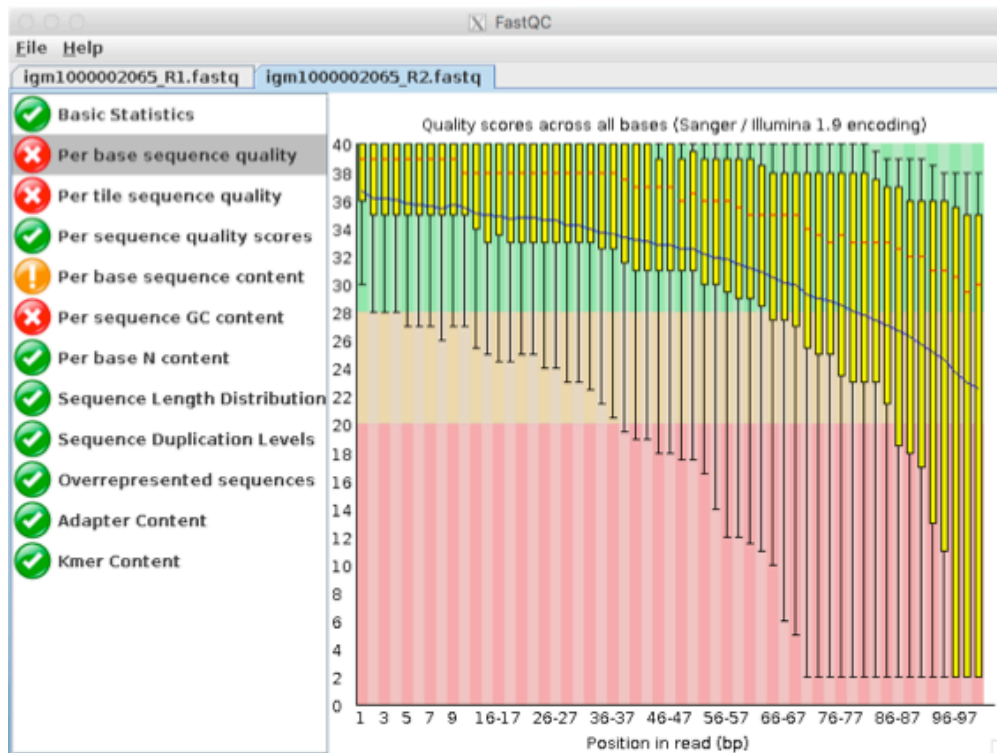
img1000002065\_R1.fastq GC Content



## Step 9.

File > Open ... img1000002065\_R1.fastq

Check the Quality Score Assessment



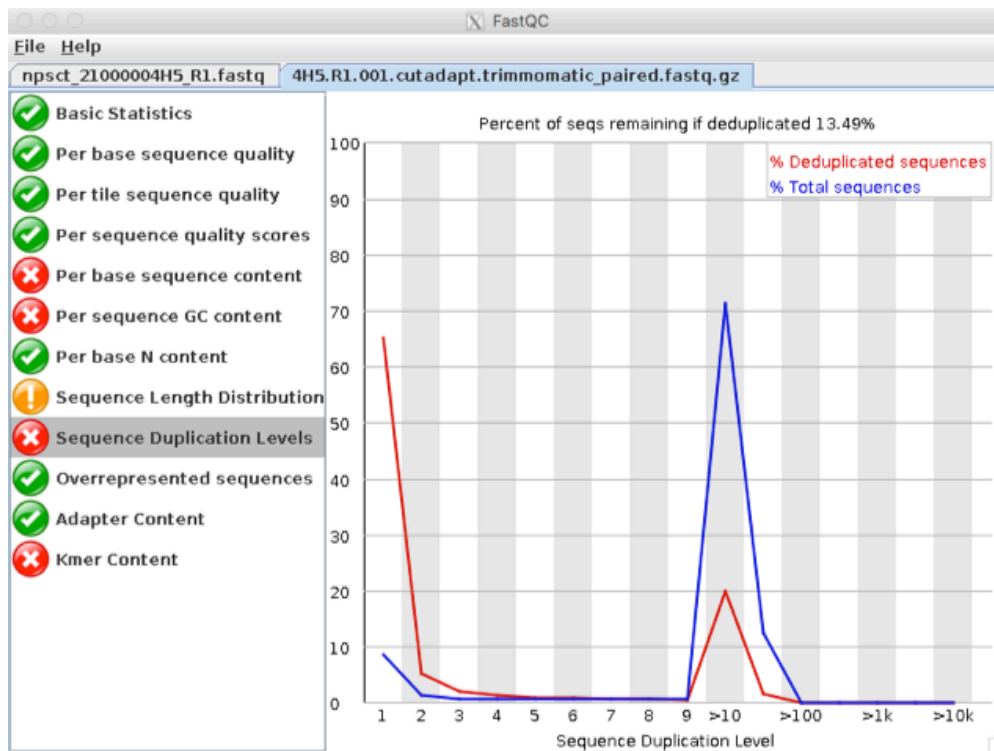
## ■ ANNOTATIONS

**Elisha Wood-Charlson** 03 Aug 2016

File > Open ... img1000002065\_R2.fastq

### Step 10.

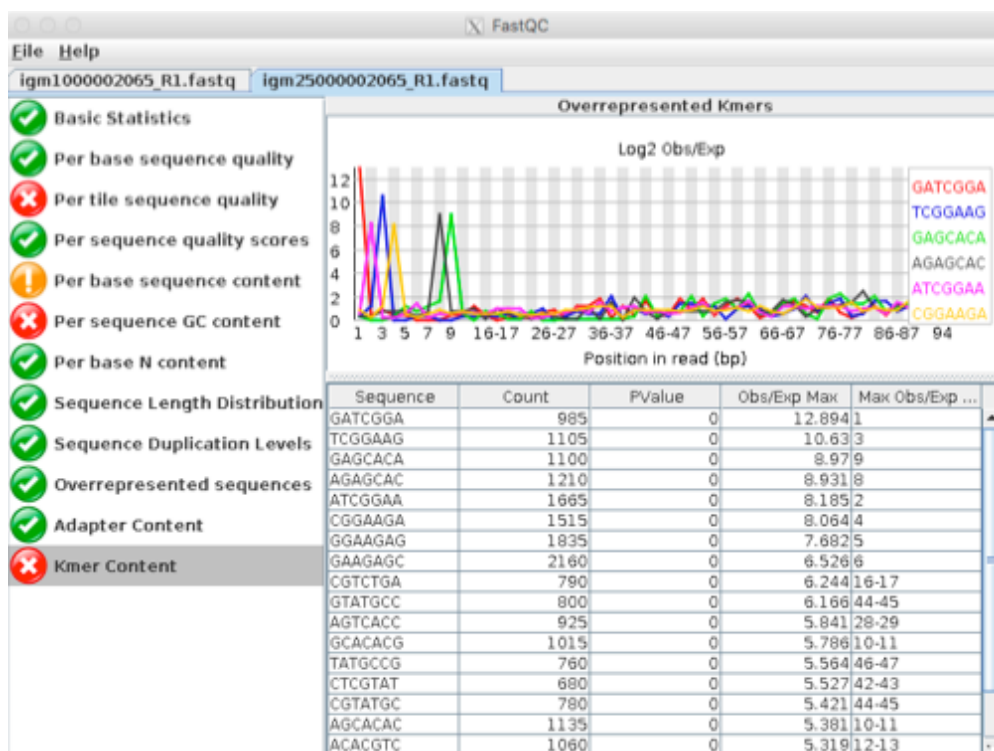
Example of Failed Sequence Duplication Levels



## Step 11.

File > Open ... img1000002065\_R1.fastq

Check the Kmer Content



📌 NOTES

**Elisha Wood-Charlson** 13 Jul 2016

Libraries which derive from random priming will nearly always show Kmer bias at the start of the library due to an incomplete sampling of the possible random primers.