

Fastq quality data.

Your Name here

May 19, 2021

Contents

1	Project characteristics	1
2	Global summaries	1
2.1	Project names and read numbers	2
3	Nucleotide patterns	2
3.1	N nucleotides	3
3.2	GC content	4
3.3	Nucleotide frequencies	5
4	Phred qualities	7
5	Hierarchical clustering	8

1 Project characteristics

Project characteristics

Contact

Phone

Institute

Mail

Start date

2 Global summaries

Input data: Summarized data on FASTQ files.

```
[fastqq] File ( 1/2) '/tmp/RtmpujJ4qq/Rinst2dd1ed32c1fcd9/seqTools/extdata/g4_1101_r
```

```
[fastqq] File ( 2/2) '/tmp/RtmpujJ4qq/Rinst2dd1ed32c1fcd9/seqTools/extdata/g5_1101_r
```

Printout of Fastqq object:

```
> fqq

Class      :      Fastqq
nFiles     :          2
maxSeqLen  :        101
k (Kmer len):         4

nReads     :        200
nr  N   nuc :          2
Min seq len :        101
Max seq len :        101
```

2.1 Project names and read numbers

```
> dfr<-data.frame(file=basename(fileName(fqq)),
+                 sample=probeLabel(fqq),
+                 reads=format(nReads(fqq), big.mark=Sys.localeconv()[7]))
> print(dfr)
```

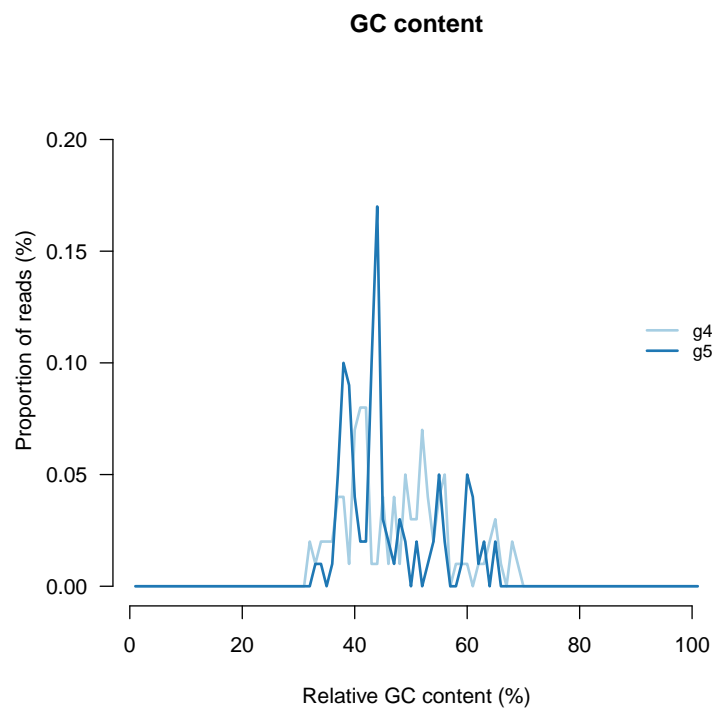
```
      file sample reads
1 g4_l101_n100.fq.gz    g4   100
2 g5_l101_n100.fq.gz    g5   100
```

3 Nucleotide patterns

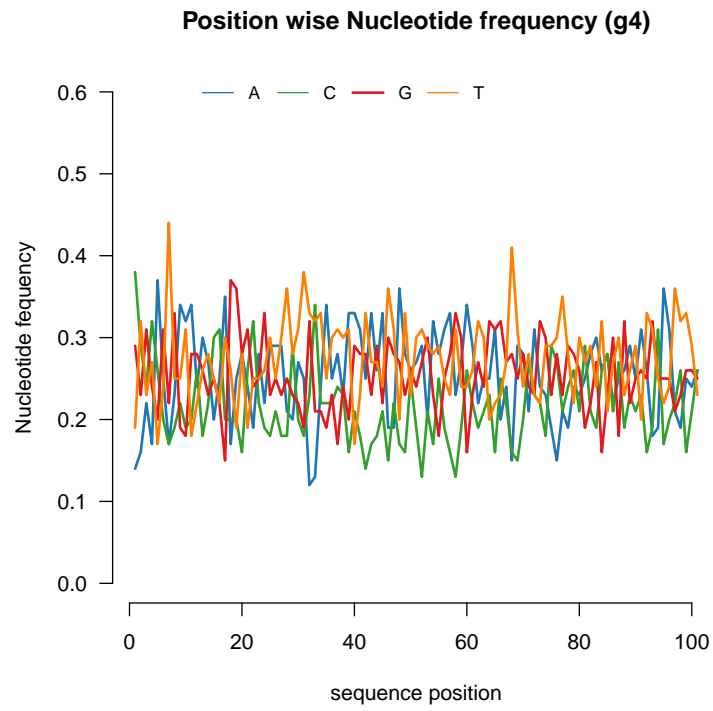
3.1 N nucleotides



3.2 GC content



3.3 Nucleotide frequencies





4 Phred qualities





5 Hierarchical clustering

1_g4	1
2_g5	2

