

Отчёт

Посмотрим статистику:

Basic Statistics

Measure	Value
Filename	SRR3824222_1.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	163183367
Sequences flagged as poor quality	0
Sequence length	19-140
%GC	22

Basic Statistics

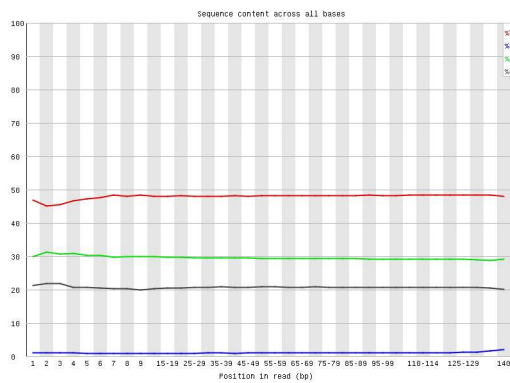
Measure	Value
Filename	SRR3414630_1.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	15244711
Sequences flagged as poor quality	0
Sequence length	65
%GC	47

BS - SEQ

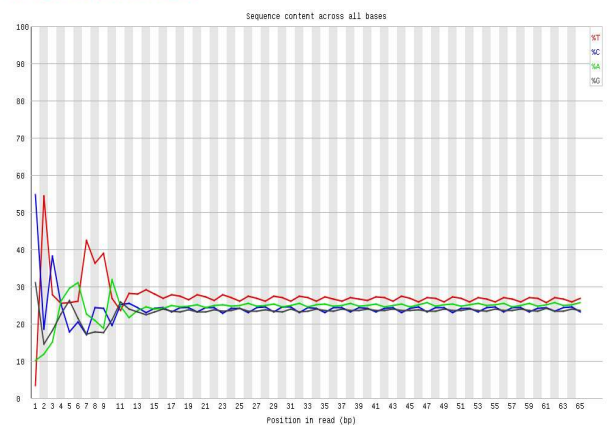
PHK

Заметим, что в PHK процент GC более чем в два раза выше, чем у BS-Seq.

Per base sequence content



Per base sequence content

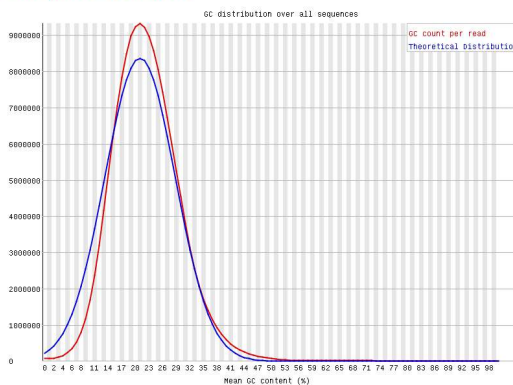


BS - SEQ

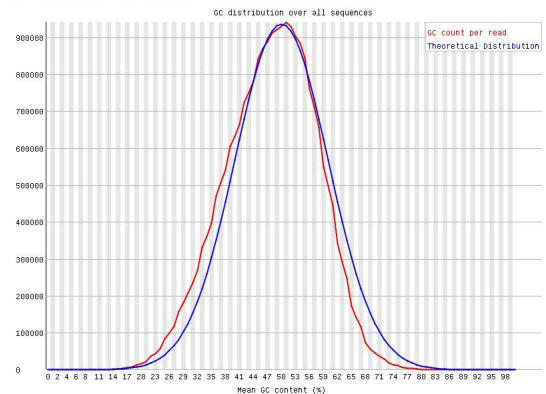
PHK

Заметим, что в первом случае почти отсутствуют Цитозины (C), содержание Гуанина (G), ниже чем в PHK, а Тимина (T) – выше.

Per sequence GC content



Per sequence GC content



BS - SEQ

PHK

Заметим явное смещение нормального распределения для GC count per read.