

# fastx-utils之统计序列文件的碱基组成以及质量值：fqchk

## 一、fastx-utils fqchk介绍

功能描述：

`fastx-utils fqchk` 统计序列文件的碱基组成以及质量值。

命令行接口：

```
1 $ fastx-utils fqchk
2
3 Usage: fastx-utils fqchk [options] <in.fq>
4 Options:
5   -q INT      offset value: 33 for Sanger quality, 64 for Illumina quality,
                default: [33]
```

可选参数：

```
1 -q 整数 33为Sanger质量，64为Illumina质量，默认为33；
```

## 二、使用场景实例及其用法

使用场景经典案例：

1. 统计fastq文件序列碱基和质量组成，绘制质量控制图。

示例演示：

示例文件： `sequence.fastq`

```
1 $ cat sequence.fastq | head -n8
```

[illegible]

运行命令：显示 `sequence.fastq` 的碱基组成和质量值。

```
1 | $ fastx-utils fqchk sequence.fastq | head -n 6
```

1	#sequence: 31; base: 9228; min_len: 32; max_len: 391; avg_len: 297.68; Q20: 0.8237; Q30: 0.3668;										
2	#POS	A	T	C	G	N	Q1	Q2	Q3	Q4	
	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13		
	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	
	Q24	Q25	Q26	Q27	Q28	Q29	Q30	Q31	Q32	Q33	
	Q34	Q35	Q36	Q37	Q38	Q39	Q40	Q41	Q42		
3	1	0	25	1	5	0	0	0	0	0	
	0	0	1	0	0	0	1	1	2	1	
	0	0	0	0	2	1	02	0	0	4	
	3	2	1	0	0	4	3	2	1	0	
	0	0	0	0	0	0	0	0			
4	2	25	4	0	2	0	0	0	0	0	
	0	0	1	0	0	0	0	1	2	1	
	0	0	0	1	0	2	01	0	0	3	
	3	0	2	2	1	4	1	3	3	0	
	0	0	0	0	0	0	0				
5	3	28	1	0	2	0	0	0	0	0	
	0	0	0	2	0	0	0	2	1	1	
	0	0	2	3	3	0	21	1	1	1	
	2	2	2	0	2	1	1	1	0	0	
	0	0	0	0	0	0	0	0			
6	4	2	24	1	4	0	0	0	0	0	
	0	0	0	0	0	0	0	1	0	3	
	0	0	0	1	1	3	10	1	6	0	
	3	1	0	1	1	4	2	0	2	0	
	0	0	0	0	0	0	0				

注意选项1: 设置 `-q` 参数, 设定 `64` 指示老的 `illumina fastq` 文件格式.

本文材料为 **BASE (Biostack Applied bioinformatic SEries)** 课程 **Linux Command Line Tools for Life Scientists** 材料, 版权归 上海逻捷信息科技有限公司 所有。

Last Update: 2020-08-10 11:56 AM