

tsv-utils之行修饰: view

一、tsv-utils view介绍

功能描述:

`tsv-utils view` 辅助操作子命令, 可以将 `double` 型转换成 `int` 类型, 去除文件中的空行以及注释行。

命令行接口:

```
1 $ tsv-utils view
2
3 Usage: tsv-utils [options] view <tsv>
4
5 Options:
6   -c  Add comment char '#' to the first line.
7   -d  Drop lines start with '#', skip first line.
8   -b  Remove the blank lines.
9   -r  Round value to Int.
10  -l  Add line number.
```

可选参数:

```
1 -c 再第一行添加'#', 如存在,跳过;
2 -b 删除空白行和注释;
3 -d 删除除了第一行外的以'#'开头的注释行;
4 -r 将数值变为整数;
5 -l 添加行号
```

二、使用场景实例及其用法

经典使用案例

`tsv-utils view` 使用场景比较多, 主要是辅助功能, 比如: `salmon` 基因丰度定量的结果 `quant.sf`。

示例演示

示例文件: `quant.sf` 文件

```
1 $ cat quant.sf | head -n 7
```

	Name	Length	EffectiveLength	TPM	NumReads	
2	TrA0001W		9003	8740.853	0.048685	1.000
3	TrA0002C		2364	2101.853	0.202465	1.000
4	TrA0003W		1065	802.853	0.000000	0.000
5	TrA0004W		783	521.202	0.000000	0.000
6	TrA0005W		654	392.544	0.000000	0.000
7	TrA0006W		921	658.964	0.000000	0.000

运行命令:

示例 1：在第一行添加 "#"

```
1 $ tsv-utils view -c quant.sf | head -n 7
```

1	#Name	Length	EffectiveLength	TPM	NumReads	
2	TrA0001W	9003	8740.853	0.048685	1.000	
3	TrA0002C	2364	2101.853	0.202465	1.000	
4	TrA0003W	1065	802.853	0.000000	0.000	
5	TrA0004W	783	521.202	0.000000	0.000	
6	TrA0005W	654	392.544	0.000000	0.000	
7	TrA0006W	921	658.964	0.000000	0.000	

示例2：添加列号。

```
1 $ tsv-utils view -c1 quant.sf | head -n 7
```

1	1	#Name	Length	EffectiveLength	TPM	NumReads	
2	2	TrA0001W	9003	8740.853	0.048685	1.000	
3	3	TrA0002C	2364	2101.853	0.202465	1.000	
4	4	TrA0003W	1065	802.853	0.000000	0.000	
5	5	TrA0004W	783	521.202	0.000000	0.000	
6	6	TrA0005W	654	392.544	0.000000	0.000	
7	7	TrA0006W	921	658.964	0.000000	0.000	

示例3：添加列号。

参数使用：设置 `-r` 参数，将数值变为整数，并删除空白行和注释行。

```
1 $ tsv-utils view -cr quant.sf | cut -f1,5 | head -n 7
```

1	#Name	NumReads
2	TrA0001W	1
3	TrA0002C	1
4	TrA0003W	0
5	TrA0004W	0
6	TrA0005W	0
7	TrA0006W	0

示例4：合并多个注释文件,删除中间的注释行。

示例文件: `alpha.A-1.txt`, `alpha.A-2.txt` alpha多样性计算文件

```
1 $ cat alpha.A-1.txt | cut -f1,2,4
```

1	#Sample	berger_parker	chao1
2	A-1	0.0313	433.2

```
1 $ cat alpha.A-2.txt alpha.A-1.txt | tsv-utils view -d - | cut -f1,2,3
```

```
1 | #Sample berger_parker    buzas_gibson
2 | A-2      0.0356  0.00743
3 | A-1      0.0313  0.00669
```

本文材料为 **BASE (Biostack Applied bioinformatic SEies)** 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: Saturday, August 29, 2020