

tsv-utils之抽取或者删除指定列：cut

一、tsv-utils cut 介绍

功能描述：

`tsv-utils cut` 根据指定的列数值抽取或者删除对应的列，抽取列按照指定顺序重新组合。

命令行接口：

```
1 $ tsv-utils cut
2
3 Usage: tsv-utils cut [options]
4 Options:
5   -f STR  field to split
6   -s STR  sperator, default: [\t]
7   -d      delete mode.
```

可选参数：

```
1 -f 指定 抽取的列
2 -s 指定 分隔符，默认"\t"
3 -d      删除指定的行
```

二、使用场景实例及其用法

示例演示示例

示例文件： `zotu_table.txt`, `zotu_table.csv`

```
1 $ cat zotu_table.txt | head -n 6
```

```
1 #OTU ID A-1      A-2      B-1      B-2      C-1      C-2
2 ZOTU_1  0        0        87       278      1829     3608
3 ZOTU_2  223      447      1268     1583     52       69
4 ZOTU_3  0        0        162      159      1116     2021
5 ZOTU_4  0        0        99       50       1250     2172
6 ZOTU_5  0        0        1        8        1216     2143
```

```
1 $ cat zotu_table.csv | head -n 6
```

```
1 #OTU ID,A-1,A-2,B-1,B-2,C-1,C-2
2 ZOTU_1,0,0,87,278,1829,3608
3 ZOTU_2,223,447,1268,1583,52,69
4 ZOTU_3,0,0,162,159,1116,2021
5 ZOTU_4,0,0,99,50,1250,2172
6 ZOTU_5,0,0,1,8,1216,2143
```

运行命令：

选项参数1: 使用 `-f` 指定选取的列, 多列使用 `,` 分割。

```
1 $ tsv-utils cut -f 1,2,3 zotu_table.txt | head -n6
```

```
1 #OTU ID A-1      A-2
2 ZOTU_1  0        0
3 ZOTU_2  223      447
4 ZOTU_3  0        0
5 ZOTU_4  0        0
6 ZOTU_5  0        0
```

选项参数2: 通过设置 `-s` 以逗号作为列之间的分隔符。

```
1 $ tsv-utils cut -s',' -f 1,2,3 zotu_table.csv | head -n6
```

```
1 #OTU ID,A-1,A-2
2 ZOTU_1,0,0
3 ZOTU_2,223,447
4 ZOTU_3,0,0
5 ZOTU_4,0,0
6 ZOTU_5,0,0
```

注意事项: `-f` 指定的列具有顺序性。

选项参数3: 设置 `-d` 参数, 删除 `zotu_table.txt` 中指定的列数据

```
1 $ tsv-utils cut -d -f2,3 zotu_table.txt | head
```

```
1 #OTU ID B-1      B-2      C-1      C-2
2 ZOTU_1  87        278      1829     3608
3 ZOTU_2  1268      1583     52        69
4 ZOTU_3  162        159      1116     2021
5 ZOTU_4  99         50       1250     2172
6 ZOTU_5  1          8        1216     2143
7 ZOTU_6  9          16       1155     1938
8 ZOTU_7  73         41       1044     1883
9 ZOTU_8  1353      1231     2         1
10 ZOTU_9  1172      1022     1         2
```

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: 8/30/2020 3:38:56 PM