

# tsv-utils之指定的文件列替换: replace

## 一、tsv-utils replace介绍

功能描述:

`tsv-utils replace` 根据字典: `key/value` 对替换指定列中可以匹配的 `key` 的 `value`

命令行接口:

```
1 $ tsv-utils replace
2
3 Usage: tsv-utils replace [options] <db> <text>
4 Options:
5   -c INT column for replace.
6   -r remove unmapped lines.
```

可选参数:

```
1 -c 整数 选择需要替换的列;
2 -r      删除不存在注释的行;
```

## 二、使用场景实例及其用法

示例演示

示例文件: `ko.txt`, `ko-module.txt`。

```
1 $ cat annotation.txt | head -n6
```

```
1 AGE75591.1      chromosomal replication initiation protein [Bacillus
   thuringiensis serovar kurstaki str. HD73]
2 AGE75592.1      DNA polymerase III, beta subunit [Bacillus thuringiensis
   serovar kurstaki str. HD73]
3 AGE75593.1      hypothetical protein HD73_0003 [Bacillus thuringiensis
   serovar kurstaki str. HD73]
4 AGE75594.1      recombination protein F [Bacillus thuringiensis serovar
   kurstaki str. HD73]
5 AGE75595.1      DNA gyrase subunit B [Bacillus thuringiensis serovar kurstaki
   str. HD73]
6 AGE75596.1      DNA gyrase, A subunit [Bacillus thuringiensis serovar
   kurstaki str. HD73]
```

```
1 $ cat convert.txt | head -n6
```

```
1 AGE75591.1      HD73_0001
2 AGE75592.1      HD73_0002
3 AGE75593.1      HD73_0003
4 AGE75594.1      HD73_0004
5 AGE75595.1      HD73_0005
6 AGE75596.1      HD73_0006
```

运行命令：

参数使用：设置 `-c` 参数，将第一列的值进行替换, 设置 `-d` 参数, 删除无法替换的行

```
1 $ tsv-utils replace -c -d convert.txt annotation.txt | head -n6
```

```
1 HD73_0001      chromosomal replication initiation protein [Bacillus
   thuringiensis serovar kurstaki str. HD73]
2 HD73_0002      DNA polymerase III, beta subunit [Bacillus thuringiensis
   serovar kurstaki str. HD73]
3 HD73_0003      hypothetical protein HD73_0003 [Bacillus thuringiensis
   serovar kurstaki str. HD73]
4 HD73_0004      recombination protein F [Bacillus thuringiensis serovar
   kurstaki str. HD73]
5 HD73_0005      DNA gyrase subunit B [Bacillus thuringiensis serovar kurstaki
   str. HD73]
6 HD73_0006      DNA gyrase, A subunit [Bacillus thuringiensis serovar
   kurstaki str. HD73]
```

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: 2020-08-10 11:56 AM