fastx-utils之截断序列：replace

一、fastx-utils replace介绍

功能描述：

`fastx-utils replace` 根据 `key/value` 键值对，修改序列的名字。

命令行接口：

```
$ fastx-utils  replace

Usage: fastx-utils replace [options] <db> <fasta>

Options:
  -r  remove unmapped sequence.
  -c  remove comment.
```

可选参数：

```
-r        移除无映射序列;
-c        移除注释信息;
```

二、使用场景实例及其用法

示例演示：

示例文件： `protein.faa  convert.txt`

```
$ cat  protein.faa | head -n  8
```

```
>AGE80385.1
MKKIIIISATTIVIGITSFAYFGSKTPLHNEAKAVESQKHNNHKKEEIPAFPKADHNAKKIDNDFSVVTNPKSNLVLINKHRKLPDGYIPEDLTRPNVPFISPKDKEKTLLRKDAAEALENMFKAAKKEGLDLTAVSGYRSYKRQKSLHDTYVRRQGKAEANSVSAIPGTSEHQTGLAMDISSKSAKFQLEPIFGETAEGKWVAEHAHEFGFVIRYLEDKTDTTEYAYEPWHLRYVGNPYATYLYKHHLTLEEAMEDKK
>AGE79558.1
MINHELVERINFLAKKAKAEGLTEEEQRERQSLREQYLKGFRQNMLNELKGIKVVNEEGTDVTPTKLKALKKQDNAKLN
>AGE81073.1
MDKVISNEILQQFKDRMRLGDDEDANLRRILFASNKDLTRVCGNYDLNIDEVFKELVFERSRYVYNDALEYFDKNFLSQINSLSIGKALEAIKLDGD
>AGE81522.1
MKKTILTTTAALTMMGTGMGINVDHIKPAEVKADTISFYDVPNNHWATKAITNLANRNIVVGYGNGQFGFGDNVTRGQVARMIYNYLKPADAGNFKNPFSDIKGHMFEKEILALAKVGIIKGYGEGKFGPDDILTREQMAQVLTNAFKFEGTKKTSFVDVDKNSWSYKAIGALEEKGVTIGTGGNMYSPTSVVTREQYSQFLFNSINVIEKETKPEEKPNTGGEVKPEEKPNTGGEVKPEEKPNTGEETKPVNIPEWLETSLATNDFTFTQAWYDGSEAINKAASTNAQQIVKNINSKYGTNLKYSEVGAIVQLVDGAREQLWLAGMNVNDFRVTFRVSNNAMIELTKELVTLVNSDLNLDQEIQEIPSAPMKIKNVEKGDYKIRISPAMADQMITIIIEKK
```

```
$ cat convert.txt  |head  -n 6
```

```
AGE75591.1      HD73_0001
AGE75592.1      HD73_0002
AGE75593.1      HD73_0003
AGE75594.1      HD73_0004
AGE75595.1      HD73_0005
AGE75596.1      HD73_0006
```

运行命令： 使用 `-r` 参数去除不能匹配的序列

```
$ fastx-utils  replace -r  convert.txt  protein.faa  | head -n 8
```

```
>HD73_4807
MKKIIIISATTIVIGITSFAYFGSKTPLHNEAKAVESQKHNNHKKEEIPAFPKADHNAKKIDNDFSVVTNPKSNLVLINKHRKLPDGYIPEDLTRPNVPFISPKDKEKTLLRKDAAEALENMFKAAKKEGLDLTAVSGYRSYKRQKSLHDTYVRRQGKAEANSVSAIPGTSEHQTGLAMDISSKSAKFQLEPIFGETAEGKWVAEHAHEFGFVIRYLEDKTDTTEYAYEPWHLRYVGNPYATYLYKHHLTLEEAMEDKK
>HD73_3980
MINHELVERINFLAKKAKAEGLTEEEQRERQSLREQYLKGFRQNMLNELKGIKVVNEEGTDVTPTKLKALKKQDNAKLN
>HD73_5496
MDKVISNEILQQFKDRMRLGDDEDANLRRILFASNKDLTRVCGNYDLNIDEVFKELVFERSRYVYNDALEYFDKNFLSQINSLSIGKALEAIKLDGD
>HD73_6041
MKKTILTTTAALTMMGTGMGINVDHIKPAEVKADTISFYDVPNNHWATKAITNLANRNIVVGYGNGQFGFGDNVTRGQVARMIYNYLKPADAGNFKNPFSDIKGHMFEKEILALAKVGIIKGYGEGKFGPDDILTREQMAQVLTNAFKFEGTKKTSFVDVDKNSWSYKAIGALEEKGVTIGTGGNMYSPTSVVTREQYSQFLFNSINVIEKETKPEEKPNTGGEVKPEEKPNTGGEVKPEEKPNTGEETKPVNIPEWLETSLATNDFTFTQAWYDGSEAINKAASTNAQQIVKNINSKYGTNLKYSEVGAIVQLVDGAREQLWLAGMNVNDFRVTFRVSNNAMIELTKELVTLVNSDLNLDQEIQEIPSAPMKIKNVEKGDYKIRISPAMADQMITIIIEKK
```

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