

# fastx-utils之对序列进行注释：comment

## 一、fastx-utils comment 介绍

功能描述：

`fastx-utils comment` 根据 `key/value` 键值对，对序列进行注释。

命令行接口：

```
1 $ fastx-utils comment
2
3 Usage: fastx-utils comment <db> <in.fa/q>
```

## 二、使用场景实例及其用法

使用场景经典案例：

1. 对序列文件进行功能描述。

示例演示：

示例文件： `protein.faa`, `annotation.txt`

```
1 $ cat protein.faa | head -n 6
```

```
1 >AGE80385.1
2 MKKIIIIISATTIVIGITSFAYFGSKTPLHNEAKAVESQKHNNHKKEEIPAFPKADHNAKKIDNDFS SVVTNP KSNLVL
  INKHKRLPDGYIPEDLTRPNVPFISPKDKEKTL LRKDAAEALENMFKA AKKEGLDLTAVSGYRSYKRQKSLHDTYVR
  RQGKAEANSVSAIPGTSEHQTGLAMDISSKSAKFQLEPIFGETAEGKWVAEHAHEFGFVIRYLEDKTDTTEYAYEPW
  HLRYVGPNPYATYLYKHHLTLEEAMEDKK
3 >AGE79558.1
4 MINHELVERINFLAKKAKAEGLTEEEQRRERQSLREQYLKGFRQNMLNELKGIKVVNEEGTDVTP TKLKALKKQDNAK
  LN
5 >AGE81073.1
6 MDKVISNEILQQFKDRMRLGDDDEDANLRRILFASNKDLTRVCGNYDLNIDEVFKELVFERSRYVYNDALEYFDKNFL
  SQINLSIGKALEAIKLDGD
```

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

```

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]

```

运行命令：

```

1 $ fastx-utils comment annotation.txt protein.faa |head -n6
2 >AGE80385.1 D-alanyl-D-alanine carboxypeptidase [Bacillus thuringiensis
   serovar kurstaki str. HD73]
3 MKKIIISATTIVIGITSFAYFGSKTPLHNEAKAVESQKHNNHKKEEIPAFPKADHNAKKIDNDFSVVTNP
   KSNLVL
   INKHRKLPDGYIPEDLTRPNVPFISPKDKEKTLRLKDAEALENMFKAAKKEGLDLTAVSGYRSYKRQK
   SLHDTYVR
   RQGKAEANSVSAIPGTSEHQTGLAMDISSKSAKFQLEPIFGETAEGKWVAEHAHEFGFVIRYLEDKTD
   TTEYAYEPW
   HLRVYGNPYATYLYKHHLTLEEAMEDKK
4 >AGE79558.1 hypothetical protein HD73_3980 [Bacillus thuringiensis serovar
   kurstaki str. HD73]
5 MINHELVERINFLAKKAKAEGLTEEEQRERQSLREQYLKGFRQNMLNELKGIKVVNEEGTDVPTKLK
   ALKKQDNAK
   LN
6 >AGE81073.1 hypothetical protein HD73_5496 [Bacillus thuringiensis serovar
   kurstaki str. HD73]
7 MDKVISNEILQQFKDRMRLGDDEDANLRRILFASNKDLTRVCGNYDLNIDEVFKELVFERSRYVYNDA
   LEYFDKNFL
   SQINLSIGKALEAIKLDGD

```

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