

fastx-utils之抽取序列文件对应的信息：view

一、fastx-utils view介绍

功能描述：

`fastx-utils view` 抽取序列文件对应的信息，以表格形式展示。

命令行接口：

```
$ fastx-utils view

Usage: fastx-utils view [options] <fasta/q>

Options:
  -a STR    print additional fileds
  -c        print comment fileds
  -l        print length fileds
  -s        print sequence fileds
```

可选参数：

-a	输出附加信息；
-c	输出注释信息；
-l	输出长度信息；
-s	输出序列信息；

二、使用场景实例及其用法

示例演示：

示例文件： `protein.faa`

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

```
>AGE80385.1 D-alanyl-D-alanine carboxypeptidase [Bacillus thuringiensis serovar
kurstaki str. HD73]
MKKIIIIISATTIVIGITSFAYFGSKTPLHNEAKAVESQKHNNHKKEEIPAFPKADHNAKKIDNDFSVVTNPKSNLVLINK
HRKLPDGYIPEDLTRPNVPFISPKDKEKTLRKDAAEALENMFKAAKKEGLDLTAVSGYRSYKRQKSLHDTYVRRQGKAE
ANSVSAIPGTSEHQTGLAMDISSSAKFQLEPIFGETAEGKWVAEHAHEFGFVIRYLEDKTDTTEYAYEPWHLRYVGPNPY
ATYLYKHHLTLEEAMEDKK
>AGE79558.1 hypothetical protein HD73_3980 [Bacillus thuringiensis serovar
kurstaki str. HD73]
MINHELVERINFLAKKAKAEGLTEEEQRERQSLREQYLKGFQRQNMNELKGIKVVNEEGTDVTPTKLKALKKKQDNAKLN
>AGE81073.1 hypothetical protein HD73_5496 [Bacillus thuringiensis serovar
kurstaki str. HD73]
MDKVISNEILQQFKDRMLGDDDEDANLRRILFASNKDLTRVCGNYDLNIDEVFKELVFERSRYVYNDALEYFDKNFLSQI
NSLSIGKALEAIKLDGD
```

运行命令：显示 `protein.faa` 文件信息。

```
$ fastx-utils view protein.faa | head -n6
AGE80385.1
AGE79558.1
AGE81073.1
AGE81522.1
AGE80137.1
AGE76491.1
```

参数选项1: 设置 `-c` 参数, 输出注释信息。

```
$ fastx-utils view -c protein.faa | head -n6
```

```
AGE80385.1      D-alanyl-D-alanine carboxypeptidase [Bacillus thuringiensis
serovar kurstaki str. HD73]
AGE79558.1      hypothetical protein HD73_3980 [Bacillus thuringiensis serovar
kurstaki str. HD73]
AGE81073.1      hypothetical protein HD73_5496 [Bacillus thuringiensis serovar
kurstaki str. HD73]
AGE81522.1      S-layer domain-containing protein (plasmid) [Bacillus
thuringiensis serovar kurstaki str. HD73]
AGE80137.1      Methylene tetrahydrofolate reductase [Bacillus thuringiensis
serovar kurstaki str. HD73]
AGE76491.1      Rhodanese domain protein [Bacillus thuringiensis serovar kurstaki
str. HD73]
```

参数选项2: 设置 `-l` 参数, 输出长度信息。

```
$ fastx-utils view -l protein.faa | head -n6
```

```
AGE80385.1      259
AGE79558.1      79
AGE81073.1      97
AGE81522.1      402
AGE80137.1      610
AGE76491.1      98
```

参数选项3: 设置 `-s` 参数, 输出序列信息。

```
$ fastx-utils view -s protein.faa | head -n6
```

AGE80385.1
MKKIIIIISATTIVIGITSFAYFGSKTPLHNEAKAVESQKHNNHKKEEIPAFPKADHNAKKIDNDFS SVVTNPKSNLVLINK
HRKLPDGYIPEDLTRPNVPFISPKDKEKTL LRKDAAEALENMFKA AKKEGLDLTAVSGYRSYKRQKSLHDTYVRRQGKAE
ANSVSAIPGTSEHQTGLAMDISSSAKFQLEPIFGETAEGKWVAEHAHEFGFVIRYLEDKTDTTEYAYEPWHLRYVGNPY
ATYLYKHHLTLEEAMEDKK
AGE79558.1
MINHELVERINFLAKKAKAEGLTEEEQRERQSLREQYLKGFRQNMNELKGIKVVNEEGTDVTPTKLKALKKQDNAKLN
AGE81073.1
MDKVISNEILQQFKDRMRLGDDANLRRILFASNKDLTRVCGNYDLNIDEVFKELVFERSRYVYNDALEYFDKNFLSQI
NSLSIGKALEAIKLDGD
AGE81522.1
MKKTILTTTAALTMMGTGMGINVDHIKPAEVKADTISFYDVPNNHWATKAITNLANRNIVVGYGNGQFGFGDNVTRGQVA
RMIYNYLKPADAGNFKNPFSDIKGHMFEKEILALAKVGIIKGYGEGKFGPDDILTREQMAQVLTNAFKFEGTKKTSFVDV
DKNSWSYKAIGALEEKGVITGTGGNMYSPTSVVTREQYSQFLFNSINVIEKETKPEEKPNTGGEVKPEEKPNTGGEVKPE
EKPNTGEETKPVNIPEWLETSLATNDFTFTQAWYDGSEAINKAASTNAQQIVKNINSKYGTNLKYSEVGAIVQLVDGARE
QLWLAGMNVNDFRVTFRVSNNAMIELTKELVTLVNSDLNLDQEIQEIPSAPMKIKNVEKGDYKIRISPAMADQMITIIE
KK
AGE80137.1
MKLLDLLSKGIVIGDGAVGTL LSHGLQSSFEELNVSDPLIISIHKQYVAAGADV IQTNTYGANEA KL RMYGLENQVTK
INRAAVKLAKASVTDKNAILGTIGGMKHIGAVTTTDMEREFMLLEQAGALLEEQVDGLLLET FYDEFELLHAVKVLRKQT
NIPIVAQLALHEAGTTQNGNDVNEILKQFIDYGANVVGLNCQLGPLHMT EAFKMISIPQNGYL SAYPNAGLP NYVEGRYV
YEGSPAYFEAMTPNFIEQGIRLLGGCCGTTPEHIQSMKRAVANITPVIEKETIQRPKVVH THEKRSKAHVTLAEKAKKQT
TVVVELDPKLTDTQRF FEGARALKRAGADAITLADNSLASPRVSNMAMGALLTKHDIPVLTHLTCRDHNVI GLQSHLLG
LSALGMEEV LALTGD PARVGD FPGATSVYDLSSIELIKMIKEMNDGRSILGKSIGPATRFSVGGAFNPHVRHLKAAVKRM
ERKIDAGAEYFLTQPIYDIK LIEEVYEATKHLEQPIFIGIMPLVSKRNADFLHFEVPGITLPEEIRERMDGHETKEAAIE
EGIRISQELVDAAMKYFNGIYLITPFLKYEITEHLVKYVREKQEVKEGIN
AGE76491.1
MKEMTAKELEEKLLRKEAVNIVDVREVEEVAEGKIPEACNIPLGLLEFRMHeldKKKEYIIVCRSGGRSARAVQFLESYG
FQAINMVGGM LAWEGKV

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