tsv-utils之汇总文件拆分: unpack

一、tsv-utils unpack介绍

功能描述：

`tsv-utils unpack` `bins` 子命令的反操作，将 `bin` 的文件转换成 `key/value` 关系，拆分指定分隔符的 `values` 数值。

命令行接口：

```
1  $ tsv-utils  unpack
2
3  Usage: tsv-utils unpack [options]  <bins>
4  Options:
5    -d char  delimiter for elements, default: [,]
6
7  Example:
8  Feature  1,2,3
```

可选参数：

```
1  -d  字符  元素之间的间隔符，默认为' , ';
```

二、使用场景实例及其用法

示例演示

示例文件： `kofams.txt`

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

```
1  #seqid  ko       threshold        score    evalue  domain_coverage
   query_coverage  definition       type
2  Y1_g_00002      K03686  415.90  500.5   1.4e-150         0.50    0.98
   molecular chaperone DnaJ
3  Y1_g_00003      K04043  830.13  895.1   7.1e-270         0.79    0.99
   molecular chaperone DnaK
4  Y1_g_00004      K03687  30.90   186.6   1.1e-55 0.43    0.92    molecular
   chaperone GrpE
5  Y1_g_00005      K03705  67.20   310.9   2.3e-93 0.79    0.98    heat-
   inducible transcriptional repressor
6  Y1_g_00006      K02495  263.57  311.1   2.2e-93 0.71    0.97    oxygen-
   independent coproporphyrinogen III oxidase [EC:1.3.98.3]
```

运行命令:

```
1  $ tsv-utils  bins  -t 2  -s 1  -d ';'  kofams.txt  | cut -f1,3 | tee
   ko.bins.txt | head -n7
```

```
1  #catalog         members
2  K13052  Y1_g_01376
3  K13292  Y1_g_00935
4  K03495  Y1_g_01211
5  K03496  Y1_g_01823
6  K01071  Y1_g_01732
7  K03497  Y1_g_01822;Y1_g_01824
```

```
1  $ tsv-utils  unpack -d  ';'  ko.bins.txt | head -n 7
```

```
1  #catalog         members
2  K13052  Y1_g_01376
3  K13292  Y1_g_00935
4  K03495  Y1_g_01211
5  K03496  Y1_g_01823
6  K01071  Y1_g_01732
7  K03497  Y1_g_01822
```

本文材料为 **BASE** (**B**iostack **A**pplied bioinformatic **SE**ies) 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 7:03:58 PM