

# fastx-utils之根据index数据拆分: demultiplex

## 一、fastx-utils demultiplex

功能描述:

`fastx-utils demultiplex` 将PE双端序列转成交叠的序列文件。

命令行接口:

```
$ fastx-utils demultiplex

Usage: fastx-utils demultiplex [options]

Options:
  -d output directory for demultiplex reads. default: [./]
  -i index sheet [sample\tindex].
  -1 forward reads.
  -2 reverse reads.
```

参数接口:

```
-d 数据输出目录
-i 信息表 [样本\t索引]
-1 正向测序数据.
-2 反向测序数据.
```

## 二、使用场景实例及其用法

示例演示:

示例文件: `A1.1.fastq.gz`, `A1.1.fastq.gz`, `index.txt`

```
$ zcat A1.1.fastq.gz | head -n 6
```

```
@A01050:39:HJ5FKDRXX:2:2135:18141:18004 1:N:0:TCCGCGAA+GGCTCTGA
GCCCTGTAGGACTACAAGGGTTTCTAATCCTGTTTCGATACCCGCACTTTCGAGCTTCAGCGTCAGTTGCGCTCCAGTGAG
CTGCCTTCGCAATCGGAGTTCTTCGTGATATCTAAGCATTTACCCGCTACACCACGAATTCGCCCCACTTTGTGCGTACT
CAAGGAAACCAGTTTCGCGCTGCAGTTCAGATGTTGAGCATCTACATTTACAACACGCTTAATCTCCGGCCTACGCTCCC
TTTAAACCCA
+
F:FF,:::FFFFFFFF:FFFFFFFFFFFFFFF:,FFFF,FF:,FFFFFFFF:FFFFF:FFFFFF:FFFF,FFFFFFFF:FFF
FFFF:FFF:FFFF,FFFFFFFF:,FFFF::FF:,FF:,FFFFFF:FFFF,F:FFFF,FFFFFF:FFFF:FFFFFFFF
FFFFFFFF:F:FFFF,FFFFFFFFFFFFFFFFFFFFFFFFFFFFF,:FFFFFFFFFFFF,F,: ,FF:,FF:FFFF:FFF
FF,F:FFFF
@A01050:39:HJ5FKDRXX:2:2116:24451:15969 1:N:0:TCCGCGAA+GGCTCTGA
CTGCTACAGTGCCAGCCGCCGCGTAATACGGAGGATGCGAGCGTTCTCCGATTTATTGGGTTTAAAGGGTGCGTAGGC
GGAAAGTTAAGTCAGCGGTAAAATCGAGAGGCTCAACCTCTTTCAGCCGTTGAACTGGTTTCTAGAGTGAGTGAGAAG
TACGCGGAATGCGAGGTGTAGCGGTGAAATGCATAGATATCACGCAGAACTCCAATTGCGAAGGCAGCGTACCGGCACTC
AACTGACGCT
```

```
$ zcat A1.2.fastq.gz | head -n 6
```

```
@A01050:39:HJ5FKDRXX:2:2135:18141:18004 2:N:0:TCCGCGAA+GGCTCTGA
TACCGGAAGTGCCAGCAGCCGCGTAATACGGAAGGTCCGGGCGTTATCCGATTTATTGGGTTTAAAGGGAGCGTAGGC
CGGAGATTAAGCGTGTGTGACATGTAGATGCTCAACATCTGAACTGCAGCGCGAACTGGTTTCTTGAGTACGCACAAA
GTGGGCGGAATTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTCACTGGAGCG
CACCTGACGC
+
FFFFFFFFFFFFFFFFFFFFFFFFFFFF,FFFFFFFFFFFF:FFFF:FFFFF:,FFFFFFFFFFFFFFFFFFFF,FFFFFFFFF
FFFFFFFFFFFFFFFF::FFFF,FFFFF:F:FFF:FFFFFFFFFFFFFFFFFFFFF,:FF:FFFFFFFFFFFFFFFFFFFFF
::FFFFFFFFFFFFFFFF,FFFFFFFFFFFF:FFFFFFFFFFFF:FFFFFFFF:FFFFF:FFFFFFFFFFFFFFFF,FFF,F
FF:FFFFFF
@A01050:39:HJ5FKDRXX:2:2116:24451:15969 2:N:0:TCCGCGAA+GGCTCTGA
TAAATGCGGGACTACAGGGTATCTAATCCTGTTTCGATCCCCACGCTTTCGTGCTTCAGCGTCAGTTGAGTGCCGCTACG
CTGCCTTCGCAATTGGAGTTCTGCGTGATATCTATGCATTTACCCGCTACACCTCGATTCCGCGTACTTCTCACTCACT
CTAGAAAACAGTTTCAACGGCTGAAAGAGGTTGAGCCTCTCGATTTTACCGCTGACTTAACTTTCCGCCTACGCACCCT
TTAAACCCAA
```

```
$ cat index.txt
```

```
D001 TCCGCGAA+GGCTCTGA
```

**运行命令：** 将双端序列转成交叠的序列文件。

```
$ fastx-utils demultiplex -d ./ -i index.txt -1 A1.1.fastq.gz -2
A1.2.fastq.gz
```

```
$ ls
A1.1.fastq.gz A1.2.fastq.gz D001_R1.fastq D001_R2.fastq index.txt
```

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: 9/2/2020 12:37:53 AM

