

# fastx-utils之双端序列转成交叠的序列文件：interleave

## 一、fastx-utils interleave介绍

功能描述：

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

命令行接口：

```
1 $ fastx-utils interleave
2
3 Usage: fastx-utils interleave [options] <in1.fq> <in2.fq>
4 Options:
5     -o old fashion interleave fastq format
```

可选参数：

```
1 -o 使用老的序列显示模式 1/, 2/
```

## 二、使用场景实例及其用法

使用场景经典案例：

1. 产生Jellyfish软件的输入数据。

示例演示：

示例文件： `A1.1.fastq`, `A1.2.fastq`

```
1 $ cat A1.1.fastq | head -n 6
```

```
1 @HISEQ:483:HLJ2LBCXY:1:2215:13265:62834 1:N:0:CACCGG
2 CCTATGGGATGCAGCAGTGGGGAATCTTGCCCAATGGGCGCAAGCCTGAGGCAGCGACGCCGCTGAGGGATGACGG
3 CCTTCGGGTTGTAAACCTCGTTCAGCAGGGACGAAGCGAAAGTGACGGTACCTGCAGAAGAAGCACCGGCTAACTAC
4 GTGCCAGCAGCCGCGGTAATACGTAGGGTGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCTCGTAGGCGGCTT
5 GTCGCGTCTGCT
6 +
7 IHIIIGGIIIEHIIIIHIIIIIIHIIIIIIIIIIIIIIHEHIIIIIIIIHIIHFIIIIHIIIIIGIIIIIIIIIIID
8 HHIIIIIIIIIIH?
9 HHFHIIHIIIIHIIIIHIIIIIIHIIHIIHIIIIIIIGIIIIHHHIIHHHIIHIIIIIIHIIHHHIIHII
10 HIGIIIIHIIHIIHHHGHCAH.?GHHFD-EAFGHIIHHIDHHGFHHIIGFFHEHHHHCH,,CHHIGIHHCHII
11 @HISEQ:483:HLJ2LBCXY:1:1210:15779:30764 1:N:0:CACCGG
12 CCTATGGGGGACCAAGTGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGGGTGAAGAAGG
13 CCTTAGGGTTGTAAACCCCTTTCAGCGGGGAAGATAATGACGGTACCCGCAGAAGAAGTCCCGGCTAACTCCGTGCC
14 AGCAGCCGCGGTAATACGGAGGGGACTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATTTGTAA
15 GTCAGGGGTGAA
16
17 $ cat A1.2.fastq | head -n 6
```

```

1 @HISEQ:483:HLJ2LBCXY:1:2215:13265:62834 2:N:0:CACCGG
2 GGACTACCCGGGTATCTAATCCTGTTGCTCCCCACGCTTTCGCTCCTCAGCGTCAGCAACGGCCCAGAGACCCGCC
  TTCGCCACCGGTGTTCTCCTGATATCTGCGCATTCCACCGCTACACCAGGAATTCGGTCTCCCCTACCGCGCTCT
  AGCCTGCCCGTATCCACTGCAGGCCCGGAGTTAAGCTCCGGGTTTTACAGCAGACGCGACAAGCCGCCTACGAGCT
  CTTTACGCCCACTA
3 +
4 HH?
  GHIHHIIIGIDHHIEDF@HHIIIIIGGHHIIIIIIIIIIIGIIIIIIIIIIIGIHHIIIIIIIIIIHHHHI
  HHHFHHIHCHEIIIEF<GHFHGHGHHIIHHGEFHIIIIHHIIIIH@GGHEHHIIIGIIHHHHH@HHGHDHHHHHHI
  IHCHIIHH?GEHHHHHHHEHIIII=HDE@6-8?E?GGIHG=@-GHHIHHH-@D,+>D:>@E++;E-6@@@C+--
  @@-6@-C:C-@@
5 @HISEQ:483:HLJ2LBCXY:1:1210:15779:30764 2:N:0:CACCGG
6 GGACTACGGGGGTATCTAATCCTGTTGCTCCCCACGCTTTCGACCTCAGCGTCAGTACCGGACCAGTGAGCCGCC
  TTCGCGACTGGTGTCTTCTAATATCTACGAATTCACCTCTACACTAGGAATTCACCTCACCTCTTCCGGACTCG
  AGATCCGCAGTATCAAAGGCAGTTCGAGGTTGAGCCTCGGGATTTACCCCTGACTTACAAATCCGCCTACGTGCG
  CTTTACGCCCACTA

```

运行命令：将双端序列转成交叠的序列文件。

```
1 $ fastx-utils interleave A1.1.fastq A1.2.fastq | head -n8
```

```

1 @HISEQ:483:HLJ2LBCXY:1:2215:13265:62834 1:N:0:CACCGG
2 CCTATGGGATGCAGCAGTGGGAATCTTGCCCAATGGGCGCAAGCCTGAGGCAGCGACGCCGCTGAGGGATGACGG
  CCTTCGGGTTGTAAACCTCGTTCAGCAGGGACGAAGCGAAAGTGACGGTACCTGCAGAAGAAGCACCGGCTAACTAC
  GTGCCAGCAGCCGCGGTAATACGTAGGGTGCAAGCGTTGTCCGAATTATTGGGCGTAAAGAGCTCGTAGGCGGCTT
  GTCGCGTCTGCT
3 +
4 IHIIIGGIIIEHHIIIIHHIIIIIIHHIIIIIIIIIIHEHHIIIIIIHHIHFIIIIHHIIIIIGIIIIIIIIIIID
  HHIIIIIIIIIIH?
  HHHFHHIIIIHHIIIIHHIIIIIIHHHHIIHHIIIIIIIIIGIIIIHHHHHHHHHHIIHHIIIIHHHHHHHHI
  HIGIIIIHHIIHHHHGHCAH.?GHHFD-EAFGHHIHHIDHHGFHHIIGFFHEHHHHCH,,CHHIGIHHCHII
5 @HISEQ:483:HLJ2LBCXY:1:2215:13265:62834 2:N:0:CACCGG
6 GGACTACCCGGGTATCTAATCCTGTTGCTCCCCACGCTTTCGCTCCTCAGCGTCAGCAACGGCCCAGAGACCCGCC
  TTCGCCACCGGTGTTCTCCTGATATCTGCGCATTCCACCGCTACACCAGGAATTCGGTCTCCCCTACCGCGCTCT
  AGCCTGCCCGTATCCACTGCAGGCCCGGAGTTAAGCTCCGGGTTTTACAGCAGACGCGACAAGCCGCCTACGAGCT
  CTTTACGCCCACTA
7 +
8 HH?
  GHIHHIIIGIDHHIEDF@HHIIIIIGGHHIIIIIIIIIIIGIIIIIIIIIIIGIHHIIIIIIIIIIHHHHI
  HHHFHHIHCHEIIIEF<GHFHGHGHHIIHHGEFHIIIIHHIIIIH@GGHEHHIIIGIIHHHHH@HHGHDHHHHHHI
  IHCHIIHH?GEHHHHHHHEHIIII=HDE@6-8?E?GGIHG=@-GHHIHHH-@D,+>D:>@E++;E-6@@@C+--
  @@-6@-C:C-@@

```

参数选项**1**：设置 `-o` 参数, 生成老格式 `fastq`。

```
1 $ fastx-utils interleave -o A1.1.fastq A1.2.fastq | head -n8
```

```

1 @HISEQ:483:HLJ2LBCXY:1:2215:13265:62834/1
2 CCTATGGGATGCAGCAGTGGGAATCTTGCCCAATGGGCGCAAGCCTGAGGCAGCGACGCCGCTGAGGGATGACGG
  CCTTCGGGTTGTAAACCTCGTTCAGCAGGGACGAAGCGAAAGTGACGGTACCTGCAGAAGAAGCACCGGCTAACTAC
  GTGCCAGCAGCCGCGGTAATACGTAGGGTGCAAGCGTTGTCCGAATTATTGGGCGTAAAGAGCTCGTAGCGCGCTT
  GTCGCGTCTGCT
3 +
4 IHIIIGGIIIEHIIIIHIIIIIIHIIIIIIIIIIIIHEHIIIIIIIIHIIHFIIIIHIIIIIGIIIIIIIIIIID
  HHIIIIIIIIIIH?
  HHFHIIHIIIIHIIIIIIHIIIIIIHIIHHIIHIIIIIIIGIIIIHHHHIHHHHIIHIIIIIIHIIHHHHIHHI
  HIGIIIIHIIIIHHHHGHCAH.?GHHFD-EAFGHIIHHIDHHGFHHIIGFFHEHHHHCH,,CHHIGIHHCHII
5 @HISEQ:483:HLJ2LBCXY:1:2215:13265:62834/2
6 GGACTACCCGGGTATCTAATCCTGTTCGCTCCCCACGCTTTCGCTCCTCAGCGTCAGCAACGGCCAGAGACCCGCC
  TTCGCCACCGGTGTTCTCCTGATATCTGCGCATTCCACCGCTACACCAGGAATTCGGTCTCCCTACCGCGCTCT
  AGCCTGCCCGTATCCACTGCAGGCCCGGAGTTAAGCTCCGGGTTTTACAGCAGACGCGACAAGCCGCCTACGAGCT
  CTTTACGCCCACTA
7 +
8 HH?
  GHIHHIIIGIDHHIEDF@HHIIIIIGGHHIIIIIIIIIIIGIIIIIIIIIIIGIHHIIIIIIIIIIHIIHHHI
  HHFHIIHCEHIIIEF<GHFHGHIIIIHGEFHIIIIHIIIIH@GGHEHHIIIGIIIIHHHH@HHGHDHHHHHHI
  IHCHIIHH?GEHHHHHHHEHIIII=HDE@6-8?E?GGIHG=@-GHIIHH-@D,+>D:>@@E++;E-6@@@C+--
  @@-6@-C:C-@@

```

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