

atlas-utils之根据标记基因拷贝数对扩增子数据进行标准化: normalization

一、atlas-utils normalization介绍

功能描述:

`atlas-utils normalization` 根据标记基因拷贝数对扩增子数据进行标准化, 做功能预测预处理。

命令行接口:

```
1 $ atlas-utils normalization
2
3 Usage: atlas-utils normalization <copy_number> <otutab>
```

二、使用场景实例及其用法

使用场景经典案例:

16S扩增子功能预测数据分析, 根据数据库 (`greengenes`) 的16S拷贝数, 对 `OTU` 表进行标准化操作。

示例演示:

示例文件: `16S.txt.gz`

```
1 $ zcat 16S.txt.gz | head -n 6
```

```
1 #OTU_IDS      16S_rRNA_Count
2 370251  5.0
3 4473812 2.0
4 3209311 3.0
5 4473818 4.0
6 4470541 2.0
```

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

```
1 #OTU ID E01L   E01S   N01L   N01S   S01L   S01S
2 564806 10      14      0       2       0       0
3 145801 7974    429     4      265     1      233
4 339087 4386    49      162    563     329    186
5 370287 197     13      47     45      36      6
6 4449244 10138  4       0      104     2       61
```

运行命令:

根据标记基因拷贝数对扩增子数据进行标准化, 做功能预测预处理

```
1 $ atlas-utils normalization 16S.txt.gz otu_table.txt | head -n 8
```

1	#OTU ID	E01L	E01S	N01L	N01S	S01L	S01S
2	564806	3.33333	4.66667	0	0.666667		0
3	145801	2658	143	1.33333	88.3333	0.333333	77.6667
4	339087	1096.5	12.25	40.5	140.75	82.25	46.5
5	370287	49.25	3.25	11.75	11.25	9	1.5
6	4449244	1448.29	0.571429		0	14.8571	0.285714
7	513445	196.8	31.8	4.6	135.2	23.4	25.4
8	351231	8	11	7	205.8	153.6	575.2

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: 2020-09-09 11:56 AM