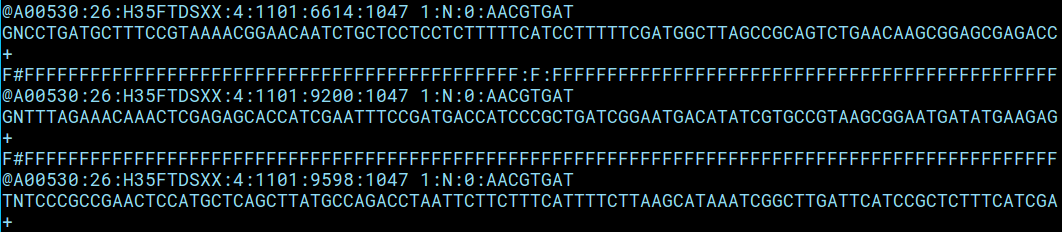
python3生信入门-提取fastq文件中的序列并转为fasta格式

测试数据及示例脚本已上传至百度盘（提取码 u869）：

<https://pan.baidu.com/s/1v-Yo_nI0ko8hYzWr-V7b6w>

## 说明

网盘中提供的测试数据“test.fastq”为一个fastq文件，其内容如下：

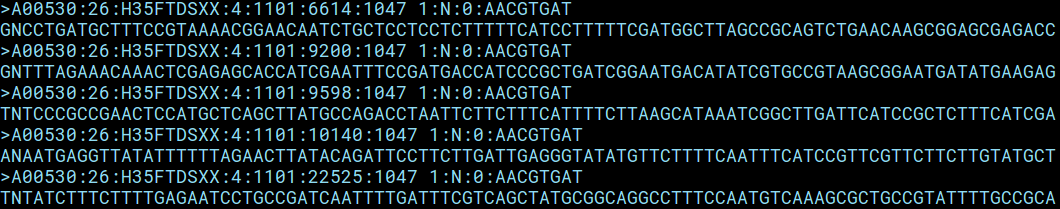


现在我们需要使用python3，对该fastq文件进行处理，要求：

（1）将该fastq文件中的碱基序列提取出来，并以fasta格式存储；

（2）同时将fastq中每条序列的id中的“@”替换为“>”后，作为新fasta文件中每条序列的id。

预期所得结果“test.fasta”内容如下：



请根据要求编写一个简单的py脚本。

## 示例

一个示例脚本如下（可参见网盘附件“fq\_fa.1.py”）。

#!/usr/bin/env python3

# -\*- coding: utf-8 -\*-

fasta = open('test.fasta', 'w')

with open('test.fastq', 'r') as fastq:

for line in fastq:

if line[0] == '@':

line = line.strip('[@\n]')

print('>' + line, file = fasta)

print(fastq.readline().strip(), file = fasta)

fastq.close()

fasta.close()

首先新建一个文件“test.fasta”，待写入内容。

其次打开“test.fastq”，使用循环逐行读取其内容并对每一行进行判断，若第一个字符为“@”，则将该行起始的“@”删除并在起始处加入“>”后，写入到文件“'test.fasta”中；同时，将其下一行的碱基序列也一并输出到文件“'test.fasta”中（写在序列id的下一行）。

处理完成后，关闭打开的文件，此时的“test.fasta”即为所得结果。



备注：

网盘附件“fq\_fa.py”为添加了命令传递行的python3脚本，可在shell中直接进行目标文件的I/O处理。其使用方式如下：

#python3 fq\_fa.py -h

python3 fq\_fa.py -i test.fastq -o test.fasta

