目录

[1 代码说明 1](#_Toc106226702)

[1.1 Master版本 1](#_Toc106226703)

[1.1.1 代码目录 1](#_Toc106226704)

[1.1.2 编译程序 2](#_Toc106226705)

[1.1.3 启动脚本 2](#_Toc106226706)

[1.2 Dp版本 2](#_Toc106226707)

[1.2.1 代码目录 2](#_Toc106226708)

[1.2.2 安装依赖包 2](#_Toc106226709)

[1.2.3 启动脚本 2](#_Toc106226710)

[2 工具说明 3](#_Toc106226711)

[2.1 模拟工具wgsim 3](#_Toc106226712)

[2.2 评价标准程序OPAL 3](#_Toc106226713)

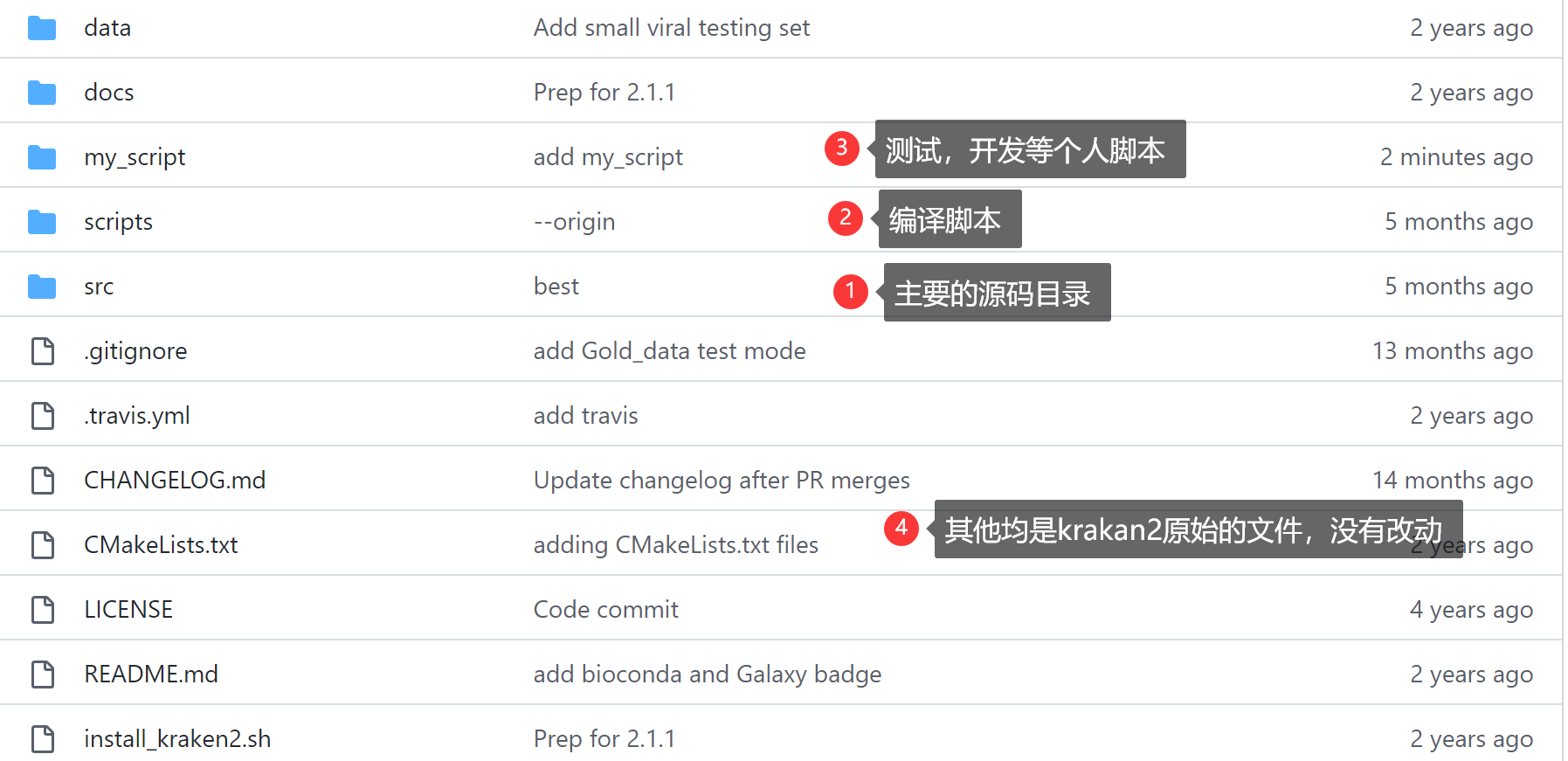
[3 其他说明 3](#_Toc106226714)

# 代码说明

## Master版本

Master版本是权重k-mer的实现版本，其代码位置在<https://github.com/zhang570221322/kraken2>的master分支。

### 代码目录





### 编译程序

需要先对程序进行编译g++11以上。

cd kraken2

./install\_kraken2.sh ../kraken2\_build 此为安装目录

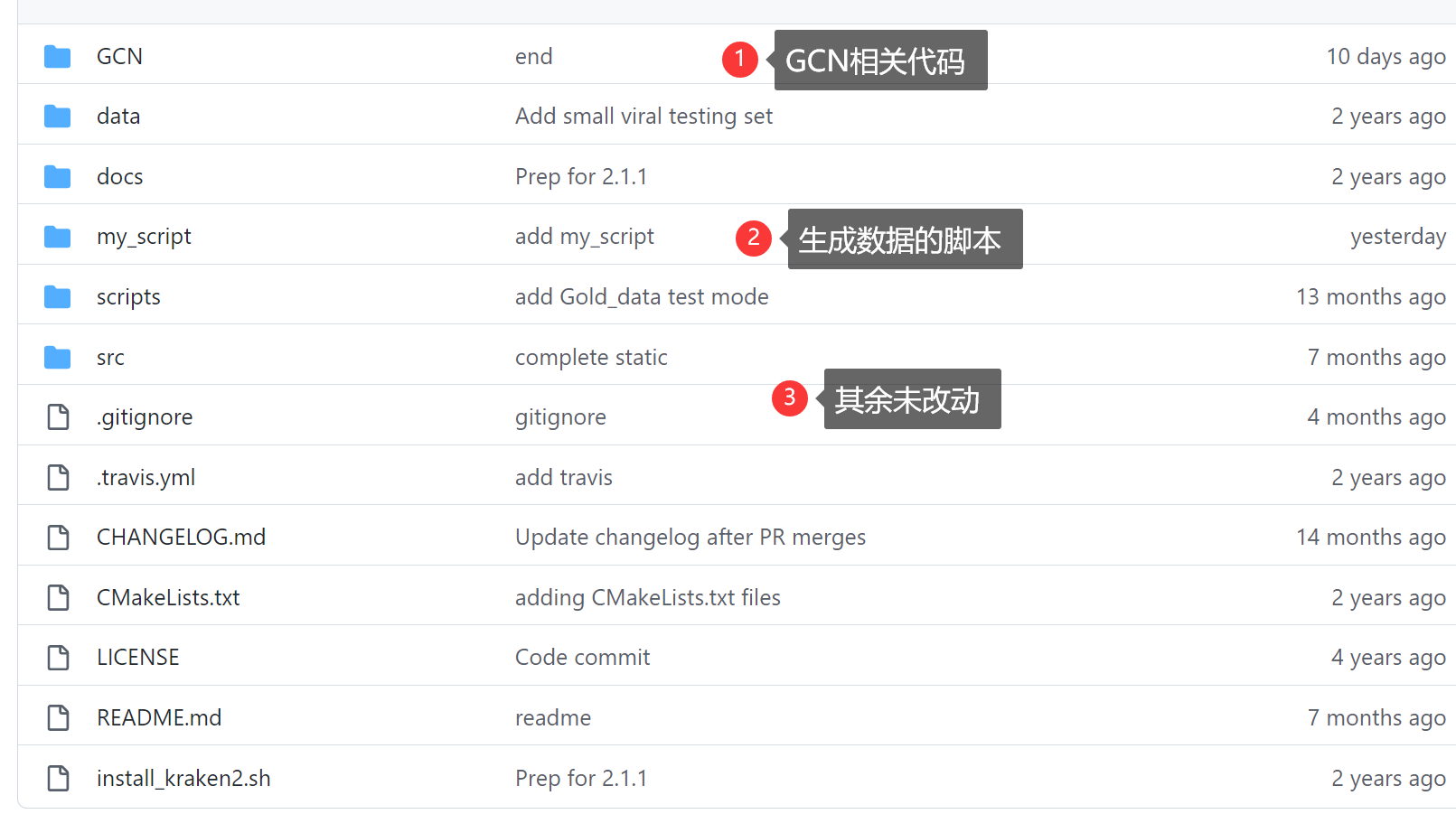
### 启动脚本

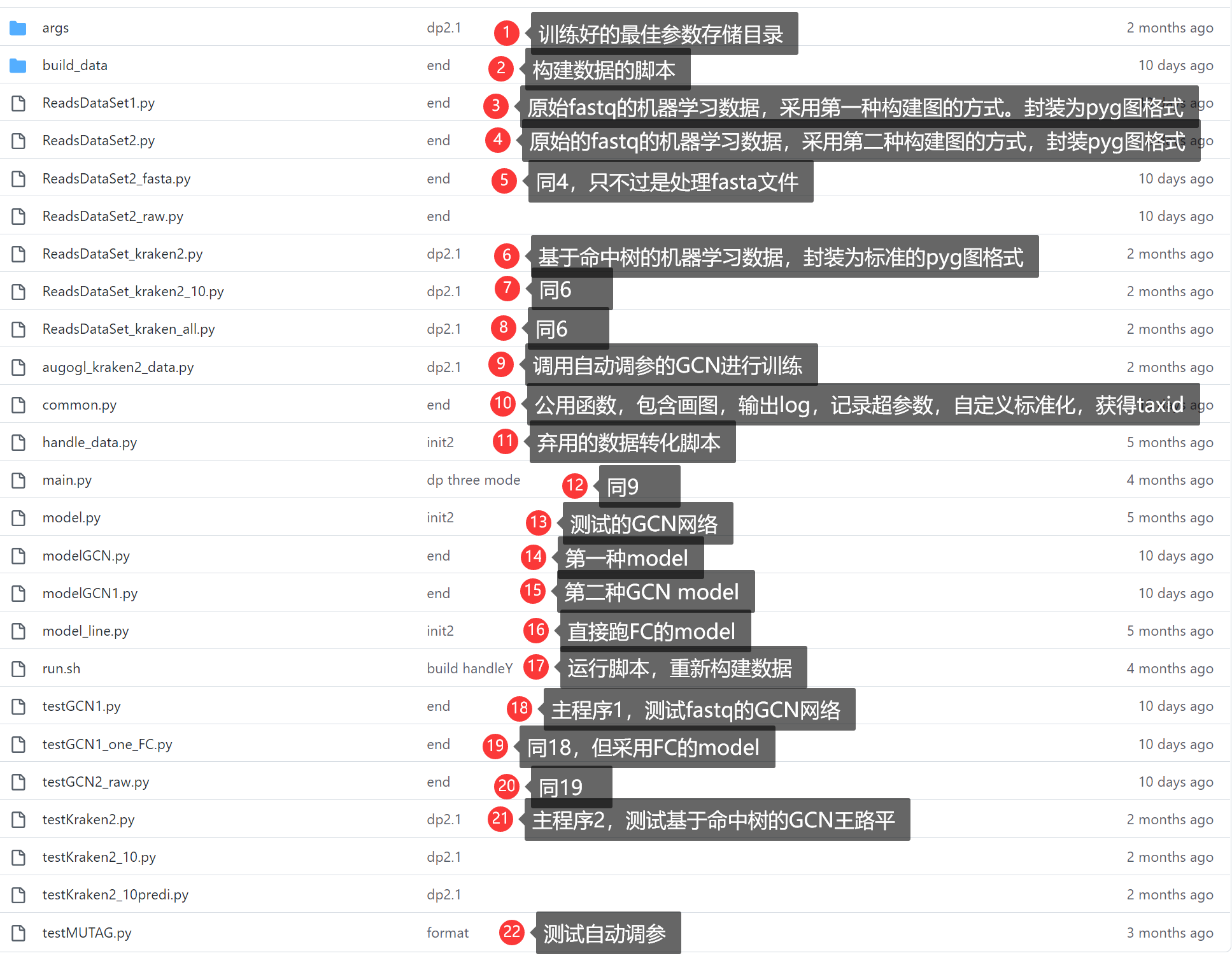
可以在my\_script/test.pbs运行启动。

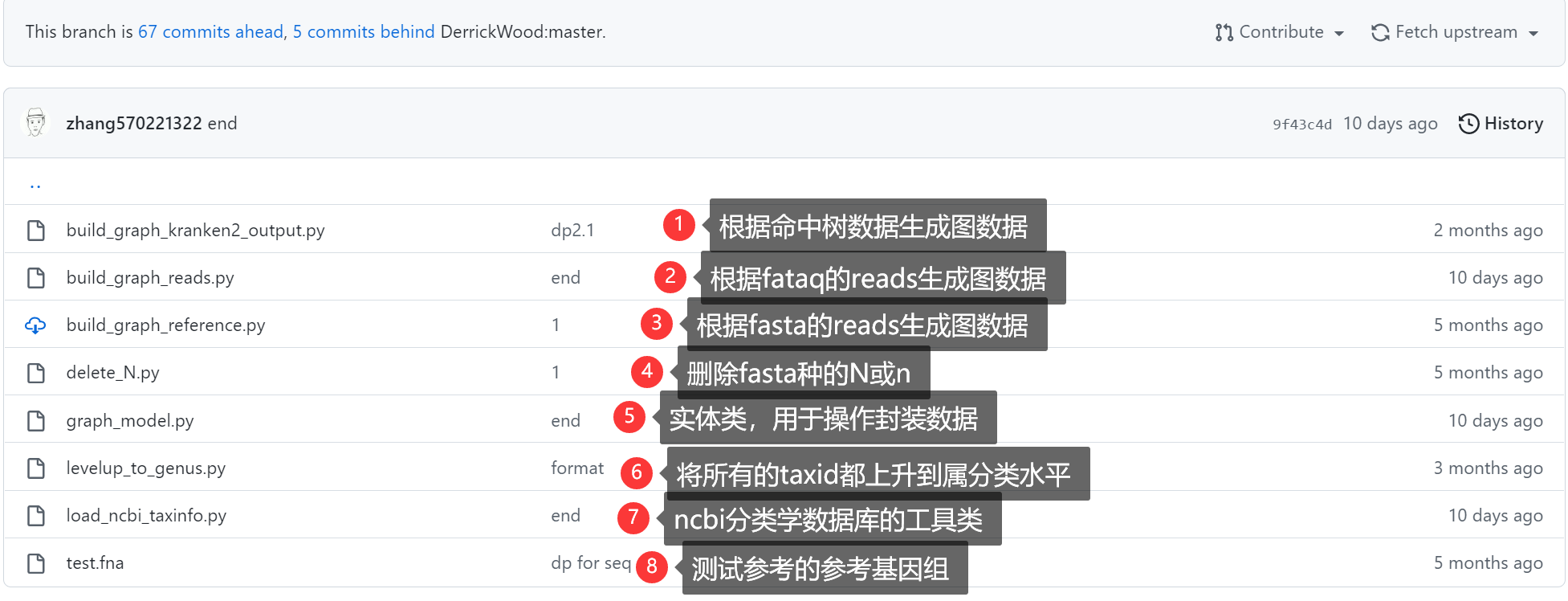
## Dp版本

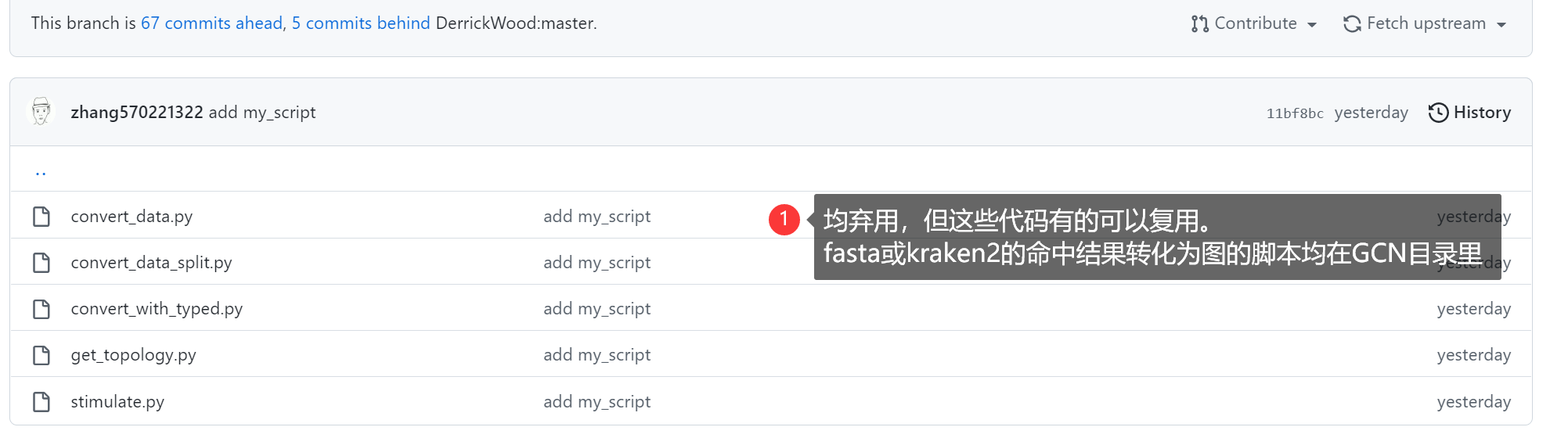
Master版本是基于GCN的实现版本，只要实现了两个方法，一个是基于命中树筛选最优标签，一个是基于原始的fastq文件进行标签分类。其代码位置在<https://github.com/zhang570221322/kraken2>的dp分支。

### 代码目录









### 安装依赖包

所有包都安装gpu电脑上的conda环境：zwl-py-gpu,包目录位于（/home/yegpu/anaconda3/envs/zwl-py-gpu），具体包描述在envs\_library文件中https://github.com/zhang570221322/kraken2/blob/dp/GCN/envs\_library

### 启动脚本

从GCN/test\*.py开始启动，包含两个主程序，testGCN.py和testkrkane2.py

# 工具说明

## 模拟工具wgsim

李衡前辈写的模拟wgs数据的模拟器，<https://github.com/lh3/wgsim>。

## 评价标准程序OPAL

CAMI挑战赛使用的标准化评价工具，<https://github.com/CAMI-challenge/OPAL>。

# 其他说明

需要预先下载参考数据库数据，然后构建kraken的数据库。

0. 下载相应的taxonomy。 https://ftp.ncbi.nlm.nih.gov/pub/taxonomy/

1. 下载所有的COVID参考序列。 https://ftp.ncbi.nlm.nih.gov/genomes/refseq/bacteria/assembly\_summary.txt

2.  下载ancestor\_taxid:186801,count:6798的所有物种 可以从脚本getURLAndDownLoad.py下载