|  |  |  |
| --- | --- | --- |
| 文件编号：SOP-XX-110 | **快速比对分析操作规范** | 拟制人： 马威锋 |
| 生效日期： 2016.X.X | 审核人： |
| 版 本：第一版 | 批准人： |

1. 目的：

根据下机样本数据，在集群下分析，生成比对分析结果，结果统计样本数据的比对情况。

1. 范围：

测序下机数据，用于产前检测等产品的数据分析和结果生成。

1. 职责：

信息部。

1. 操作步骤：
   1. 公共密钥

如果已经做过设置，则忽略公共密钥设置教程。

* + 1. 注意事项：
       1. 修改密码

首次登录需在10.20.4.11上进行，如果修改时有交互，回车即可。

* + - 1. 投放任务

登录到节点上，目前可以在此节点上跑程序。

* + - 1. 保存数据

申请开通工作目录后，今后的工作在工作目录进行，/home下的个人目录中仅允许放小的程序，不允许放大数据，空间有限。

* + 1. 秘钥生成
       1. 生成密钥 (可忽略)

## ssh-keygen -t rsa

输入命令后一直按回车

* + - 1. 切换到家目录的.ssh目录

## cd ~/.ssh

## cat id\_rsa.pub > authorized\_keys

具体操作如下图内容：

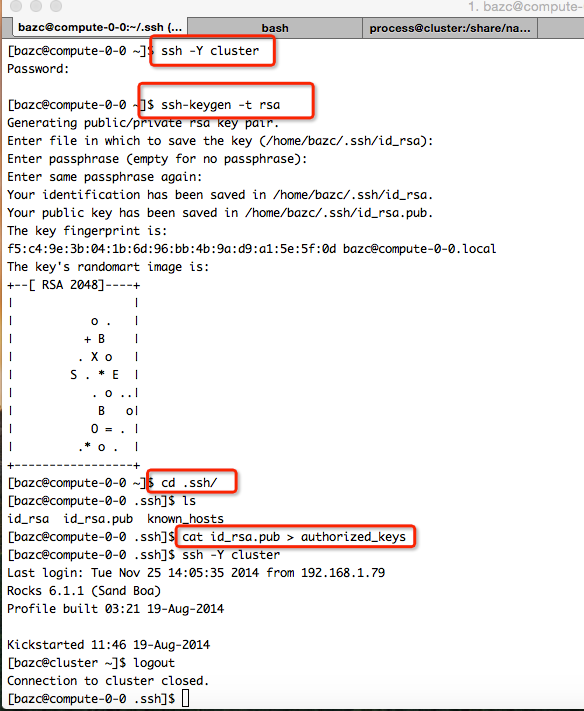


图1

公共密钥配置成功才能在集群上做下面的操作。

* 1. 操作方法：
     1. 准备包

需要软件包hash ，hash的内容如下图２

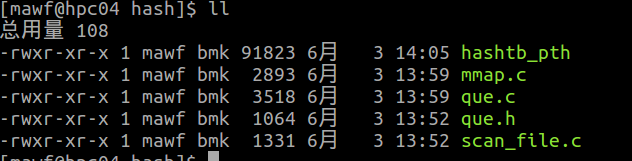


图2

* + 1. 准备文件

参考基因组文件hg19.fa.

下机测序fq格式文件，如：

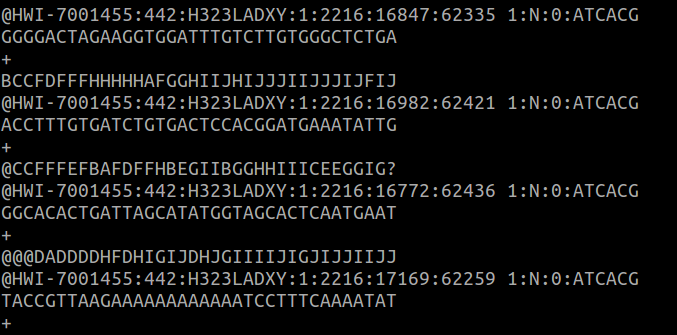


图３

说明：

每4行记录一条reads信息：

第一行：以@开头，标示一条序列，包括测序仪名字，标记位置，单/双端测序以及过滤情况和引物接头等相关信息描述；

第二行：reads序列；

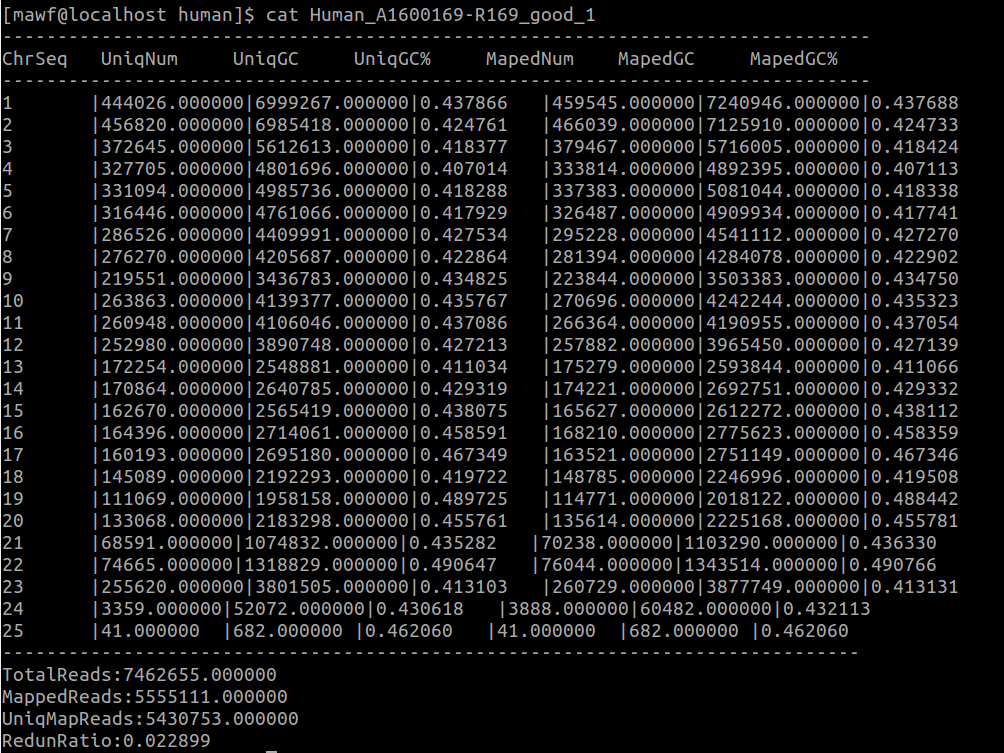
第三行：以‘+’开头，后面是序列标示符、描述信息，或者什么也不加；

第四行：是reads质量信息，和第二行的序列相对应。

* + 1. 运行和投放任务

在命令行运行主程序./hashtb\_pth即可，它可以实时监控下机数据目录，一旦有数据出现，能够实时处理。

4.2.4比对分析结果：



1. 相关文件：

**无。**

1. 相关记录：

无。