### LEB课程内容总结 沈南奇2300934024

#### EMBOSS

##### EMBOSS软件安装

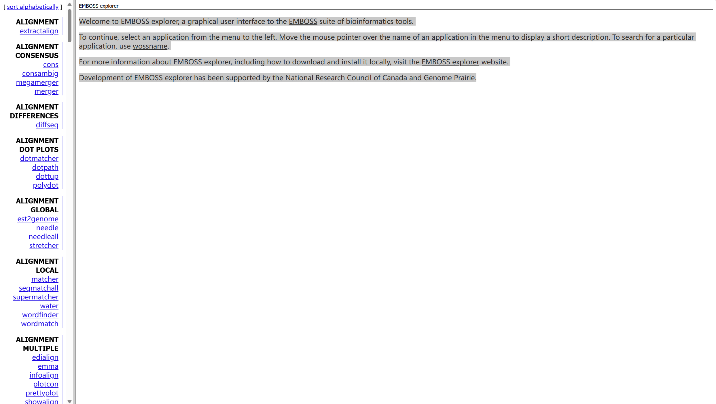
EMBOSS软件包主网站  
<http://emboss.open-bio.org/>

进入网站没有直接找到下载方式

同组廖一岩分享了安装包，在运行该exe文件前提示需要先安装java <https://www.java.com/zh-CN/download/manual.jsp>

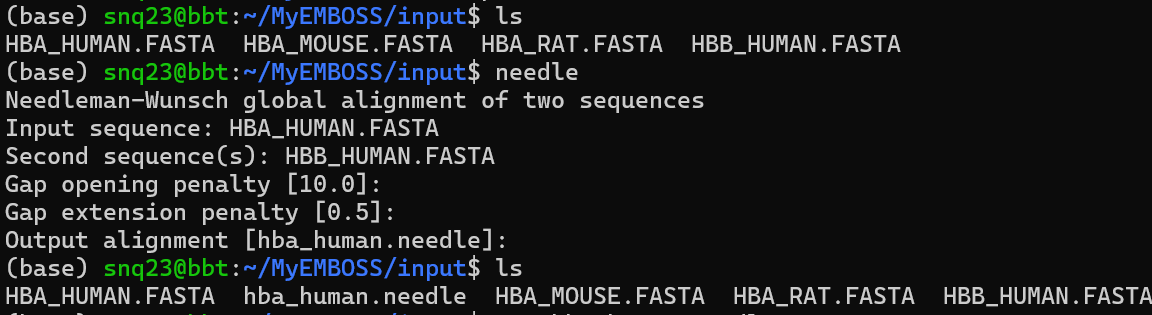
##### 另一种EMBOSS使用方式

荷兰瓦格宁根大学EMBOSS Explorer  
<http://www.bioinformatics.nl/emboss-explorer/>

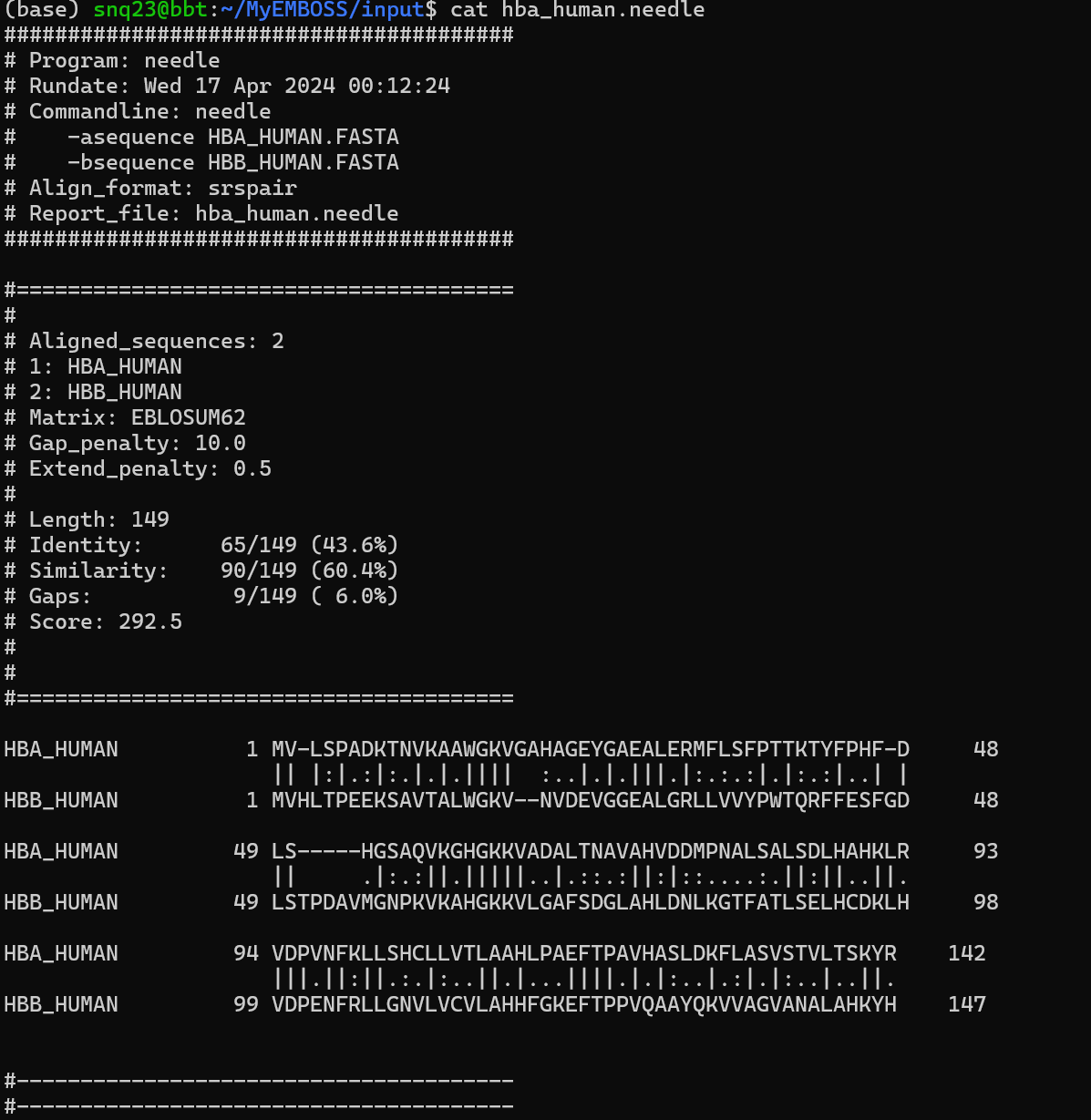
通过浏览器访问该网站，打开后页面右侧可以选择工具打开使用

###### 运行needle程序

###### 交互式



**操作流程：**

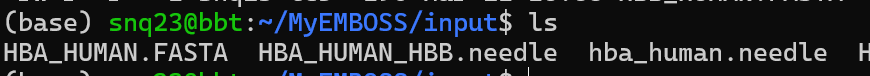
1. 准备需要比对的两个序列，以人血红蛋白HBA和HBB亚基的氨基酸序列为例。
2. 输入needle开始运行程序。
3. 输入第一个研究序列FASTA格式文件的文件名。
4. 输入第二个文件名。
5. 系统给出起始空位罚分值Gap opening penalty，默认10.0，可输入自定分数后回车，若直接回车则取用默认值。
6. 系统给出延伸空位罚分值Gap extension penalty，默认0.5，输入自定分数后回车，若直接回车则取用默认值。
7. 系统提示输出文件名，默认小写，回车确定，也可自行输入。
8. ****当前目录下多出hba\_human.needle文件，可cat命令查看

###### 参数式

**操作：**

输入needle HBA\_HUMAN.FASTA HBB\_HUMAN.FASTA -gapo 10.0 -gape 0.5 -out HBA\_HUMAN\_HBB.needle

**备注：**

红字部分为要比对的序列文件，-gapo 10.0是起始空位罚分，-gape 0.5是延伸空位罚分，-out HBA\_HUMAN\_HBB.needle是输出名为HBA\_HUMAN\_HBB.needle的文件（下图第二个）

##### 运行dottup点阵图程序

###### 点阵图程序

知乎搜到资料：

dottup：一个word要在两个序列里面完全一模一样才有dot

dotmatcher：不完全匹配，只要在threshold以上，就有dot

dotpath：类似dottup（但是默认的word size比较小），但是序列里面的任意一段只能match到一个地方（不能对到两个不同的地方）。换句话说，没有overlap的[alignment](https://www.zhihu.com/search?q=alignment&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A41580497%7D" \t "_blank)，也就是从图上看，任意x值上都只有一个dot，任意y值上也只有一个dot。

polydot：一个[fasta文件](https://www.zhihu.com/search?q=fasta%E6%96%87%E4%BB%B6&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A41580497%7D)，里面所有序列对（all pairwise）的dottup。比如有3条序列就有3\*3个dottup  
作者：Philip Yang  
链接：<https://www.zhihu.com/question/28623438/answer/41580497>

点阵图程序dottup多用于核酸序列，而dotmatcher则可用于蛋白质序列（EMBOSS软件包序列分析程序应用实例，罗静初，DOI：10.12113/202008002，生物信息学）

###### 交互式

**操作流程**：

1.准备HBA\_HUMAN.FASTA和HBB\_HUMAN.FASTA文件

2.输入dottup运行程序

3.输入第一步准备文件的文件名

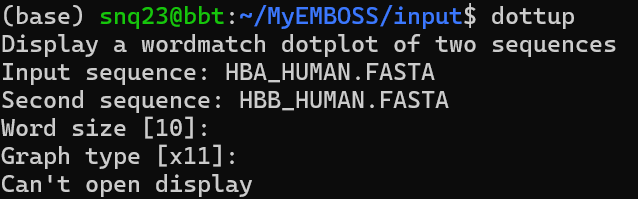
4.修改字长或选取默认值10

5.输出结果图形格式默认X11，输入svg（[SVG：可缩放矢量图形](https://developer.mozilla.org/zh-CN/docs/Web/SVG)）

6.生成dottup.svg文件

###### 参数式

操作：输入dottup HBA\_HUMAN.FASTA HBB\_HUMAN.FASTA -word 3 -graph svg -goutfile HBA-HBB

备注：输入dottup，两个文件名，设置字长（-word）和输出结果图形格式（-graph），修改输出结果文件名（-goutfile）。默认输出图形格式是x11，但是默认格式生成不成功，需修改成svg格式

#### GitHub

##### GitHub重要性

软件开发的核心工具

项目管理的优选平台

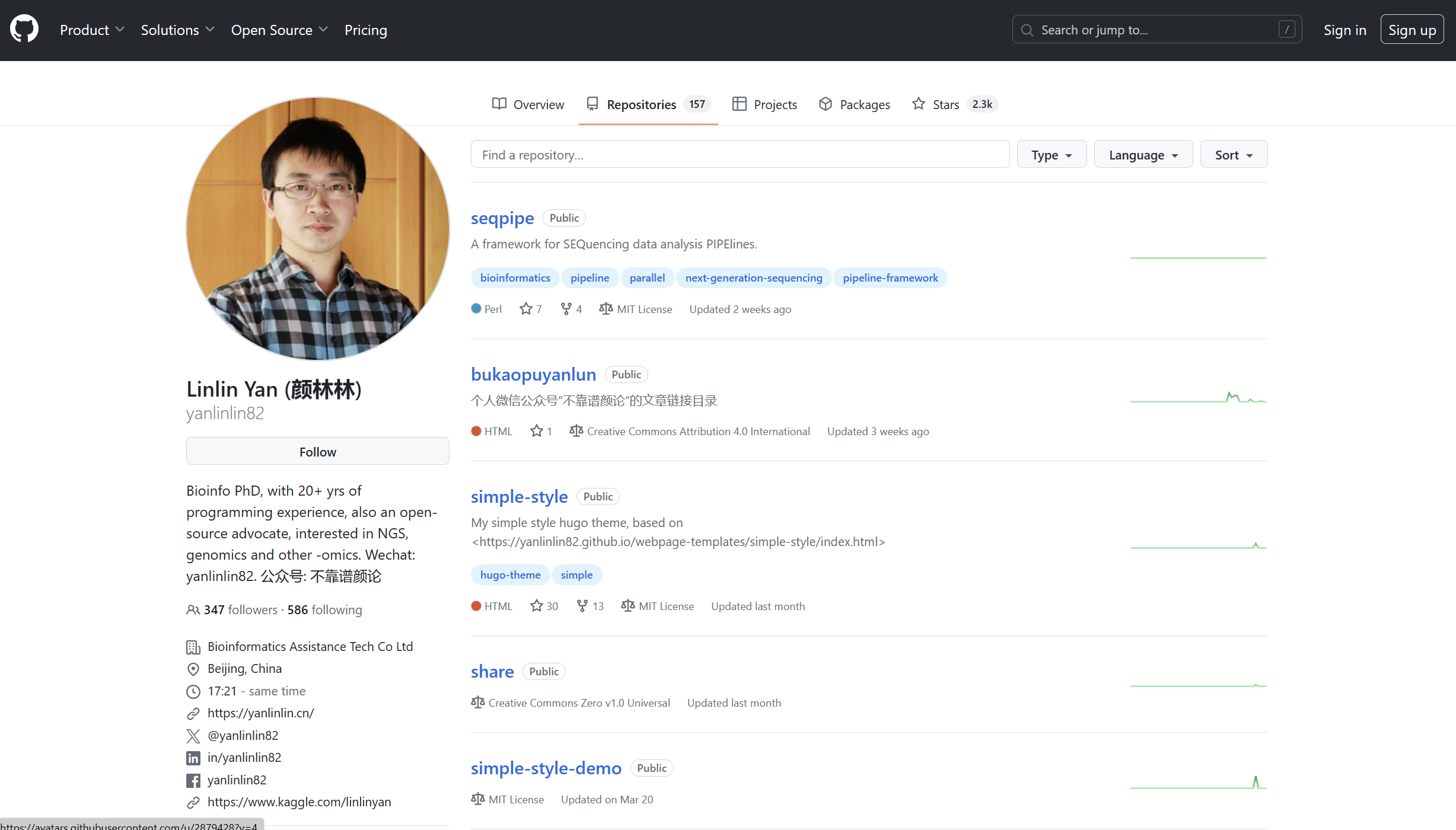
开源协作的中心舞台

GitHub 是一个存放软件代码的网站，将代码放在网站上，其他人可以在现有代码的基础上进行二次开发，不用从头开始工作。如果是团队合作，团队成员可以上传自己的工作，也可以拉取他人的。

##### 基本概念

**仓库 & 分支：**

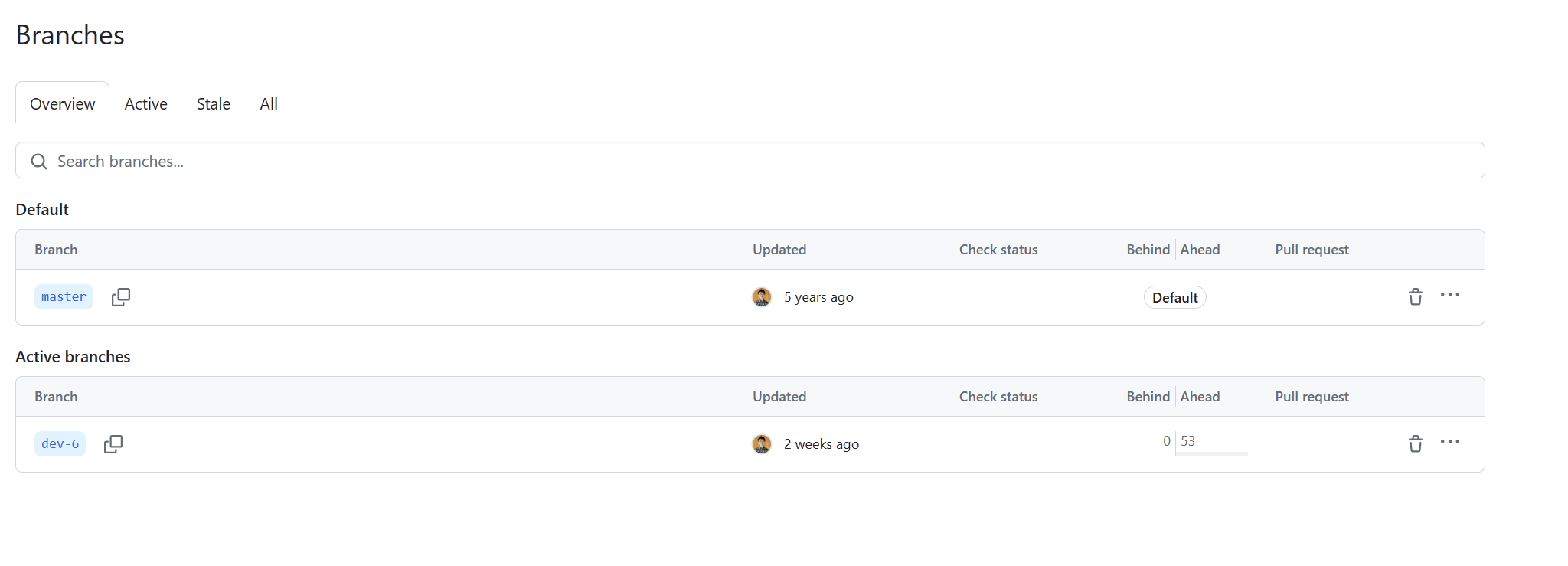
仓库(Repository): 项目的存储地点



分支(Branch): 不同版本的开发线路

主分支: 主开发线

特性分支: 新功能开发



**提交 & 拉取请求：**

提交(Commit):

保存更改的快照

提交历史: 更改的记录

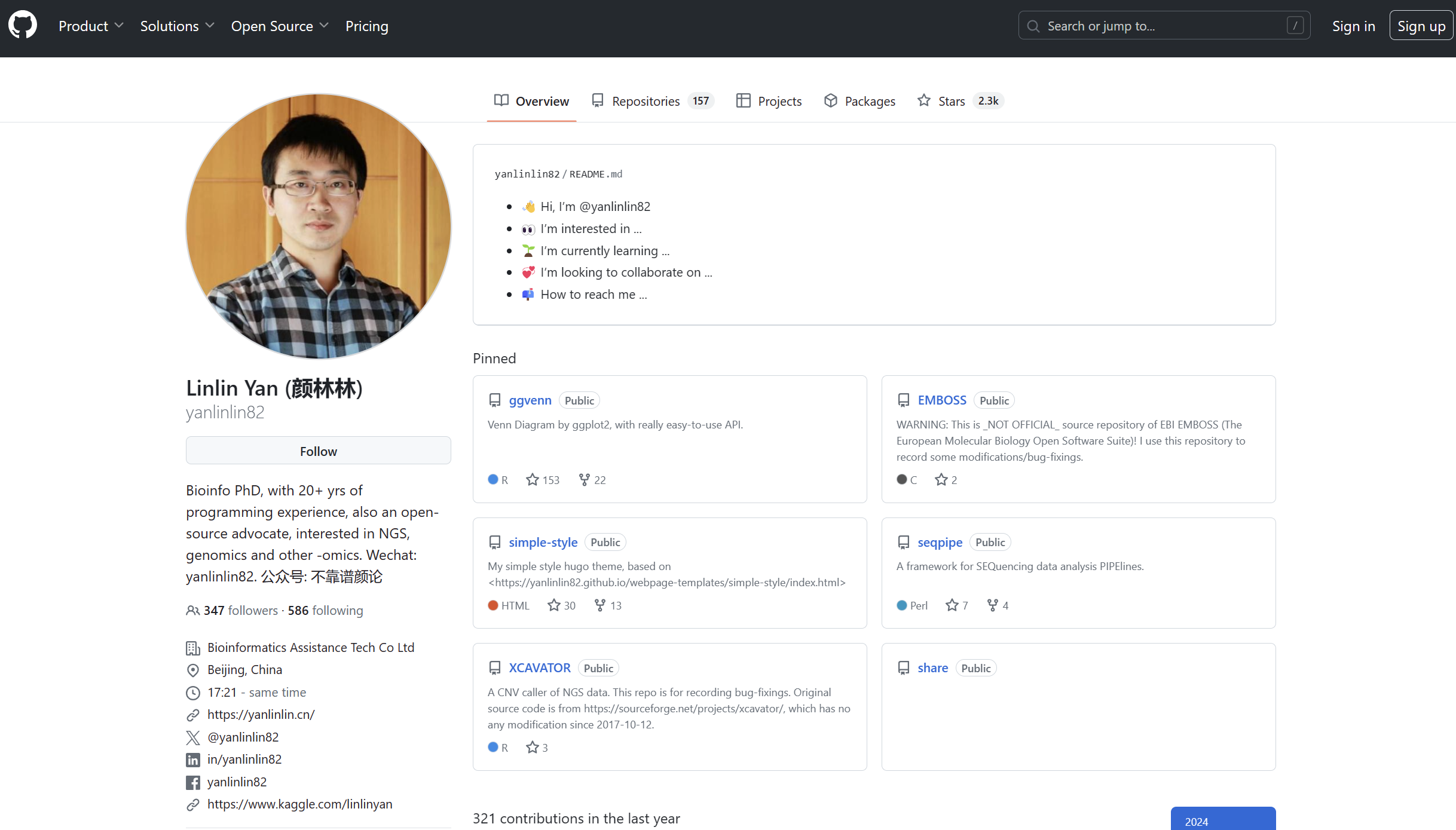
拉取请求(Pull Request):

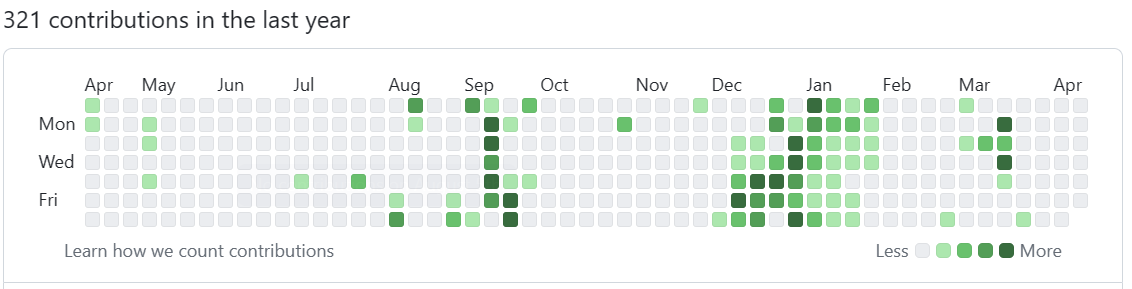
合并分支的请求

代码审查: 团队成员评审更改

合并: 将更改并入主分支

##### 主页overview





上图展示在github的贡献和活跃度

#### Git

##### 参考资料

[Git - Book (git-scm.com)](https://git-scm.com/book/zh/v2)

[Learn Git Branching](https://learngitbranching.js.org/?locale=zh_CN)小游戏

##### 配置Git

服务器中有git，省略安装步骤，安装可以参照[Git - 安装 Git (git-scm.com)](https://git-scm.com/book/zh/v2/%E8%B5%B7%E6%AD%A5-%E5%AE%89%E8%A3%85-Git)

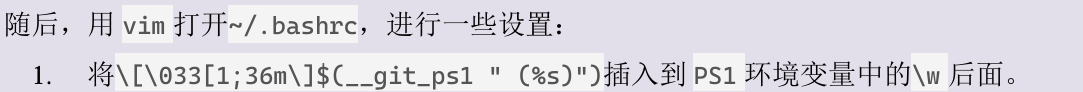
###### 确认安装完成：

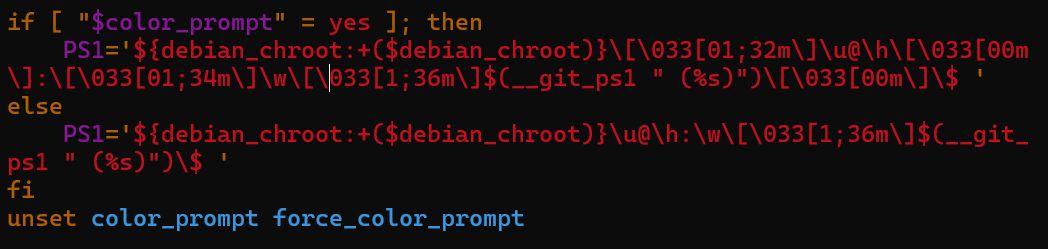
###### 配置用户邮箱和用户名



###### 下载并设置一些Git的辅助插件



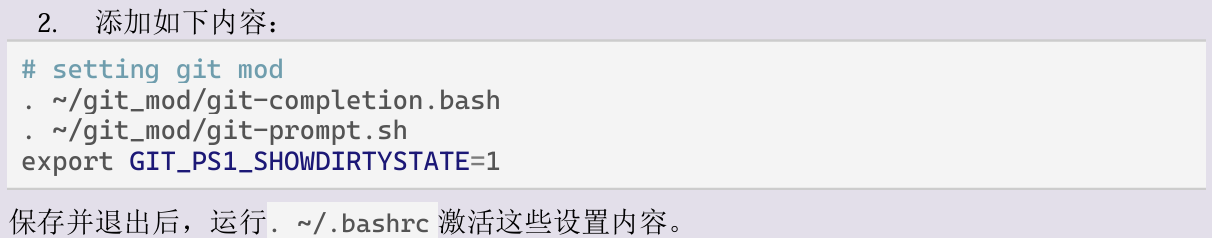




一共添加在三处

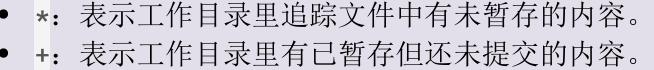
第一次进行完修改.bashrc的操作后，无法在进行任何操作，退出之后出现登录不上服务器的问题，助教认为可能是有操作失误影响了文件中的其他部分，用未修改的文件替换我家目录下文件后，可以正常登陆以及使用。所以修改.bashrc的操作还是比较“危险”的，之后助教又提供了官方教程的方法供学习。

<https://git-scm.com/book/zh/v2/%E9%99%84%E5%BD%95-A%3A-%E5%9C%A8%E5%85%B6%E5%AE%83%E7%8E%AF%E5%A2%83%E4%B8%AD%E4%BD%BF%E7%94%A8-Git-Bash-%E4%B8%AD%E7%9A%84-Git>



##### 建立一个使用Git进行版本控制的项目

###### 初始化



###### git status

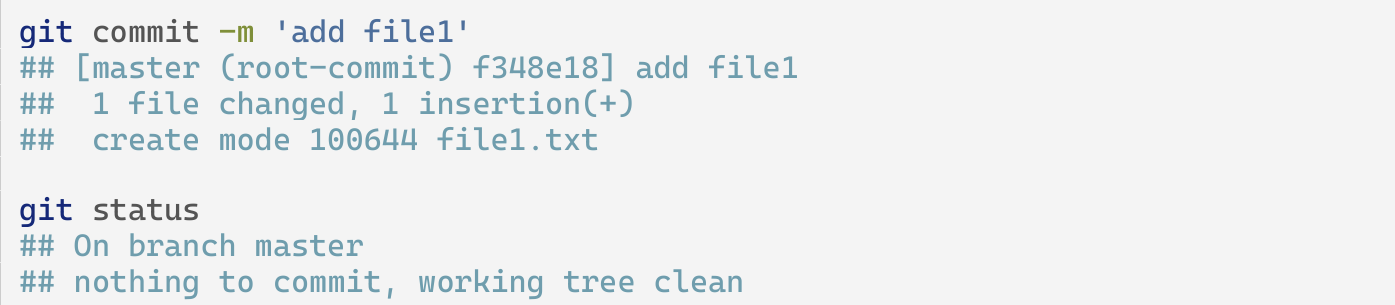
使用git status命令可以查看当前工作目录的状态，可以看到被识别出的文件是untracked的状态，是红字，表示还没有被Git所追踪。并且提示我们可以 使用git add对其进行追踪，以进行提交。

###### git status

对这个文件运行git add命令，Git会把它放到一个“暂存区”，git status检查，变成绿字了

###### git commit

这个命令表示将所有被暂存的改动储存起来，变成一次提交给记录下来。



###### git log

通过git log命令查看提交的记录



如想撤销修改，可以回溯至之前的提交

#### Anaconda

##### 使用Conda

###### 安装

服务器上有下载好的Conda存放于： /rd1/home/public/RNA-Seq/software/Miniconda3-latest-Linux-x86\_64.sh

cp /rd1/home/public/RNA-Seq/software/Miniconda3-latest-Linux-x86\_64.sh ~/

cd

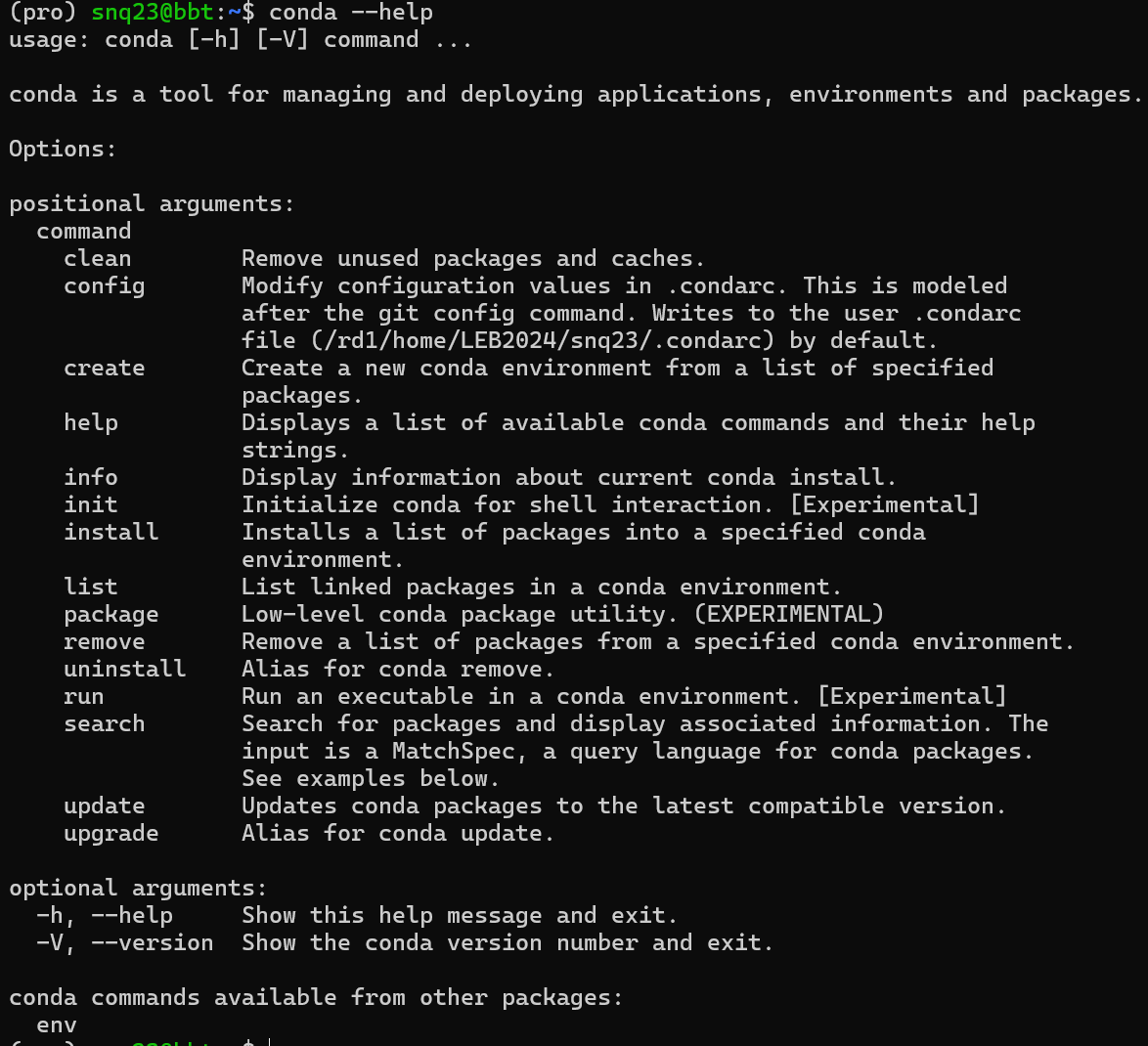
bash Miniconda3-latest-Linux-x86\_64.sh

一直按回车键，直到出现：

输入yes再回车

重新登陆服务器后左下角会出现base则表示安装完成

###### conda --help



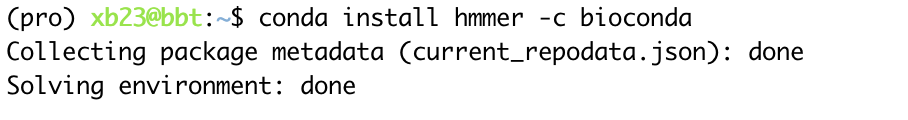
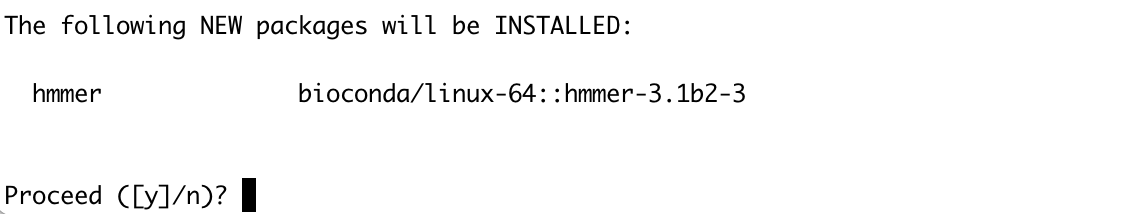
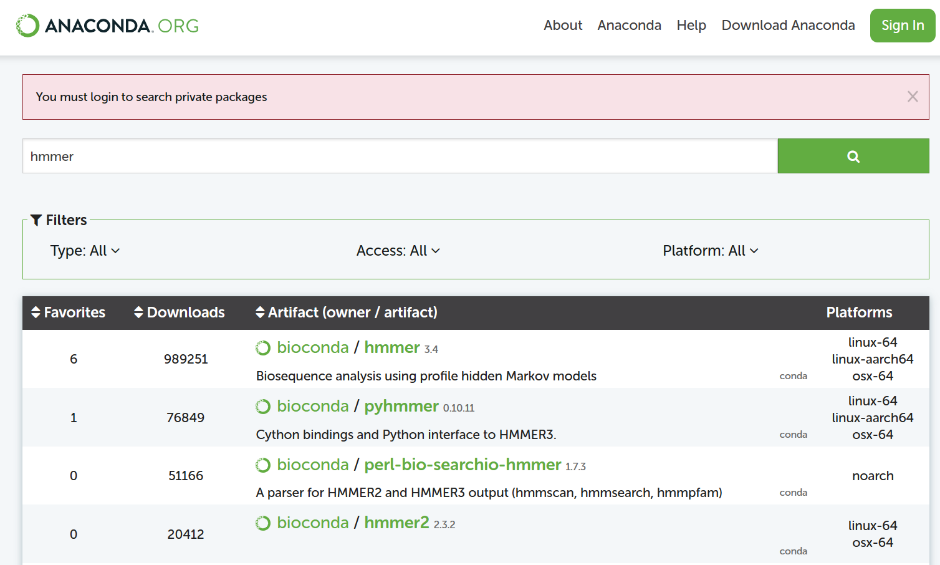
###### Conda镜像站

[anaconda | 镜像站使用帮助 | 清华大学开源软件镜像站 | Tsinghua Open Source Mirror](https://mirror.tuna.tsinghua.edu.cn/help/anaconda/?eqid=a38a4a4800068d0600000003645b7087) 

###### 创建激活新的环境

conda env remove删除名为pro的环境

###### 安装hmmer工具包

[:: Anaconda.org](https://anaconda.org/)可以搜索需要安装的包是否存在

conda remove命令用来删除已经安装的工具

##### mamba

[Welcome to Mamba’s documentation! — documentation](https://mamba.readthedocs.io/en/latest/index.html)

mamba是一个C++语言实现的conda，具有完全一样的功能且速度是conda的几十倍！

使用方式与conda完全一样，只需要把命令中的conda改为micromamba

建议直接一开始直接使用mamba