

生物信息学软件的安装与维护

人生苦短，我用 conda

伊现富 (Yi Xianfu)

2017-11-18



图 1: <https://github.com/Yixf-Education/demo4bx>

- 其他操作系统：Windows；Mac OS X
- Linux 操作系统
 - 服务器 vs. 本地
 - 单系统 vs. 双系统
 - U 盘系统： **MultiBootUSB**

MultiBootUSB is a cross platform software written in python which allows you to install multiple live linux on a USB disk non destructively and option to uninstall distros.

生活 生信的无奈

- Linux: `http`, `lftp`, `aria2c`, `dos2unix`, `git`
 - `apt`; `wget`, `./configure`, `make`, `make install`
- Perl: `Getopt::Long`, `Data::Dumper`, `List::Util`
 - `cpan`
- R: `tidyverse`, `data.table`, `knitr`
 - `install.package()`
- Bioconductor: `ComplexHeatmap`, `clusterProfiler`, `ggtree`
 - `source("..."); biocLite()`
- Python: `NumPy`, `SciPy`, `Pandas`, `Matplotlib`, `SciKit-Learn`
 - `pip`
- Bioinformatics: `BWA`, `Bowtie2`, `bedtools`, `samtools`
 - `wget`, `make install`; `git clone`

通用包管理系统——conda

Conda 是开源的包管理系统和环境管理系统，可以安装软件包的多个版本和依赖，而且各环境之间可以很方便得切换。Conda 支持 Linux, Mac OS X 和 Windows 系统。Conda 主要为 Python 程序所创建，但是可以打包和分布任意软件。Conda 有多个版本，包括 Anaconda, Anaconda Server 和 Miniconda。

Package, dependency and environment management for any language—Python, R, Ruby, Lua, Scala, Java, JavaScript, C/ C++, FORTRAN.

Conda is an open source package management system and environment management system that runs on Windows, macOS and Linux. Conda quickly installs, runs and updates packages and their dependencies. Conda easily creates, saves, loads and switches between environments on your local computer. It was created for Python programs, but it can package and distribute software for any language.

抉择——X vs. Y

- Anaconda vs. Miniconda
 - Anaconda: 重量级, 适合使用 Python 进行数据分析的用户
 - **Miniconda**: 轻量级, 适合主要使用 conda 功能的用户
 - 按需选择, 无关紧要
- Python 2.7 vs. 3.6
 - Python 2.7: 成熟
 - **Python 3.6**: 最新
 - 至关重要 无关紧要

conda: 安装与配置

❶ 下载 Miniconda(Python 3.6)

```
wget -c https://URLtoMiniconda/Miniconda3-latest-Linux-x86_64.sh
```

❷ 安装 Miniconda

```
bash Miniconda3-latest-Linux-x86_64.sh
```

❸ 配置 Miniconda: 修改 .bashrc

```
export PATH="PathToMiniconda/bin:$PATH"
```

❹ 配置 .bashrc: 修改 .profile

```
source $HOME/.bashrc
```

bioconda: 安装

❶ 添加 channels

```
(conda config --add channels r)
conda config --add channels defaults
conda config --add channels conda-forge
conda config --add channels bioconda
```

❷ 查看是否生效

```
conda info
```

❸ 注意事项

- 添加顺序很重要
- r channel 是为了向后兼容
- 本质是修改配置文件: \$HOME/.condarc

国内镜像——TUNA（清华）

- TUNA="https://mirrors.tuna.tsinghua.edu.cn"
- Anaconda

```
conda config --add channels TUNA/anaconda/pkgs/main/  
conda config --add channels TUNA/anaconda/pkgs/free/  
conda config --set show_channel_urls yes
```

- Conda Forge

```
conda config --add channels TUNA/anaconda/cloud/conda-forge/
```

- R & bioconda

```
conda config --add channels TUNA/anaconda/pkgs/r/  
conda config --add channels TUNA/anaconda/cloud/bioconda/
```

国内镜像——USTC (中科大)

- USTC="https://mirrors.ustc.edu.cn"
- Anaconda

```
conda config --add channels USTC/anaconda/pkgs/main/  
conda config --add channels USTC/anaconda/pkgs/free/
```

- Conda Forge

```
conda config --add channels USTC/anaconda/cloud/conda-forge/
```

- R & bioconda

```
conda config --add channels USTC/anaconda/pkgs/r/  
conda config --add channels USTC/anaconda/cloud/conda-forge/
```

conda: 基本使用

- 安装软件

```
conda install bwa
```

- 查找软件

```
conda search igraph
```

- 删除软件

```
conda remove ucsc-liftover
```

conda: 基本使用 (续)

- 罗列软件

```
conda list
```

- 软件升级

```
conda update conda  
conda update r-essentials  
conda update --all
```

- 帮助信息

```
conda --help  
conda -h  
conda create -h
```

conda: 高级应用

- 安装特定版本的软件
 - |: OR
 - ,: AND
 - *: 等同于正则表达式中的.*
 - < > <= >= == !=: 针对版本号的运算

```
conda install bedtools=2.20 # 2.20*  
conda install bedtools==2.20.1 # 2.20.1
```

conda: 高级应用 (续)

- 使用环境/容器
 - 创建环境 (同时安装软件)

```
conda create --name py2
conda create --name py2 python=2.7
```

- 查看环境

```
conda info --envs
```

- 使用/切换环境

```
conda install --name py2 tophat
source activate py2
conda install tophat
conda install/remove/... ...
source deactivate py2
```

基于 conda 的生信软件管理

- 策略一：为特定工作创建“单元”环境

```
conda create --name aligners
source activate aligners
conda install bwa bowtie bowtie2 hisat2
source deactivate aligners
```

- 策略二：为整个流程创建“流水线”环境

```
conda create --name tuxedo
source activate tuxedo
conda install samtools bowtie2 tophat cufflinks bioconductor-cumtuxedo
source deactivate tuxedo
```

特殊情况

如果需要的工具不在 conda 中.....

- Perl, R, Python (sudo)

```
cpan  
install.package()  
pip
```

- Others
 - ① 按要求进行安装
 - ② 链接至 ~~CONDA/bin~~
 - ③ 照常使用

参考资料

- Conda 使用手册
- bioconda
- Anaconda 镜像使用帮助