

RNA-seq生信分析环境搭建

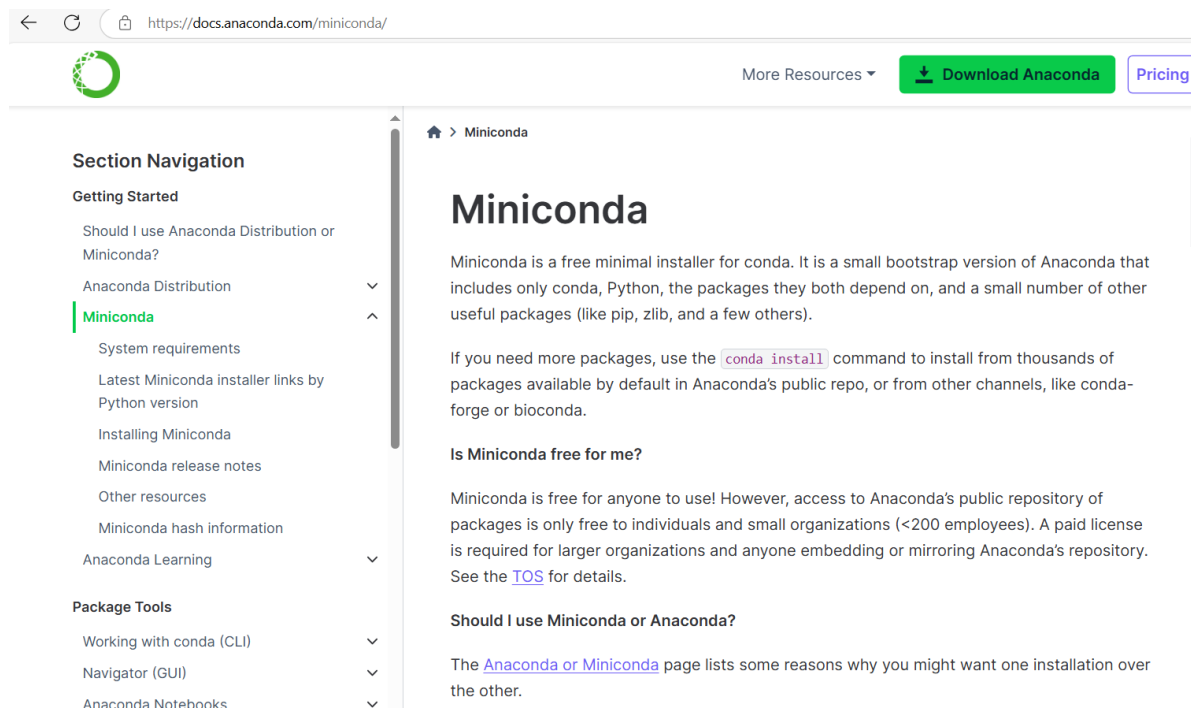
陈明杰

202410

搭建分析环境

- conda虚拟环境
- 质控软件: fastqc
- 去接头软件: fastp
- 比对软件: hisat2、STAR
- 其他软件: samtools、sambamba
- R及常见包: ggplot2, pheatmap, BiocManager, DESeq2, clusterProfiler
- Python及相关包: matplotlib, numpy, scipy, pandas

conda环境安装（下载安装包）



conda是一个包和环境管理工具，用于创建、管理和切换Python的虚拟环境。隔离环境，避免环境冲突。

Anaconda和Miniconda是conda的发行版，前者包含大量预装包，适合初学者，后者精简，仅包含基本组件。

安装Miniconda后，可通过命令行创建、列出、激活、删除环境，以及安装和卸载包。对于开发者，conda提供了一种高效管理Python环境和包的方式

https://repo.anaconda.com/miniconda/Miniconda3-latest-Linux-x86_64.sh

约145M

```
vboxuser@ubuntu: /mnt/test
vboxuser@ubuntu:/mnt/test$ ls
Miniconda3-latest-Linux-x86_64.sh
vboxuser@ubuntu:/mnt/test$ sh Miniconda3-latest-Linux-x86_64.sh

Welcome to Miniconda3 py312_24.7.1-0

In order to continue the installation process, please review the license
agreement.
Please, press ENTER to continue
>>>
```

```
vboxuser@ubuntu: /mnt/test
ANACONDA TERMS OF SERVICE

Please read these Terms of Service carefully before purchasing, using, accessing
, or downloading any Anaconda Offerings (the "Offerings"). These Anaconda Terms
of Service ("TOS") are between Anaconda, Inc. ("Anaconda") and you ("You"), the
individual or entity acquiring and/or providing access to the Offerings. These T
OS govern Your access, download, installation, or use of the Anaconda Offerings,
which are provided to You in combination with the terms set forth in the applica
ble Offering Description, and are hereby incorporated into these TOS. Except wh
ere indicated otherwise, references to "You" shall include Your Users. You herab
y acknowledge that these TOS are binding, and You affirm and signify your consen
t to these TOS by registering to, using, installing, downloading, or accessing t
he Anaconda Offerings effective as of the date of first registration, use, insta
ll, download or access, as applicable (the "Effective Date"). Capitalized defini
tions not otherwise defined herein are set forth in Section 15 (Definitions). If
You do not agree to these Terms of Service, You must not register, use, install
, download, or access the Anaconda Offerings.

1. ACCESS & USE
1.1 General License Grant. Subject to compliance with these TOS and any applicab
le Offering Description, Anaconda grants You a personal, non-exclusive, non-tran
sferable, non-sublicensable, revocable, limited right to use the applicable Anac
onda Offering strictly as detailed herein and as set forth in a relevant Offerin
--More--
```

```
vboxuser@ubuntu: /mnt/test
es in source (if provided to You by Anaconda as source) and binary forms, with o
r without modification subject to the requirements set forth below.
4. Updates. Anaconda may, at its option, make available patches, workarounds or
other updates to Miniconda(R). Unless the updates are provided with their separa
te governing terms, they are deemed part of Miniconda(R) licensed to You as prov
ided in this Agreement.
5. Support. This Agreement does not entitle You to any support for Miniconda(R).
6. Redistribution. Redistribution and use in source and binary forms, with or wi
thout modification, are permitted provided that the following conditions are met
: (1) Redistributions of source code must retain the above copyright notice, thi
s list of conditions and the following disclaimer; (2) Redistributions in binary
form must reproduce the above copyright notice, this list of conditions and the
following disclaimer in the documentation and/or other materials provided with
the distribution.
7. Intellectual Property Notice. You acknowledge that, as between You and Anacon
da, Anaconda owns all right, title, and interest, including all intellectual pro
perty rights, in and to Miniconda(R) and, with respect to third-party products d
istributed with or through Miniconda(R), the applicable third-party licensors ow
n all right, title and interest, including all intellectual property rights, in
and to such products.

Do you accept the license terms? [yes|no]
>>> yes
```

```
vboxuser@ubuntu: /mnt/test
: (1) Redistributions of source code must retain the above copyright notice, thi
s list of conditions and the following disclaimer; (2) Redistributions in binary
form must reproduce the above copyright notice, this list of conditions and the
following disclaimer in the documentation and/or other materials provided with
the distribution.
7. Intellectual Property Notice. You acknowledge that, as between You and Anacon
da, Anaconda owns all right, title, and interest, including all intellectual pro
perty rights, in and to Miniconda(R) and, with respect to third-party products d
istributed with or through Miniconda(R), the applicable third-party licensors ow
n all right, title and interest, including all intellectual property rights, in
and to such products.

Do you accept the license terms? [yes|no]
>>> yes

Miniconda3 will now be installed into this location:
/home/vboxuser/miniconda3

- Press ENTER to confirm the location
- Press CTRL-C to abort the installation
- Or specify a different location below

[/home/vboxuser/miniconda3] >>>
```

```
vboxuser@ubuntu: /mnt/test

wheel          pkgs/main/linux-64::wheel-0.43.0-py312h06a4308_0
xz             pkgs/main/linux-64::xz-5.4.6-h5eee18b_1
yaml-cpp       pkgs/main/linux-64::yaml-cpp-0.8.0-h6a678d5_1
zlib           pkgs/main/linux-64::zlib-1.2.13-h5eee18b_1
zstandard      pkgs/main/linux-64::zstandard-0.22.0-py312h2c38b39_0
zstd           pkgs/main/linux-64::zstd-1.5.5-hc292b87_2

Downloading and Extracting Packages:

Preparing transaction: done
Executing transaction: done
installation finished.
Do you wish to update your shell profile to automatically initialize conda?
This will activate conda on startup and change the command prompt when activated
.
If you'd prefer that conda's base environment not be activated on startup,
  run the following command when conda is activated:

conda config --set auto_activate_base false

You can undo this by running `conda init --reverse $SHELL`? [yes|no]
[no] >>> 
```

```
vboxuser@ubuntu: /mnt/test

run the following command when conda is activated:

conda config --set auto_activate_base false

You can undo this by running `conda init --reverse $SHELL`? [yes|no]
[no] >>> yes
no change      /home/vboxuser/miniconda3/condabin/conda
no change      /home/vboxuser/miniconda3/bin/conda
no change      /home/vboxuser/miniconda3/bin/conda-env
no change      /home/vboxuser/miniconda3/bin/activate
no change      /home/vboxuser/miniconda3/bin/deactivate
no change      /home/vboxuser/miniconda3/etc/profile.d/conda.sh
no change      /home/vboxuser/miniconda3/etc/fish/conf.d/conda.fish
no change      /home/vboxuser/miniconda3/shell/condabin/Conda.psm1
no change      /home/vboxuser/miniconda3/shell/condabin/conda-hook.ps1
no change      /home/vboxuser/miniconda3/lib/python3.12/site-packages/xontrib/conda.xsh
no change      /home/vboxuser/miniconda3/etc/profile.d/conda.csh
modified       /home/vboxuser/.bashrc

==> For changes to take effect, close and re-open your current shell. <==

Thank you for installing Miniconda3!
vboxuser@ubuntu: /mnt/test$ 
```

Yes, 直接配置好了

安装好后

```
vboxuser@ubuntu: ~/Desktop
(base) vboxuser@ubuntu:~/Desktop$ conda deactivate 退出conda环境
vboxuser@ubuntu:~/Desktop$ conda activate 激活conda环境
(base) vboxuser@ubuntu:~/Desktop$ R R没有装
Command 'R' not found, but can be installed with:
sudo apt install r-base-core
(base) vboxuser@ubuntu:~/Desktop$ python Python已装
Python 3.12.4 | packaged by Anaconda, Inc. | (main, Jun 18 2024, 15:12:24) [GCC
11.2.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> 
```

创建虚拟环境

conda create -n jimmy37 python=3.7

注意：装的主版本相同，子版本不一定相同
例如上周我装的是3.10.14，这周就变成了3.10.15

```
vboxuser@ubuntu: ~/Desktop
(base) vboxuser@ubuntu:~/Desktop$ conda create --name jimmy3 python=3.10
Channels:
 - defaults
Platform: linux-64
Collecting package metadata (repodata.json): |
```

```
vboxuser@ubuntu: ~/Desktop

libstdc++-ng      pkgs/main/linux-64::libstdc++-ng-11.2.0-h1234567_1
libuuid           pkgs/main/linux-64::libuuid-1.41.5-h5eee18b_0
ncurses           pkgs/main/linux-64::ncurses-6.4-h6a678d5_0
openssl           pkgs/main/linux-64::openssl-3.0.15-h5eee18b_0
pip               pkgs/main/linux-64::pip-24.2-py310h06a4308_0
python            pkgs/main/linux-64::python-3.10.14-h955ad1f_1
readline          pkgs/main/linux-64::readline-8.2-h5eee18b_0
setuptools        pkgs/main/linux-64::setuptools-75.1.0-py310h06a4308_0
sqlite            pkgs/main/linux-64::sqlite-3.45.3-h5eee18b_0
tk                pkgs/main/linux-64::tk-8.6.14-h39e8969_0
tzdata            pkgs/main/noarch::tzdata-2024a-h04d1e81_0
wheel             pkgs/main/linux-64::wheel-0.44.0-py310h06a4308_0
xz                pkgs/main/linux-64::xz-5.4.6-h5eee18b_1
zlib              pkgs/main/linux-64::zlib-1.2.13-h5eee18b_1

Proceed ([y]/n)?

Downloading and Extracting Packages:
python-3.10.14      | 26.8 MB | #####9 | 97%
openssl-3.0.15     | 5.2 MB | ##### | 100%
pip-24.2            | 2.3 MB | ##### | 100%
setuptools-75.1.0  | 1.7 MB | ##### | 100%
wheel-0.44.0        | 109 KB | ##### | 100%
```

删除环境：conda remove -n jimmy3 --all

退出环境：conda deactivate

安装fastqc – 质控

conda install bioconda::fastqc

fastqc 直接弹出软件窗口
fastqc -h 查看帮助

```
vboxuser@ubuntu: ~/Desktop

Preparing transaction: done
Verifying transaction: done
Executing transaction: done
(jimmy3) vboxuser@ubuntu:~/Desktop$ fastqc
(jimmy3) vboxuser@ubuntu:~/Desktop$ fastqc -h

FastQC - A high throughput sequence QC analysis tool

SYNOPSIS

fastqc seqfile1 seqfile2 .. seqfileN

fastqc [-o output dir] [--(no)extract] [-f fastq|bam|sam]
[-c contaminant file] seqfile1 .. seqfileN

DESCRIPTION

FastQC reads a set of sequence files and produces from each one a quality
control report consisting of a number of different modules, each one of
which will help to identify a different potential type of problem in your
data.
```

```
vboxuser@ubuntu: ~/Desktop

-?, --help                print this message

(jimmy3) vboxuser@ubuntu:~/Desktop$
(jimmy3) vboxuser@ubuntu:~/Desktop$ fastqc -h

FastQC - A high throughput sequence QC analysis tool

SYNOPSIS

fastqc seqfile1 seqfile2 .. seqfileN

fastqc [-o output dir] [--(no)extract] [-f fastq|bam|sam]
[-c contaminant file] seqfile1 .. seqfileN

DESCRIPTION

FastQC reads a set of sequence files and produces from each one a quality
control report consisting of a number of different modules, each one of
which will help to identify a different potential type of problem in your
data.

If no files to process are specified on the command line then the program
will start as an interactive graphical application. If files are provided
on the command line then the program will run with no user interaction
```

<https://anaconda.org/bioconda/fastqc>

安装fastp – 去接头

conda install bioconda::fastp

```
vboxuser@ubuntu: ~/Desktop
(base) vboxuser@ubuntu:~/Desktop$ conda activate jimmy3
(jimmy3) vboxuser@ubuntu:~/Desktop$ conda install bioconda::fastp

Channels:
- defaults
- bioconda
Platform: linux-64
Collecting package metadata (repodata.json): |
```

fastp -h 查看帮助

```
vboxuser@ubuntu: ~/Desktop
--add-chrname      add 'chr' to reference names in alignment
--version          print version information and quit
-h/--help          print this usage message
(ERR): hisat2-align exited with value 1
(jimmy3) vboxuser@ubuntu:~/Desktop$ fastp
fastp: an ultra-fast all-in-one FASTQ preprocessor
version 0.23.2
usage: fastp [options] ...
options:
  -i, --in1          read1 input file name (string [=])
  -o, --out1          read1 output file name (string [=])
  -I, --in2          read2 input file name (string [=])
  -O, --out2          read2 output file name (string [=])
  --unpaired1         for PE input, if read1 passed QC but read
2 not, it will be written to unpaired1. Default is to discard it. (string [=])
  --unpaired2         for PE input, if read2 passed QC but read
1 not, it will be written to unpaired2. If --unpaired2 is same as --unpaired1 (d
efault mode), both unpaired reads will be written to this same file. (string [=]
)
  --overlapped_out    for each read pair, output the overlapped
region if it has no any mismatched base. (string [=])
  --failed_out        specify the file to store reads that cann
ot pass the filters. (string [=])
  -m, --merge         for paired-end input, merge each pair of
```

<https://anaconda.org/bioconda/fastp>

安装hisat2 – 比对

conda install bioconda::hisat2

```
vboxuser@ubuntu: ~/Desktop
(jimmy3) vboxuser@ubuntu:~/Desktop$ conda install bioconda::hisat2
Channels:
  - defaults
  - bioconda
Platform: linux-64
Collecting package metadata (repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: /home/vboxuser/miniconda3/envs/jimmy3

  added / updated specs:
    - bioconda::hisat2

The following packages will be downloaded:



| package      | build      | size    | channel  |
|--------------|------------|---------|----------|
| hisat2-2.2.1 | h87f3376_4 | 16.5 MB | bioconda |
| Total:       |            | 16.5 MB |          |


```

hisat2 -h 查看帮助

```
vboxuser@ubuntu: ~/Desktop
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
(jimmy3) vboxuser@ubuntu:~/Desktop$ hisat2
No index, query, or output file specified!
HISAT2 version 2.2.1 by Daehwan Kim (infphilo@gmail.com, www.ccb.jhu.edu/people/infphilo)
Usage:
  hisat2 [options]* -x <ht2-idx> {-1 <m1> -2 <m2> | -U <r>} [-S <sam>]

<ht2-idx>  Index filename prefix (minus trailing .X.ht2).
<m1>      Files with #1 mates, paired with files in <m2>.
           Could be gzip'ed (extension: .gz) or bzip2'ed (extension: .bz2).
<m2>      Files with #2 mates, paired with files in <m1>.
           Could be gzip'ed (extension: .gz) or bzip2'ed (extension: .bz2).
<r>       Files with unpaired reads.
           Could be gzip'ed (extension: .gz) or bzip2'ed (extension: .bz2).
<sam>     File for SAM output (default: stdout)

<m1>, <m2>, <r> can be comma-separated lists (no whitespace) and can be
specified many times.  E.g. '-U file1.fq,file2.fq -U file3.fq'.

Options (defaults in parentheses):
```

<https://anaconda.org/bioconda/hisat2>

安装STAR – 比对

conda install bioconda::star

```
vboxuser@ubuntu: ~/Desktop
--version      What version of bedtools are you using?.
--contact      Feature requests, bugs, mailing lists, etc.

(jimmy3) vboxuser@ubuntu:~/Desktop$
(jimmy3) vboxuser@ubuntu:~/Desktop$
(jimmy3) vboxuser@ubuntu:~/Desktop$ conda install bioconda::star

Channels:
- defaults
- bioconda
Platform: linux-64
Collecting package metadata (repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: /home/vboxuser/miniconda3/envs/jimmy3

  added / updated specs:
    - bioconda::star

The following packages will be downloaded:
```

human比对至少35G内存

STAR 查看帮助

```
vboxuser@ubuntu: ~/Desktop

SOLUTION: use non-empty value for this parameter

Sep 20 12:56:32 ..... FATAL ERROR, exiting
(jimmy3) vboxuser@ubuntu:~/Desktop$ STAR -h
Usage: STAR [options]... --genomeDir /path/to/genome/index/  --readFilesIn R1.
fq R2.fq
Spliced Transcripts Alignment to a Reference (c) Alexander Dobin, 2009-2022

STAR version=2.7.10b
STAR compilation time,server,dir=2022-11-01T09:53:26-04:00 :/home/dobin/data/STA
R/STARcode/STAR.master/source
For more details see:
<https://github.com/alexdobin/STAR>
<https://github.com/alexdobin/STAR/blob/master/doc/STARmanual.pdf>
### versions
versionGenome          2.7.4a
      string: earliest genome index version compatible with this STAR release. Ple
ase do not change this value!

### Parameter Files
parametersFiles        -
      string: name of a user-defined parameters file, "-": none. Can only be defin
ed on the command line.
```

<https://anaconda.org/bioconda/star>

安装samtools – 比对格式转换

conda install bioconda::samtools

samtools 查看帮助

```
vboxuser@ubuntu: ~/Desktop
length is 7 if not specified.

-q --quiet      Suppress all progress messages on stdout and only report errors.

-d --dir        Selects a directory to be used for temporary files written when
                generating report images. Defaults to system temp directory if
                not specified.

BUGS

Any bugs in fastqc should be reported either to simon.andrews@babraham.ac.uk
or in www.bioinformatics.babraham.ac.uk/bugzilla/

(jimmy3) vboxuser@ubuntu:~/Desktop$
(jimmy3) vboxuser@ubuntu:~/Desktop$ conda install bioconda::samtools
Channels:
- defaults
- bioconda
Platform: linux-64
Collecting package metadata (repodata.json): /
```

```
vboxuser@ubuntu: ~/Desktop
(jimmy3) vboxuser@ubuntu:~/Desktop$ samtools

Program: samtools (Tools for alignments in the SAM format)
Version: 1.6 (using htslib 1.6)

Usage:  samtools <command> [options]

Commands:
-- Indexing
    dict      create a sequence dictionary file
    faidx     index/extract FASTA
    index     index alignment

-- Editing
    calmd     recalculate MD/NM tags and '=' bases
    fixmate   fix mate information
    reheader  replace BAM header
    rmdup     remove PCR duplicates
    targetcut cut fosmid regions (for fosmid pool only)
    addreplacerg adds or replaces RG tags
    markdup   mark duplicates

-- File operations
    collate   shuffle and group alignments by name
```

<https://anaconda.org/bioconda/samtools>

安装featureCounts – 定量

conda install bioconda::subread

featureCounts 查看帮助

```
vboxuser@ubuntu: ~/Desktop
(jimmy3) vboxuser@ubuntu:~/Desktop$ conda install bioconda::subread
Channels:
  - defaults
  - bioconda
Platform: linux-64
Collecting package metadata (repodata.json): \
```

```
vboxuser@ubuntu: ~/Desktop
(jimmy3) vboxuser@ubuntu:~/Desktop$ featurecounts -h
featurecounts: command not found
(jimmy3) vboxuser@ubuntu:~/Desktop$ featureCounts

Version 2.0.1

Usage: featureCounts [options] -a <annotation_file> -o <output_file> input_file1
[input_file2] ...

## Mandatory arguments:

-a <string>      Name of an annotation file. GTF/GFF format by default. See
                 -F option for more format information. Inbuilt annotations
                 (SAF format) is available in 'annotation' directory of the
                 package. Gzipped file is also accepted.

-o <string>      Name of output file including read counts. A separate file
                 including summary statistics of counting results is also
                 included in the output ('<string>.summary'). Both files
                 are in tab delimited format.

input_file1 [input_file2] ...  A list of SAM or BAM format files. They can be
                               either name or location sorted. If no files provided,
                               <stdin> input is expected. Location-sorted paired-end read
```

备选：htseq, 慢

<https://anaconda.org/bioconda/subread>

安装R

```
vboxuser@ubuntu: ~/Desktop
(base) vboxuser@ubuntu:~/Desktop$ ls
(base) vboxuser@ubuntu:~/Desktop$ conda activate jimmy3
(jimmy3) vboxuser@ubuntu:~/Desktop$ conda install conda-forge::r-base
Retrieving notices: ...working... done
Channels:
 - defaults
 - conda-forge
Platform: linux-64
Collecting package metadata (repodata.json): /
```

```
vboxuser@ubuntu: ~/Desktop
Downloading and Extracting Packages:
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
(jimmy3) vboxuser@ubuntu:~/Desktop$ R

R version 4.2.3 (2023-03-15) -- "Shortstop Beagle"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-conda-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

安装R/Python包

```
options(repos=structure(c(CRAN="https://mirrors.tuna.tsinghua.edu.cn/CRAN/")))
```

```
r
```

```
install.packages("ggplot2")
```

```
r
```

```
install.packages("devtools")  
devtools::install_github("hadley/ggplot2")
```

```
r
```

```
if (!requireNamespace("BiocManager", quietly = TRUE))  
  install.packages("BiocManager")  
BiocManager::install("BiocGenerics")
```

```
options(Bioc_mirror="https://mirrors.tuna.tsinghua.edu.cn/bioconductor")
```

```
bash
```

```
conda install r-ggplot2
```

```
r
```

本地安装

```
install.packages("ggplot2_4.0.0.tar.gz", repos=NULL, type="source")
```

library("DESeq2") 无报错，安装成功*

```
bash
```

```
pip install matplotlib  
-i https://mirrors.tuna.tsinghua.edu.cn/pypi/web/simple
```

```
bash
```

```
conda install matplotlib
```

```
bash
```

本地安装

```
pip install /path/to/matplotlib-3.1.1-cp37-cp37m-manylinux1_x86_64.whl
```

import matplotlib 无报错，安装成功*

常见包安装（缺啥装下，见招拆招）

- R: Matrix, ggplot2, pheatmap, BiocManager, DESeq2, clusterProfiler
- Python: matplotlib (numpy), scipy, pandas, seaborn

```
vboxuser@ubuntu: ~/Desktop
(base) vboxuser@ubuntu:~/Desktop$ conda activate jimmy3
(jimmy3) vboxuser@ubuntu:~/Desktop$ R

R version 4.2.3 (2023-03-15) -- "Shortstop Beagle"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-conda-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> install.packages('ggplot2')
--- Please select a CRAN mirror for use in this session ---
```



```
vboxuser@ubuntu: ~/Desktop

** byte-compile and prepare package for lazy loading
** help
*** installing help indices
*** copying figures
** building package indices
** installing vignettes
** testing if installed package can be loaded from temporary location
** checking absolute paths in shared objects and dynamic libraries
** testing if installed package can be loaded from final location
** testing if installed package keeps a record of temporary installation path
* DONE (tibble)
ERROR: dependencies 'MASS', 'mgcv' are not available for package 'ggplot2'
* removing '/home/vboxuser/miniconda3/envs/jimmy3/lib/R/library/ggplot2'

The downloaded source packages are in
      '/tmp/RtmpDeJXhV/downloaded_packages'
Updating HTML index of packages in '.Library'
Making 'packages.html' ... done
Warning messages:
1: In install.packages("ggplot2") :
  installation of package 'mgcv' had non-zero exit status
2: In install.packages("ggplot2") :
  installation of package 'ggplot2' had non-zero exit status
>
```

排查错误 (文件没下载完整, 缺少头文件等)

```
In file included from caches.h:7,  
                 from caches.cpp:1:  
ft_cache.h:9:10: fatal error: ft2build.h: No such file or directory  
9 | #include <ft2build.h>  
  | ^~~~~~
```

<https://www.jianshu.com/p/228629667957>

<https://www.jianshu.com/p/bbc2624e8217>

conda install freetype

sudo apt install libtiff-dev

sudo apt-get install libfreetype6-dev

sudo ln -s /usr/include/freetype2/ft2build.h /usr/include/

conda install r-systemfonts

下载新版ggtree, 本地安装

```
(jimny3) vboxuser@ubuntu:~$ pip install matplotlib -i https://mirrors.tuna.tsinghua.edu.cn/pypi/web/simple  
  
Looking in indexes: https://mirrors.tuna.tsinghua.edu.cn/pypi/web/simple  
Collecting matplotlib  
  Downloading https://mirrors.tuna.tsinghua.edu.cn/pypi/web/packages/8d/9d/d06860390f9d154fa884f1740a5456378fb153ff57443c91a4a32bab7092/matplotlib-3.9.2-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (8.3 MB)  
    5.8/8.3 MB 131.1 kB/s eta 0:00:20
```

Python安装包

用conda安装，或者下载安装

```
vboxuser@ubuntu: ~/Desktop
Updating HTML index of packages in '.Library'
Making 'packages.html' ... done
Warning message:
In install.packages("mgcv") :
  installation of package 'mgcv' had non-zero exit status
> install.packages('Matrix')
Warning message:
package 'Matrix' is not available for this version of R
'Matrix' version 1.7-0 is in the repositories but depends on R (>= 4.4.0)
'Matrix' version 1.7-0 is in the repositories but depends on R (>= 4.5)

A version of this package for your version of R might be available elsewhere,
see the ideas at
https://cran.r-project.org/doc/manuals/r-patched/R-admin.html#Installing-package
s
>
[1]+  Stopped                  R
(jimmy3) vboxuser@ubuntu:~/Desktop$ conda install conda-forge::r-matrix

Channels:
 - defaults
 - conda-forge
Platform: linux-64
Collecting package metadata (repodata.json): \
```

```
vboxuser@ubuntu: ~/Desktop

The following packages will be downloaded:

package | build | size | channel
-----|-----|-----|-----
r-lattice-0.20_45 | r42h76d94ec_0 | 1.1 MB | 
r-matrix-1.6_5 | r42h316c678_0 | 3.7 MB | conda-forge
-----|-----|-----|-----
Total: 4.8 MB

The following NEW packages will be INSTALLED:

r-lattice | pkgs/r/linux-64::r-lattice-0.20_45-r42h76d94ec_0
r-matrix | conda-forge/linux-64::r-matrix-1.6_5-r42h316c678_0

Proceed ([y]/n)?

Downloading and Extracting Packages:

Preparing transaction: done
Verifying transaction: done
Executing transaction: done
(jimmy3) vboxuser@ubuntu:~/Desktop$
(jimmy3) vboxuser@ubuntu:~/Desktop$
```

经验

- base环境不要动， 可以创建多个环境
- 新环境相当于复制了一份， 注意磁盘空间
- 一般默认参数即可。但是建议， 根据自己数据设置参数（搞科研）
- 版本越新， 不一定越好（旧代码， 新版本； 新代码， 旧版本）

Conda vs docker

- 1. 隔离级别：** Conda在用户空间提供隔离，主要针对Python环境和包管理； Docker则在操作系统层面提供隔离，可以包含整个操作系统的用户空间和内核空间。
- 2. 用途：** Conda主要用于包管理和环境管理，特别是对于Python项目； Docker则更侧重于整个应用程序的部署和分发。
- 3. 性能：** Docker容器通常比Conda环境启动更快，占用资源更少，因为它们不需要为每个环境复制整个操作系统。
- 4. 可移植性：** Docker提供了更好的可移植性，因为它的容器可以在不同的机器和云平台上无缝运行。
- 5. 系统资源：** Conda环境可能占用更多磁盘空间，因为每个环境都可能包含完整的Python解释器和库的副本； Docker镜像可以共享底层操作系统的内核和库，从而节省空间。
- 6. 使用场景：** 如果你主要使用Python并且需要管理多个项目的环境，Conda可能是更好的选择。如果你需要部署复杂的应用程序或者需要在多个平台上保持一致性，Docker可能更适合。

安装bedtools

conda install bioconda::bedtools

```
vboxuser@ubuntu: ~/Desktop
depad          convert padded BAM to unpadded BAM

(jimmy3) vboxuser@ubuntu:~/Desktop$
(jimmy3) vboxuser@ubuntu:~/Desktop$ conda install bioconda::bedtools
Channels:
  - defaults
  - bioconda
Platform: linux-64
Collecting package metadata (repodata.json): done
Solving environment: done

## Package Plan ##

environment location: /home/vboxuser/miniconda3/envs/jimmy3

added / updated specs:
  - bioconda::bedtools

The following packages will be downloaded:

package | build | size | channel
-----|-----|-----|-----
bedtools-2.30.0 | h7d7f7ad_2 | 17.9 MB | bioconda
```

bedtools 查看帮助

```
vboxuser@ubuntu: ~/Desktop

Verifying transaction: done
Executing transaction: done
(jimmy3) vboxuser@ubuntu:~/Desktop$
(jimmy3) vboxuser@ubuntu:~/Desktop$ bedtools
bedtools is a powerful toolset for genome arithmetic.

Version:      v2.30.0
About:        developed in the quinlanlab.org and by many contributors worldwide.
Docs:         http://bedtools.readthedocs.io/
Code:         https://github.com/arq5x/bedtools2
Mail:         https://groups.google.com/forum/#!forum/bedtools-discuss

Usage:        bedtools <subcommand> [options]

The bedtools sub-commands include:

[ Genome arithmetic ]
intersect      Find overlapping intervals in various ways.
window         Find overlapping intervals within a window around an interval.
closest        Find the closest, potentially non-overlapping interval.
coverage       Compute the coverage over defined intervals.
map            Apply a function to a column for each overlapping interval.
genomecov      Compute the coverage over an entire genome.
merge          Combine overlapping/nearby intervals into a single interval.
```

<https://anaconda.org/bioconda/bedtools>