

BIOCONDA TUTORIAL

ANACONDA

Anaconda is a free and open-source distribution of the Python and R programming languages for scientific computing (data science, machine learning applications, large-scale data processing, predictive analytics, etc.), that aims to simplify package management and deployment. The distribution includes data-science packages suitable for Windows, Linux, and macOS.

[SOURCE WIKIPEDIA]

For more information please visit the following web links:

<https://www.anaconda.com/>

<https://github.com/Anaconda-Platform>

MINICONDA

Miniconda is a free minimal installer for conda. It is a small, bootstrap version of Anaconda that includes only conda, Python, the packages they depend on, and a small number of other useful packages, including pip, zlib and a few others. Use the **conda install command** to install 720+ additional conda packages from the Anaconda repository.

[SOURCE:<https://docs.conda.io/en/latest/miniconda.html>]

For more information please visit the following web link:

<https://docs.conda.io/en/latest/miniconda.html>

CONDA

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

[SOURCE:<https://docs.conda.io/en/latest/>]

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.

[SOURCE:<https://docs.conda.io/en/latest/>]

For more information please visit the following web link:

<https://docs.conda.io/en/latest/>

“Anaconda can be thought of the data scientists hardware store. It’s got everything you need. From tools for exploring datasets, to tools for modeling them, to tools for visualizing what you’ve found. Everyone can access the hardware store and all the tools inside.”

“Miniconda is the workbench of a data scientist. Every workbench starts clean with only the bare necessities. But as a project grows, so do the number of tools on the workbench. They get used, they get changed, they get swapped. Each workbench can be customized however a data scientist wants. One data scientists workbench may be completely different to another, even if they’re on the same team.”

“Conda helps to organize all of these tools. Although Anaconda comes with many of them ready to go, sometimes they’ll need changing. Conda is like the assistant who takes stock of all the tools. The same goes for Miniconda.”



[SOURCE:<https://www.mrdbourke.com/>]

BIOCONDA

Bioconda (<https://bioconda.github.io>), is a distribution of bioinformatics software for the lightweight, multi-platform and language-agnostic package manager Conda.

Bioconda offers a collection of over 3000 software packages, which is continuously maintained, updated, and extended by a growing global community of more than 200 contributors.

Bioconda improves analysis reproducibility by allowing users to define isolated environments with defined software versions, all of which are easily installed and managed without administrative privileges.

References:

Grüning, Bjeřrn, et al. "Bioconda: sustainable and comprehensive software distribution for the life sciences." *Nature methods* 15.7 (2018): 475-476.

<https://github.com/bioconda/bioconda-recipes>

<https://anaconda.org/bioconda>

NCGR CONDA ENVIRONMENTS:

NCGR has several environments, one for every software. To get a list of all the environments, you can use the conda info command. (see below)

In order to use these programs, their environment(s) have to be activated. Please type the command below to activate the relevant environments:

For example, if you were to run a sw to check quality of reads, you would use FastQC to do the same. In order to activate the environment you would,

```
#login to logrus and go to your home directory
cd /home/asundara/
#list the different environments on NCGR's conda
conda info --envs
```

```
[asundara@logrus ~]$ cd /home/asundara/
[asundara@logrus ~]$ conda info --envs
# conda environments:
#
qiime2-2020.6      /home/asundara/.conda/envs/qiime2-2020.6
base               * /home/condasw/miniconda3
HISAT              /home/condasw/miniconda3/envs/HISAT
HTSEQ              /home/condasw/miniconda3/envs/HTSEQ
STAR               /home/condasw/miniconda3/envs/STAR
assembly           /home/condasw/miniconda3/envs/assembly
centrifuge         /home/condasw/miniconda3/envs/centrifuge
deseq2-r           /home/condasw/miniconda3/envs/deseq2-r
fastQC             /home/condasw/miniconda3/envs/fastQC
fastqdump          /home/condasw/miniconda3/envs/fastqdump
featurecounts      /home/condasw/miniconda3/envs/featurecounts
```

```
#activate the relevant environment to run the sw
source activate fastQC
```

```
[asundara@logrus ~]$ source activate fastQC
(fastQC) [asundara@logrus ~]$ █
```

To exit out of the environment, please type the command below,

```
conda deactivate
```

```
[(fastQC) [asundara@logrus ~]$ conda deactivate  
[asundara@logrus ~]$ █
```