Sep 20, 2019

Use IPython to run bioconda tools in jupyter V.3

Alise J. Ponsero[1]

[1]University of Arizona

1 *Works for me* dx.doi.org/10.17504/protocols.io.7iqhkdw

Metagenomics class 2019


Alise J. Ponsero
University of Arizona

ABSTRACT

Learn how to run shell commands directly in your jupyter notebook using IPython

BEFORE STARTING

For this protocol, you'll need Anaconda/miniconda and jupyter installed on your machine.

Looking at bioconda

1 Meet Bioconda. Bioconda is a project aiming to integrate bioinformatic tools in the conda package management system. They have over 6000 bioinformatic tools available and a repository of recipe on how to use them.

Yes, Bioconda is your new best friend.

You can browse for your tool of interest in Bioconda through their tool repository. For this tutorial we will install and run Mothur.

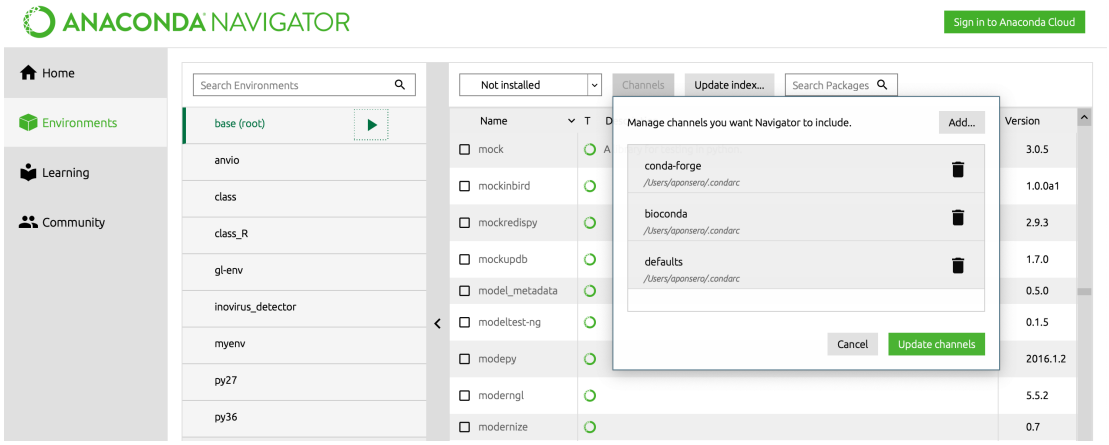
installing Mothur

2 Open Anaconda Navigator and go to your environment manager. Select the class environment and search for the Mothur package in the "not installed" packages list.

Install Mothur in the environment. If you don't remember how to do this. Do not panic and go back to the Anaconda protocol (https://www.protocols.io/view/first-steps-using-a-jupyter-notebook-6uphevn).

The installation will take some time, therefore you can read this very serious paper about Inuit poop knives.

If you don't find mothur in the "not installed" or the "all" list of packages in the conda navigator look at your channels in the conda navigator :



If you don't see these three channels, go to step 2.2

2.1 For Miniconda users, open your terminal and type :

```
source activate class_2019
```

You should have your environment correctly loaded. Then type :

```
conda config --add channels defaults
conda config --add channels bioconda
conda config --add channels conda-forge
```

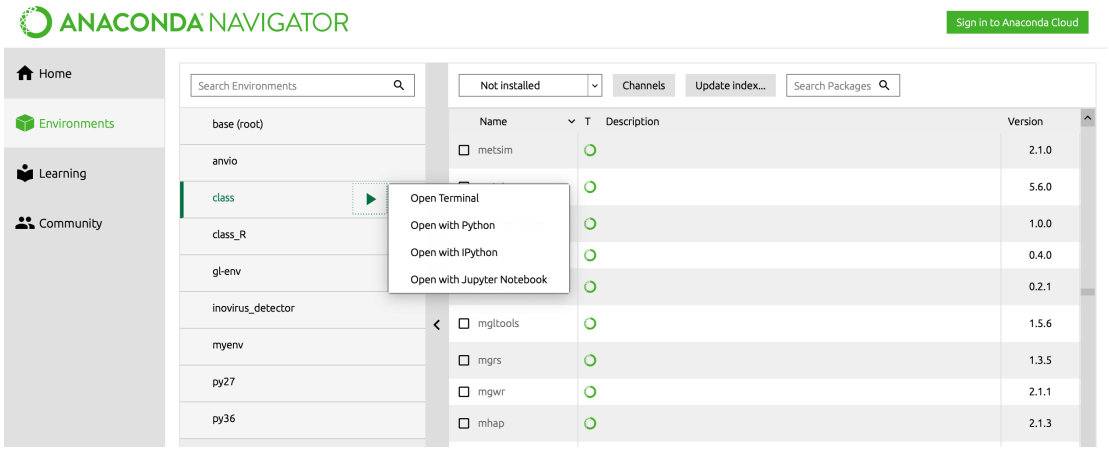
That should have set up your conda channels correctly. Now, let's install mothur.

```
conda install Mothur
```

Confirm the installation by typing "y" and wait patiently for the installation.

2.2 Let's fix your channels (these are the channels that conda look at when they look for packages to install).

In the conda navigator, load your "class_2019" environment, then click on the "triangle" logo and ask to open a terminal :



In the terminal type :

```
conda config --add channels defaults
conda config --add channels bioconda
conda config --add channels conda-forge
```

then :

```
conda install mothur
```

When the prompt asks you if you really want to install the packages, type "y". Once the install is complete, close the terminal and go back to the anaconda navigator.

Using IPython

3 Is Mothur is installed in your conda environment? Great! Now open a Python jupyter notebook in your Document folder.

4 Think of IPython as a bridge between Python and shell commands. It's basically allowing you to run shell command lines directly into your Python Notebook ! Are you excited? Ok, let's try it!

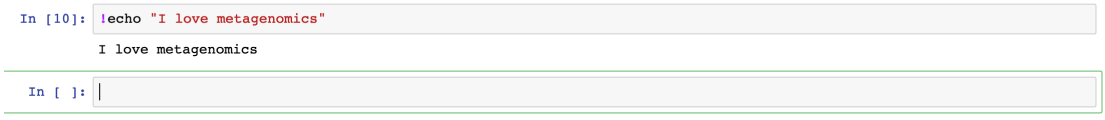
As you (should) know, in shell you can print a text using the "echo" command. Like this :

```
echo "I love metagenomics"
```

Well, let's say you REALLY want to use the "echo" command in your notebook, then IPython allows you to just type :

```
!echo "I love metagenomics"
```

Now, try it and run the cell ! You should see



And the same idea applies to ls, pwd, mkdir... all your favorite commands are there! All? Well, almost all...

5 If you play with IPython long enough, you'll quickly notice that !cd doesn't exactly work as it is supposed to. If you try to run

```
!cd ..
```

You won't see an error, but you won't have changed directory either... This is because each cell from the notebook is run as a temporary sub-shell. It opened a shell, ran the cd .. and then closed it. So you're back to square one.

To run these commands, you need to use a % instead of the !

```
%cd ..
```

Now, it should work ! And even better, if you have the "automagic" IPython function installed, you don't need to use the "%". You can simply run :

```
cd ..
```

You have automagic functions for the main shell commands (ls, cd, cat, man, mkdir, mv, pwd, rm and rmdir).

running Mothur

6 So let's use IPython to run Mothur. To do so, you need to find where conda installs your packages. The path to this folder should look like this :

'/Users/[USERNAME]/anaconda3/envs/class_2019/bin'

Once you know where is you bin folder, you can run Mothur using the command (here showing the output of the help command in Mothur) :

```
!/Users/[USERNAME]/anaconda3/envs/class_2019/bin/mothur "#help();"
```

Note that Mothur is by default an interactive tool. You cannot use this feature in a notebook, however, you can use Mothur command line mode.

7 The above command can become very long and painful to write. It would be awesome to be able to save "/Users/[USERNAME]/anaconda3/envs/class_2019/bin" in a variable called "bin_dir", right? Well, let's do this !
First, let's create the variable:

```
bin_dir="/Users/[USERNAME]/anaconda3/envs/class_2019/bin"
```

Now, we can to pass this variable in our IPython command using the {variable} syntax:

```
!{bin_dir}/mothur "#help();"
```