

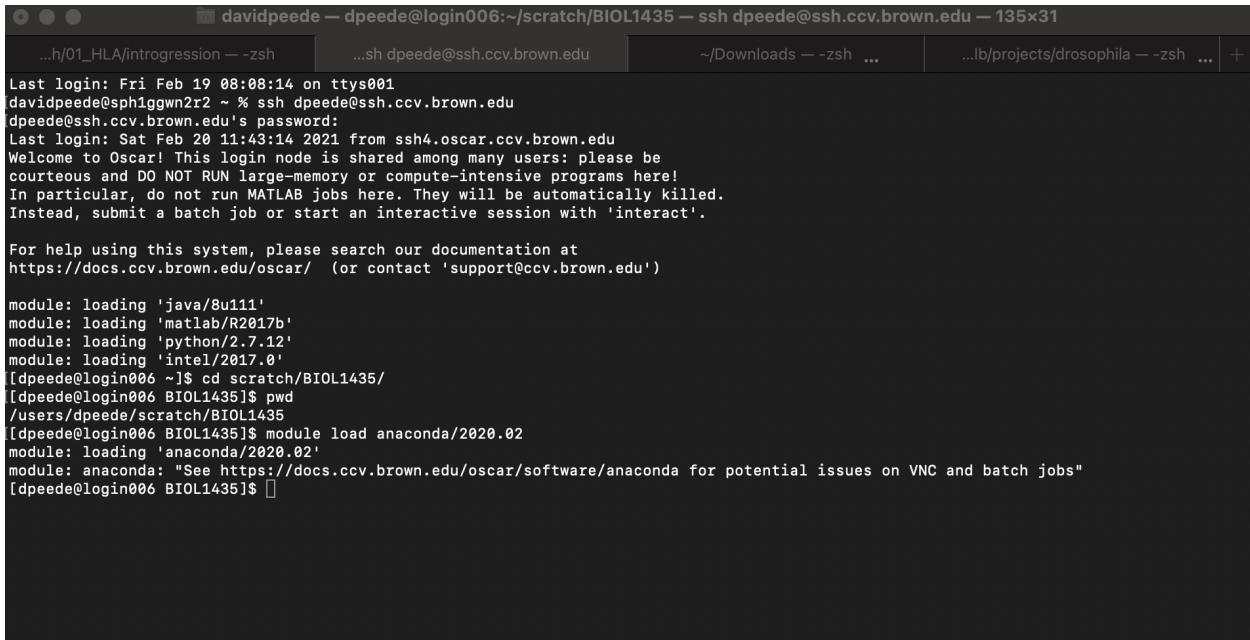
Anaconda on OSCAR

David Peede

February 21, 2021

1 Introduction

Anaconda environments allow for easy package installation and management. I like to use Anaconda environments as opposed to loading modules on OSCAR since Anaconda takes care of all my package dependencies as well as updating packages! So lets log on to OSCAR and navigate to our BIOL1435 directory in our scratch space that we created during the introduction to OSCAR tutorial. Once you are there type `module load anaconda/2020.02` and press enter to load the anaconda module. If you did everything correctly your terminal should look similar to Figure 1.



The screenshot shows a terminal window with four tabs. The active tab displays a shell session for user dpeede@login006. The session starts with a login message from a shared node, followed by a welcome message from the Oscar system. The user then navigates to their scratch directory and loads the anaconda module using the command `module load anaconda/2020.02`. A note at the bottom of the screen provides a link for potential issues on VNC and batch jobs.

```
davidpeede - dpeede@login006:~/scratch/BIOL1435 - ssh dpeede@ssh.ccv.brown.edu - 135x31
...h/01_HLA/introgression -- zsh ... .sh dpeede@ssh.ccv.brown.edu ... ~/Downloads -- zsh ... ...lb/projects/drosophila -- zsh ... +
Last login: Fri Feb 19 08:08:14 on ttys001
[davidpeede@spfh1ggwn2r2 ~ % ssh dpeede@ssh.ccv.brown.edu
[dpeede@ssh.ccv.brown.edu's password:
Last login: Sat Feb 20 11:43:14 2021 from ssh4.oscar.ccv.brown.edu
Welcome to Oscar! This login node is shared among many users: please be
courteous and DO NOT RUN large-memory or compute-intensive programs here!
In particular, do not run MATLAB jobs here. They will be automatically killed.
Instead, submit a batch job or start an interactive session with 'interact'.

For help using this system, please search our documentation at
https://docs.ccv.brown.edu/oscar/ (or contact 'support@ccv.brown.edu')

module: loading 'java/8u111'
module: loading 'matlab/R2017b'
module: loading 'python/2.7.12'
module: loading 'intel/2017.0'
[dpeede@login006 ~]$ cd scratch/BIOL1435/
[dpeede@login006 BIOL1435]$ pwd
/users/dpeede/scratch/BIOL1435
[dpeede@login006 BIOL1435]$ module load anaconda/2020.02
module: loading 'anaconda/2020.02'
module: anaconda: "See https://docs.ccv.brown.edu/oscar/software/anaconda for potential issues on VNC and batch jobs"
[dpeede@login006 BIOL1435]$
```

Figure 1: Loading the anaconda module on OSCAR.

2 Creating Your Conda Environment

Now let's create our Conda environment! Type `conda create -n BIOL1435env` and press enter. After a second or two you will be asked if you want to proceed type `y` and press enter, your terminal should look similar to Figure 2.

```

davidpeede@...:~/scratch/BIOL1435$ ssh dpeede@ssh.ccv.brown.edu
module: loading 'matlab/R2017b'
module: loading 'python/2.7.12'
module: loading 'intel/2017.0'
[dpeede@login006 ~]$ cd scratch/BIOL1435/
[dpeede@login006 BIOL1435]$ pwd
/users/dpeede/scratch/BIOL1435
[dpeede@login006 BIOL1435]$ module load anaconda/2020.02
module: loading 'anaconda/2020.02'
module: anaconda: "See https://docs.ccv.brown.edu/oscar/software/anaconda for potential issues on VNC and batch jobs"
[dpeede@login006 BIOL1435]$ conda create -n BIOL1435env
Collecting package metadata (current_repodata.json): done
Solving environment: done

==> WARNING: A newer version of conda exists. <==
  current version: 4.8.2
  latest version: 4.9.2

Please update conda by running

$ conda update -n base -c defaults conda

## Package Plan ##

environment location: /users/dpeede/anaconda/BIOL1435env

Proceed ([y]/n)? y

```

Figure 2: Creating a Conda environment.

After you successfully created your Conda environment your terminal should look similar to Figure 3.

```

davidpeede@...:~/scratch/BIOL1435$ ssh dpeede@ssh.ccv.brown.edu
...h01_HLA/introgression -- zsh ... sh dpeede@ssh.ccv.brown.edu ... ~/Downloads -- zsh ... ...lb/projects/drosophila -- zsh ... +
==> WARNING: A newer version of conda exists. <==
  current version: 4.8.2
  latest version: 4.9.2

Please update conda by running

$ conda update -n base -c defaults conda

## Package Plan ##

environment location: /users/dpeede/anaconda/BIOL1435env

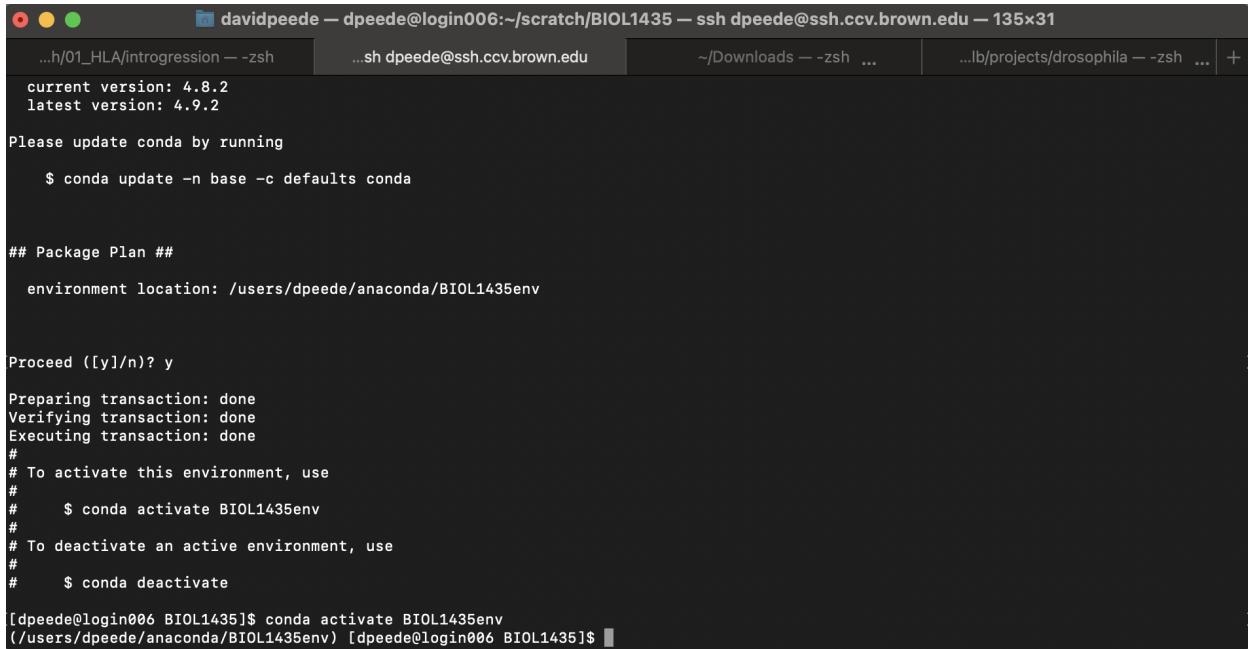
Proceed ([y]/n)? y

Preparing transaction: done
Verifying transaction: done
Executing transaction: done
#
# To activate this environment, use
#
# $ conda activate BIOL1435env
#
# To deactivate an active environment, use
#
# $ conda deactivate
[dpeede@login006 BIOL1435]$

```

Figure 3: Successfully creating a Conda environment.

Now that we have created our Conda environment let's activate our newly created environment, type `conda activate BIOL1435env` and then press enter. Your terminal should look similar to Figure 4.



The screenshot shows a terminal window with the title bar "davidpeede — dpeede@login006:~/scratch/BIOL1435 — ssh dpeede@ssh.ccv.brown.edu — 135x31". The terminal content is as follows:

```
...h/01_HLA/introgression — zsh ...sh dpeede@ssh.ccv.brown.edu ~/Downloads — zsh ...lb/projects/drosophila — zsh ... +  
current version: 4.8.2  
latest version: 4.9.2  
Please update conda by running  
$ conda update -n base -c defaults conda  
  
## Package Plan ##  
environment location: /users/dpeede/anaconda/BIOL1435env  
  
Proceed ([y]/n)? y  
Preparing transaction: done  
Verifying transaction: done  
Executing transaction: done  
#  
# To activate this environment, use  
#  
# $ conda activate BIOL1435env  
#  
# To deactivate an active environment, use  
#  
# $ conda deactivate  
[dpeede@login006 BIOL1435]$ conda activate BIOL1435env  
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login006 BIOL1435]$
```

Figure 4: Activating your Conda environment.

3 Installing Packages

Now let's load msprime to our environment! First we need to set up the conda-forge channel, which will allow us to download packages that are not on the default channel of Anaconda (like msprime), to configure the conda-forge channel type `conda config --add channels conda-forge` and then press enter. You should only have to do this once after creating your first Conda environment so if you are new to Anaconda your terminal should look different than Figure 5.

```

Downloads — dpeede@login005:~/scratch/BIOL1435 — ssh dpeede@ssh.ccv.brown.edu — 135x31
...h/01_HLA/introgression — -zsh | ~/Downloads — -zsh | ...dpeede@ssh.ccv.brown.edu | ...lb/projects/drosophila — -zsh ... + 

Last login: Sun Feb 21 08:18:24 on ttys003
davidpeede@sph1ggwn2r2 Downloads % ssh dpeede@ssh.ccv.brown.edu
dpeede@ssh.ccv.brown.edu's password:
Last login: Sat Feb 20 08:58:35 2021 from ssh4.oscar.ccv.brown.edu
Welcome to Oscar! This login node is shared among many users: please be
courteous and DO NOT RUN large-memory or compute-intensive programs here!
In particular, do not run MATLAB jobs here. They will be automatically killed.
Instead, submit a batch job or start an interactive session with 'interact'.

For help using this system, please search our documentation at
https://docs.ccv.brown.edu/oscar/ (or contact 'support@ccv.brown.edu')

module: loading 'java/8u111'
module: loading 'matlab/R2017b'
module: loading 'python/2.7.12'
module: loading 'intel/2017.0'
[dpeede@login005 ~]$ cd scratch/BIOL1435/
[dpeede@login005 BIOL1435]$ module load anaconda/2020.02
module: loading 'anaconda/2020.02'
module: anaconda: "See https://docs.ccv.brown.edu/oscar/software/anaconda for potential issues on VNC and batch jobs"
[dpeede@login005 BIOL1435]$ conda activate BIOL1435env
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ conda config --add channels conda-forge
Warning: 'conda-forge' already in 'channels' list, moving to the top
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ 

```

Figure 5: Configuring the conda-forge channel.

Now type `conda install msprime` and press enter. Your terminal should first look similar to Figure 6 and after prompted to proceed type `y` and press enter, as seen in Figure 7. If everything downloaded properly your terminal should look similar to Figure 8.

```

Downloads — dpeede@login005:~/scratch/BIOL1435 — ssh dpeede@ssh.ccv.brown.edu — 135x31
...h/01_HLA/introgression — -zsh | ~/Downloads — -zsh | ...dpeede@ssh.ccv.brown.edu | ...lb/projects/drosophila — -zsh ... + 

(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ conda config --add channels conda-forge
Warning: 'conda-forge' already in 'channels' list, moving to the top
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ conda install msprime
Collecting package metadata (current_repodata.json): done
Solving environment: done

==> WARNING: A newer version of conda exists. <==
  current version: 4.8.2
  latest version: 4.9.2

Please update conda by running

$ conda update -n base -c defaults conda

## Package Plan ##

environment location: /users/dpeede/anaconda/BIOL1435env
added / updated specs:
- msprime

The following packages will be downloaded:

  package          |      build
  -----|-----
  gsl-2.6           |  he838d99_2      3.2 MB  conda-forge
  ld_impl_linux-64-2.35.1 |  hea4e1c9_2      618 KB  conda-forge

```

Figure 6: Installing msprime part 1.

```
Downloads — dpeede@login005:~/scratch/BIOL1435 — ssh dpeede@ssh.ccv.brown.edu — 135x31
...h01_HLA/introgression -- zsh          ~/Downloads -- zsh          ...dpeede@ssh.ccv.brown.edu          ...lb/projects/drosophila -- zsh  ...
libgfortran5      conda-forge/linux-64::libgfortran5-9.3.0-hff62375_18
libgomp           conda-forge/linux-64::libgomp-9.3.0-h2828fa1_18
liblapack          conda-forge/linux-64::liblapack-3.9.0-0-openblas
libnghttp2         conda-forge/linux-64::libnghttp2-1.43.0-h812cca2_0
libopenblas        conda-forge/linux-64::libopenblas-0.3.12-pthreads_h4812303_1
libssh2            conda-forge/linux-64::libssh2-1.9.0-hab1572f_5
libstdcxx-ng       conda-forge/linux-64::libstdcxx-ng-9.3.0-h6de172a_18
msprime           conda-forge/linux-64::msprime-0.7.4-py39h0263c30_2
ncurses            conda-forge/linux-64::ncurses-6.2-h58526e2_4
numpy              conda-forge/linux-64::numpy-1.20.1-py39hdbf815f_0
openssl            conda-forge/linux-64::openssl-1.1.1j-h7f98852_0
pip                conda-forge/noarch::pip-21.0.1-pyhd8ed1ab_0
pyparsing          conda-forge/noarch::pyparsing-2.4.7-pyh9f0adid_0
pyrsistent         conda-forge/linux-64::pyrsistent-0.17.3-py39h3811e60_2
python              conda-forge/linux-64::python-3.9.2-hffdb5ce_0_cpython
python_abi          conda-forge/linux-64::python_abi-3.9-1_cp39
readline           conda-forge/linux-64::readline-8.0-he28a2e2_2
setuptools         conda-forge/linux-64::setuptools-49.6.0-py39hf3d152e_3
six                conda-forge/noarch::six-1.15.0-pyh9f0adid_0
sqlite             conda-forge/linux-64::sqlite-3.34.0-h74cdb3f_0
svgwrite           conda-forge/noarch::svgwrite-1.4.1-pyhd8ed1ab_0
tk                 conda-forge/linux-64::tk-8.6.10-h21135ba_1
tskit              conda-forge/linux-64::tskit-0.3.4-py39hce5d2b2_0
tzdata             conda-forge/noarch::tzdata-2021a-he74cb21_0
wheel              conda-forge/noarch::wheel-0.36.2-pyhd3deb0d_0
xz                 conda-forge/linux-64::xz-5.2.5-h516909a_1
zipp               conda-forge/noarch::zipp-3.4.0-py_0
zlib               conda-forge/linux-64::zlib-1.2.11-h516909a_1010

Proceed ([y]/n)? y
```

Figure 7: Installing msprime part 2.

Figure 8: Successfully installing msprime.

4 Using msprime from your Conda Environment

Now that we have installed msprime let's use it! First type `python` and press enter so that we move out of our bash shell and start a python shell. Copy the following code line by line followed by pressing the enter

key. Your terminal should look similar to Figure 9 (NOTE: since we did not specify a random seed number your tree topology might look a tad different than mine).

```
1 import msprime
2 tree = msprime.simulate(5).first()
3 print(tree.draw(format="unicode"))
```

Downloads — dpeede@login005:~/scratch/BIOL1435 — ssh dpeede@ssh.ccv.brown.edu — 135x31

	...h01_HLA/introgression -- zsh	~/Downloads -- zsh	...dpeede@ssh.ccv.brown.edu	...lb/projects/drosophila -- zsh ...
pip-21.0.1	1.1 MB	##### # #####	##### # #####	##### # #####
python-3.9.2	27.3 MB	##### # #####	##### # #####	##### # #####
libblas-3.9.0	11 KB	##### # #####	##### # #####	##### # #####
tzdata-2021a	121 KB	##### # #####	##### # #####	##### # #####
ld_impl_linux-64-2.3	618 KB	##### # #####	##### # #####	##### # #####
libnghttp2-1.43.0	808 KB	##### # #####	##### # #####	##### # #####
liblapack-3.9.0	11 KB	##### # #####	##### # #####	##### # #####
gsl-2.6	3.2 MB	##### # #####	##### # #####	##### # #####
numpy-1.20.1	5.8 MB	##### # #####	##### # #####	##### # #####
libcblas-3.9.0	11 KB	##### # #####	##### # #####	##### # #####
Preparing transaction: done				
Verifying transaction: done				
Executing transaction: done				
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]\$ python				
Python 3.9.2 packaged by conda-forge (default, Feb 21 2021, 05:02:46)				
[GCC 9.3.0] on linux				
Type "help", "copyright", "credits" or "license" for more information.				
>>> import msprime				
>>> tree = msprime.simulate(5).first()				
>>> print(tree.draw(format="unicode"))				
8				
7				
6				
5				
2				
3				
0				
1				
>>>				

Figure 9: Running msprime in a python shell.

Now let's exit our python shell and return to the default bash shell. To do so type `exit()` and press enter. Next, type `vi my_first_sim.py` and press enter. If you remember from the previous walk through this opens a text editor so that we can write our python script! Type `i` to enter insert mode and copy and paste the following python code into the document, after you are done type `:x` and press enter to save and exit your python script. Your terminal should look similar to Figure 10 and Figure 11.

```
1 import msprime
2 tree_sequence = msprime.simulate(sample_size=6, Ne=1000)
3 tree = tree_sequence.first()
4 print(tree.draw(format="unicode"))
```

The screenshot shows a terminal window with four tabs. The active tab displays the command `python` followed by a series of imports and a tree simulation. A phylogenetic tree diagram is shown, rooted at node 8, with nodes numbered 0 through 7. The tree structure is as follows:

```

graph TD
    8 --- 1
    1 --- 7
    1 --- 6
    6 --- 5
    6 --- 5
    5 --- 0
    5 --- 4
    5 --- 1
    5 --- 2
    5 --- 3
  
```

Below the tree, the command `>>> exit()` is entered, followed by the prompt `(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ vi my_first_sim.py`.

Figure 10: Exiting the python shell.

The screenshot shows a terminal window with four tabs. The active tab displays a Python script named `my_first_sim.py`. The script contains the following code:

```

import msprime
tree_sequence = msprime.simulate(sample_size=6, Ne=1000)
tree = tree_sequence.first()
print(tree.draw(format="unicode"))
  
```

The script is 4 lines long and 136 characters wide. The status bar at the bottom right indicates "4,1 All".

Figure 11: Your python script.

Now let's make this python script executable by typing `chmod +x my_first_sim.py` followed by pressing enter. To actually execute our script type `python my_first_sim.py` and press enter. Your terminal should look similar to Figure 12 and Figure 13 (NOTE: since we did not specify a random seed number your tree topology might look a tad different than mine).

Figure 12: Making our python script executable.

```
Downloads — dpeede@login005:~/scratch/BIOL1435 — ssh dpeede@ssh.ccv.brown.edu — 135x31
...h01_HLA/introgression -- zsh          ~/Downloads -- zsh          ...dpeede@ssh.ccv.brown.edu      ...lb/projects/drosophila -- zsh  ...
Type "help", "copyright", "credits" or "license" for more information.
>>> import msprime
>>> tree = msprime.simulate(5).first()
>>> print(tree.draw(format="unicode"))
    8
   / \
  7   6
 / \ / \
0  4 1  2 3
>>> exit()
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ vi my_first_sim.py
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ chmod +x my_first_sim.py
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ python my_first_sim.py
    10
   / \
   9   8
  / \ / \
  7   8 4
 / \ / \
 6   7 3 4
 / \ / \
 5   2 3 1
0  5 2 3 4 1
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$
```

Figure 13: Output from our python script

5 Exiting Your Conda Environment

Lastly to exit your Conda environment simply type `conda deactivate` and press enter. After exiting your Conda environment your terminal should look similar to Figure 14.

```
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ cd Downloads
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ ls
...h01_HLA/introgression --zsh  ~/Downloads --zsh  ...dpeede@ssh.ccv.brown.edu  ...lb/projects/drosophila --zsh  ...
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ cd ...h01_HLA/introgression
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ python my_first_sim.py
10
9
8
7
6
5
4
3
2
1
0
(/users/dpeede/anaconda/BIOL1435env) [dpeede@login005 BIOL1435]$ conda deactivate
[dpeede@login005 BIOL1435]$
```

Figure 14: Exiting your Conda environment

6 Notes

As always feel free to email me david_peede@brown.edu if you run into any issues! To learn more about Anaconda visit <https://anaconda.org/> and here is a cheat sheet that I found very helpful https://conda.io/projects/conda/en/latest/_downloads/843d9e0198f2a193a3484886fa28163c/conda-cheatsheet.pdf. Lastly visit <https://msprime.readthedocs.io/en/stable/introduction.html> to learn more about msprime. Happy simulating!