New instructions 2022-02-02:

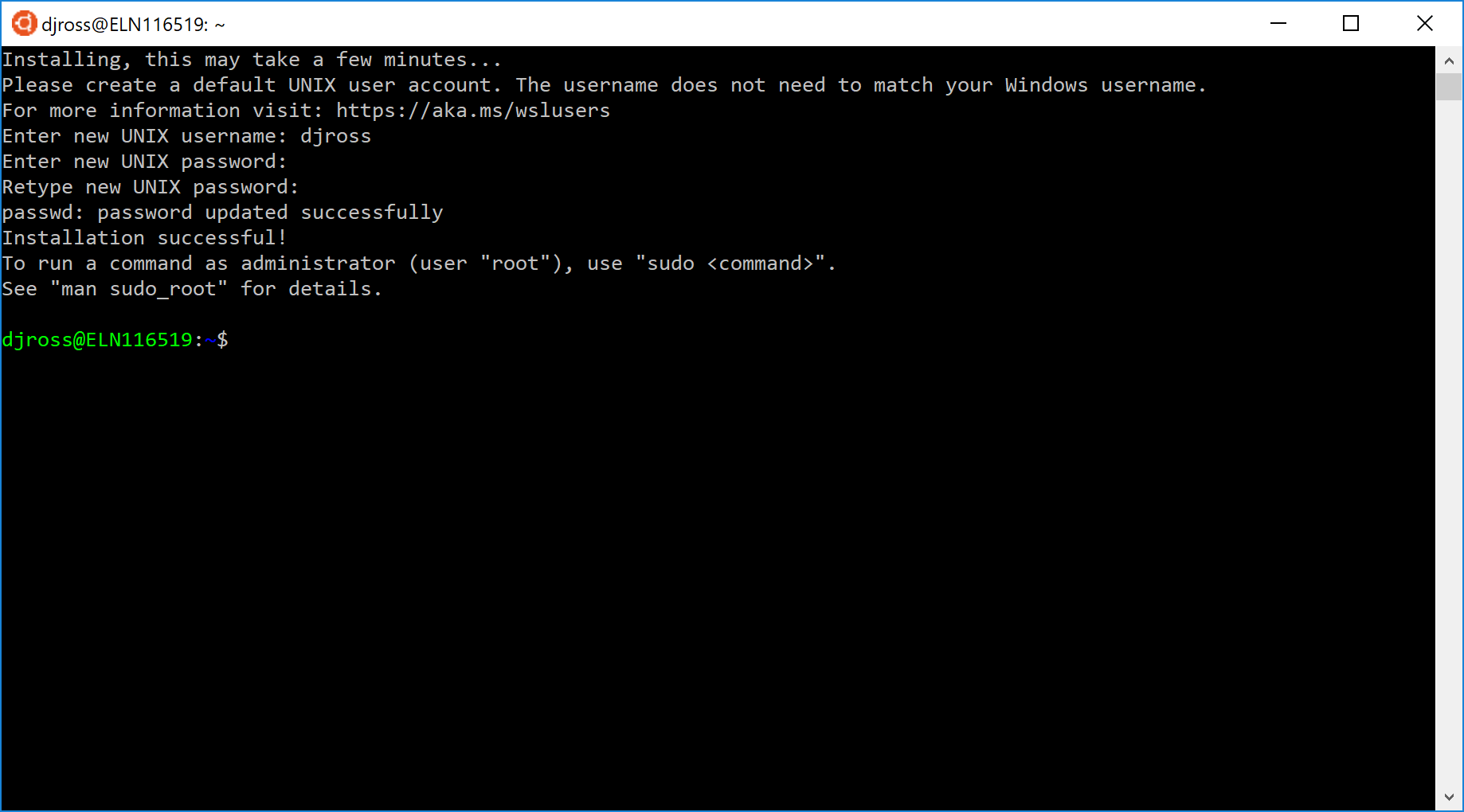
In Windows command, install WSL and Ubuntu with:

wsl --install

Details: <https://docs.microsoft.com/en-us/windows/wsl/setup/environment>

Installation will require a reboot (probably).

When it is done, follow the instructions to set up your UNIX username and password:



There is a problem with the way Windows sets things up though. So, the WSL-Ubuntu environment won’t be able to access anything from the internet (needed to install stuff below).

The problem might be related to the fact that I’m logging into the computer via the NIST VPN.

So, try this:

**While on VPN** in Powershell/Cmd:

ipconfig /all

In the outpour, search for the adapter that is linked to the VPN connection:

“Hyper-V Virtual Ethernet Adapter”

Under that, get the list of DNS servers:

fec0:0:0:ffff::1%1

fec0:0:0:ffff::2%1

fec0:0:0:ffff::3%1

In Powershell:

Get-NetAdapter | Where-Object {$\_.InterfaceDescription -Match " Hyper-V Virtual Ethernet Adapter”"} | Set-NetIPInterface -InterfaceMetric 6000

Then on WSL:

sudo echo fec0:0:0:ffff::1%1 > /etc/resolv.conf

Doesn’t work: permission denied…

To run the Levy lab Bartender code, you'll need python 2.7 and the g++ compiler:

sudo apt-get update (updates all installed packages on Ubuntu)

sudo apt-get install python (installs python 2.7)

~~sudo apt-get install python-pip~~

update on installing pip for python 2.7:

curl [https://bootstrap.pypa.io/pip/2.7/get-pip.py -o get-pip.py](https://bootstrap.pypa.io/pip/2.7/get-pip.py%20-o%20get-pip.py)

then:

python get-pip.py

Also, the latest version of biopython doesn’t work with python 2.7, so install old version (I don’t think it actually needs Biopython the way I use bartender):

python -m pip install biopython==1.76

sudo apt install g++ (installs gcc compiler needed to compile c++ bartender code)

Then, to compile and install the bartender code:

cd bartender-1.1 (edit this line to navigate to the "bartender-1.1" directory)

sudo apt install make

make all

sudo make install

If you have copied the already compiled files, when you try to run "make all", you may get the message, "Nothing to be done for 'all'". This is ok; it just means that the compiled files you have copied are still ok (I think).

To test the installation (using the example data provided as part of the bartender-1.1 repository):

cd example

cd random\_small\_data

bartender\_extractor\_com -f 2M\_test.fq -o 2M\_extracted -q ? -p TACC[4-7]AA[4-7]AA[4-7]TT[4-7]ATAA -m 2

Or, on an AWS instance, you may need to explicitly call python with a pointer to the relevant bartender "\_com" wrapper:

python ~/.local/bin/bartender\_extractor\_com -f 2M\_test.fq -o 2M\_extracted -q ? -p TACC[4-7]AA[4-7]AA[4-7]TT[4-7]ATAA -m 2

or

sudo python ~/bartender-1.1/bartender\_extractor\_com -f 2M\_test.fq -o 2M\_extracted -q ? -p TACC[4-7]AA[4-7]AA[4-7]TT[4-7]ATAA -m 2

ubuntu ubuntu 93371MB 
na 
total 187331MB 
-rw-rw r 
-rw-rw r 
1 
1 
1 
1 
1 
ubuntu root 
ubuntu root 
571MB 
2MB 
-16/HiSeq/barcode_ana1ysis 
Nov 16 16:24 2019-10-16_HiSeq_ 
Nov 16 18:14 2019-10-16_HiSeq_ 
Nov 16 18:13 2019-10-16_HiSeq_ 
Nov 16 15:55 2019-10-16_HiSeq_ 
Nov 16 18:13 2019-10-16_HiSeq_ 
-16/HiSeq/barcode_ana1ysis$ 
-1 
-block-size=MB 
forward_lintags . txt 
barcode . csv 
reverse 
cluster. csv 
reverse 
ubuntu ubuntu 93370MB 
ubuntu root 
18MB 
reverse 
reverse 
_lintags . txt 
_ quality . csv 
ubuntu@ip 
-10-200-221-221:æ/2019-10 

The output should look like:

Machine generated alternative text:
ELNI 16519: /m nt/c/users/dj ross/sou rce/repos/bartender-l .1 [exam ple/random_smal l_data 
-c -03 -std=c++ll -pthread src/Sing1eReadProcessorWithUmi.cpp -o src/Sing1eReadProcessorWithUmi.o 
g++ ./src/kmerdecoder.o ./src/kmers_bitwisetransform.o ./src/kmers_dictionary.o ./src/sequence.o ./src/util.o ./src/sing 
lereadsprocessor.o ./src/bartender_extractor.o ./src/barcodeextractor.o ./src/pattern.o ./src/fastapattern.o ./src/fastq 
pattern. o ./src/UmiExtractor.o ./src/Sing1eReadProcessorWithUmi.o -o bartender _ extractor 
-q 
djross@ELN116519: 
[sudo] password for djross: 
cp bartender _ single 
/usr/local/bin/ 
cp bartender _ extractor /usr/local/bin/ 
cp bartender _ combiner /usr/local/bin/ 
cp bartender_single.com /usr/local/bin/ 
cp bartender_extractor.com /usr/local/bin/ 
cp bartender_combiner_com /usr/local/bin/ 
djross@ELN116519: 
Python 2.7.15+ (default, Jul 9 2019, 16:51:35) 
[GCC 7.4. e ] on linux2 
Type "help" , 
'copyright", "credits" or "license" 
exit() 
djross@ELN116519: 
djross@ELN116519: 
djross@ELN116519: 
test.fq -0 2M _ extracted 
for 
more 
$ sudo make install 
$ python 
information . 
$ cd example/ 
$ cd random_small_data/ 
$ bartender_extractor.com -f 2M 
-7]ATAA -m 2 
Running bartender extractor 
bartender _ extractor 2M _ test . fq 
2M_extracted 63 " (TAC. ITA.clT.ccl . Acc) ( 7}) (AA) ( 7}) (AA) ( 
IAT.AIA.AAI . TAA)" TACC ATAA 9 1 
Totally there are leee reads in 2M_test.fq file! 
Totally there are 976 valid barcodes from 2M_test.fq file 
Totally there are 924 valid barcodes whose quality pass the quality condition 
The estimated sequence error from the prefix and suffix parts is 0.00153689 
djross@ELN116519: 

To run the Bartender clustering on AWS:

Set up and start AWS instance.

Connect to instance using ssh and PEM file.

example:

ssh -i "XXXXXXXXXXX.pem" ubuntu@NN.NNN.NNN.NNN

Install Bartender following instructions above.

Move files to AWS instance with scp

example:

scp -r -i "C:/Users/user/XXXXXXXXXXX.pem" example\_data ubuntu@ NN.NNN.NNN.NNN:~

Run forward and reverse barcode clustering on AWS instance.

examples:

sudo python ~/.local/bin/bartender\_single\_com -f ~/example\_data /barcode\_analysis/example\_reverse\_lintags.txt -o ~/example\_data /barcode\_analysis/example\_reverse -c 500 -z 8 -l 5 -s 1 -t 64 -d 4 &> example\_reverse.clustering.log

sudo python ~/.local/bin/bartender\_single\_com -f ~/example\_data /barcode\_analysis/example\_forward\_lintags.txt -o ~/example\_data /barcode\_analysis/example\_forward -c 500 -z 8 -l 5 -s 1 -t 64 -d 4 &> 2019-example\_forward.clustering.log

Since Bartender was run with sudo, you have to change the ownership of the resulting output files before you can transfer them back to your local machine.

example:

sudo chown ubuntu example\_reverse\_barcode.csv

Then transfer the Bartender output files and the log files back to your local machine:

examples:

scp -r -i "C:/Users/user/XXXXXXXX.pem" ubuntu@ NN.NNN.NNN.NNN:~/example\_data/barcode\_analysis\{example\_reverse\_barcode.csv,example\_reverse\_cluster.csv,example\_reverse\_quality.csv\} "C:\Users\user\Documents\barcode\_analysis"

scp -r -i "C:/Users/user/Documents/AWS/XXXXXXXXXXXXX.pem" ubuntu@ NN.NNN.NNN.NNN:~/example\_data/barcode\_analysis\{example\_forward\_barcode.csv,example\_forward\_cluster.csv,example\_forward\_quality.csv\} "C:\Users\user\Documents\barcode\_analysis"

scp -r -i "C:/Users/user/Documents/AWS/XXXXXXXXXXXXX.pem" ubuntu@ NN.NNN.NNN.NNN:~/example\_forward.clustering.log "C:\Users\user\Documents\barcode\_analysis"

scp -r -i "C:/Users/user/Documents/AWS/XXXXXXXXXXXXX.pem" ubuntu@ NN.NNN.NNN.NNN:~/example\_reverse.clustering.log "C:\Users\user\Documents\barcode\_analysis"