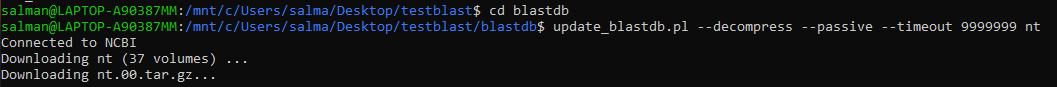
**Automating BLAST Analysis**

**Use Linux.**  
You will need a Linux terminal. If you are working on a Windows computer, open the Microsoft Store from the start menu and search for “Ubuntu” and install Ubuntu 20.04 LTS (or similar name). This will allow you to run a Linux terminal on your Windows computer.  
  
**Download the command line BLAST program.**  
Go to <https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/> and click on   
ncbi-blast-2.11.0+-x64-linux.tar.gz (or similar name) to start the download. Move this file to a folder called ‘blast’ (or whatever name you’d like).   
  
**Unzip the archive and set the PATH variable.**Open up the terminal and navigate to your ‘blast’ folder. ‘cd’ is the command to change which directory you are in, and you can see what options you have by pressing the tab key after typing in ‘cd ’. If you are using Ubuntu on Windows, start by typing ‘cd /mnt/c/Users/’ and pressing tab to see your options, and make your way to the ‘blast’ folder (mine is called ‘testblast’ below).  
  
After entering the second command, you will see all the files that have been unzipped in the ‘blast’ folder. Next, make a new folder called ‘blastdb’ which will hold your database files and open up a file called ‘profile’ with the following commands.  
  
This will open up the ‘profile’ file in a text editor. If this opens up a blank file, your profile file may be called something else; try googling “where is the bash profile file”. Navigate to the bottom of the file using the arrow keys and type in the following lines (make sure you are using the correct directory names/version numbers/etc. that you have on your computer).   
Text

Description automatically generated  
Press Ctrl+S to save your changes and Ctrl+X to exit the text editor. <enter ‘source ~/.profile’>. Try entering ‘blastn’ in the command line. You should get an error message. This is to check if setting the PATH worked.  
Graphical user interface, text

Description automatically generated  
  
**Download the database you want to search.**To run BLAST on your own computer without any limitations on job size/run time, you must have the database on your computer. Entering ‘update\_blastdb.pl --showall’ will show you a list of databases available for download (it may take a few moments to connect to NCBI). ‘nt’ is the nonredundant nucleotide database. Go to your ‘blastdb’ folder and download the database by entering the following:   
  
As of May 2021, the nt database files are collectively around 200 GB. There is a way to get by with less space but it’s not ideal since you won’t be able to automatically update your copy of the database in the future. The update\_blastdb.pl script will download the database in volumes and this will take a long time. It’s okay if you can’t do it in one go. Just come back to the ‘blastdb’ folder and enter the same command as above, and the download will continue where it left off.  
  
**Create folder for Python/bash scripts.**Download the scripts and put all of them in one folder.   
Text

Description automatically generated  
Here, I created a new folder within ‘testblast’ and dragged the scripts into that folder (not shown). The ‘ls’ command simply lists the contents of the current folder.

**Run BLAST.**The setup is now complete and you can run BLAST with the following command:  
  
Enter: bash blast\_summary.sh <assembled\_contigs.fa> <blast\_results\_output.txt> <sample\_ID>

<blast\_results\_output> is any name you would like to give the raw results file. The extracted viral hits summary will be written to a file called <sample\_ID>\_blast\_summary.csv, which can be opened in Excel or any other spreadsheet program (or text editor, too).  
Text

Description automatically generated

Note: you need to ‘cd’ to your blast\_scripts folder when running the command, and make sure your  
assembled contigs file is also in the same folder.