Make sure you have python installed. Open a terminal and type or copy and paste without the brackets [python –version] If you have python it will tell you what version it is, if you don’t have python installed or it isn’t fully installed you will get an error.

[Note that most versions of MacOS come with Python 2. You should [install Python 3](https://docs.python-guide.org/starting/install3/osx/) in order to run the script. Once Python 3 is installed, you may need to use the command ‘python3’ instead of ‘python’ in order for it to run the correct version.]

Install the libraries for python: [pip install pandas] and ~~[pip install biopython]~~ i(not needed anymore) in the terminal window without the brackets. If they are already installed it would say a bunch of lines of requirement already met, if it needs to install it will spit out a bunch of stuff and then go back to the command prompt.

If you are having trouble with python, we will need to troubleshoot it. I am happy to help, but try asking Chat GPT and copy and paste the errors into it and it can help troubleshoot.

Assuming python is working:

Make a new folder in the directory where you store your GenBank uploads

Name it such that you will always know what it means i.e. CA\_FUNDIS\_PL05\_Oct\_5\_2023\_GenBank01

You can name it something shorter if you want, but if someone asks you for the upload you did for fundis plate05 in 5 years from now, you should be able to find it on your computer using the folder name alone.

Copy the python script (select the file and control c or apple c) and paste it in the folder you made.

Download the most recent iNaturalist data from the google drive and place it in that folder

Download the 2 files from Mycomap and place them in the folder. Rename them something easy to type. Make sure the data file is a .tsv and the sequence file is .fasta . If the sequence file is .fa, just rename it to say .fasta.

Open the terminal window by right-clicking on the folder and looking for a menu option that says open in terminal or something like that

Type or copy and paste into the command line without the brackets [python GB\_format6.py]

It will ask you to enter the names of the files with the extensions, so that means you have to put file\_name.tsv or other\_file.fasta whatever you name them plus the extension.

If it works, then you have your stuff to continue with.

Spot check one of the sequences and make sure it matches the record in the spreadsheet. Make sure that names are in the format Lyophyllum sp ‘CA01’ or Lyophyllum sp. ‘decastes CA01’ for any provisional names. Always use the sp. And always use single quotes around the provisional part.