***cell\_LineScrape.py***

This program is an easy to use program which allows the user to search ATCC’s 1211 human cell line products found on 122 pages. Pages 1-121 have ten entries per page and page 122 has one entry. This program creates the soup object from the ATCC cell line database, runs the user interface, creates a cells table in cell.db (if one does not currently exist), and creates a json file titled cell\_LineScrape.json (if one does not exist). This program is required for cell\_Interface.py, cell\_toDB.py, and json\_LineScrape.py to work.

Upon loading, the first page appears with the day and time of the search followed by the prompt “Enter choices separated by spaces, select page number by typing ‘p’ followed by the number (ex. p9), press enter to select all, or type ‘esc’ to exit: “. Continue searching single entries or multiple entries for name, organism, tissue, cell type, and disease until you decide to exit the program.

To run the program, change to the directory containing the file in Windows Powershell. Next, run ‘python cell\_LineScrape.py’. This program has a run time of approximately 15 minutes.

***cell\_Interface.py***

This program requires cell\_LineScrape.py. The sole purpose of this function is to run the user interface. It is an easy to use program which allows the user to search ATCC’s 1211 human cell line products found on 122 pages. This program does not create a json file or cell database. The user interface is identical to the interface of cell\_LineScrape.py.

To run the program, change to the directory containing the file in Windows Powershell. Next, run ‘python cell\_Interface.py’. This program has a run time of approximately 15 minutes.

***cell\_toDB.py***

This program requires cell\_LineScrape.py. The sole purpose of this function is to create a cells table in cell.db (if one does not currently exist). To create a new cell database, delete the original database and then run cell\_toDB.py or cell\_LineScrape.py. The cell table created by cell\_toDB.py is identical to that created in cell\_LineScrape.py.

To run the program, change to the directory containing the file in Windows Powershell. Next, run ‘python cell\_toDB.py’. This program has a run time of approximately 15 minutes.

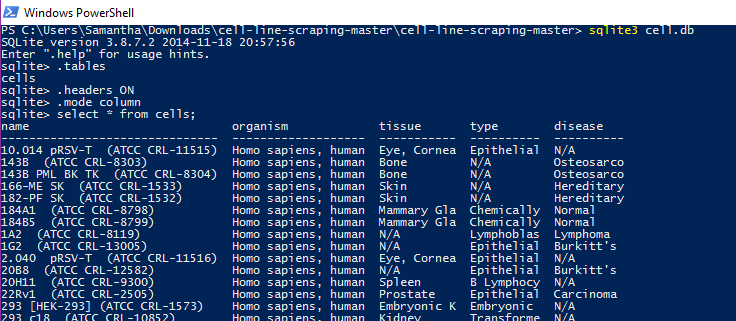
***json\_LineScrape.py***

This program requires cell\_LineScrape.py. The sole purpose of this function is export the soup object to a json file titled cell\_LineScrape.json (if one does not currently exist). To create a new json file, delete the original file and then run json\_LineScrape.py or cell\_LineScrape.py. The json file created by json\_LineScraep.py is identical to that created in cell\_LineScrape.py.

To run the program, change to the directory containing the file in Windows Powershell. Next, run ‘python json\_LineScrape.py’. This program has a run time of approximately 15 minutes.

***cell.db***

To access the cells table in cell.db in an easy to read format run:



***cell\_LineScrape.json***

To open the json file, run the following code in a Python interpreter or python code. The result is a list of dictionaries containing cell line information.

