Guide for spot_calling_batch_processing.py algorithm

Author: Austin Chiappetta, achiapp2@jh.edu

Description: this algorithm automates the processing of spots data files generated from Airlocalize. The analyzes the .loc files produced by Airlocalize and generates processed_data.txt files for each. The processed_data.txt files are tab-separated values files which list the x, y, and z pixel coordinates, the intensities, scaled intensities, and number of molecules quantified for each spot. The processed_data.txt files can then be used for downstream colocalization analysis using the colocalization_calculator.py algorithm. Also generates a summary file which lists the number of spots and molecules quantified for each file analyzed and the averages for both.
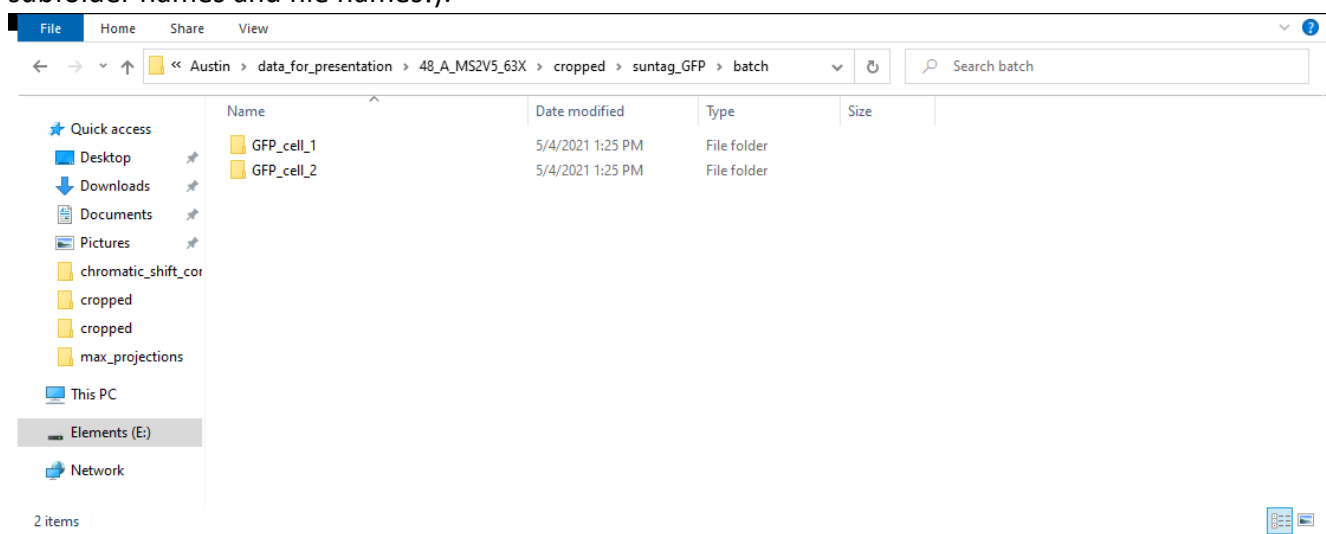
Function call usage: in the command line, type: "python spot_calling_batch_processing.py PATH/TO/FOLDER/CONTAINING/THRESHOLD/SUBFOLDERS/" (all one line, without quotation marks, as shown in tutorial below)

Note: the folder path can be either the absolute path or the relative path, and the path must contain forward slashes ("/") instead of back slashes.

VERY important! Note that none of the folders in the path can contain spaces! Please use underscore ("_") instead of spaces.

Step-by-step tutorial:

1. Organize the output files from Airlocalize batch spot calling (from a single channel, i.e., don't include data from both a GFP channel and an MCHERRY channel together. Run the algorithm separately for each channel) into subfolders which are named with a descriptive title used as shown in the pictures below (avoid the use of spaces in the subfolder names and file names!):
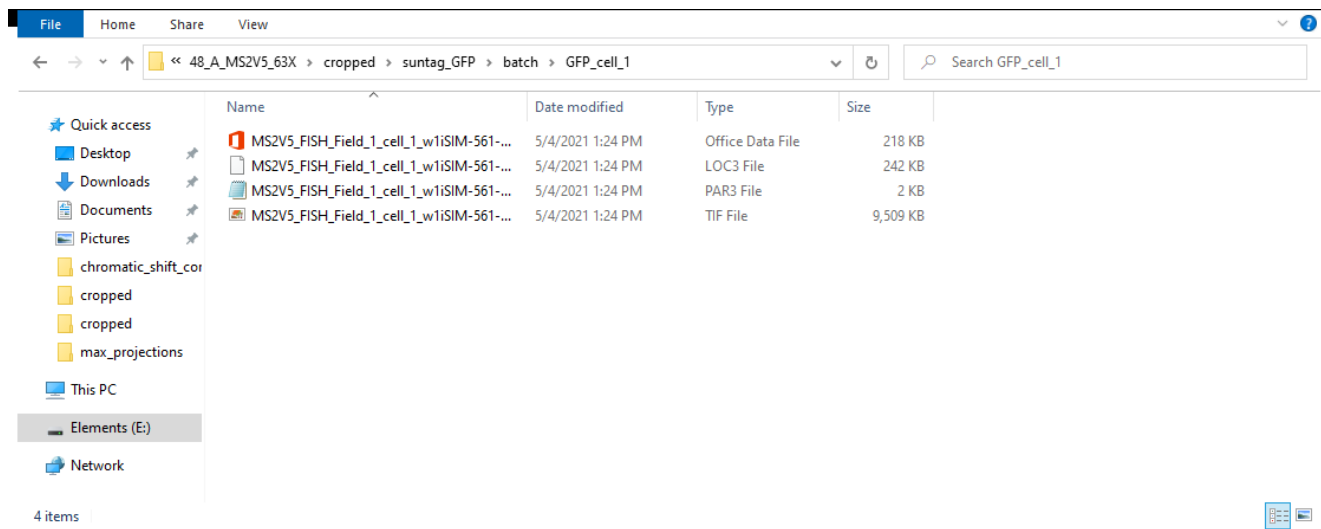
2. In the command line, navigate to the directory which contains the python script (If using Windows Git Bash, you can right click in the folder and select "Git Bash Here").

3. Type "python spot_calling_batch_processing.py PATH/TO/FOLDER/CONTAINING/THRESHOLD/SUBFOLDERS/" (all one line, without the quotation marks) into the terminal.

4. Press Enter.

5. After about 1 seconds, a message will appear in the terminal reporting the average number of spots called per field, the average number of molecules quantified per field, and the location of where the summary file has been saved. In addition, processed_data.txt files have been generated for each field and saved in each subfolder where the .loc files are.
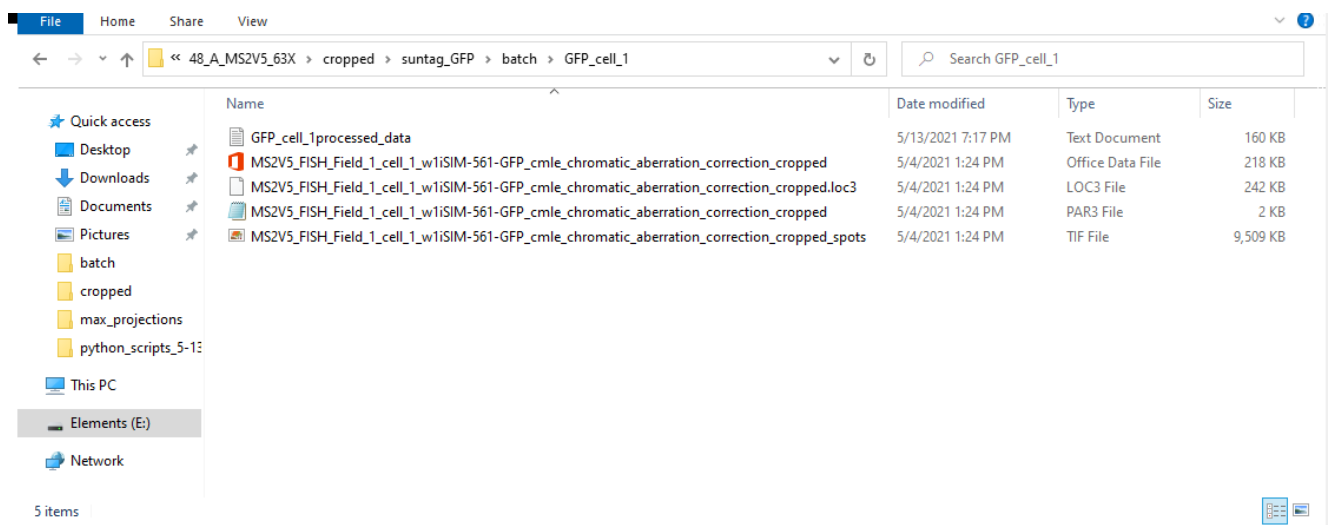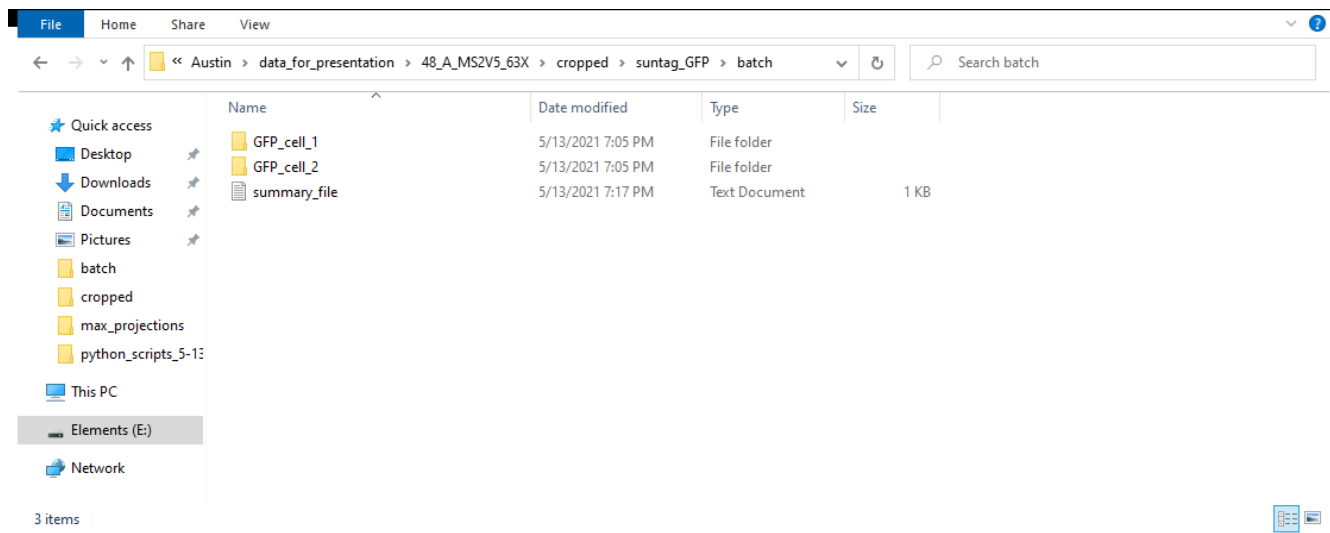
6. The end.