User Manual of EpiProfile

Garcia Lab at UPenn
(bgarci@pennmedicine.upenn.edu, zuoyuan@pennmedicine.upenn.edu)
Ref: Molecular & Cellular Proteomics 2015, 14 (6): 1696-707

Contents

A. Requirements:

1. Windows system (Windows 7 or later version, any Intel or AMD x86-64 processor, RAM with 2 GB or more).
2. Xcalibur and Matlab should be installed firstly.

B. Instructions:

1. modify the input parameters (open the folder 'EpiProfile', open the file 'paras.txt', put your data path after 'raw_path', set other parameters following the instructions).

2. start Matlab (in the folder 'EpiProfile', doubly click the file 'EpiProfile.m').

3. run EpiProfile (in the Matlab Command Window input “EpiProfile” and press “Enter”).

The results are under the data path (histone_layouts and histone_ratios.xls, or histone_ratios_SILAC.xls, or histone_ratios_C13.xls, or histone_ratios_N15.xls, or histone_ratios_13CD3.xls).

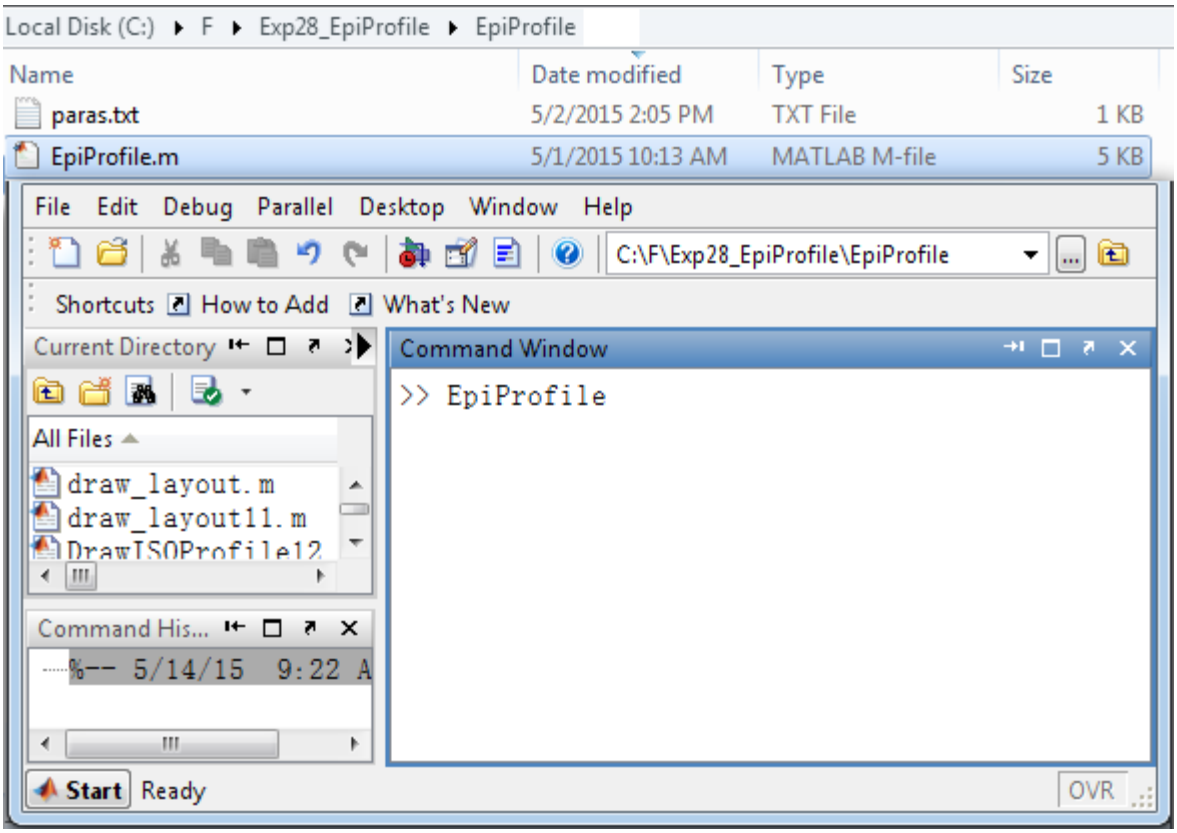
C. Steps:

Step 1 (modify the input parameters)

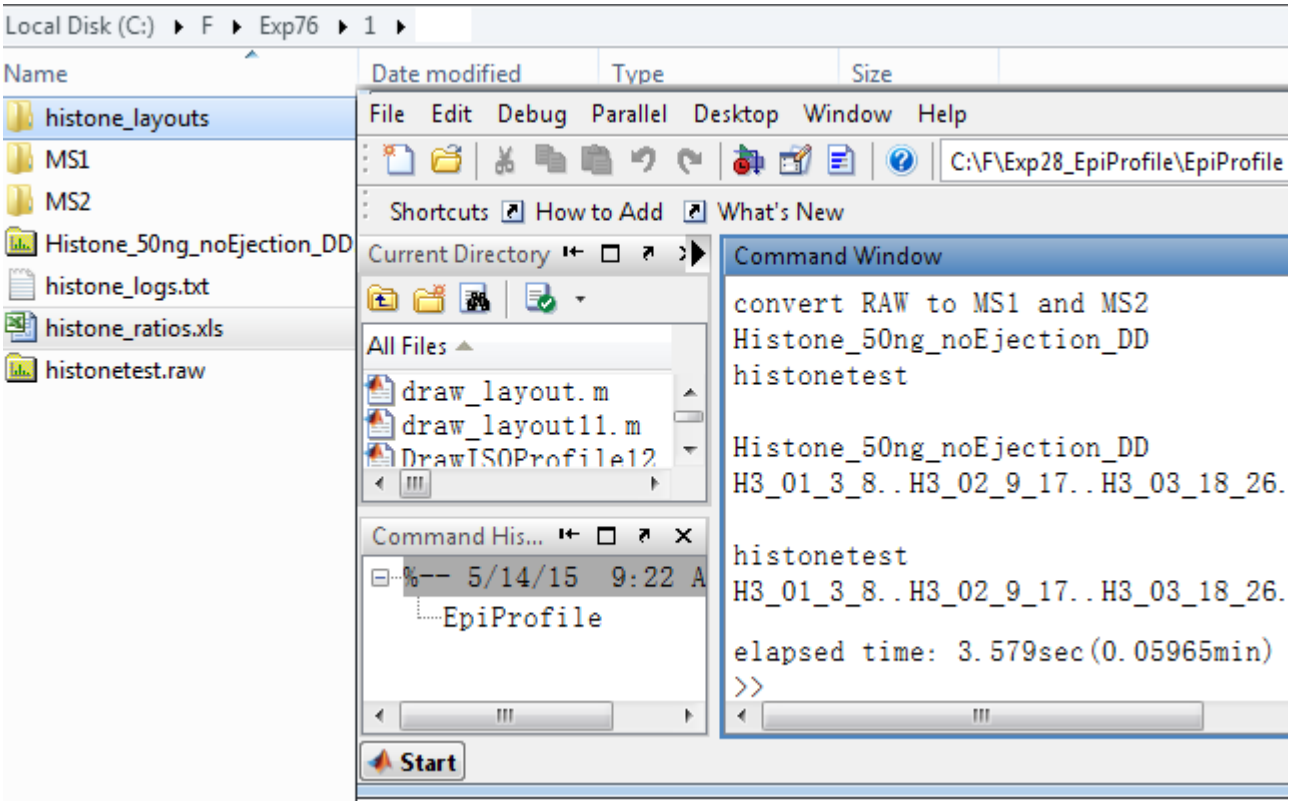
```
[EpiProfile]
% the datapath of raw files
raw_path=C:\F\Exp76\1

% 1: Human, 2: Mouse
norganism=1

% 1: histone_normal, 2: histone_SILAC, 3: histone_C13, 4: histone_N15, 5: histone_13CD3
nsource=1

% if histone_N15, 0: N14 light Mods, 1: N15 light Mods, 2: N14 heavy Mods, 3: N15 heavy Mods, 4: 0+1, 5: 0+3
nsubtype=0
```

Step 2&3 (start Matlab and run EpiProfile)

Start and Run



Results



D. Advanced options

Note: these are advanced options which might need the author's help (zuoyuan@pennmedicine.upenn.edu).

1. Some options in 'paras.txt'. If "nsource=2", set "nsubtype=0" for SILAC of light and heavy R, and set "nsubtype=1" for SILAC of light and heavy K and R. "nsource=3" means two (2) 13C on acetylation. "nsource=4" and if "nsubtype" contains 15N means all N in each amino acid is labeled by 15N. "nsource=5" means 13CD3 on methylation and Methionine.

2. How to set the peptide mass tolerance? In 'check_otherparas.m', the default setting is "def_ptol = 10", which means 10 ppm. If the mass tolerance is shifted to 20 or 30 ppm, then the value of "def_ptol" should be changed.

3. How to check phosphorylation on H3? In 'check_otherparas.m', the default setting is "soutput = '11'", which means no phosphorylation on H3. If H3 S10ph is needed, it should be "soutput = '21'". If H3 S10ph and S28ph are needed, it should be "soutput = '31'".

4. Whether or not to use reference retention time? In 'check_otherparas.m', the default setting is "ndebug = 0'", which means the reference retention time will be extracted first. If the RAW files are generated by the same instruments in near days, the same peptide in different RAW files will elute at close time (e.g. 2 or 3 mins shift). However, if the RAW files are generated in different instruments or in different days (e.g. the interval is several weeks or months), there are two solutions: (1) keep "ndebug = 0'", divide the RAW files to groups by instruments and days, and run each group individually; (2) set "ndebug = 2'", keep the RAW files in one folder, run each RAW file individually without reference retention time.

5. How to do manual validation? It needs *.lyt files which are the layouts applied to RAW files. The *.lyt files can be downloaded from https://github.com/zfyuan/EpiProfile2.0_Family/histone lyt files.zip. Then you can open RAW files, apply these *.lyt files, and compare to layout files generated by EpiProfile.

6. How to correct wrong extracted ion chromatographic (XIC) peaks? If wrong XIC peaks are found by manual validation, then they can be corrected through: (1) set "ndebug = 1'" in 'check_otherparas.m', (2) put the correct retention time into "His.rt_ref" in the corresponding *.m file (e.g. H3_01_3_8.m), (3) in the folder of "histone_layouts" and under the specific RAW file search the peptide (e.g. H3_01_3_8) and delete the searched files, and (4) rerun EpiProfile. The correction is too complex for general users, so please contact the author!!!

7. How to add a new PTM? In 'GetMods.m', the form and mass of a PTM can be added (e.g., cr is on K with the mass of 68.026215). In the corresponding *.m file (e.g. H3_02_9_17.m), the new PTM (e.g. K9cr) can be added into the function of 'init_histone', 'calculate_layout', 'relocate', and 'relocate2'. Especially, in the function of 'init_histone', 'relocate', and 'relocate2', it is better to determine the retention time relationship between the ned PTM and propionylation (e.g. cr elutes later than pr). Again, it is better to ask the author to add a new PTM!!!

E. Trouble shooting

1. The folder of 'EpiProfile' cannot be put under a path contains a space (e.g. C:/my folder), through the data path can contain a space.
2. If Xcalibur is installed but EpiProfile cannot use Xcalibur to convert RAW to MS1 and MS2, then MSFileReader need be installed, which can be downloaded from https://thermo.flexnetoperations.com.
3. In addition to MATLAB, EpiProfile2.0 uses the following 3 toolboxes: Statistics and Machine Learning Toolbox, Curve Fitting Toolbox, and Bioinformatics Toolbox.

4. The application file ‘xtract.exe’ in the folder of 'EpiProfile' will expire on the end day of each year. Therefore, it needs be updated in the beginning of each year.