General Information

This README file pertains to code hosted at:

https://github.com/uomsystemsbiology/epidermal_data

It is designed to run for the analysis of data hosted on GigaDB, the GigaScience data repository:

Cursons, J; Angel, C, E; Hurley, D, G; Print, C, G; Dunbar, P; Jacobs, M, D; Crampin, E, J (2015): Supporting data for "Spatially-transformed fluorescence image data for ERK-MAPK and selected proteins within human epidermis". *GigaScience Database*.

http://dx.doi.org/10.5524/100168

Detailed technical information on data collection can be found within the corresponding Data Note:

Cursons *et al.* (2015). Spatially-transformed fluorescence image data for ERK-MAPK and selected proteins within human epidermis. *GigaScience*. Accepted Nov 2015.

doi: to-be-known

These MATLAB scripts have dependencies for several MATLAB Toolboxes and MATLAB File Exchange functions, as detailed in the Script/Function Overview (*p*2). Users who can accept the MathWorks Academic Usage Agreement, these scripts are available as a Virtual Reference Environment

For further information please contact The University of Melbourne Systems Biology Laboratory:

Edmund Crampin - edmund.crampin@unimelb.edu.au

Joe Cursons - joseph.cursons@unimelb.edu.au

Last Modified: 10/11/15

Table of Contents

Script/Function Overview

Within the base code folder, there are a number of scripts:

/code/sample_to_loessDiscCentroids.m

This MATLAB script reads in the sample location data, extracts pixel intensities from the raw image data, maps to the normalised distance co-ordinate, performs loess smoothing, and outputs .csv/.tiff at various stages.

Described in further detail on page 3

/code/recreate_pMEK_heterogeneity_fig.m

This MATLAB script reproduces Fig. AF4.1 of the GigaScience Data Note described above (NB: this output was tidied in GIMP to produce the final figure).

Described in further detail on page 5

There are also a number of functions used by the scripts above:

/functions/calculateDistancesToBoundary.m

/functions/calculateLoessCurve.m

/functions/calculateSpatialDivisions.m

/functions/convertSmpAnaToAbundVector.m

/functions/extractSampleInformation.m

/functions/initSamplingKernel.m

/functions/loadImageStack.m

/functions/loadImageStackAsSparse3DCellArrays.m

/functions/produceSmpAna.m

/functions/rescaleNormDist.m
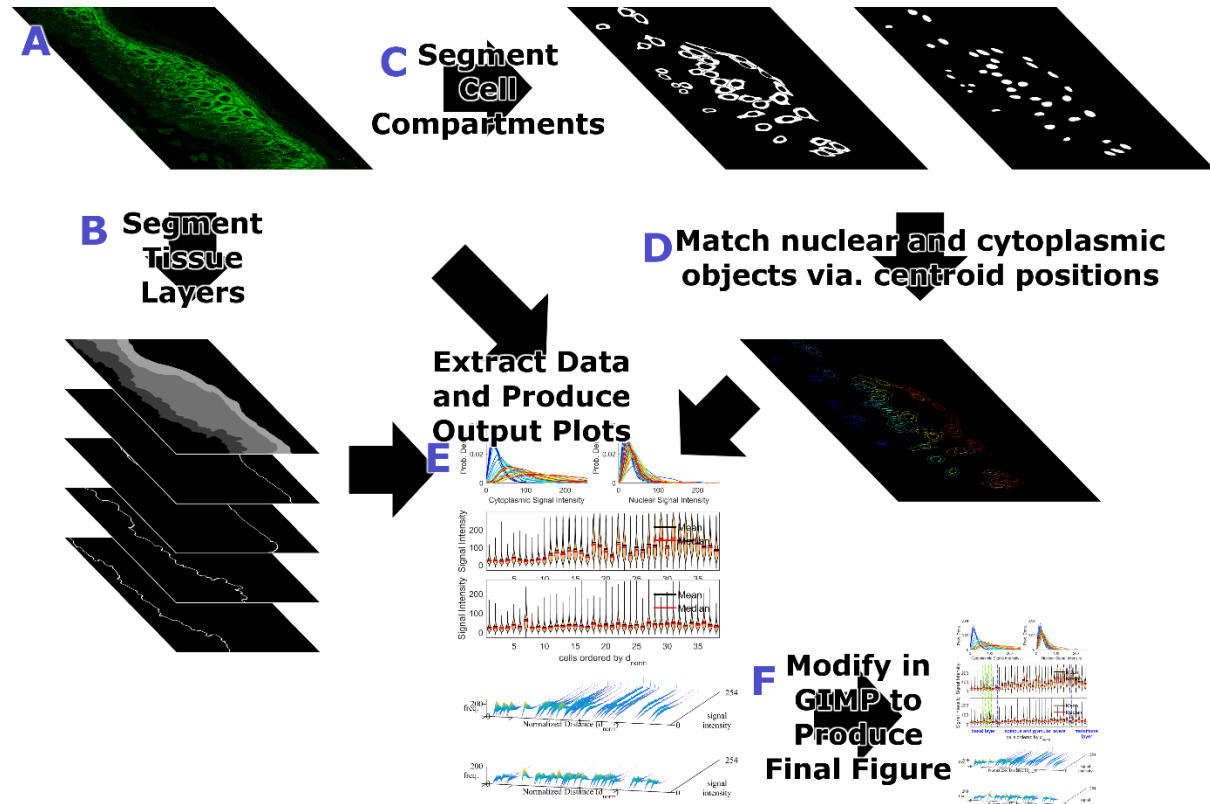
sample_to_loessDiscCentroids.m

This MATLAB script reads in the sample location data, extracts pixel intensities from the raw image data, maps to the normalised distance co-ordinate, performs loess smoothing, and outputs .csv/.tiff at various stages. A graphical overview of this script, the functions used, and the outputs at each stage is given below:

Dependencies:

Place holder

recreate_pMEK_heterogeneity_fig.m

This MATLAB script reproduces Fig. AF4.1 of the GigaScience Data Note noted above [<mark>ref-not-yet-created</mark>]. A Graphical overview of this script is given with cross-references to the corresponding lines of code within `recreate_pMEK_heterogeneity_fig.m`:

Lines in Script	Description
	(**A**) Image data were loaded into GIMP for segmentation of (**B**) tissue layers and (**C**) sub-cellular compartments (*note that segmentation masks have been provided for this figure and these can be easily modified by users wishing to perform their own analysis*).
115 - 185	(**D**) Centroids of segmented cytoplasm and nuclei were identified using functions from the image processing toolbox, and the nearest cytoplasm/nuclei were matched using the `dist` function from the Neural Network Toolbox
150 - 180	Segmented objects were tracked using the Image Processing Toolbox, allowing signal intensity data at specific locations to be extracted and tracked.
185 - 510	The nucleus centroid of each cell was mapped on to the normalised-distance co-ordinate [1], and (**E**) the distribution of pixel intensities is displayed using: probability density functions (*line color reflects relative tissue depth*); violin plots (*ranked horizontally by relative tissue depth*); and 'histogram surface plots' (*with probability density functions mapped on to the normalised distance co-ordinate*). **NB**: If executing this script through the Virtual Reference Environment alpha-layers/transparency cannot be used due to issues with software OpenGL rendering.
	(**F**) To create Fig. 4A of the GigaScience Data Note, the output image underwent minor modifications within GIMP (*removing overlap of axis labels etc*).

Dependencies:

This script executes several functions with MATLAB ToolBox dependencies:

- Image Processing Toolbox
    - `bwlabel, regionprops, bwmorph`
    - `calculateDistancesToBoundary` → `bwlabel, regionprops`
- Neural network Toolbox
    - `dist`
- Statistics and Machine Learning Toolbox
    - `ksdensity`

This script also executes `violin` from the MATLAB File Exchange (File ID: #45134), created by Holger Hoffman: hhoffmann@uni-bonn.de
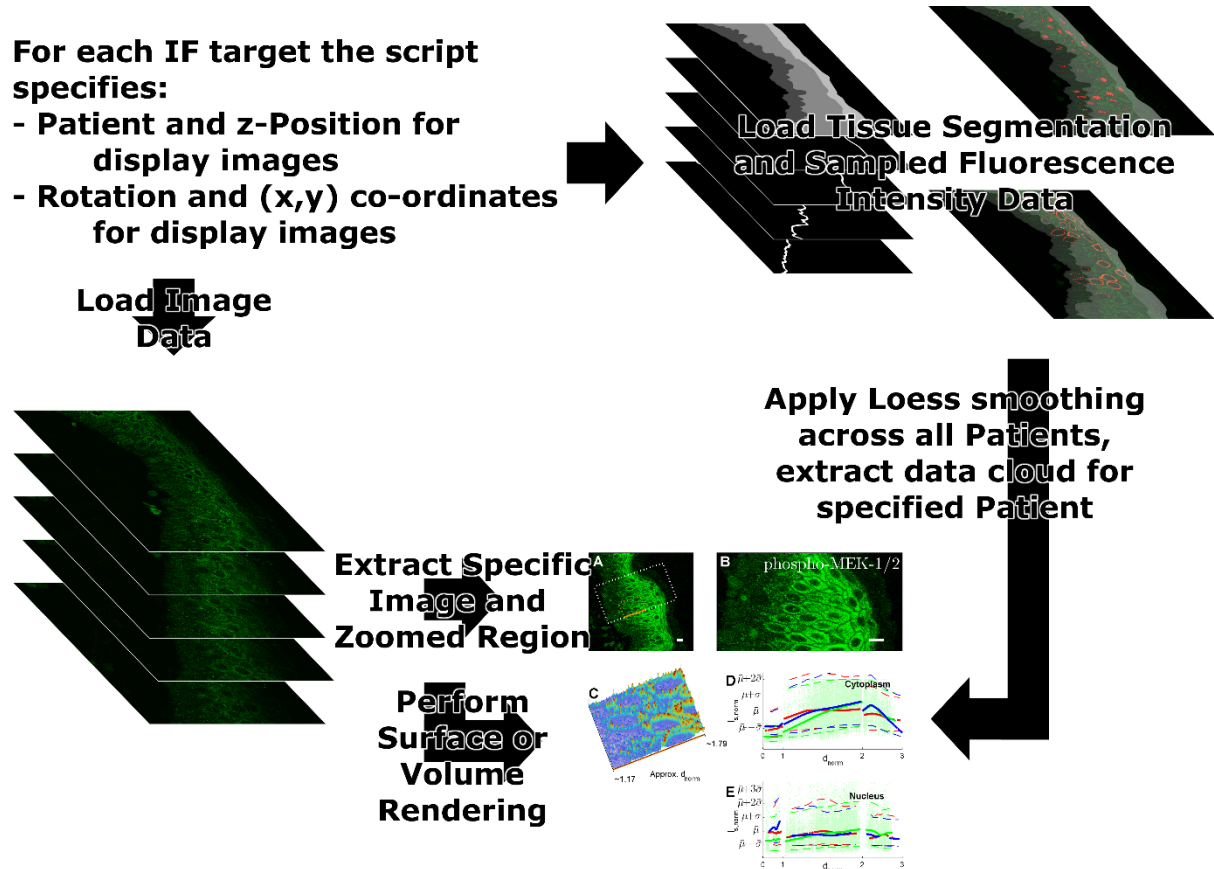
- http://www.mathworks.com/matlabcentral/fileexchange/45134-violin-plot-based-on-kernel-density-estimation

create_IF_data_summary.m

This MATLAB script creates figures (one per immunofluorescence target) which give a summary of the image and sampled data. These figures are used within Additional File 3 of the associated GigaScience Data Note [ref-not-yet-created].

For each IF target the script specifies:
- Patient and z-Position for
    display images
- Rotation and (x,y) co-ordinates
    for display images

Load Image Data

Load Tissue Segmentation and Sampled Fluorescence Intensity Data

Apply Loess smoothing across all Patients, extract data cloud for specified Patient

Extract Specific Image and Zoomed Region

Perform Surface or Volume Rendering

At the top of this script there are a large number of arrays and variables which define the specific information to be displayed within the data summary file:

Dependencies:

This script executes several functions with MATLAB ToolBox dependencies:

References

1. Cursons J, Gao J, Hurley DG, Print CG, Dunbar PR, Jacobs MD et al. Regulation of ERK-MAPK signaling in human epidermis. BMC systems biology. 2015;9:41. doi:10.1186/s12918-015-0187-6.