Annotate Cell-Level MEP-LINCs Data

Mark Dane

2015-07-16

Summary

This script prepares cell-level data and metadata for the LINCs Analysis Pipeline.

In the code, the variable ss determines which staining set (SS1, SS2 or SS3) to merge. All .txt data files in the ./Cell/Raw Data folder will be merged with the well (xlsx) and log (XML) data from the ./Metadata folder.

This merging assumes that the actual, physical B row wells (B01-B04) have been printed upside-down. That is, rotated 180 degrees resulting in the spot 1,1 being in the lower right corner instead of the upper left corner. The metadata is matched to the actual printed orientation.

The data is filtered to remove objects with area less than 1000 pixels.

In addition to merging the metadata with the cell-level data, several types of derived parameters are added. These include:

Cell Density based on the number of nuclear centers within a radius around each nucleus. Normalized intensity values calculated by dividing each intensity value by the median of the corresponding intensity value in the control well of the same plate.

Some parameters are clustered or gated into two or more classifications. For example, cells are classified as EdU+ or EdU-, 2N and less or 4N and more DNA. Each cell does not have every derived parameter as each staining set has unique endpoints.

The data is summarized by the median of the parameter at each spot. Then additional derived values are added to this spot-level dataset:

The Spot Cell Count for the number of cells at each spot Loess models of the spot cell count that capture regional information about cell seeding.

The staining sets are summarized by the medians of their normalized replicate values then combined to the MEP-level within each cell line.