



## Vizgen MERFISH Mouse Brain Receptor Map

### Data Set Description:

This dataset contains a MERFISH measurement of a gene panel containing 483 total genes including canonical brain cell type markers, GPCRs, and RTKs measured on 3 full coronal slices across 3 biological replicates. Below are details describing what you will find within each data file.

#### 1. Information on all transcripts detected in the sample (List of Detected Transcripts)

- These .csv files (detected\_transcripts.csv) contain the list of all detected transcripts in the sample where each row is a detected transcript. The column names and descriptions of what the column represents are:
  - A. barcode\_id - internally used id of the gene. Each gene has a unique barcode\_id
  - B. global\_x, global\_y - The global micron x and y coordinates of this transcript
  - C. x, y - The pixel coordinates of this transcript within the field of view
  - D. global\_z - The index of the z slice that this transcript was detected in. Each z slice is separated by 1.5 microns
  - E. fov - the index of the field of view where this transcript was detected
  - F. gene - the gene name of this detected transcript

#### 2. Single cell information from cell segmentation (Cell by Gene Matrix)

- Cell by gene count matrix
- These .csv files (cell\_by\_gene.csv) contain the count of detected transcripts within the segmentation boundaries for each cell for each gene. Each row corresponds to a cell and the first column contains the cells unique ID. Each remaining column corresponds to a gene. This matrix is not filtered for segmentation artifacts and we recommend removing cells with fewer than 50 total counts.
- B. Cell spatial metadata
- These .csv files (cell\_metadata.csv) contain the spatial metadata for each of the detected cells. Each row corresponds to a cell. The first column is the unique cell ID and the remaining columns are:
  - A. fov - the field of view containing the cell
  - B. volume - The volume of the cell in  $\mu\text{m}^3$
  - C. center\_x - the x coordinate of the center of the cell in global micron coordinates
  - D. center\_y - the y coordinate of the center of the cell in global micron coordinates
  - E. min\_x, max\_x - the minimum and maximum of the bounding box containing this cell in x in global micron coordinates
  - F. min\_y, max\_y - the minimum and maximum of the bounding box containing this cell in y in global micron coordinates

## C. Cell boundary polygons

- These .hdf5 files (cell\_boundaries/feature\_data\_X.hdf5) contain the polygon boundaries for cells identified in field of view X. If no cells were detected within field of view X then there will be no file feature\_data\_X.hdf5. The boundaries are stored in hdf5 format indexed by the unique cell ID.

## 3. Mosaic images of the full tissue

- These .tif files (images/mosaic\_X\_zY.tif) contain mosaic images of additional stains where X indicates the name of the stain (DAPI, polyT) and Y is the z index of the image. The mosaic image is a composite of all fields of view.
- These .csv files (images/micron\_to\_mosaic\_pixel\_transform.csv) contain the transformation matrix for converting from real-world micron coordinates to pixels in the mosaic images.

Questions? Contact [info@vizgen.com](mailto:info@vizgen.com) for support.