

SMI Data File ReadMe

Transcript File

fov	cell_ID	cell	x_local_px	y_local_px	x_global_px	y_global_px	z	target	CellComp
1	80	c_1_1_80	761	730	101383	730	6	AQP4	None
1	80	c_1_1_80	714	731	101430	731	4	ATP9A	None
1	80	c_1_1_80	730	731	101414	731	5	PLP1	None
1	80	c_1_1_80	740	731	101404	731	3	NBL1	None
1	80	c_1_1_80	744	731	101400	731	6	ST8SIA1	None
1	80	c_1_1_80	717	732	101427	732	4	SCARB2	None
1	80	c_1_1_80	717	732	101427	732	5	RBM5	None
1	80	c_1_1_80	747	732	101397	732	2	MBP	None
1	80	c_1_1_80	760	732	101384	732	2	CYP2R1	None
1	80	c_1_1_80	706	733	101438	733	5	ABCA2	None

- Column descriptions:
 - fov
 - The Field Of View (FOV) the transcript is in. All FOVs from the sample are contained within this transcript file.
 - cell_ID
 - Unique identifier for a single cell within a given FOV. To make a unique identifier for a cell within the whole sample use both the "fov" and "cell_ID" columns. Transcripts not assigned to cells have a value of 0.
 - x_local_px
 - The x position of this transcript within the FOV, measured in pixels. To convert to microns multiply the pixel value by 0.12 um per pixel. For FOV-local coordinates, the origin (0,0) is the top left of the FOV.
 - y_local_px
 - Same as "x_local_px" but for the y dimension. Due to differences in Cartesian and imaging plotting frames of reference, the local y-axis coordinates should be inverted to align the transcript positions to images.
 - x_global_px
 - See "x_local_px" description. The global positions describes the relative position of this transcript within the large sample reference frame. For global coordinates, the origin (0,0) is the bottom left of the capture area.
 - y_global_px
 - Same as "x_global_px" but for the y dimension
 - z
 - The Z-slice image number. This has an arbitrary unit.
 - target
 - The gene detected. HUGO gene symbols are used.
 - CellComp
 - The subcellular compartment (Nuclear, Membrane, Cytoplasmic, or None) in which the transcript was detected as determined by cell segmentation algorithm. In this dataset, subcellular locations were not calculated.

Cell Expression File

S3_exprMat_file.csv

	fov	cell_ID	cell	A1BG	A2M	AAAS	AAK1	NegPrb1	NegPrb10	NegPrb11	FalseCode1	FalseCode10	FalseCode100
2	3	86	c_1_3_86	1	0	1	1	0	0	0	0	0	0
3	3	87	c_1_3_87	0	0	0	0	0	0	0	0	0	0
4	3	88	c_1_3_88	0	0	0	0	0	0	0	0	0	0
5	3	89	c_1_3_89	0	0	0	0	0	0	0	0	0	0
6	3	90	c_1_3_90	3	0	0	0	0	0	0	0	0	0
7	3	91	c_1_3_91	1	0	0	0	0	0	0	0	0	0
8	3	92	c_1_3_92	1	0	0	0	0	0	0	0	0	0
9	3	93	c_1_3_93	0	0	0	0	0	0	0	0	0	0
10	3	94	c_1_3_94	0	0	0	0	0	0	0	0	0	0
11	3	95	c_1_3_95	0	0	0	0	0	0	0	0	0	0

- Column descriptions:
 - fov
 - The Field Of View (FOV) the transcript is in. All FOVs from the sample are contained within this Cell Expression file.

- **cell_ID**
 - Unique identifier for a single cell within a given FOV. To make a unique identifier for a cell within the whole sample use both the "fov" and "cell_ID" columns. All transcripts not assigned to a cell are show with a "cell_ID" value of 0.
- **Gene expression target (AATK, ABL1, ..., etc)**
 - The number of transcripts observed for a given gene target for a given cell.
- **Negative probes (i.e. NegPrb##)**
 - Negative probes, which do not match any sequence within the transcriptome or genome, are included. These can be used to assess background levels.
- **False codes (i.e. FalseCode##)**
 - False codes, which represent barcodes that are not included in the probe mix.

Cell Metadata File

S3_metadata_file.csv

slide_ID	Run_Tissue_name	fov	cell_ID	AspectRatio	CenterX_local_px	CenterY_local_px	CenterX_global_px	CenterY_global_px	Width	Height	Area	Mean.Histone	Max.Histone	Mean.G	Max.G	Mean.rRNA	Max.rRNA	Mean.GFAP	Max.GFAP	Mean.DAPI	Max.DAPI	cell
1	S3	1	2	1.64	492	34	492.07794	93597.99	113	69	5760	2190	8824	98	1440	4342	15388	7123	55904	406	1264	c_1_1_2
1	S3	1	3	0.90	734	38	734.11628	93593.99	69	77	4433	22881	45304	21	344	1465	7948	1522	37744	548	1308	c_1_1_3
1	S3	1	4	1.98	2566	21	2566.40651	93610.99	85	43	2743	9047	29624	3	288	1151	7148	891	21056	119	440	c_1_1_4
1	S3	1	5	2.18	3358	19	3358.53198	93612.99	85	39	2609	13435	38648	48	896	2175	6812	9699	46224	111	512	c_1_1_5
1	S3	1	6	0.94	3869	69	3869.61293	93562.99	132	140	7551	9807	43080	222	1504	2113	7164	3246	33648	154	680	c_1_1_6
1	S3	1	7	0.91	3812	58	3812.60390	93573.99	105	116	7657	8957	28200	215	4120	1406	4536	2222	25472	216	720	c_1_1_7
1	S3	1	8	1.00	66	78	66.01046	93553.98	77	77	4565	6881	14648	73	956	1342	8444	2606	24848	606	1260	c_1_1_8
1	S3	1	9	0.82	296	75	296.04689	93556.98	45	55	1955	261	1364	23	356	288	4784	13069	55904	409	1032	c_1_1_9
1	S3	1	10	0.85	138	91	138.02186	93540.98	61	72	2824	13185	26168	49	2860	1441	5280	2711	41696	457	992	c_1_1_10
1	S3	1	11	2.06	4189	111	4189.66362	93520.98	132	64	5896	9077	20344	245	1344	5583	28780	922	14800	293	712	c_1_1_11

- **Column descriptions:**
 - **slide_ID**
 - The slide the cell is on.
 - **Run_Tissue_name**
 - An annotation for the tissue.
 - **fov**
 - The Field Of View (FOV) the cell is in. All FOVs from the sample are contained within this Cell Metadata file.
 - **cell_ID**
 - Unique identifier for a single cell within a given FOV. To make a unique identifier for a cell within the whole sample use both the "fov" and "cell_ID" columns.
 - **Aspect Ratio**
 - width divided by height.
 - **CenterX_local_px**
 - The x position of this transcript within the FOV, measured in pixels. The pixel edge length is 120nm. Thus, to convert to microns multiply the pixel value by 0.12 um per pixel. For FOV-local coordinates, the origin (0,0) is the top left of the FOV.
 - **CenterY_local_px**
 - Same as "CenterX_local_px" but for the y dimension. Due to differences in Cartesian and imaging plotting frames of reference, the local y-axis coordinates should be inverted to align the cell positions to images.
 - **CenterX_global_px**
 - See "CenterX_local_px" description. The global positions describes the relative position of this transcript within the large sample reference frame. For global coordinates, the origin (0,0) is the bottom left of the capture area.
 - **CenterY_global_px**
 - Same as "CenterX_global_px" but for the y dimension
 - **Width**
 - Cell's maximum length in x dimension (pixels)
 - **Height**
 - Cell's maximum length in y dimension (pixels)
 - **Area**
 - Number of pixels assigned to a given cell.
 - **Mean.Histone**
 - Mean fluorescence intensity within a given cell.
 - **Max.Histone**
 - Maximum fluorescence intensity within a given cell.
 - **Mean.G**
 - Mean fluorescence intensity within a given cell.
 - This channel was left blank in this experiment.
 - **Max.G**
 - Maximum fluorescence intensity within a given cell.
 - **Mean.rRNA**
 - Mean fluorescence intensity within a given cell.
 - **Max.rRNA**
 - Maximum fluorescence intensity within a given cell.
 - **Mean.GFAP**
 - Mean fluorescence intensity within a given cell.
 - **Max.GFAP**
 - Maximum fluorescence intensity within a given cell.
 - **Mean.DAPI**
 - Mean fluorescence intensity within a given cell.

- **Max.DAPI**
 - Maximum fluorescence intensity within a given cell.
- **cell**
 - Identifier for the cell that comprises the slide_ID, fov, and cell_ID

FOV Positions File

This file is used to reference each FOV within the global structure of the sample.

S3_fov_positions_file.csv

Slide	FOV	X_mm	Y_mm
1	1	12.28598	0.0000000
1	2	12.28598	0.5119162
1	3	12.28598	1.0238316
1	4	12.28598	1.5357473
1	5	12.28598	2.0476630
1	6	12.28598	2.5595787
1	7	12.28598	3.0714943
1	8	12.28598	3.5834100
1	9	12.28598	4.0953257
1	10	12.28598	4.6072414

- **Column descriptions:**
 - **Slide**
 - The slide number.
 - **FOV**
 - The Field Of View (FOV) number.
 - **X_mm**
 - The relative x position of this the FOV, measured in millimeters. The FOV coordinate corresponds to the top right of the FOV.
 - **Y_mm**
 - See "X_mm" description, but in for the y dimension.

Cell Polygons File

This file is a simple polygon descripton of cell boundaries. These are provided as a graphical convenience but are only approximations of the cell boundries, which are fully defined by the cellLabels mask images.

S3-polygons.csv

fov	cellID	cell	x_local_px	y_local_px	x_global_px	y_global_px
1	2	c_1_1_2	542	0	101602	0
1	2	c_1_1_2	545	1	101599	1
1	2	c_1_1_2	547	3	101597	3
1	2	c_1_1_2	548	6	101596	6
1	2	c_1_1_2	548	22	101596	22
1	2	c_1_1_2	543	37	101601	37
1	2	c_1_1_2	539	45	101605	45
1	2	c_1_1_2	525	59	101619	59
1	2	c_1_1_2	509	67	101635	67
1	2	c_1_1_2	506	68	101638	68

- **Column descriptions:**
 - **fov**
 - The Field Of View (FOV) number.
 - **cellID**
 - Unique identifier for a single cell within a given FOV. To make a unique identifier for a cell within the whole sample use both the "fov" and "cell_ID" columns.
 - **cell**
 - Identifier for the cell that comprises the slide_ID, fov, and cell_ID
 - **x_local_px**
 - The x position of this transcript within the FOV, measured in pixels. To convert to microns multiply the pixel value by 0.12 um per pixel. For FOV-local coordinates, the origin (0,0) is the top left of the FOV.
 - **y_local_px**
 - Same as "x_local_px" but for the y dimension. Due to differences in Cartesian and imaging plotting frames of reference, the local y-axis coordinates should be inverted to align the cell positions to images.
 - **x_global_px**

- The relative x position of this the FOV, measured in pixels. To convert to microns multiply the pixel value by 0.12 um per pixel.
For global coordinates, the origin (0,0) is the bottom left of the capture area.
- **y_global_px**
 - See "x_global_px" description, but in for the y dimension.