

# Visium HD preprint datasets

<https://www.10xgenomics.com/products/visium-hd-spatial-gene-expression/dataset-human-crc>

## Characterization of immune cell populations in the tumor microenvironment of colorectal cancer using high definition spatial profiling

### Dataset downloads

#### Files

[Visium HD, Sample P1 CRC](#)

[Visium HD, Sample P2 CRC](#)

[Visium HD, Sample P5 CRC](#)

[Visium HD, Sample P3 NAT](#)

[Visium HD, Sample P5 NAT](#)

[Visium CytAssist v2, Sample P2 CRC](#)

[Visium CytAssist v2, Sample P3 NAT](#)

[Visium CytAssist v2, Sample P5 NAT](#)

[Xenium In Situ, Sample P1 CRC](#)

[Xenium In Situ, Sample P2 CRC](#)

[Xenium In Situ, Sample P5 CRC](#)

[Chromium Single Cell Flex, aggregated](#)

[Chromium Single Cell 5' Immune Profiling, aggregated](#)

## Visium HD, Sample P1 CRC

### Overview

This dataset is associated with the preprint titled *Characterization of immune cell populations in the tumor microenvironment of colorectal cancer using high definition spatial profiling*. See the preprint for full details on methods and results.

### Tissue sectioning

5  $\mu$ m sections were taken from the FFPE tissue blocks with a microtome (Epredia HM355S). Sections were adjacent or near-adjacent (within 5-10  $\mu$ m of each other). Sectioning followed the Visium CytAssist Spatial Gene Expression for FFPE – Tissue Preparation Guide (CG000518, Rev C).

### Library preparation & sequencing

We placed FFPE tissue sections on plain glass slides for deparaffinization, H&E staining and imaging following the Visium HD FFPE Tissue Preparation Handbook (CG000684). Probe hybridization, probe ligation, slide preparation, probe release, extension, library construction, and sequencing followed the Visium HD Spatial Gene Expression Reagent Kits User Guide (CG000685). Sequencing was performed on an Illumina NovaSeq 6000 with paired-end reads (43 cycles Read 1, 10 cycles i7, 10 cycles i5, 50 cycles Read 2).

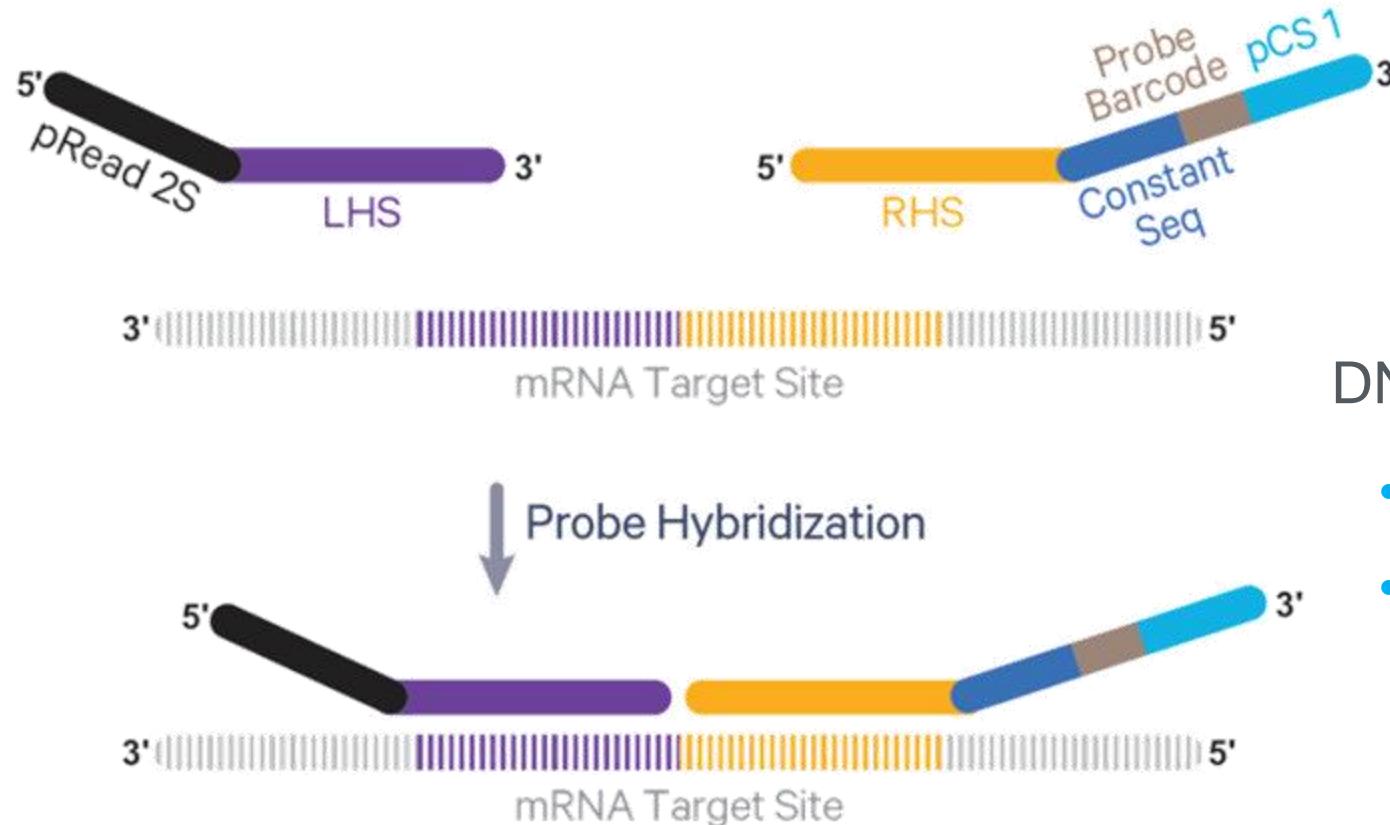


# Chromium Single Cell FLEX Data

---

# Fixed RNA Profiling uses probes to detect mRNA

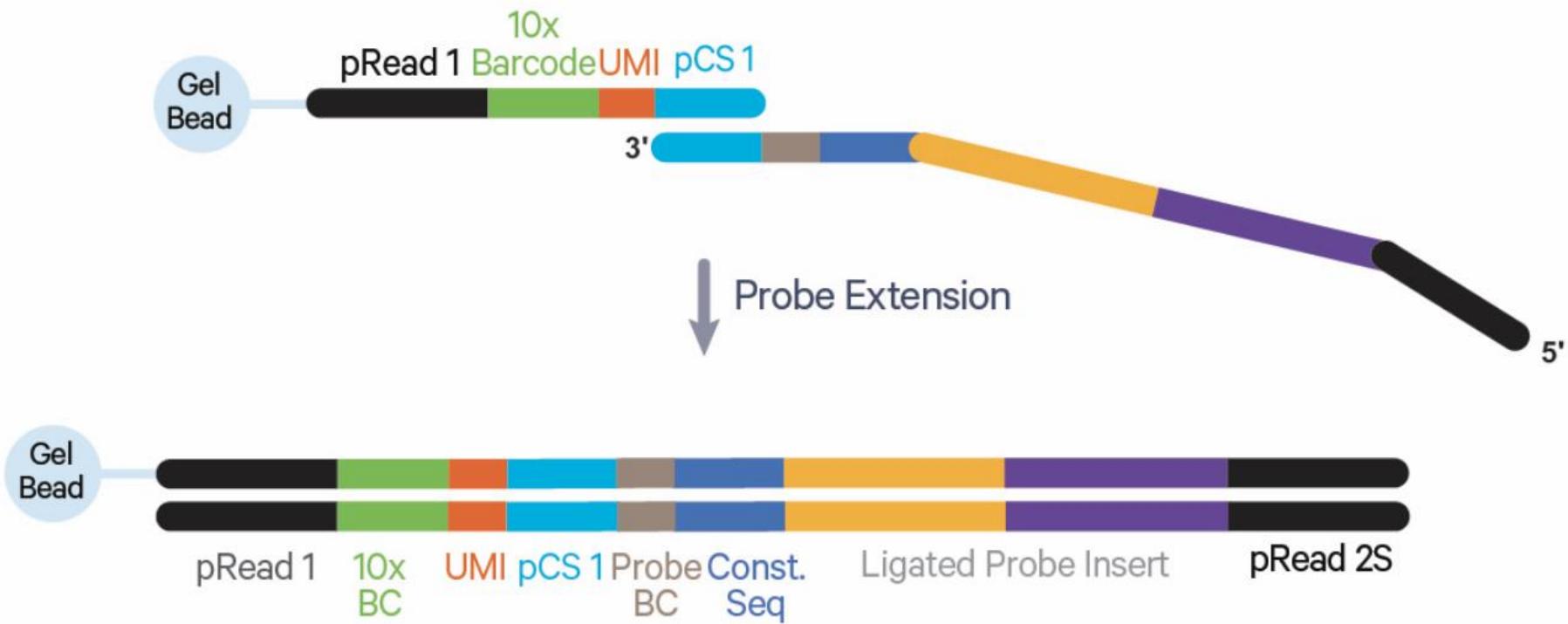
- Sample fixation is a common practice, but RNA tends to become fragmented after fixation
- RT-based chemistries performs poorly on fragmented and crosslinked RNA



## DNA-based Probes

- LHS (25 bp) has partial Read 2 seq
- RHS (25 bp) has partial capture seq 1, probe barcode, and constant seq

# Fixed RNA Profiling uses probes to detect mRNA



- Probe-based assay is neither a 5' nor 3' assay
- No polyA capture
- **It is the ligated probes being sequenced, not cDNA from RNA**
- No classes of intronic, intergenic, or antisense reads as in regular GEX analysis

# Chromium single cell data

<https://www.10xgenomics.com/products/visium-hd-spatial-gene-expression/dataset-human-crc>

Input Files	Size	md5sum	
Aggregation Input (CSV)	4.76 kB	42a7d406b241eea006309e2b2f78ed1a	Multi-sample aggregation file
Feature Reference (CSV)	1.98 MB	04ad9d02b505ea7c9db36afaf630f54d	
Output Files	Size	md5sum	
Clustering analysis	104 MB	010059a26c7280820cc25a62f217d1ed	
Feature barcode matrix HDF5 (filtered)	426 MB	c353717870a812b997e7f66defd8e408	Filtered gene-cell matrix
Feature barcode matrix (filtered)	1.17 GB	92f2284f82d2bd364a90318b39f612d7	
Loupe Browser file	1.38 GB	0a3b6bab997818ab347f03c8a90b6e6e	
Summary JSON	35.1 kB	33e51486a98372531a5004892241def4	

# Chromium single cell data

<https://www.10xgenomics.com/products/visium-hd-spatial-gene-expression/dataset-human-crc>

Multi-sample aggregation file

```
1 | sample_id,molecule_h5
2 | 20k_HumanColonCancer_Flex_Multiplex_P1L1_P2CRC_BC1,/path/to/aggr_inputs/
| 20k_HumanColonCancer_Flex_Multiplex_P1/L1_P2CRC_BC1/sample_molecule_info.h5
3 | 20k_HumanColonCancer_Flex_Multiplex_P1L1_P4CRC_BC2,/path/to/aggr_inputs/
| 20k_HumanColonCancer_Flex_Multiplex_P1/L1_P4CRC_BC2/sample_molecule_info.h5
4 | 20k_HumanColonCancer_Flex_Multiplex_P1L2_P2CRC_BC1,/path/to/aggr_inputs/
| 20k_HumanColonCancer_Flex_Multiplex_P1/L2_P2CRC_BC1/sample_molecule_info.h5
5 | 20k_HumanColonCancer_Flex_Multiplex_P1L2_P4CRC_BC2,/path/to/aggr_inputs/
| 20k_HumanColonCancer_Flex_Multiplex_P1/L2_P4CRC_BC2/sample_molecule_info.h5
6 | 20k_HumanColonCancer_Flex_Multiplex_P1L3_P2CRC_BC1,/path/to/aggr_inputs/
| 20k_HumanColonCancer_Flex_Multiplex_P1/L3_P2CRC_BC1/sample_molecule_info.h5
7 | 20k_HumanColonCancer_Flex_Multiplex_P1L3_P4CRC_BC2,/path/to/aggr_inputs/
| 20k_HumanColonCancer_Flex_Multiplex_P1/L3_P4CRC_BC2/sample_molecule_info.h5
8 | 20k_HumanColonCancer_Flex_Multiplex_P1L4_P2CRC_BC1,/path/to/aggr_inputs/
| 20k_HumanColonCancer_Flex_Multiplex_P1/L4_P2CRC_BC1/sample_molecule_info.h5
```

Cell barcodes in aggregated matrix

```
> cells[1:10]
[1] "AAACAAGCAACAGCACACTTAGG-1"
[4] "AAACAAGCAAGGCCTGACTTTAGG-1"
[7] "AAACAAGCAGGGCTATACTTTAGG-1"
[10] "AAACCAATCACTGTTACTTTAGG-1"
> cells[15000:15010]
[1] "GCTCACTAGGGTTGCAAACGGGAA-2"
[4] "GCTCAGGCATTGTTGAACGGGAA-2"
[7] "GCTCCATAGGTTAGTAAACGGGAA-2"
[10] "GCTCCTGAGGTGTCCTAACGGGAA-2"
> cells[20000:20010]
[1] "ATATTCGGTACCAGCACTTTAGG-3"
[4] "ATATTCGGTCCATCCTACTTTAGG-3"
[7] "ATATTCGGTTACCGTGACTTTAGG-3"
[10] "ATCAGCACACTGACTTACTTTAGG-3"
```

# Single-cell FLEX annotation results

[https://github.com/10XGenomics/HumanColonCancer\\_VisiumHD/tree/main/MetaData](https://github.com/10XGenomics/HumanColonCancer_VisiumHD/tree/main/MetaData)

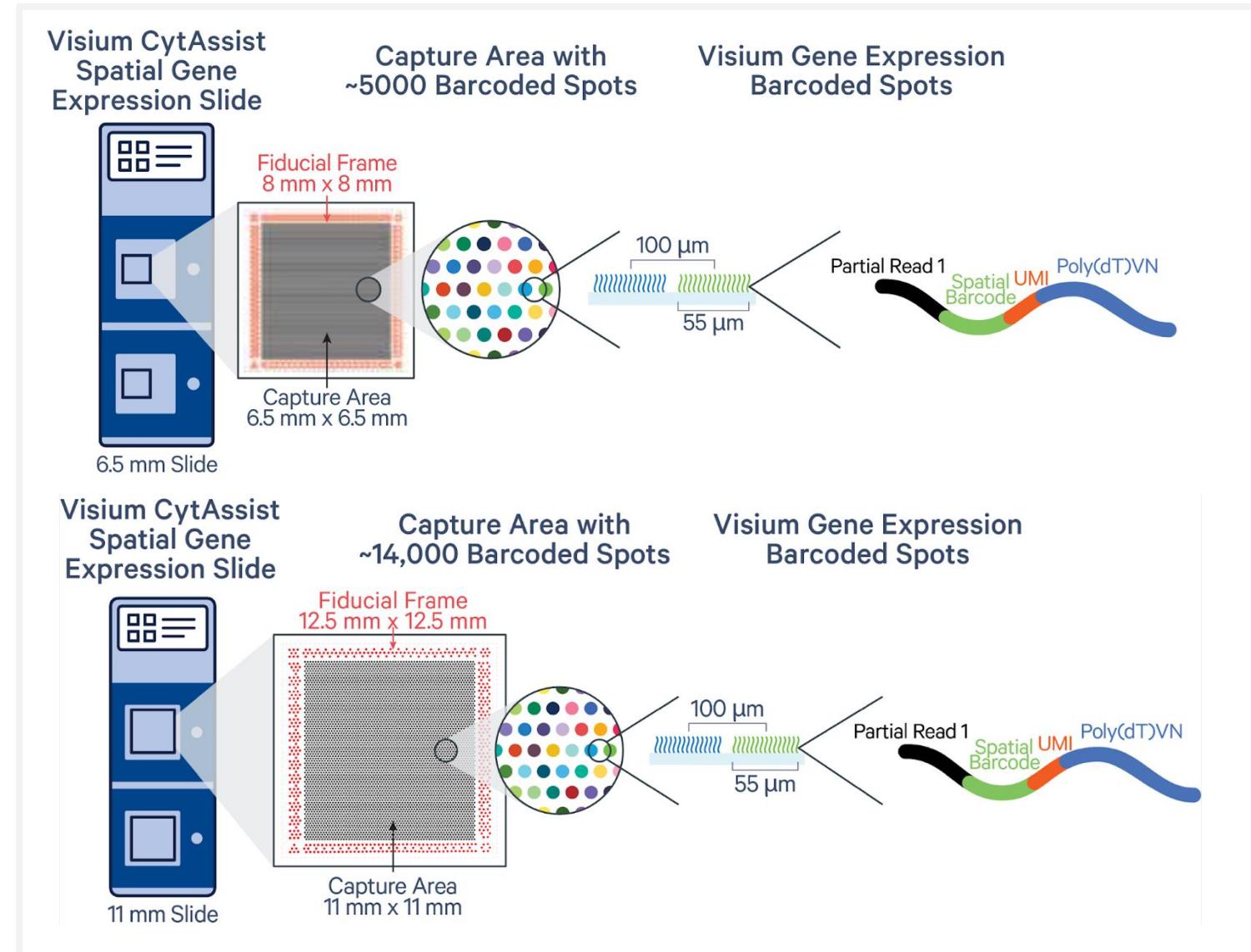
Name
..
DeconvolutionResults_P1CRC.csv.gz
DeconvolutionResults_P2CRC.csv.gz
DeconvolutionResults_P5CRC.csv.gz
SingleCell_MetaData.csv.gz



# Visium CytAssist Data

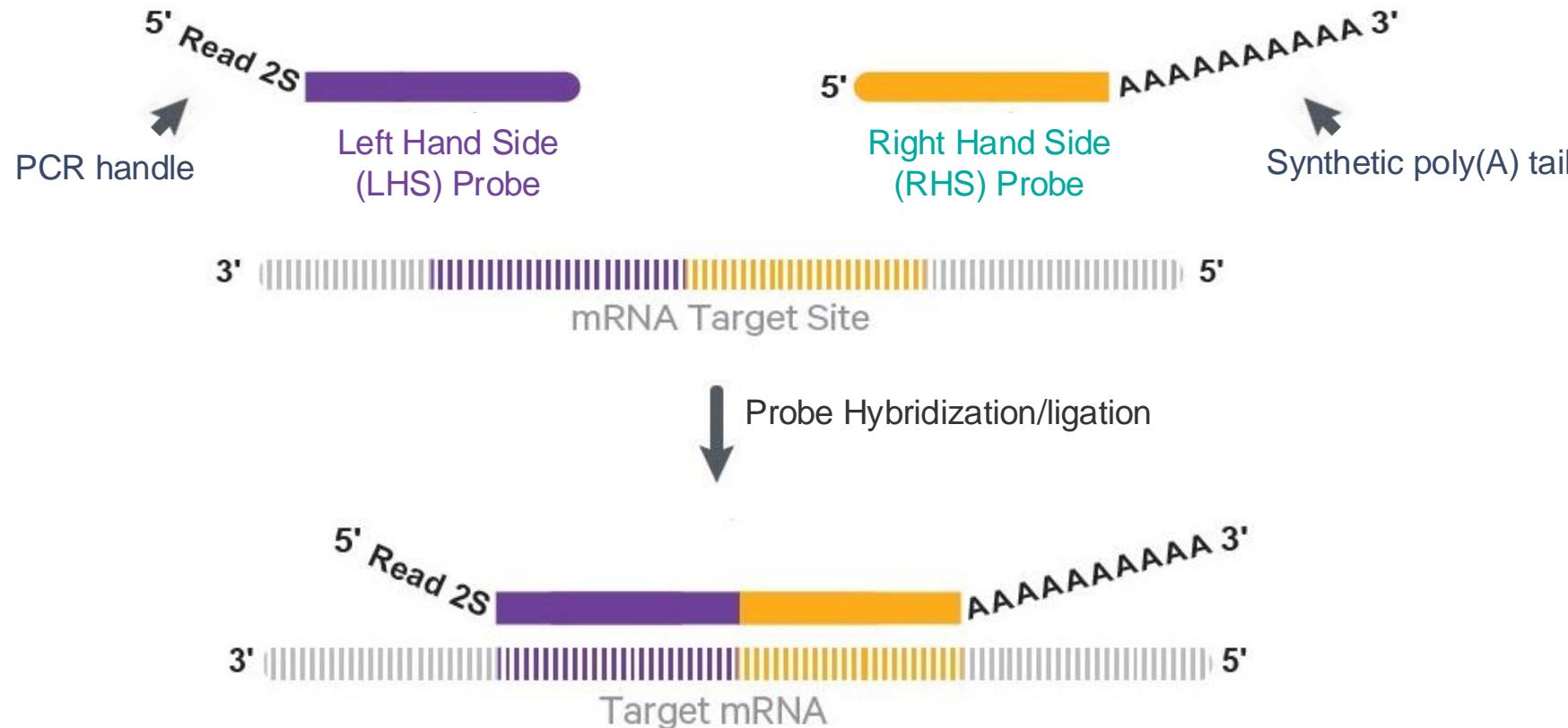
---

# Visium CytAssist Gene Expression Slide Architecture



# Visium for FFPE uses probes to detect mRNAs

Probe pairs Hybridize to mRNA Targets



## DNA-based Probes

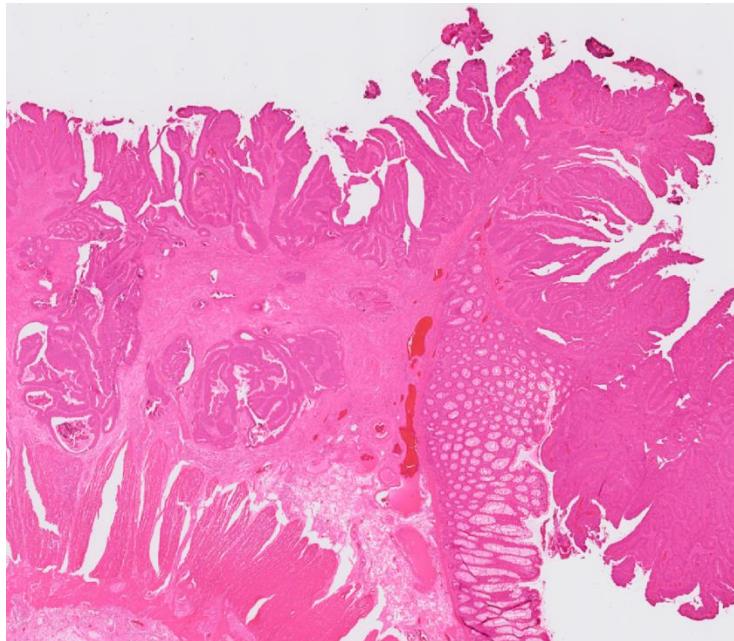
- LHS has partial Read 2 sequence
- RHS has synthetic poly(A) sequence

CytAssist assays are all probe based for FF, FFPE and FxF samples

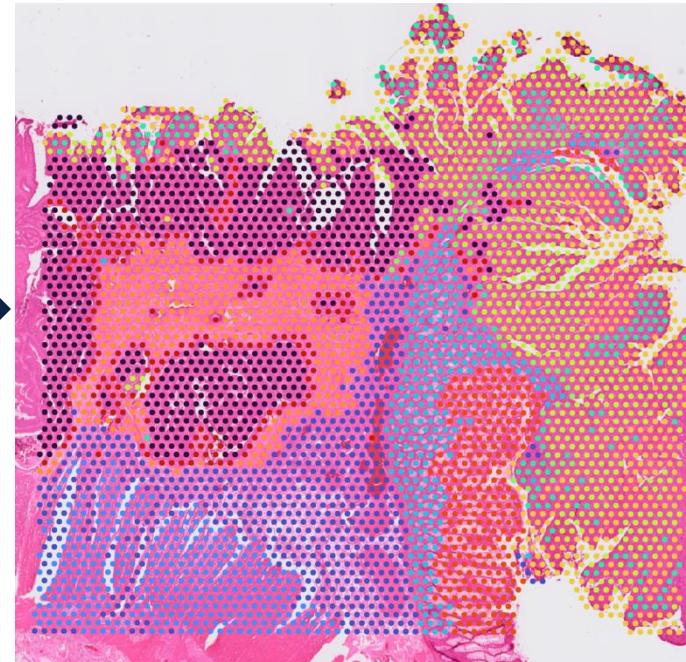
# Two types of information make up Visium data

Tissue image and gene expression

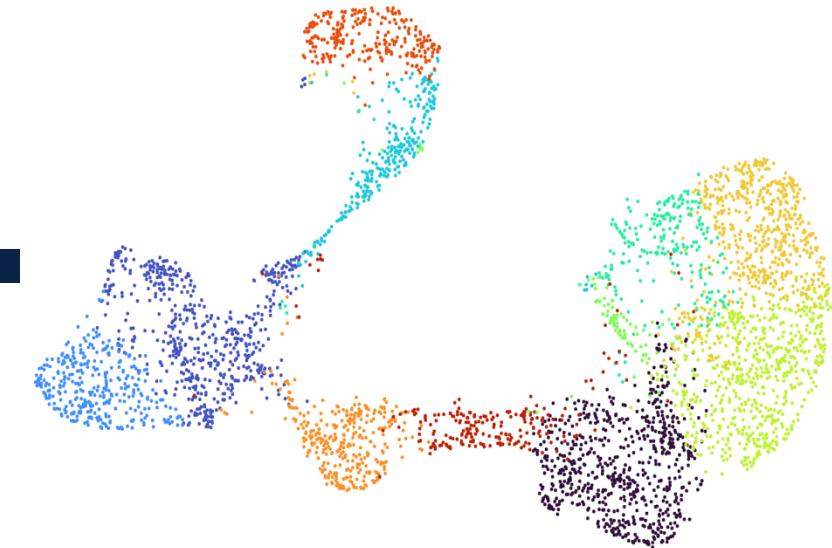
Histology image



Histology + Gene expression



Gene expression



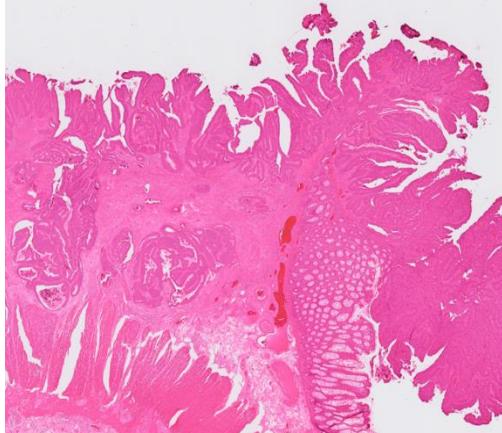
# Visium CytAssist data

Visium CytAssist image (Required)



Sequencing data .fastq.gz (Required)

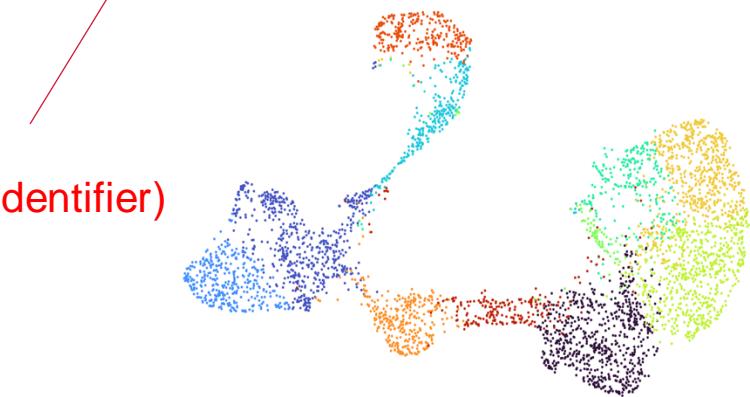
High-resolution microscope image (Optional)



Space  
Ranger

	Gene1	Gene2	Gene3
Barcode1	1	0	1
Barcode2	0	2	1
Barcode3	1	1	0
Barcode4	1	0	1
Barcode5	1	1	1

UMI  
(Unique Molecule Identifier)



# Visium CytAssist data

<https://www.10xgenomics.com/products/visium-hd-spatial-gene-expression/dataset-human-crc>

Visium input files:

Input Files	Size	md5sum	
FASTQs	130 GB	ec6b67507249f7e64745e670c0bc40cb	Sequencing data
Probe Set (CSV)	6.13 MB	a74f1583689de3c9f95164b2fc090194	
CytAssist Image (TIFF)	25.1 MB	8bd2b1f2020d525da2fe2c85855eb718	Visium CytAssist image
Microscope image (BTF)	11.2 GB	a3d267bad4d2b32d1ee4825ce05c1bd1	High-resolution microscope image
Alignment file (JSON)	1.48 MB	22f9f07499009516ba7e95913ef08a9d	

# Visium CytAssist data

<https://www.10xgenomics.com/products/visium-hd-spatial-gene-expression/dataset-human-crc>

Visium output files from spaceranger: more details in <https://www.10xgenomics.com/support/software/spaceranger/latest/analysis/outputs/output-overview#count-visium-v1-v2>

Output Files	Size	md5sum	
Clustering analysis	36.6 MB	563572b5aad1cd7252e1a72dcd1d98f0	
Feature barcode matrix HDF5 (filtered)	46.9 MB	0a9733cf0dfd3dfa3b6c1c5c24bcbed	
Feature barcode matrix HDF5 (raw)	162 MB	6ef05a15bc17e9107a1547644e458b2a	
Feature barcode matrix (filtered)	156 MB	1af187605b0b0d6505b8a918f0bf81a0	Filtered barcode-gene matrix
Feature barcode matrix (raw)	446 MB	b14269874bb62d94c2404259c964cb0a	
Probe barcode matrix HDF5 (raw)	259 MB	f621992a1bdb61835c2333175b2c6570	
Spatial outputs	44.8 MB	a32edc825ab7089fca7848fc80cd5113	Spatial image information
Per-molecule read information	4.15 GB	27893884bfcded523a473bf332874987	
Loupe Browser file	6.3 GB	c5726e203c21a2fc3432831ea7c74252	cloupe file for Loupe Browser
Spot deconvolution	10.6 MB	e621fa722d6cc7e181c0da359681ab8b	
Summary CSV	1.05 kB	1c9e73f90cd018dd8c975bb64818b31a	

# Visium CytAssist data

<https://www.10xgenomics.com/products/visium-hd-spatial-gene-expression/dataset-human-crc>

Visium output files from spaceranger:

Spatial outputs	44.8 MB	a32edc825ab7089fca7848fc80cd5113	Spatial image information
-----------------	---------	----------------------------------	---------------------------

Link to understand spatial image information:

<https://www.10xgenomics.com/support/software/space-ranger/latest/analysis/outputs/spatial-outputs>

-  aligned\_fiducials.jpg
-  aligned\_tissue\_image.jpg
-  cytassist\_image.tiff
-  detected\_tissue\_image.jpg
-  scalefactors\_json.json
-  spatial\_enrichment.csv
-  tissue\_hires\_image.png
-  tissue\_lowres\_image.png
-  tissue\_positions.csv

# Visium CytAssist data

Visium output files from spaceranger:

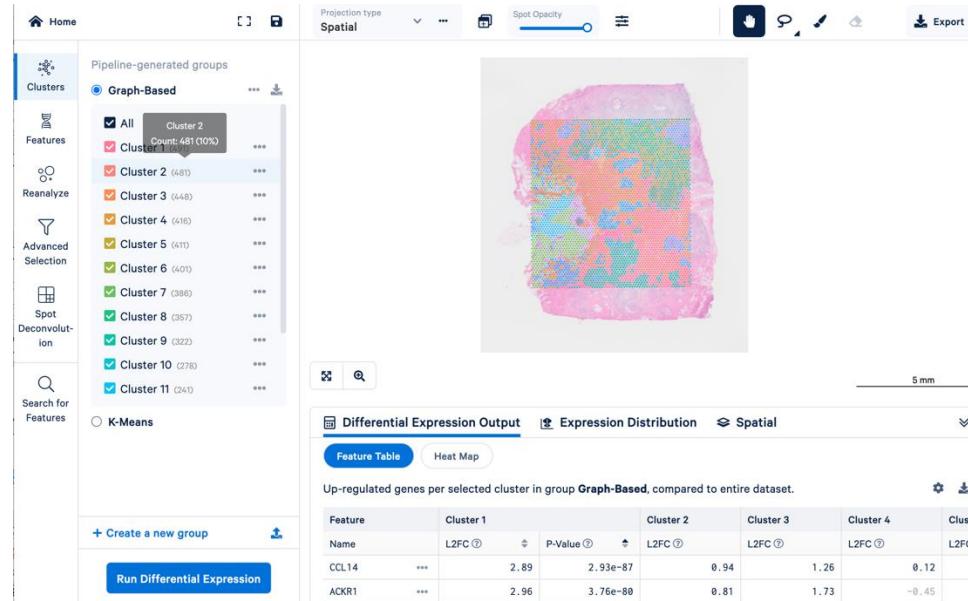
Loupe Browser file

6.3 GB

c5726e203c21a2fc3432831ea7c74252

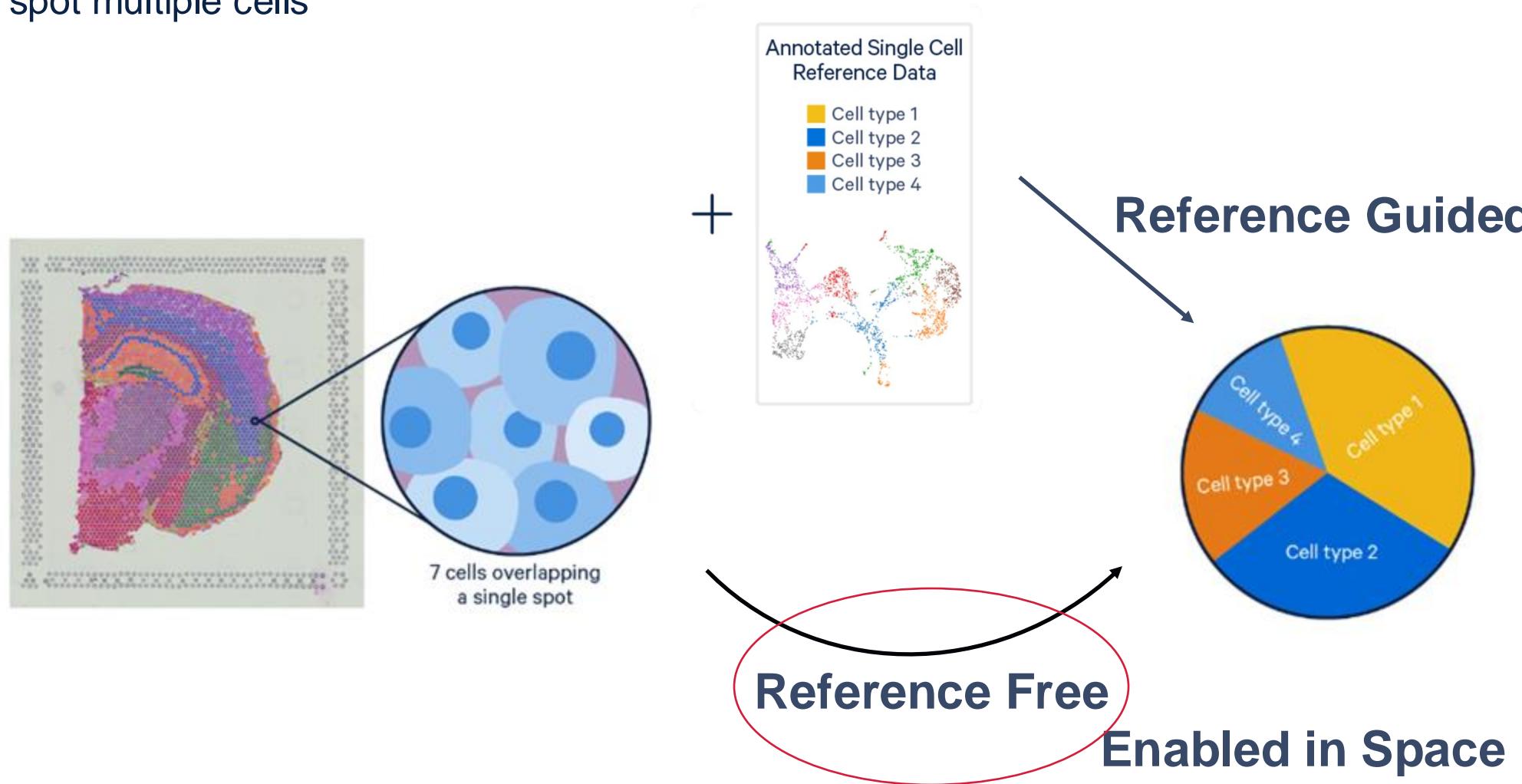
cloupe file for Loupe Browser

Loupe Browser download link: <https://www.10xgenomics.com/support/software/loupe-browser/downloads>  
Loupe Browser learning link: <https://www.10xgenomics.com/videos/vt55rn7399?autoplay=true>



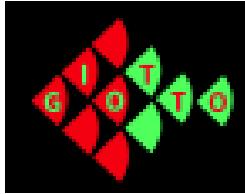
# Spot deconvolution

One spot multiple cells



# Spot Cell-type Deconvolution: Community Developed Software

## Reference



SpatialDWLS



Spacexr  
(RCTD)



## Reference-free



STdeconvolve

Enabled in Space Ranger v2.1

See our Analysis Guides to learn more about deconvolution:

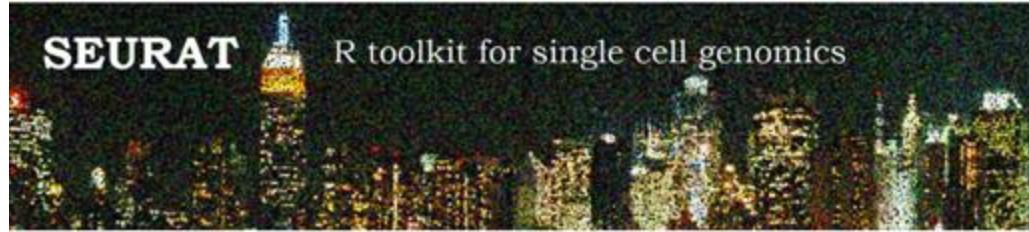
[Publication Highlight: Benchmarking Methods to Integrate Spatial and Single-cell Transcriptomics Data](#)

[Integrating 10x Visium and Chromium data with R](#)

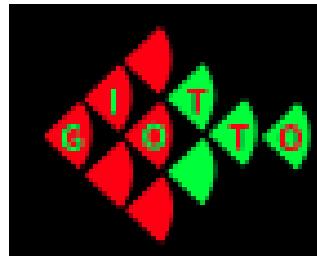
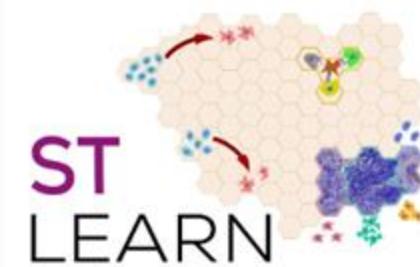
[Integrating Single Cell and Visium Spatial Gene Expression Data](#)

# Ecosystem of community developed tools

## Comprehensive Toolkits



R toolkit for single cell genomics



Scanpy tutorial for Visium: <https://scanpy.readthedocs.io/en/stable/tutorials/spatial/basic-analysis.html>  
Seurat tutorial for Visium: [https://satijalab.org/seurat/articles/spatial\\_vignette#x-visium](https://satijalab.org/seurat/articles/spatial_vignette#x-visium)

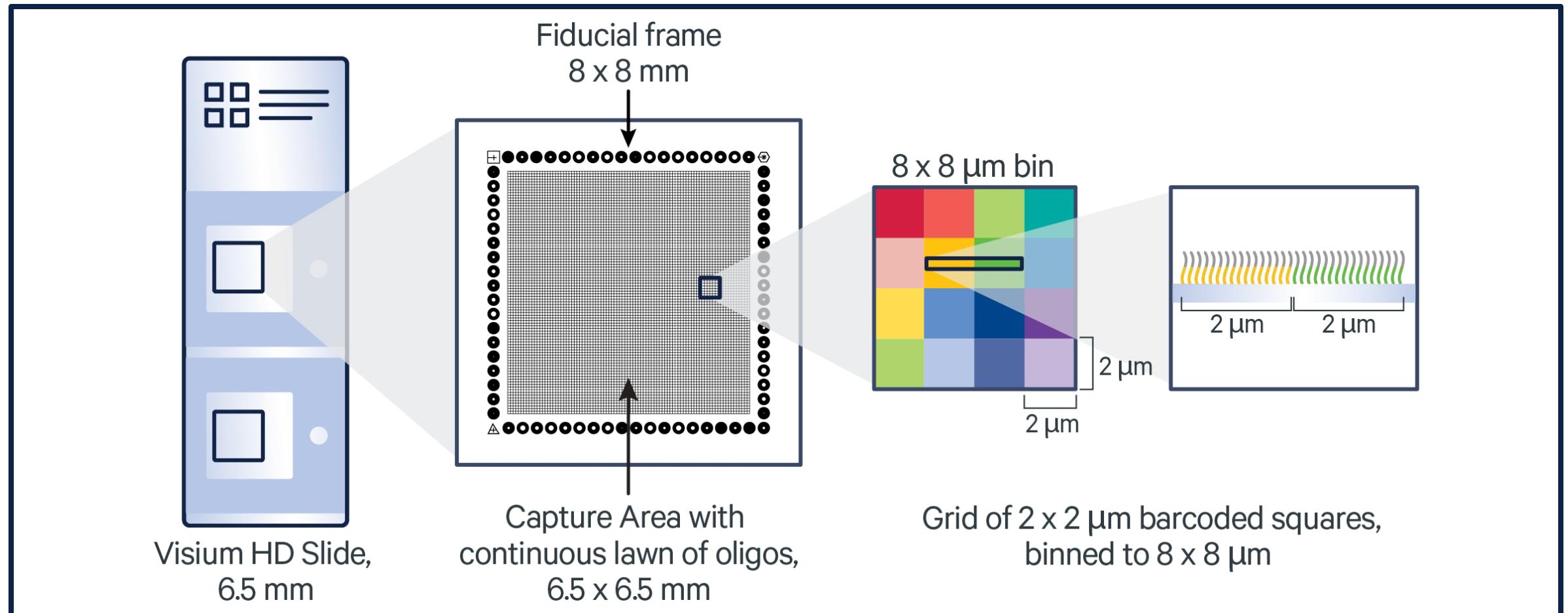


# Visium HD Data

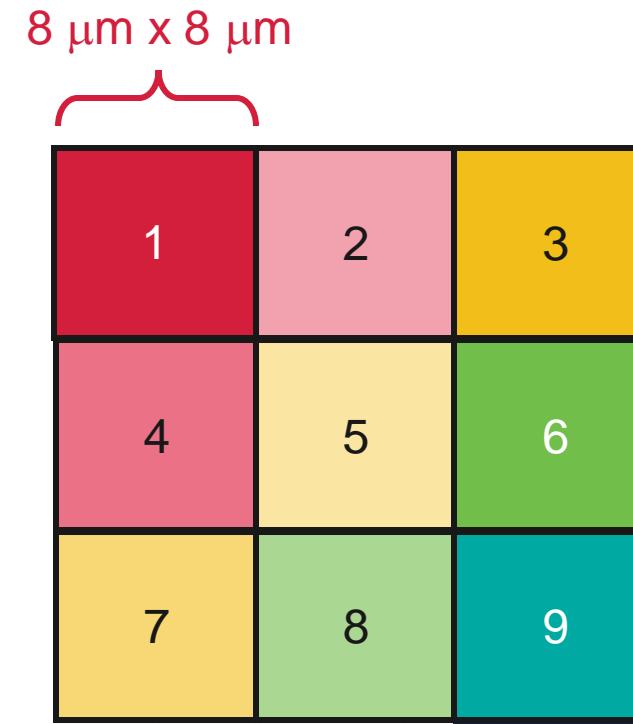
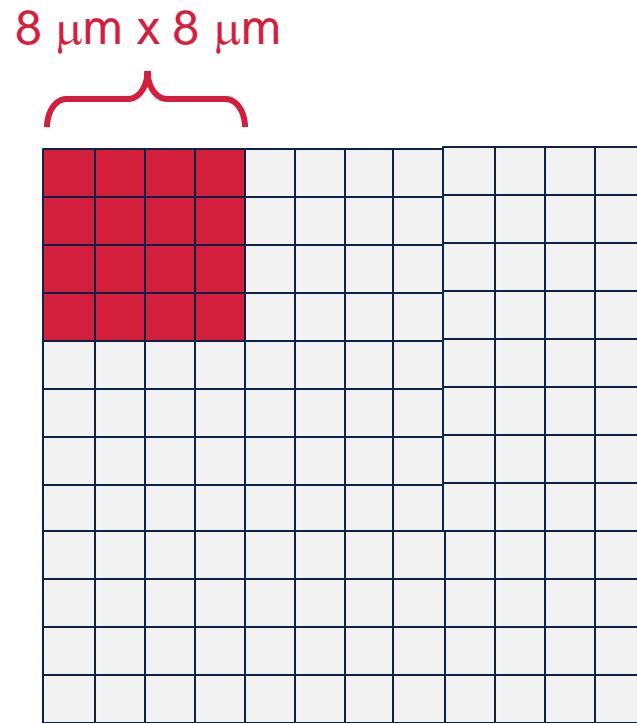
---

# Visium HD Spatial Gene Expression

## Visium HD Slide Design



# Space Ranger Barcode Binding



	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9
Gene 1	24	0	20	1	0	10	40	4	7
Gene 2	30	30	5	1	0	15	5	8	8
Gene 3	8	10	10	1	30	20	1	16	9

Repeat across the entire grid.

# Visium HD data

<https://www.10xgenomics.com/products/visium-hd-spatial-gene-expression/dataset-human-crc>

Visium input files:

Input Files	Size	md5sum	
FASTQs	62.1 GB	e73640ee6bdc62b186b7bbb0732a39ad	Sequencing data
Probe Set (CSV)	6.13 MB	a74f1583689de3c9f95164b2fc090194	
CytAssist Image (TIFF)	29 MB	50ae8e50b2f313fdbd16b7d59082608f	Visium CytAssist image
Microscope image (BTF)	12.6 GB	81f5f4b49b47755d64b5b8d20eb29305	High-resolution microscope image
Alignment file (JSON)	118 kB	4c8bd95e5fecb4f9cbf6d8707f8124c5	
Slide file (VLF)	546 B	46ef2d8466c239f5e253931916ec3347	

# Space Ranger Output Structure

“binned\_outputs” folder contents

```
outs/
└── binned_outputs
    ├── cloupe_008um.cloupe
    ├── feature_slice.h5
    ├── metrics_summary.csv
    ├── molecule_info.h5
    ├── probe_set.csv
    └── spatial
    web_summary.html
Binned Outputs Directory
```

- 2  $\mu\text{m}$  x 2  $\mu\text{m}$  raw data
- 8  $\mu\text{m}$  x 8  $\mu\text{m}$ , 16  $\mu\text{m}$  x 16  $\mu\text{m}$  bins produced by default
- Feature barcode matrices (raw and filtered) produced for each binned output
- Spatial results generated for each binned output
- Custom binning results produced when specified

```
binned_outputs/
└── square_002um
    ├── square_002um
    ├── square_008um
    └── square_016um
```

```
square_002um/
    ├── filtered_feature_bc_matrix
    ├── filtered_feature_bc_matrix.h5
    ├── raw_feature_bc_matrix
    ├── raw_feature_bc_matrix.h5
    ├── raw_probe_bc_matrix.h5
    └── spatial

square_008um/
    ├── analysis
    ├── cloupe.cloupe
    ├── filtered_feature_bc_matrix
    ├── filtered_feature_bc_matrix.h5
    ├── raw_feature_bc_matrix
    ├── raw_feature_bc_matrix.h5
    └── spatial

square_016um/
    ├── analysis
    ├── filtered_feature_bc_matrix
    ├── filtered_feature_bc_matrix.h5
    ├── raw_feature_bc_matrix
    ├── raw_feature_bc_matrix.h5
    └── spatial
```

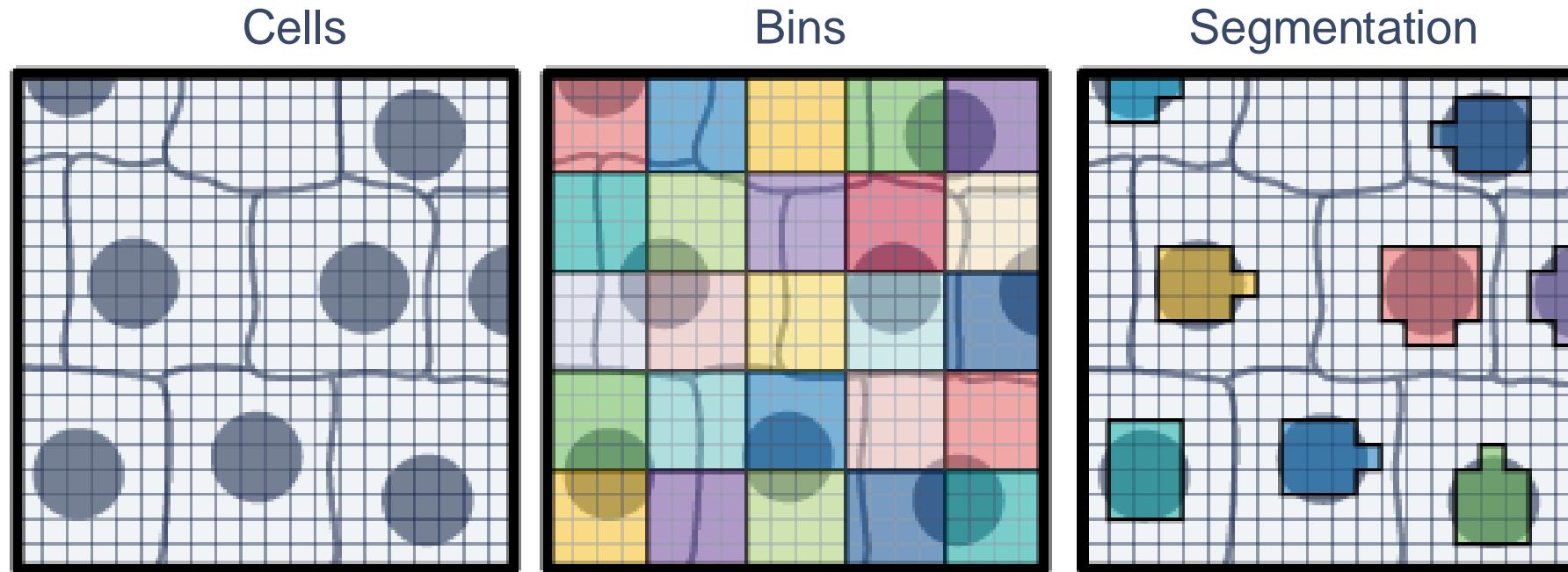
# Visium HD data

<https://www.10xgenomics.com/products/visium-hd-spatial-gene-expression/dataset-human-crc>

Visium output files from spaceranger: more details in <https://www.10xgenomics.com/support/software/spaceranger/latest/analysis/outputs/output-overview#hd-outputs>

Output Files	Size	md5sum
Binned outputs	12.3 GB	7e66f54d272b626ba22e438d7e7cd8cb
Spatial outputs	77.2 MB	ead014aa0fa2234b937b09672c75e362
Per-molecule read information	761 MB	5999a9a037a30af5340bbac17ca5a8fb
Loupe Browser file	8.61 GB	c5d80d498b9c4d0c442c9da4479d70ee
Feature slice (HDF5)	512 MB	2929e805901d562b18b6a915ef16aef0
Summary CSV	1.73 kB	363a1c0a09e6d36c1c5904c50d13e04d

# Nuclei segmentation of H&E images to inform Visium HD segmentation at 2 $\mu$ m resolution



<https://www.10xgenomics.com/analysis-guides/segmentation-visium-hd>

# Community developed tools for HD

[https://satijalab.org/seurat/articles/visiumhd\\_analysis\\_vignette](https://satijalab.org/seurat/articles/visiumhd_analysis_vignette)

## Analysis, visualization, and integration of Visium HD spatial datasets with Seurat

Compiled: 2024-05-06

Source: vignettes/visiumhd\_analysis\_vignette.Rmd

### Visium HD support in Seurat

We have [previously released support](#) Seurat for sequencing-based spatial transcriptomic (ST) technologies, including 10x visium and SLIDE-seq. We have now updated Seurat to be compatible with the Visium HD technology, which performs profiling at substantially higher spatial resolution than previous versions.

Users can install the Visium HD-compatible release from Github. Existing Seurat workflows for [clustering, visualization, and downstream analysis](#) have been updated to support both Visium and Visium HD data.

We note that Visium HD data is generated from spatially patterned oligonucleotides labeled in 2um x 2um bins. However, since the data from this resolution is sparse, adjacent bins are pooled together to create 8um and 16um resolutions. 10x recommends the use of 8um binned data for analysis, but Seurat supports the simultaneous loading of multiple binnings - and stores them in a single object as multiple assays.

In this vignette, we provide an overview of some of the spatial workflows that Seurat supports for analyzing Visium HD data, in particular:



# Xenium Data

---

# Xenium data

<https://www.10xgenomics.com/products/visium-hd-spatial-gene-expression/dataset-human-crc>

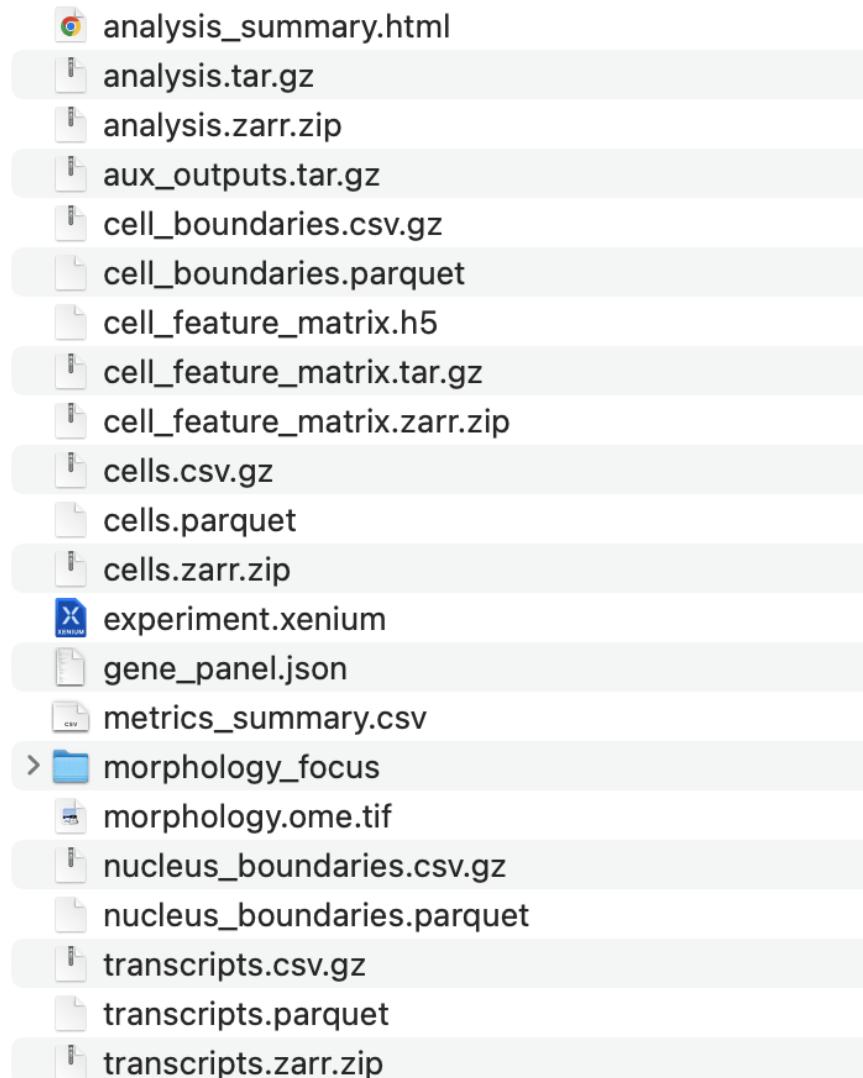
If the file size is large, we suggest using [batch download](#) instead.

Input Files	Size	md5sum
Panel (JSON)	222 kB	2d614474a7ca8e21c30734c7f4a1bfdf
Post-Xenium H&E image (OME-TIFF)	8.66 GB	0673d535420a344cab766f6ac9daf474
H&E Image Alignment File (CSV)	638 B	822890860598d8d0ba5e7940486fd920

Output Files	<a href="#">Format details</a>	Size	md5sum
Analysis Summary HTML		14.8 MB	d4a2dfd74622bd014043e340e2042aff
Xenium Output Bundle		13 GB	d12df7c313c808c80a628c232a731012

# Xenium data

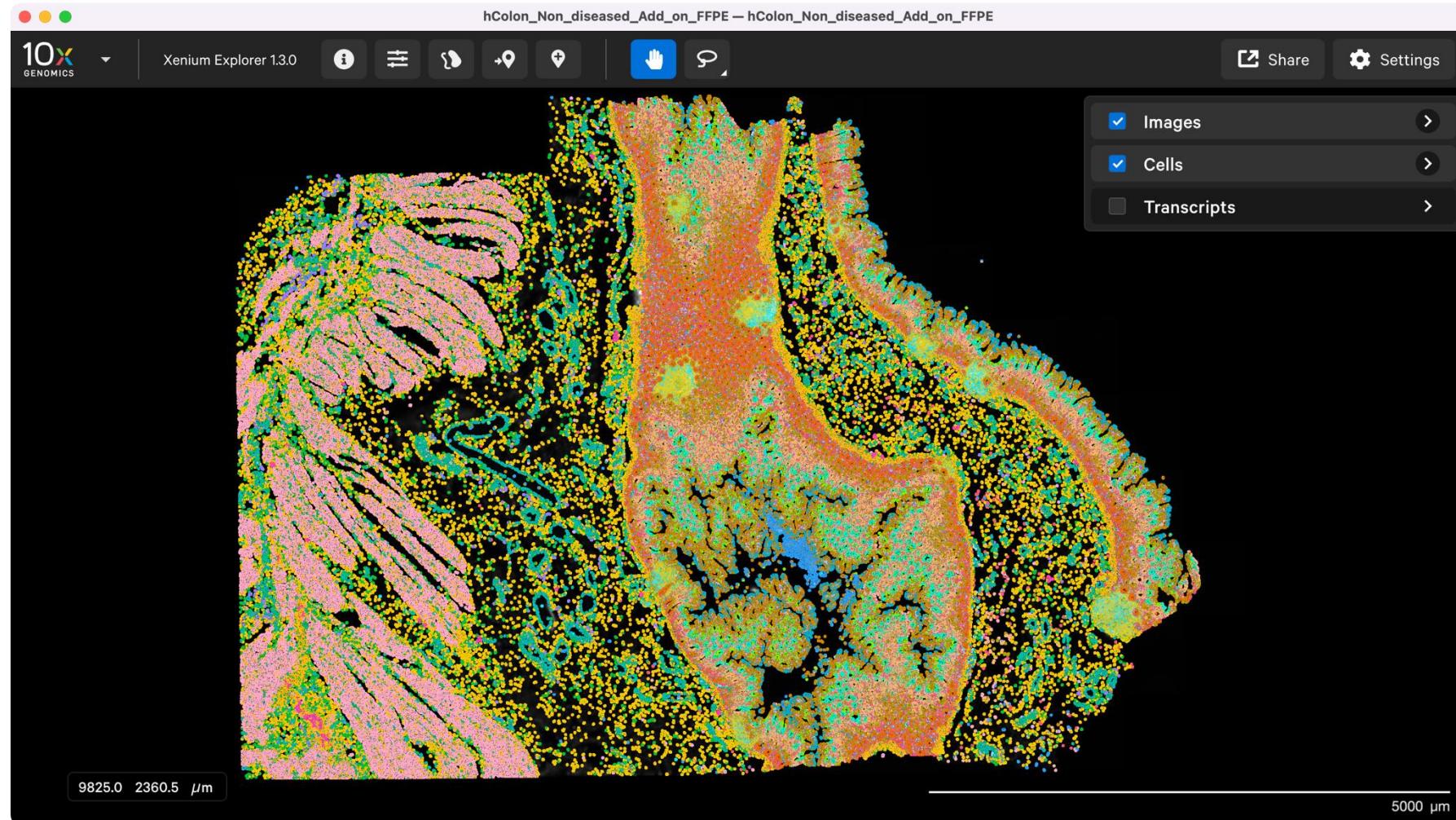
Understand details of Xenium outputs: <https://www.10xgenomics.com/support/software/xenium-onboard-analysis/latest/analysis/xoa-output-understanding-outputs>



# Xenium Explorer

Download link: <https://www.10xgenomics.com/support/software/xenium-explorer/downloads>

Tutorial link: <https://www.10xgenomics.com/support/software/xenium-explorer/latest/tutorials>



# Xenium Community developed tools

## Community-developed tools written in R

- [Seurat](#): QC, cropping images and cell boundaries, and unsupervised clustering
- [Giotto](#): add statistics, normalize expression, calculate high variable features, dimension reduction, clustering, and subcellular visualization
- [Voyager](#): QC, spatial autocorrelation of QC metrics, Moran's I (correlation coefficient, measures how one object is similar to others surrounding it), dimension reduction, differential expression, and local spatial statistics of marker genes

## Community-developed tools written in Python

- [stLearn](#) (two tutorials, [CCI](#), [PSTS](#)): normalization, gridding, permutation test for high co-expression, [cell-cell interactions](#), clustering, and [pseudo-time-space spatial trajectory analysis](#),
- [Squidpy](#): QC, dimension reduction, spatial statistics, neighbors enrichment analysis, and compute Moran's I score