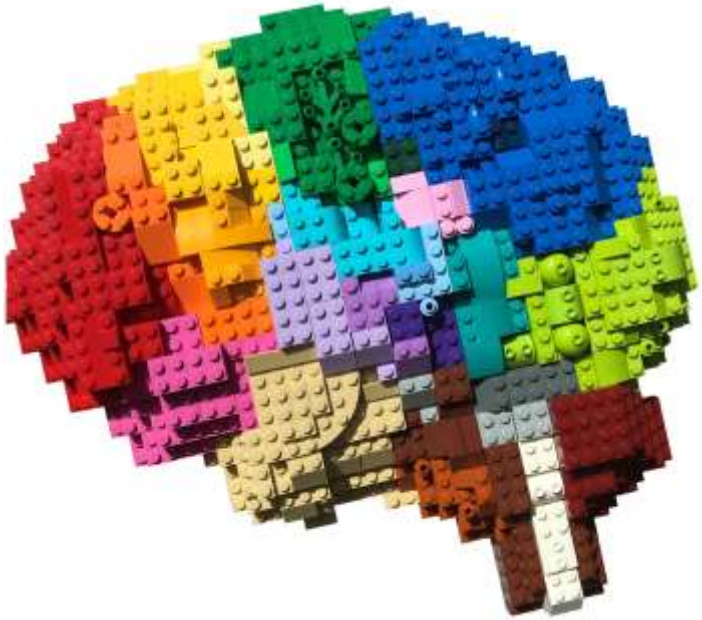
生物資訊與數據分析系列

Spatial Transcriptomics II

2026-08-11 | 吳啟智 | 中研院植微所細胞核心實驗室

Spatial gene-expression maps help reveal cellular differentiation, development, and tissue microenvironments

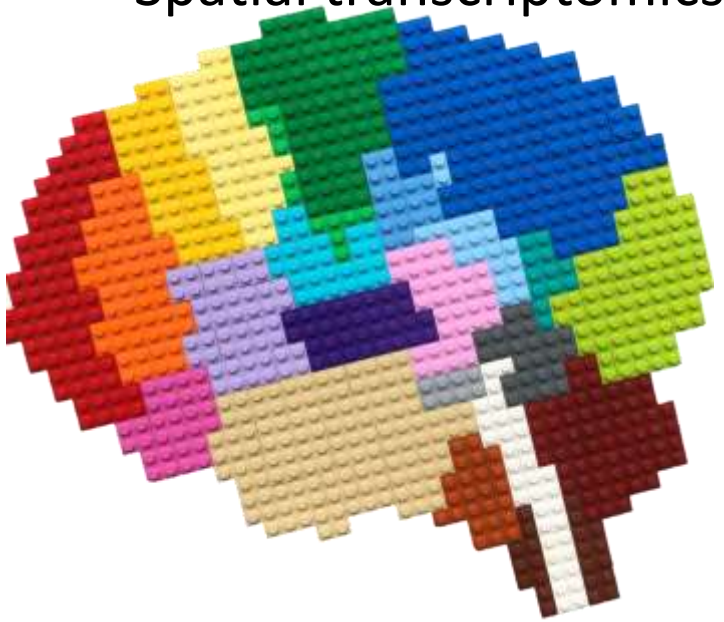
Sample



Bulk RNA-seq

Single-cell RNA-seq

Spatial transcriptomics



By the end of this workshop, participants will be able to:

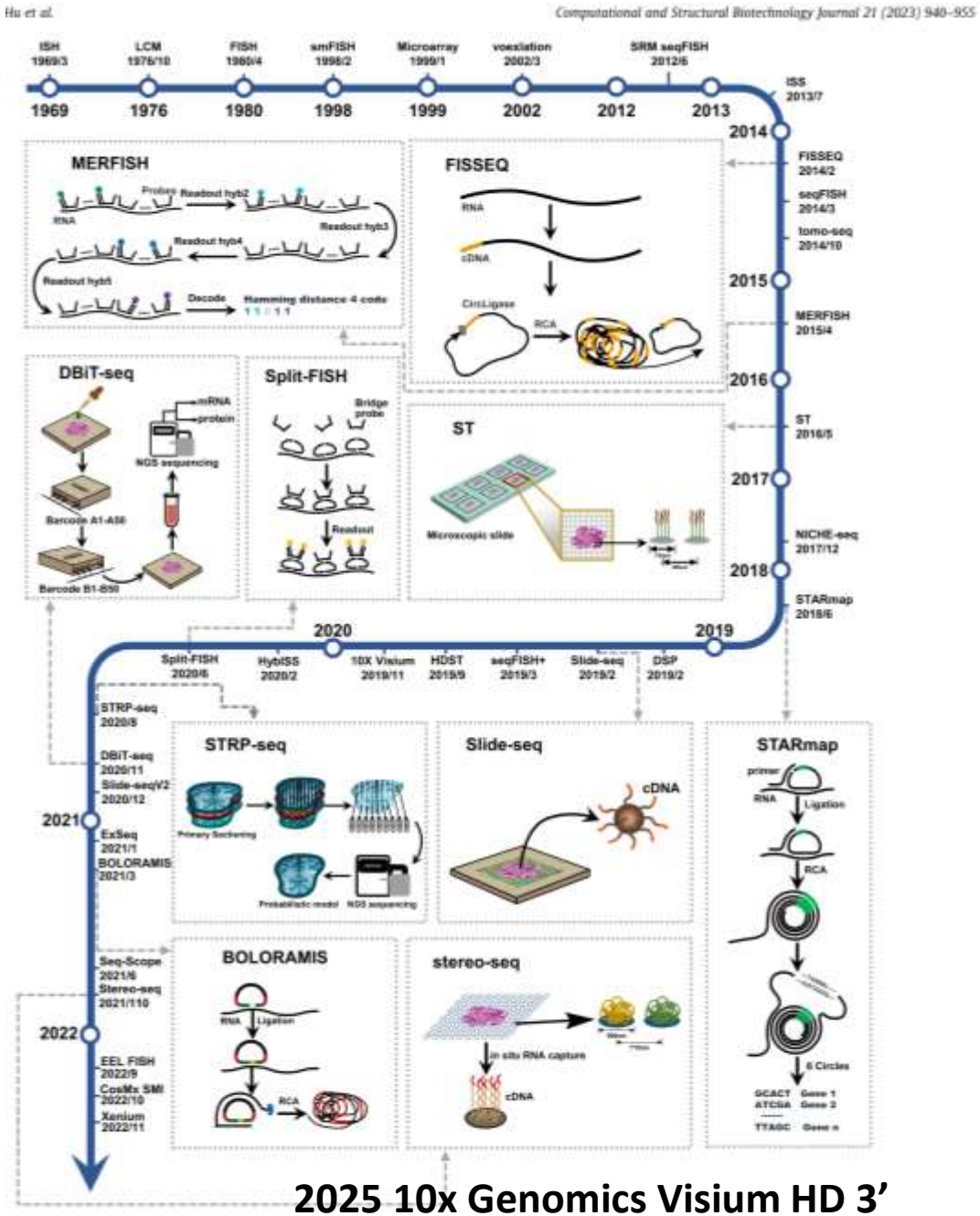
- Explain the basic principles of 10x Genomics Visium spatial transcriptomics.
- Identify the major files generated by Space Ranger.
- Describe the principal components of a Seurat spatial object.
- Perform basic quality control, normalization, clustering, and visualization.
- Add and validate spot- and domain-level annotations.
- Integrate an annotated scRNA-seq reference with spatial transcriptomics data.

Development of spatial gene-expression technologies

In situ hybridization–based methods

In situ sequencing–based methods

Next-generation sequencing–based methods



patial transcriptomics and related technological developments. The dates marked are referenced to the date of acceptance of the paper, whereas preprint articles to the date of online publication. Visium is refereneed to the press release on the official 10X Genomics website. The ten technologies selected are representative ment of spatial transcriptomic technologies, and together with the explanations in the text, they provide an accurate picture of the evolution of spatial tran-

Selected references and optional reading

1. Mapping biology in space: from spatial transcriptomics platforms to analytical tools and databases. *Engineering*, 2026.
    1. 完整分析工具

2. Wu et al. A Guide for Spatial Omics Technologies: Innovation, Evaluation, and Application. *Advanced Science*, 2026. https://doi.org/10.1002/advs.202520806
    1. 技術與平台比較

3. Yan G, Hua SH, Li JJ. Categorization of **34 computational methods** to detect spatially variable genes from spatially resolved transcriptomics data. *Nature Communications*. 2025;16:1141. https://doi.org/10.1038/s41467-025-56080-w
    1. 說明 spatial analysis 與一般 scRNA-seq analysis 的差異

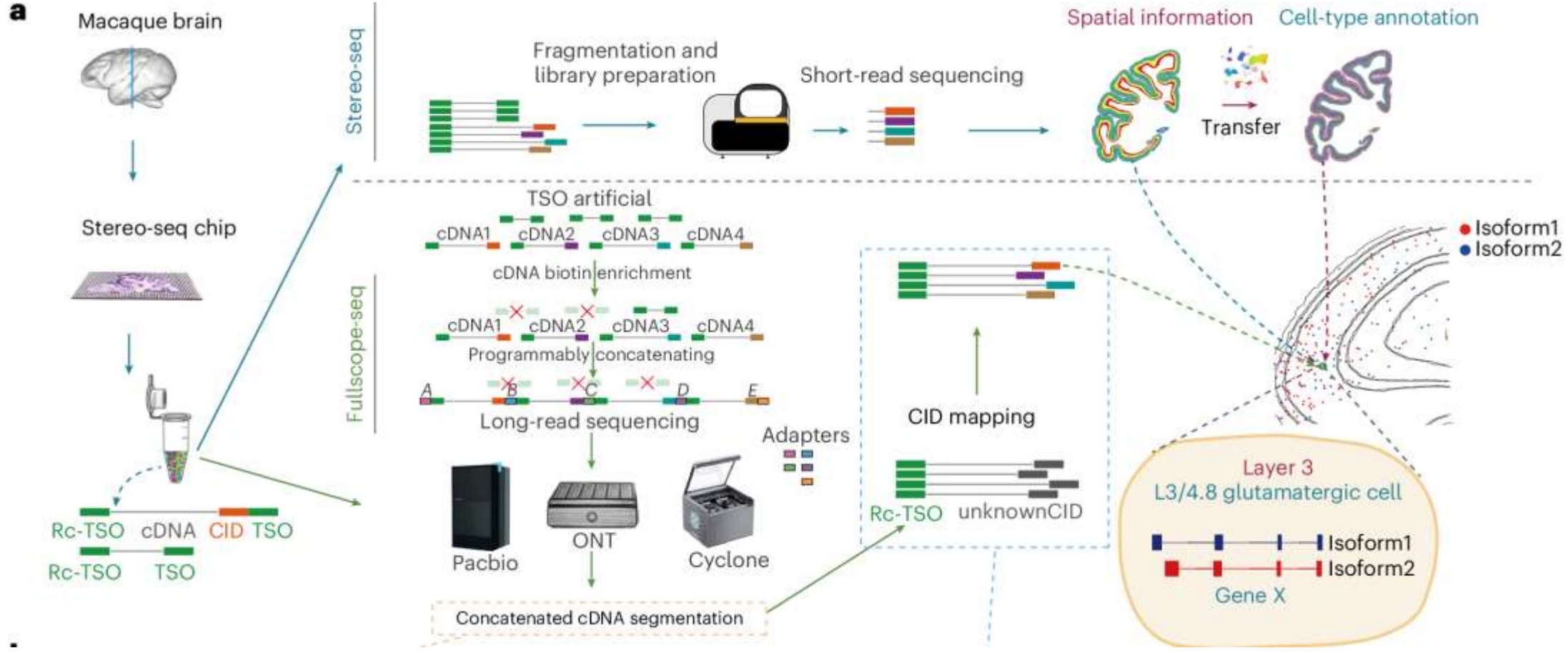
Year in review 2025 (Nature methods)

Spatial transcriptomics continues to be hot

Multiplexed sequencing offers more than the sum of the parts

Full-length single-cell spatial transcriptomics reveals spatial and cell-type-specific transcript isoforms in the primate brain

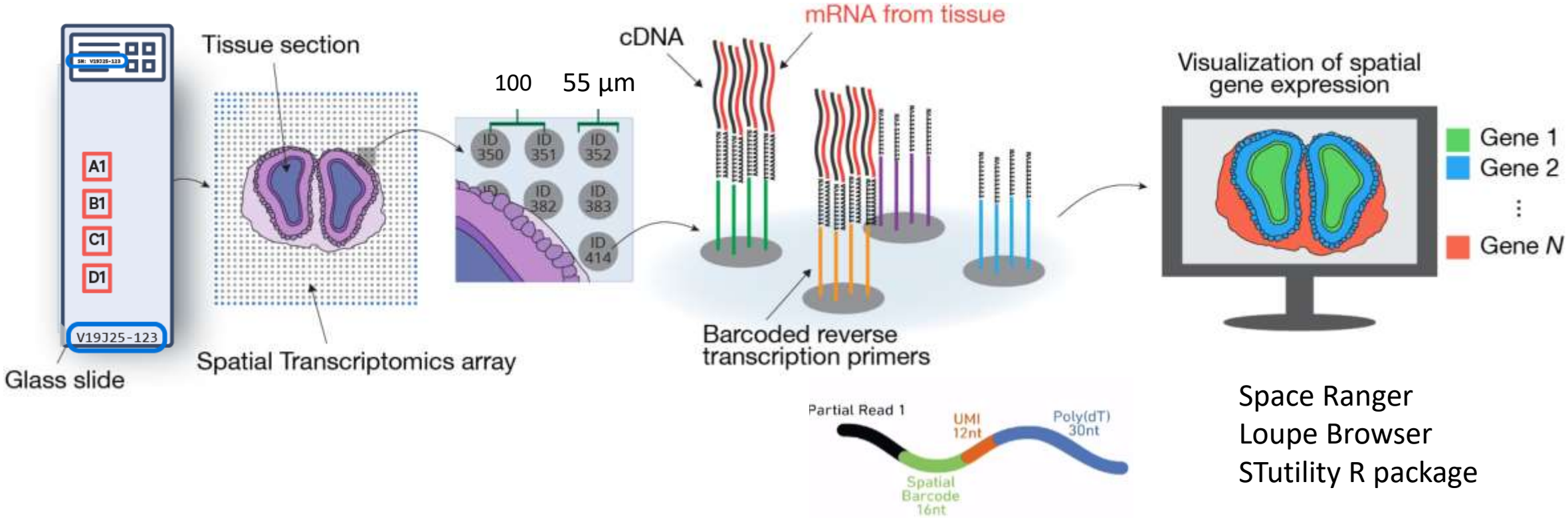
Liu, et. al., Nature methods, July 2026



From spatial transcriptomics to 10x Genomics Visium

Spatial transcriptomics, Lundeberg's lab, KTH, (Science, 2016)
Resolution: spot size 100 micrometer, center-to-center: 200 micrometer

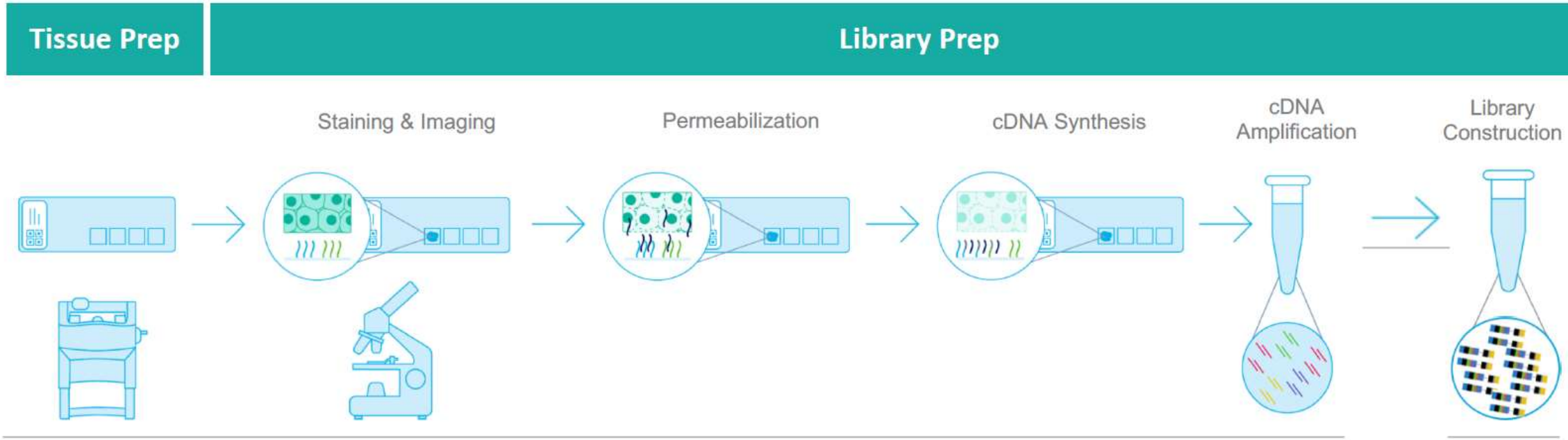
10x Visium V1 Spatial Gene Expression (2019.09)
Resolution: spot size 55 micrometer, center-to-center: 100 micrometer



Space Ranger
Loupe Browser
STutility R package

Visium Spatial Gene Expression

Complete in 1 Working Day with Standard Tools for Tissue Studies



**** Cryofixation***

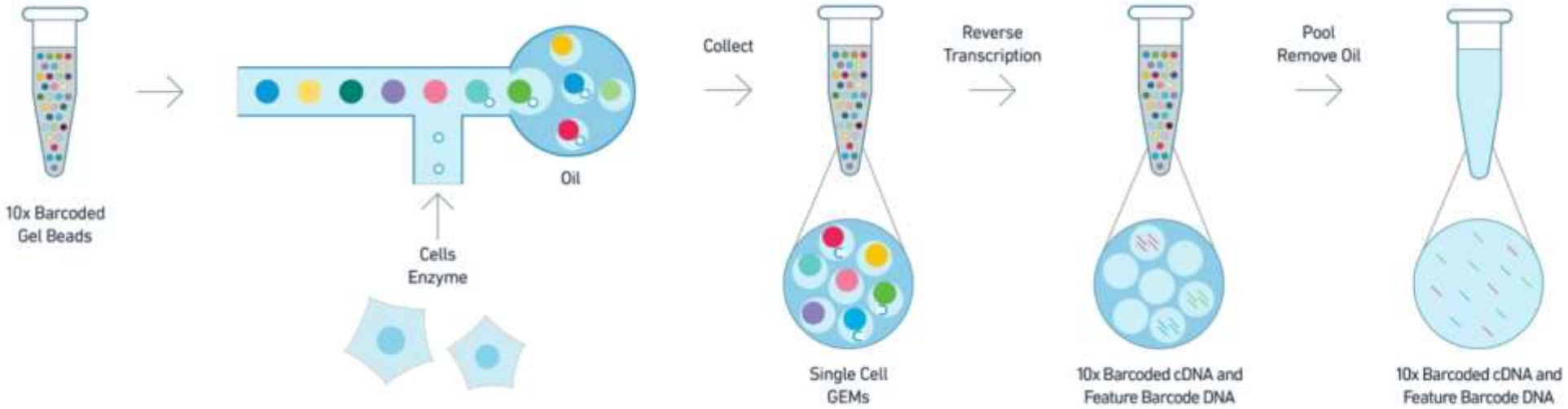
- *Section tissue*
- *Place on slide*

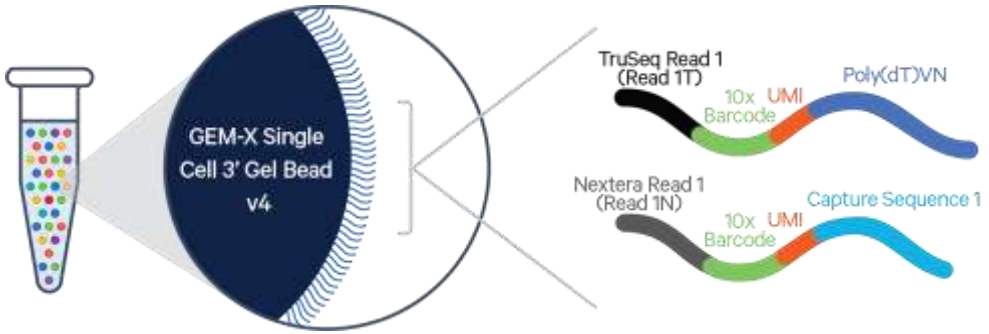
- *Fix Tissue*
- *H&E Stain*
- *Brightfield Image*

- *Barcoded Spot Capture*

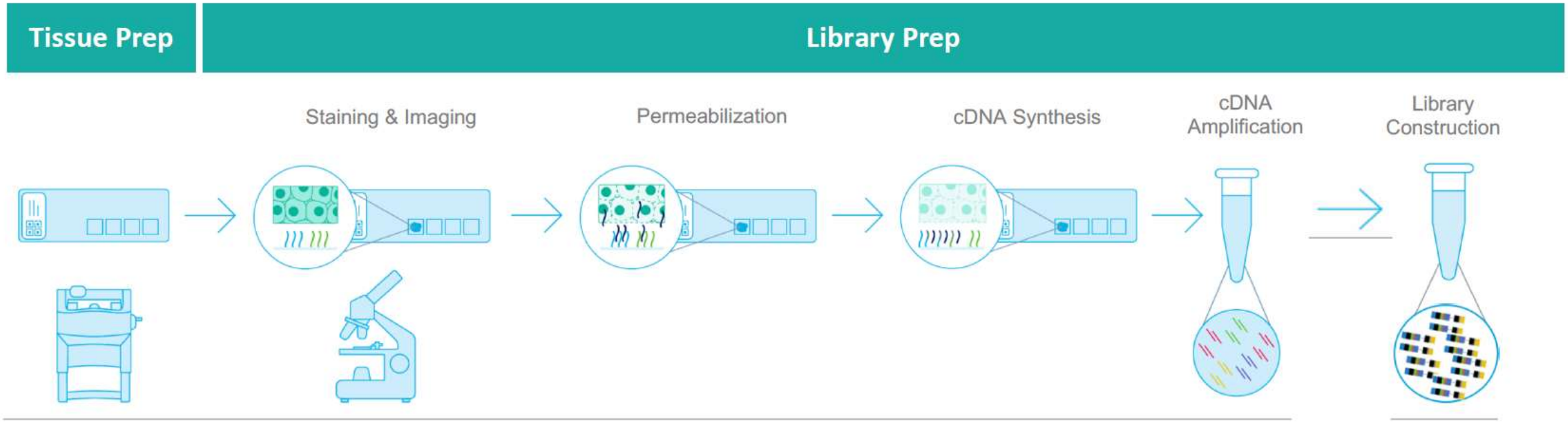
- *Generate cDNA*
- *Denature cDNA*

10x Genomics Chromium Single Cell Gene Expression





Inputs for raw sequencing-data processing



(1) A High-resolution histological image, such as H&E-stained RGB image

(3) A reference genome and GTF annotation compatible with Space Ranger

(2) Paired end FASTQ files
Read 1: spatial barcode and UMI
Read 2: transcript sequence
PE 28 x 90

```
spaceranger mkgtf Zm-B73-Reference-NAM-5.0.gtf Zm-B73-Reference-NAM-5.0_filtered.gtf \
                --attribute=gene_biotype:protein_coding \
                --attribute=gene_biotype:lncRNA \
                --attribute=gene_biotype:antisense
```

Inputs for spaceranger count

Supported workflows

1. Visium v1
    1. For fresh frozen (Poly-dT)
    2. For FFPE (**probe-based**)
2. Visium v2 + **CytAssist**
    1. For fresh frozen (**probe-based**)
    2. For FFPE (**probe-based**)
3. Visium HD + **CytAssist** (**probe-based**)
4. Visium 3’ +**CytAssist** (Poly-dT)

Inputs for spaceranger count

1. Sequencing FASTQ files
2. Prebuilt reference transcriptome
3. Slide-layout file (if required)
4. Slide information
    1. Slide serial number
    2. Capture area
5. Histological image
6. ***Probe-set information (--probe-set=)***
7. ***CytAssist image + JASON file***

V1 Fresh Frozen (Poly-dT)

```
spaceranger count --id=sample345 \ #Output directory
                  --transcriptome=/home/jdoe/refdata/GRCh38-2020-A \ #Path to Reference
                  --fastqs=/home/jdoe/runs/HAWT7ADXX/outs/fastq_path \ #Path to FASTQs
                  --sample=mysample \ #Sample name from FASTQ filename
                  --image=/home/jdoe/runs/images/sample345.tiff \ #Path to brightfield image
                  --slide=V19J01-123 \ #Slide ID
                  --area=A1 \ #Capture area
                  --localcores=8 \ #Allowed cores in localmode
                  --localmem=64  #Allowed memory (GB) in localmode
```

```
spaceranger count --id=sample345 \ #Output directory
            --transcriptome=/home/jdoe/refdata/GRCh38-2020-A \ #Path to Reference
            --fastqs=/home/jdoe/runs/HAWT7ADXX/outs/fastq_path \ #Path to FASTQs
            --sample=mysample \ #Sample name from FASTQ filename
            --image=/home/jdoe/runs/images/sample345.tiff \ #Path to brightfield image
            --slide=V19J01-123 \ #Slide ID
            --area=A1 \ #Capture area A1, B1, C1, D1
            --localcores=8 \ #Allowed cores in localmode
            --localmem=64  #Allowed memory (GB) in localmode
```

[Sample Name]_S1_L00[Lane Number]_[Read Type]_001.fastq.gz
or
[Sample Name]_S1_[Read Type]_001.fastq.gz

Where Read Type is one of:

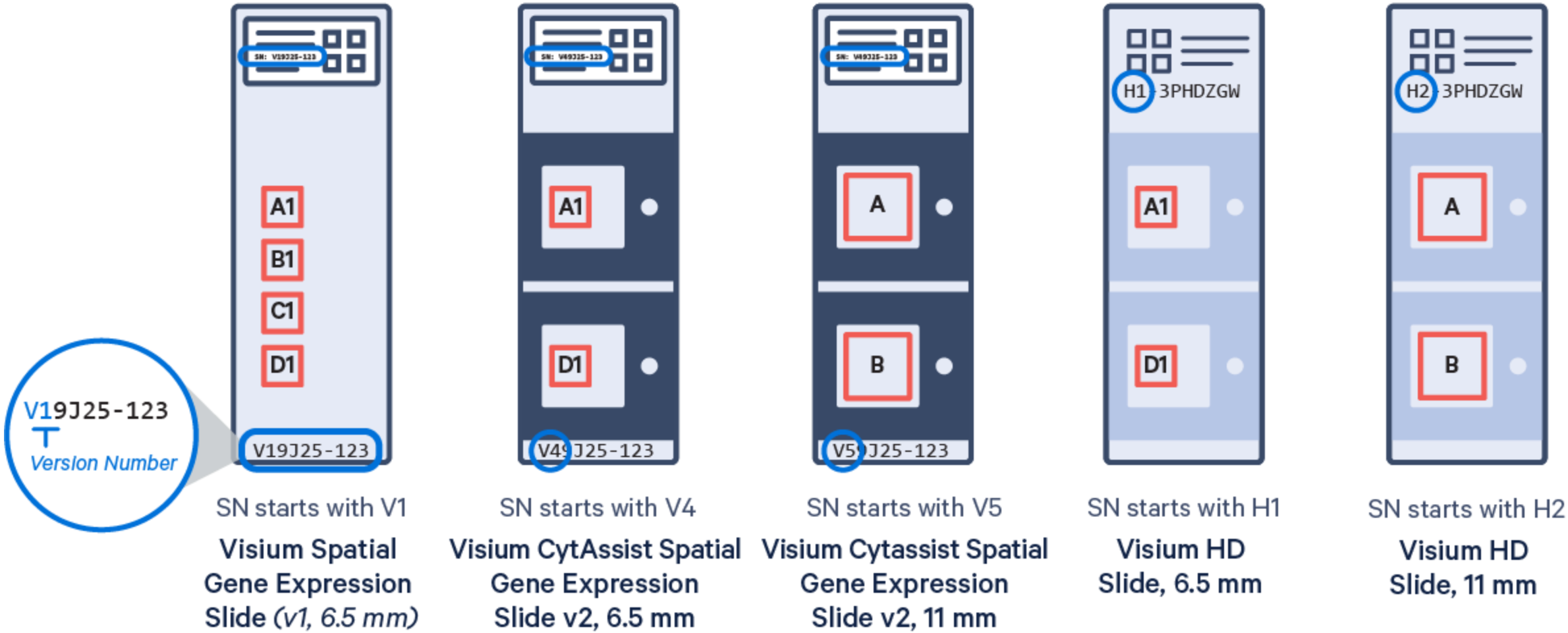
- **R1: Read 1**
- **R2: Read 2**
- I1: Sample index read (optional)
- I2: Sample index read (optional)

```
spaceranger count --fastqs=/flowcell_1,/flowcell_2 --sample=Sample-GA-A1
```

```
├── flowcell_1
│   ├── Sample-GA-A1
│   │   ├── Sample-GA-A1_S1_L001_I1_001.fastq.gz
│   │   ├── Sample-GA-A1_S1_L001_R1_001.fastq.gz
│   │   ├── Sample-GA-A1_S1_L001_R2_001.fastq.gz
│   │   └── Sample-GA-A1_S1_L001_I2_001.fastq.gz
├── flowcell_2
│   ├── Sample-GA-A1
│   │   ├── Sample-GA-A1_S1_L001_I1_001.fastq.gz
│   │   ├── Sample-GA-A1_S1_L001_R1_001.fastq.gz
│   │   ├── Sample-GA-A1_S1_L001_R2_001.fastq.gz
│   │   └── Sample-GA-A1_S1_L001_I2_001.fastq.gz
```

Layouts of 10x Gene Expression Slide

Serial Number locations
Capture Areas



SN starts with V1

Visium Spatial Gene Expression Slide *(v1, 6.5 mm)*

SN starts with V4

Visium CytAssist Spatial Gene Expression Slide v2, 6.5 mm

SN starts with V5

Visium Cytassist Spatial Gene Expression Slide v2, 11 mm

SN starts with H1

Visium HD Slide, 6.5 mm

SN starts with H2

Visium HD Slide, 11 mm

Batch_Run.sh

```
#S210_B73V5_short_MTCL
./spaceranger count --id=XGE21_VR03_B73V5S210_short_MTCL --transcriptome=/home/macchihlee/ST/spaceranger/spaceranger-2.1.0/B73_V5_S210_MT_short
--fastqs=/home/macchihlee/ST/spaceranger/spaceranger-
1.2.2/maizeST_04192021/20210408_S1_nx003_Wen_Hsiung_Li/Project_20210408_S1_nx003_Wen_Hsiung_Li/XGE21-VR03 --sample=XGE21-VR03 --
image=/home/macchihlee/ST/spaceranger/spaceranger-
1.2.2/maizeST_04192021/20210408_S1_nx003_Wen_Hsiung_Li/Project_20210408_S1_nx003_Wen_Hsiung_Li/GEImages/Slide4_Region_01020210324_GE_C1.ti
f --slide=V10F04-109 --area=C1 --localcores=8 --localmem=64

./spaceranger count --id=XGE21_VR02_B73V5S210_short_MTCL --transcriptome=/home/macchihlee/ST/spaceranger/spaceranger-2.1.0/B73_V5_S210_MT_short
--fastqs=/home/macchihlee/ST/spaceranger/spaceranger-
1.2.2/maizeST_04192021/20210408_S1_nx003_Wen_Hsiung_Li/Project_20210408_S1_nx003_Wen_Hsiung_Li/XGE21-VR02 --sample=XGE21-VR02 --
image=/home/macchihlee/ST/spaceranger/spaceranger-
1.2.2/maizeST_04192021/20210408_S1_nx003_Wen_Hsiung_Li/Project_20210408_S1_nx003_Wen_Hsiung_Li/GEImages/Slide4_Region_00920210324_GE_B1.ti
f --slide=V10F04-109 --area=B1 --localcores=8 --localmem=64

./spaceranger count --id=XGE21_VR04_B73V5S210_short_MTCL --transcriptome=/home/macchihlee/ST/spaceranger/spaceranger-2.1.0/B73_V5_S210_MT_short
--fastqs=/home/macchihlee/ST/spaceranger/spaceranger-
1.2.2/maizeST_04192021/20210408_S1_nx003_Wen_Hsiung_Li/Project_20210408_S1_nx003_Wen_Hsiung_Li/XGE21-VR04 --sample=XGE21-VR04 --
image=/home/macchihlee/ST/spaceranger/spaceranger-
1.2.2/maizeST_04192021/20210408_S1_nx003_Wen_Hsiung_Li/Project_20210408_S1_nx003_Wen_Hsiung_Li/GEImages/Slide4_Region_01120210324_GE_D1.ti
f --slide=V10F04-109 --area=D1 --localcores=8 --localmem=64
.
.
.
.
.
```

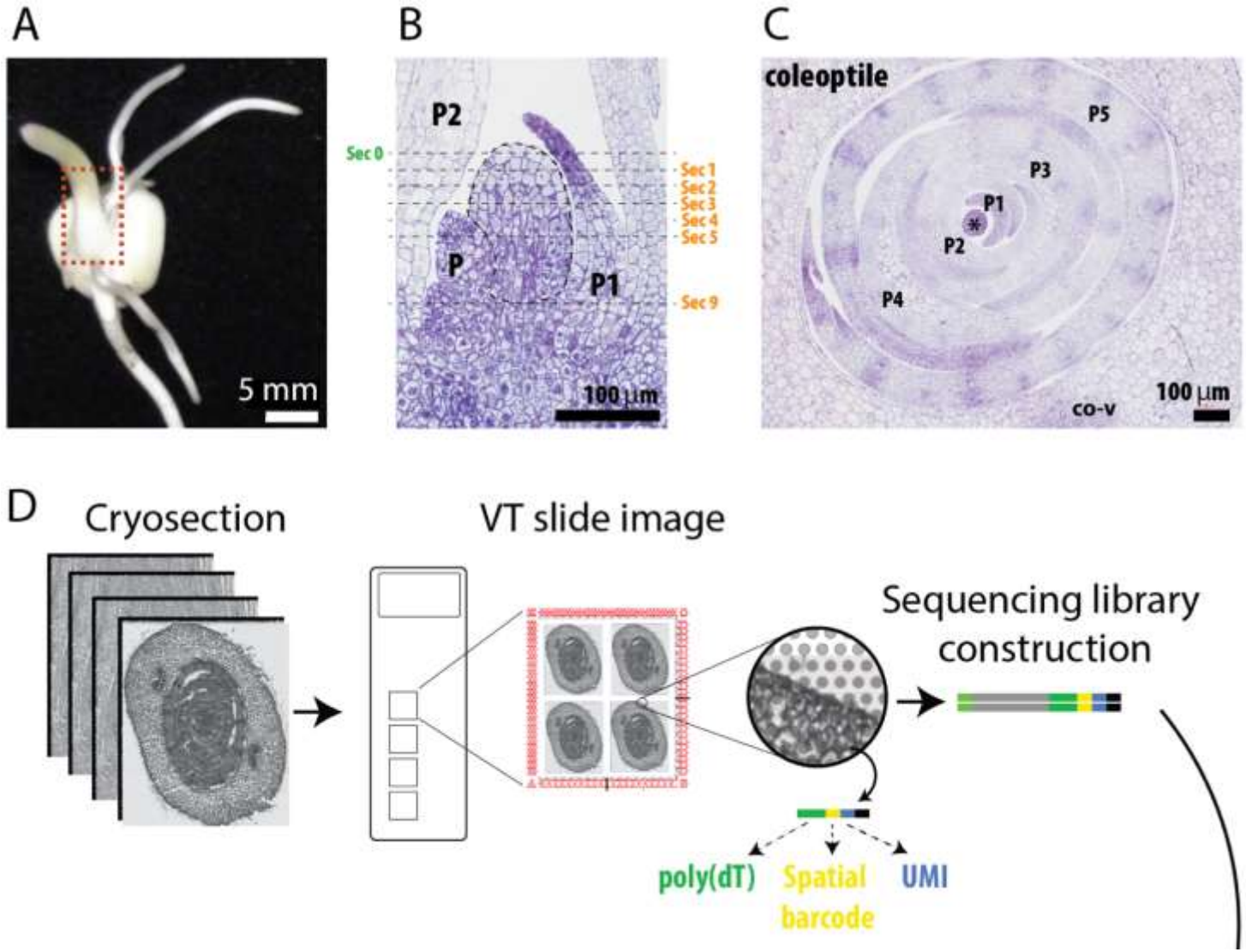
Major outputs from spaceranger count

Visium (55)

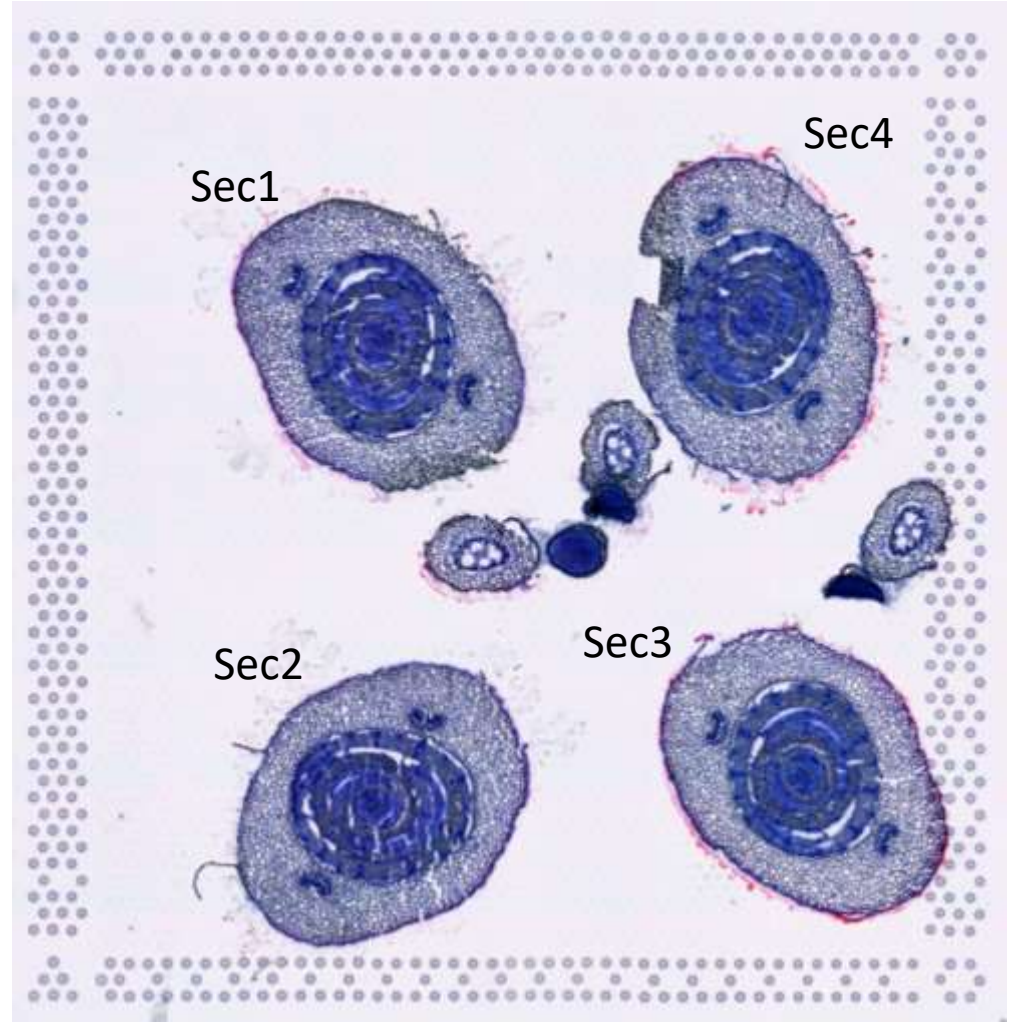
```
sample_id/outs/
│
├── web_summary.html              QC report
├── metrics_summary.csv           QC metrics
├── cloupe.cloupe                 Loupe Browser
│
├── filtered_feature_bc_matrix.h5 Tissue spots: gene × spot counts
├── raw_feature_bc_matrix.h5      All barcodes
├── molecule_info.h5              Per-molecule data; used by aggr
├── possorted_genome_bam.bam      Aligned sequencing reads
│
├── spatial/
│   ├── tissue_positions.csv      Spot coordinates
│   ├── scalefactors_json.json    Image scaling
│   ├── tissue_hires_image.png    High-resolution image
│   └── tissue_lowres_image.png   Low-resolution image
│
└── analysis/
    ├── pca/
    ├── umap/
    ├── clustering/
    └── diffexp/
```

Visium HD (2)

```
sample_id/outs/
│
├── web_summary.html          # 分析結果與 QC 報告
├── metrics_summary.csv       # QC 指標表
├── molecule_info.h5          # 每個分子的資訊
├── feature_slice.h5          # HD 基因表現空間切片
├── barcode_mappings.parquet  # 2/8/16 µm bins 與細胞的對應
├── spatial/                  # 影像、組織定位與配準結果
│
├── binned_outputs/           # 方形 bin 層級
│   ├── square_002um/         # 原生 2 µm resolution
│   │   ├── filtered_feature_bc_matrix.h5
│   │   ├── raw_feature_bc_matrix.h5
│   │   └── spatial/
│   │
│   ├── square_008um/         # 常用分析尺度
│   │   ├── filtered_feature_bc_matrix.h5
│   │   ├── raw_feature_bc_matrix.h5
│   │   ├── analysis/
│   │   ├── spatial/
│   │   └── cloupe.cloupe
│   │
│   └── square_016um/         # 較高 sensitivity、較低解析度
│       ├── filtered_feature_bc_matrix.h5
│       ├── raw_feature_bc_matrix.h5
│       ├── analysis/
│       └── spatial/
│
└── segmented_outputs/        # 啟用細胞分割時產生
```


A
5 mm
B
P2
P
P1
Sec 0
Sec 1
Sec 2
Sec 3
Sec 4
Sec 5
Sec 9
100 µm
C
coleoptile
P5
P3
P1
P2
P4
co-v
100 µm
D
Cryosection
VT slide image
Sequencing library construction
poly(dT)
Spatial barcode
UMI

VR03


Sec1
Sec4
Sec2
Sec3

Review Space Ranger outputs before analysis

1. Before opening R, inspect:
    1. web_summary.html for sequencing, mapping, tissue-detection, and spot-level QC.
    2. Cloupe.cloupe file in Loupe Browser for basic exploration.

XGE21_VR03

Summary | Analysis

1,935
Number of Spots Under Tissue

62,345
Mean Reads per Spot

4,862
Median Genes per Spot

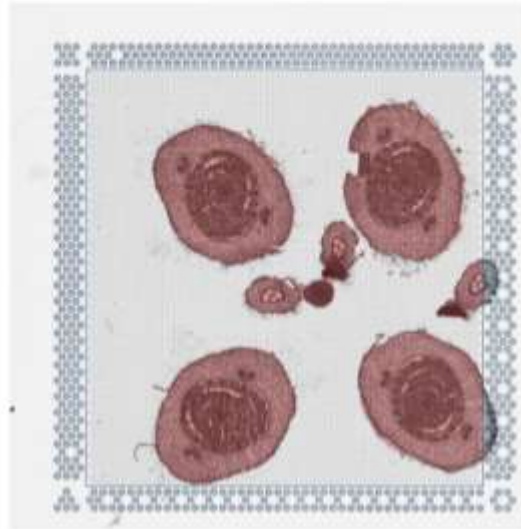
Sequencing

Number of Reads	120,637,490
Valid Barcodes	97.2%
Valid UMIs	100.0%
Sequencing Saturation	61.6%
Q30 Bases in Barcode	95.4%
Q30 Bases in RNA Read	94.3%
Q30 Bases in UMI	95.5%

Mapping

Reads Mapped to Genome	93.4%
Reads Mapped Confidently to Genome	89.1%
Reads Mapped Confidently to Intergenic Regions	9.9%
Reads Mapped Confidently to Intronic Regions	0.8%
Reads Mapped Confidently to Exonic Regions	78.4%
Reads Mapped Confidently to Transcriptome	73.5%
Reads Mapped Antisense to Gene	0.4%

Spots

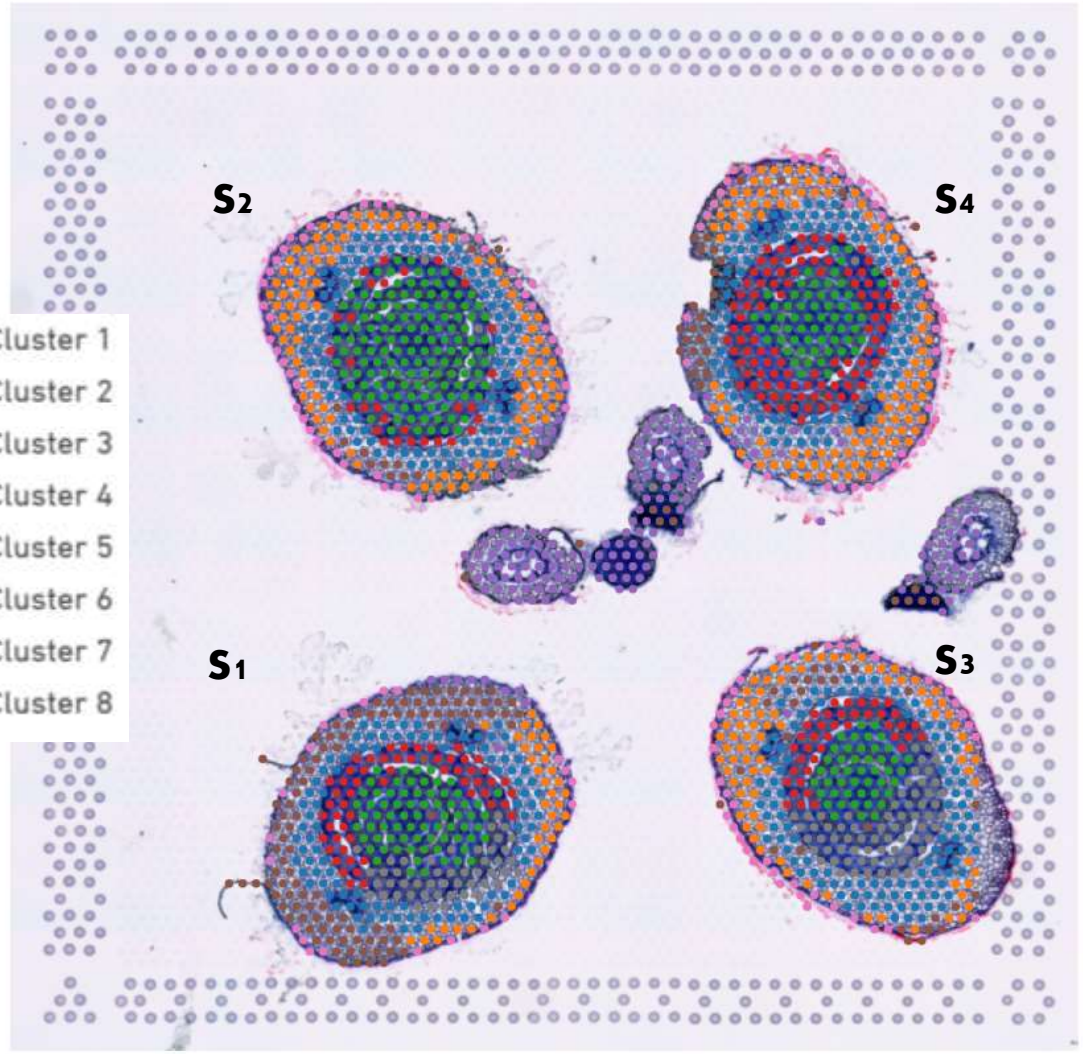


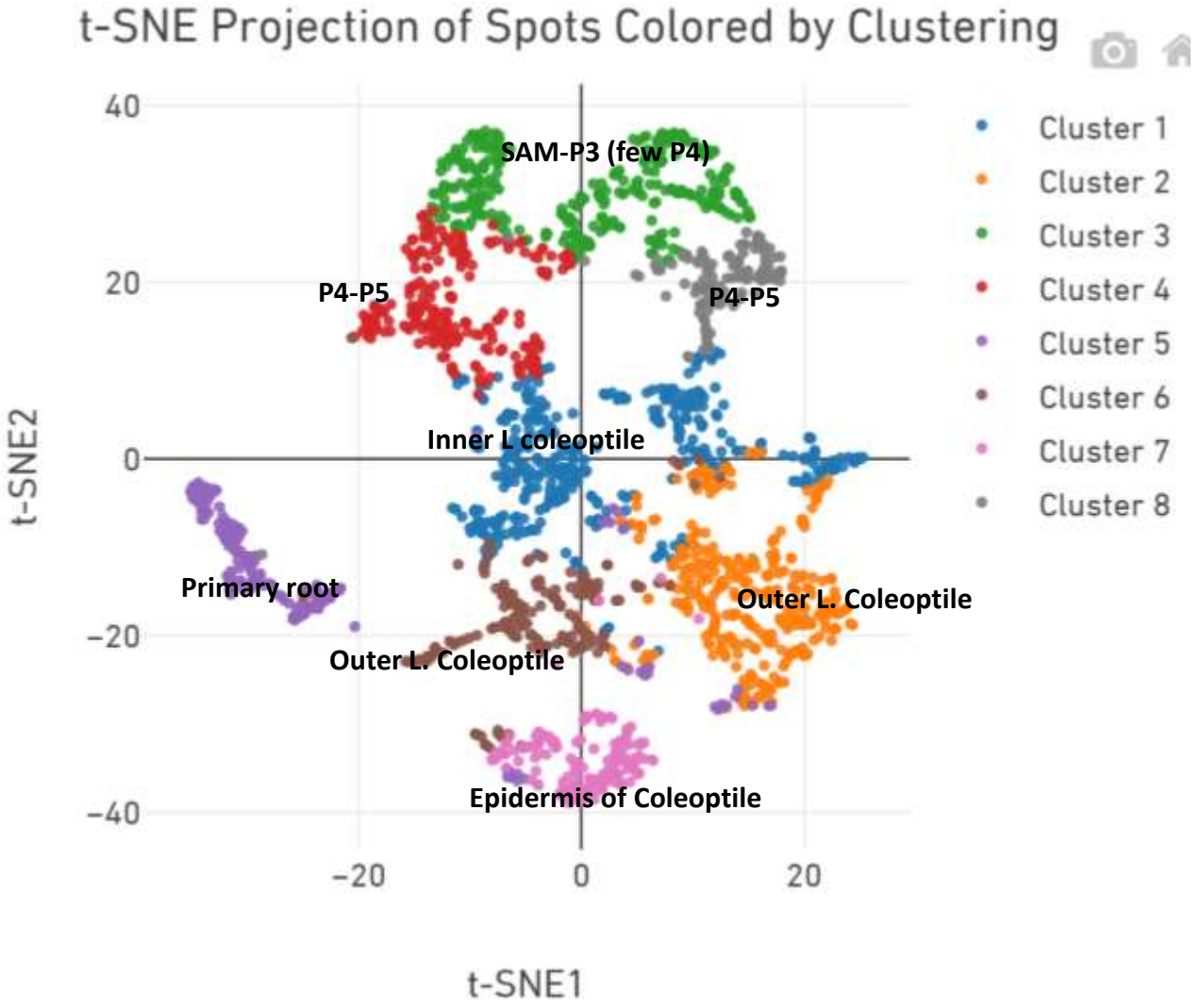
Fraction Reads in Spots Under Tissue	78.9%
Mean Reads per Spot	62,345
Mean Reads Under Tissue per Spot	47,237
Median Genes per Spot	4,862
Total Genes Detected	27,989
Median UMI Counts per Spot	11,936

Sample

Sample ID	XGE21_VR03
Sample Description	
Chemistry	Spatial 3' v1
Slide Serial Number	V10F04-109-C1
Reference Path	..aceranger/spaceranger-1.2.2/B73_v499
Transcriptome	B73_v499-
Pipeline Version	spaceranger-1.2.2

XG21_VR3


S2
S4
S1
S3
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
1 mm


t-SNE Projection of Spots Colored by Clustering
SAM-P3 (few P4)
P4-P5
P4-P5
Inner L coleoptile
Primary root
Outer L. Coleoptile
Outer L. Coleoptile
Epidermis of Coleoptile
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
t-SNE2
t-SNE1
40
20
0
−20
−40
−20
0
20

Load Space Ranger output into Seurat

```
XGE21_VR03/
 ├── filtered_feature_bc_matrix.h5
 └── spatial/
    ├── scalefactors_json.json
    ├── tissue_positions_list.csv
    └── tissue_hires_image.png
    └── tissue_lower_image.png
```

```
data_dir<-“C:/path…/XGE21_VR03/"
```

```
se <- Load10X_Spatial(
  data_dir,
  filename =
"filtered_feature_bc_matrix.h5",
image.name =
"tissue_hires_image.png",
)
```

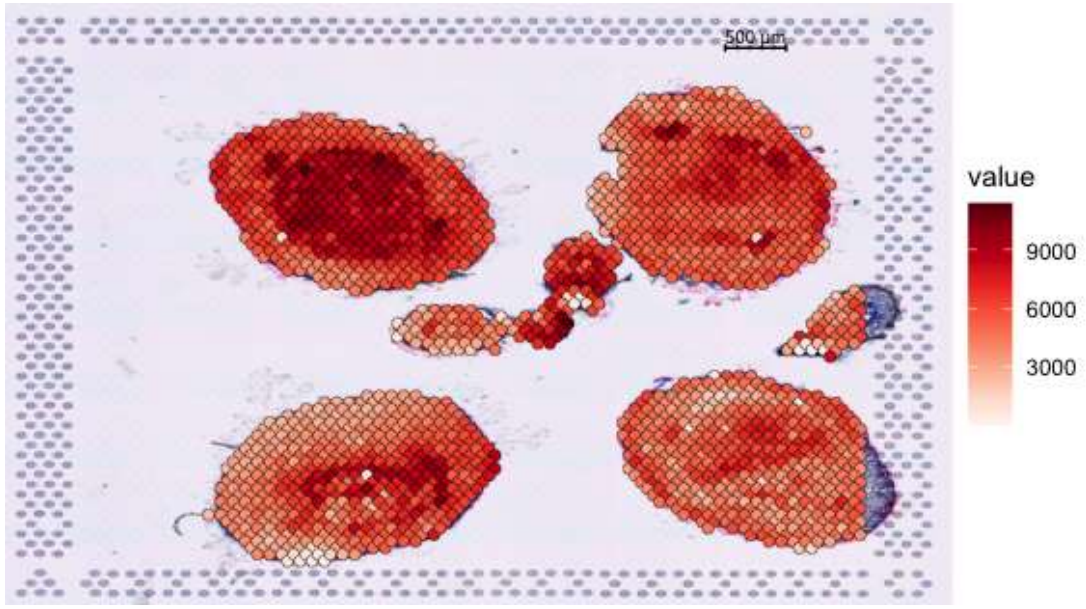
VR03_S5 | VR03_SE

Show Attributes

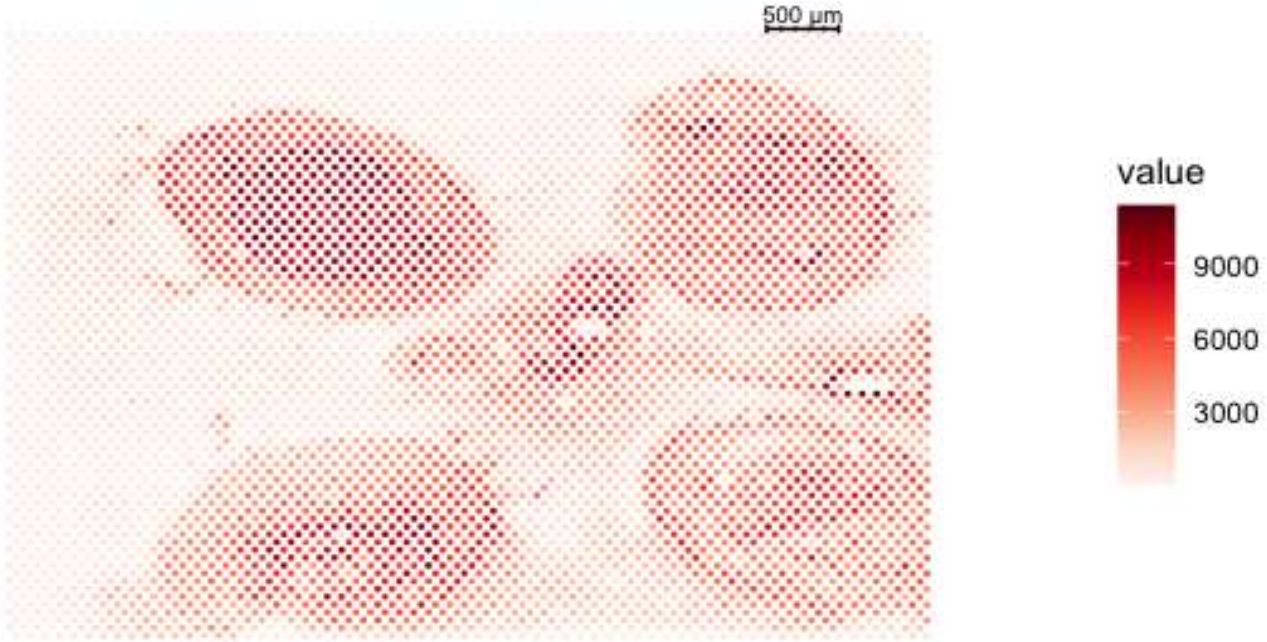
Name	Type	Value
VR03_S5	S4 [40109 x 1935] (SeuratObject:	S4 object of class Seurat
assays	list [1]	List of length 1
RNA	S4 [40109 x 1935] (SeuratObject:	S4 object of class Assay5
meta.data	list [1935 x 3] (S3: data.frame)	A data.frame with 1935 rows and 3 columns
orig.ident	factor	Factor with 1 level: "SeuratProject"
nCount_RNA	double [1935]	12424 16600 16725 10985 6994 23964 …
nFeature_RNA	integer [1935]	4689 5887 6027 4455 3145 6713 …
active.assay	character [1]	'RNA'
active.ident	factor	Factor with 1 level: "SeuratProject"
graphs	list [0]	List of length 0
neighbors	list [0]	List of length 0
reductions	list [0]	List of length 0
images	list [1]	List of length 1
VR03	S4 [600 x 585] (Seurat::VisiumV1)	S4 object of class VisiumV1
project.name	character [1]	'SeuratProject'
misc	list [0]	List of length 0
version	list [1] (S3: package_version, num	List of length 1
commands	list [0]	List of length 0
tools	list [0]	List of length 0

VR03_S5

Inspect raw read spatial profile and tissue coverage

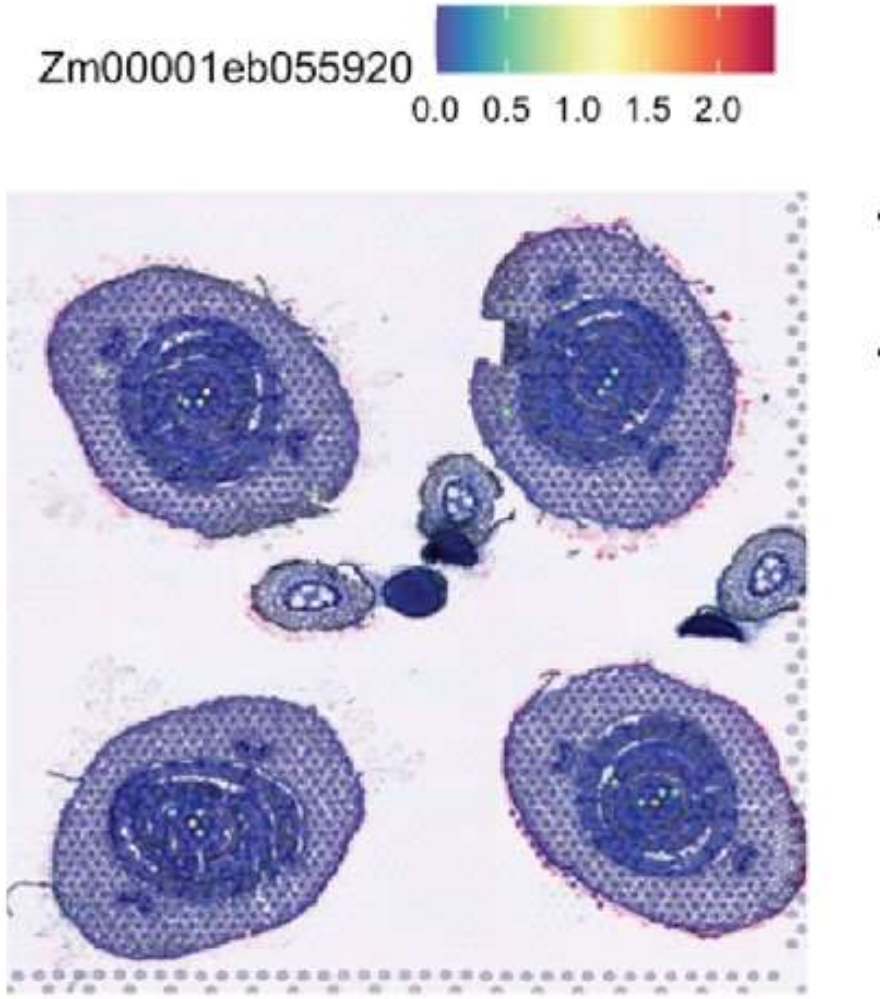


FeatureOverlay (STu)

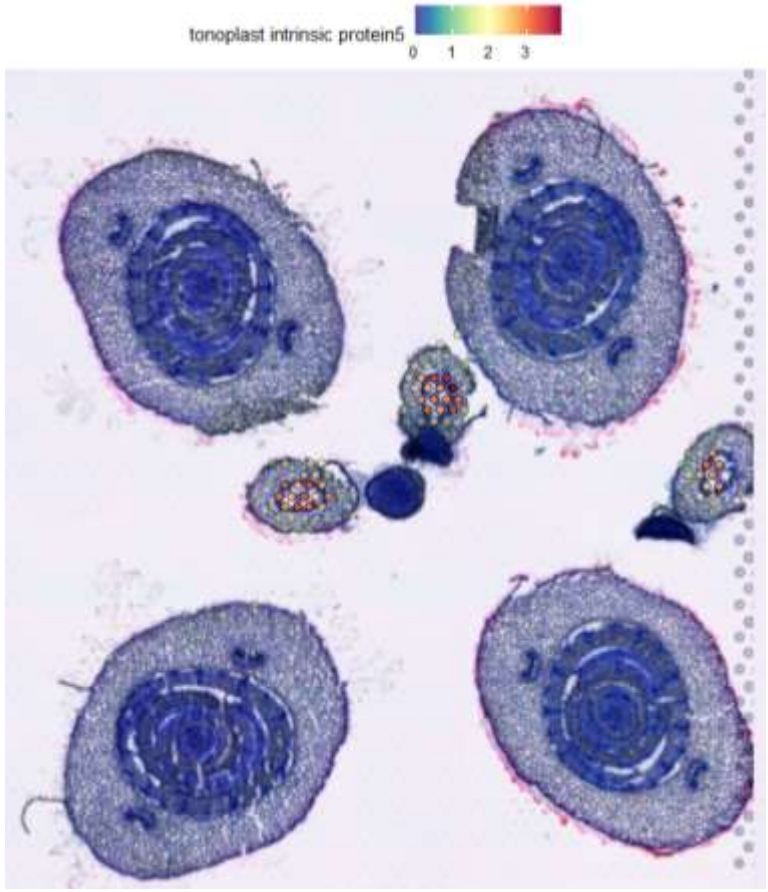


ST.FeaturePlot

Validate the dataset with known marker genes

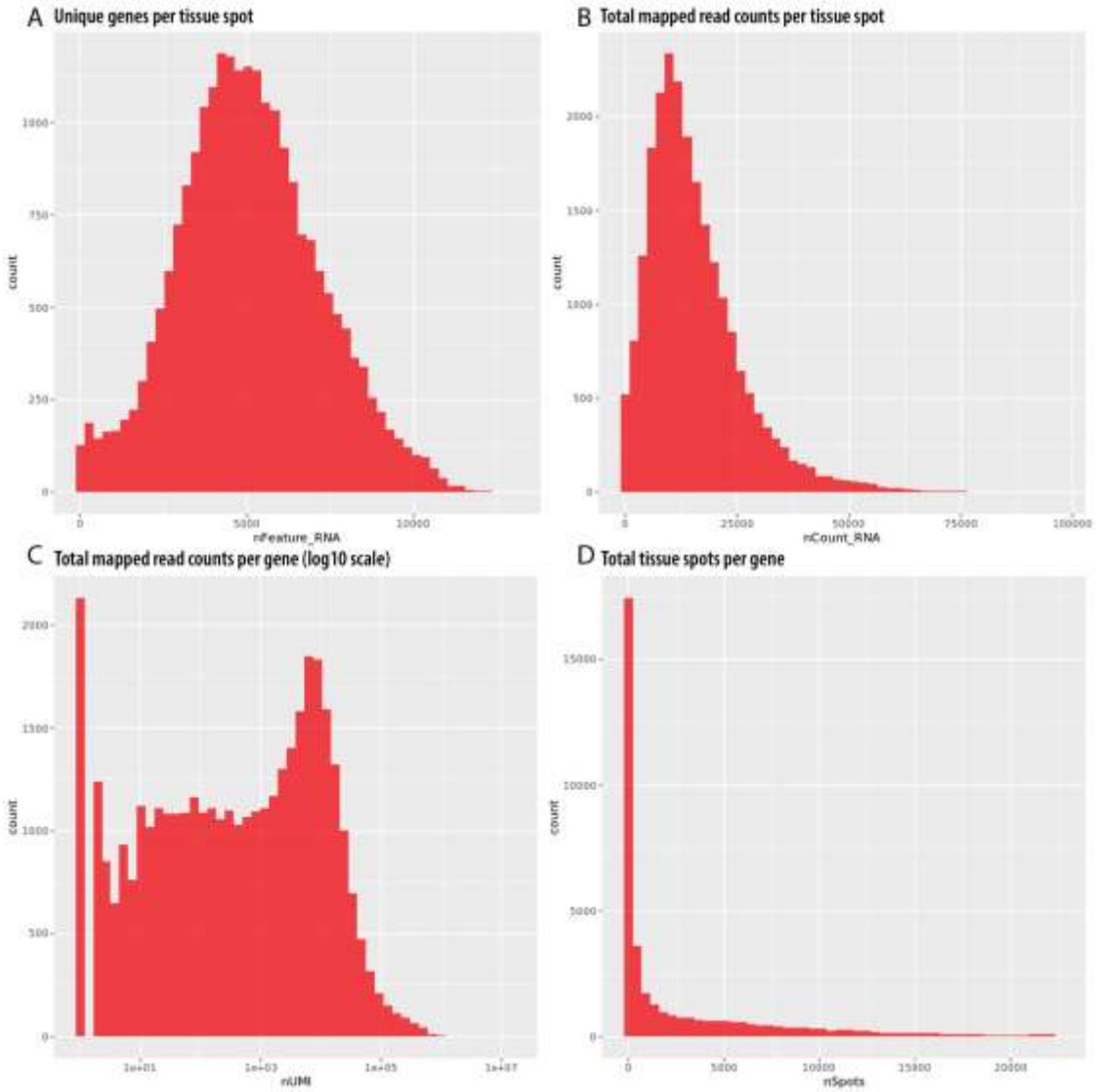


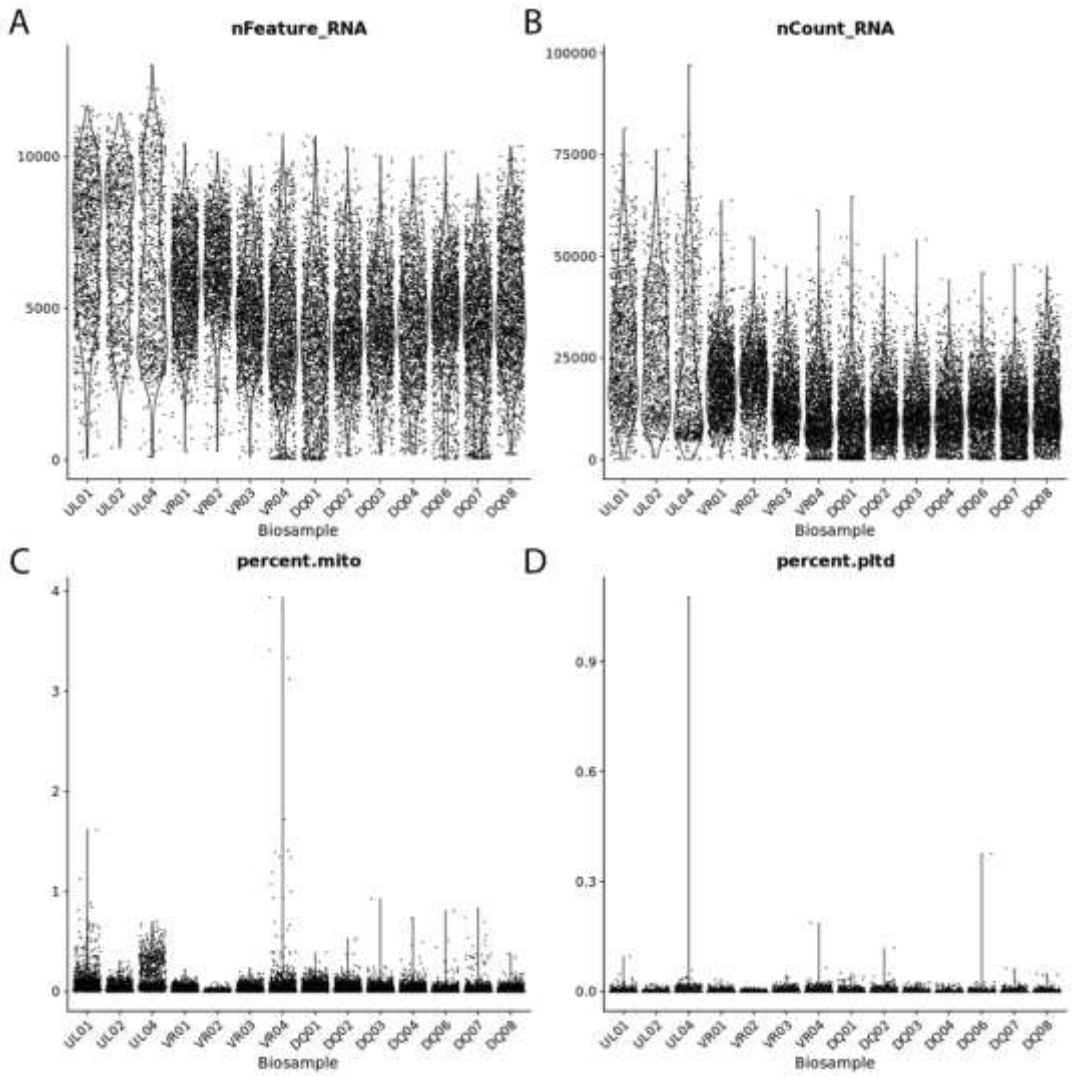
KNOTTED1 (*KN1*)
known marker of the shoot apical meristem.



Tonoplast intrinsic proteins (TIPs), marker for vascular compartments (Gattolin et al., 2009)

3. Inspect count and feature QC metrics





Merging, normalization, and integration answer different questions

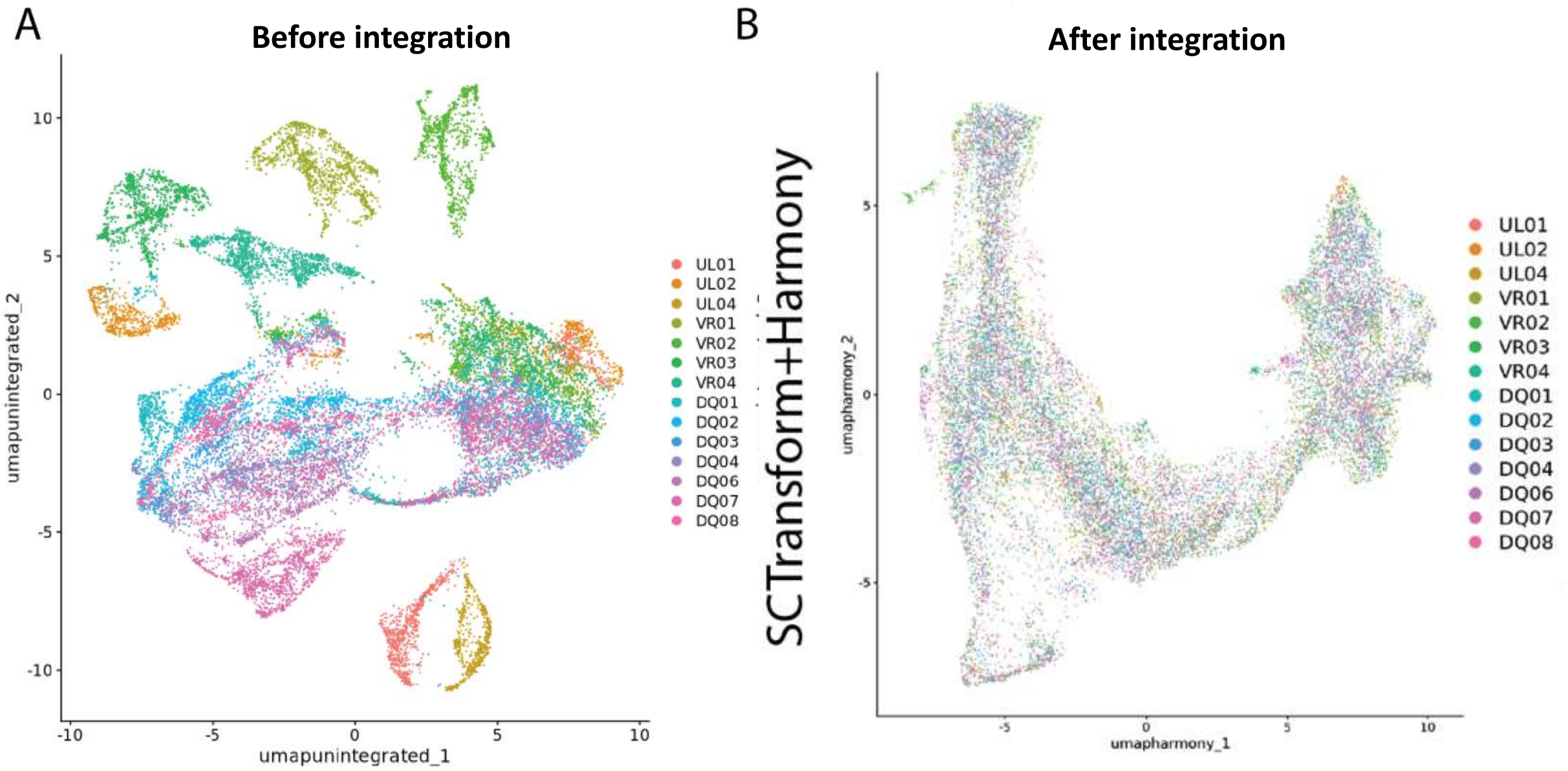
Function	Primary purpose	Effect on expression data	Batch correction?
merge()	Combines capture areas into one Seurat object	Combines existing count/data matrices; does not transform their values	No
SCTransform()	Normalizes and variance-stabilizes UMI data	Creates a new SCT assay with corrected counts, log-corrected data, and Pearson residuals; original raw counts remain unchanged	Not explicit batch correction
IntegrateLayers()	Aligns datasets/layers across batches	Usually creates a batch-corrected dimensional reduction; does not normally replace raw or SCT expression values	Yes, for unwanted dataset-level variation

```
# merge

# IntegrateLayers
```

https://satijalab.org/seurat/articles/seurat5_integration

Integration should reduce technical separation without removing biological structure



Check the active assay before running downstream analysis

Assays(dataset)

DefaultAssay(dataset)

DefaultAssay(dataset)<-"RNA"

DefaultAssay(dataset)<-"SCT"

Name	Type	Value
XGE202122_S5_leaf_SCT_...	S4 [26364 x 4701] (SeuratObject:	S4 object of class Seurat
assays	list [2]	List of length 2
RNA	S4 [40109 x 4701] (SeuratObject:	S4 object of class Assay5
SCT	S4 [26364 x 4701] (Seurat::SCTAs	S4 object of class SCTAssay
meta.data	list [4701 x 19] (S3: data.frame)	A data.frame with 4701 rows and 19 columns
orig.ident	factor	Factor with 14 levels: "UL01", "UL02", "UL04", "VR01", "VR02", "VR03", ...
nCount_RNA	double [4701]	48360 54010 64139 59049 62863 48968 ...
nFeature_RNA	integer [4701]	10983 11156 12248 10915 12182 11038 ...
sample_id	factor	Factor with 53 levels: "DQ01_S1", "DQ01_S2", "DQ01_S3", "DQ02_S1", "DQ02_S3", "D, ...
labels	character [4701]	'Section2' 'Section1' 'Section3' 'Section2' 'Section3' 'Section4' ...
domains	factor	Factor with 5 levels: "SAM", "P1_P2", "P3", "P4", "P5"
sample	factor	Factor with 14 levels: "UL01", "UL02", "UL04", "VR01", "VR02", "VR03", ...
section	factor	Factor with 9 levels: "Section1", "Section2", "Section3", "Section4", "Section5", ...
spot_inS5	character [4701]	'AAACCCGAACGAAATC-1_1_1' 'AAACGTGTTCGCCCTA-1_1_1' 'AAAGGGCAGCTTGAAT-1_
spot_inSE	character [4701]	'AAACCCGAACGAAATC-1' 'AAACGTGTTCGCCCTA-1' 'AAAGGGCAGCTTGAAT-1' 'AAATCTA
nCount_SCT	double [4701]	30208 30108 29934 29881 29997 30089 ...
nFeature_SCT	integer [4701]	10180 9559 9576 8661 9557 10092 ...
seurat_clusters	factor	Factor with 11 levels: "0", "1", "5", "6", "8", "10", ...

Add Loupe Browser spot annotations to Seurat metadata

Annotate spots in Loupe Browser

↓

Export section_number.csv

↓

Load the CSV and Seurat object

↓

Replace blank labels with "other"

↓

Match annotations by barcode

↓

Add "section" to Seurat metadata

↓

Verify counts and spatial locations

↓

Save a new annotated Seurat object

Write a complete, copy-ready R script for importing manual spatial-spot annotations exported from 10x Genomics Loupe Browser into a Seurat object.
Inputs:
•Seurat object: VR03_S5.rds
•Loupe annotation file: section_number.csv
•The CSV contains the columns Barcode and section.
•Valid section labels are Section01, Section02, Section03, and Section04.
•Blank, NA, N/A, or n/a section labels must be converted to "other".

```
setwd("C:/Users/user/Dropbox/2025_IPMB/classes/2026_08_SciLife_library_bioinformatics/supp_files")

source( "C:/Users/user/Dropbox/2025_IPMB/classes/2026_08_SciLife_library_bioinformatics/supp_files/workshop_add_loupe_section_metadata.R" )
```

Integrate single-cell RNA-seq (scRNA-seq) data with spatial transcriptomics data

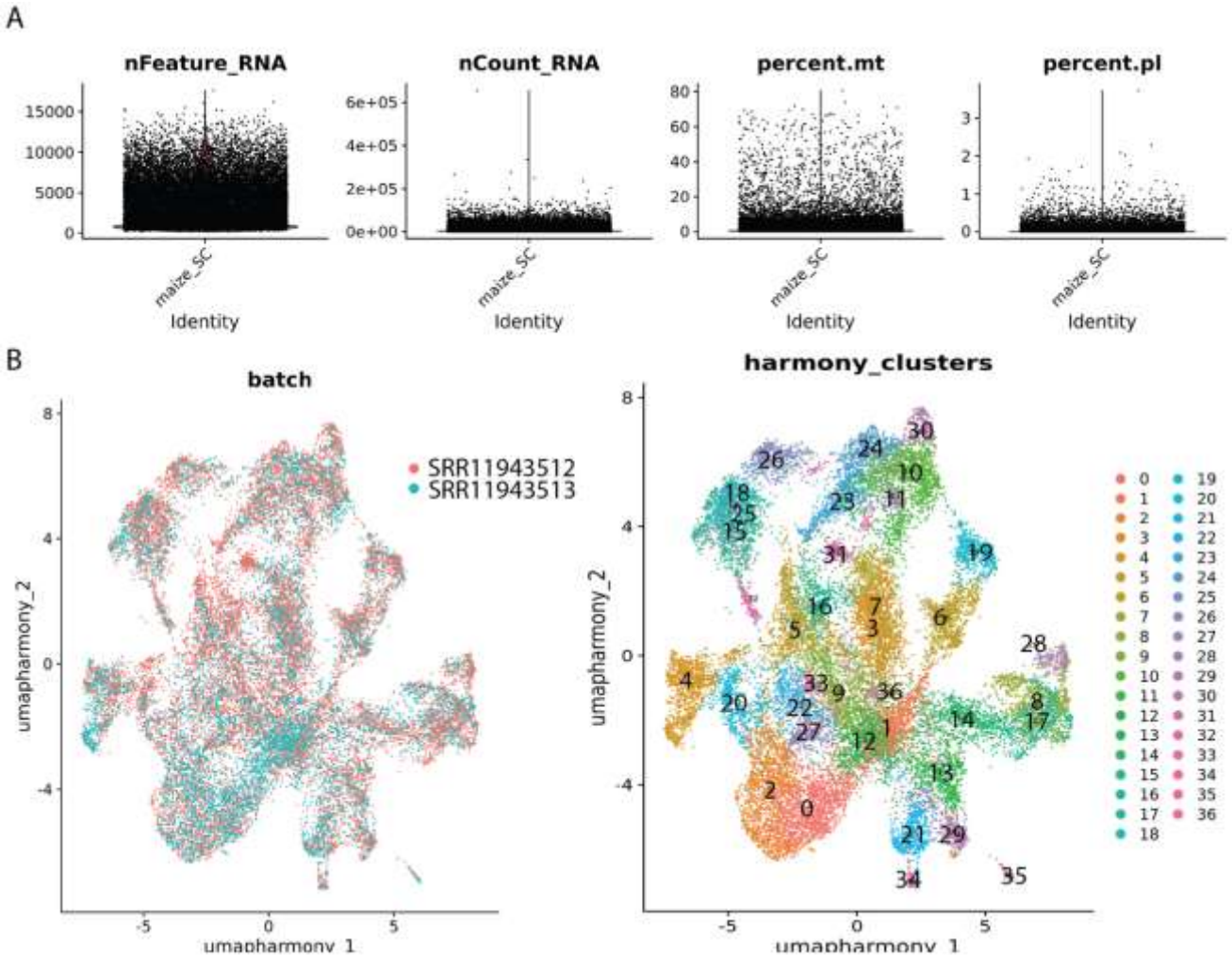
1. Prepare and Perform scRNA-seq quality control
   - Filter low-quality cells using gene counts, UMI counts, and mitochondrial/plastid percentages.
   - Detect doublets separately in each library with scDblFinder.

2. Annotate scRNA-seq cell types
   - Prepare curated **marker-gene signatures**.
   - Run **SCINA** to assign cell identities and probabilities.
   - Review assignments using clusters, UMAP, and marker expression.

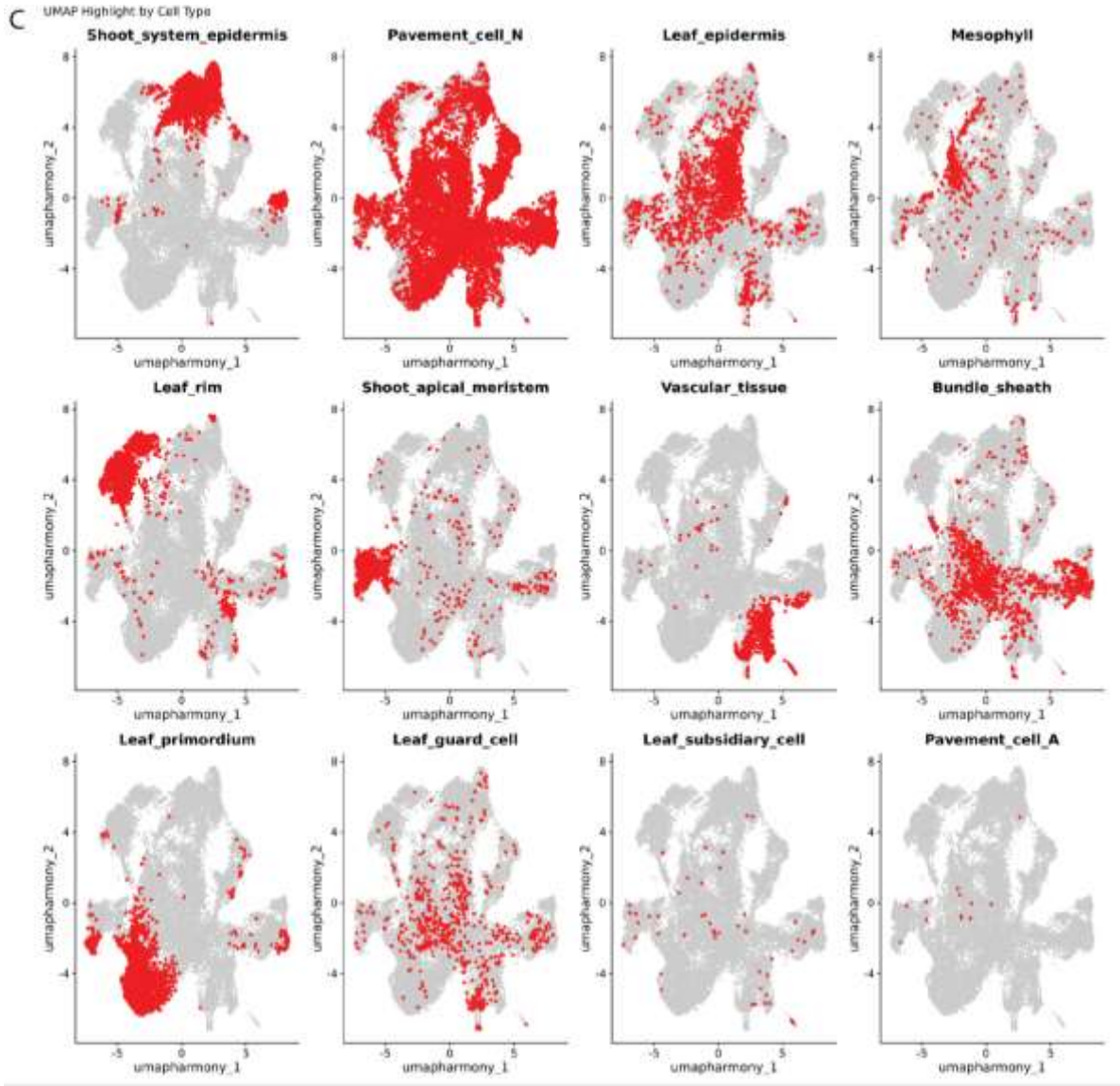
3. Deconvolve spatial spots
   - Harmonize gene identifiers between scRNA-seq and Visium data.
   - Select cell-type markers and highly variable genes.
   - Run SPOTlight using annotated scRNA-seq cells as the reference.
   - Add estimated **cell-type proportions** to Visium metadata.

4. Visualize spatial cell-type distributions
   - Plot continuous **cell-type proportions**.
   - Display estimates on spatial coordinates.

1. Prepare and Perform scRNA-seq quality control

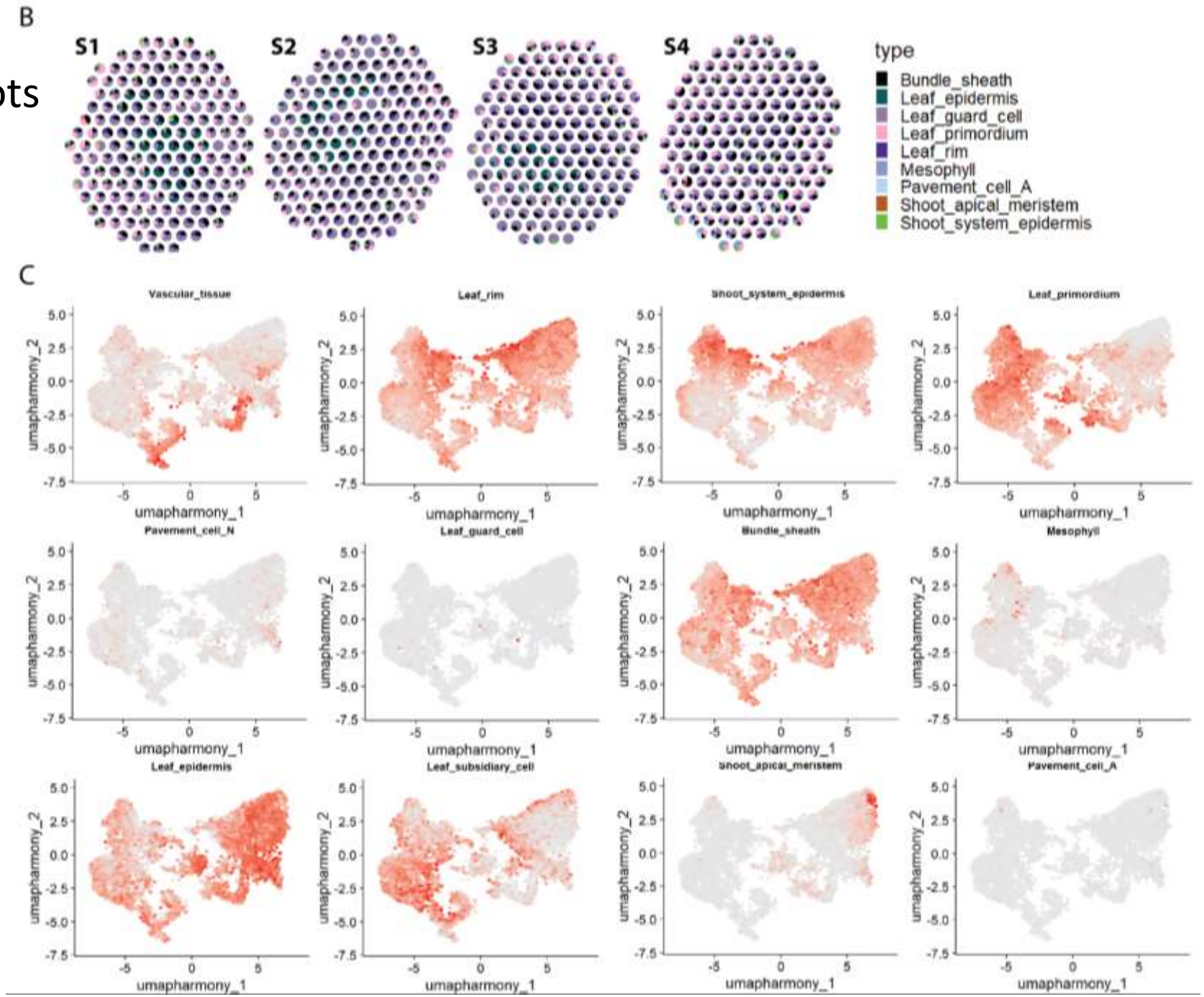


Annotate scRNA-seq cell types

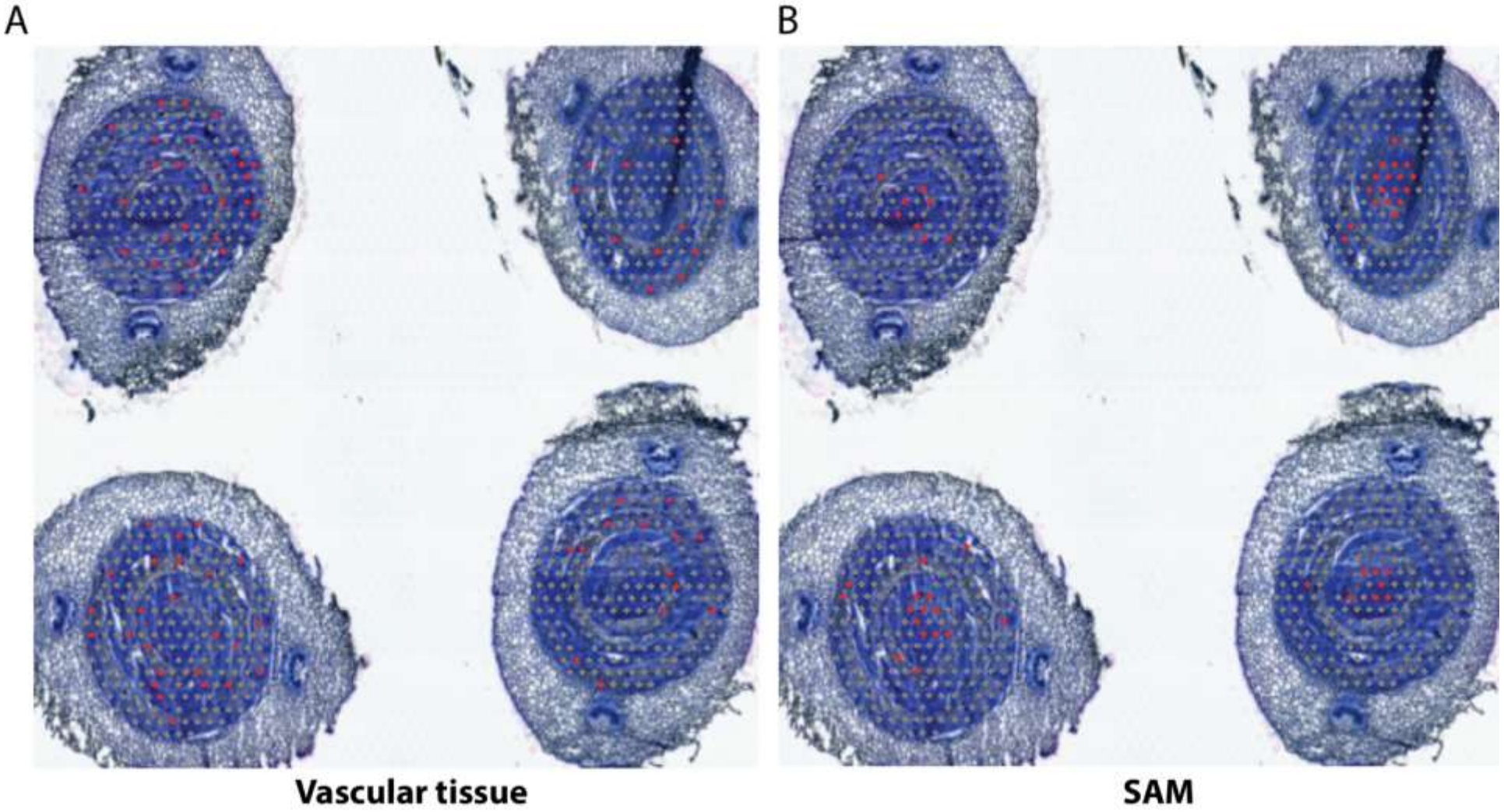


3. Deconvolve spatial spots

4. Plot continuous **cell-type proportions**



Display annotated cell-type on spatial coordinates



Seurat object 的資料層

Assays
Spatial / SCT / integrated

Cells / spots
barcode 是共同索引

Metadata
sample / section / QC / domain

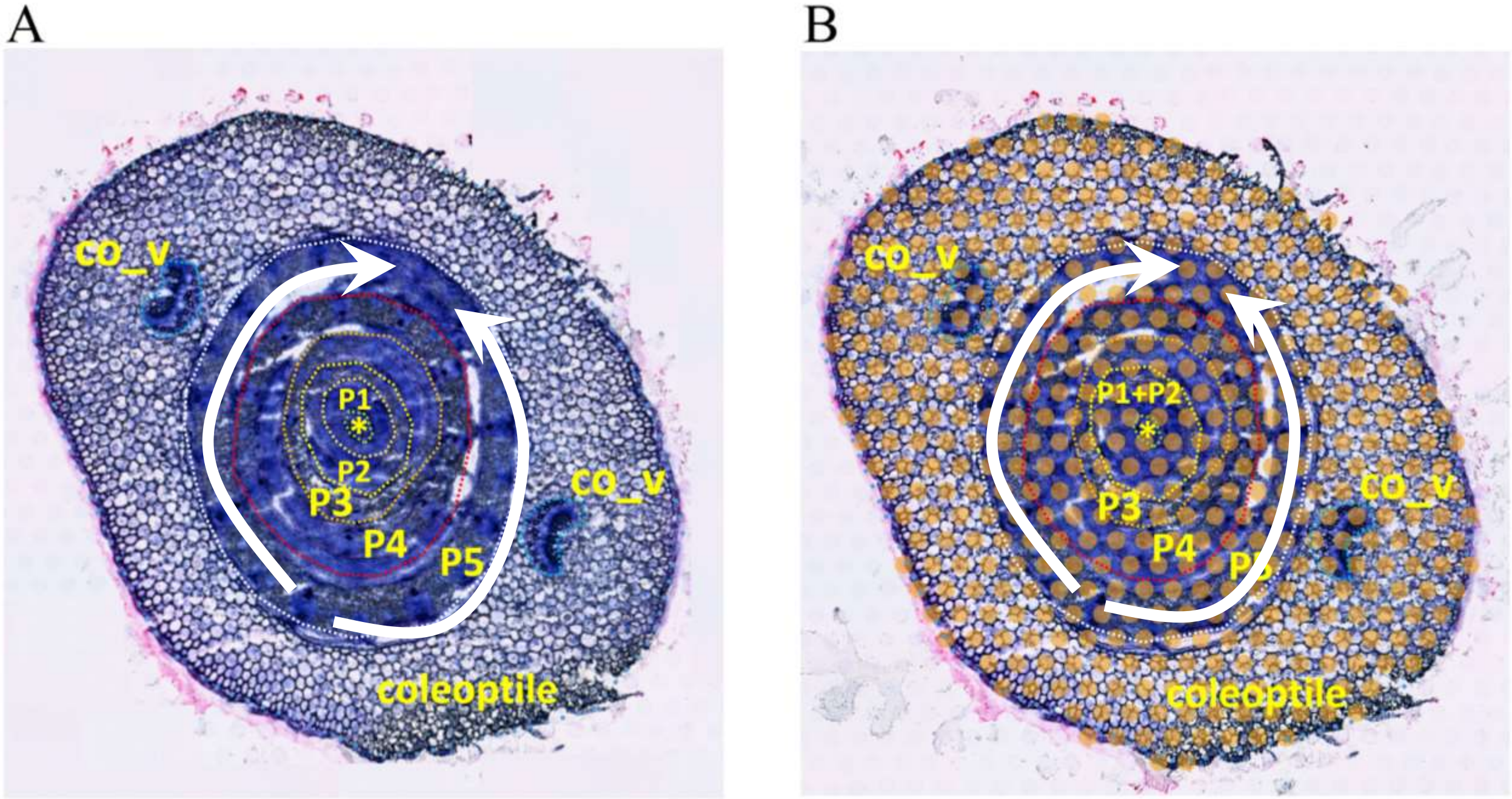
Images
tissue + coordinates

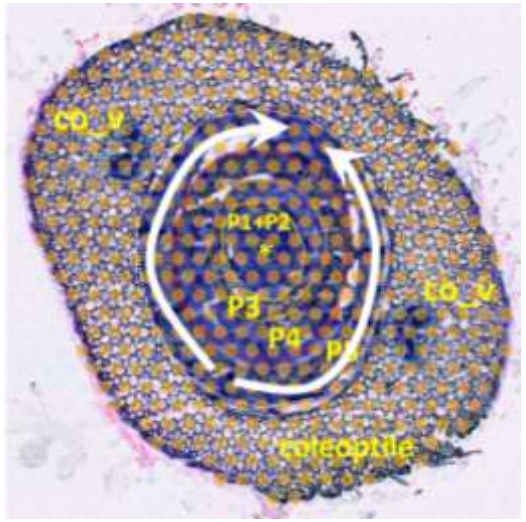
Reductions
PCA、UMAP、Harmony

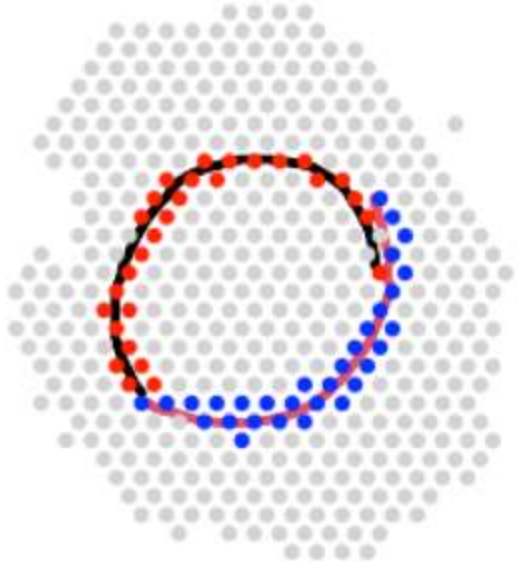
Graphs
neighbors / clusters

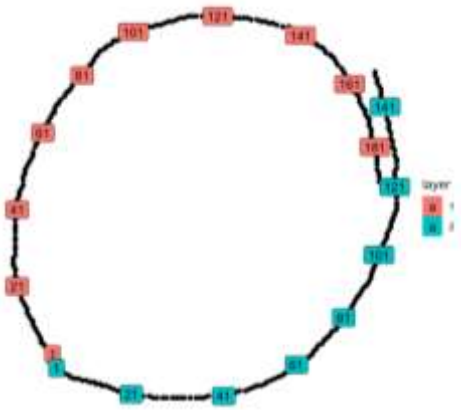
Features
gene identifiers 必須一致

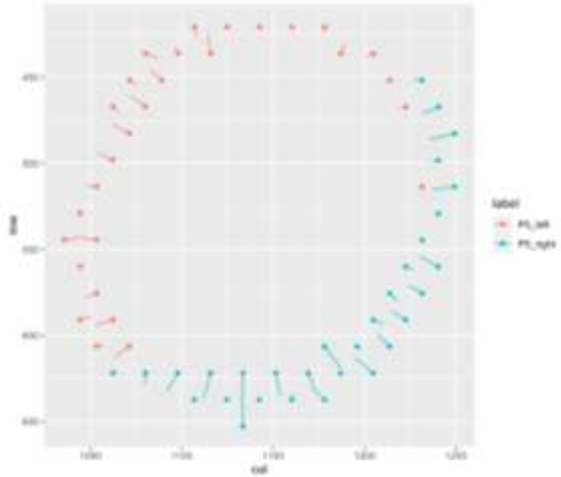
Provenance
參數、版本、輸入輸出

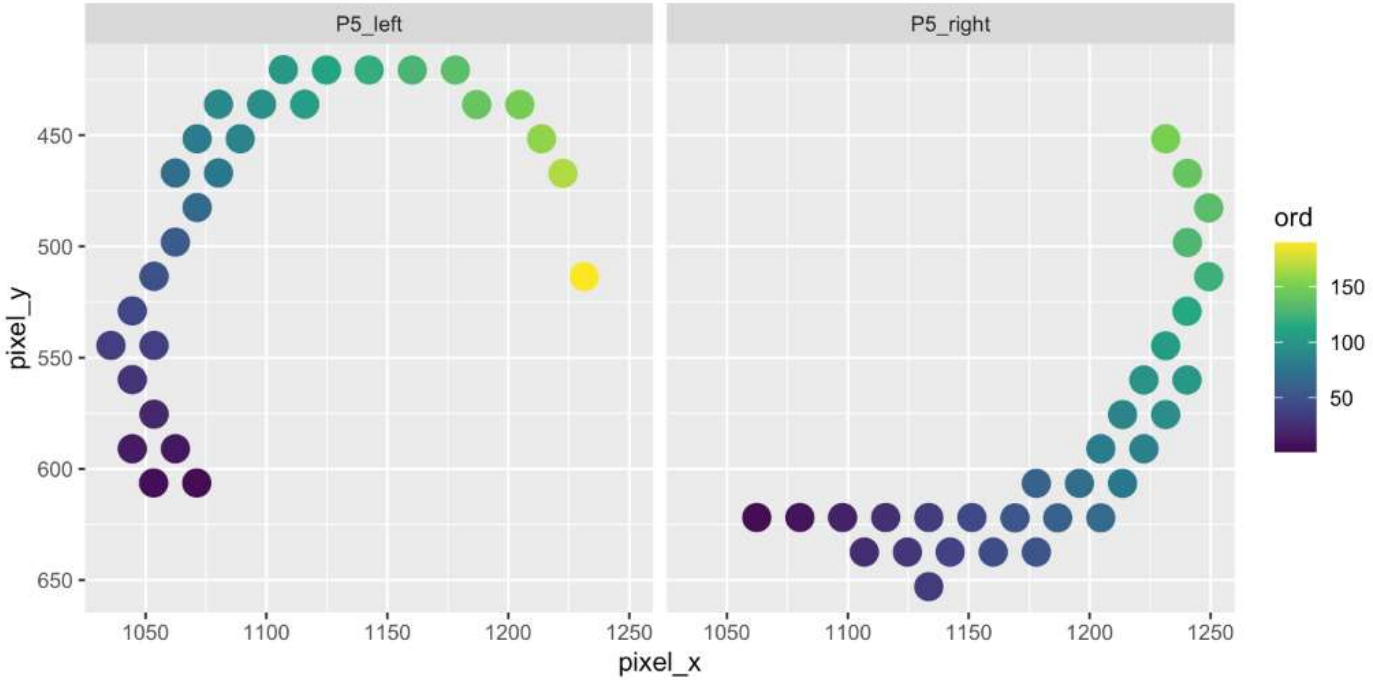

A
CO_V
P1
P2
P3
P4
P5
CO_V
coleoptile
B
CO_V
P1+P2
P3
P4
CO_V
coleoptile

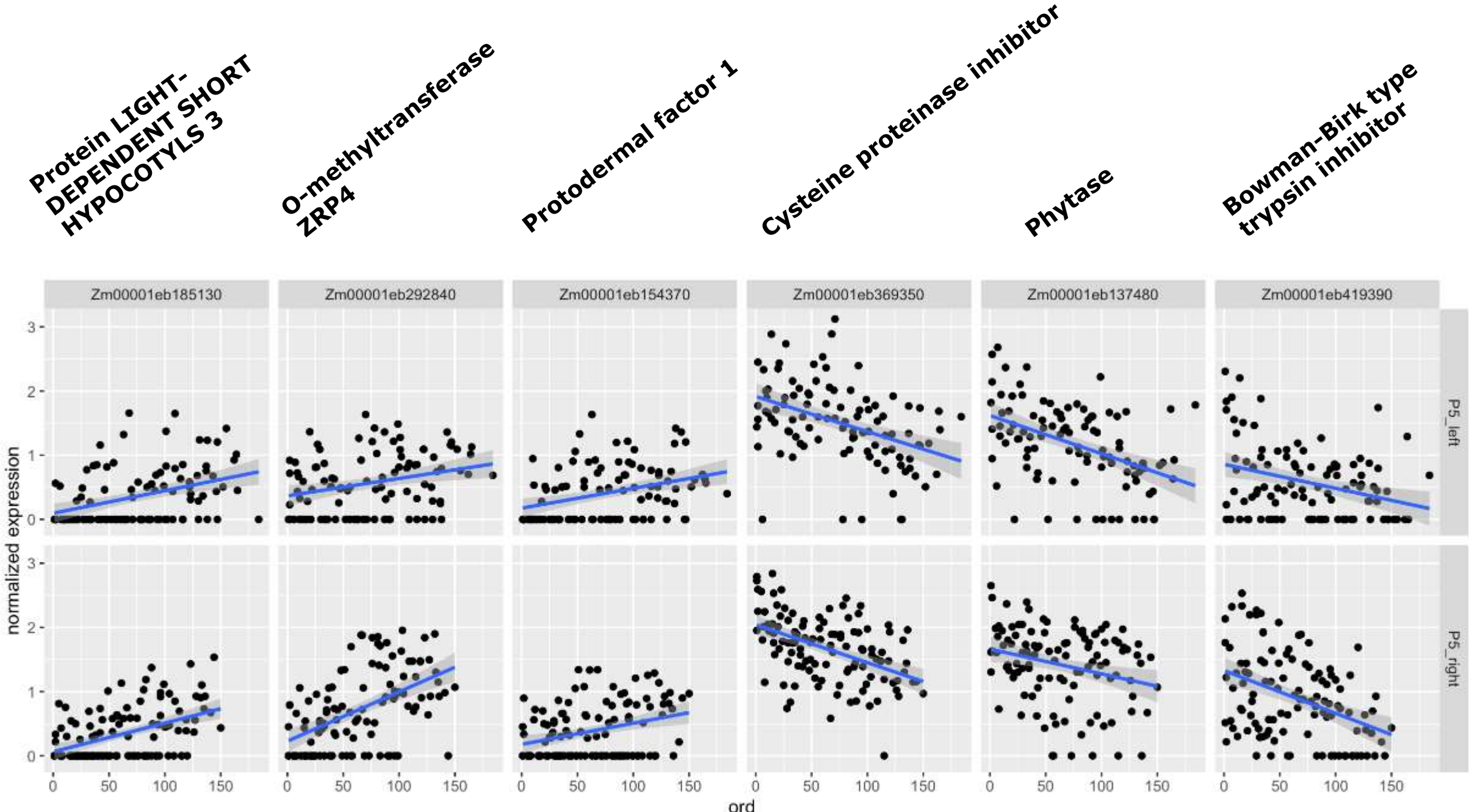

CO_V
P1+P2
P3
P4
P5
CO_V
coleoptile








P5_left
P5_right
pixel_y
pixel_x
450
500
550
600
650
1050
1100
1150
1200
1250
ord
150
100
50


Protein LIGHT-DEPENDENT SHORT HYPOCOTYLS 3
O-methyltransferase ZRP4
Protodermal factor 1
Cysteine proteinase inhibitor
Phytase
Bowman-Birk type trypsin inhibitor
Zm00001eb185130
Zm00001eb292840
Zm00001eb154370
Zm00001eb369350
Zm00001eb137480
Zm00001eb419390
P5_left
P5_right
normalized expression
ord
0
1
2
3
0
50
100
150

Q&A