

BIOF3 组学数据分析

09 空间转录组学

导出日期：2026年5月13日

09 空间转录组学

空间转录组在 scRNA-seq 的基础上保留了一个关键维度：每条表达信号来自切片的哪个位置。不同技术的空间分辨率差别很大：10x Visium 把组织打在微珠芯片上，每个 spot 约 55 μm 包含数个细胞；Slide-seq、MERFISH、seqFISH+ 能做到单细胞或亚细胞分辨率，代价是数据量更大、分析工具更专。

本节用 10x Visium 小鼠脑数据演示最常用的一条流程：读入 \rightarrow 质控 \rightarrow 空间感知标准化 (SCT) \rightarrow 聚类 \rightarrow 空间变异基因 \rightarrow 与参考 scRNA-seq 数据做 cell type 反卷积。

主流空间技术对比

技术	分辨率	数据量	常用工具
10x Visium	spot (~55 μm)	中等	Seurat、Scanpy + Squidpy
Slide-seq V2	~10 μm	中等	Seurat、RCTD
Stereo-seq	亚微米	大	StereoPy
MERFISH / seqFISH+	单细胞	大	Squidpy、Giotto
Xenium	亚细胞	大	10x pipeline、Seurat

Visium 仍然是门槛最低的选择：一张切片产生的 spot 数在几千级，普通笔记本能直接分析，教材、社区工具链也最完整。

用 Seurat 分析 Visium 小鼠脑

从 10x Genomics 下载 Visium 小鼠脑切片数据后（或用 SeuratData 里的 stxBrain），基本流程和 scRNA-seq 很像，只是多了 spatial 这一层坐标和组织图片：

```
library(Seurat)
# 若要一键获取示例数据:
# InstallData("stxBrain")
# brain <- LoadData("stxBrain", type = "anterior1")

# 或者从本地目录读取
brain <- Load10X_Spatial(
  data.dir = "~/biof3-data/visium-mouse-brain/outs",
  filename = "filtered_feature_bc_matrix.h5",
  assay = "Spatial",
  slice = "slice1"
)

# 质控: 在组织图上看每个 spot 的总表达量
VlnPlot(brain, features = "nCount_Spatial", pt.size = 0.1) + NoLegend()
SpatialFeaturePlot(brain, features = "nCount_Spatial")
```

两张 QC 图放在一起看：有没有“测序偏低的区域”集中在切片边缘（组织脱落），或者局部异常高（气泡、污染）。

标准化这一步建议用 `sctransform`，它对 spot 间差异较大的空间数据表现稳定：

```
brain <- SCTransform(brain, assay = "Spatial", verbose = FALSE)
brain <- RunPCA(brain, assay = "SCT")
brain <- FindNeighbors(brain, reduction = "pca", dims = 1:30)
brain <- FindClusters(brain, verbose = FALSE)
brain <- RunUMAP(brain, reduction = "pca", dims = 1:30)

# 聚类结果同时画在 UMAP 和组织切片上
DimPlot(brain, label = TRUE)
SpatialDimPlot(brain, label = TRUE, label.size = 3)
```

小鼠脑 Visium 在 `SpatialDimPlot` 里会呈现皮层、海马、胼胝体等解剖结构对应的聚类，不做任何 cell type 注释也能看出大致分区。

空间变异基因

常规的 `FindVariableFeatures` 只考虑表达变异，不考虑位置。空间变异基因是“表达分布不是随机的”，通常也是真正有解剖学意义的基因：

```
brain <- FindSpatiallyVariableFeatures(
  brain,
  assay = "SCT",
  features = VariableFeatures(brain)[1:1000],
  selection.method = "moransi"
)

top_features <- head(
  SpatiallyVariableFeatures(brain, selection.method = "moransi"),
  6
)

SpatialFeaturePlot(brain, features = top_features, ncol = 3, alpha = c(0.1, 1))
```

"moransi" (Moran's I) 衡量的是邻近 spot 的表达相似度，值越高说明越空间聚集。alpha 是作图参数，用来突出高表达区域、弱化背景。

真实示例：Visium 小鼠脑 Anterior 切片

配套脚本 [module10_spatial_sci.R](#) 用 `SeuratData::stxBRAIN` 的 `anterior1` 切片 (10x Visium 小鼠脑前部，约 2700 个 spot) 把上面的流程跑完。首次运行会通过 `InstallData("stxBRAIN")` 把数据装到本地 (~136 MB)，之后每次复用不再下载。

```
Rscript scripts/single-cell/sc10_spatial_sci.R
```

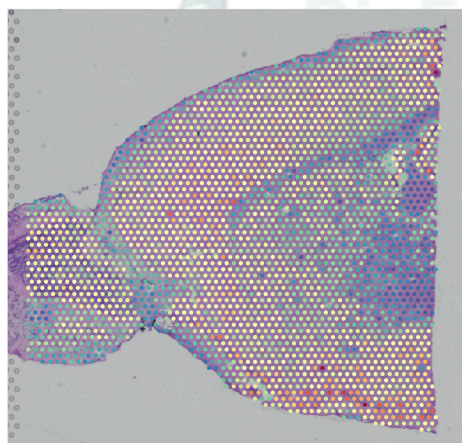
脚本顺序：LoadData 加载 Seurat 对象 → 在组织图上画 QC → `sctransform` 标准化 → PCA / Leiden 聚类 / UMAP → 在 UMAP 和切片上各看一遍 cluster → 画四个已知解剖学 marker 的空间表达 → 用 Moran's I 找空间变异基因 → 单基因在 UMAP 和切片上的同步视图。

每张图看什么

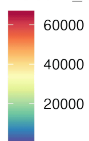
Spatial QC: sequencing depth across tissue

Check for edge dropout or local artifacts (bubbles, contamination)

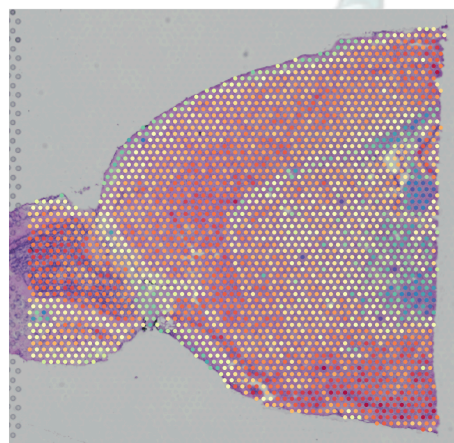
Total UMI per spot



nCount_Spatial



Genes detected per spot



nFeature_Spatial

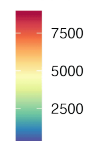


图 1: 左图是每个 spot 的总 UMI 数 (测序深度), 右图是检测到的基因数。切片边缘深度偏低是正常的 (组织少), 内部如果出现一块局部异常高的亮斑就要警惕气泡、折叠或污染。这一步直接看切片比看小提琴图更直观。

UMAP: SCT-normalized clustering

11 clusters

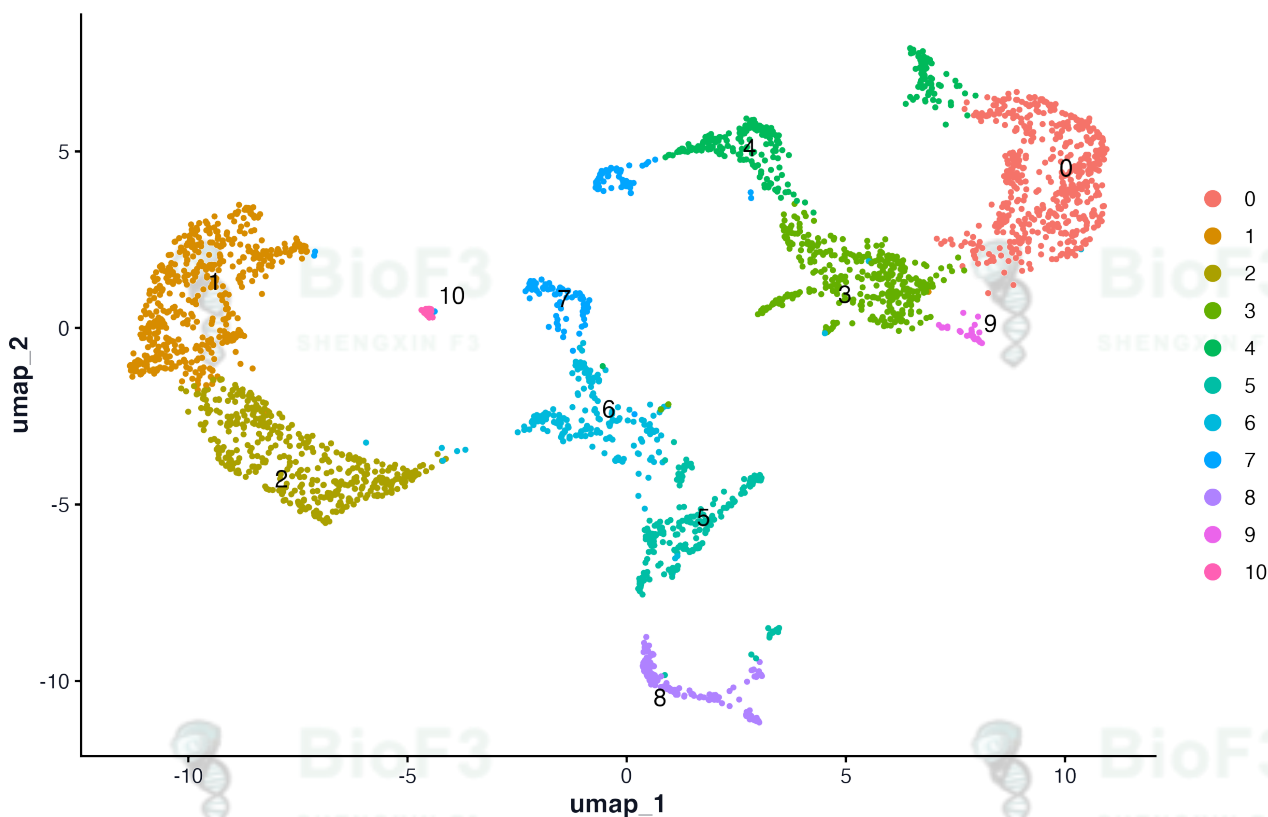


图 2: SCT 归一化后在 PCA 上跑 Leiden 得到的聚类, 投到 UMAP 上。和 scRNA-seq 不同, Visium 的 cluster 数量不会太多 (10 几个), 因为一个 spot 代表的是几个细胞的混合。

Cluster distribution on tissue section

Cortex, hippocampus, corpus callosum, and choroid plexus map to distinct clusters

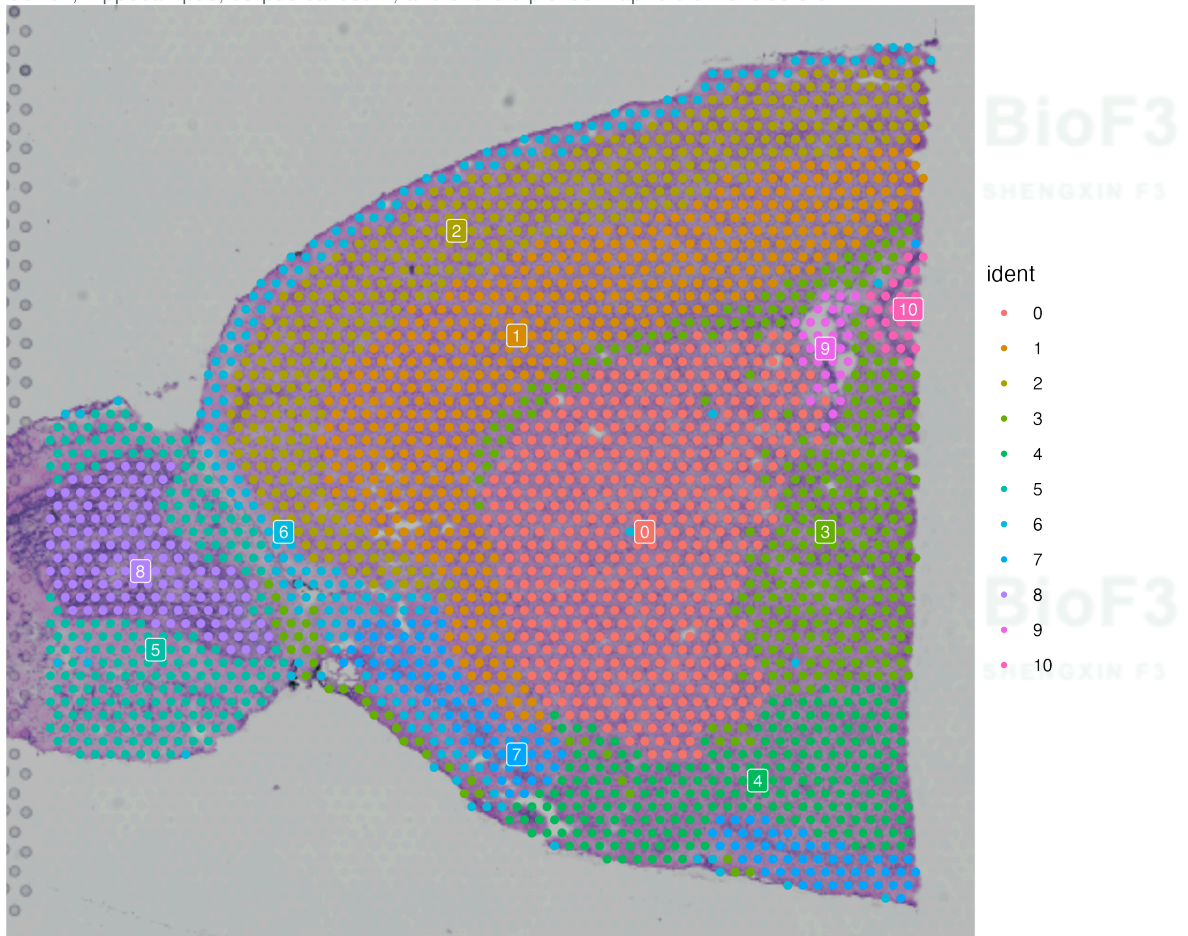


图 3: 同一个 `seurat_clusters` 直接画到组织切片上。不用做任何 cell type 注释, 已经能看到皮层 (同心环层)、海马 (马蹄形)、胼胝体 (中央带状)、脉络丛 (点状) 自然对应上不同的 cluster。这是 Visium 最“震撼”的一张图: 空间聚类 = 解剖学结构。

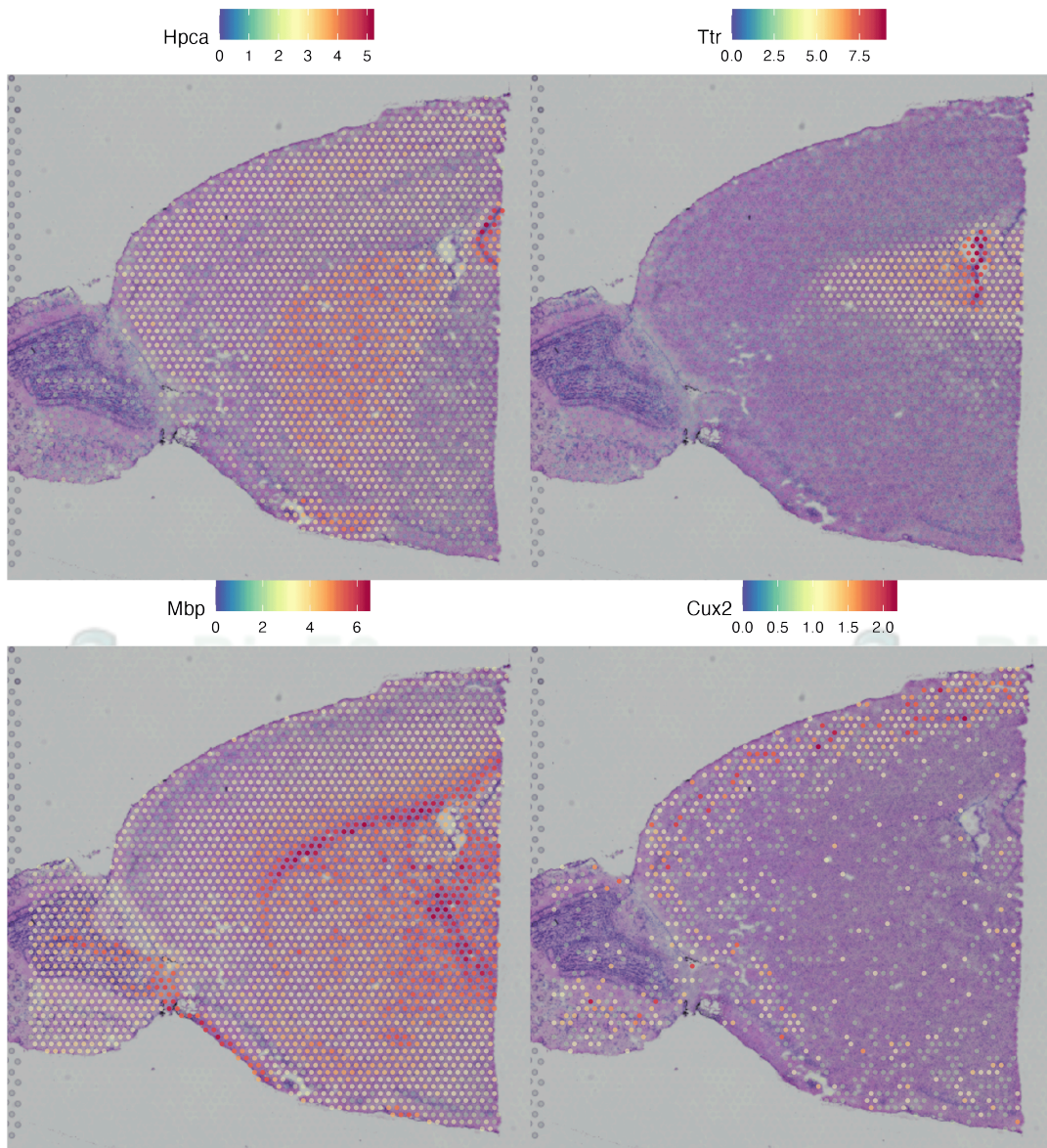


图 4：四个经典小鼠脑 marker 的空间表达 —— *HpcA*（海马）、*Ttr*（脉络丛）、*Mbp*（胼胝体 / 白质）、*Cux2*（皮层 II/III 层）。把这张图和图 3 一起看，能验证上一步的聚类是否确实落到了对应的解剖结构上。真实项目里换上自己关心的基因就能直接读出组织定位。

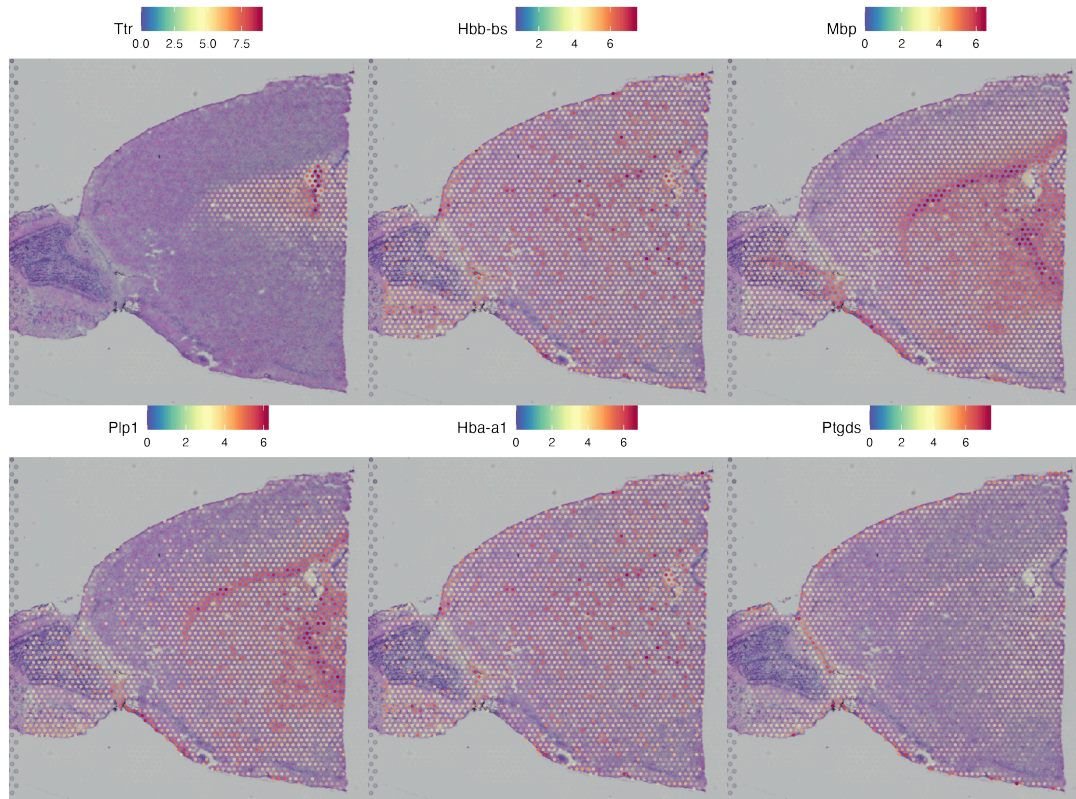
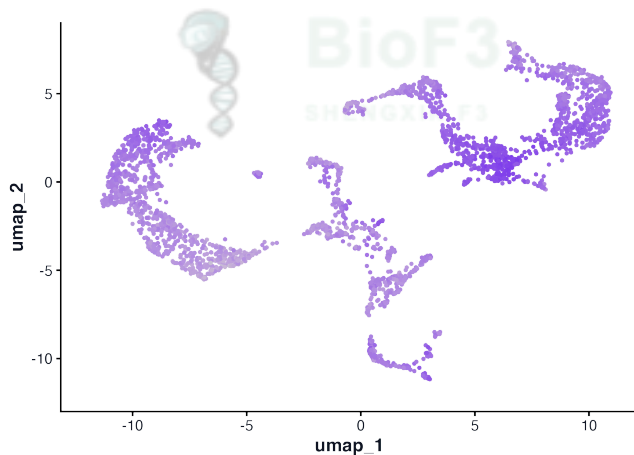


图 5：用 Moran's I 在前 1000 个高变基因里排名，取 Top 6 画空间分布。脚本实际跑出来是 `Ttr`, `Hbb-bs`, `Mbp`, `Plp1`, `Hba-a1`, `Ptgds` —— 脉络丛、红细胞、胼胝体、少突胶质、红细胞、脑膜/神经干细胞 —— 每一个都对应清晰的解剖结构，而不是随机散布，这正是空间变异基因要挑的信号。

Mbp: UMAP vs tissue section

Linking UMAP clusters to anatomical structures

Mbp (UMAP)



Mbp (tissue section)

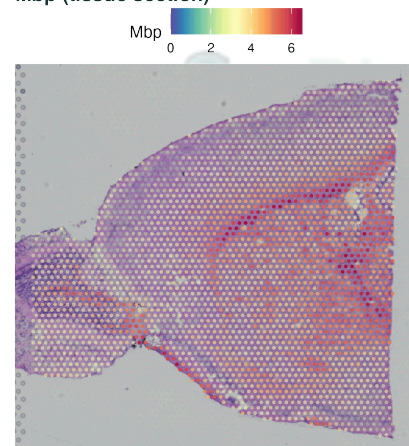


图 6：把 `Mbp` 在 UMAP 和组织切片上各画一张放一起。左边看它对应 UMAP 里的哪一个 cluster，右边看它对应切片上的什么结构。当一个未知基因的 UMAP 分布集中但你还不知道它的生物学意义时，拿它跟切片版一起看，经常能马上定位到“这是 XX 层 / XX 核团的 marker”。

套到自己数据上

脚本对小鼠脑做了演示。自己的切片（人脑、肿瘤、肝、皮肤等）换 `LoadData` 那行为

`Load10X_Spatial("~/path/to/outs")`，其余不动就行。`marker_genes` 按组织换：比如做肿瘤切片，可以换成 `EPCAM`（上皮）、`PTPRC`（免疫）、`COL1A1`（基质）；做肝切片可以换成 `ALB`（肝细胞）、`GLUL`（周中静脉）、`CYP2E1`（三区带）。分辨率更高的技术（Stereo-seq、Xenium）数据量更大，配合 Python 的 `Squidpy` 更流畅。

把 scRNA-seq 当作参考做反卷积

Visium 的每个 spot 通常包含多个细胞，直接做 cell type 注释会模糊。常用做法是把已经注释好的 scRNA-seq 数据当参考，把 spot 内的细胞组成估出来。常用工具有 RCTD、SPOTlight、CARD：

```
# SPOTlight 示例 (输入: scRNA 参考 + Visium 对象)
library(SPOTlight)

sc_ref <- readRDS("pbmc_reference.rds") # 假设已分析并注释
markers <- Seurat::FindAllMarkers(sc_ref)

spotlight_res <- SPOTlight(
  x      = sc_ref,
  y      = brain,
  groups = sc_ref$cell_type,
  mgs    = markers
)

plotSpatialScatterpie(
  x      = brain,
  y      = spotlight_res,
  cell_types = unique(sc_ref$cell_type),
  img    = FALSE
)
```

得到的 scatterpie 在每个 spot 位置画一个小饼图，显示各 cell type 的估计比例。常见验证方式：查看某种 cell type 的比例在组织切片上是否和已知解剖学位置一致。

用 Scanpy + Squidpy 做 Python 版

Python 生态下 scanpy 配合 squidpy 可以覆盖类似流程，并提供更丰富的空间统计 (nhood enrichment、co-occurrence 等)：

```
import scanpy as sc
import squidpy as sq

adata = sc.read_visium("~/biof3-data/visium-mouse-brain/outs")
adata.var_names_make_unique()

sc.pp.calculate_gc_metrics(adata, inplace=True)
sc.pl.spatial(adata, color="total_counts")

sc.pp.normalize_total(adata, inplace=True)
sc.pp.log1p(adata)
sc.pp.highly_variable_genes(adata, flavor="seurat", n_top_genes=2000)

sc.pp.pca(adata)
sc.pp.neighbors(adata)
sc.tl.umap(adata)
sc.tl.leiden(adata)
sc.pl.spatial(adata, color="leiden")

# 空间邻域富集: 看哪些 cluster 在空间上彼此相邻
sq.gr.spatial_neighbors(adata)
sq.gr.nhood_enrichment(adata, cluster_key="leiden")
sq.pl.nhood_enrichment(adata, cluster_key="leiden")
```

Scanpy/Squidpy 侧最适合的场景是高分辨率空间数据 (Stereo-seq、MERFISH)，它们对大矩阵更友好。Visium 级别数据两套都能跑，选择看个人生态。

下载资源

`module10_spatial_sci.R`
12 KB

[下载 Visium 小鼠脑空间分析完整脚本 ↗](#)

参考资源

- [10x Visium 官方文档](#)
- [Seurat Spatial Vignette](#)
- [Squidpy 文档](#)
- [SPOTlight](#)
- [Giotto Suite](#)



扫码关注微信公众号【生信F3】

获取文章完整内容，分享生物信息学最新知识。