

SDAS应用案例

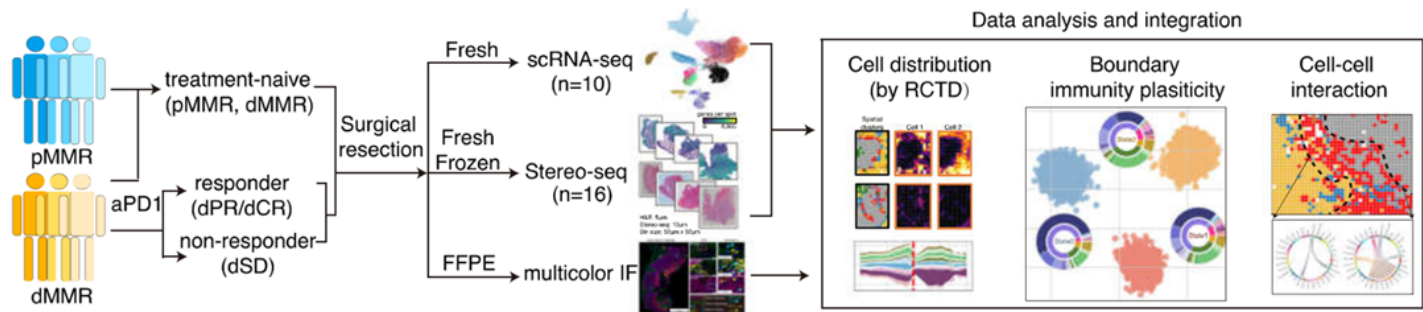
1. 文献说明

文章标题: Spatially organized tumor-stroma boundary determines the efficacy of immunotherapy in colorectal cancer patients

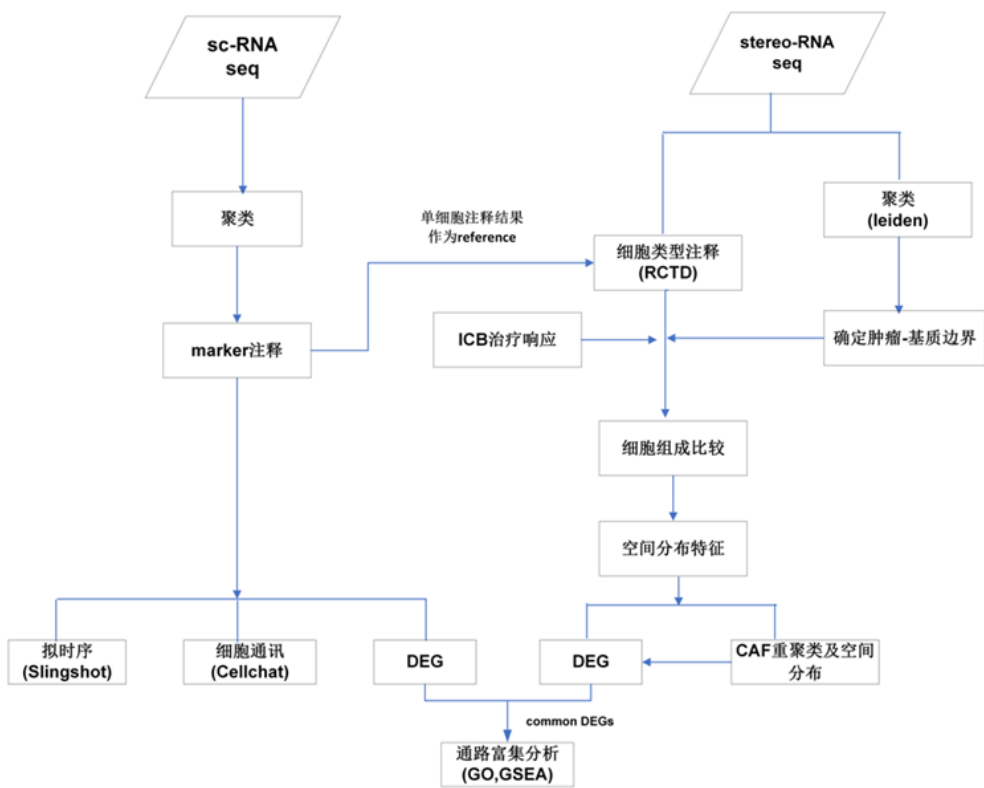
期刊: Nature Communications **影响因子:** 14.7

摘要: 本研究应用Stereo-seq和scRNA-seq，绘制CRC患者的高分辨率TME空间图谱，重点分析肿瘤-基质边界的细胞与分子特征及其在ICB响应中的作用机制。

样本分组: 1) 未做治疗：pMMR vs. dMMR 2) ICB治疗(dMMR)：dPR/dCR vs. dSD



下图为该文献的生信分析流程图，其中scRNA-seq聚类 and 注释直接使用文章提供的数据，其他分析均进行了复现。



2. 分析准备



SDAS软件包: SDAS_beta.tar.gz

文章数据: ./Application_cases/Application_test_data

single_cell.h5ad: 单细胞数据, 包含注释信息

stereo_bin100.h5ad: 16个样本stereo-seq整合的数据, bin100, 数据格式和SAW输出的h5ad格式相同

运行脚本和Jupyter代码: ./Application_cases/Application_scripts

3. 分析步骤

脚本参数说明

- 1 `${tool_dir}`为软件解压后的文件夹路径
- 2 `${output_dir}`为输出结果的文件夹路径
- 3 `${single_cell_h5ad}`为文章的单细胞数据single_cell.h5ad的绝对路径
- 4 `${stereo_h5ad}`为文章的stereo数据stereo_bin100.h5ad的绝对路径

3.1 将stereo-seq数据进行格式处理得到SDAS的标准输入（必做）

- 1 `#输出文件名为stereo_bin100_standard.h5ad`
- 2 `${tool_dir}/SDAS Data_Process input2h5ad -i ${stereo_h5ad} --mode single -o ${output_dir}`

3.2 CCA去批次, 聚类, 根据marker定义tumor-stromal boundary bins

- 1 `#在seruat环境中, 读入stereo_bin100_standard.h5ad文件转成seruat对象, 进行CCA去批次, 存储integrated后的表达矩阵和meta.data信息 (运行时间: 12h, 最大内存: 324G)`
- 2 `Rscript ./Application_cases/Application_scripts/Figure1_CCA_in_seurat.r`
- 3
- 4 `#在scanpy环境中, 读入CCA后的表达矩阵和meta.data信息创建h5ad对象 (运行时间: 1h17m, 最大内存: 169G)`
- 5 `python`
- 6 `./Application_cases/Application_scripts/Figure1_generate_h5ad_after_CCA.py`

```
7 #在scanpy环境中, 进行leiden聚类, 可自行调整resolution参数; 根据文献中提供的marker
  list, 绘制表达heatmap图, 并对cluster进行细胞类型定义
8 ./Application_cases/Application_scripts/Figure1_clustering_in_scanpy.ipynb
```

3.3 RCTD注释

```
1 #reference为文章的单细胞数据single_cell.h5ad, 结果文件为
  stereo_bin100_standard_anno_rctd.h5ad
2 ${tool_dir}/SDAS cellAnnotation rctd -i stereo_bin100_standard.h5ad -o
  $output_dir --reference ${single_cell_h5ad} --label_key annotation2 --
  bin_size 100 --filter_rare_cell 0 --input_gene_symbol_key _index --slice_key
  id
```

3.4 stereo-seq数据提取boundary的bins, 进行DEG和prerank分析

```
1 ##提取聚类为boundary的bin
2 ${tool_dir}/SDAS dataProcess subset -i stereo_bin100_standard.h5ad --
  label_key level1 --list_include boundary -o $output_dir
3
4 ##对dMMR和pMMR两个分组进行DEG分析
5 ${tool_dir}/SDAS DEG -i ${subset_h5ad} --diff_plan boundary_deg_plan.csv -o
  $output_dir
6
7 ##prerank分析 (运行时间43min, 最大运行内存25.4G) , 绘制文章指定通路的prerank图
8 ${tool_dir}/SDAS geneSetEnrichment prerank -i
  ${output_dir}/dMMR.vs.pMMR.deg.csv --species human --gmt
  $gmt_db_dir/GO_Biological_Process_2023.gmt --pathways boundary_pathway.txt -o
  $output_dir
```

3.5 单细胞提取样本类型为dmmr的数据进行细胞通讯分析

```
1 #提取样本类型为dmmr的单细胞数据, 输出文件为single_cell_subset.h5ad
2 ${tool_dir}/SDAS dataProcess subset -i ${single_cell_h5ad} -o $output_dir --
  label_key type --list_include dmmr
3
4 #将h5ad转成rds
5 ${tool_dir}/SDAS dataProcess h5ad2rds -i single_cell_subset.h5ad -o
  ${output_dir}
6
7 #进行细胞通讯分析 (运行时间: 16min, 运行最大内存: 8.8G)
```

```

8  ${tool_dir}/SDAS CCI cellchat --input ${rds_file} --label_key annotation2 --
   species h   uman --output $output_dir
9
10 #Fig3b画图, 绘制CD274-PDCD1配受体图
11 ${tool_dir}/cci/anaconda/lib/R/bin/R --no-echo --no-restore --
   file=cellchat_pic.R --args ${cellchatRDS}
   ${output_dir}/Fig3b_cell_cell_interaction.pdf

```

3.6 单细胞提取CAF细胞进行拟时序分析

```

1  #提取单细胞数据中CAF的细胞类型
2  ${tool_dir}/SDAS dataProcess subset -i ${single_cell_h5ad} -o $output_dir --
   label_key annotation2 --list_include
   CAF_ADAMDEC1,CAF_CXCL8,CAF_CXCL14,CAF_KCNN3,CAF_PI16
3
4  #过滤基因数<350的细胞, 存成adata, 使用BBKNN去批次, scanpy聚类后生成UMAP信息, 将UMAP信
   息导出为umap_coordinates.csv
5  ./Application_cases/Application_scripts/CAF_umap.ipynb
6
7  #将过滤后的h5ad转成rds
8  ${tool_dir}/SDAS dataProcess h5ad2rds -i single_cell_subset_filter.h5ad -o
   ${output_dir}
9
10 #进行拟时序分析, 其中UMAP信息使用scanpy聚类后生成UMAP信息 (运行时间3min, 最大内存
   1.4G)
11 ${tool_dir}/SDAS trajectory monocle3 --input single_cell_subset_filter.rds --
   root_key annotation2 --root CAF_PI16 --resolution 0.004 --umap
   umap_coordinates.csv --output $output_dir

```

3.7 单细胞提取CAF细胞, 进行DEG和prerank分析

```

1  #提取单细胞数据中CAF的细胞类型
2  ${tool_dir}/SDAS dataProcess subset -i ${single_cell_h5ad} -o $output_dir --
   label_key annotation2 --list_include
   CAF_ADAMDEC1,CAF_CXCL8,CAF_CXCL14,CAF_KCNN3,CAF_PI16
3
4  #对CAF_CXCL14和CAF_CXCL8CAF两组细胞类型进行DEG分析
5  ${tool_dir}/SDAS DEG -i ${h5ad_file} --diff_plan CAF_deg_plan.csv -o
   $output_dir
6
7  #对DEG进行GO富集

```

```
8  ${tool_dir}/SDAS geneSetEnrichment enrichr -i
   ${output_dir}/CAF_CXCL14.vs.CAF_CXCL8.deg_filtered.csv --species human --gmt
   $gmt_db_dir/GO_Biological_Process_2023.gmt --cut_off 0.05 -o $output_dir
9
10 #prerank分析(运行时间16min, 最大内存 4.7G), 绘制文章指定通路的prerank图
11 ${tool_dir}/SDAS geneSetEnrichment prerank -i
   ${output_dir}/CAF_CXCL14.vs.CAF_CXCL8.deg.csv --species human --gmt
   $gmt_db_dir/GO_Biological_Process_2023.gmt --pathways CAF_pathway.txt -o
   $output_dir
```

3.8 使用SDAS的分析结果进行figure绘制

- 1 文章中的图使用notebook进行绘制，数据均来源于上述SDAS的分析结果，环境可用
\${tool_dir}/anno/anaconda/bin/python
- 2 ./Application_cases/Application_scripts/stereo_plot.ipynb

4. 结果解读

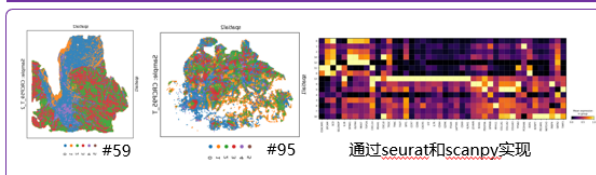
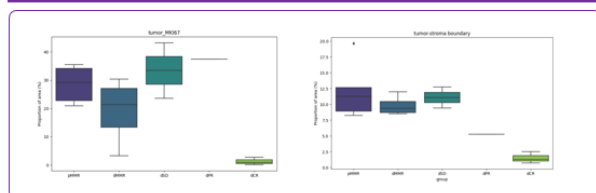
4.1 CRC空间转录组图谱构建

肿瘤-基质边界的显著特征

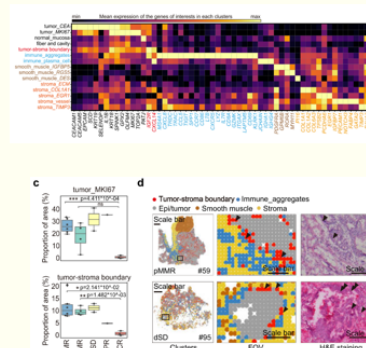
- dMMR患者边界区域缺乏明确组织，与pMMR患者相比，dMMR患者更倾向于无序的免疫细胞聚集。这种结构特点可能与更高的免疫浸润能力相关。
- dCR患者表现出肿瘤区域中增殖相关基因（如MKI67）的表达显著减少，表明ICB疗效可能抑制了肿瘤细胞增殖。
- 说明：据文章Methods中描述，多片聚类使用Seurat CCA方法进行多片去批次，然后使用Scanpy进行PCA降维, neighbors网络构建，leiden算法进行聚类。由于SDAS目前不支持该方法，我们提供了分析代码（见3.2分析步骤）。由于算法中的随机部分和参数细节未知，聚类结果与原文存在一定的差异，后续的分析 and 解读均使用文章提供的h5ad文件的obs['level2']进行。

结果1: CRC空间转录图谱构建, leiden聚类, 确定肿瘤-基质边界

dCR患者表现出肿瘤区域增殖相关基因(如MKI67)的表达显著减少, 表明ICB疗效可能抑制肿瘤细胞增殖
dCR和dPR的肿瘤边界比例降低, 可能与更高的免疫浸润能力相关

通过leiden聚类, 鉴定肿瘤-基质边界**dCR患者的肿瘤区域增殖相关基因 (MKI67) 及边界比例较低**

文章原图



4.2 肿瘤-基质边界的免疫细胞的功能特性

1) 免疫细胞在边界的动态分布

- 沿边界的免疫细胞分布曲线显示, pMMR患者的免疫细胞分布呈间断性聚集, 而dMMR患者具有连续且丰富的免疫细胞分布, 尤其在边界区域 ($0 \pm 150\mu\text{m}$) 内。

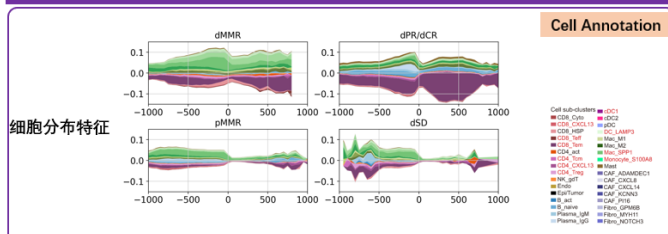
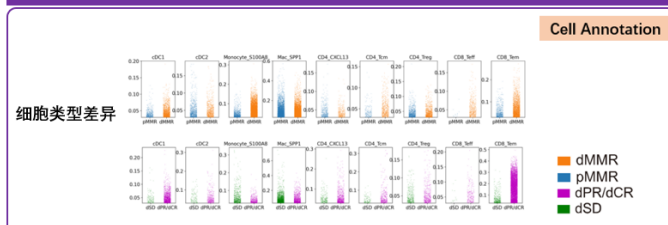
2) 免疫细胞富集与ICB疗效相关

- dMMR患者中CD8+效应T细胞 (CD8_Teff)、记忆T细胞 (CD8_Tem) 以及DC_LAMP3显著富集于边界区域, 这种局部免疫细胞聚集可能增强抗肿瘤免疫反应。
- dSD患者中SPP1+巨噬细胞的比例更高, 与较低的ICB疗效相关, 提示这种髓系细胞可能通过免疫抑制机制阻碍治疗反应。

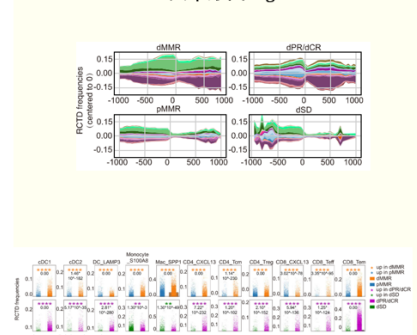
结果2 肿瘤-基质边界的免疫细胞特征

不同患者对ICB治疗的预后效果差异可能是由于肿瘤-基质边界的空间特征导致, 局部免疫细胞聚集可能增强抗肿瘤免疫反应

- dMMR和dPR/dCR患者具有连续且丰富的免疫细胞分布
- pMMR和dSD患者的免疫细胞分布呈间断性聚集

**DC和T细胞在dMMR和dPR/dCR患者中比例更高**

文章原图 Fig. 2



4.3 LAMP3+树突状细胞（DC）与T细胞的相互作用

1) 对dMMR和pMMR的边界区域进行DEG及富集

dMMR在正向趋化调节通路富集，该通路会触发信号转导，说明存在细胞串扰。

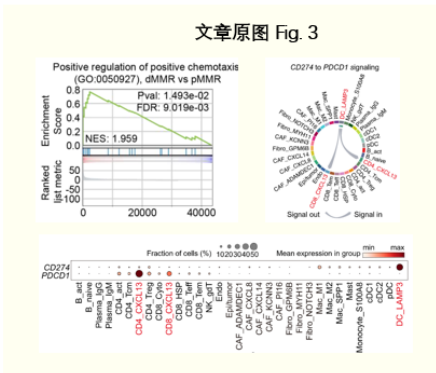
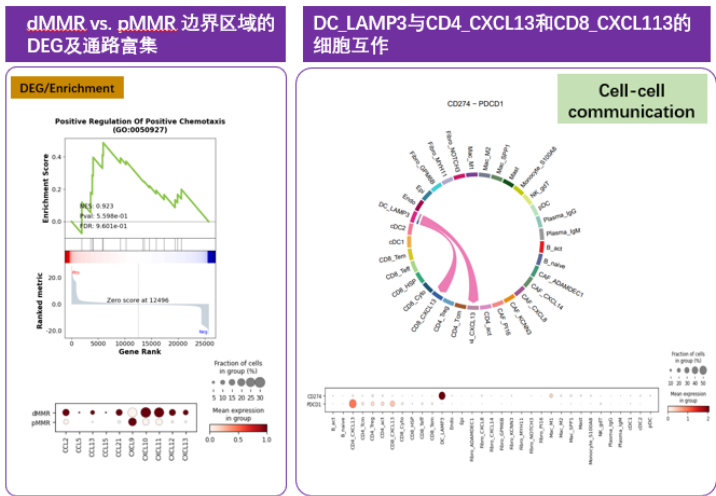
2) DC与T细胞的空间邻近性

- DC_LAMP3和CXCL13+T细胞在边界区域表现出高度邻近性（<200μm），且相互作用通过PD1-PD-L1轴介导。
- 边界区域中DC_LAMP3的高密度分布有助于增强抗肿瘤T细胞活性。

■ 结果3 LAMP3+ DC与T细胞的相互作用

dMMR在正向趋化调节通路富集，该通路会触发细胞信号转导，说明存在细胞串扰

DC_LAMP3与CD4_CXCL13、CD8_CXCL113通过LR-pair（CD274和PDCD1）进行细胞互作，有助于增强抗肿瘤T细胞活性



4.4 CAF塑性与免疫排斥的关联

1) 肿瘤-基质边界免疫状态的多样性

- 定义了三种免疫状态（状态0、1、2），状态0在pMMR和dSD患者中占主导地位，表现为免疫排斥，而状态1和2主要与免疫活化相关，见于dMMR和dPR/dCR患者
- 免疫排斥状态与肿瘤-基质边界区域内的CAF（尤其是CXCL14+CAF）比例升高密切相关。

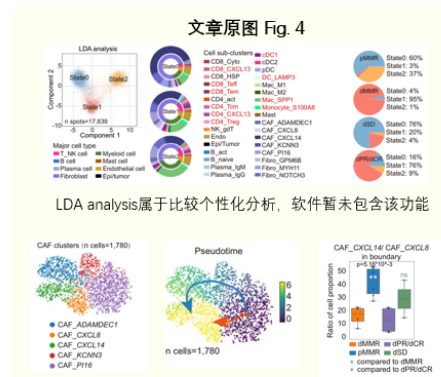
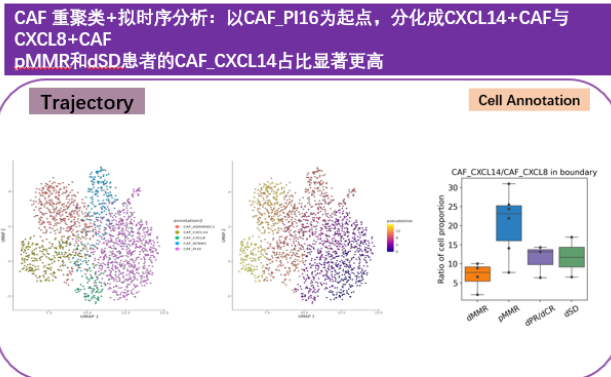
2) CAF亚群的功能特性

- CXCL14+CAF与CXCL8+CAF的分化路径分析表明，CXCL14+CAF通过增强细胞外基质（ECM）组织化，抑制T细胞浸润，从而形成免疫排斥的边界结构。

■ 结果4 CAF可塑性与免疫排斥的关联

肿瘤基质边界的免疫状态与ICB治疗响应相关：state 0在pMMR和dSD患者中占主导地位，表现为免疫排斥，而state1和2主要与免疫活化相关，见于dMMR和dPR/dCR患者

免疫排斥状态与肿瘤-基质边界区域内的CAF（尤其是CXCL14+CAF）比例升高密切相关



4.5 CAF在基质屏障形成中的作用

1) CXCL14+CAF的屏障作用

- CXCL14+CAF在pMMR和dSD患者的肿瘤边界显著富集，伴随更高的ECM组织化特征，这种屏障结构显著限制了免疫细胞的浸润。

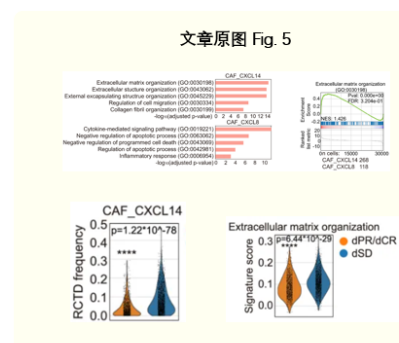
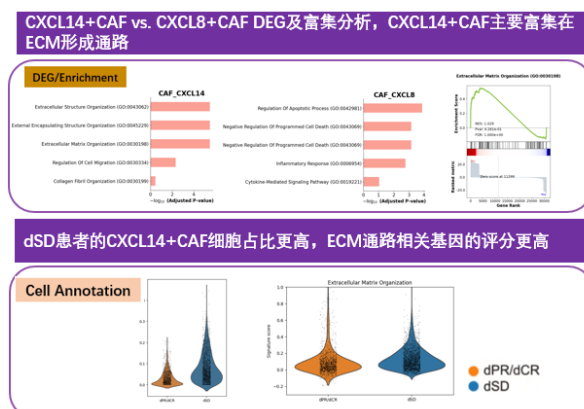
2) ECM组织化的免疫抑制效应

- CXCL14+CAF调节ECM结构，形成有序的基质屏障，排斥T细胞进入肿瘤核心区域，降低ICB疗效。

结果5 CAF在基质屏障形成中的作用，其中CAF_CXCL14与ECM形成相关

CXCL14+CAF通过调节ECM结构，形成有序的基质屏障，排斥T细胞进入肿瘤核心区域

ICB治疗响应低的患者（dSD）中有较高比例的CXCL14+CAF细胞类型，导致对免疫细胞的排斥，影响免疫治疗效果



5. reference

Feng, Y., Ma, W., Zang, Y. *et al.* Spatially organized tumor-stroma boundary determines the efficacy of immunotherapy in colorectal cancer patients. *Nat Commun* **15**, 10259 (2024).

<https://doi.org/10.1038/s41467-024-54710-3>