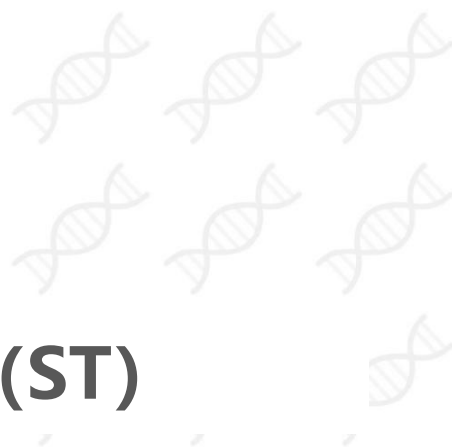


10X 空间转录组 常见问题 (FAQ) Spatial Transcriptomics (ST)



10X 空间转录组常见问题（FAQ）

10X 空间转录组常见问题（FAQ）	2
一、 前期常见问题	4
（一） 技术选择	4
1. 10x Genomics 公司的两个空转平台（p100 捕获和探针捕获）的区别是什么？	4
2. 空间转录组测序需要生物学重复吗？	4
3. 临床样本无法一次把所有样本取到，可以分批取样吗？	4
4. 研究的物种没有参考基因组，可以做吗？	4
5. FFPE 样本时间太长，是否适合做空转？	5
6. 组织太小可以拼片吗？	5
7. 想要和其他组学联合，可以考虑哪些组学？	5
（二） 样本准备	5
1. 新鲜样本的包埋是否需要客户自己完成？	5
2. 切片是客户自己做还是公司完成？	5
3. FF 和 FFPE 的切片厚度是否有差异？	6
4. FF 和 FFPE 分别需要多少张切片？	6
5. 样本如何收集？	6
6. 如果空间转录组和空间代谢组一起做，该如何取样？	6
二、 关于 spaceranger 的问题	6
1. 空间样本需要覆盖多少 spots 才能分析？	6
2. 空间转录组测序的 reads 能否进行拼接？	6
3. Sequencing Saturation 需要达到多少才能继续分析？	7
4. 是否可以和单细胞 cellranger 一样进行 force cell？	7
5. 如果组织中存在比较大的非细胞结构（类似囊腔等），spaceranger 将这部分对应的 spots 识别出来继续分析，是否可行？	7
三、 关于报告和后续分析问题	7
1. 空间转录组测序结果，除了 spaceranger 软件处理之外，是否还进行其他质控？	7
2. 空转可以调整 cluster 的数目吗？	7
3. 如何判断是否存在批次效应？	7
4. 如何判断过度矫正？	8

5. 常用的批次矫正方法	8
6. 空间转录组可以进行细胞类型鉴定吗？常用的方法是什么？	8
7. 为什么有的报告有 celltype 结果，有的没有？	8
8. marker 基因的鉴定原则是什么，top10 是如何筛选的？	8
9. 空间 spots 的成分比较复杂，那么分 cluster 的意义在哪里？	8
10. 差异基因的筛选标准是什么？	9
11. 空间切片上存在不同的组织结构，如癌和癌旁、不同脑区，是否需要单独分析或讨论？	9
12. 空间转录组可以进行细胞亚型分析吗？	9
13. 如果利用单细胞鉴定到了细胞亚型，想在空间上进行分析，是否可行？	9
14. 空间测序数据可以与 bulk RNA-seq 联合分析吗，有何好处？	9
15. 空转是否可以检测融合基因？	9
16. 关注的基因在测序结果中检测不到为什么？	9
17. 10X Genomics 单细胞转录组测序是否可以分析可变剪切？	9
18. 10X Genomics 单细胞转录组测序是否可以分析 SNP 和 CNV？	9
四、关于高级分析的内容	10
(一) 拟时序的相关问题	10
1. 拟时序分析的方法有几种？	10
2. 拟时序分析什么情况下适用？	10
3. 拟时序分析需要人为判定方向吗？如何判断？	10
(二) Scenic 的相关问题	10
1. 空间转录组是否可以做 scenic 分析？	10
2. 是否可以按照分组/cluster 进行 regulon 比较？	10
3. 什么情况下按照分组分析 regulon？	10
4. 如果 RAS 和 RSS 的数据不一致，怎么办？	11
5. 活性高是否代表这个基因表达量也高？	11
6. 挑选的某个 cluster 里面活性高的 regulon，是否下游靶基因也是上调的？	11
7. 确定了 regulon 之后还能做些什么？	11
(三) 细胞间通讯的相关问题	11
1. 空间通讯分析一定要有细胞类型的信息才能做吗？	11
答：不一定。有两种情况。	11
2. cellphoneDB 和 cellchat 该如何选择？	11

3. 如何筛选关键的受配体基因?	11
4. cellphoneDB 没有收录通路信息, 是否意味着不能关注通路?	12
(四) CNV 的相关问题	12
1. 是否存在专门的数据库进行 CNV 预测?	12
2. 空间转录组一个 spots 可能包含多个细胞类型, 肿瘤细胞和正常细胞共定位的话, 正常细胞如何指定?	12
3. CNV group 应该怎么用?	12

一、前期常见问题

(一) 技术选择

1. 10x Genomics 公司的两个空转平台 (plobA 捕获和探针捕获) 的区别是什么?

答: 10x Visum (PlobA 捕获) 和 Cytassist (探针捕获) 平台的主要区别如下:

	10x Visum平台	Cytassist平台
涉及物种类型	真核生物均可	人、小鼠
样本处理方式	新鲜样本OCT包埋(FF)	新鲜样本OCT包埋 (FF)、石蜡包埋 (FFPE)
RNA捕获方式	plobA捕获	探针捕获
质检方式	RIN>7	DV200>30%
是否需要透化优化	需要	不需要
非编码RNA捕获	可以捕获部分LncRNA	仅捕获mRNA
未知基因检出	可以检出未知基因	仅针对已知基因
基因数检出 (中位值)	视样本具体类型, 一般4000以下	视样本具体类型, 一般是visum的1.5-2倍
是否可以和单细胞联合分析	可以	可以

2. 空间转录组测序需要生物学重复吗?

答: 建议要做生物学重复。对于小鼠的样本, 背景比较单一, 可以考虑每组 3 个样本; 对于临床样本, 个体差异比较大, 建议每组 3-5 个甚至更多。

3. 临床样本无法一次把所有样本取到, 可以分批取样吗?

答: 可以分批提供。空间转录组后续分析过程中可以通过软件进行批次矫正。

4. 研究的物种没有参考基因组, 可以做吗?

答：不建议。没有参考基因的样本，后续空转测完之后无法对基因进行比对和注释，因此后续分析可能无法达到老师的目的。

5. FFPE 样本时间太长，是否适合做空转？

答：按照官方的建议，最好 5 年以内的 FFPE 蜡块，不过正式实验之前要检测 DV200 是否大于 30%，如果质检合格，也是可以尝试做空转的。

6. 组织太小可以拼片吗？

答：理论上是可以的，不过要视切片大小具体定夺，尽量避免两张切片部分重叠导致数据无法拆分的情况。此外，FF 样本建议不同样本直接包埋在一起，避免后期分别贴片造成更多的风险。

7. 想要和其他组学联合，可以考虑哪些组学？

答：组学联合可以考虑一下

- 1) DNA 层面：可以联合 scATAC-seq, WGS, WES (找到关键基因或转录因子进行空间定位分析)
- 2) 转录层面：可以联合 bulk RNA-seq, scRNA-seq 针对关键基因/转录因子/细胞类型/细胞亚型进行空间定位和空间微环境分析
- 3) 蛋白层面：DIA、label-free, TMT 检测 (至少每组 3 重复，临床至少 5 对以上)，差异蛋白或特征通路等进行空间分析
- 4) 代谢层面：GC-MS/LC-MS (可以通过通路等进行联合分析)，或可以联合空间代谢 (建议邻近切片测序)

(二) 样本准备

1. 新鲜样本的包埋是否需要客户自己完成？

答：

- 1) 做 FF 的样本：建议客户包埋好送公司切片。主要原因是防止直接冻存时间久，可能会导致组织内冰晶太多 (比如肺组织，皮肤及其他有空腔结构的组织)
- 2) 做 FFPE 的样本：直接提供蜡块即可

2. 切片是客户自己做还是公司完成？

答：两个平台要求不同

10x Visium 平台：切片由公司完成，直接提供组织或包埋好的组织块即可

Cytassist：可以由公司完成，也可提供切好的切片。如果客户自己切片，需要按照样本准备指南进行准备，可以联系销售索取。

3. FF 和 FFPE 的切片厚度是否有差异？

答：有差异。FF 样本在 10 μm 左右，部分样本 15 μm；FFPE 样本 5 μm 厚度。

4. FF 和 FFPE 分别需要多少张切片？

答：

FF：大约 20 张左右，每张 10 μm 厚度（分别需要进行 RNA 质检，HE 染色，透化优化，正式实验）。

FFPE：

样本类型	用途	厚度 (μm)	提供量	备注
石蜡卷及石蜡切片	RNA 抽提质检	5-20	60μm < 总厚度 ≤ 80μm (总厚度=厚度*张数)	新鲜石蜡卷放置于 Ep 管中 (每管 2 卷)
	HE 染色质检	5	2-3 张	用于框选正式实验区域
	Cytassist 正式实验	5	2 张，客户确认包含 ROI (Region of Interest)	务必参考样本准备指南贴于玻片上有效转片区域内

5. 样本如何收集？

答：

- 1) FF 样本：干冰运输和保存
- 2) FFPE：4℃运输和保存

6. 如果空间转录组和空间代谢组一起做，该如何取样？

答：注意事项：

- 1) 仅针对 FF 样本可行，石蜡包埋样本（FFPE）不能做空间代谢组
- 2) 包埋胶只能用莱卡特用胶，不能用 OCT

二、关于 spaceranger 的问题

1. 空间样本需要覆盖多少 spots 才能分析？

答：空间样本的大小取决能覆盖多少 spots，一般建议目标区域覆盖的 spots 尽量不要低于 50 个，否则后续无法完成聚类及相应分析。

2. 空间转录组测序的 reads 能否进行拼接？

答：不能，空间转录组测序技术有两种（探针捕获和 ployA 捕获），都只能针对特定带有标签的序列进行测序和基因定量，并不会进行拼接

3. Sequencing Saturation 需要达到多少才能继续分析？

答：空转目前对于测序饱和度没有特殊要求，但是这个指标可以反应样本的基因丰富度，所以可以对比单细胞，一般达到 40%以上为最佳。

4. 是否可以和单细胞 cellranger 一样进行 force cell？

答：不可以，单细胞是根据基因和 UMI 来预测是否为细胞，指标设定的不同，可以对于细胞数进行调整，但是空转的结果 spots 的多少是和组织部位有关的，理论上 spots 的数量是一定的，进行 force cell 并不合理。

5. 如果组织中存在比较大的非细胞结构（类似囊腔等），spaceranger 将这部分对应的 spots 识别出来继续分析，是否可行？

答：如果囊腔部分很小，比如覆盖半个 spots，是可以继续留用的，但如果面积很大，多个 spots 都未被组织覆盖，建议手动调整对应关系，将非组织部位的 spots 剔除为最佳。

三、关于报告和后续分析问题

1. 空间转录组测序结果，除了 spaceranger 软件处理之外，是否还进行其他质控？

答：不会，空间转录和单细胞不同，不会对部分 spots 进行剔除，避免造成结构不完整。但是会对 spots 的 UMI 和 gene 等进行可视化，和标准化分析。

2. 空转可以调整 cluster 的数目吗？

答：cluster 的划分是根据各 spot 基因表达模式相似性进行分析的，一般表达模式相似性高的细胞会聚在同一个 cluster，cluster 分出来的多少，和参数（resolution）设置有关，必要情况下，可以针对 resolution 参数调整，从而实现 cluster 分群数目的调整。

3. 如何判断是否存在批次效应？

答：空转也会存在批次，但是空转的批次判断，可能由于切片位置不同和个体差异，需要结合病理特征进行判断

- 1) 同一个组的样本是否处在不同的疾病阶段；
- 2) 两个组 cluster 完全不同，甚至是同一个 cluster 的不同分组 spots 完全不重合；
- 3) 同一个组的重复样本完全分开

4. 如何判断过度矫正？

答：一般总的细胞聚类做批次矫正，对于是否造成过度，并没有严格的标准作为参考。但是根据项目经验，可能出现的表现：

- 1) UMAP 会缩在一起，各 cluster 在 HE 上的分布没有明显的边界，很均匀；
- 2) 筛选不到明显的 marker 基因；
- 3) 聚类的 cluster 和病理结构（如癌和癌旁）完全没有对应关系；

5. 常用的批次矫正方法

答：MNN, Harmony, CCA, 不矫正批次为 PCA 分析

6. 空间转录组可以进行细胞类型鉴定吗？常用的方法是什么？

答：可以的，一般常用的是三种方法

- 1) 利用单细胞数据反卷积空间结果，例如 RCTD, SPOTlight, MIA 等
- 2) 利用 marker 基因在空间进行 addmodule score, GSEA 评分等
- 3) 利用 marker 基因解析 spots 的细胞类型组成，例如 CARDfree（不常用，结果比较差）

7. 为什么有的报告有 celltype 结果，有的没有？

答：一般情况如果对应的组织类型有单细胞的公开数据，或老师自己做了单细胞结果并知道细胞类型，空转的报告会默认提供 celltype 信息，如果没有上述的数据，会默认不提供。

8. marker 基因的鉴定原则是什么，top10 是如何筛选的？

答：一般认为某个 cluster 的 marker 基因为在某特定 cluster 中高表达，在其他 cluster 中低表达或不表达的基因，可以根据 pct1 和 pct2 的值来评估，top10 的 marker 基因是按照 Gene Diff 由大到小来排列的。

Pct1: clusterA 中表达 a 基因的 spots 数除以 clusterA 中 spots 的比值

Pct2: 除了 clusterA，其他 cluster 中表达 a 基因的 spots 除以其他 cluster 总 spots 的比值

$$\text{Gene Diff} = \frac{pct1}{pct2}$$

9. 空间 spots 的成分比较复杂，那么分 cluster 的意义在哪里？

答：空间聚类结果，也是具有一定的价值的

- 1) 如果关注空间的某个部位，可以通过聚类探究内部是否存在异质性；
- 2) 如果不能根据 HE 结果划分病理结构，可以考虑通过聚类探究是否存在不同的基因表达模式；

3) 可以根据聚类, 筛选特征基因, 从而推断不同位置的基因表达模式和功能特征。

10. 差异基因的筛选标准是什么?

答: 通常按照 $P < 0.05$, $FC > 2$ 来筛选差异基因, 但由于某些样本差异基因比较少, 可调整为 $FC > 1.2$ 或 1.5 , $P < 0.05$ 来筛选差异基因, 也是比较认可的。

11. 空间切片上存在不同的组织结构, 如癌和癌旁、不同脑区, 是否需要单独分析或讨论?

答: 建议分开, 这样不同的区域特征可以分别筛选, 但不一定要分开讨论, 比如癌和癌旁, 可以看到肿瘤, 边界和癌旁的微环境差异。

12. 空间转录组可以进行细胞亚型分析吗?

答: 可以的, 和单细胞一样, 如果认为某个区域存在较强的异质性, 可以考虑进行亚型分析。

13. 如果利用单细胞鉴定到了细胞亚型, 想在空间上进行分析, 是否可行?

答: 可以的, 具体方法可以参考问题 6

14. 空间测序数据可以与 bulk RNA-seq 联合分析吗, 有何好处?

答: 可以。与 bulk RNA-seq 联合分析的目的有两个:

- 1) 与空间转录组测序进行联合分析, 计算 person 相关性指数, 观察两组数据的相似性;
- 2) 针对 bulk 中筛选的特定基因和通路, 可以在空间上进行可视化, 分析同部位的共表达基因及微环境的细胞组成等。

15. 空转是否可以检测融合基因?

答: 不一定:

- 1) ployA 捕获平台, 是利用 mRNA 3' 末端捕获的 ployA 尾来捕获 mRNA, 因此检测到的序列也是靠近 3' 末端的序列, 所以如果其中一个基因距离 3' 末端较远, 可能检测不到;
- 2) 探针捕获平台, 是针对某些基因特定的保守其余设计的引物, 可能也涉及不到融合区域, 因此也有可能检测不到。

16. 关注的基因在测序结果中检测不到为什么?

答: 有三种情况, 1) 关注基因表达的丰度太低了, 导致捕获不到; 2) 有些基因的半衰期很短, mRNA 水平停留的时间很短, 导致捕获不到; 3) 用的蛋白 name 进行搜索, 可以转换 gene name 试一下。

17. 10X Genomics 单细胞转录组测序是否可以分析可变剪切?

答: 不可以。空间转录组测序只能捕获 mRNA 的 3' 端或特定保守区域进行测序, 因此无法分析可变剪切。

18. 10X Genomics 单细胞转录组测序是否可以分析 SNP 和 CNV?

答：目前空间转录组还不能分析 SNP， 可以进行 CNV 分析。

四、关于高级分析的内容

（一）拟时序的相关问题

1. 拟时序分析的方法有几种？

答：有两种情况：1) 沿用单细胞的拟时序分析方法，如 Monocle2；2) 空间拟时序方法，如 stLearn。

2. 拟时序分析什么情况下适用？

答：以下几种情况下可以考虑应用

- 1) 探究生长发育过程可以考虑用；
- 2) 肿瘤侵袭和不同亚克隆的相互转变可以考虑用。

3. 拟时序分析需要人为判定方向吗？如何判断？

答：需要。软件自动预测的方向，可能和实际生物学意义不符。判断方向的原则参考如下

- 1) 先验知识的积累（常用）
- 2) 正常疾病的样本，一般会指定正常组定位的细胞为起点（常用）
- 3) 根据不同分支基因的功能/通路富集情况进行推测（不常用）
- 4) 根据特定基因的表达，如凋亡、初始相关基因定位的分支，认为是起点（不常用）
- 5) 根据 HE 染色结果判断（如肿瘤的侵袭会有不同的体现）

（二）Scenic 的相关问题

1. 空间转录组是否可以做 scenic 分析？

答：可以

2. 是否可以按照分组/cluster 进行 regulon 比较？

答：可以的。

3. 什么情况下按照分组分析 regulon？

答：理论上什么情况都可以，

- 1) 一般针对 subcluster 分析时，如果不同的 subcluster 在组间差异比较大的情况，建议按照 cluster 之间的差异 regulon 来分析
- 2) 如果针对很重要的区域，但是细胞数很少，做 subcluster 效果不好的情况，可以考虑直接做组间的差异 regulon 分析

3) 如果某些区域在组间 cluster 分布差异不大, subcluster 看不到想要的信息, 也可以考虑直接分析组间的 regulon

4. 如果 RAS 和 RSS 的数据不一致, 怎么办?

答: RAS 和 RSS 侧重点不一样, 通常情况下, RAS 高不代表 RSS 高, 因此如果数据不一致, 可以重点参考 RAS 指标。

5. 活性高是否代表这个基因表达量也高?

答: 不一定

6. 挑选的某个 cluster 里面活性高的 regulon, 是否下游靶基因也是上调的?

答: 不一定, 有些转录因子可以抑制靶基因的转录, 造成靶基因下调

7. 确定了 regulon 之后还能做些什么?

答:

- 1) 分析层面, 可以针对 CSI 模块的 regulon 或重点关注的 regulon 及靶基因做富集分析, 这样不仅可以关注到具体的基因, 还能靶向相关的 pathway
- 2) 验证层面, 可以对相关 TF 或靶基因进行免疫荧光或 KO 等实验验证

(三) 细胞间通讯的相关问题

1. 空间通讯分析一定要有细胞类型的信息才能做吗?

答: 不一定。有两种情况。

- 1) 如果知道每个 spots 的细胞类型组成, 可以更好的评估这些 spots 之间的通讯主要和哪些细胞类型相关;
- 2) 如果不知道细胞类型的情况, 可以按照 spot cluster 做通讯分析。

2. cellphoneDB 和 cellchat 该如何选择?

答: 两个方法都能用, 但目前 cellchat 应用的广泛一些。两个方法主要有几点不同

- 1) cellphoneDB 收录的受配体对相对较多, cellchat 相对较少
- 2) cellchat 可视化的图更多样 cellchat 具有通路的信息

3. 如何筛选关键的受配体基因?

答: 如果没有关注的基因, 可参考如下原则

- 1) 先锁定细胞类型
- 2) 筛选 $P < 0.05$ 的关系对
- 3) 在 2) 的基础上, 筛选 prob(通讯概率, cellchat) 或 express(L-R 的表达量, cellphoneDB) 由大到小

4. cellphoneDB 没有收录通路信息，是否意味着不能关注通路？

答：cellphoneDB 本身没有通路信息，但如果老师想要关注一些细胞类型之间的受体配体信息，可以考虑针对 $P < 0.05$ 的关系对，或特定关注的关系对，进行 GO/KEGG 富集。

（四）CNV 的相关问题

1. 是否存在专门的数据库进行 CNV 预测？

答：目前没有数据库，一般是基本项目中正常的细胞，作为对照进行 CNV 分析。

2. 空间转录组一个 spots 可能包含多个细胞类型，肿瘤细胞和正常细胞共定位的话，正常细胞如何指定？

答：可以考虑从 3 个方面来指定：

- 1) 根据 HE 结果，将 stroma 区域划分出来，作为 Normal
- 2) 根据 RCTD、Spotlight 结果，将非肿瘤的 spots 指定成 Normal
- 3) 软件自动预测（准确性会差）

3. CNV group 应该怎么用？

答：CNV group 是 inferCNV 软件通过判定肿瘤细胞的 CNV 水平进行的分组，一般情况下，如果想要探究是否肿瘤具有不同的恶性程度，及不同恶性细胞之间的特征基因等信息，可以考虑用 CNV group 进行分类。此外，也可以通过原始聚类的 cluster 进行比较分析，不用 CNV group，也是可行的。