slingshot:细胞轨迹分析

项目名称: slingshot:细胞轨迹分析

项目负责:北京寻因生物科技有限公司

撰写人:谢卓明，信息分析工程师

2021 年 08 月 02 日

特别声明

本文档属于内部资料，仅提供给公司正式员工参阅;未经允许限制资料的传输，禁止公开传阅。

版本历史记录

|  |  |  |  |
| --- | --- | --- | --- |
| 版本号 | 日期 | 提交人 | 备注 |
| V1 | 2021.8.02 | 谢卓明 | slingshot:细胞轨迹分析 |
|  |  |  |  |

版权所有:北京寻因生物科技有限公司保留所有权利。

本产品或文档受版权保护，其使用、复制、发行和反编译均受许可证限制。未经北京寻因生物科技有限公司及其授权者的事先书面许可，不得以任何形式、任何手段复制本产品或文档的任何部分。

**目录**

[slingshot:细胞轨迹分析 1](#_Toc78807804)

[项目名称: slingshot:细胞轨迹分析 1](#_Toc78807805)

[项目负责:北京寻因生物科技有限公司 1](#_Toc78807806)

[撰写人:谢卓明，信息分析工程师 1](#_Toc78807807)

[2021 年 07 月 30 日 1](#_Toc78807808)

[1 文档说明 4](#_Toc78807809)

[1.1 关于手册 4](#_Toc78807810)

[1.2 面向的使用者 4](#_Toc78807811)

[1.3 信息反馈 4](#_Toc78807812)

[2 背景介绍 5](#_Toc78807813)

[3 流程使用说明 6](#_Toc78807814)

[3.1 模块简介 6](#_Toc78807815)

[3.2 模块参数描述 6](#_Toc78807816)

[3.2.1 配置文件 7](#_Toc78807817)

[3.2.2 脚本配置 7](#_Toc78807818)

[3.3 执行过程 7](#_Toc78807819)

[4 结果展示及解读 8](#_Toc78807820)

[5 文献实例 9](#_Toc78807821)

[6 参考文献 10](#_Toc78807822)

[7 问题反馈及建议 10](#_Toc78807823)

# 文档说明

## 关于手册

该文档旨在基于含有序列标签单细胞转录(sc-RNA)数据，利用基因表达矩阵进行机器学习分析，进而模拟时间发育过程的动态变化。

## 面向的使用者

寻因信息分析工程师。

## 信息反馈

如果对本文档有任何意见和建议，请联系:xiezhuoming@seekgene.com

# 背景介绍

在整个生命生长发育过程中，细胞都在不断从一种功能“状态”过渡到另一种功能“状态”。处于不同状态的细胞表达不同的基因，产生蛋白质和代谢物的动态重复，从而完成它们的工作。当细胞在不同的状态间转化时，会经历转录重组的过程，一些基因被沉默，而另一些则被激活。这些瞬态通常很难描述，因为在更稳定的端点状态之间进化细胞可能很困难或不可能。但由于这个过程是连续发生的，我们可以使用轨迹推断(TI，trajectory inference)的方法可以根据测序的细胞(瞬时状态)之间表达模式的相似性对单细胞沿着轨迹进行排序，以此来模拟细胞动态变化的过程。也就是重建分化轨迹或者拟时间轴。

Slingshot 旨在模拟单细胞 RNA 测序数据中的发育轨迹，作为降维和聚类后下游分析组件。 它足够灵活，可以处理任意多个分支事件，并允许通过监督图构建合并先验知识。

# 流程使用说明

## 模块简介

脚本名称: slingshot.R

脚本版本:v1.0

脚本路径: /PROJ/development/Seek\_Script/slingshot

## 模块参数描述

slingshot.R

|  |  |
| --- | --- |
| 参数 | 描述 |
| --seurat\_obj | 经过降维的 seurat 对象，RDS文件。 |
| --inference\_clusters | 设置用于轨迹推断的分组变量，默认是seurat\_clusters。 |
| --reduction\_space | 设置需要推断轨迹的降维空间，默认是umap。 |
| --start\_cluster | 设置轨迹推断起始 cluster，非必须。 |
| --end\_cluster | 设置轨迹推断结束 cluster，非必须。 |
| --outdir | 默认是./output。 |

### 配置文件

slingshot 主要用于处理聚类、降维后的单细胞对象，轨迹推断的基本单位是 cluster。

因此 slingshot.R 至少需要 seurat\_obj、inference\_cluster 和 reduction\_space 三个参数。

此外，和其他类似软件不同，slingshot 还可以指定轨迹推断的起始和终止 cluster，用于生物学意义层面的人为监督，适合于生物学背景已知的情况。

### 脚本配置

run.sh：

#!bin/bash

source /PROJ/development/xiezhuoming/soft/miniconda2/bin/activate /PROJ/development/xiezhuoming/soft/miniconda2/envs/Slingshot

Rscript slingshot.R \

--seurat\_obj /PROJ/development/xiezhuoming/proj/velocyto/input/demo/seekone\_demo/demo\_seurat.rds \

--inference\_cluster seurat\_clusters \

--outdir .

## 执行过程

运行 run.sh 。

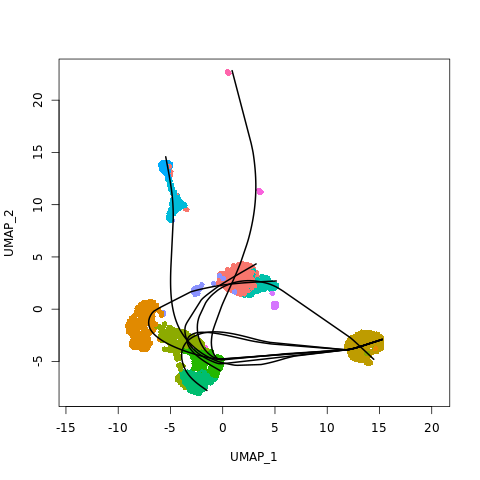
投递 qsub -cwd -l vf=50G,p=10 -V

# 结果展示及解读

(1) slingshot\_obj.rds

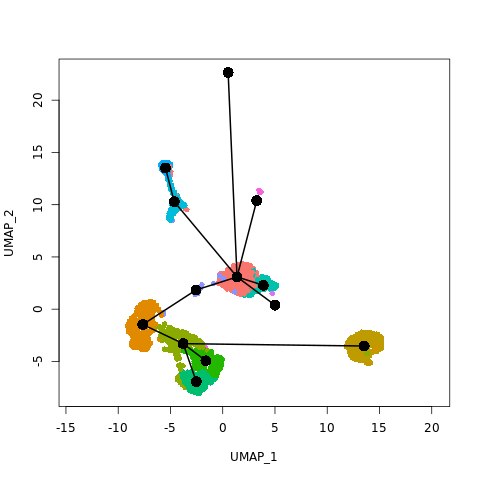
slingshot 对象，包含轨迹推断结果和画图坐标，可用函数 slingPseudotime 提取 Pseudotime value。

（2）curv.png



轨迹推断图，拟合曲线，更改轨迹推断开始和结束 cluster 可以修正拟合结果。

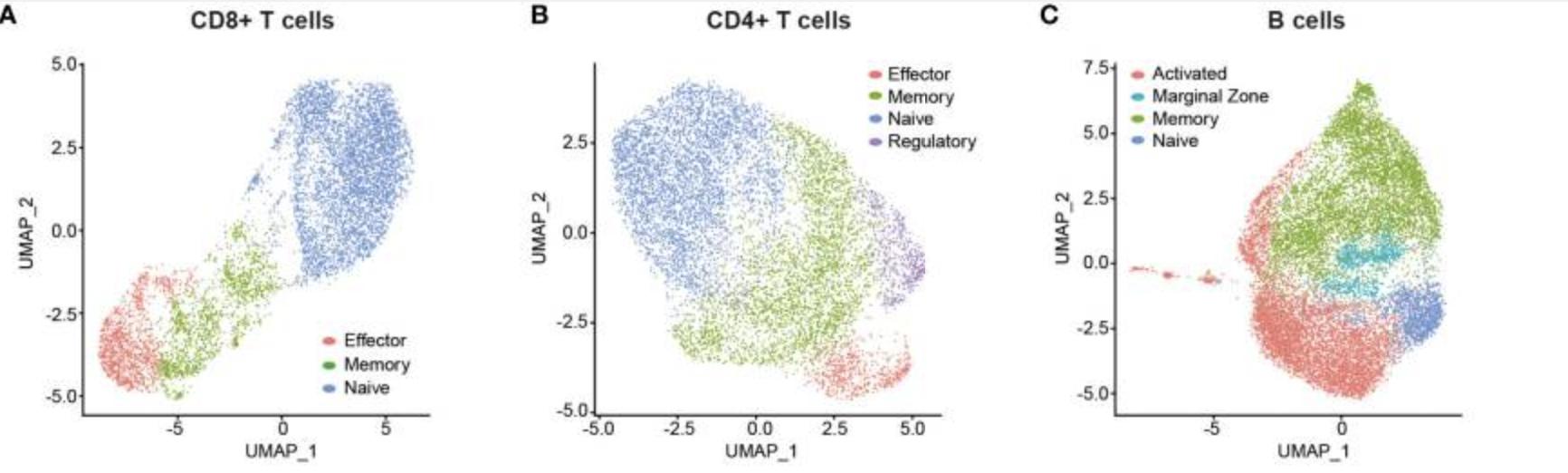
（3）line.png

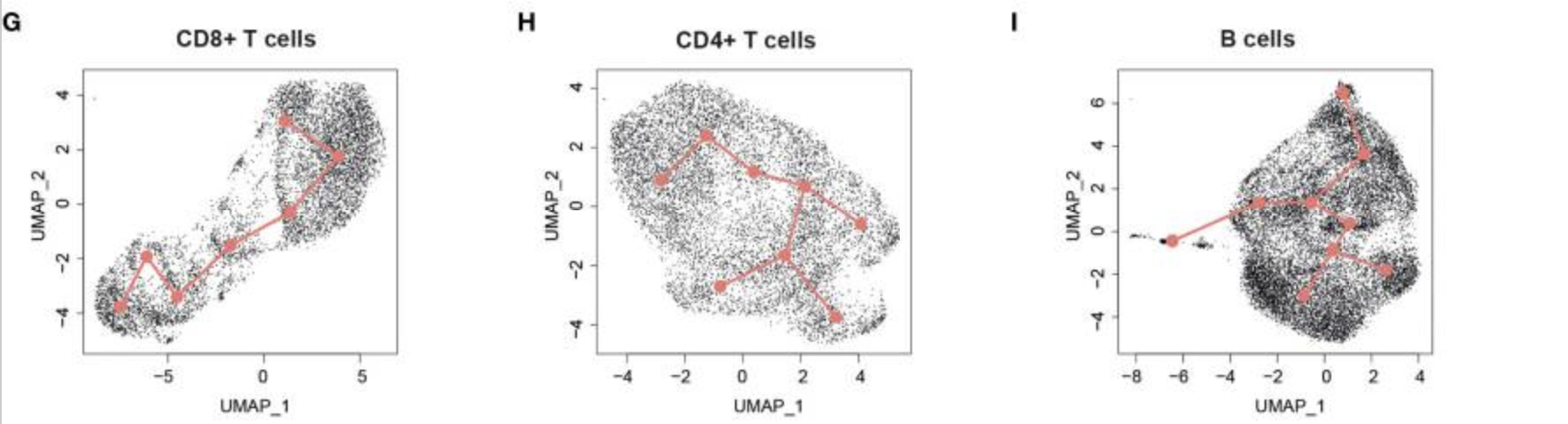


轨迹推断图，每个节点代表一个 cluster 。

# 文献实例

文章研究了康复期 COVID-19 患者的适应性免疫组库和转录组图谱。





在定义免疫细胞后，使用 slingshot 轨迹推断，验证了细胞定义结果，并且除了常规 naïve-MZ-memory 和 naïve-MZ-activated 轨迹外，还发现 MZ-memory-activated 轨迹，证明重新激活的记忆 B 细胞的存在。

# 参考文献

1. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, Purdom E, Dudoit S. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics. 2018 Jun 19;19(1):477. doi: 10.1186/s12864-018-4772-0. PMID: 29914354; PMCID: PMC6007078.
2. Bieberich F, Vazquez-Lombardi R, Yermanos A, Ehling RA, Mason DM, Wagner B, Kapetanovic E, Di Roberto RB, Weber CR, Savic M, Rudolf F, Reddy ST. A Single-Cell Atlas of Lymphocyte Adaptive Immune Repertoires and Transcriptomes Reveals Age-Related Differences in Convalescent COVID-19 Patients. Front Immunol. 2021 Jul 12;12:701085. doi: 10.3389/fimmu.2021.701085. PMID: 34322127; PMCID: PMC8312723.

# 问题反馈及建议

|  |  |  |  |
| --- | --- | --- | --- |
| 联系人 | 邮箱 | QQ | 备注 |
| 谢卓明 | xiezhuoming@seekgene.com |  |  |