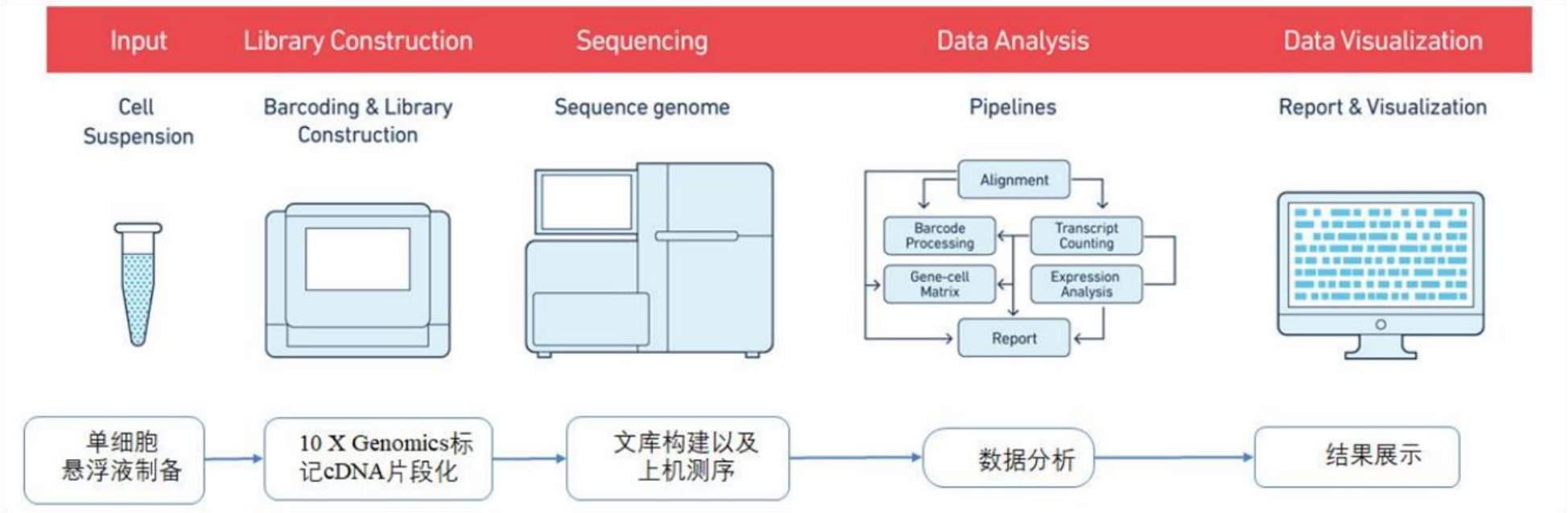
ScRNAPip：系统、动态的单细胞转录组数据分析工作流

徐莉敏[1,2,#]，张敬[3,#]，何依倩[1]，杨倩倩[1]，慕天浩[3]，郭秋实[1,2]，
李映樯[3]，童田[3]，陈实富[3,*]，叶德全[1,4,*]

[1]香港中文大学（深圳）福田生物医药创新研发中心
[2]香港中文大学（深圳）医学院
[3]深圳市海普洛斯生物科技有限公司
[4]香港中文大学（深圳）科比尔卡创新药物开发研究院

Xu, Limin, Jing Zhang, Yiqian He, Qianqian Yang, Tianhao Mu, Qiushi Guo, Yingqiang Li, Tian Tong, Shifu Chen, and Richard D.Ye. 2023. “ScRNAPip: A Systematic and Dynamic Pipeline for Single-Cell RNA Sequencing Analysis.” iMeta e132. https://doi.org/10.1002/imt2.132

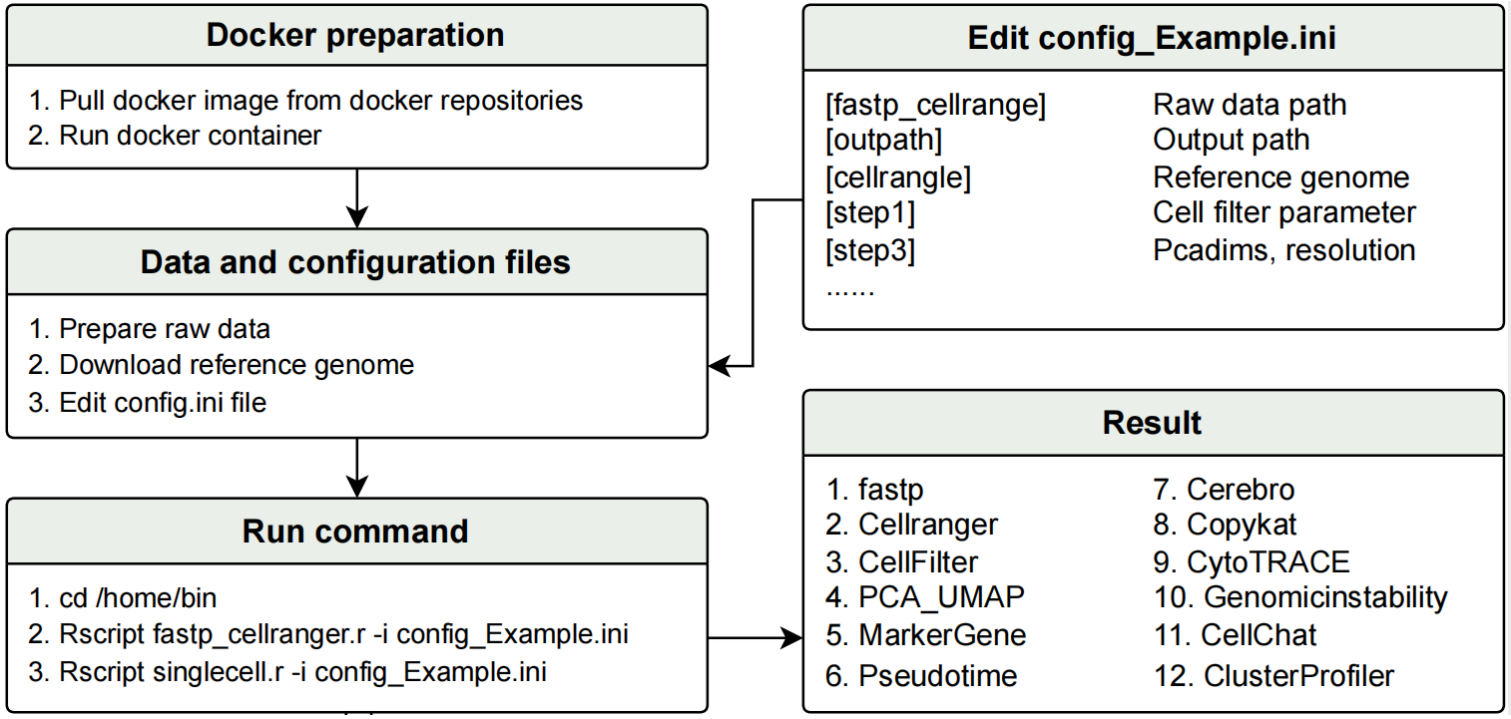
引言



- 目前，用于 scRNA-seq 数据分析的工具已经开发了 1,500 多种，涉及 30 多个类别。这可能对研究人员在选择适当的数据分析工具方面构成很大挑战。

- 在本研究中，我们旨在建立一个系统化、动态和可重复的工作流程，并指导用户完成 scRNA-seq 分析的关键步骤。

软件流程



第一步： 拉取 Docker 镜像

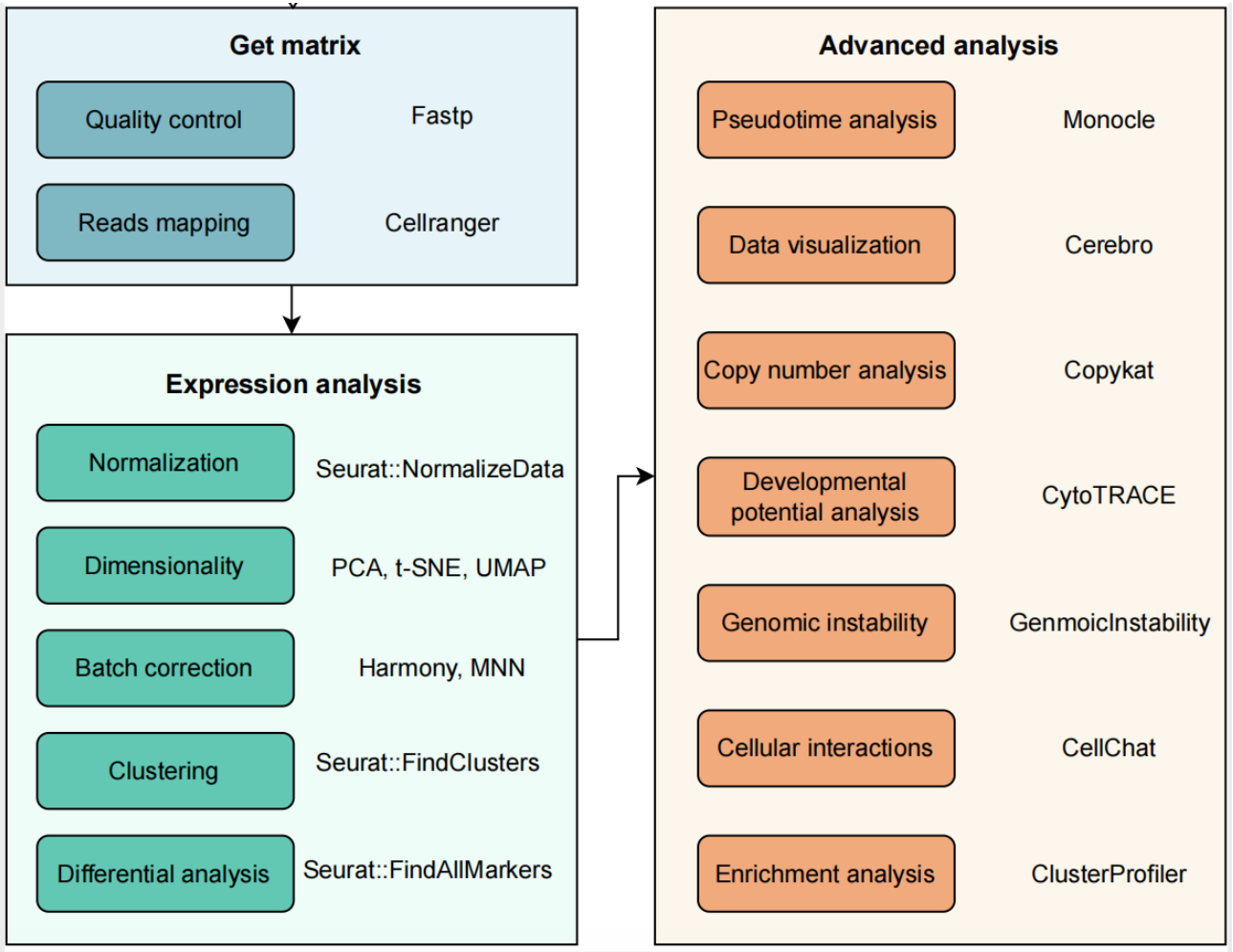
第二步： 数据和配置文件准备

第三步： 编辑配置文件

第四步： 流程执行

第五步： 结果查看

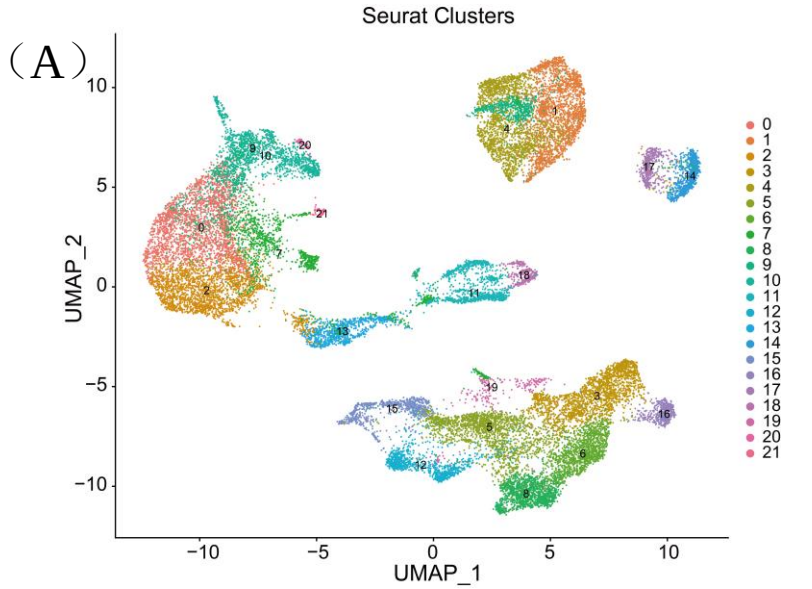
结果

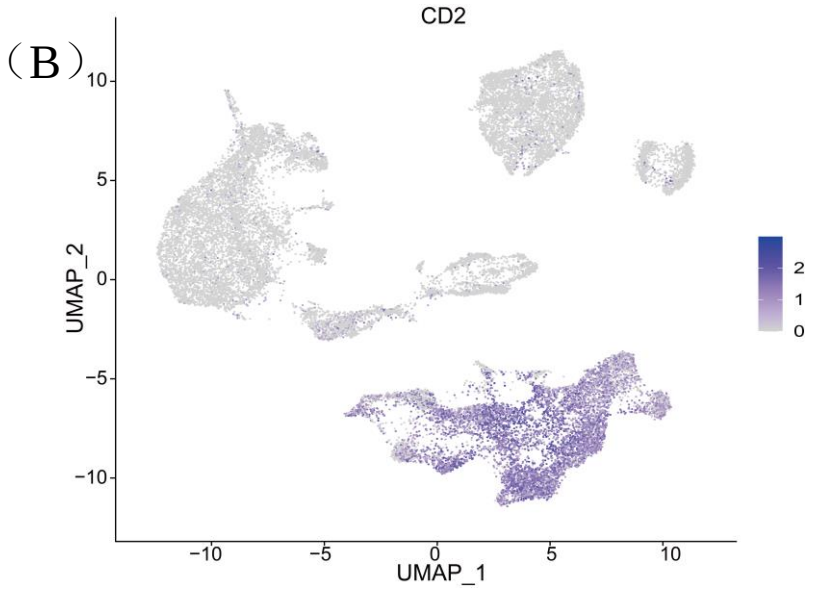


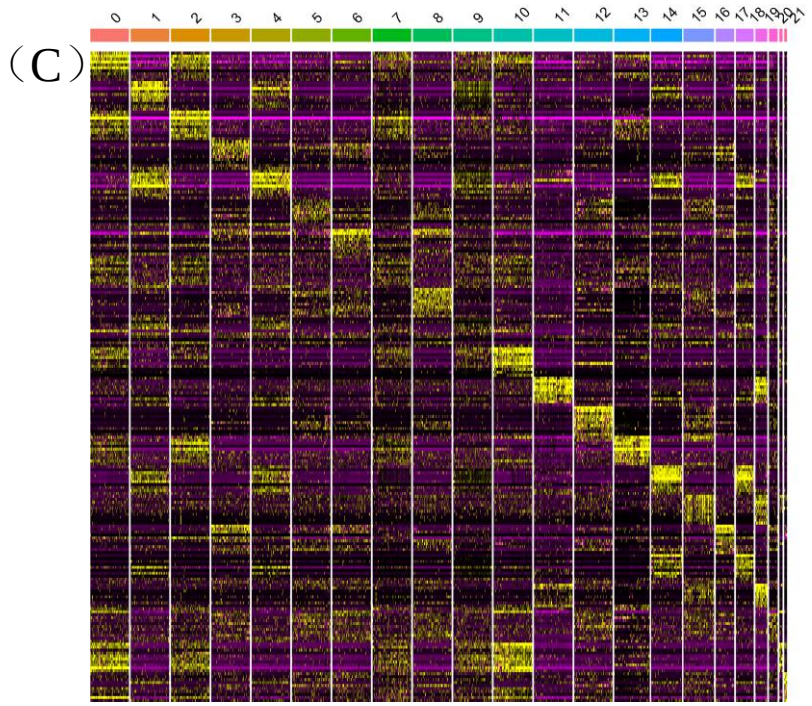
ScRNA-seq 分析包括以下步骤：

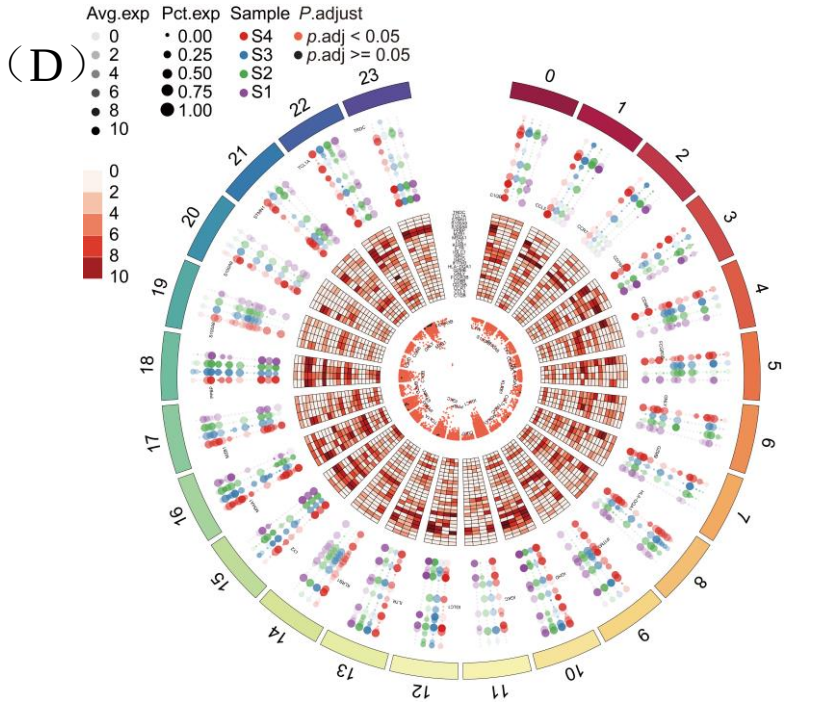
1）原始数据过滤
2）映射到参考基因组生成表达矩阵
3）过滤多个细胞、凋亡细胞和质量较差的细胞
4）归一化
5）降维和聚类
6）差异分析
7）数据可视化
8）环状图分析
9）拷贝数分析
10）发育潜力分析
11）基因组不稳定性分析
12）细胞相互作用分析
13）富集分析

结果









（A）UMAP 图：显示单细胞数据中的细胞聚类。

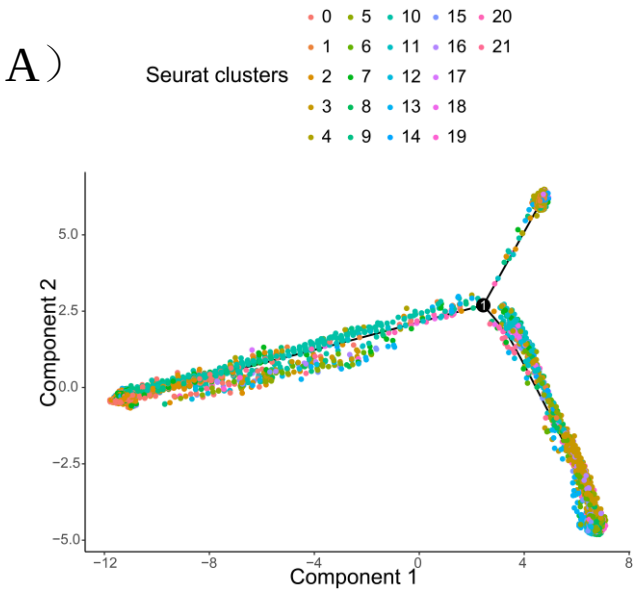
（B）基因表达染色图，图中展示 CD2 标记基因在单个细胞中的表达水平。

（C）Marker 基因热图，每个 cluster 根据 *p* 值挑选前 10 的 marker 基因绘制热图。

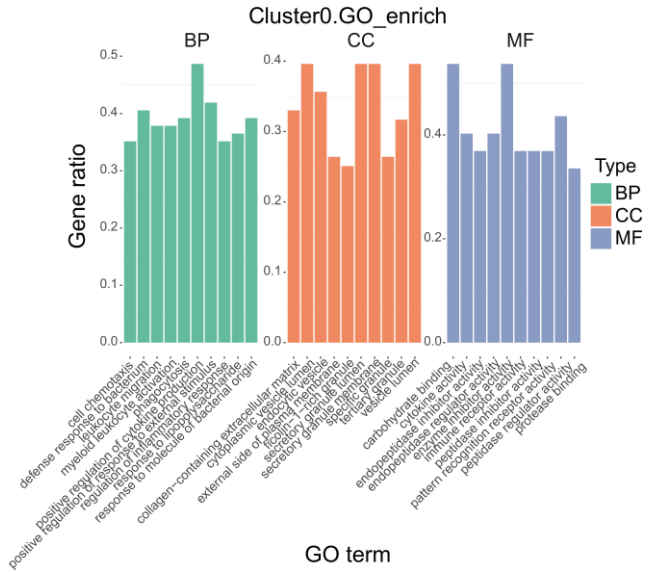
（D）Circos 图，选择每个 cluster 或细胞类型的关键 marker 基因，使用气泡图，热图和散点图进行展示并将他们排列在不同的轨道上。

结果

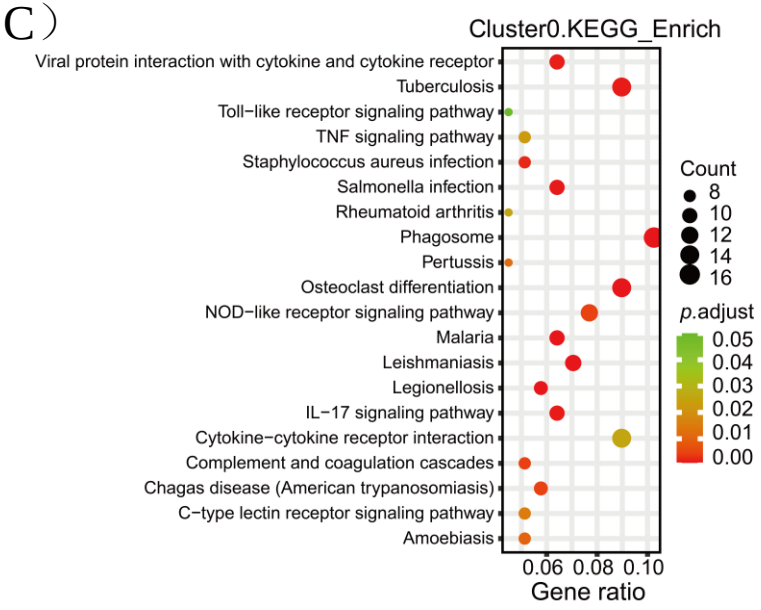
（A）



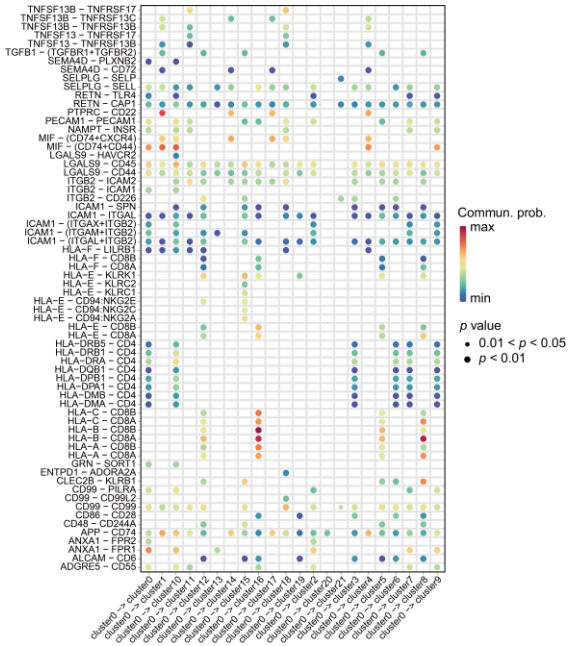
（B）



（C）



（D）



（A）使用 Monocle 构建的细胞轨迹图。

（B）GO 柱状图，展示显著富集的 GO 条目。

（C）KEGG 气泡图，展示显著富集的 KEGG 通路，点的大小代表该通路显著富集的基因数，颜色表示 *p* adjust。

（D）细胞通讯气泡图，所有以 cluster0 为配体的显著配体-受体对。点的颜色和大小表示计算的通讯概率和 *p* 值。

软件操作

拉取 ScRNAPip 的镜像：

```
@gm55b-1:~>docker pull zhangjing12/scrnapip
```

确认自己 ID 信息：

```
@gm55b-1:~>id
uid=20022(usr) gid=20022(usr) groups=20022(usr),979(docker)
@gm55b-1:~>
```

运行镜像：

```
@gm55b-1:~>docker run -d -e USERID=20022 -e GROUPID=20022 -v /thinker:/thinker zhangjing12/scrnapip
```

进入镜像：

```
@gm55b-1:~>docker exec -ti --user 20022:20022 7b4e02613359 bash
```

软件操作

修改配置文件：

注：配置文件示例在镜像中 /home/bin/config_Example.ini 路径下

```
[fastp_cellrange]
SR39.R1=["/scRNAtest/data/rawdata/SRR8448139_S1_L001_R1_001.fastq.gz"]
SR39.R2=["/scRNAtest/data/rawdata/SRR8448139_S1_L001_R2_001.fastq.gz"]
SR40.R1=["/scRNAtest/data/rawdata/SRR8448140_S1_L001_R1_001.fastq.gz"]
SR40.R2=["/scRNAtest/data/rawdata/SRR8448140_S1_L001_R2_001.fastq.gz"]
SR41.R1=["/scRNAtest/data/rawdata/SRR8448141_S1_L001_R1_001.fastq.gz"]
SR41.R2=["/scRNAtest/data/rawdata/SRR8448141_S1_L001_R2_001.fastq.gz"]
SR42.R1=["/scRNAtest/data/rawdata/SRR8448142_S1_L001_R1_001.fastq.gz"]
SR42.R2=["/scRNAtest/data/rawdata/SRR8448142_S1_L001_R2_001.fastq.gz"]


[indata]
SR39="/scRNAtest/outfile/00.cellranger/SR39/SR39/outs/filtered_feature_bc_matrix"
SR40="/scRNAtest/outfile/00.cellranger/SR40/SR40/outs/filtered_feature_bc_matrix"
SR41="/scRNAtest/outfile/00.cellranger/SR41/SR41/outs/filtered_feature_bc_matrix"
SR42="/scRNAtest/outfile/00.cellranger/SR42/SR42/outs/filtered_feature_bc_matrix"
#       NC4="/celltest/00.cellranger/NC4/outs/filtered_feature_bc_matrix"
[outpath]
outpath="/scRNAtest/outfile"
[tempdata]
#传入的rds的路径，优先考虑默认路径下，其次再在该路径下寻找
tempdata="workout"

[run]
#运行那些步骤，需要运行的为true，不需要的为false或删除
fastp=true
cellrangle=true
step1=true
step2=true
step3=true
step4=true
step5=true
step6=true
step7=true
step8=true
step9=true
step10=true
step11=true
```

软件操作

运行 cellranger 命令：

```
rstudio@245c9361c1d0:/$ nohup Rscript /home/bin/fastp_cellranger.r  -i config_Example.ini 2>&1 > nohup.log &
```

运行分析步骤命令:

```
rstudio@245c9361c1d0:/$ nohup Rscript /home/bin/singlecell.r  -i config_Example.ini  2>&1 > nohup.log &
```

结果查看：

```
@gm55b-1:~>cd workout
@gm55b-1:~>ls
01.Fastp        03.CellFilter  05.MarkerGene  08.copykat     10.genomicinstably  12.ClusterProfiler  temp
02.Cellranger   04.PCA_UMAP    07.Cerebro     09.cytoTRACE   11.cellchat         shell
```

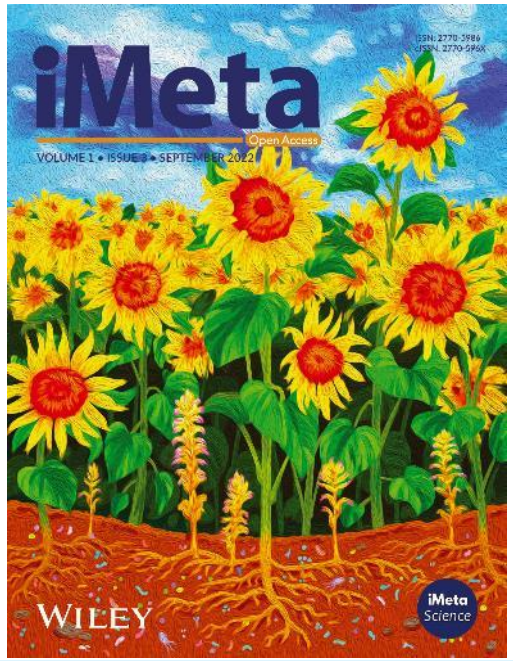
总结

- ScRNAPip 提供了一站式、用户友好的工作流程；
- 全面涵盖从原始测序数据到个性化分析的整个过程，生成易于阅读和解析的结果；
- 分析参数的选择具有灵活性，适用于多个癌种。

◆ ScRNAPip 可以通过 GitHub（https://github.com/OpenGene/scrnapip）公开获取。软件的使用方法和分析结果的详细描述可以在该网站上找到。

iMeta: 整合宏组学重新认识生命和环境科学

iMeta WILEY



"*iMeta*" 是由威立、肠菌分会和本领域数百位华人科学家合作出版的开放获取期刊，主编由中科院微生物所刘双江研究员和荷兰格罗宁根大学傅静远教授共同担任。目的是发表原创研究、方法和综述以促进宏基因组学、微生物组和生物信息学发展。目标是发表前10%(IF > 15)的高影响力论文。期刊特色包括视频投稿、可重复分析、图片打磨、青年编委、前3年免出版费、50万用户的社交媒体宣传等。2022年的三月、六月和九月期已正式在线出版发行!

主页: http://www.imeta.science
出版社: https://wileyonlinelibrary.com/journal/imeta

投稿: https://mc.manuscriptcentral.com/imeta

office@imeta.science

iMeta 宣传片