

BIOF3 FIGCODE · SCI 绘图代码集

UMAP 聚类图 (UMAP)

导出日期: 2026年6月27日



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。

分类: 降维可视化 | 依赖包: ggplot2、ggrepel

解决的生物学问题

单细胞数据中存在哪些细胞亚群? 各亚群之间的关系如何?

应用场景

- 单细胞聚类结果展示
- 细胞类型注释可视化
- 不同样本/条件的细胞分布对比
- 多模态整合 (WNN) 降维展示

输入数据格式

Seurat 对象 (已跑 RunUMAP + FindClusters)

关键参数

- resolution 聚类分辨率 (0.1-2.0)
- n.neighbors UMAP 邻居数 (30)
- min.dist 最小距离 (0.3)
- dims 使用的 PC 数 (1:30)
- label.size 标签大小

快速开始

1. 直接在线运行

打开 [FigCode 在线绘图](#)，点击本工具卡片的"在线绘图"按钮即可用内置示例数据出图，零环境配置。

2. 本地复现

下载脚本和示例数据，本地 RStudio 运行：

```
# 下载
curl -O https://<your-site>/figcode/scripts/umap.R
curl -O https://<your-site>/figcode/data/umap.csv
```

3. 安装依赖

```
# CRAN 包
install.packages(c("ggplot2", "ggrepel"))

# Bioconductor 包 (如需)
# BiocManager::install(c())
```

完整代码

```
library(ggplot2)
library(ggrepel)

# --- Demo: simulate UMAP coordinates ---
set.seed(42)
n_cells <- 800
clusters <- sample(paste0("C", 0:7), n_cells, replace = TRUE, prob = c(0.2,0.18,0.15,
# Generate cluster centers
centers <- data.frame(
  cluster = paste0("C", 0:7),
  cx = c(-5, 5, 0, -3, 4, -6, 6, 0),
  cy = c(3, 3, -4, -2, -3, -5, 0, 5)
)
umap_df <- data.frame(cluster = clusters)
umap_df <- merge(umap_df, centers, by = "cluster")
umap_df$UMAP1 <- umap_df$cx + rnorm(n_cells, 0, 0.8)
umap_df$UMAP2 <- umap_df$cy + rnorm(n_cells, 0, 0.8)

# Cluster labels at centroids
label_df <- aggregate(cbind(UMAP1, UMAP2) ~ cluster, data = umap_df, FUN = mean)

# --- Plot ---
ggplot(umap_df, aes(x = UMAP1, y = UMAP2, color = cluster)) +
  geom_point(size = 0.5, alpha = 0.7) +
  geom_label_repel(data = label_df, aes(label = cluster), size = 3.5, fontface = "bold",
  labs(title = "UMAP Clustering", x = "UMAP 1", y = "UMAP 2") +
```

```
theme_classic() +  
theme(legend.position = "none")
```

替换为自己的数据

脚本中以 `# --- Demo data ---` 标注的段落是示例数据生成代码。替换为自己的数据时，**保持列名一致**即可：

- 输入格式：Seurat 对象（已跑 RunUMAP + FindClusters）
- 使用 `read.csv()` / `readRDS()` 读取本地文件

延伸阅读

- 相关教程：[UMAP 聚类图 完整流程](#)
- 出现 bug？欢迎在 [FigCode 页面](#) 点击对应卡片，在评论区留言。