

BIOF3 FIGCODE · SCI 绘图代码集

表达分布直方图 (Expression Distribution)

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: 分布可视化 | 依赖包: ggplot2

解决的生物学问题

目标基因在高/低表达组中的分布是否清晰分离? 截断值是否合理?

应用场景

- 生存分析分组验证
- 基因表达二分类合理性检查
- 截断值敏感性分析

输入数据格式

数据框: gene_expr, group

关键参数

- bins 直方图柱数
- cutoff 截断值
- fill 分组颜色
- alpha 透明度

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/expr-dist.R
curl -O https://<your-site>/figcode/data/expr-dist.csv
```

3. 安装依赖

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

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

完整代码

```
library(ggplot2)

# --- Demo data ---
set.seed(42)
n <- 100
df <- data.frame(
  gene_expr = rnorm(n, mean = 8, sd = 2),
  group = factor(rep(c("Low", "High"), each = 50), levels = c("Low", "High"))
)
cutoff_val <- median(df$gene_expr)

# --- Plot ---
ggplot(df, aes(x = gene_expr, fill = group)) +
  geom_histogram(bins = 30, alpha = 0.7, position = "identity") +
  geom_vline(xintercept = cutoff_val, linetype = "dashed", color = "gray30", linewidth = 2) +
  scale_fill_manual(values = c(Low = "#2563eb", High = "#dc2626")) +
  labs(x = "Gene Expression", y = "Count",
       title = "Expression Distribution",
       subtitle = paste0("Cutoff = ", round(cutoff_val, 3))) +
  theme_classic() +
  theme(plot.title = element_text(size = 13, face = "bold"), legend.position = "top")
```

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：数据框：gene_expr, group
- 使用 `read.csv()` / `readRDS()` 读取本地文件

延伸阅读

- 相关教程：[表达分布直方图 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。