

BIOF3 FIGCODE · SCI 绘图代码集

表达分布图 (Violin + Boxplot)

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

宝儿，这个图真好看！！！目标基因在不同实验条件或细胞类型中的表达分布是否存在差异？

应用场景

- 候选基因在各组中的表达验证
- 单细胞 marker 基因表达展示
- 蛋白丰度分布对比
- 药物处理前后表达变化

输入数据格式

长格式: gene, condition, expression

关键参数

- 展示基因列表
- fill 分组变量
- scales 是否自由 y 轴 (free_y)
- jitter width 抖动宽度
- 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/violin.R
curl -O https://<your-site>/figcode/data/violin.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)
genes <- c("TP53", "BRCA1", "MYC", "EGFR")
conditions <- c("Control", "Treatment")
df <- expand.grid(gene = genes, condition = conditions, rep = 1:20)
df$expression <- rnorm(nrow(df), mean = 8, sd = 1.5)
# Add treatment effect
df$expression[df$condition == "Treatment" & df$gene %in% c("MYC", "EGFR")] <-
  df$expression[df$condition == "Treatment" & df$gene %in% c("MYC", "EGFR")] + 2.5

# --- Plot ---
ggplot(df, aes(x = condition, y = expression, fill = condition)) +
  geom_violin(width = 0.8, alpha = 0.7, color = NA) +
  geom_boxplot(width = 0.15, fill = "white", outlier.shape = NA) +
  geom_jitter(width = 0.1, size = 1.5, alpha = 0.6) +
  facet_wrap(~ gene, scales = "free_y") +
  scale_fill_manual(values = c(Control="#475569", Treatment="#dc2626"), guide = "none") +
  labs(x = NULL, y = "Expression", title = "Gene Expression Distribution") +
  theme_classic()
```

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：长格式：gene, condition, expression
- 使用 `read.csv()` / `readRDS()` 读取本地文件

延伸阅读

- 相关教程：[表达分布图 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。