

2025 年国际定向进化大赛实验记录

实验时间：2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

17. 转录组学蛋白组学联合分析

一、实验目的

- 1.1 通过转录组与蛋白质组联合分析，系统解析人膀胱移行细胞癌 T24-RC48 耐药细胞模型中基因表达与蛋白质水平的关联性与差异性。
- 1.2 筛选在转录和蛋白层面共同显著变化的基因/蛋白，挖掘其参与的生物学功能及信号通路，揭示耐药相关的潜在调控机制。
- 1.3 验证转录本水平变化是否直接决定蛋白质丰度变化，探究耐药表型形成的多组学协同调控网络。

二、实验流程

2.1 样本信息与分组

转录组样本：T24_R_high（高度耐药）、T24_R_mode（中度耐药）、T24_R_low（低度耐药）、T24_R_control（野生型）

蛋白质组样本：T24_R_high（高度耐药）、T24_R_mode（中度耐药）、T24_R_low（低度耐药）、T24_R_control（野生型）

2.2 差异分析标准

转录组： $|\log_2FC| \geq 1.0$, $p_{adj} < 0.05$

蛋白质组： $|\log_2FC| \geq 0.263$, $p_{value} < 0.05$

2.3 联合分析流程

总体结果比较：Venn 分析、基因表达分布、相关性分析（线性拟合 R^2 值评估转录与蛋白水平一致性）。

差异结果比较：差异基因/蛋白统计、Venn 分析、聚类热图。

2.4 数据挖掘分析：

GO 功能富集：共有 GO 功能（BP、MF、CC）挖掘，通过 Venn 图、热图、气泡图、柱状图展示。

KEGG 通路富集：共有通路筛选，结合通路图及网络图解析关键信号通路。

关键基因筛选：基于共表达趋势（如 UP-UP、DOWN-DOWN 等）及 PPI 网络分析，鉴定核心耐药相关分子。

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

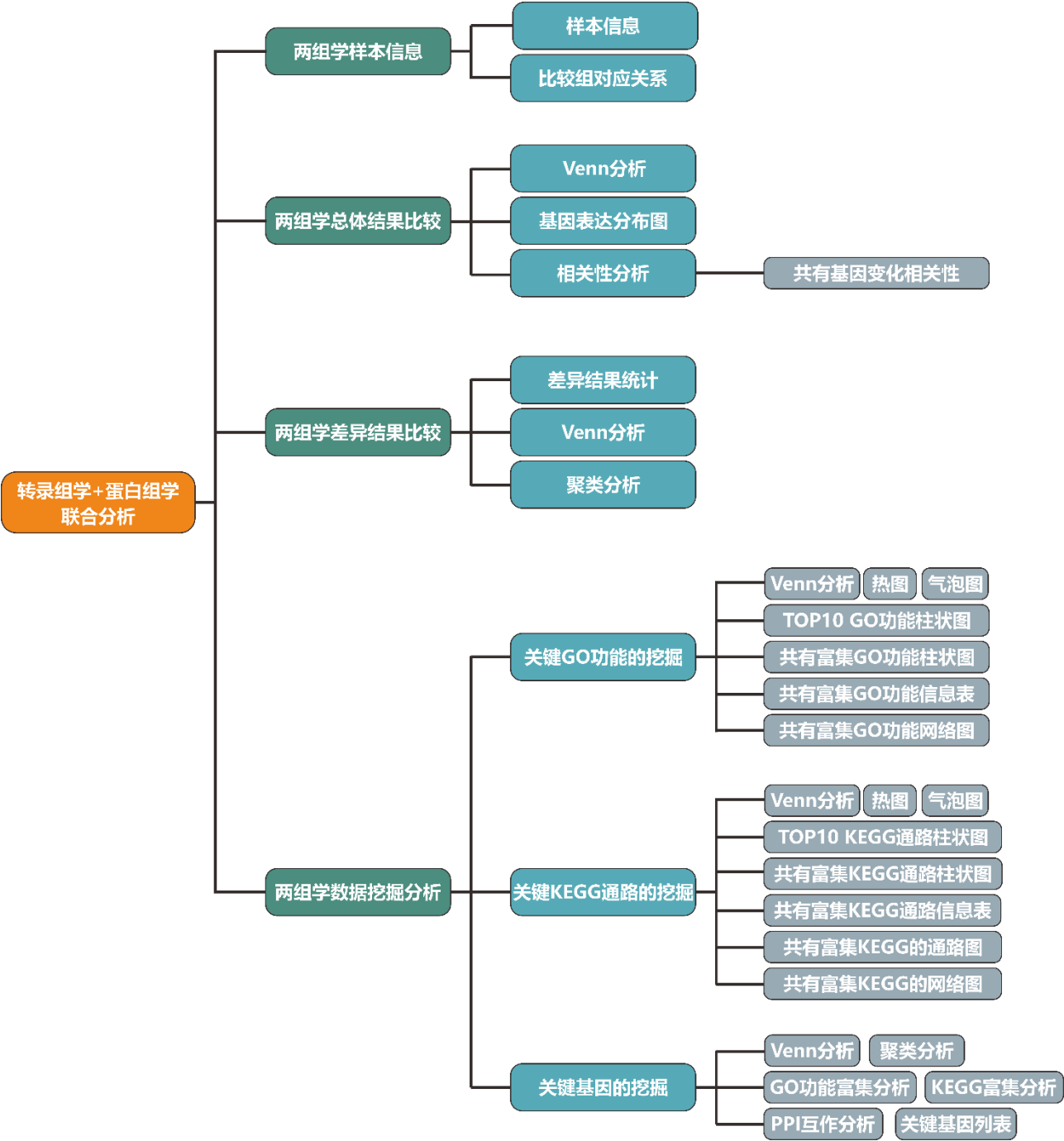


图 1 联合分析流程及方法

三、实验结果

3.1 两组学总体结果比较

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

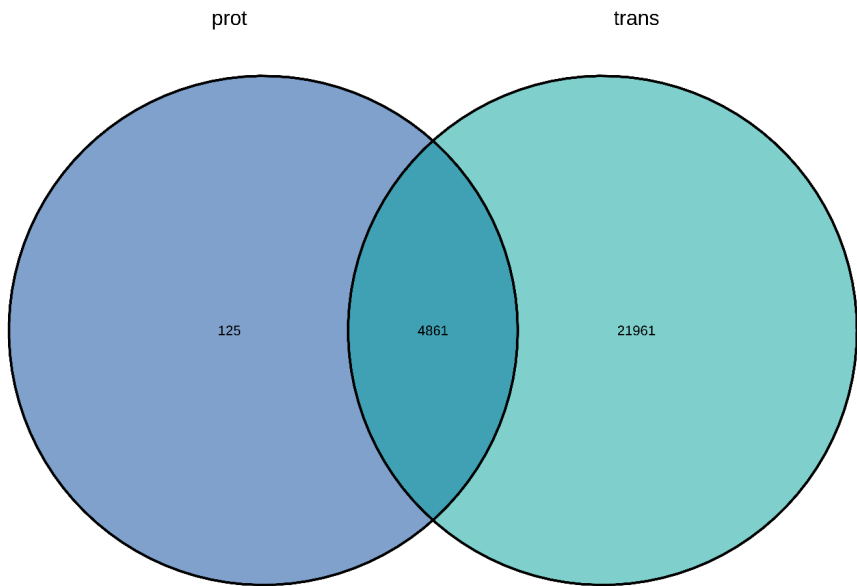


图 2 差异基因 Venn 图

基因表达分布图（图 3）显示，转录组基因表达覆盖范围更广（蓝色分布），而蛋白质组数据集中分布于中低表达区域（灰色分布）。

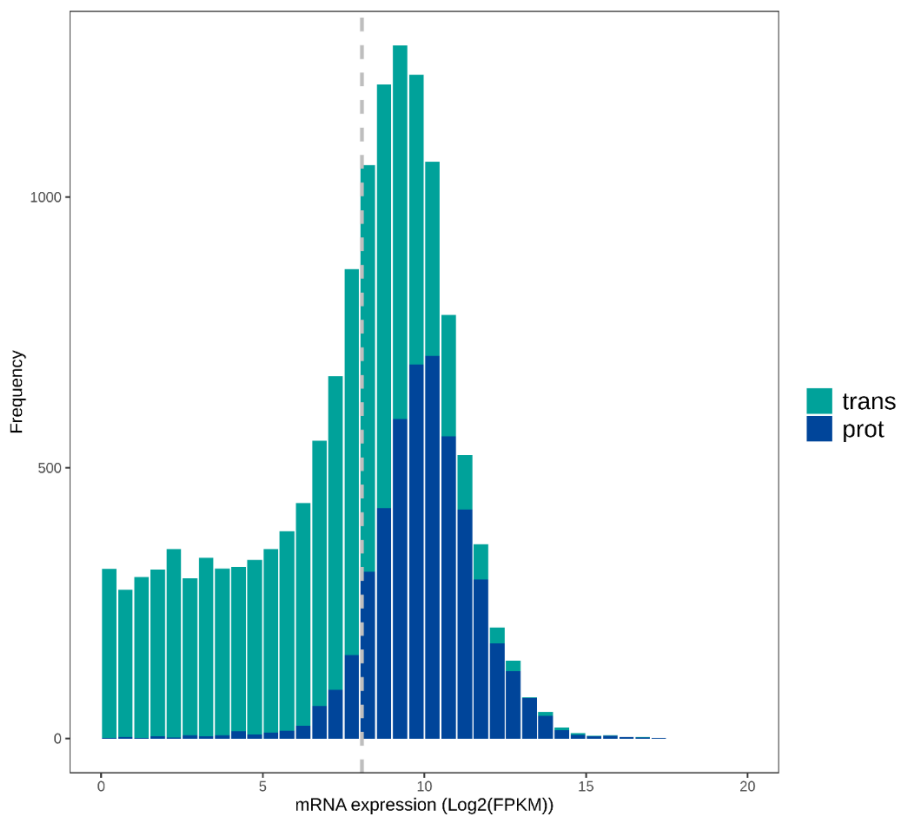


图 3 基因表达分布图

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

相关性分析是指对两个或多个具备相关性的变量元素进行分析，从而衡量两个变量因素的相关密切程度。考虑到“基因的表达在转录水平和蛋白质水平的一致性如何”是很多研究者关注的问题。我们采用线性拟合的方式（线性方程 $y=kx+b$ ）对转录组学和蛋白质组学的共有基因进行相关性分析，帮助我们分析在相同生理病理条件下实验组与对照组的基因在转录水平和蛋白质水平的整体相关性。相关性系数用 R 来表示。 R^2 为判定系数，用于评价试验数据与拟合函数之间的吻合程度。 R^2 值越接近 1，吻合程度越高，表明共有基因在转录和蛋白表达水平之间的相关性越高； R^2 越接近 0，则吻合程度越低，表明共有基因在转录和蛋白表达水平之间的相关性越低。

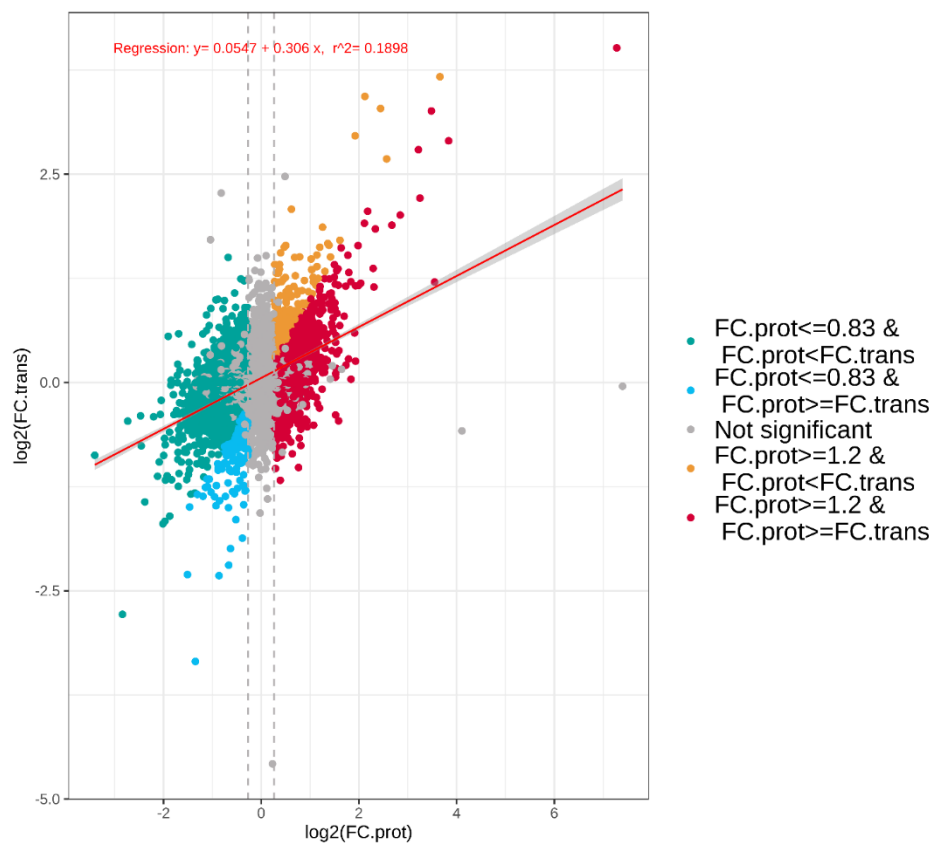


图 4 基因变化相关性图

3.2 两组学差异结果比较

我们将 bulk 转录水平上的组间差异基因表达结果与蛋白质水平上的组间差异分子表达结果进行分析比较。本项目中转录组学差异基因筛选标准为： $|\log_2 FC| \geq 1.0$, $p_{adj} < 0.05$ ；蛋白组学差异蛋白筛选标准为： $|\log_2 FC| \geq 0.263$, $p_{value} < 0.05$ 。两组学的差异数量统计结果，如下表所示：

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

表 1 差异结果统计表

联合分析方案名	转录组学			蛋白组学		
	UP	DOWN	ALL	UP	DOWN	ALL
uniteT24_high_vs_uniteT24_con	1118	941	2059	1229	1327	2556
uniteT24_high_vs_uniteT24_low	413	355	768	1260	1266	2526
uniteT24_high_vs_uniteT24_mode	360	212	572	1094	1081	2175
uniteT24_mode_vs_uniteT24_low	359	484	843	723	820	1543
uniteT24_mode_vs_uniteT24_con	1089	1129	2218	894	1049	1943
uniteT24_low_vs_uniteT24_con	569	379	948	424	519	943

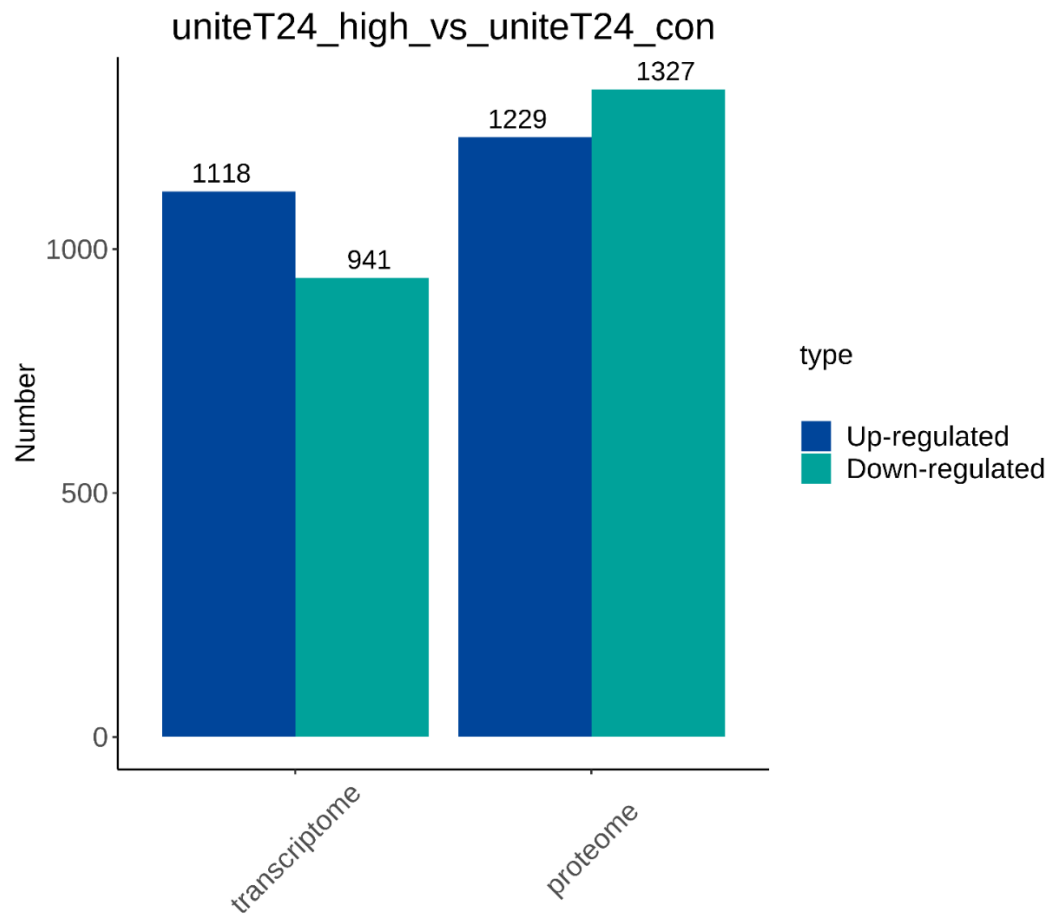


图 5 差异结果柱状图

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

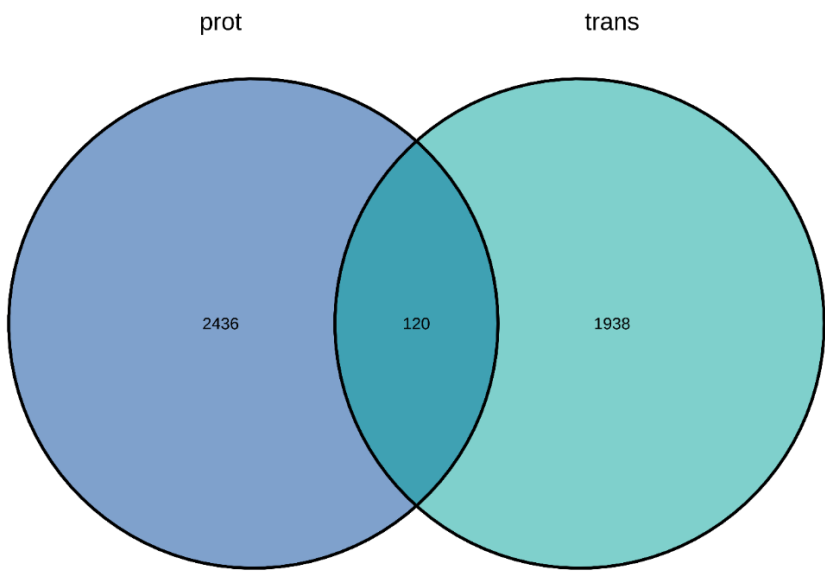


图 6 差异结果 Venn 分析图

对在上述 Venn 分析中得到的共有差异基因在两个组学中的 FC 值 (\log_2FC) 进行聚类热图分析，结果展示如下：

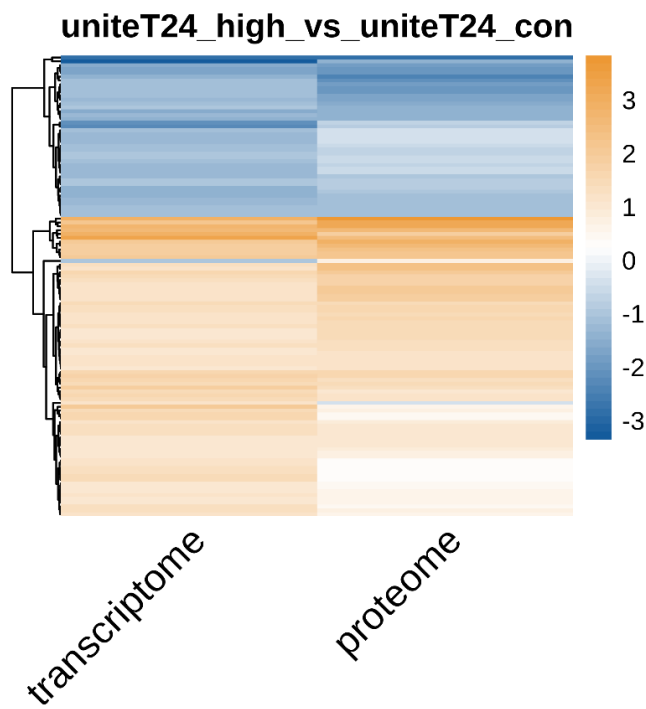


图 7 聚类分析热图

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

3.3 两组学数据挖掘分析

3.3.1 关键 GO 功能富集

基因本体论（Gene ontology, GO）是一种系统地对物种基因及其产物属性进行注释的方法和过程。它的目标是：1）维护和发展有限的基因及其产物属性描述的词汇；2）注释基因及其产物，同化和传播注释数据；3）提供方便的工具访问数据；4）实现在实验数据的基础上，使用 GO 进行程式解析，例如基因富集组分分析。它主要包括三个分支：细胞组件、分子功能和生物过程。本项目对两组学的差异 mRNA 和差异蛋白质分别富集的 GO 功能结果进行比较分析，以确定关键变化的 GO 功能，更全面深入的理解差异分子参与调控的生物学功能。

对两组学差异分子分别富集到的显著 GO 功能（转录组学：p_value<0.05；蛋白组学：p_value<0.05）进行 Venn 分析，结果如下图所示：

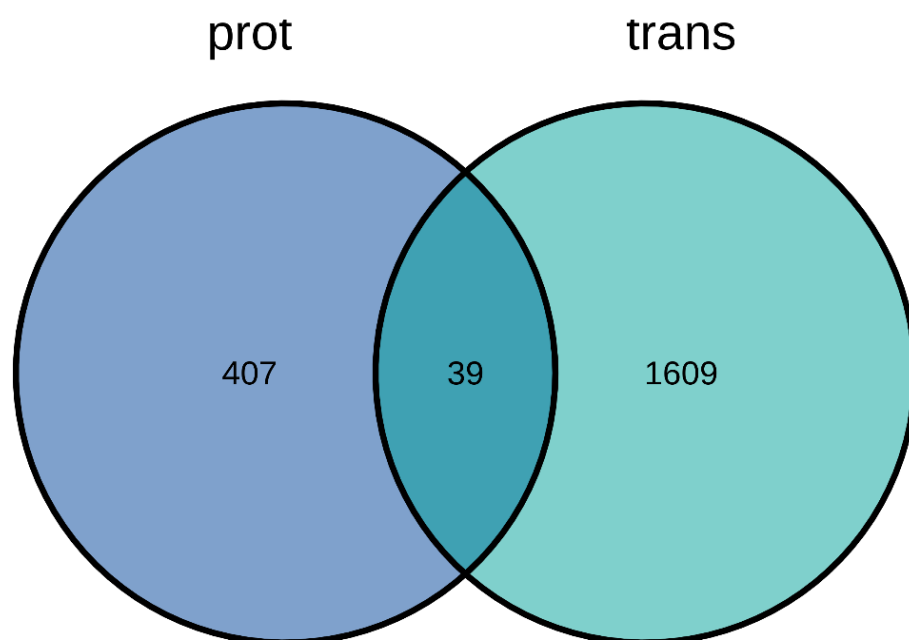


图 8 GO 功能富集 Venn 图

利用并集两组学 GO 功能富集结果（p_value<0.05）的 p_value 值进行聚类热图分析。结果展示如下：

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

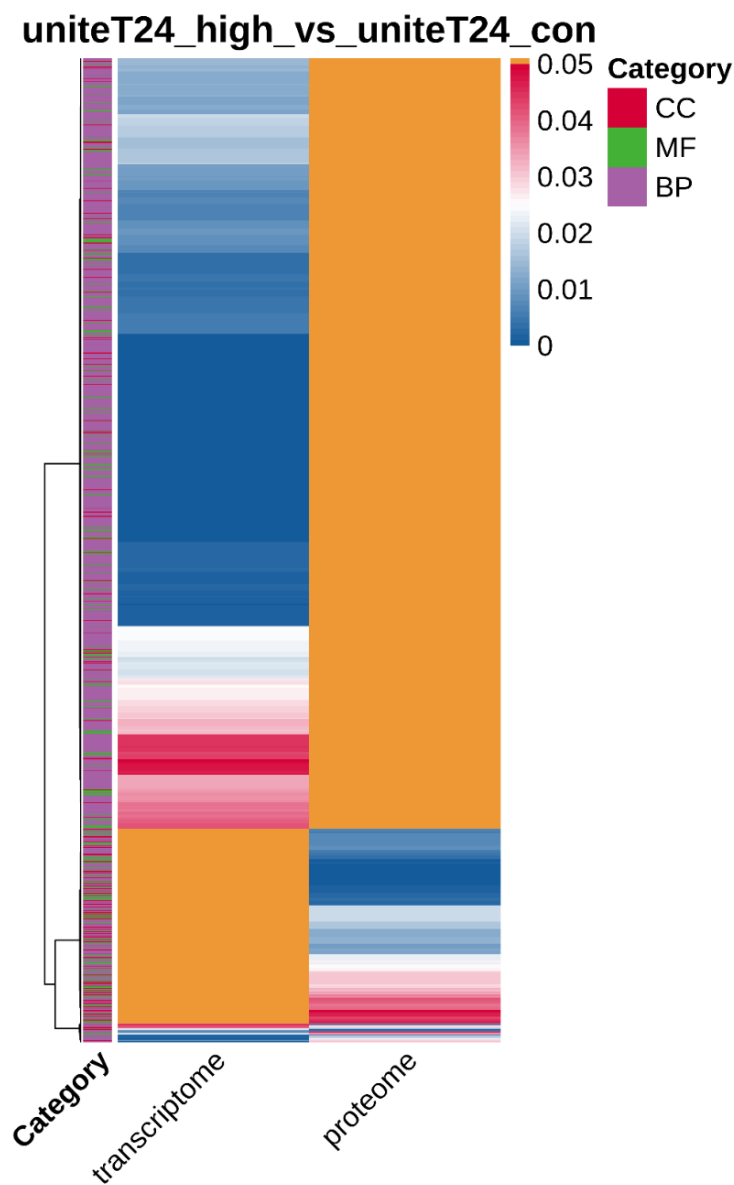


图 9 GO 功能热图

气泡图不仅直观展示了各组学显著性富集到的 GO 功能，同时也直观反映出两组学 GO 功能富集结果的共性和特异性。

基于两组学差异分子富集到的 GO 功能，我们以富集显著性 p_value 值从小到大进行排序，选择各组学 p_value 值最小的前 10 条通路，即合计至多 20 条通路以气泡图形式呈现如下。

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

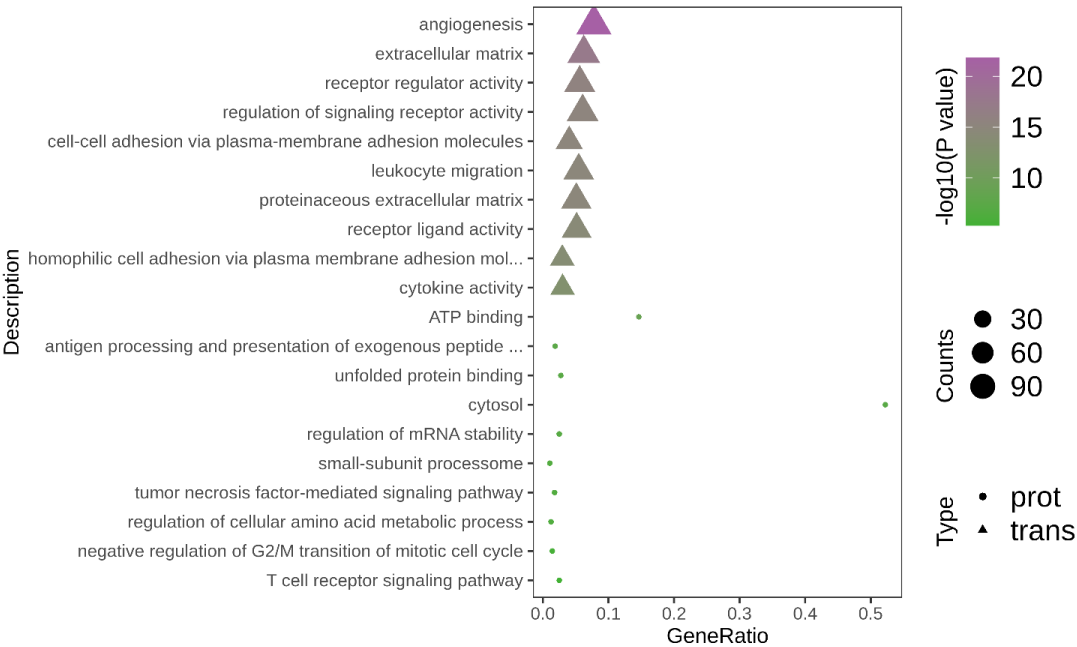


图 10 GO 功能气泡图

究竟哪些 GO 功能富集到了更多数量的差异分子（mRNA 或蛋白）呢？我们对各 GO 功能富集到的差异 mRNA 数目和差异蛋白数目进行加和统计，并优先选择总差异分子数（差异 mRNA 数+差异蛋白数）最多的前 10 条 GO 功能进行展示，以便帮助研究者了解本项目差异分子集中注释到哪些 GO 功能上了。当 GO 功能数不足 10 条时，则有多少条 GO 功能展示多少条。

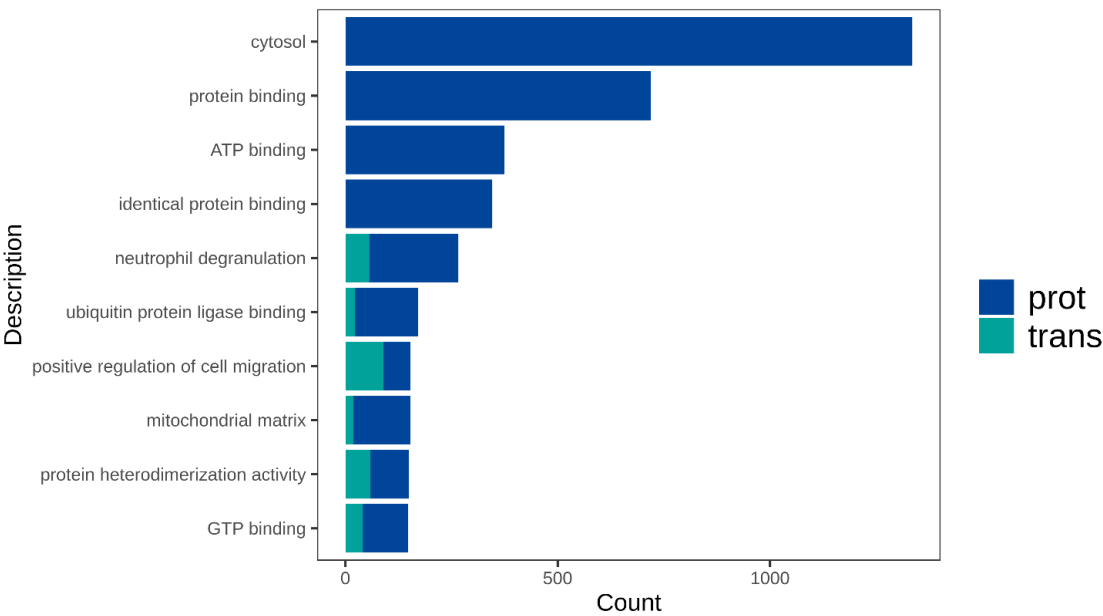


图 11 TOP10 GO 功能柱状图

两组学共有 GO 功能通常也是为研究者所关注。我们以柱状图对共有 GO 功能进行直观

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

展示。默认在共有 GO 功能中，优先展示 TOP10（共有通路分别在两组学中按照富集显著性 p 值升序排列，并将在两组学中排序位置进行加和，加和最小的前 10 作为 TOP10（即按照 p_value_rank 挑出共有富集 GO 功能）；柱子展示差异 mRNA 和差异蛋白 count 数的累积加和）共有富集 GO 功能。

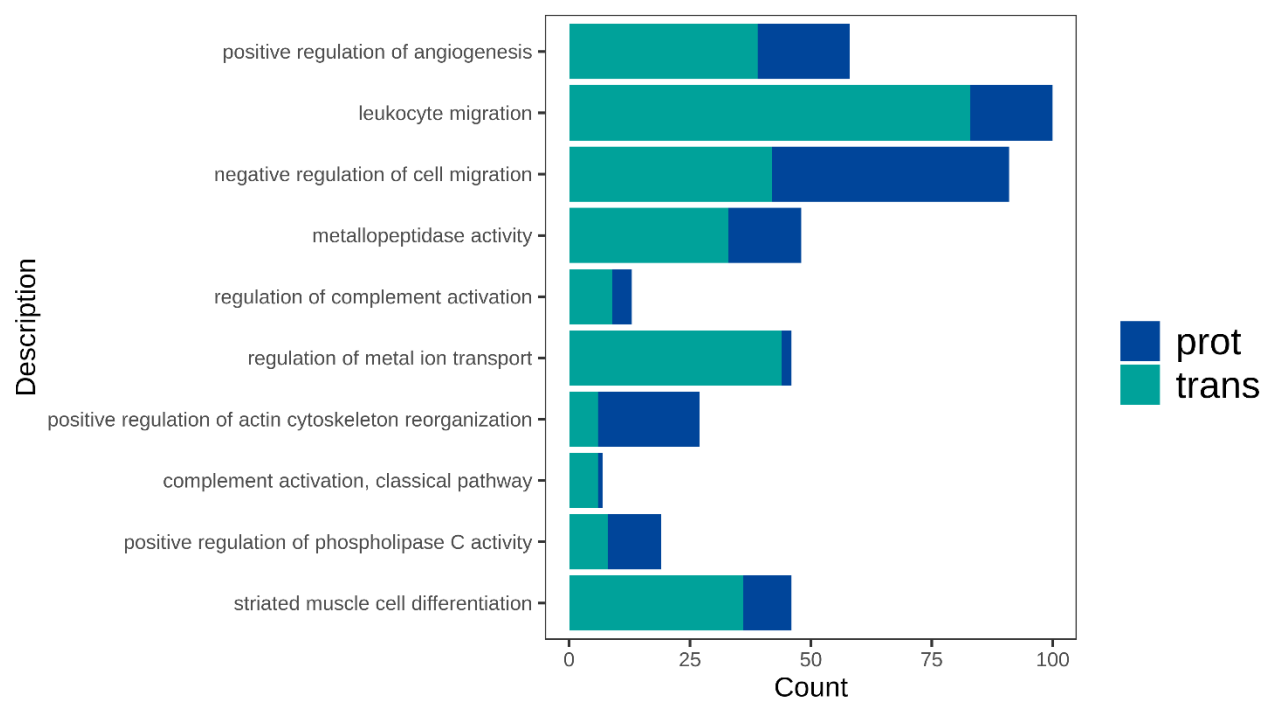


图 12 共有富集 GO 功能柱状图

3.3.2 关键 KEGG 通路富集

KEGG（Kyoto Encyclopedia of Genes and Genomes，<http://www.kegg.jp/>）数据库收录了新陈代谢、遗传信息加工、环境信息加工、细胞过程、生物体系统、人类疾病以及药物开发等多个方面的通路信息，将基因组信息和代谢物功能信息有机地结合了起来。通过对转录组、蛋白组分析得到的差异分子进行 KEGG 通路整合分析，从信息流层面来探究生物学问题，相互验证，有助于从海量的数据中确定关键变化的信号通路，构建机体调控网络。

对两组学差异分子分别富集到的显著 KEGG 通路（转录组学：p_value<0.05；蛋白组学：p_value<0.05）进行 Venn 分析。

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

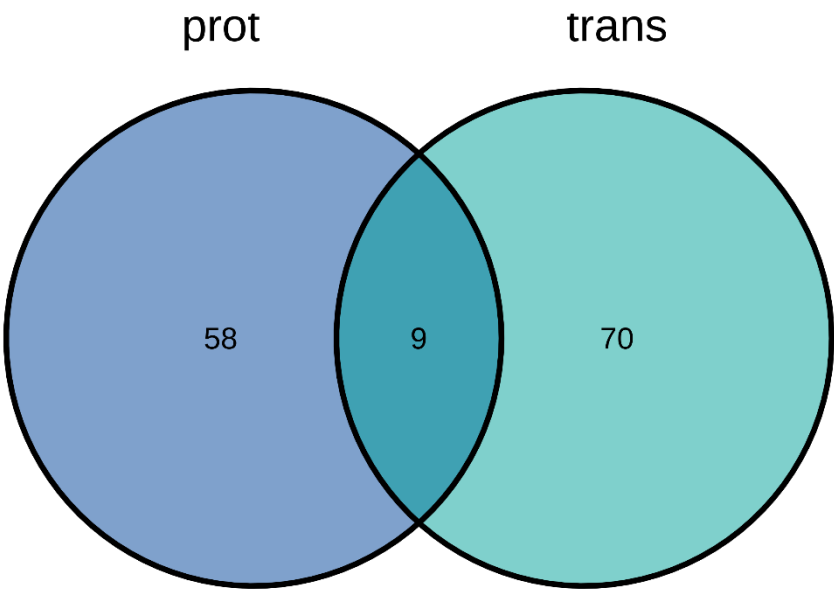


图 13 KEGG 通路富集 Venn 图

利用并集两组学 KEGG 通路富集结果 ($p_value < 0.05$) 的 p_value 值进行聚类热图分析。

结果展示如下：

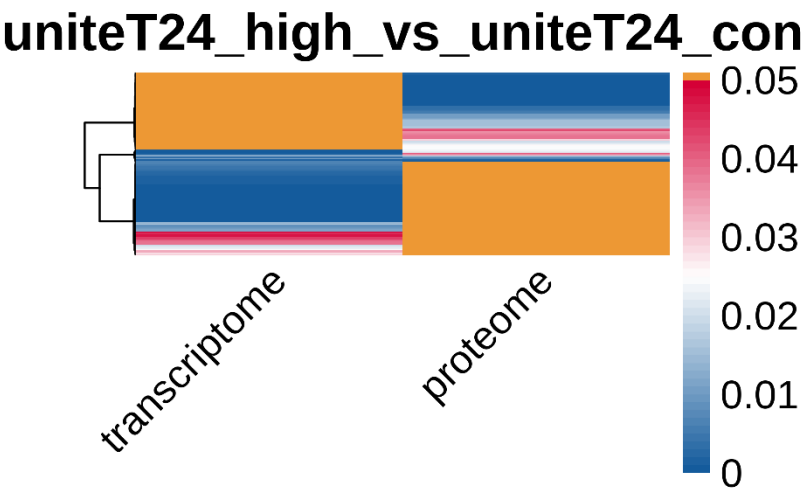


图 14 KEGG 通路热图

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

气泡图不仅直观展示了各组学显著性富集到的 KEGG 通路，同时也直观反映出两组学 KEGG 通路富集结果的共性和特异性。

基于两组学差异分子富集到的 KEGG 通路，我们以富集显著性 p_value 值从小到大进行排序，选择各组学 p_value 值最小的前 10 条通路，即合计至多 20 条通路以气泡图形式呈现如下。

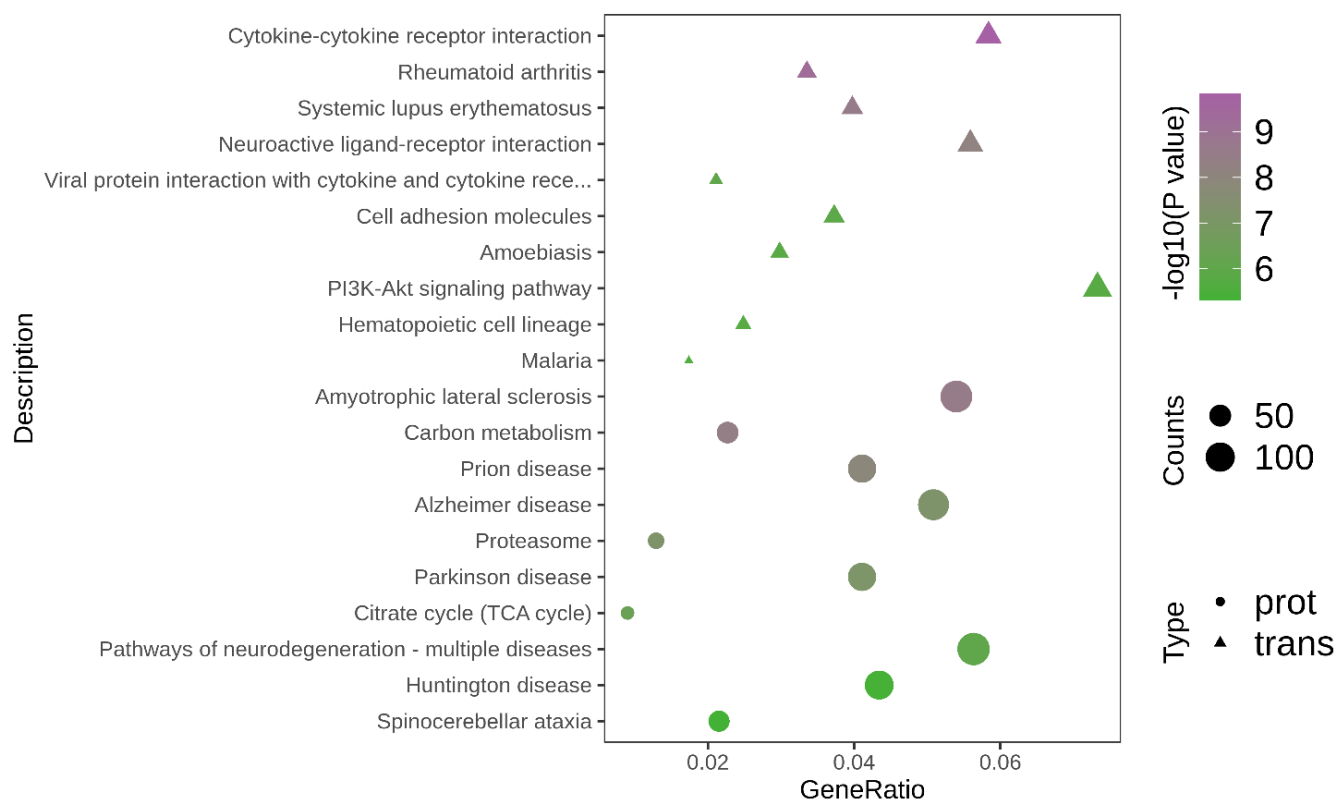


图 15 KEGG 通路气泡图

究竟哪些 KEGG 通路富集到了更多数量的差异分子（mRNA 或蛋白）呢？我们对各 KEGG 通路富集到的差异 mRNA 数目和差异蛋白数目进行加和统计，并优先选择总差异分子数（差异 mRNA 数+差异蛋白数）最多的前 10 条 KEGG 通路进行展示，以便帮助研究者了解本项目差异分子集中注释到哪些 KEGG 通路上了。当 KEGG 通路数不足 10 条时，则有多少条 KEGG 通路展示多少条。结果如下图所示：

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

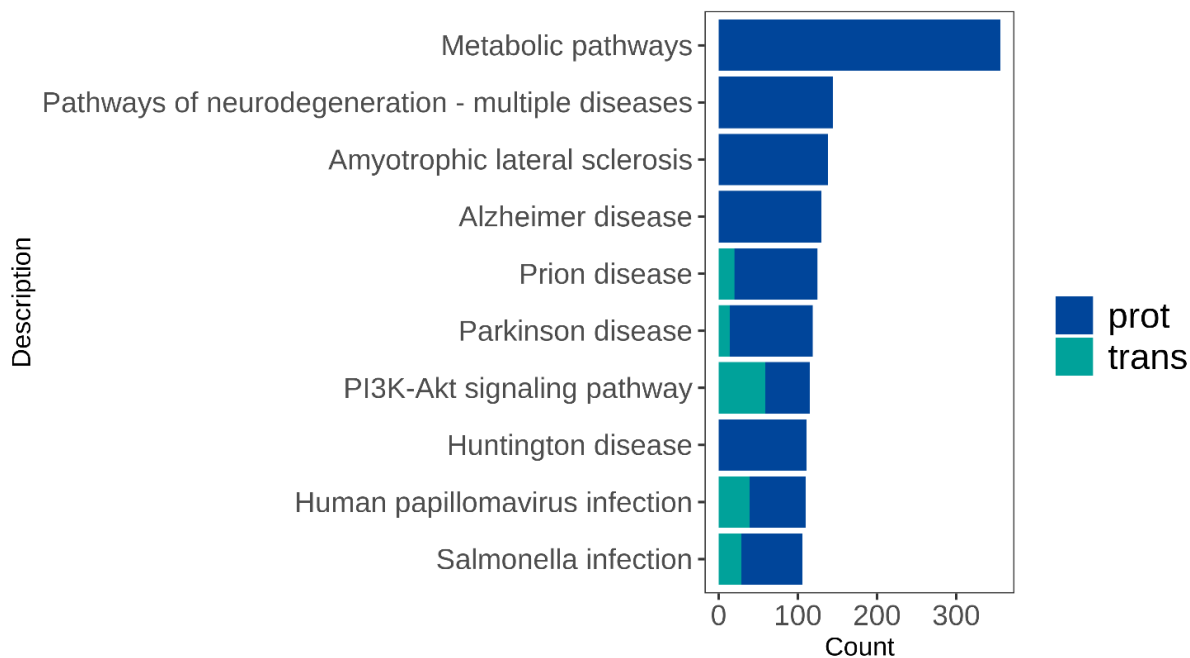


图 16 TOP10 KEGG 通路柱状图

两组学共有 KEGG 通路通常也是为研究者所关注。我们以柱状图对共有 KEGG 通路进行直观展示。默认在共有 KEGG 通路中，优先展示 TOP10（共有通路分别在两组学中按照富集显著性 p 值升序排列，并将在两组学中排序位置进行加和，加和最小的前 10 作为 TOP10（即按照 p_value_rank 挑出共有富集 KEGG 通路）；柱子展示差异 mRNA 和差异蛋白 count 数的累积加和）共有富集 KEGG 通路。

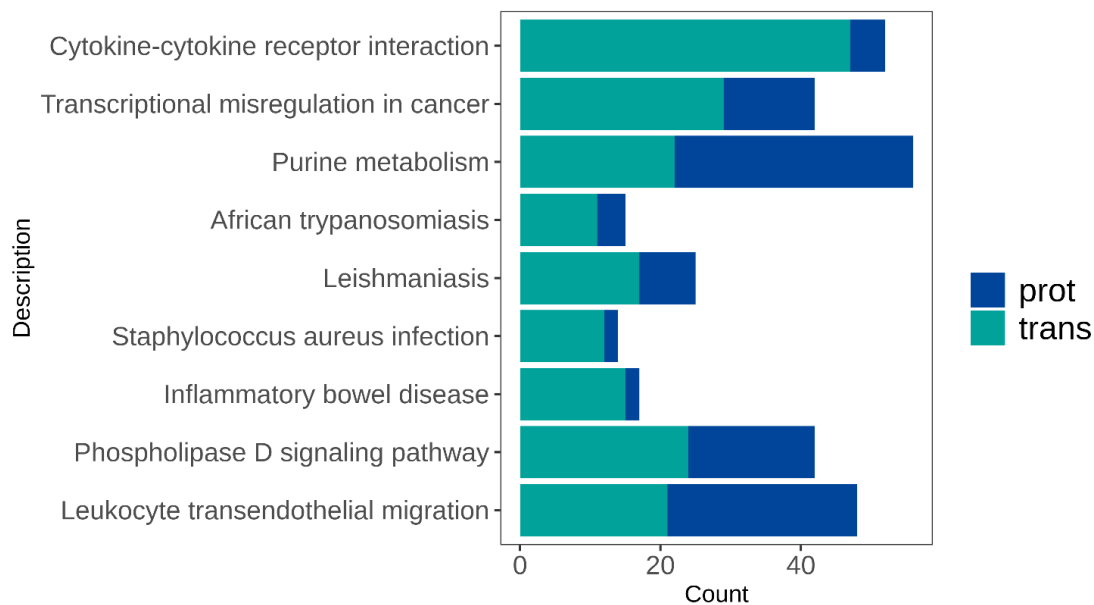


图 17 共有富集 KEGG 通路柱状图

3.3.3 关键基因分析

在转录组和蛋白组中具有共同变化趋势的基因往往参与调控相同的信号通路，而在不同组学数据中共有的信号通路往往扮演着重要的角色。本项目通过对关键基因的挖掘分析，以发现关键的 mRNA 或蛋白，探究更深入的生物学过程。

对两组学的差异分子进行 Venn 分析，鉴定有着相同或相反变化趋势的差异分子。

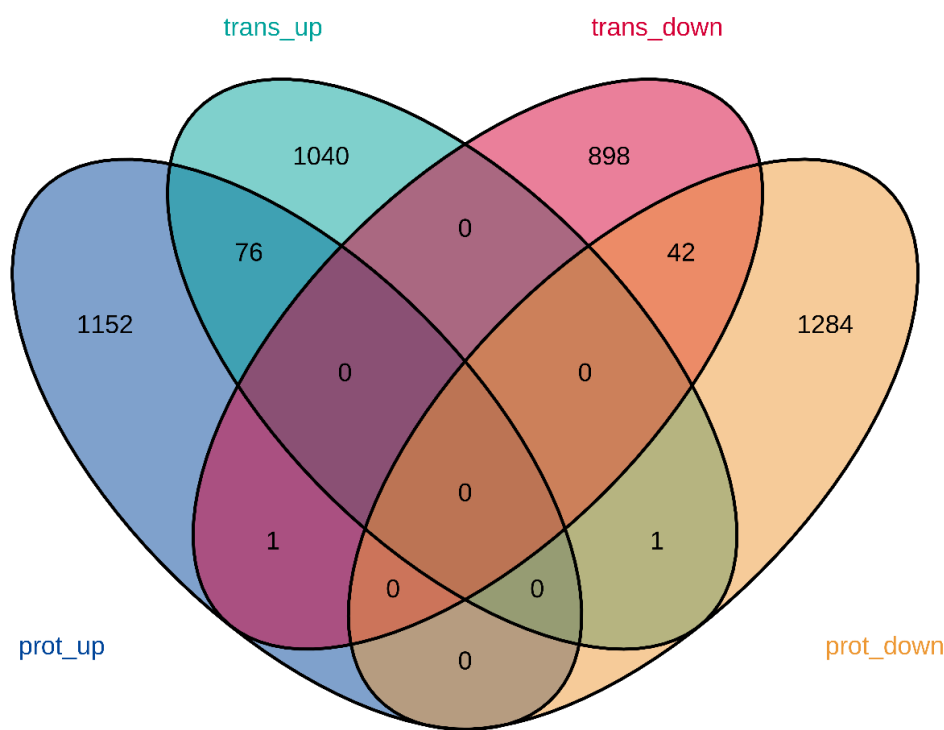


图 18 差异基因变化趋势 Venn 图

同时，我们提供了四象限图来展示两组学的共有差异分子。转录组学和蛋白组学均表达上调的基因，简称 UP-UP；转录组学表达上调，蛋白组学表达下调的基因，简称 UP-DOWN；转录组学表达下调，蛋白组学表达上调的基因，简称 DOWN-UP；转录组学和蛋白组学均表达下调的基因，简称 DOWN-DOWN。四象限图如下图所示：

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

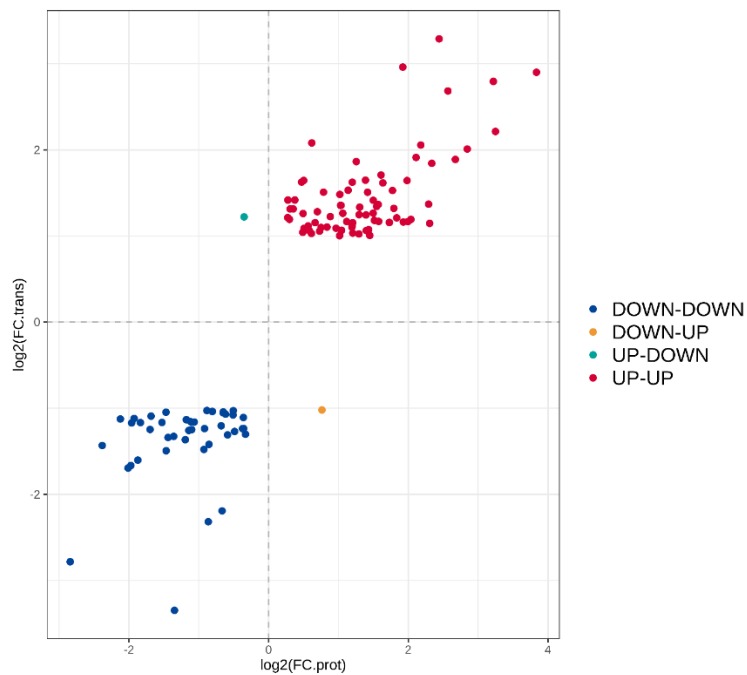


图 19 共有差异基因四象限图

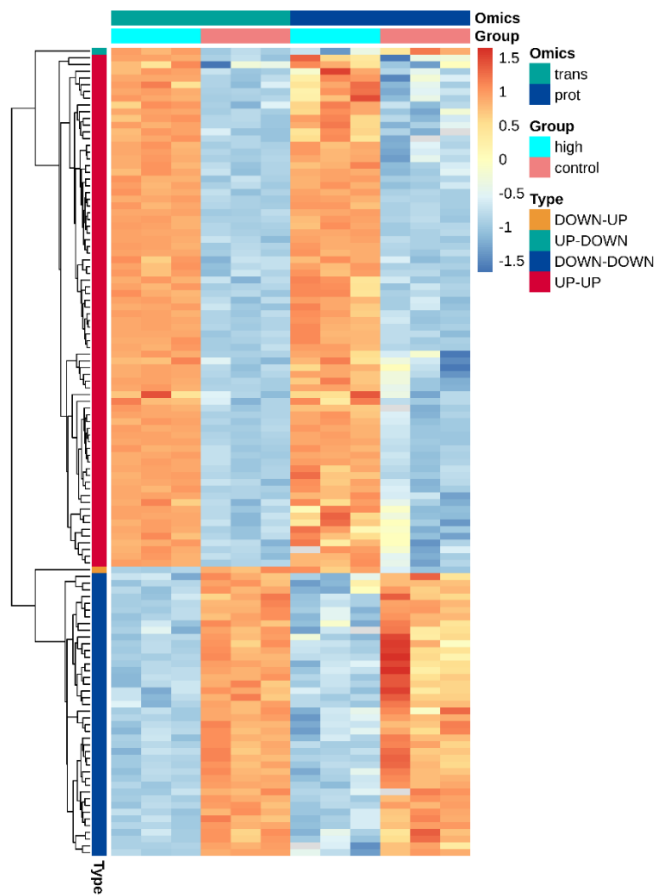


图 20 聚类分析热图

对上述的 Up-Up、Down-Down、Up-Down 和 Down-Up 4 类基因分别进行 GO 功能富集分析，推测其可能参与的生物学功能。结果以柱状图 and 气泡图分别展示如下：

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

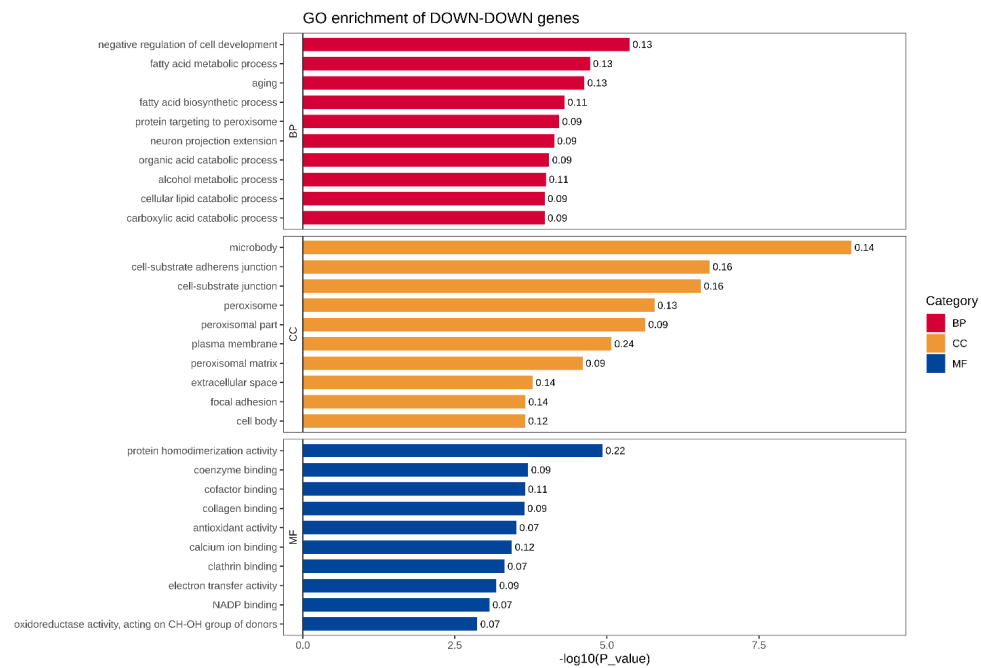


图 21 GO 功能富集柱状图

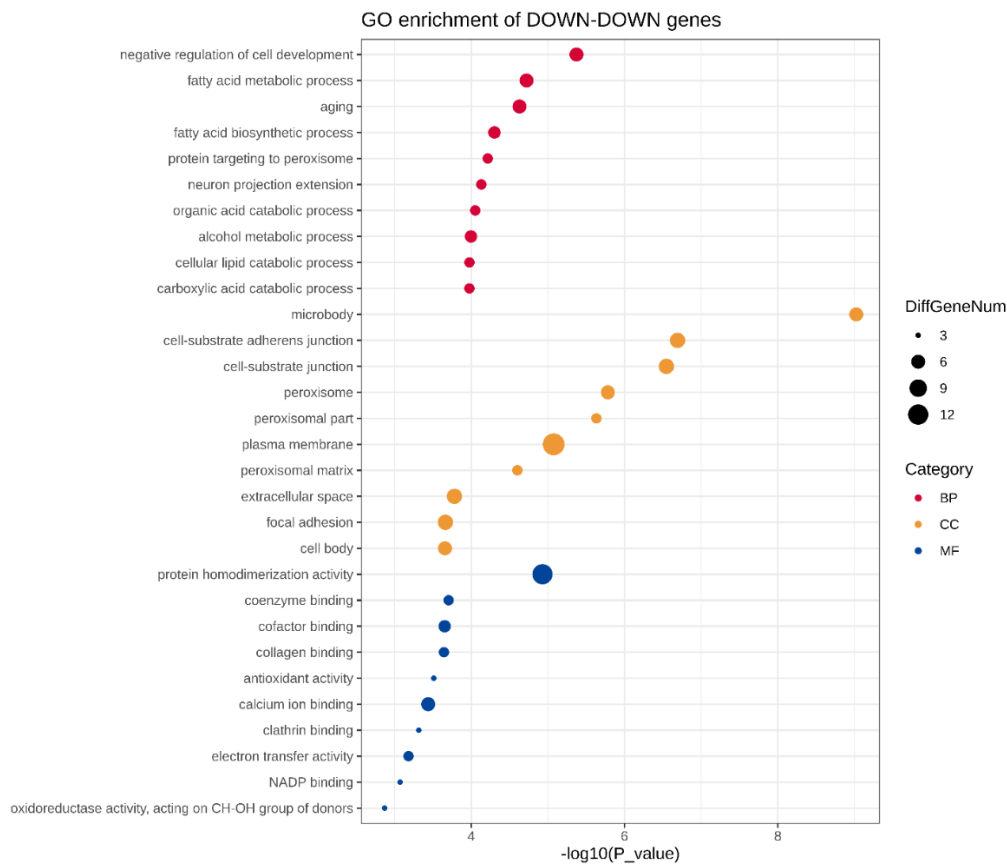


图 22 GO 功能富集气泡图

对上述的 Up-Up、Down-Down、Up-Down 和 Down-Up 4 类基因分别进行 KEGG 通路富集分析，推测其可能参与的生物学通路。结果以柱状图和气泡图分别展示如下：

2025 年国际定向进化大赛实验记录

实验时间： 2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

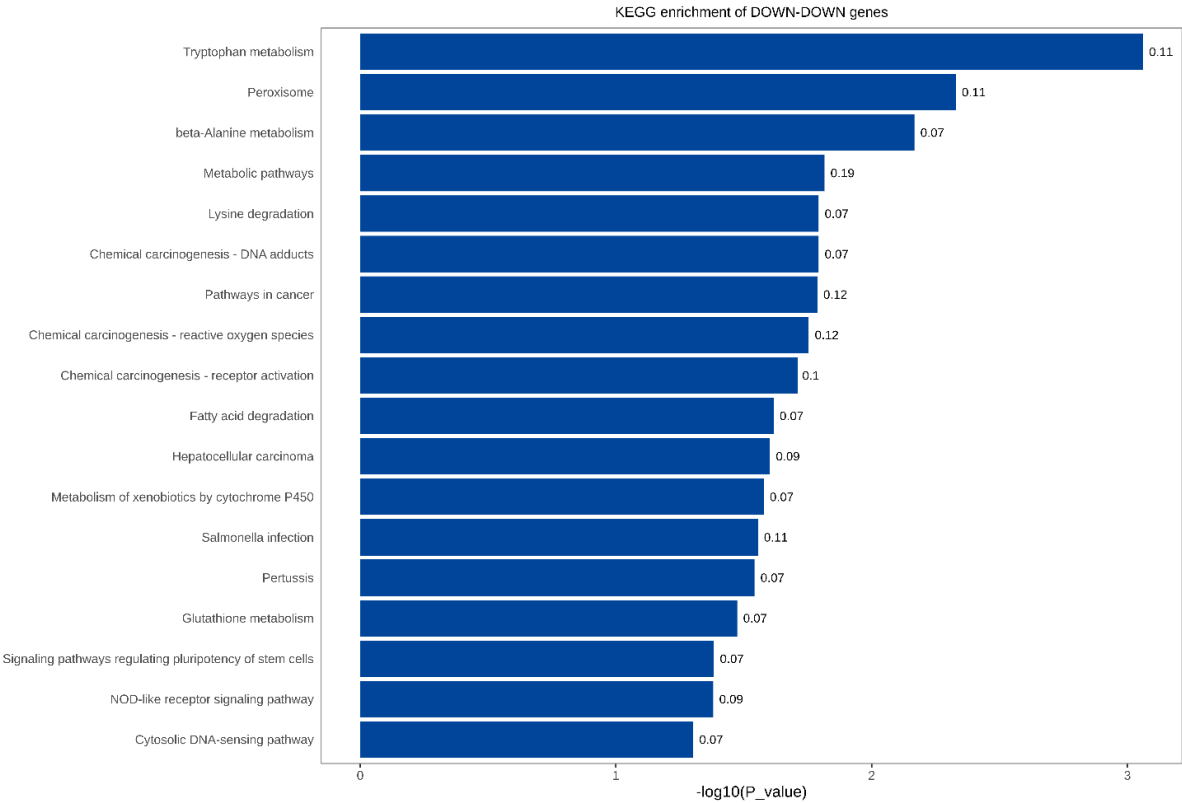


图 23 KEGG 通路富集柱状图

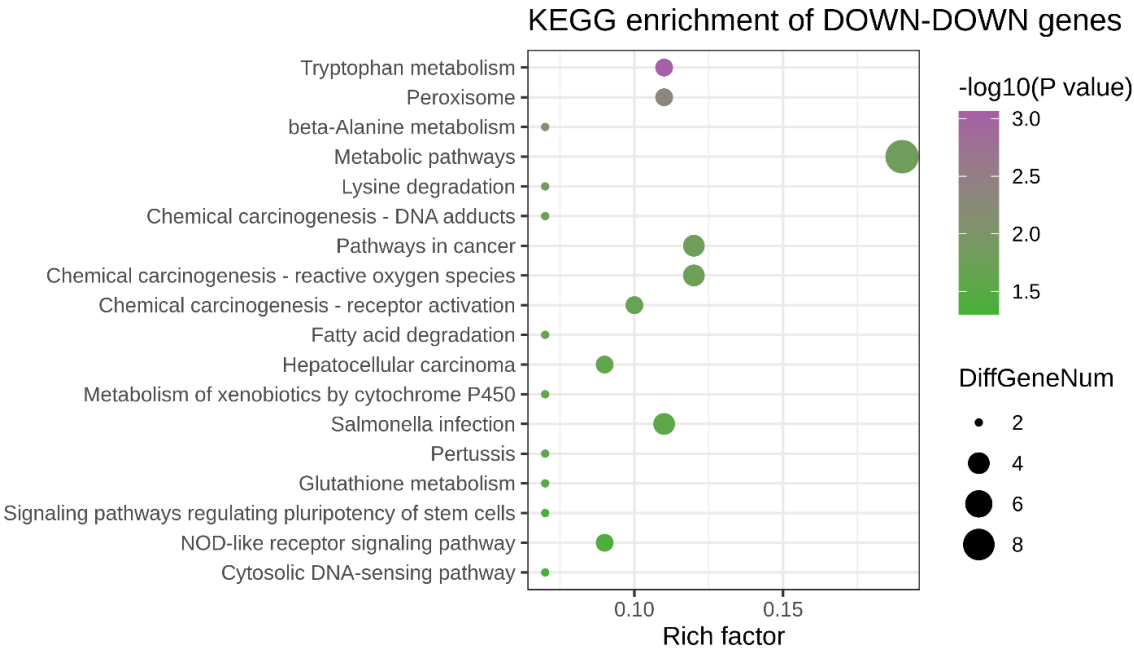


图 24 KEGG 通路富集气泡图

四、结果分析

本研究通过整合转录组与蛋白质组数据，系统解析了人膀胱移行细胞癌 T24-RC48 耐药模型的分子调控特征。结果显示，两组学数据在基因表达层面存在显著差异，且转录与蛋白

2025 年国际定向进化大赛实验记录

实验时间：2025 年 6 月 17 日 18:00 — 2025 年 6 月 19 日 21:00

水平的变化趋势呈现一定非同步性，提示耐药机制可能涉及转录后调控及蛋白翻译修饰等复杂过程。功能富集分析进一步揭示了耐药相关分子在关键生物学过程（如细胞迁移调控、代谢重编程）及信号通路（如 PI3K-Akt、细胞因子互作）中的协同作用，表明耐药表型的形成是多维度分子事件动态调控的结果。上述发现为深入解析 RC48 耐药机制提供了重要线索，其多组学联合分析策略也为后续耐药机制研究奠定了数据基础，具有潜在的科学价值与应用前景。