**生信分析报告**

**项目标题： 多酚五种成分与高脂血症网络药理分析 ;**

**单 号： BSXG250204 ;**

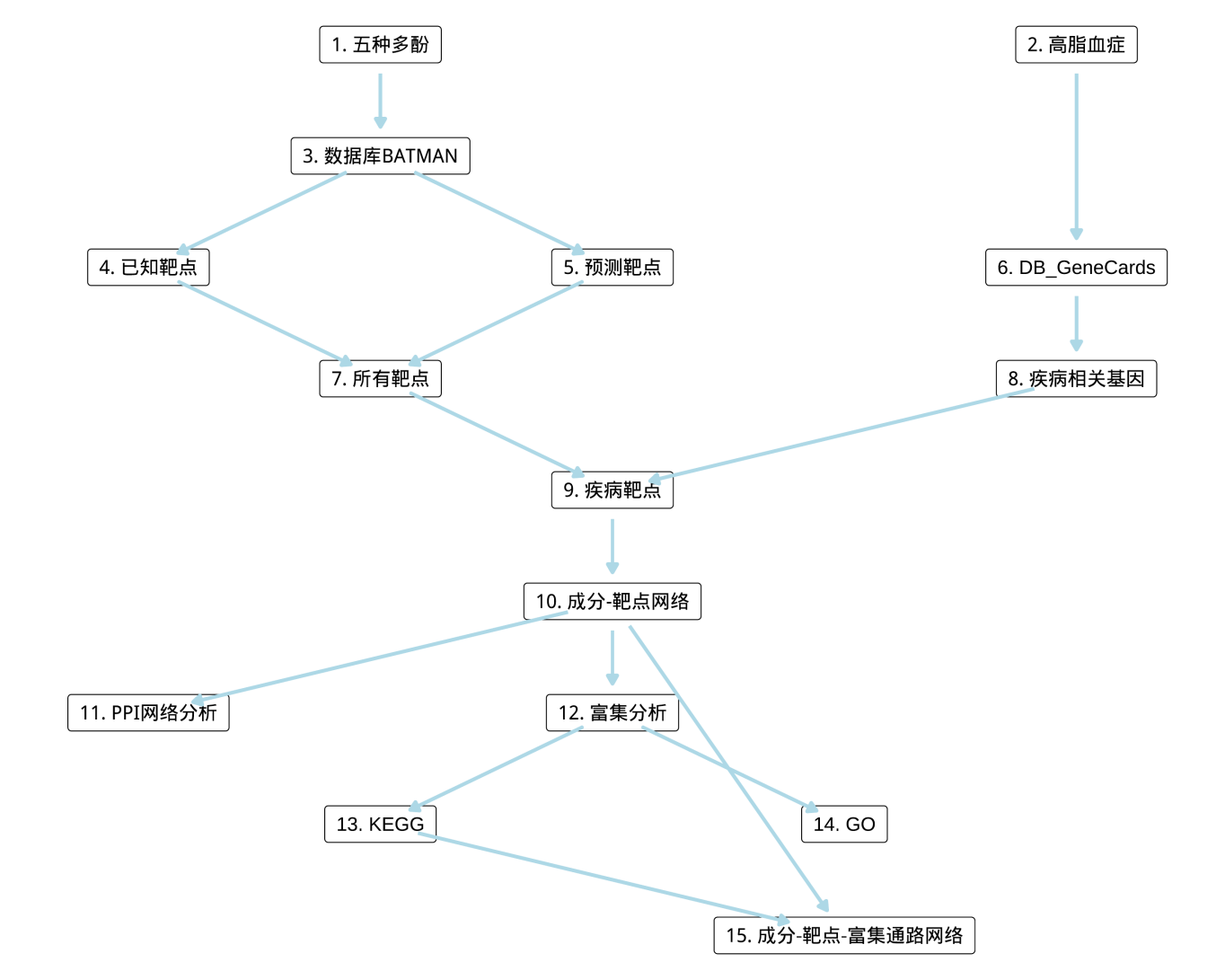
**分析人员： 黄礼闯 ;**

**分析类型： 生信分析 ;**

**委 托 人： 公司内部 ;**

**受 托 人： 杭州铂赛生物科技有限公司 .**

# 1 分析流程



**Fig.** **1** Route

**(File path: Figure+Table/1.0\_分析流程\_{#abstract}/Route.pdf)**

# 2 材料和方法

## 2.1 数据分析平台

在 Linux pop-os x86\_64 (6.9.3-76060903-generic) 上，使用 R version 4.4.2 (2024-10-31) (<https://www.r-project.org/>) 对数据统计分析与整合分析。

## 2.2 BATMAN 网络药理学 (Dataset: PHENOL)

从数据库 BATMAN-TCM ((2024, **IF:16.6**, Q1, Nucleic acids research)1) 中获取 Pseudo\_herb 等中药的成分、靶点数据。(即中药：Pseudo\_herb)。 使用 BATMAN-TCM 数据库中的 known\_target\_proteins 作为成分靶点。使用 BATMAN-TCM 数据库中的 known\_target\_proteins 作为成分靶点。此外，还使用了 BATMAN-TCM 数据库中的 predicted\_target\_proteins 作为成分靶点，并设定 分数 cut-off 为 0.5。合并靶点数据。以 BiomaRt ((2009, **IF:13.1**, Q1, Nature protocols)2) 对靶点信息的 entrez\_id 转化为基因 Symbol (hgnc\_symbol) 。 以 PubChemR 获取化合物同义名 (Synonym)，按正则表达式 (Regex) 匹配化合物简短的同义名用以化合物注释。

## 2.3 GeneCards 基因获取 (Dataset: HL)

从 GeneCards 数据库 (2016, Current protocols in bioinformatics)3 获取 Hyperlipidemia 相关的基因集，得分 cut-off 为 1。

## 2.4 STRINGdb PPI 分析 (Dataset: PHENOL)

以 R 包 STEINGdb (2.18.0) (2021, **IF:16.6**, Q1, Nucleic Acids Research)4 构建 PPI 网络。数据版本为 12.0，互作类型为 physical。以 Cytohubba (2014, BMC Systems Biology)5 的算法计算 MCC score (在 R 中计算) 。随后，以 ggraph 可视化网络 (2.2.1)。

## 2.5 ClusterProfiler 富集分析 (Dataset: PHENOL)

以 ClusterProfiler R 包 (4.15.0.2) (2021, **IF:33.2**, Q1, The Innovation)6进行 KEGG 和 GO 富集分析。以 p.adjust 表示显著水平。

# 3 分析结果

## 3.1 BATMAN 网络药理学 (PHENOL)

从数据库 BATMAN-TCM 中药的成分、靶点数据 (详见方法章节) 。各中药的化合物组成统计 (可能有交叉涵盖)：Pseudo\_herb (n=5) 。 共 5 个化合物 (注：根据唯一 PubChem CID 统计)， 其中，含有靶点信息记录的化合物共 5 个 (非重复)。共包含靶点 366 个 (非重复)。 Tab. **[1](#Herbs-compounds-and-targets)** 为所有中药、化合物以及对应靶点汇总表格。

**Tab.** **1** Herbs compounds and targets

| Ingredient.id | Herb pinyin name | Ingredient.name | Target.name |
| --- | --- | --- | --- |
| 5280000 | Pseudo herb | Quercetin | LPCAT1 |
| 5280000 | Pseudo herb | Quercetin | LPCAT2 |
| 5280000 | Pseudo herb | Quercetin | MMP2 |
| 5280000 | Pseudo herb | Quercetin | TP53 |
| 5280000 | Pseudo herb | Quercetin | MMP1 |
| ... | ... | ... | ... |

**(File path: Figure+Table/3.1\_BATMAN\_网络药理学\_(PHENOL)/Herbs-compounds-and-targets.csv)**

## 3.2 GeneCards 基因获取 (HL)

从 GeneCards 搜索 Hyperlipidemia, 获取对应靶点数据，统计为 Functional Element (n=12) , Genetic Locus (n=4) , Protein Coding (n=742) , Pseudogene (n=2) , RNA Gene (lncRNA) (n=12) , RNA Gene (miRNA) (n=23) , RNA Gene (ncRNA) (n=1) , RNA Gene (snoRNA) (n=1) , RNA Gene (tRNA) (n=2) 。共 799 个靶点。 Tab. **[2](#HL-disease-related-targets-from-GeneCards)**

**Tab.** **2** HL disease related targets from GeneCards

| Symbol | Description | Category | UniProt ID | GIFtS | GC id | Score |
| --- | --- | --- | --- | --- | --- | --- |
| APOB | Apolipopro... | Protein Co... | P04114 | 59 | GC02M020956 | 188.9 |
| LPL | Lipoprotei... | Protein Co... | P06858 | 63 | GC08P019901 | 61.71 |
| LIPC | Lipase C, ... | Protein Co... | P11150 | 59 | GC15P058410 | 53.88 |
| LOC106560211 | APOB 5' Re... | Functional... |  | 9 | GC02P021041 | 46.12 |
| LDLR | Low Densit... | Protein Co... | P01130 | 64 | GC19P142248 | 39.86 |
| ... | ... | ... | ... | ... | ... | ... |

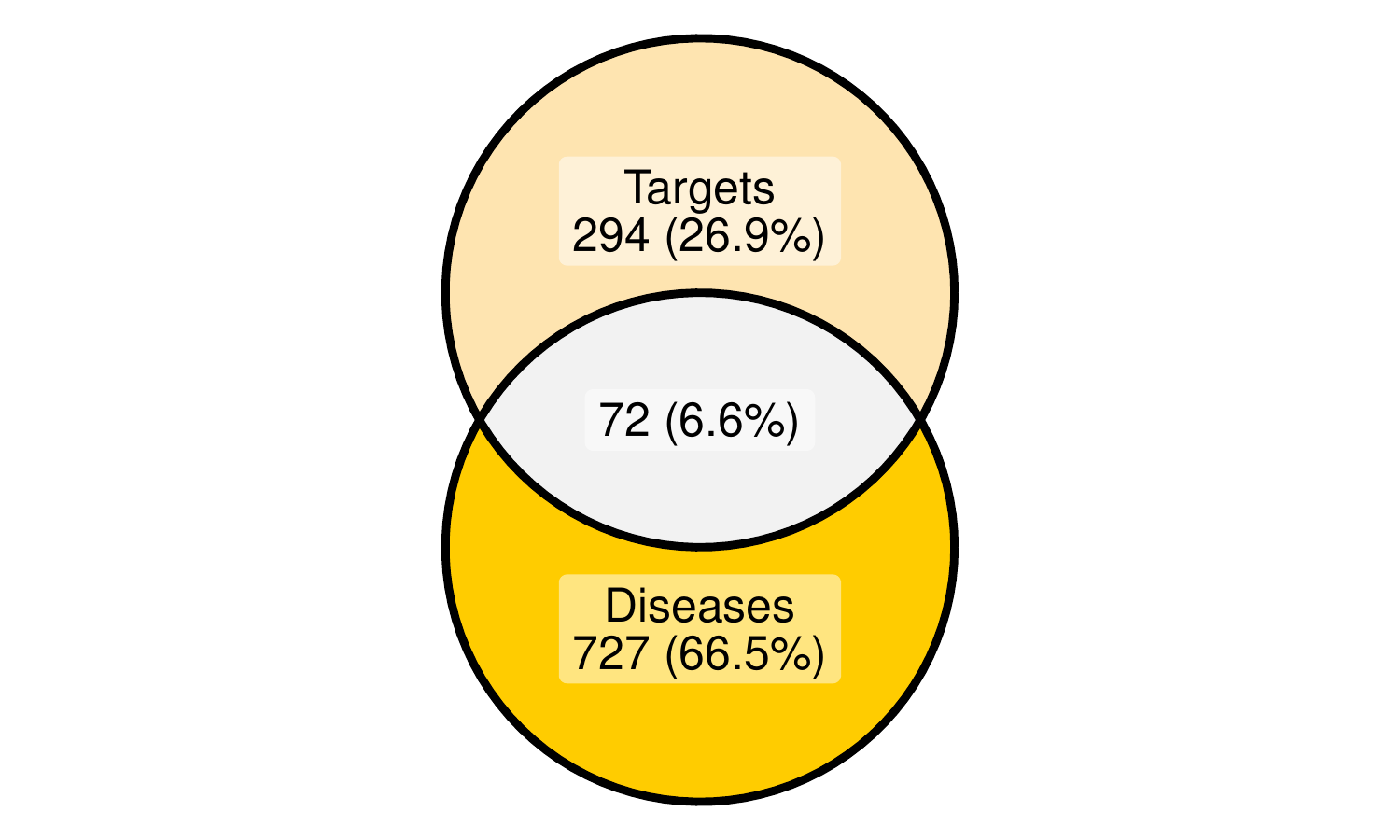
**(File path: Figure+Table/3.2\_GeneCards\_基因获取\_(HL)/HL-disease-related-targets-from-GeneCards.xlsx)**

* The GeneCards data was obtained by querying: Hyperlipidemia
* Restrict (with quotes): TRUE
* Filtering by Score:: Score > 1

## 3.3 Network 中药-成分-疾病-靶点网络 (PHENOL)

将 疾病 的靶点与中药靶点取交集，随后过滤中药成分与靶点数据，形成中药-成分-疾病-靶点网络。

Fig. **[2](#PHENOL-Targets-intersect-with-related-targets-dis)** 展示了中药的靶点与疾病靶点基因集 (来自于GeneCards 基因获取[Section: HL]) 的交集数目。 两者共含有 72 个交集靶点。 Fig. **[3](#PHENOL-network-pharmacology-with-filtered-type-dis)** 展示了中药、成分、靶点 (中药靶点与疾病靶点的交集) 的网络图。 该图对中心度 (centrality\_degree) 较高的节点 (成分或靶点) 做了名称标注。 图中的图例标注了节点的所属类型：中药、化合物、靶点。 化合物可分为该中药唯一所含的化合物，或者与其他中药共有的化合物。 共有的化合物环绕在靶点周围，而唯一的化合物则环绕在中药周围。 Tab. **[3](#PHENOL-network-pharmacology-with-filtered-type-original-data-dis)** 为用于绘制中药-成分-疾病-靶点网络的数据集。

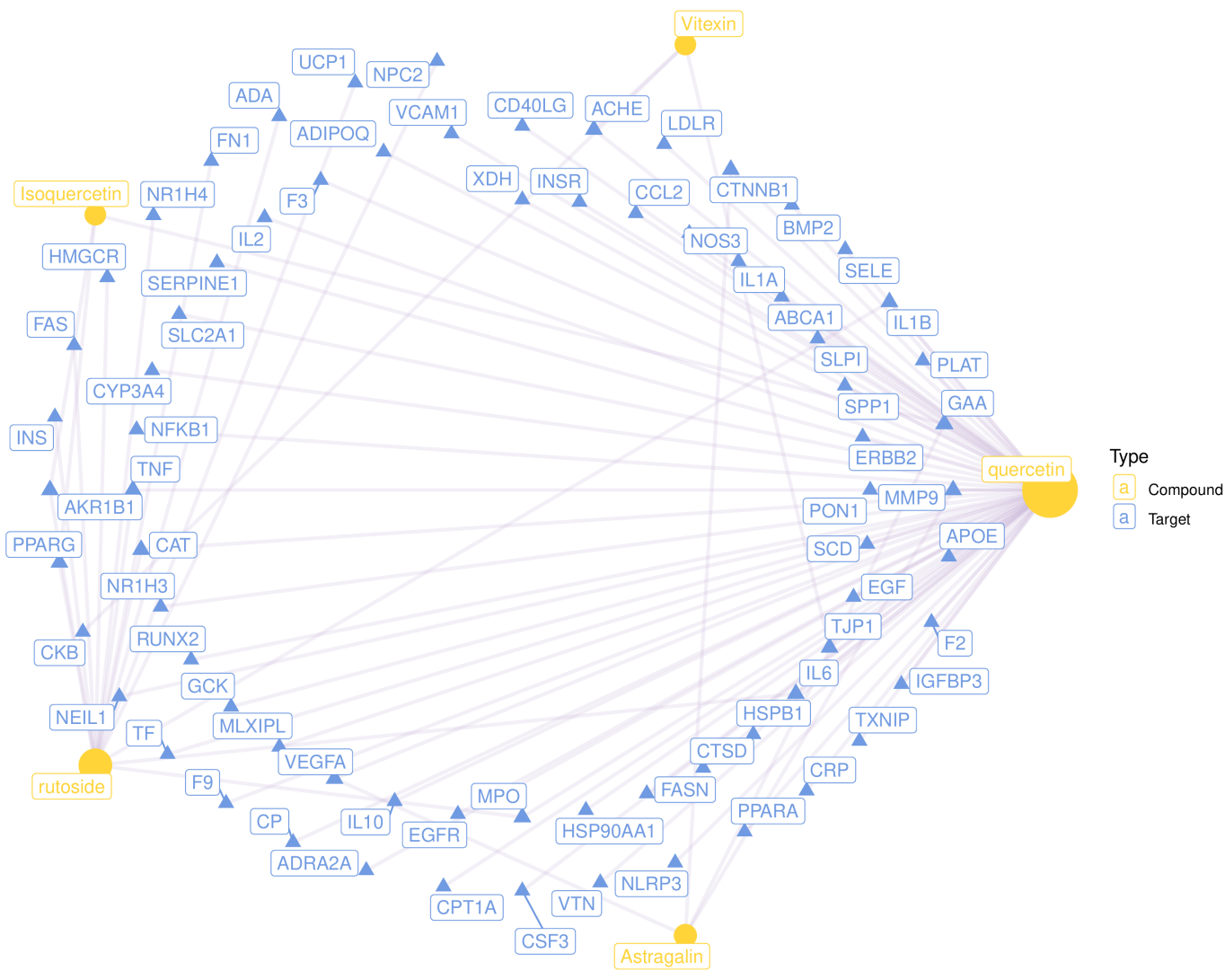


**Fig.** **2** PHENOL Targets intersect with related targets dis

**(File path: Figure+Table/3.3\_Network\_中药-成分-疾病-靶点网络\_(PHENOL)/PHENOL-Targets-intersect-with-related-targets-dis.pdf)**

* All\_intersection: LDLR, APOE, HMGCR, ABCA1, INS, PPARA, ADIPOQ, PON1, GCK, PPARG, TXNIP, CRP, GAA, SERPINE1, INSR, ADRA2A, FN1, VEGFA, SLC2A1, NOS3, NR1H4, CYP3A4, TNF, VTN, CCL2, CP, SCD, NR1H3, IL6, FAS, TF, CSF3, EGFR, SELE, CKB, F2, MMP9, RUNX2, NEIL1, UCP1, PLAT, F9, XDH, CAT, CTSD, IGFBP3, IL1A, SPP1, AKR1B1, IL1B, MLXIPL, HSPB1, IL10, NFKB1, IL2, CPT1A, ADA, ACHE, TJP1, EGF, MPO, NLRP3, NPC2, FASN, F3, VCAM1, CD40LG, SLPI, CTNNB1, BMP2, ERBB2, HSP90AA1

**(See: Figure+Table/3.3\_Network\_中药-成分-疾病-靶点网络\_(PHENOL)/PHENOL-Targets-intersect-with-related-targets-dis-content)**



**Fig.** **3** PHENOL network pharmacology with filtered type dis

**(File path: Figure+Table/3.3\_Network\_中药-成分-疾病-靶点网络\_(PHENOL)/PHENOL-network-pharmacology-with-filtered-type-dis.pdf)**

**Tab.** **3** PHENOL network pharmacology with filtered type original data dis

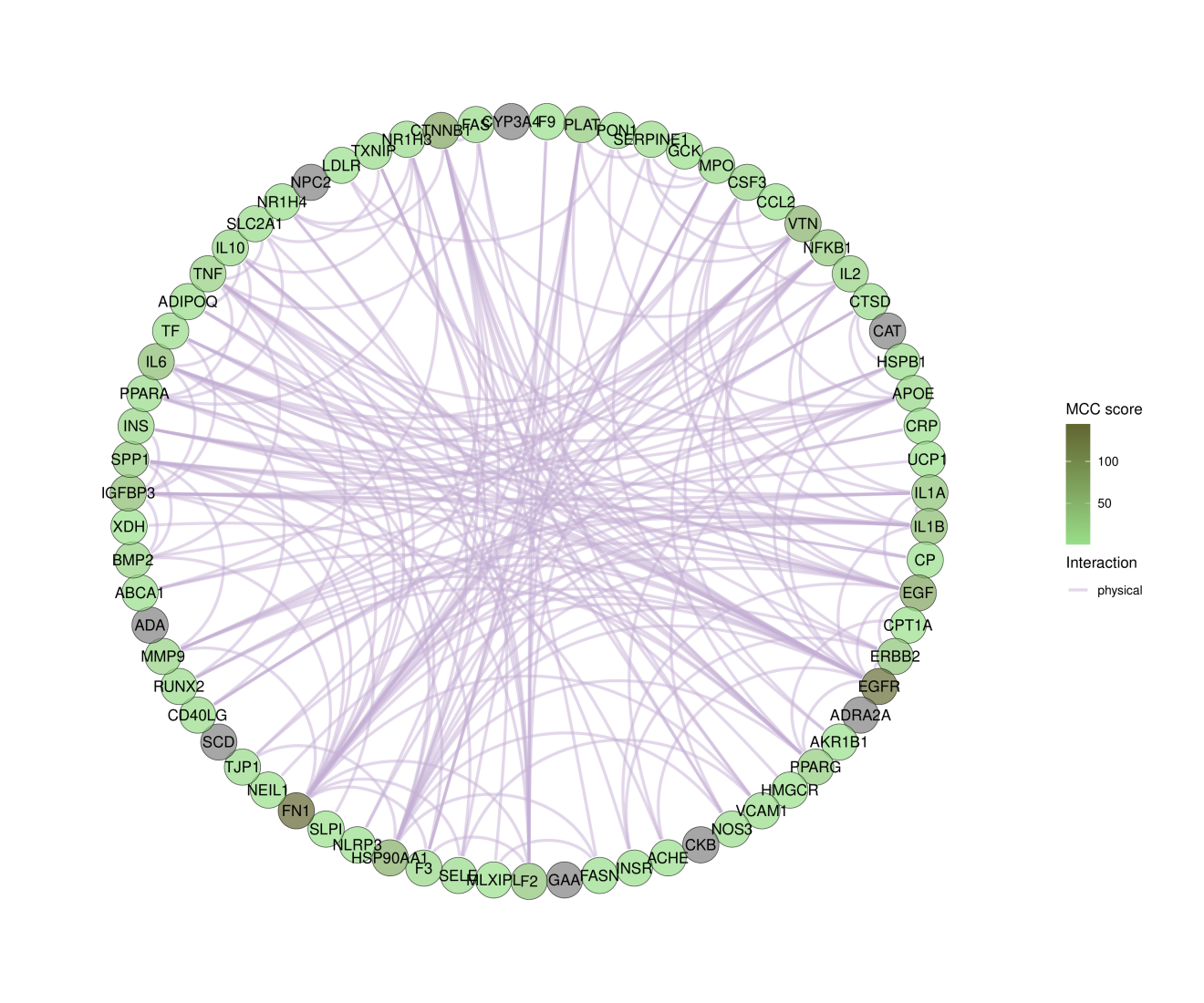
| Herb pinyin name | Ingredient.name | Target.name |
| --- | --- | --- |
| Pseudo herb | Quercetin | XDH |
| Pseudo herb | Quercetin | INSR |
| Pseudo herb | Quercetin | CCL2 |
| Pseudo herb | Quercetin | NOS3 |
| Pseudo herb | Quercetin | IL1A |
| ... | ... | ... |

**(File path: Figure+Table/3.3\_Network\_中药-成分-疾病-靶点网络\_(PHENOL)/PHENOL-network-pharmacology-with-filtered-type-original-data-dis.csv)**

## 3.4 STRINGdb PPI 分析 (PHENOL)

对基因集 (LDLR, APOE, HMGCR, …[n = 72], 来自于中药-成分-疾病-靶点网络[Section: PHENOL]) 进行STRINGdb PPI 分析。

Fig. **[4](#PHENOL-Top-MCC-score)** PPI (带有 Cytohubba (2014, BMC Systems Biology)5 MCC 得分) 网络图 Tab. **[4](#PHENOL-graph-MCC-layout-data)** PPI (带有 Cytohubba (2014, BMC Systems Biology)5 MCC 得分) 附表



**Fig.** **4** PHENOL Top MCC score

**(File path: Figure+Table/3.4\_STRINGdb\_PPI\_分析\_(PHENOL)/PHENOL-Top-MCC-score.pdf)**

**Tab.** **4** PHENOL graph MCC layout data

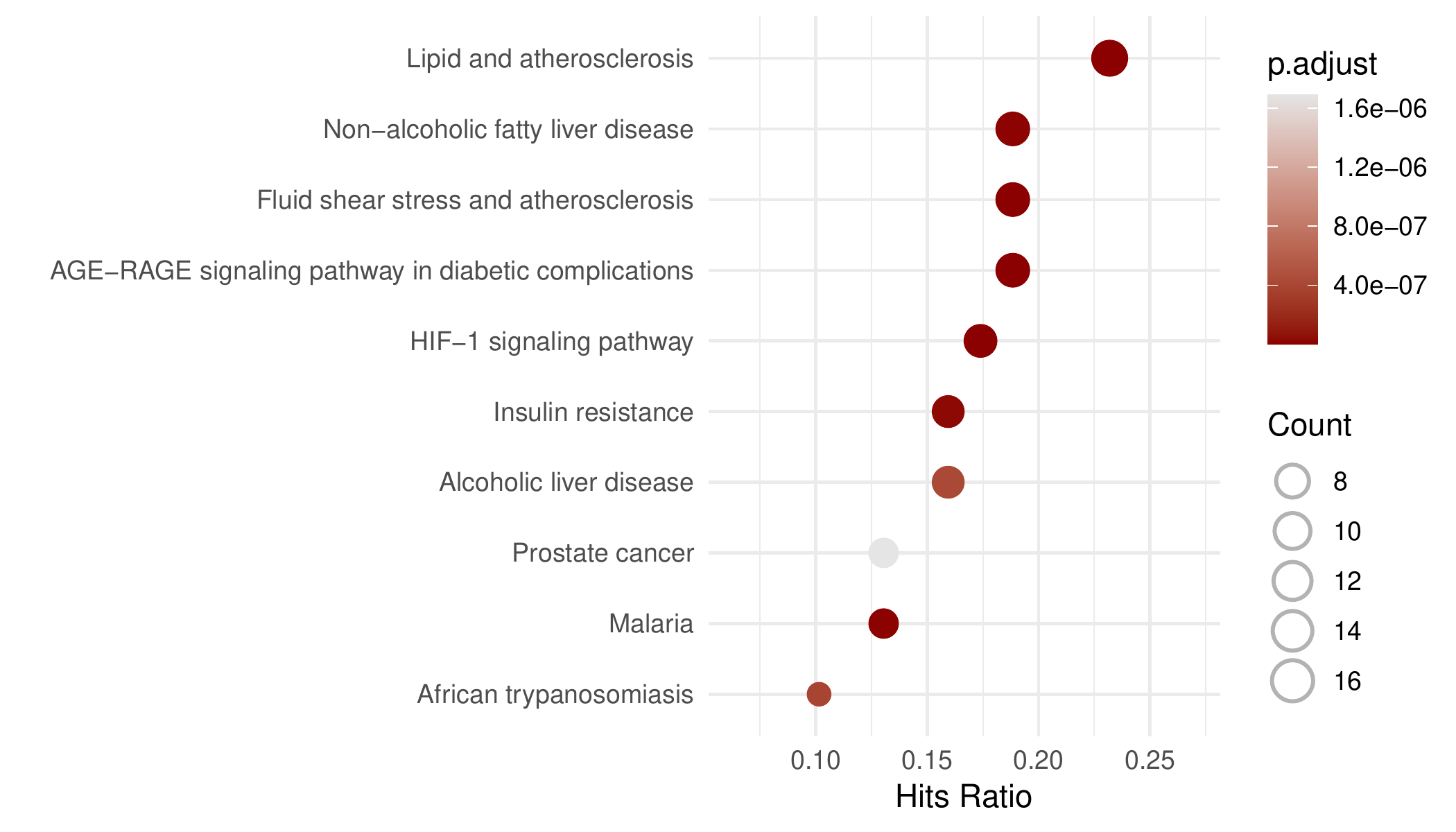
| X | Y | Width | Start | End | R0 | Name | Symbol | MCC score | Centra... |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0.04423 | 0.999 | 1 | 0 | 0.0885 | 1 | 9606.E... | F9 | 2 | 2 |
| 0.1324 | 0.9912 | 1 | 0.0885 | 0.177 | 1 | 9606.E... | PLAT | 28 | 8 |
| 0.2194 | 0.9756 | 1 | 0.177 | 0.2655 | 1 | 9606.E... | PON1 | 3 | 3 |
| 0.3048 | 0.9524 | 1 | 0.2655 | 0.354 | 1 | 9606.E... | SERPINE1 | 8 | 5 |
| 0.3878 | 0.9217 | 1 | 0.354 | 0.4425 | 1 | 9606.E... | GCK | 1 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

**(File path: Figure+Table/3.4\_STRINGdb\_PPI\_分析\_(PHENOL)/PHENOL-graph-MCC-layout-data.csv)**

## 3.5 ClusterProfiler 富集分析 (PHENOL)

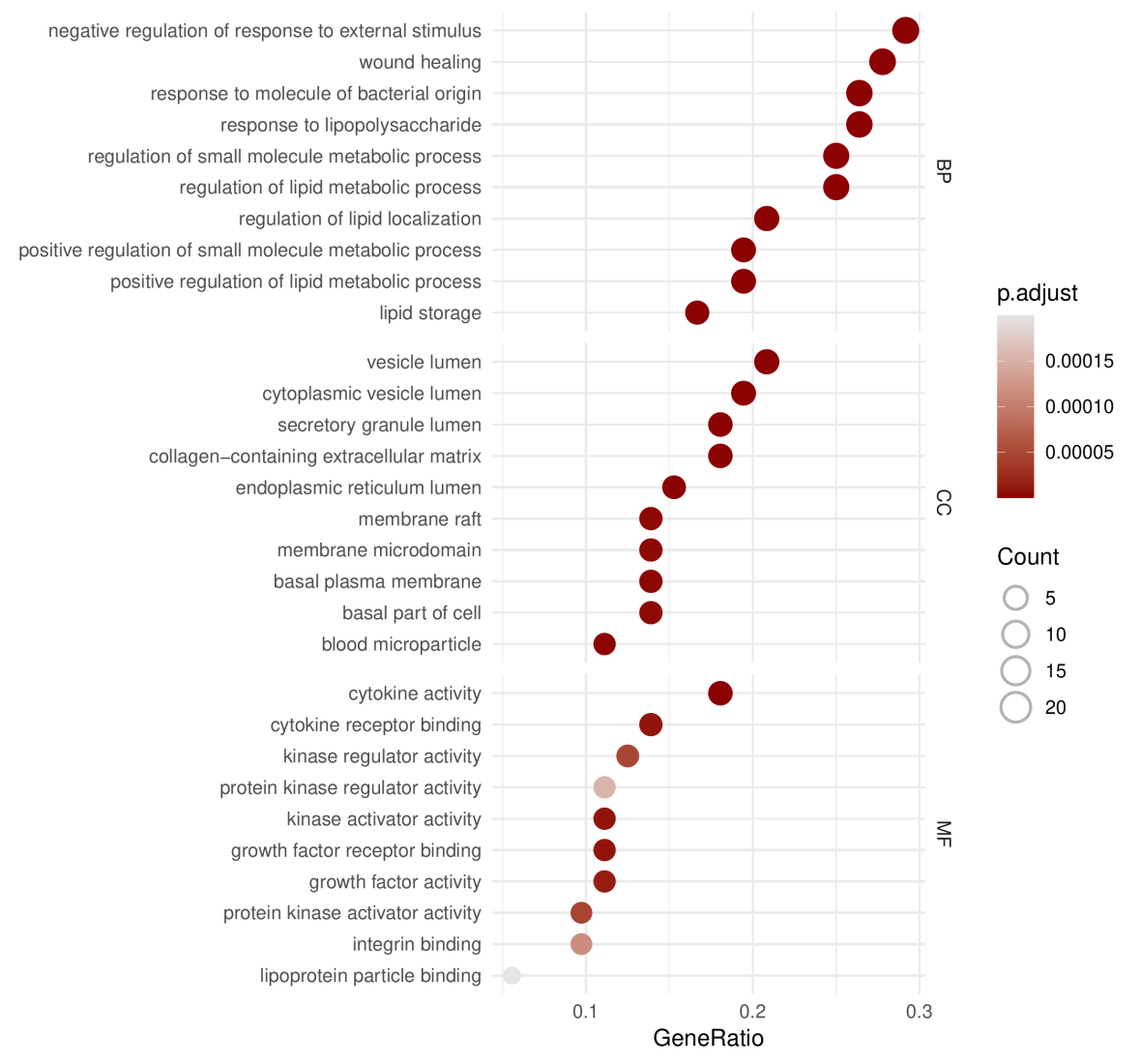
对基因集 (LDLR, APOE, HMGCR, …[n = 72], 来自于中药-成分-疾病-靶点网络[Section: PHENOL]) 进行ClusterProfiler 富集分析。

Fig. **[5](#PHENOL-KEGG-enrichment)** 为 GO 富集分析气泡图。 Fig. **[6](#PHENOL-GO-enrichment)** 为 GO 富集分析气泡图。 Tab. **[5](#PHENOL-KEGG-enrichment-data)** 为 KEGG 富集分析统计表。 Tab. **[6](#PHENOL-GO-enrichment-data)** 为 GO 富集分析统计表。



**Fig.** **5** PHENOL KEGG enrichment

**(File path: Figure+Table/3.5\_ClusterProfiler\_富集分析\_(PHENOL)/PHENOL-KEGG-enrichment.pdf)**



**Fig.** **6** PHENOL GO enrichment

**(File path: Figure+Table/3.5\_ClusterProfiler\_富集分析\_(PHENOL)/PHENOL-GO-enrichment.pdf)**

**Tab.** **5** PHENOL KEGG enrichment data

| Category | Subcat... | ID | Descri... | GeneRatio | BgRatio | Pvalue | P.adjust | Qvalue | GeneID |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Human ... | Endocr... | Hsa04933 | AGE-RA... | 13/69 | 101/8541 | 8.974e-13 | 1.983e-10 | 9.73e-11 | 5054/2... |
| Human ... | Cardio... | Hsa05417 | Lipid ... | 16/69 | 216/8541 | 9.406e-12 | 1.039e-09 | 5.099e-10 | 3949/1... |
| Enviro... | Signal... | Hsa04066 | HIF-1 ... | 12/69 | 110/8541 | 5.404e-11 | 3.753e-09 | 1.841e-09 | 3630/5... |
| Human ... | Cardio... | Hsa05418 | Fluid ... | 13/69 | 141/8541 | 6.792e-11 | 3.753e-09 | 1.841e-09 | 7422/4... |
| Human ... | Infect... | Hsa05144 | Malaria | 9/69 | 50/8541 | 1.649e-10 | 7.287e-09 | 3.575e-09 | 7124/6... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

**(File path: Figure+Table/3.5\_ClusterProfiler\_富集分析\_(PHENOL)/PHENOL-KEGG-enrichment-data.xlsx)**

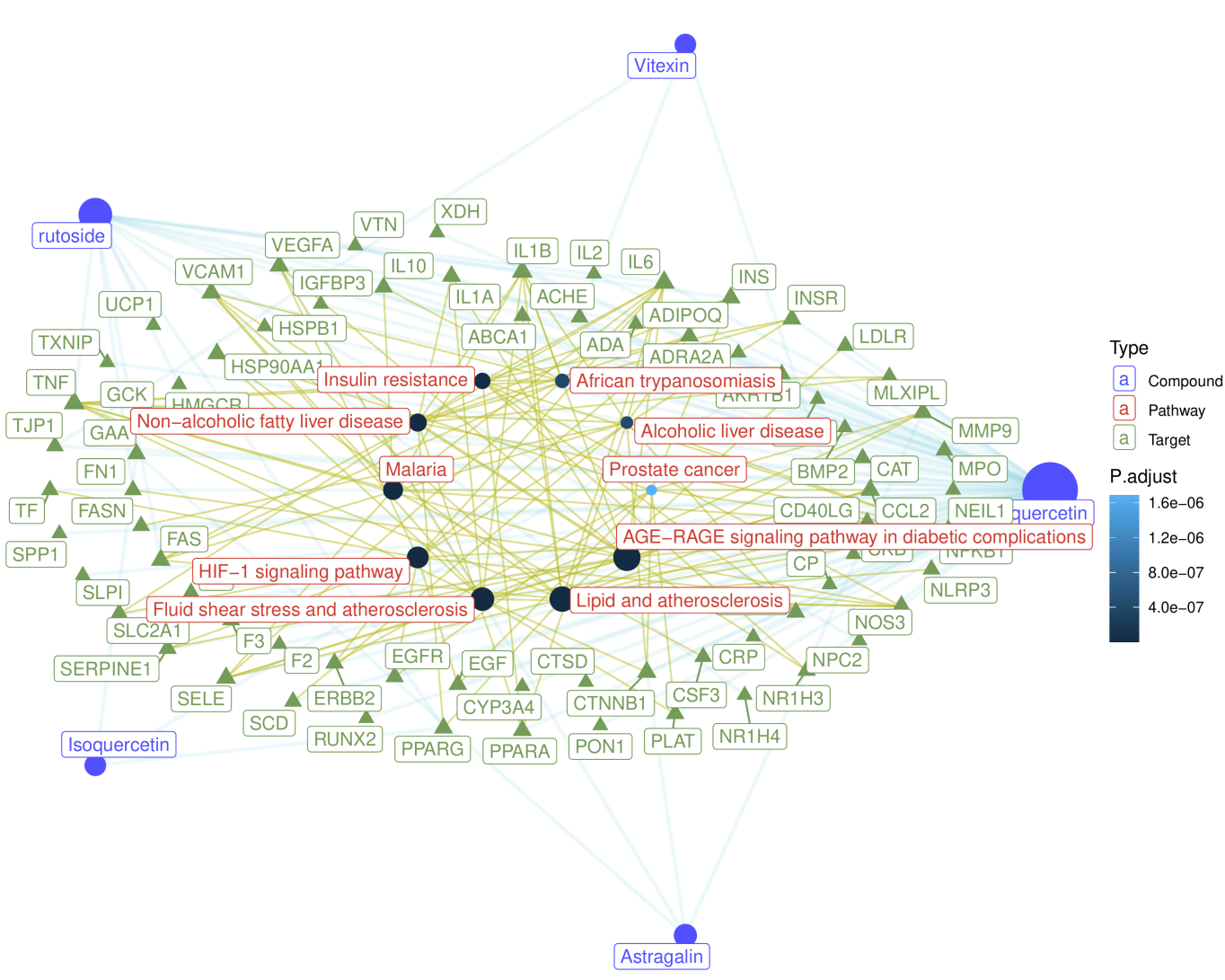
**Tab.** **6** PHENOL GO enrichment data

| Ont | ID | Descri... | GeneRatio | BgRatio | Pvalue | P.adjust | Qvalue | GeneID | Count |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| BP | GO:003... | Negati... | 21/72 | 437/18986 | 6.532e-18 | 2.095e-14 | 9.227e-15 | 3949/3... | 21 |
| BP | GO:003... | Respon... | 19/72 | 354/18986 | 4.098e-17 | 6.572e-14 | 2.895e-14 | 19/505... | 19 |
| BP | GO:004... | Positi... | 14/72 | 132/18986 | 6.465e-17 | 6.911e-14 | 3.044e-14 | 3949/3... | 14 |
| BP | GO:000... | Respon... | 19/72 | 375/18986 | 1.19e-16 | 9.544e-14 | 4.204e-14 | 19/505... | 19 |
| BP | GO:004... | Wound ... | 20/72 | 444/18986 | 1.596e-16 | 1.023e-13 | 4.508e-14 | 348/36... | 20 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

**(File path: Figure+Table/3.5\_ClusterProfiler\_富集分析\_(PHENOL)/PHENOL-GO-enrichment-data.xlsx)**

## 3.6 Network 中药-成分-疾病-靶点-富集通路网络 (PHENOL)

Fig. **[7](#PHENOL-network-pharmacology-with-filtered-type-enrich)** 展示了中药、成分、靶点 (中药靶点与疾病靶点的交集) 的网络图。 该图对中心度 (centrality\_degree) 较高的节点 (成分或靶点) 做了名称标注。 图中的图例标注了节点的所属类型：中药、化合物、靶点。 化合物可分为该中药唯一所含的化合物，或者与其他中药共有的化合物。 共有的化合物环绕在靶点周围，而唯一的化合物则环绕在中药周围。此外，图中标注了靶点基因富集的通路。



**Fig.** **7** PHENOL network pharmacology with filtered type enrich

**(File path: Figure+Table/3.6\_Network\_中药-成分-疾病-靶点-富集通路网络\_(PHENOL)/PHENOL-network-pharmacology-with-filtered-type-enrich.pdf)**

# 4 总结

见上述分析结果。

# Reference

1. Kong, X. *et al.* BATMAN-tcm 2.0: An enhanced integrative database for known and predicted interactions between traditional chinese medicine ingredients and target proteins. *Nucleic acids research* **52**, D1110–D1120 (2024).

2. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the r/bioconductor package biomaRt. *Nature protocols* **4**, 1184–1191 (2009).

3. Stelzer, G. *et al.* The genecards suite: From gene data mining to disease genome sequence analyses. *Current protocols in bioinformatics* **54**, 1.30.1–1.30.33 (2016).

4. Szklarczyk, D. *et al.* The string database in 2021: Customizable proteinprotein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research* **49**, D605–D612 (2021).

5. Chin, C.-H. *et al.* CytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Systems Biology* **8**, S11 (2014).

6. Wu, T. *et al.* ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**, (2021).