**生信分析报告**

**项目标题： 清心莲子饮网络药理学分析 ;**

**单 号： BSJF240734 ;**

**分析人员： 黄礼闯 ;**

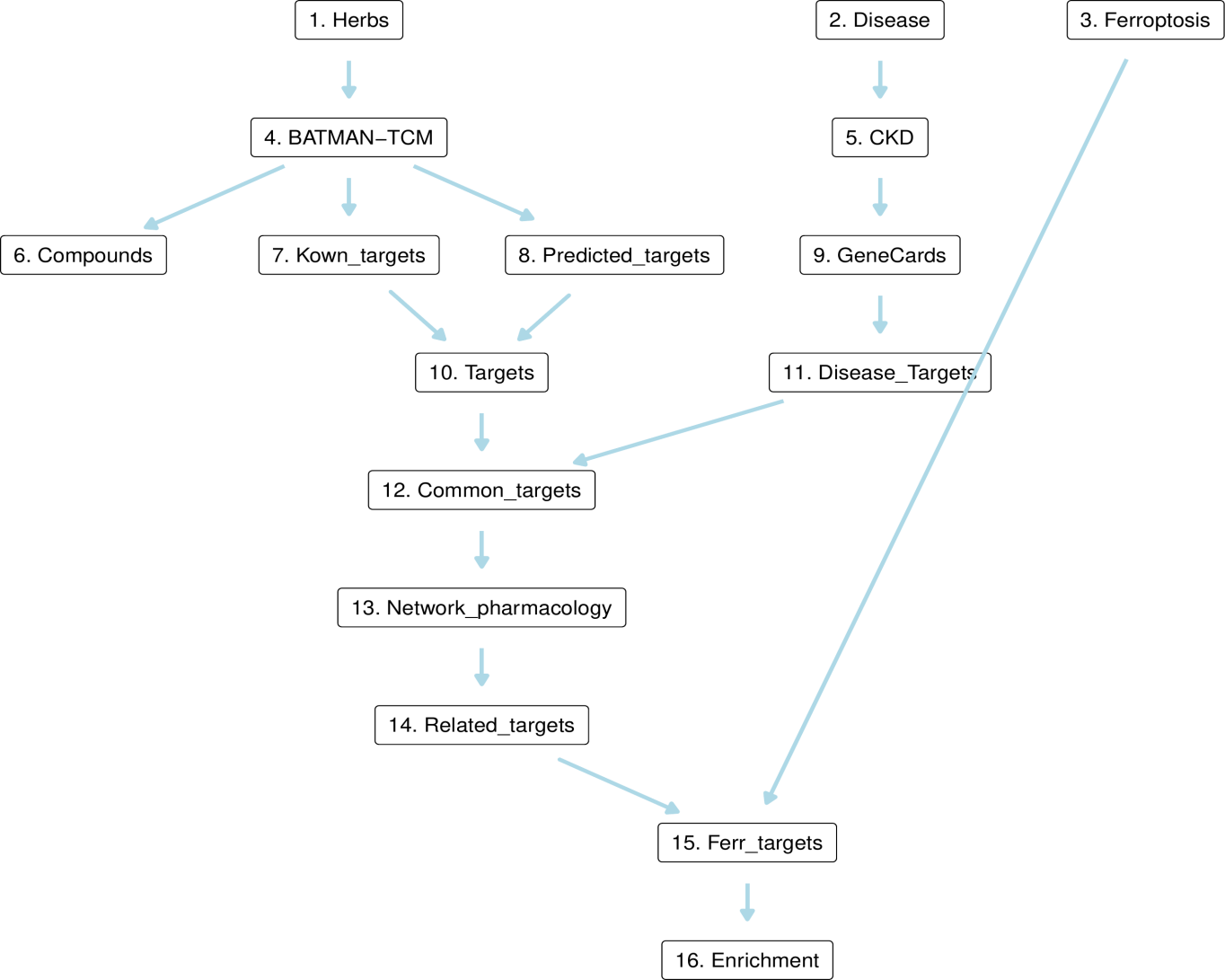
**分析类型： 生信分析 ;**

**委 托 人： 林波 ;**

**受 托 人： 杭州铂赛生物科技有限公司 .**

# 1 分析流程

1. 通过网络药理学筛选出清心莲子饮和CKD的共同靶点
2. 根据上一步靶点的相关信号通路与铁死亡相关信号通路进行通路富集



**Fig.** **1** Unnamed 1

**(File path: Figure+Table/1.0\_分析流程\_{#abstract}/Unnamed-1.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: QINGXIN)

从数据库 BATMAN-TCM ((2024, **IF:16.6**, Q1, Nucleic acids research)1) 中获取 CHI SHAO, MAI DONG, CHE QIAN ZI, REN SHEN, DI GU PI, LIAN ZI, HUANG QIN, GAN CAO, HUANG QI 等中药的成分、靶点数据。(即中药：莲子, 麦冬, 黄芩, 地骨皮, 人参, 车前子, 甘草, 赤芍, 黄芪)。以 Python 工具 HOB (2022, **IF:7.1**, Q1, Journal of Cheminformatics)2 预测化合物人类口服利用度 (20%)。 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 0.9。合并靶点数据。以 BiomaRt ((2009, **IF:13.1**, Q1, Nature protocols)3) 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: CKD)

从 GeneCards 数据库 (2016, **IF:NA**, NA, Current protocols in bioinformatics)4 获取 Chronic kidney disease 相关的基因集，得分 cut-off 为 3。

## 2.4 FerrDb 铁死亡调控因子 (Dataset: FERR)

从数据库 FerrDb V2 (2023, **IF:16.6**, Q1, Nucleic acids research)5 获取与铁死亡相关的调控因子或铁死亡与疾病之间的关联信息 <http://www.zhounan.org/ferrdb/current/>。

## 2.5 富集分析 (Dataset: COMMON)

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。 以 pathview R 包 (1.46.0) 对选择的 KEGG 通路可视化。

## 2.6 Limma 差异分析 (Dataset: CKD)

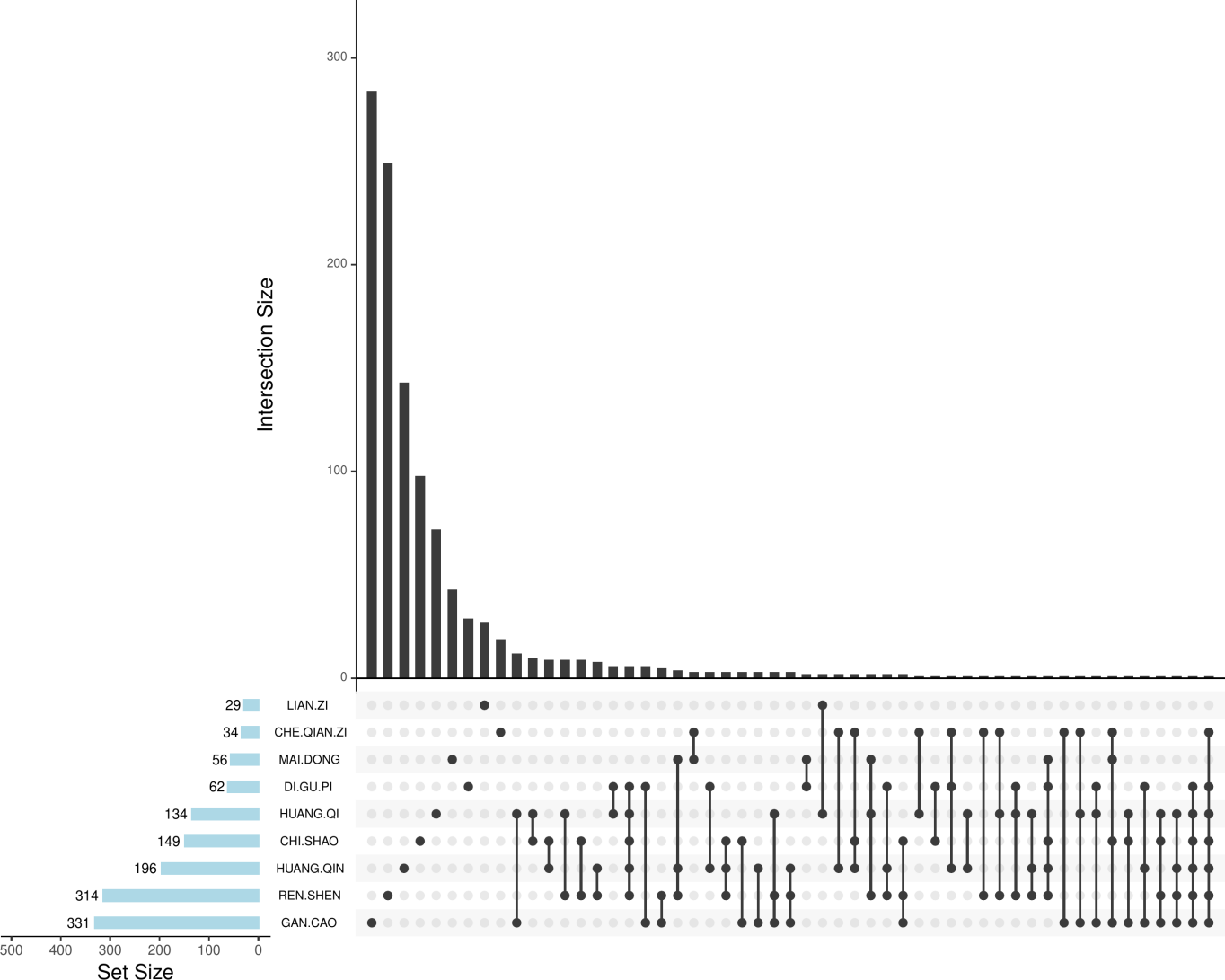
以 R 包 limma (3.62.1) (2005, **IF:**, , )7 进行差异分析。使用 log2 和 limma::normalizeBetweenArrays 对数据标准化。以 公式 ~ 0 + group + batch 创建设计矩阵 (design matrix) 用于线性分析。 创建对比矩阵，差异分析：CKD vs Control。使用 limma::lmFit, limma::contrasts.fit, limma::eBayes 拟合线形模型。以 limma::topTable 提取所有结果，并过滤得到 adj.P.Val 小于 0.05，|Log2(FC)| 大于 1 的统计结果。

# 3 分析结果

## 3.1 BATMAN 网络药理学 (QINGXIN)

从数据库 BATMAN-TCM 中药的成分、靶点数据 (详见方法章节) 。通过 PubChemR 获取化合物的结构式 (SMILES)。以化合物结构式 (SMILES) 通过 HOB 程序预测是否达到口服利用度 20%。HOB 预测结果，所有用于预测的化合物 (含有结构式信息的) 1900 个，达到口服利用度标准的有 1101 个 (注：根据唯一结构式统计)。经过筛选，各中药的化合物组成统计 (可能有交叉涵盖)：CHE QIAN ZI (n=34) , CHI SHAO (n=149) , DI GU PI (n=62) , GAN CAO (n=331) , HUANG QI (n=134) , HUANG QIN (n=196) , LIAN ZI (n=29) , MAI DONG (n=56) , REN SHEN (n=314) 。 共 1102 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 171 个 (非重复)。共包含靶点 1486 个 (非重复)。

Fig. **[2](#Intersection-of-herbs-compounds)** UpSet 图展示了中药各组成成分之间的交集 (中药名为拼音)。左侧横柱状图展示了各中药对应化合物数量。右侧展示了交集数目。各中药的化合物数目为数据库包含的化合物数量。数目最多的为 GAN CAO (n=331)；数目最少的为 LIAN ZI (n=29)；平均数目 145。 Fig. **[3](#Intersection-of-herbs-all-targets)** UpSet 图展示了中药的所有靶点 (中药名为拼音) 。该靶点是通过加和所有中药成分的靶点实现的。数目最多的为 REN SHEN (n=972)；数目最少的为 LIAN ZI (n=50)；平均数目 447。

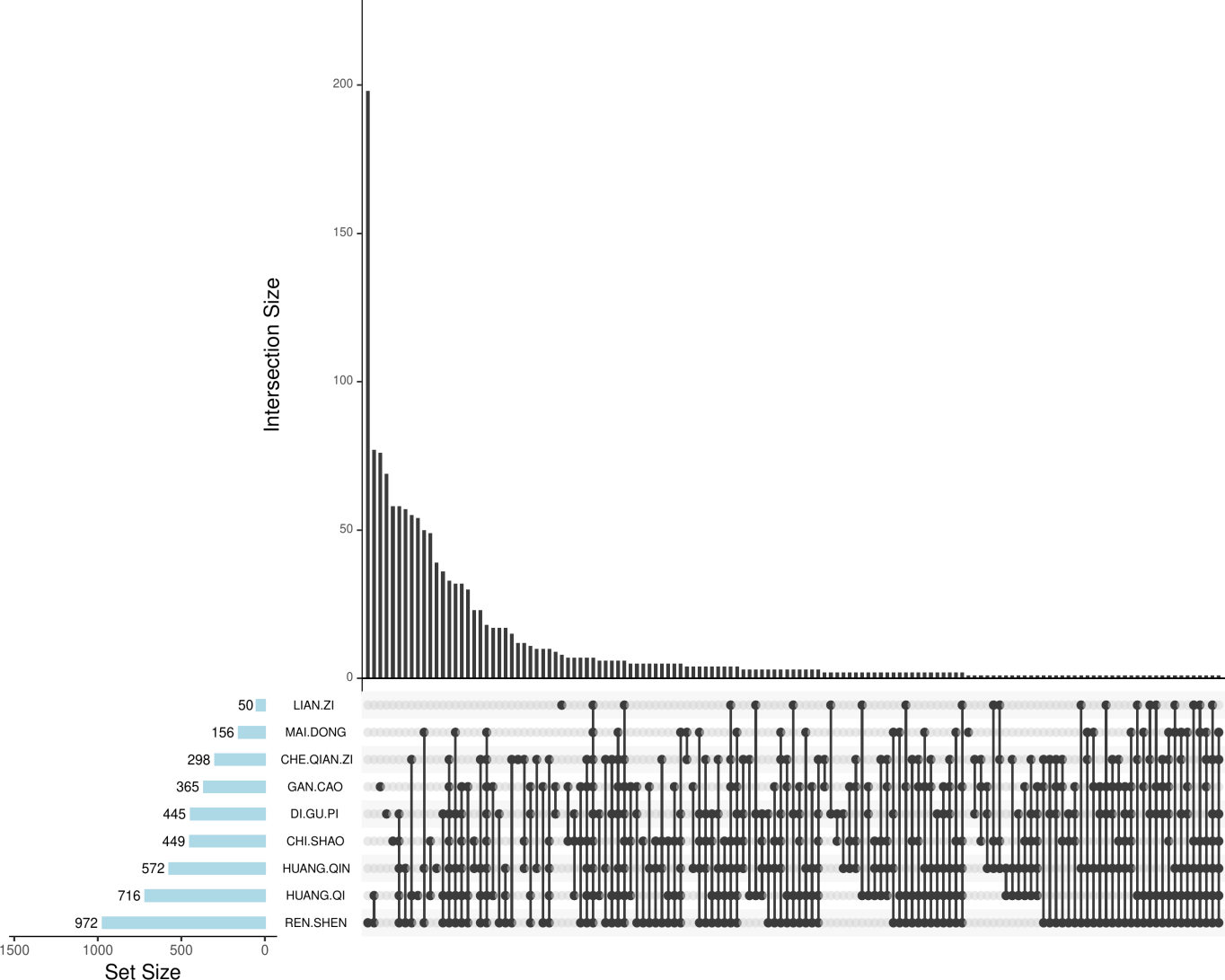


**Fig.** **2** Intersection of herbs compounds

**(File path: Figure+Table/3.1\_BATMAN\_网络药理学\_(QINGXIN)/Intersection-of-herbs-compounds.pdf)**

* All\_intersection:

**(See: Figure+Table/3.1\_BATMAN\_网络药理学\_(QINGXIN)/Intersection-of-herbs-compounds-content)**



**Fig.** **3** Intersection of herbs all targets

**(File path: Figure+Table/3.1\_BATMAN\_网络药理学\_(QINGXIN)/Intersection-of-herbs-all-targets.pdf)**

* All\_intersection: CASP3, NA, CASP9, HMOX1, MAPK1, MAPK3, AKT1

**(See: Figure+Table/3.1\_BATMAN\_网络药理学\_(QINGXIN)/Intersection-of-herbs-all-targets-content)**

## 3.2 GeneCards 基因获取 (CKD)

从 GeneCards 搜索 Chronic kidney disease, 获取对应靶点数据，统计为 Functional Element (n=11) , Genetic Locus (n=1) , Protein Coding (n=675) , Pseudogene (n=1) , RNA Gene (lncRNA) (n=20) , RNA Gene (miRNA) (n=28) , RNA Gene (tRNA) (n=2) 。共 738 个靶点。

**Tab.** **1** CKD disease related targets from GeneCards

| Symbol | Description | Category | UniProt ID | GIFtS | GC id | Score |
| --- | --- | --- | --- | --- | --- | --- |
| COL4A4 | Collagen T... | Protein Co... | P53420 | 55 | GC02M226973 | 34.39 |
| NPHS2 | NPHS2 Stom... | Protein Co... | Q9NP85 | 54 | GC01M179554 | 31.78 |
| COL4A5 | Collagen T... | Protein Co... | P29400 | 56 | GC0XP108439 | 30.73 |
| UMOD | Uromodulin | Protein Co... | P07911 | 55 | GC16M027663 | 29.38 |
| PKD2 | Polycystin... | Protein Co... | Q13563 | 60 | GC04P088007 | 26.96 |
| ... | ... | ... | ... | ... | ... | ... |

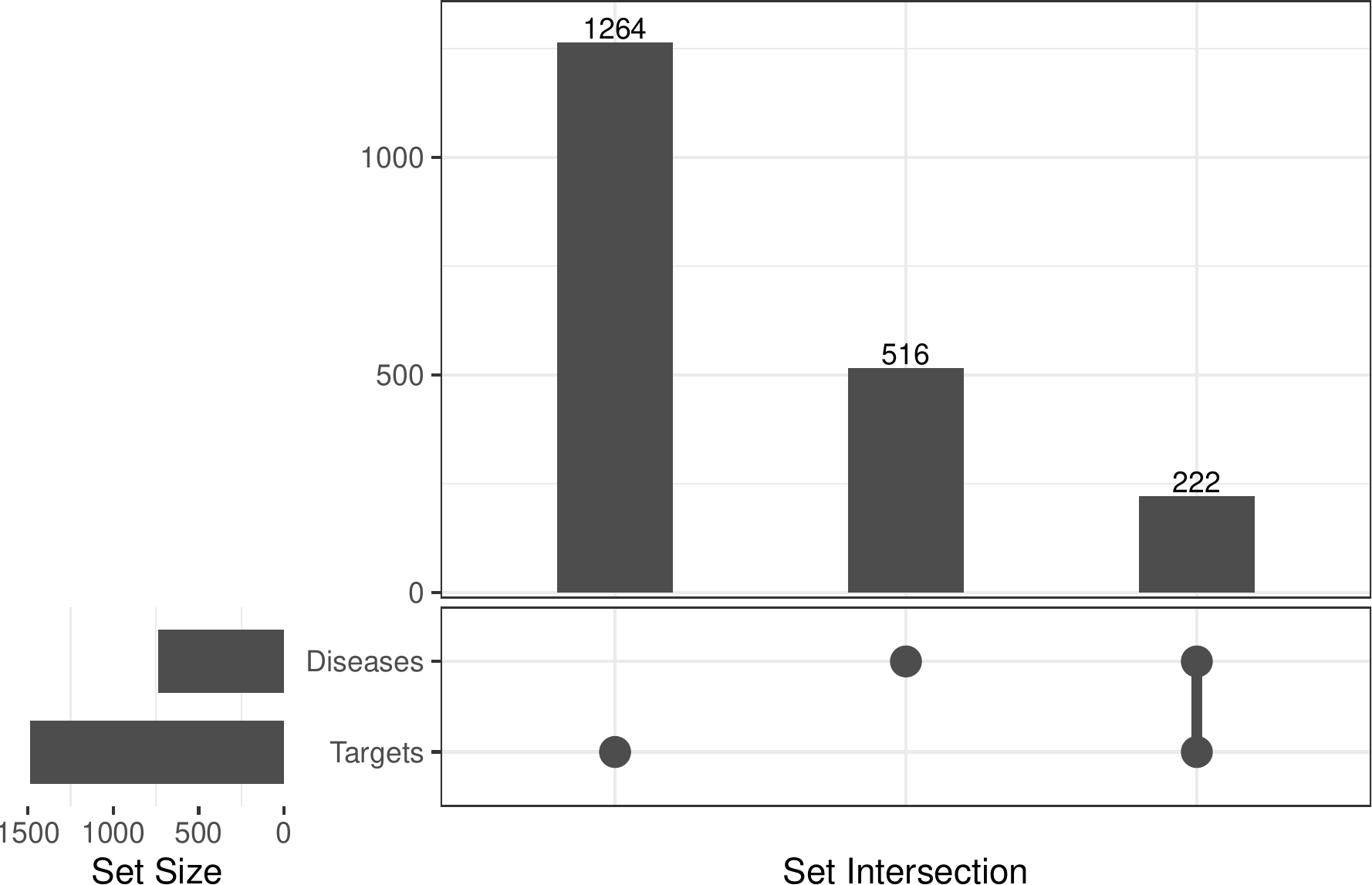
**(File path: Figure+Table/3.2\_GeneCards\_基因获取\_(CKD)/CKD-disease-related-targets-from-GeneCards.xlsx)**

* The GeneCards data was obtained by querying: Chronic kidney disease
* Restrict (with quotes): TRUE
* Filtering by Score:: Score > 3

## 3.3 Network 中药-成分-疾病-靶点网络 (QINGXIN)

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. **[4](#QINGXIN-Targets-intersect-with-related-targets)** Fig. **[5](#QINGXIN-network-pharmacology-with-filtered-type)** 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。唯一化合物统计为：CHE QIAN ZI (n=3) , CHI SHAO (n=15) , DI GU PI (n=7) , GAN CAO (n=20) , HUANG QI (n=6) , HUANG QIN (n=21) , LIAN ZI (n=4) , REN SHEN (n=23)  
Tab. **[2](#QINGXIN-network-pharmacology-with-filtered-type-original-data)** 为用于绘制中药-成分-疾病-靶点网络的数据集。

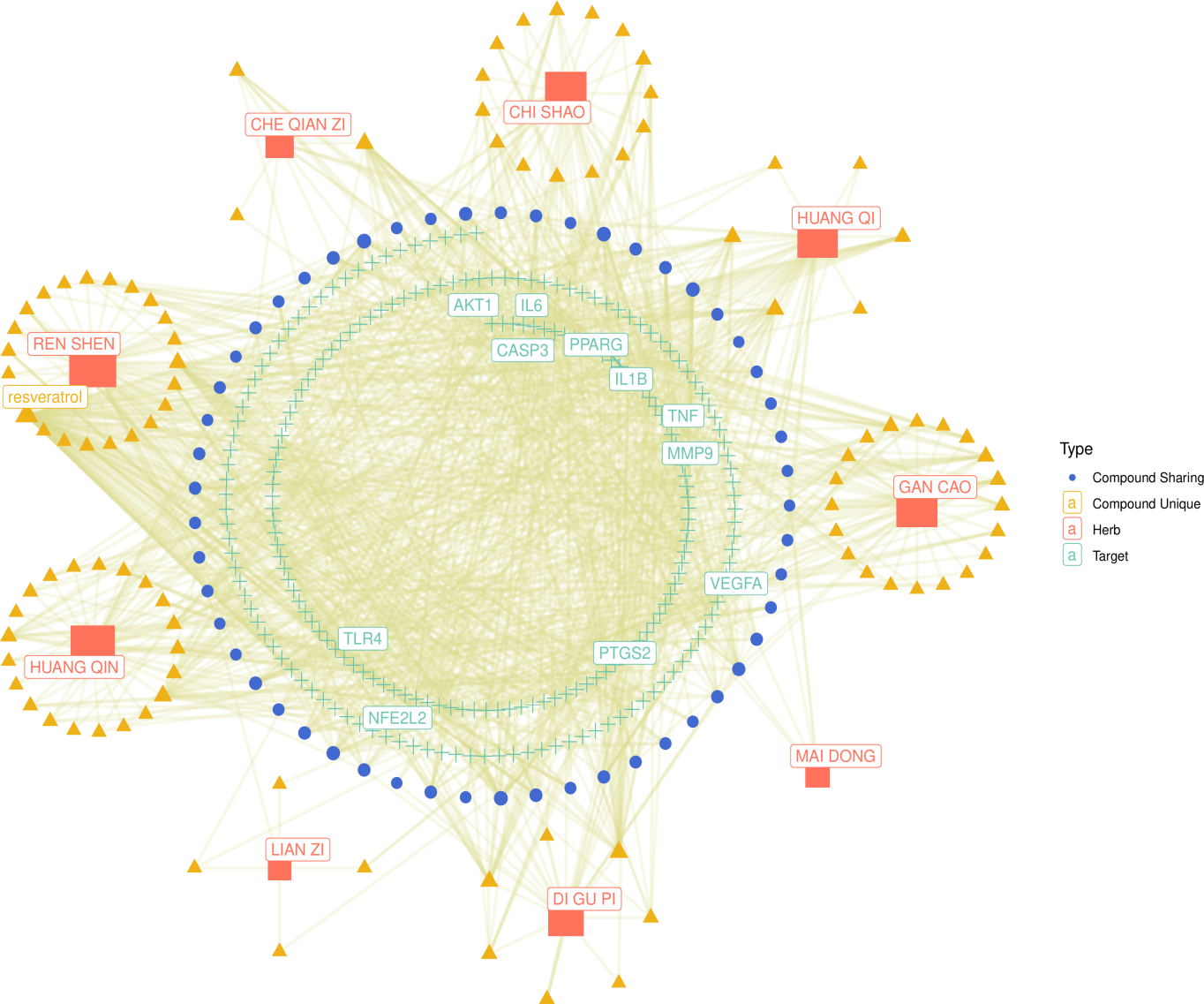


**Fig.** **4** QINGXIN Targets intersect with related targets

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

* All\_intersection: NPHS2, PKD2, FGF23, CASR, AGT, KL, ACE, CRP, ALB, REN, TNFRSF11B, APOA1, NPHS1, LCN2, IL6, ADIPOQ, PTH, VDR, AGTR1, TNF, MTHFR, FN1, LEP, TGFB1, NOS3, INS, BGLAP, SPP1, NPPB, IGF1, MUC1, CBS, TF, EDN1, PON1, APOB, AGXT, CD36, VEGFA, CCL2, GATM, TNFSF11, CETP, NPPA, TTR, CTNNB1, APRT, APOE, HMOX1, IGFBP3, CYP27B1, IL1B, FABP1, NR3C2, MIR21, ALPL, SLC12A3, LPL, HGF, MMP3, MMP1, TGFBR1, HMGB1, RBP4, HP, CALCA, AQP1, SLPI, MIR155, MMUT, MMP9, HAMP, SOD1, AGER, PLA2G7, RUNX2, MMP2, ESR1, GPX3, DNA…

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



**Fig.** **5** QINGXIN network pharmacology with filtered type

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

**Tab.** **2** QINGXIN network pharmacology with filtered type original data

| Herb pinyin name | Ingredient.name | Target.name |
| --- | --- | --- |
| GAN CAO | Protocatechuate | AKT1 |
| HUANG QI | Aminalon | GATM |
| HUANG QI | Aminalon | IL6 |
| HUANG QI | Aminalon | PDX1 |
| HUANG QI | Aminalon | GCK |
| ... | ... | ... |

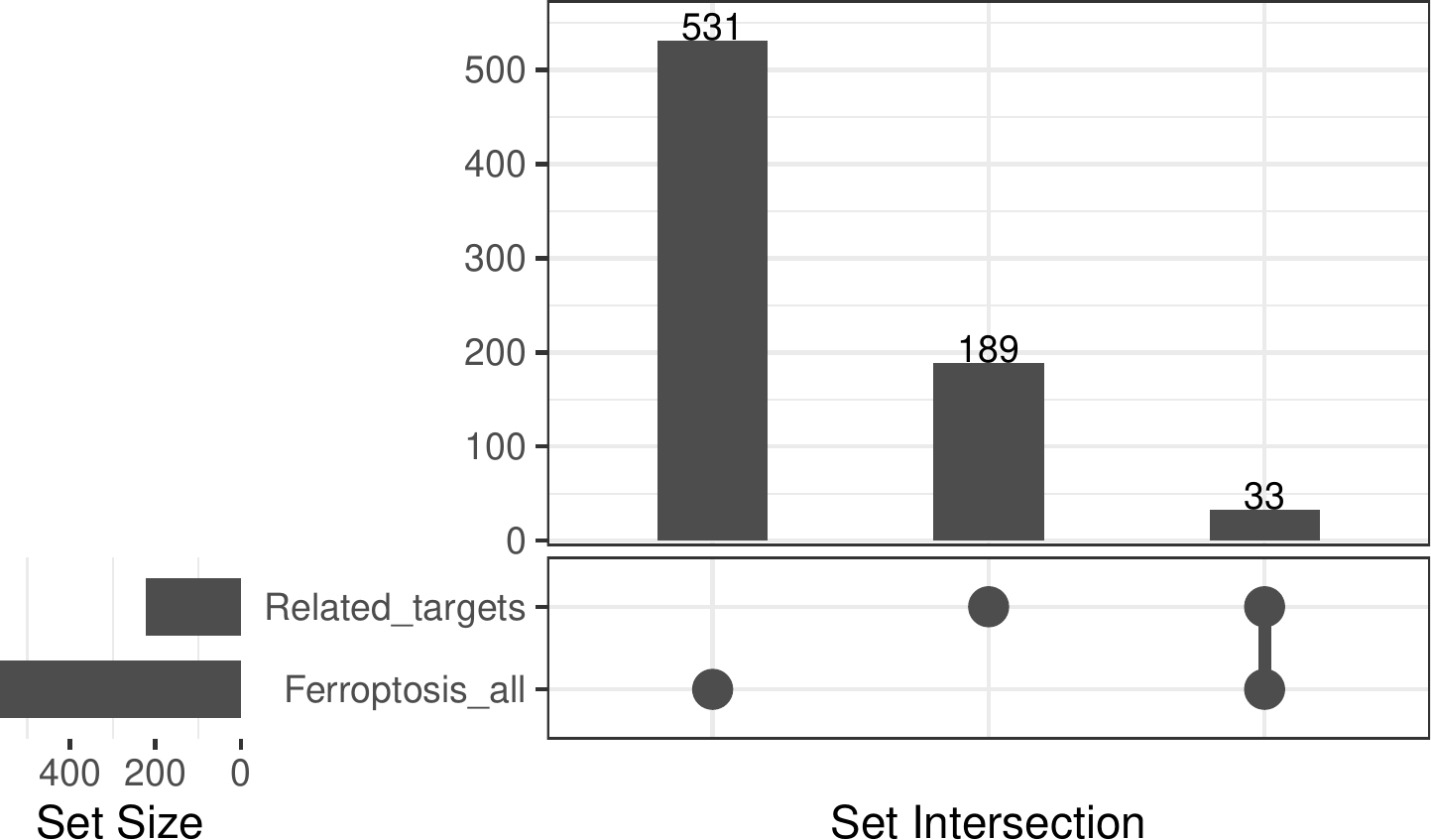
**(File path: Figure+Table/3.3\_Network\_中药-成分-疾病-靶点网络\_(QINGXIN)/QINGXIN-network-pharmacology-with-filtered-type-original-data.xlsx)**

## 3.4 FerrDb 铁死亡调控因子 (FERR)

从数据库 FerrDb V2 获取与铁死亡相关的调控因子或铁死亡与疾病之间的关联信息。铁死亡相关调控因子统计：marker (n=9) , driver (n=264) , suppressor (n=238) , unclassifier (n=110)

## 3.5 FerrDb 与铁死亡相关基因的交集 (FERR)

将基因集 (基因集 来自于 中药-成分-疾病-靶点网络 (QINGXIN)) 与 FerrDb 数据库中铁死亡调控因子取交集。得到共 33 个交集基因。 Fig. **[6](#Intersection-of-Related-targets-with-Ferroptosis-all)** 展示了与铁死亡相关调控因子的交集。数目最多的为 Ferroptosis\_all (n=564)；数目最少的为 Related\_targets (n=222)；平均数目 393。



**Fig.** **6** Intersection of Related targets with Ferroptosis all

**(File path: Figure+Table/3.5\_FerrDb\_与铁死亡相关基因的交集\_(FERR)/Intersection-of-Related-targets-with-Ferroptosis-all.pdf)**

* All\_intersection: ALB, LCN2, IL6, ADIPOQ, VDR, TGFB1, MUC1, CBS, TF, VEGFA, HMOX1, IL1B, TGFBR1, HMGB1, HAMP, IFNG, TLR4, NFE2L2, HIF1A, TIMP1, PPARG, GSTM1, HNF4A, TP53, DPP4, MTOR, CDH1, NOX4, IDO1, PTGS2, GSK3B, SLC11A2, KEAP1

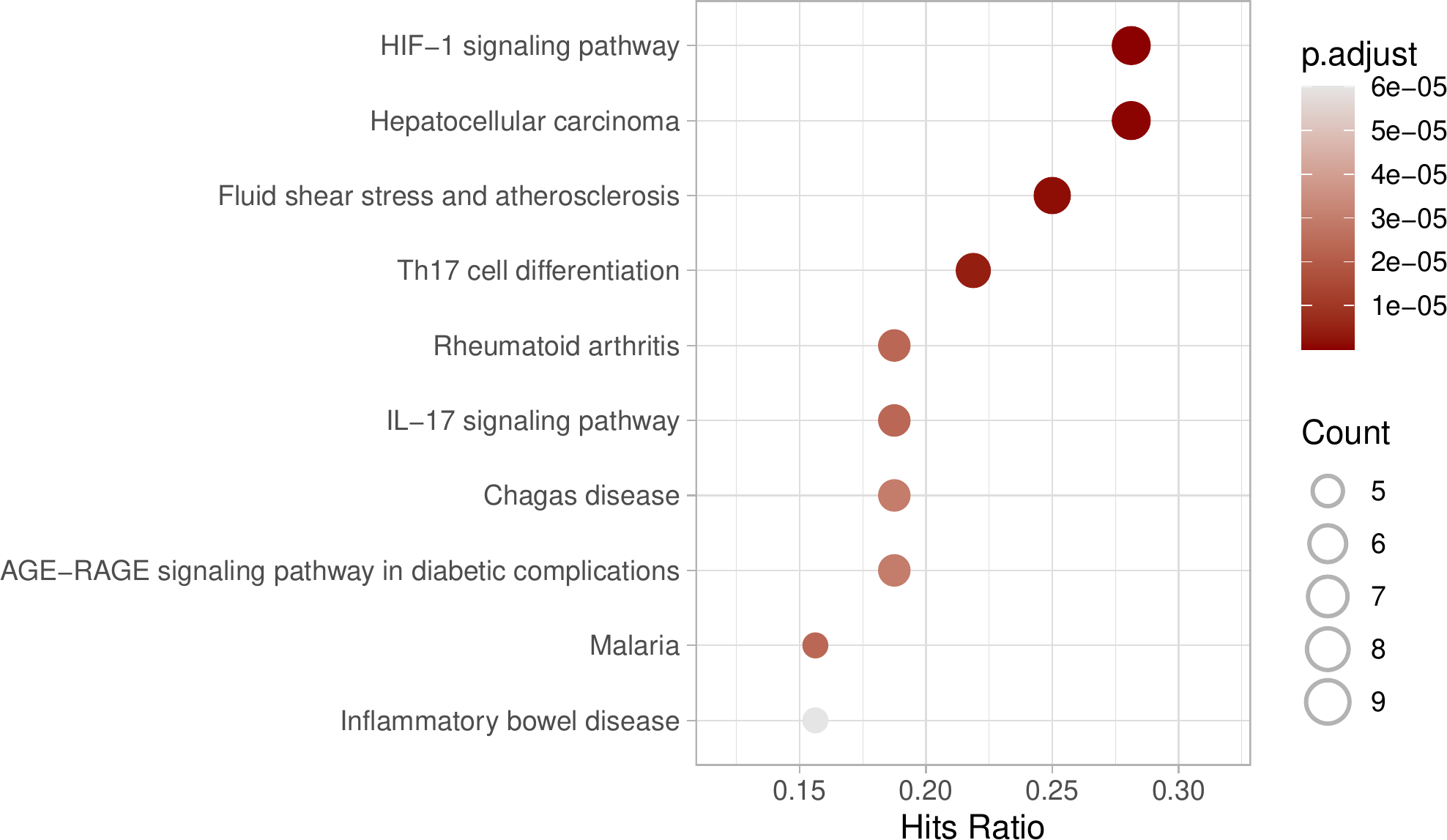
**(See: Figure+Table/3.5\_FerrDb\_与铁死亡相关基因的交集\_(FERR)/Intersection-of-Related-targets-with-Ferroptosis-all-content)**

## 3.6 富集分析 (COMMON)

对基因集 (基因集来自于铁死亡交集基因(FERR)) 进行 KEGG 和 GO 富集分析。以 pathview 探究基因集在通路 hsa04066 中的上下游关系。

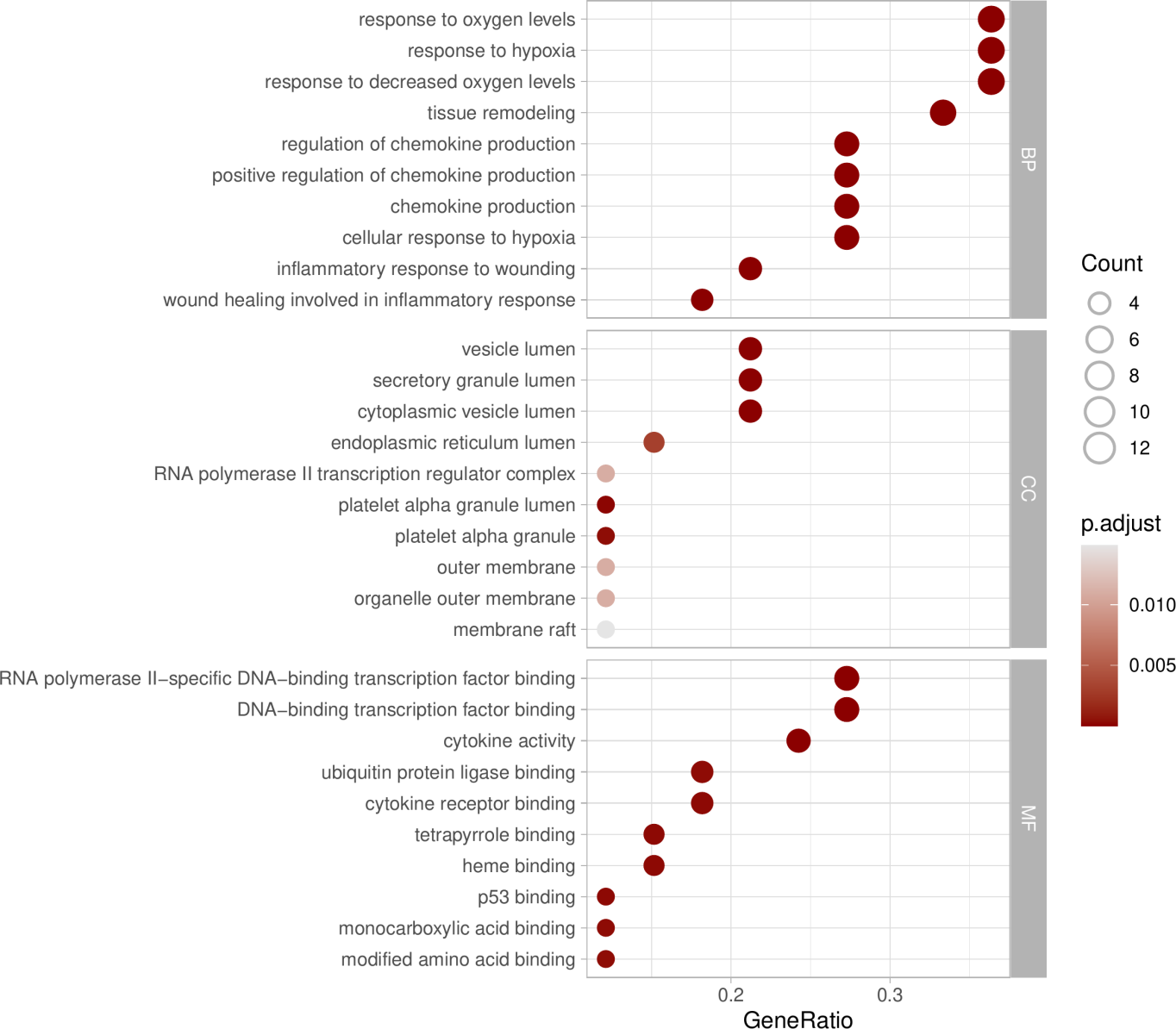
Fig. **[7](#COMMON-KEGG-enrichment)** KEGG 富集图展示了以 p.adjust 排序，前 10 的富集通路。 Fig. **[8](#COMMON-GO-enrichment)** GO 富集图展示了基因集在 GO 的 BP (Biological Process), MF (Molecular Function), CC (Cellular Component) 组中的富集结果 (以 p.adjust 排序，各自展示前 10 的富集通路) 。 Fig. **[9](#COMMON-hsa04066-visualization)** KEGG 通路可视化 (hsa04066) 展示了富集基因在该通路的上下游关系。通路图中的基因的映射颜色表示是否显著富集。

HIF-1 signaling pathway 已在多种疾病中被报道与铁死亡的关系 (2023, **IF:3.9**, Q2, Frontiers in endocrinology)8 (2022, **IF:4.4**, Q1, Frontiers in pharmacology)9 (2023, **IF:3.5**, Q2, Brain research bulletin)10。



**Fig.** **7** COMMON KEGG enrichment

**(File path: Figure+Table/3.6\_富集分析\_(COMMON)/COMMON-KEGG-enrichment.pdf)**



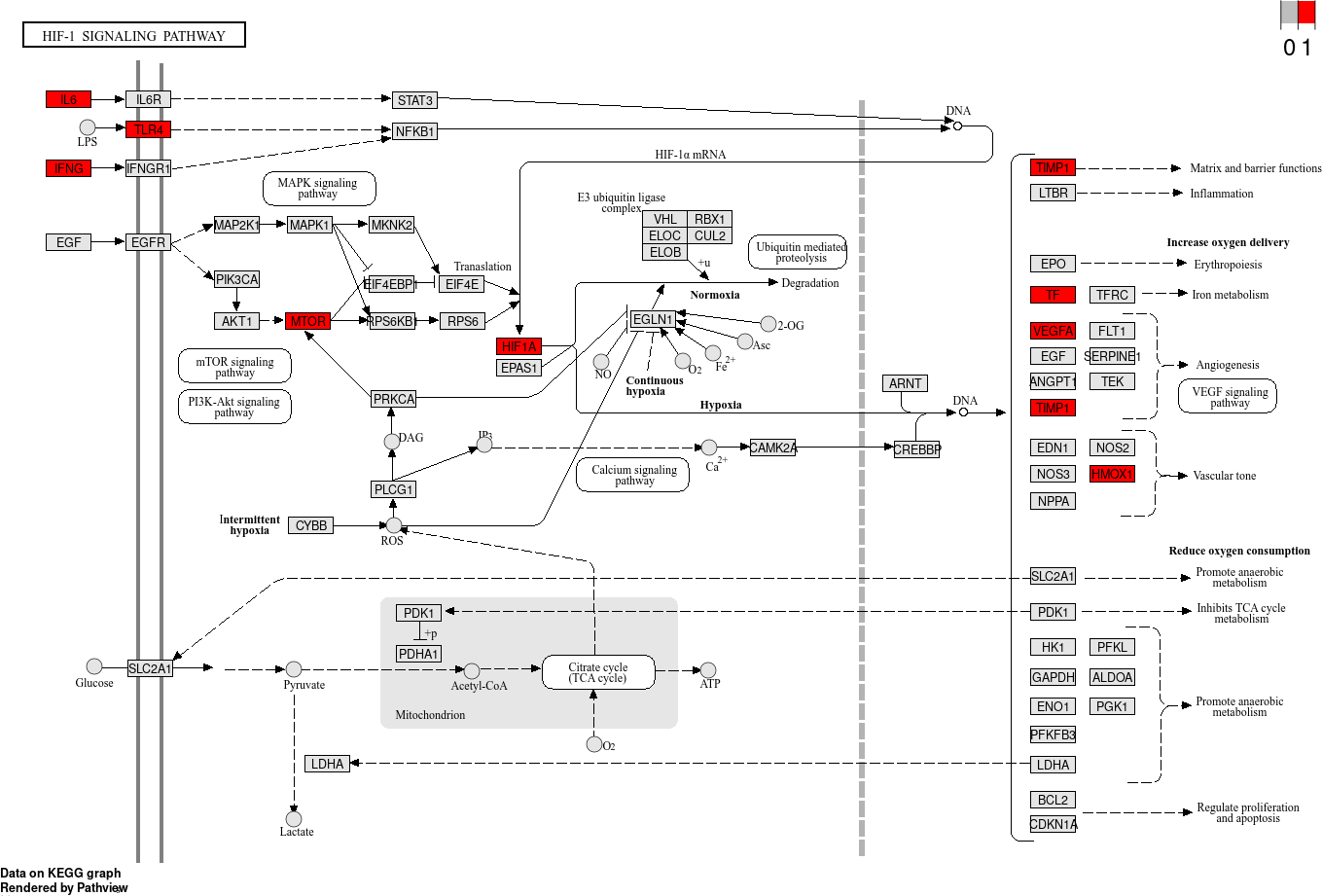
**Fig.** **8** COMMON GO enrichment

**(File path: Figure+Table/3.6\_富集分析\_(COMMON)/COMMON-GO-enrichment.pdf)**

**Tab.** **3** COMMON KEGG enrichment data

| Category | Subcat... | ID | Descri... | GeneRatio | BgRatio | Pvalue | P.adjust | Qvalue | GeneID |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Enviro... | Signal... | Hsa04066 | HIF-1 ... | 9/32 | 110/8878 | 1.0927... | 1.9560... | 9.0872... | 3091/3... |
| Human ... | Cancer... | Hsa05225 | Hepato... | 9/32 | 170/8878 | 5.3817... | 4.8166... | 2.2376... | 2932/2... |
| Human ... | Cardio... | Hsa05418 | Fluid ... | 8/32 | 141/8878 | 2.5292... | 1.5091... | 7.0109... | 2944/3... |
| Organi... | Immune... | Hsa04659 | Th17 c... | 7/32 | 109/8878 | 9.0534... | 4.0514... | 1.8821... | 3091/3... |
| Human ... | Infect... | Hsa05144 | Malaria | 5/32 | 50/8878 | 8.2896... | 2.3703... | 1.1012... | 3458/3... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

**(File path: Figure+Table/3.6\_富集分析\_(COMMON)/COMMON-KEGG-enrichment-data.xlsx)**



**Fig.** **9** COMMON hsa04066 visualization

**(File path: Figure+Table/3.6\_富集分析\_(COMMON)/COMMON-hsa04066-visualization.png)**

* Interactive figure: <https://www.genome.jp/pathway/hsa04066>
* Enriched genes: VEGFA, HIF1A, MTOR, TF, TIMP1, HMOX1, TLR4, IFNG, IL6

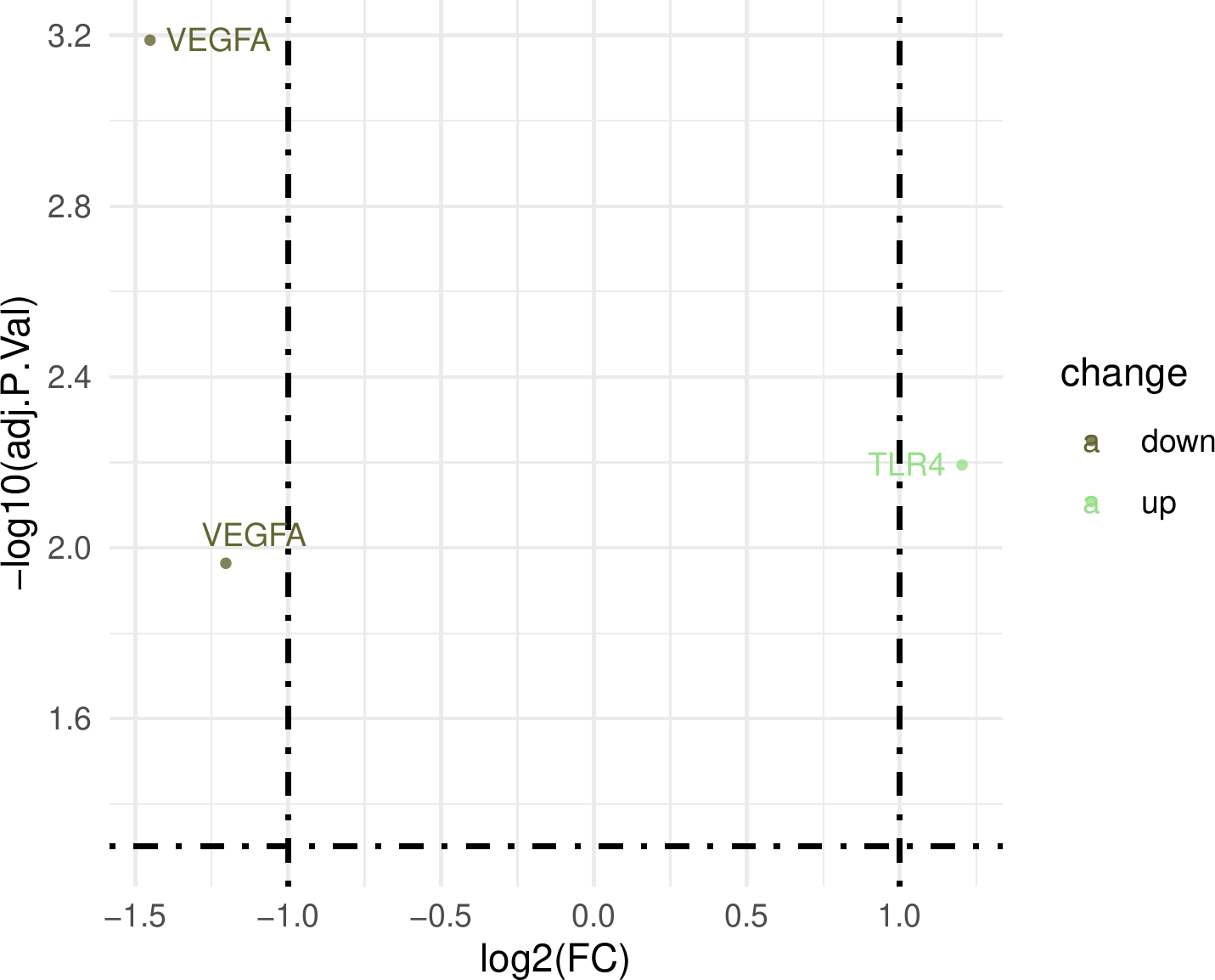
## 3.7 GEO 数据获取 (CKD)

* Data Source ID: GSE66494
* data\_processing: The scanned images were analyzed with Feature Extraction Software (Agilent) using default parameters (protocol GE1\_107\_Sep09 and Grid: 014850\_D\_20070820) to obtain background subtracted and spatially detrended Processed Signal intensities. Features flagged in Feature Extraction as Feature Non-uniform outliers were excluded.
* data\_processing.1: The signals were normalized with each median.

**(见Figure+Table/3.7\_GEO\_数据获取\_(CKD)/CKD-GSE66494-content)**

## 3.8 Limma 差异分析 (CKD)

样本分组：CKD (n=53) , Control (n=8) 。以 公式 ~ 0 + group + batch 创建设计矩阵 (design matrix) 用于线性分析(Batch: Discovery (n=53) , Validation (n=8) )。差异分析：CKD vs Control。(若 A vs B，则为前者比后者，LogFC 大于 0 时，A 表达量高于 B)。 针对 VEGFA, HIF1A, MTOR, TF, TIMP1, HMOX1, TIMP1, TLR4, IFNG, IL6 (基因集来自于通路 (hsa04066) 中的富集基因(Section: COMMON)) 差异分析。 上调或下调 DEGs 统计：up (n=1) , down (n=1) 以 wilcox.test 检验少量基因 (VEGFA, TLR4) 的表达。

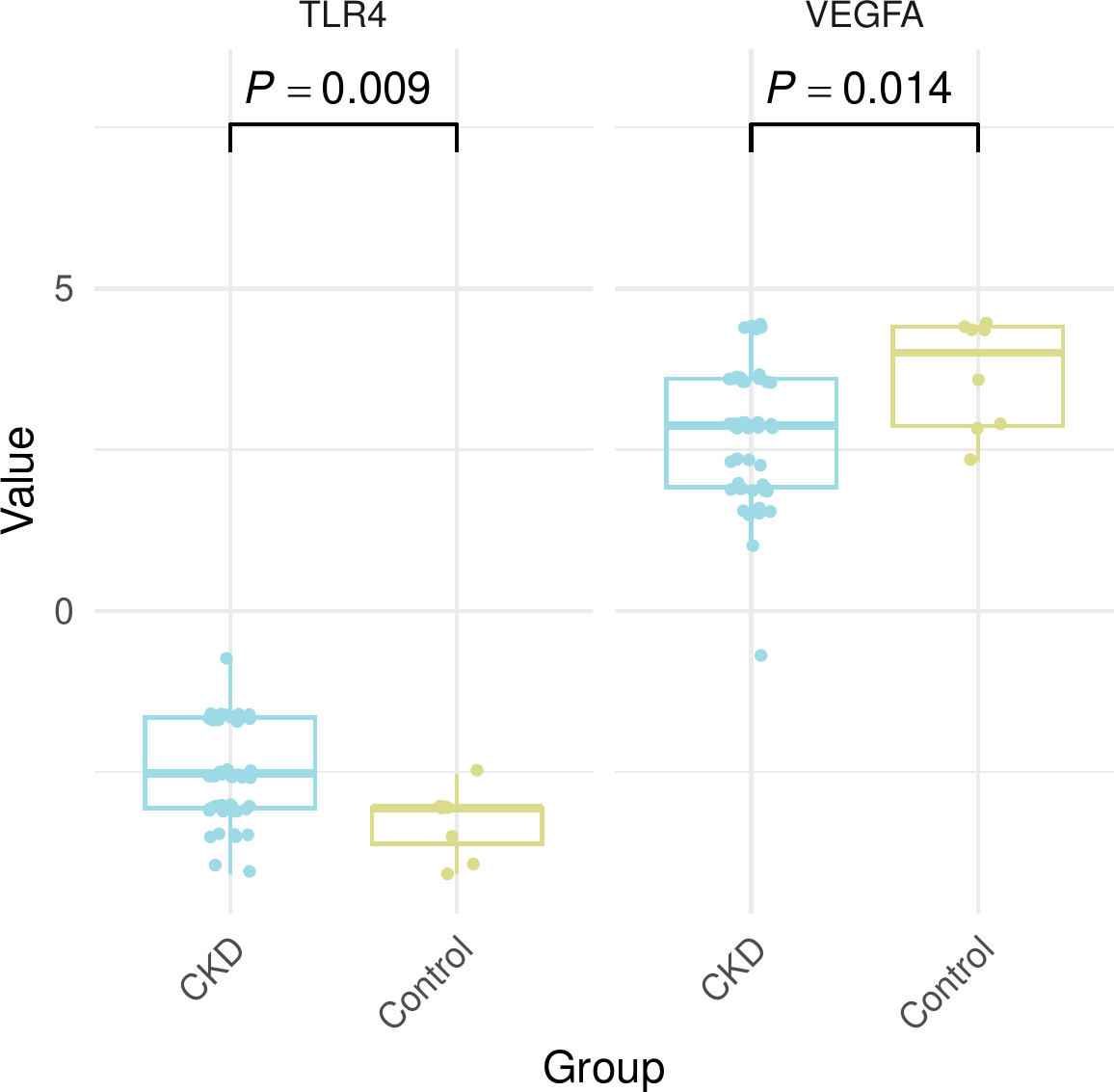


**Fig.** **10** CKD CKD vs Control

**(File path: Figure+Table/3.8\_Limma\_差异分析\_(CKD)/CKD-CKD-vs-Control.pdf)**

* adj.P.Val cut-off: 0.05
* Log2(FC) cut-off: 1

**(See: Figure+Table/3.8\_Limma\_差异分析\_(CKD)/CKD-CKD-vs-Control-content)**



**Fig.** **11** CKD wilcox test for some genes

**(File path: Figure+Table/3.8\_Limma\_差异分析\_(CKD)/CKD-wilcox-test-for-some-genes.pdf)**

# 4 总结

调控铁死亡的通路可能是 HIF-1 ，见Fig. **[9](#COMMON-hsa04066-visualization)**。 以 GEO 数据验证后发现，VEGFA, TLR4 为差异表达基因。 其他可能的通路，见Tab. **[3](#COMMON-KEGG-enrichment-data)**

# 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. Wei, M. *et al.* HobPre: Accurate prediction of human oral bioavailability for small molecules. *Journal of Cheminformatics* **14**, (2022).

3. 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).

4. 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).

5. Zhou, N. *et al.* FerrDb v2: Update of the manually curated database of ferroptosis regulators and ferroptosis-disease associations. *Nucleic acids research* **51**, D571–D582 (2023).

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

7. Smyth, G. K. Limma: Linear models for microarray data. in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* (eds. Gentleman, R., Carey, V. J., Huber, W., Irizarry, R. A. & Dudoit, S.) 397–420 (Springer-Verlag, 2005). doi:[10.1007/0-387-29362-0\_23](https://doi.org/10.1007/0-387-29362-0_23).

8. Xiang, Q., Zhao, Y. & Li, W. Identification and validation of ferroptosis-related gene signature in intervertebral disc degeneration. *Frontiers in endocrinology* **14**, (2023).

9. Liu, C., Li, Z. & Xi, H. Bioinformatics analysis and <i>in vivo</i> validation of ferroptosis-related genes in ischemic stroke. *Frontiers in pharmacology* **13**, (2022).

10. Dong, H. *et al.* Ferroptosis related genes participate in the pathogenesis of spinal cord injury via hif-1 signaling pathway. *Brain research bulletin* **192**, 192–202 (2023).