

## · 研究报告 ·

**黄坚** 首都医科大学附属北京友谊医院研究员、教授、博士生导师，北京市临床医学研究所分子生物学实验室主任。中国医药教育协会医学基因组学与生物信息学专业委员会副主任委员，国家药品标准物质委员会医疗器械（体外诊断试剂）分委会委员，中国研究型医院生物标志物专业委员会委员，中国医药生物技术协会生物诊断技术分会委员，中国研究型医院分子诊断专业委员会委员。先后主持国家“十三五”科技重大专项任务、国家自然科学基金等课题 10 余项。在 *Clin Cancer Res*、*EBioMedicine*、*J Med Genet*、*Human Mutat*、*Liver Int* 等杂志上以第一作者或通讯作者发表 SCI 论文近 30 篇，获得发明专利授权 5 项，其中 2 项专利已成功转化。



**伯晓晨** 军事科学院军事医学研究院研究员、博士生导师，全军转化医学委员会委员，中国药理学会网络药理学专业委员会常务委员，北京自动化学会理事。2007 年入选北京市“科技新星”，2016 年入选国家科技部创新人才推进计划“中青年科技领军人才”，获得中国科协“求是”杰出青年奖。2003 年开始从事生物信息学研究，目前研究方向为生物医学大数据处理与分析。作为负责人和主要研究者先后承担国家重点研发专项、国家重大新药创制、863 计划、国家自然科学基金重大研究计划等项目。在 *Nature*、*Nature Microbiology*、*Molecular Cell*、*Nucleic Acids Research*、*Bioinformatics* 等期刊发表 SCI 论文 80 余篇。



# 基于转录组的肝豆状核变性调控网络的构建和分析

杨晓曦<sup>1</sup>，何松<sup>2</sup>，李潇瑾<sup>1</sup>，周冬虎<sup>1</sup>，伯晓晨<sup>2</sup>，黄坚<sup>1</sup>

1 首都医科大学附属北京友谊医院 北京市临床医学研究所，北京 100050

2 军事科学院军事医学研究院 卫生勤务与血液研究所 生物信息研究室，北京 100850

杨晓曦，何松，李潇瑾，周冬虎，伯晓晨，黄坚. 基于转录组的肝豆状核变性调控网络的构建和分析. 生物工程学报, 2022, 38(10): 3844-3858.

YANG XX, HE S, LI XJ, ZHOU DH, BO XC, HUANG J. Construction and analysis of transcriptome-based hepatolenticular degeneration regulatory network. Chin J Biotech, 2022, 38(10): 3844-3858.

Received: June 25, 2022; Accepted: September 14, 2022

Supported by: Beijing Postdoctoral Sustentation Fund (2021-ZZ-012); Scientific Research Foundation of Beijing Friendship Hospital (yyqdk2020-52)

Corresponding authors: HUANG Jian. E-mail: huangj1966@hotmail.com  
BO Xiaochen. E-mail: boxiaoc@163.com

基金项目: 北京市博士后工作经费资助项目 (2021-ZZ-012); 北京友谊医院科研启动基金资助项目 (yyqdk2020-52)

**摘要:** 本研究旨在利用生物信息学方法构建经铜诱导的 *ATP7B* 基因敲除 HepG2 细胞系的转录调控网络。探讨关键转录因子在肝豆状核变性发生、发展中的潜在作用机制。收集公共基因表达数据库 (gene expression omnibus, GEO) 中包含野生型、*ATP7B* 基因敲除型、铜诱导的野生型和铜诱导的 *ATP7B* 基因敲除型 HepG2 细胞系数据。筛选由铜诱导产生的差异表达基因 (differentially expressed genes, DEGs) 后进行基因本体论 (gene ontology, GO)、京都基因和基因组百科全书 (Kyoto encyclopedia of genes and genomes, KEGG) 富集分析。基于蛋白相互作用网络, 识别疾病关键基因和功能模块, 并对关键功能模块中的基因进行富集分析。最后, 构建转录调控网络, 筛选核心转录因子。共筛选出 1 034 个差异表达基因, 其中上调 525 个, 下调 509 个。上、下调关键功能模块分别包括了 3 785 个和 3 931 个基因。关键功能模块中的基因主要定位于细胞-基质连接、染色体、剪接复合体、核糖体等区域, 共同参与了 mRNA 加工、组蛋白修饰、RNA 剪切、DNA 代谢调节、蛋白磷酸化等生物学过程, 且与转录共调控活性、DNA 转录因子结合、泛素样蛋白连接酶结合等分子功能相关。KEGG 分析表明功能模块中的基因显著富集的通路包括乙型肝炎、有丝分裂原活化蛋白激酶 (mitogen-activated protein kinase, MAPK) 信号通路、细胞衰老和凋亡、神经营养信号通路和神经变性途径等。肝豆状核变性转录调控网络包括 11 个差异表达转录因子和 96 个差异表达基因, 其中 U2AF1、NFRKB、FUS、MAX、SRSF1、CEBPA 和 RXRA 为核心差异表达转录因子。该研究为肝豆状核变性转录调控相关分子的生物学功能研究提供了重要的参考依据。

**关键词:** 肝豆状核变性; *ATP7B* 基因敲除; 差异表达基因; 转录调控网络; 差异表达转录因子

## Construction and analysis of transcriptome-based hepatolenticular degeneration regulatory network

YANG Xiaoxi<sup>1</sup>, HE Song<sup>2</sup>, LI Xiaojin<sup>1</sup>, ZHOU Donghu<sup>1</sup>, BO Xiaochen<sup>2</sup>, HUANG Jian<sup>1</sup>

<sup>1</sup> Beijing Institute of Clinical Medicine, Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China

<sup>2</sup> Department of Bioinformatics, Institute of Health Service and Transfusion Medicine, Academy of Military Medicine Sciences, Academy of Military Sciences, Beijing 100850, China

**Abstract:** A transcriptional regulatory network for wild-type and *ATP7B*-knockout HepG2 cells exposed to copper was constructed by bioinformatics methods to explore the potential mechanism of key transcription factors in the pathogenesis of hepatolenticular degeneration. The differentially expressed genes (DEGs) for wild-type and *ATP7B*-knockout HepG2 cell lines without copper and exposed to copper were collected from the gene expression omnibus (GEO) database. Gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis were performed for DEGs induced by copper. The key functional modules and genes were identified based on the protein-protein interaction (PPI) network. Moreover, the enrichment analysis of genes in functional modules was performed. Finally, a transcriptional regulatory network was constructed to screen the core transcription factors. A total of 1 034 genes, including 509 down-regulated genes and 525 up-regulated genes, were selected as DEGs. The

up-regulated and down-regulated functional modules based on PPI network included 3 785 and 3 931 genes, respectively. Genes in key functional modules were enriched in cell-substrate junction, chromosomal region, spliceosomal complex and ribosome. They were involved in mRNA processing, histone modification, RNA splicing, regulation of DNA metabolic process, protein phosphorylation and other biological processes. Moreover, they were correlated to transcriptional coregulator activity, DNA-binding transcription factor binding, ubiquitin-like protein ligase binding and other molecular functions. KEGG analysis showed that genes in key functional modules were significantly enriched in hepatitis B, MAPK signaling pathway, cellular senescence and apoptosis, neurotrophin signaling pathway and pathways of neurodegeneration-multiple diseases. The transcriptional regulatory network contained 11 differentially expressed transcription factors and 96 DEGs. Among them, U2AF1, NFRKB, FUS, MAX, SRSF1, CEBPA and RXRA were the core transcription factors, which may facilitate the study of the biological function of relevant molecules in transcriptional regulation of hepatolenticular degeneration.

**Keywords:** hepatolenticular degeneration; *ATP7B* gene knockout; differentially expressed gene; transcriptional regulatory network; differentially expressed transcription factors

肝豆状核变性 (hepatolenticular degeneration), 又名威尔森疾病 (Wilson disease, WD), 是一种常染色体隐性遗传性铜代谢障碍疾病。铜离子转运 ATP 酶  $\beta$  肽 (ATPase  $\text{Cu}^{2+}$  transporting beta polypeptide, ATP7B) 基因突变是肝豆状核变性的主要致病原因<sup>[1-2]</sup>。典型的肝豆状核变性通常伴随肝脏受累的表现 (包括转氨酶升高、急/慢性肝炎、肝硬化和肝功能衰竭) 或神经系统症状 (如运动障碍、肌张力障碍和精神状态等)<sup>[3]</sup>。然而, 部分患者在疾病发生的早期并不具备典型的肝豆状核变性症状, 甚至约 20% 的患者检测不到 *ATP7B* 基因的纯合或复合杂合突变<sup>[4-5]</sup>。这为疾病的早期诊断带来困难。

目前, 国内外关于肝豆状核变性致病机制的基础研究多局限于 *ATP7B* 基因突变谱和突变热点的发现, 较少考虑转录调控对疾病发生发展的影响。然而, 越来越多的证据表明, 转录调控和表观遗传机制在研究肝豆状核变性表型发生和发展中具有重要意义<sup>[6]</sup>。到目前为止, 仅有极少数的研究报道使用小鼠模型对肝豆状核变性的调控模式进行了初步的探索, 研究人

员利用小鼠肝脏组织分析了差异 lncRNA 和 mRNA 的表达, 并讨论了转录调控和信号通路的变化<sup>[7-8]</sup>。此外, 有研究发现, *ATP7B* 基因缺失的肝细胞通过激活细胞自噬降低细胞凋亡水平, 能够缓解过量铜诱导的肝细胞损伤<sup>[9]</sup>。然而, 在疾病发生的过程中, 与肝豆状核变性密切相关的转录调控因子和疾病本身涉及的转录调控机制尚不清楚。本研究基于转录组测序数据和生物信息学分析方法探讨了铜诱导的 *ATP7B* 基因敲除 HepG2 细胞系转录调控的分子机制, 筛选出肝豆状核变性转录调控网络中的关键转录因子, 为揭示疾病潜在的调控模式和促进疾病的分子功能机制研究提供了参考依据。

## 1 材料与方法

### 1.1 数据来源

以“Wilson disease”为关键词检索 GEO 数据库 (<https://www.ncbi.nlm.nih.gov/geo/>), 根据以下几条具体原则进行数据筛选工作: (1) 数据来源为人源; (2) 基于高通量转录组测序; (3) 至

少 3 个生物学重复。最终筛选到 GSE107323 数据集进行后续分析<sup>[9]</sup>。该数据集共包括 12 例样本, 包括野生型 HepG2 细胞系和经 *ATP7B* 基因敲除的 HepG2 细胞系, 以及铜诱导的野生型 HepG2 细胞系和铜诱导的 *ATP7B* 基因敲除的 HepG2 细胞系各 3 例, 其测序平台为 GPL18573。

## 1.2 差异表达基因筛选

分别利用 R 语言 (version 4.0.3) 的 *limma*<sup>[10]</sup>、*edgeR*<sup>[11]</sup>和 *DESeq2*<sup>[12]</sup> 3 个软件包对上述野生型和经 *ATP7B* 基因敲除的 HepG2 细胞系进行差异表达基因 (differentially expressed genes, DEGs) 分析, 设置同时满足  $P < 0.05$ 、 $|\log_2FC| > 1$  且  $FDR < 0.05$  为 DEGs 的筛选阈值。对 3 个软件包的分析结果取交集, 获得高可信度的、经基因敲除产生的 DEGs。再采用同样的计算方法和阈值设定, 对铜诱导的野生型 Hep2G 细胞系和铜诱导的 *ATP7B* 基因敲除的 HepG2 细胞系进行 DEGs 分析, 获得可信度高的且铜诱导的 DEGs。对两类差异基因取补集, 扣除 *ATP7B* 基因敲除导致的差异基因, 最终获得因铜诱导产生的差异表达基因。并使用 R 语言 *heatmap* 软件包 (version 1.0.12. <https://CRAN.R-project.org/package=heatmap>) 绘制差异基因热图。

## 1.3 DEGs 的富集分析

利用基因本体论 (gene ontology, GO) 数据库对上述 DEGs 进行 GO 功能富集分析, 包括生物学过程 (biological process, BP)、细胞组分 (cellular component, CC) 和分子功能 (molecular function, MF) 3 个方面。京都基因与基因组百科全书 (Kyoto encyclopedia of genes and genomes, KEGG) 通路富集分析用于识别 DEGs 参与的主要生物学通路。本研究使用 R 语言 *clusterProfiler*<sup>[13]</sup> 软件包对上述筛选出的 DEGs 进行 GO 和 KEGG 富集分析, 设定  $P < 0.05$

并利用 Benjamini-Hochberg 法进行  $P$  值校正, 识别显著富集的 GO term 和 KEGG 通路。

## 1.4 基于蛋白-蛋白相互作用网络识别肝豆状核变性关键基因和功能模块

本研究利用 Barabási 等<sup>[14]</sup>构建的包含了 15 970 个蛋白共 217 160 种相互作用的蛋白-蛋白相互作用 (protein-protein interaction, PPI) 网络 (命名为 Barabási-PPI), 分别提取上、下调 DEGs 和 Barabási-PPI 网络中度分布排名在前 5% 节点的交集, 定义为上、下调关键基因。然后, 提取上、下调关键基因以及它们在 Barabási-PPI 网络中的邻居节点, 作为疾病的关键功能模块。

## 1.5 肝豆状核变性关键功能模块中基因的富集分析

进一步对上述筛选出的关键功能模块中的基因进行 GO 和 KEGG 富集分析, 研究模块中基因的功能特性, 具体方法同 1.3 节中的研究方法, 设置  $P < 0.05$  为阈值并采用 Benjamini-Hochberg 法进行  $P$  值校正, 对富集结果进行筛选。

## 1.6 肝豆状核变性转录调控网络的构建和分析

Cistrome DB 数据库是目前最全面的研究 ChIP-seq 和 DNase-seq 的数据库, 主要收录了人和小鼠的转录和表观遗传基因调控数据<sup>[15]</sup>。从该数据库中下载了与 HepG2 细胞系相关的转录因子列表, 将转录因子和 DEGs 取交集筛选出差异表达转录因子。并利用 R 语言 *Hmisc* 软件包 (version 4.5-0. <http://CRAN.R-project.org/package=Hmisc>) 中的 *rcorr* 函数计算差异表达转录因子和差异表达基因之间的相关性, 定义  $cor > 0.5$  且  $P < 0.001$  为正相关,  $cor < -0.5$  且  $P < 0.001$  为负相关。最后, 将节点度分布大于 5 的差异表达转录因子和差异表达基因导入 *Cytoscape*<sup>[16]</sup> (version 3.6.1) 中绘制出肝豆状核变性转录调

控网络图。

## 2 结果与分析

### 2.1 DEGs 筛选

利用 R 语言 limma、edgeR 和 DESeq2 软件包，对野生型和经 *ATP7B* 基因敲除的 HepG2 细胞系，以及铜诱导的野生型和铜诱导的 *ATP7B* 基因敲除的 HepG2 细胞系分别进行差异

表达基因分析。对 3 个软件包的结果取交集后，再对这两类 DEGs 取补集，共筛选出 1 034 个由铜诱导产生的 DEGs (图 1A)，其中 509 个基因表达下调，525 个基因表达上调，并进一步绘制出差异表达基因热图 (图 1B)，可见铜诱导的野生型和铜诱导的 *ATP7B* 基因敲除型细胞系的 mRNA 表达出现显著差异。

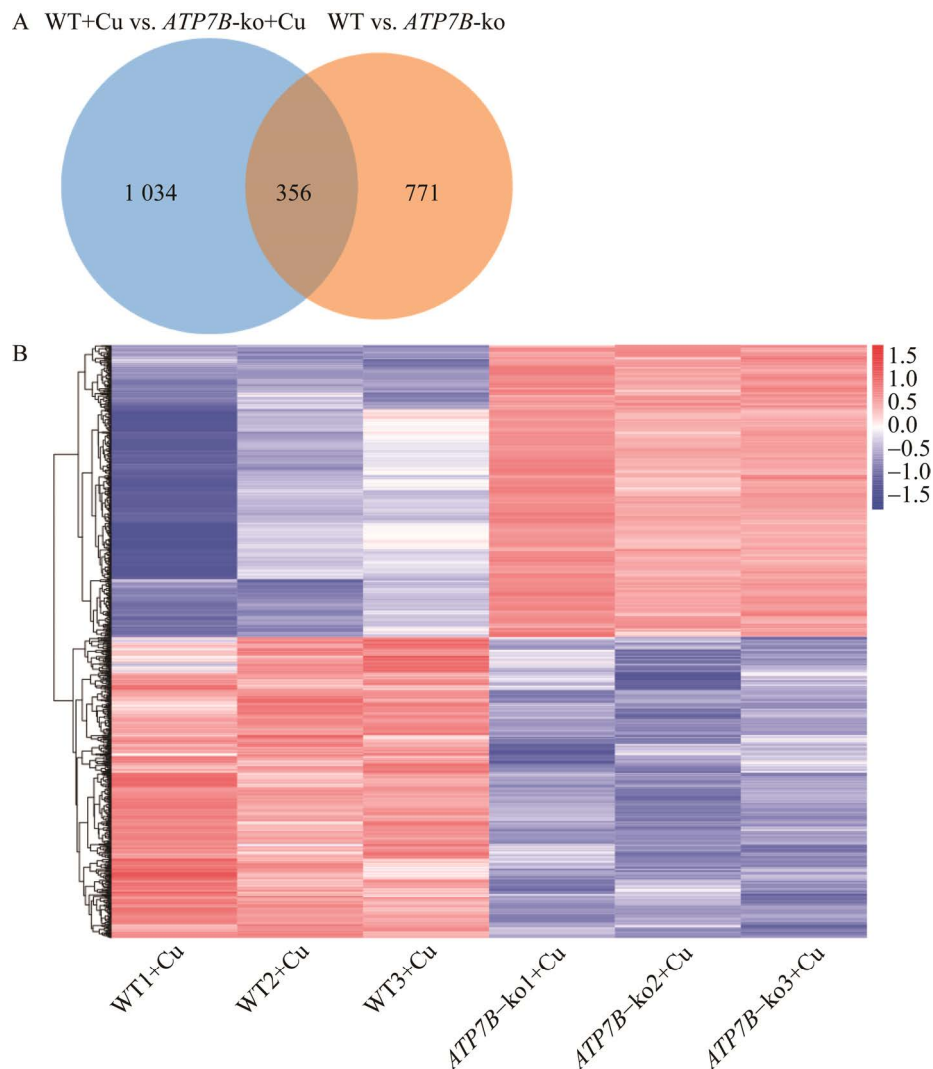


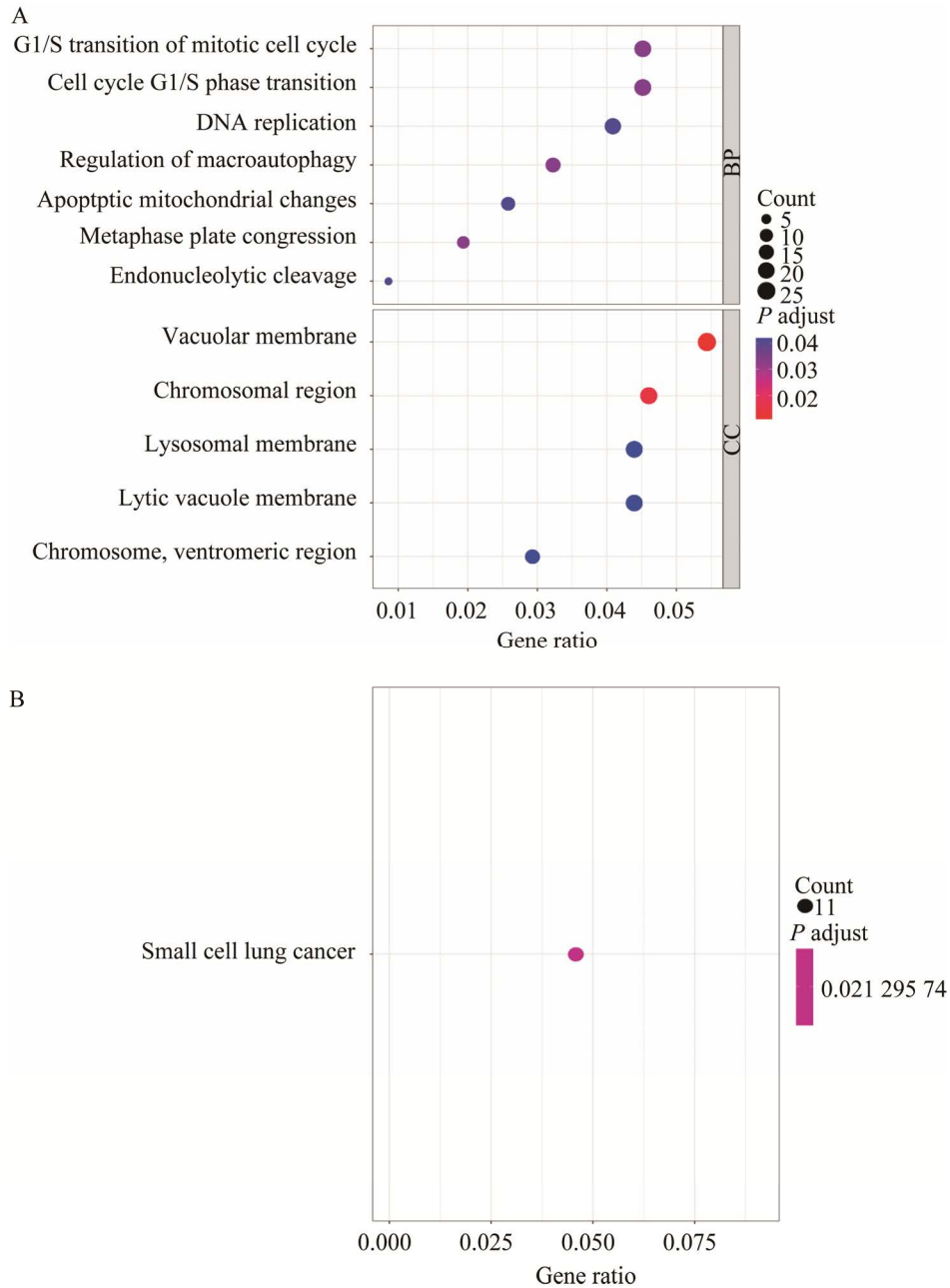
图 1 DEGs 筛选

Figure 1 Screening of DEGs. (A) Venn diagram showing overlap of DEGs by limma, edgeR and DESeq2 packages. (B) Heatmap of DEGs between wild-type and *ATP7B*-knockout HepG2 cells exposed to copper. WT1+Cu~WT3+Cu: wild-type HepG2 cells with Cu exposure; *ATP7B*-ko1+Cu~*ATP7B*-ko3+Cu: *ATP7B*-knockout HepG2 cells with Cu exposure; blue scale: down-regulated genes; red scale: up-regulated genes.

## 2.2 DEGs 富集分析

为了对 DEGs 影响的生物学功能和通路进行描述, 分别对上、下调 DEGs 进行 GO 和 KEGG 富集分析, 结果如图 2A-D 所示。下调 DEGs 主要富集在细胞周期、细胞自噬、DNA 修复和线粒体凋亡等 GO term 上 (图 2A-B)。而上调 DEGs 则显著富集到糖代谢、小分子代

谢、脂质代谢、醇代谢等 GO term 上, 以及碳代谢、糖代谢等、丙酸盐代谢信号通路上 (图 2C-D)。值得注意的是, 从总体水平上看, 上、下调 DEGs 显著富集的 GO-CC 和 KEGG 信号通路较少, 并且在 GO 分子功能方面没有显著富集到任何术语。因此, 需要进一步对这些 DEGs 进行研究。



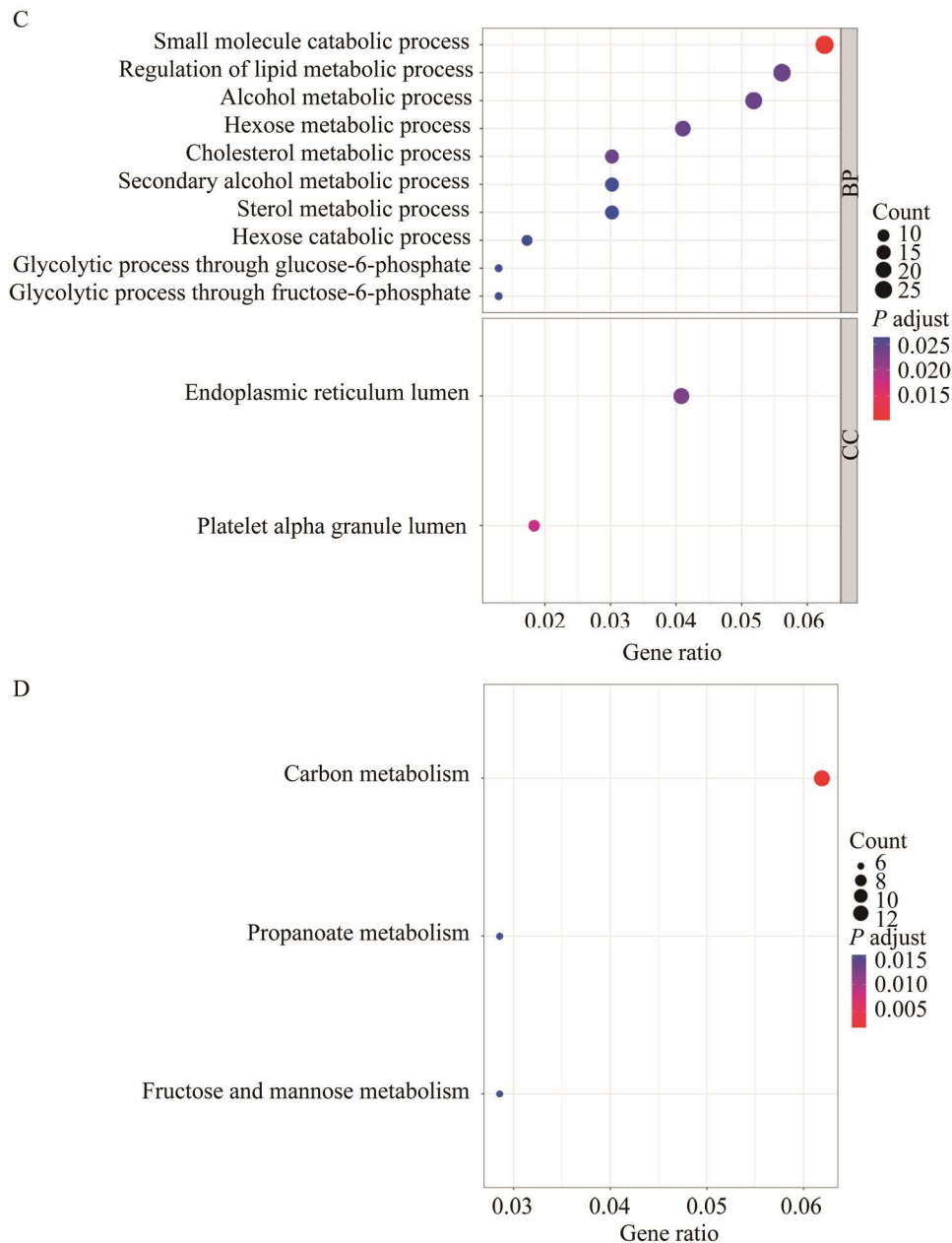


图 2 差异表达基因的 GO 和 KEGG 富集分析

Figure 2 GO and KEGG enrichment analysis of DEGs. GO enrichment results of down-regulated genes (A) and up-regulated genes (C); KEGG enrichment results of down-regulated genes (B) and up-regulated genes (D).

### 2.3 肝豆状核变性关键基因和关键功能模块的富集分析

根据方法 1.4 中的描述筛选出上、下调关键基因和关键功能模块，分别得到关键下调基因 37 个 (表 1) 和关键上调基因 17 个 (表 2)。上、下调关键功能模块分别包含了 3 785 个和

3 931 个基因。为了进一步研究关键功能模块中基因的生物学功能和通路富集情况，分别对这 2 个关键功能模块中的基因进行 GO 和 KEGG 富集分析，富集结果如图 3A–D 所示。值得注意的是，在 DEGs 的富集结果中，仅仅是差异表达基因无法显著富集到更多相关的生物学功



能和信号通路, 但通过 PPI 网络分析后, 关键功能模块中的基因均能够显著地富集到某些 GO term 和 KEGG 通路上。其中, GO 富集分析结果显示, 上、下调关键功能模块中的基因主

表 1 下调关键基因

Table 1 Down-regulated key genes

Gene symbol	Gene entrez ID	Degree of the node
AKT1	207	499
HNRNPU	3 192	479
EED	8 726	468
YWHAQ	10 971	442
BAG3	9 531	439
MAP1LC3B	81 631	365
CBL	867	285
RPL10	6 134	278
H2AX	3 014	233
KAT5	10 524	222
MAGOH	4 116	198
DYNLL1	8 655	195
CASP8	841	188
SDCBP	6 386	183
KPNA2	3 838	182
RNPS1	10 921	175
ACTA1	58	170
CHUK	1 147	169
MLH1	4 292	165
SRSF1	6 426	161
NFKBIA	4 792	158
TUBA1C	84 790	152
E2F3	1 871	148
LIMA1	51 474	146
FUS	2 521	145
LRIF1	55 791	142
TUBA4A	7 277	129
CEBPA	1 050	117
FXR2	9 513	117
U2AF1	7 307	116
PSMD11	5 717	115
MAX	4 149	112
SRSF2	6 427	111
TUBA1B	10 376	110
SF3B4	10 262	108
ITSN1	6 453	106
TFCP2	7 024	106

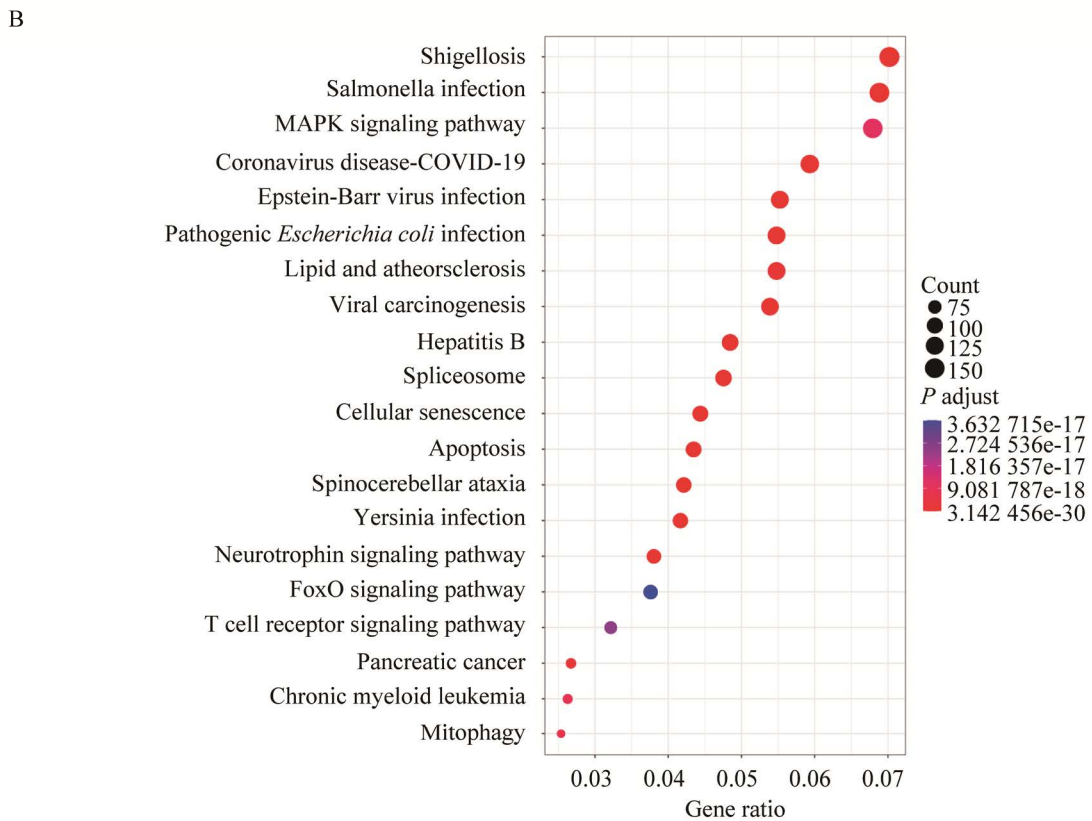
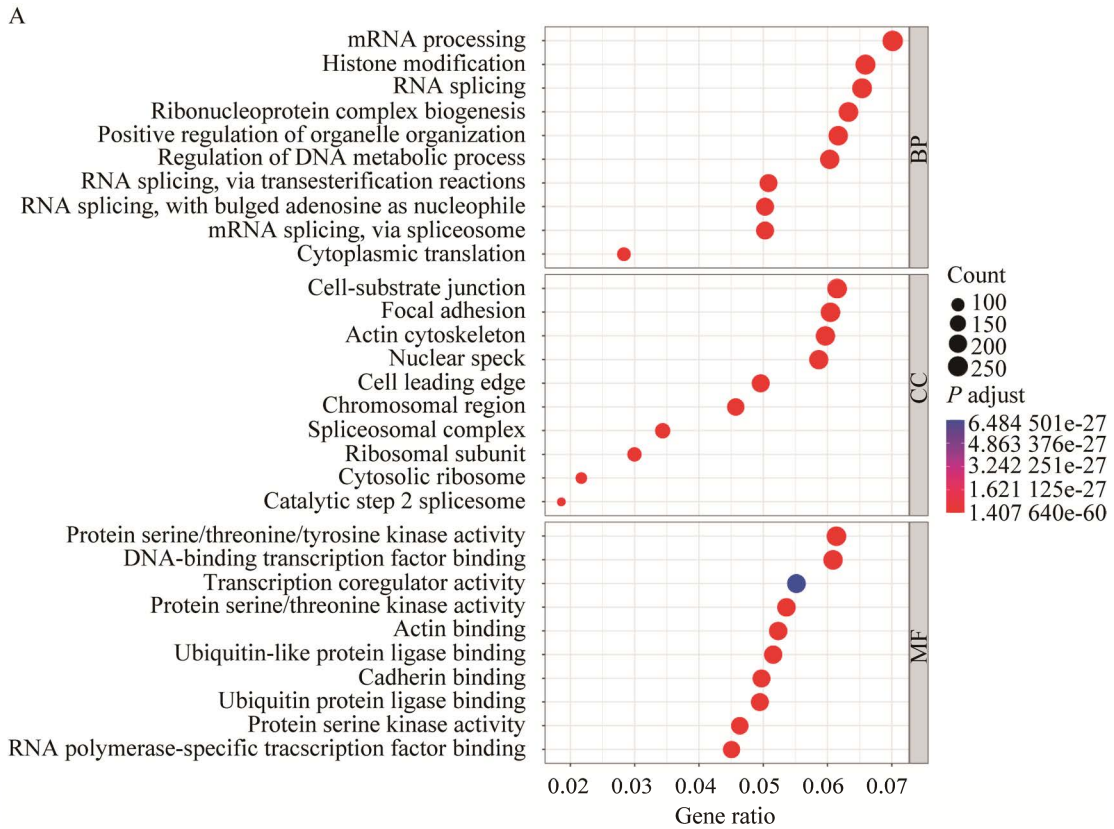
表 2 上调关键基因

Table 2 Up-regulated key genes

Gene symbol	Gene entrez ID	Degree of the node
APP	351	1 969
CSNK1E	1 454	361
HDAC5	10 014	341
EEF1A1	1 915	337
GOLGA2	2 801	304
ERBB2	2 064	277
TUBB	203 068	264
GAPDH	2 597	217
FBL	2 091	161
SUV39H1	6 839	151
TLE5	166	149
RXRA	6 256	141
ICAM1	3 383	140
PKM	5 315	137
SMAD9	4 093	123
PIK3R2	5 296	114
MYO18A	399 687	113

要定位于细胞-基质连接、染色体区域、剪接复合体、核糖体亚基、胞质核糖体等区域 (图 3A、3C); 在生物学过程方面, DEGs 主要参与 mRNA 加工、组蛋白修饰、RNA 剪切、DNA 代谢的调节、胞质翻译、蛋白磷酸化、蛋白定位、核糖核蛋白复合物发生 (图 3A、3C); 在分子功能方面, DEGs 主要参与了蛋白激酶活性、DNA 转录因子结合、转录共调控活性、泛素样蛋白连接酶结合、钙粘蛋白结合等 GO term (图 3A、3C)。KEGG 分析结果显示, DEGs 主要富集在乙型肝炎、MAPK 信号通路、病毒性致癌、细胞衰老和凋亡、神经营养信号通路、神经变性途径、FoxO 信号通路、T 细胞受体信号通路、线粒体自噬等相关信号通路 (图 3B、3D)。由于上、下调关键功能模块中存在部分重叠的基因, 因此关键功能模块富集分析得到的 GO term 和 KEGG 通路存在部分重叠。但无论从上调功能模块的角度还是下调功能模块的角度分析, 这部分 GO term 和通路均受到了影响, 表明这些 GO 术语和 KEGG 通路对于肝豆状核变性具有重要的生物学意义。





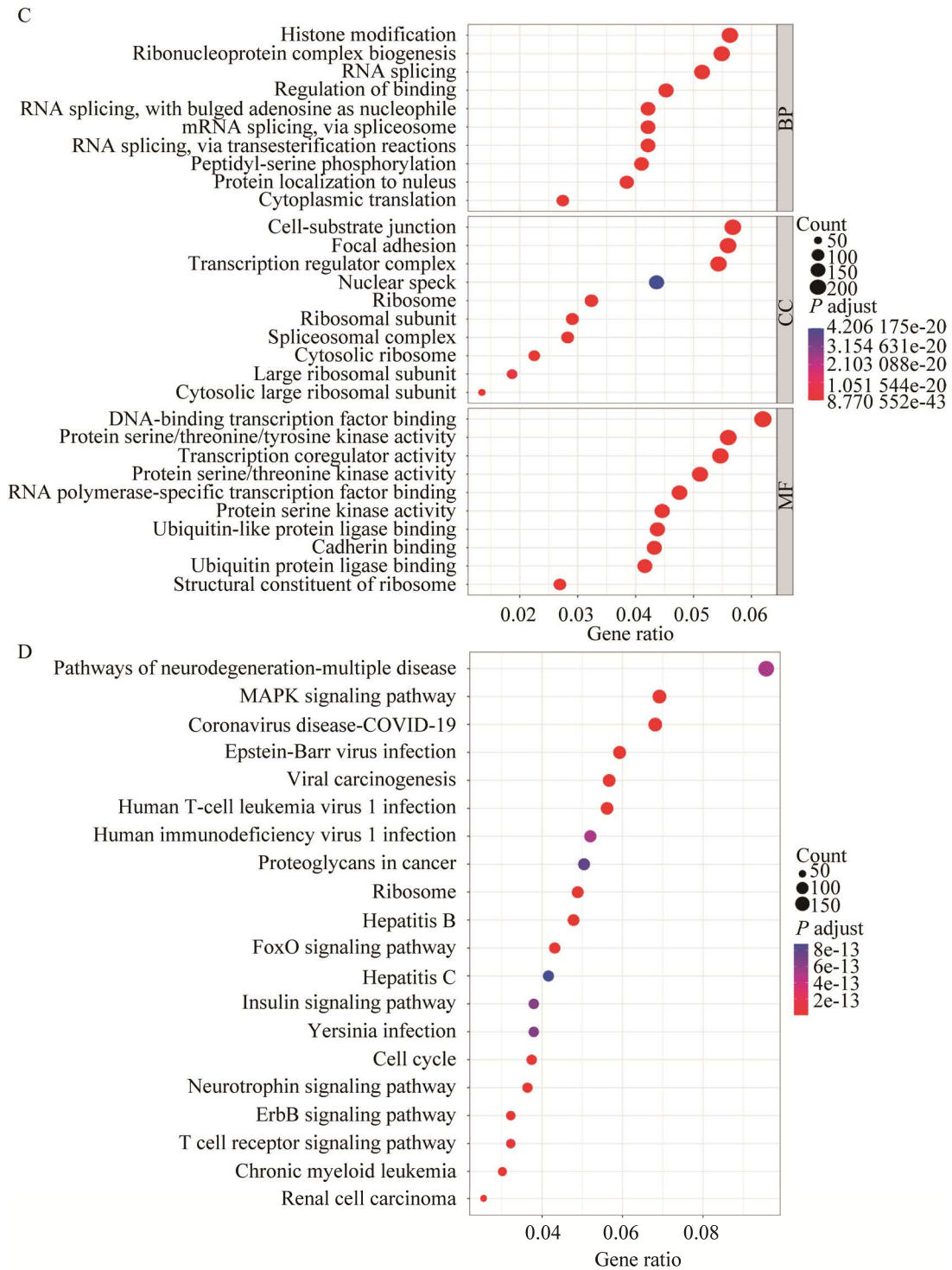


图 3 关键功能模块中基因的 GO 和 KEGG 富集分析

Figure 3 GO and KEGG enrichment analysis of genes in key functional modules. GO enrichment results of genes in down-regulated key functional modules (A) and in up-regulated key functional modules (C); KEGG enrichment results of genes in down-regulated key functional modules (B) and in up-regulated key functional modules (D).

## 2.4 肝豆状核变性转录调控网络的构建和分析

为了减少对单个基因过度关注的局限性,本研究从整体角度系统分析了肝豆状核变性差

异表达基因的转录调控方式,详细讨论了差异表达转录因子和差异表达基因之间的调控关系,并绘制了转录调控网络图(图4)。最终得到的网络共包括11个差异表达转录因子和96个差

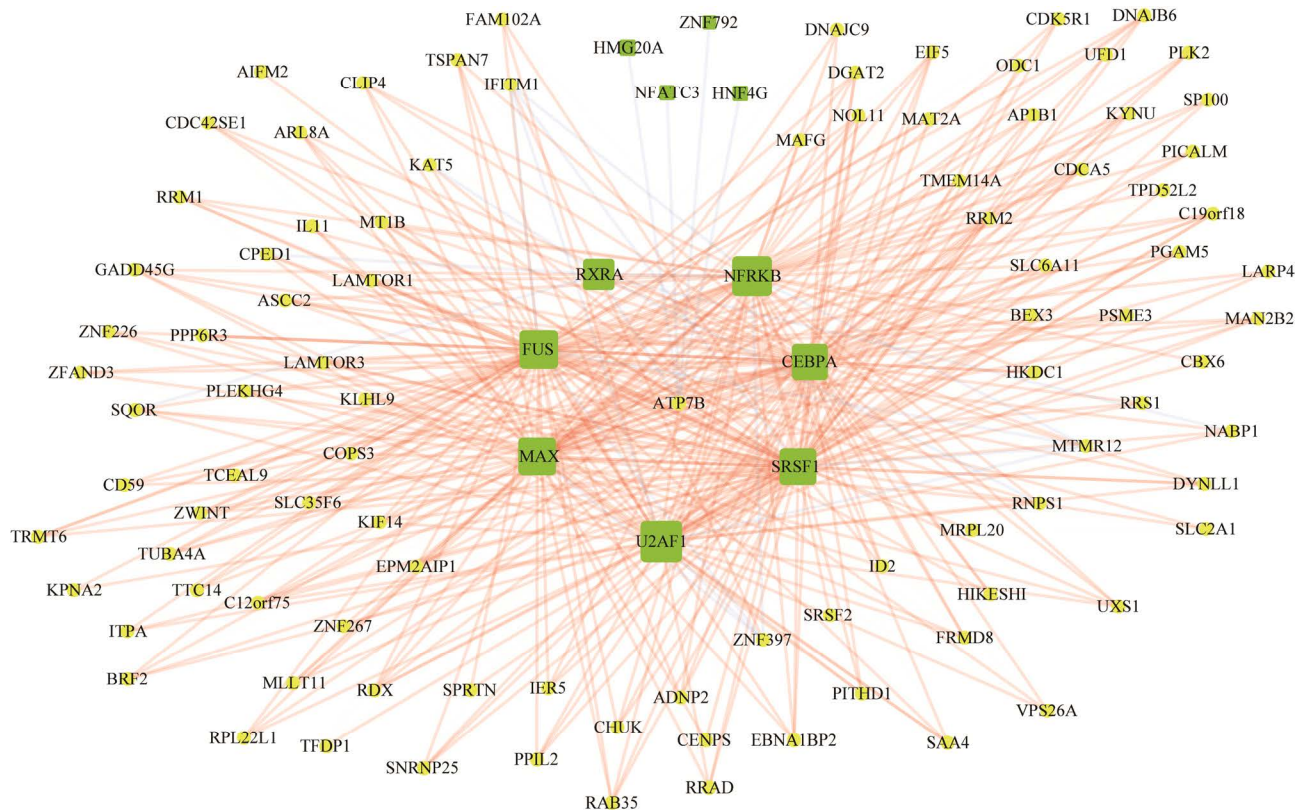


图4 差异表达转录因子和差异表达基因调控网络

Figure 4 Differentially expressed transcription factors and differentially expressed genes regulatory network. Green square represents differentially expressed transcription factors, and yellow circle represents differentially expressed genes. The size of a node represents its degree. Red lines represent up-regulated, and blue lines represent down-regulated.

异表达基因之间的547个调控关系。其中,U2AF1、NFRKB、FUS、MAX、SRSF1、CEBPA和RXRA这7个转录因子位于调控网络的中心,这些核心差异表达转录因子共同调控了大多数DEGs的表达(表3)。此外,图4显示,肝豆状核变性关键基因 $ATP7B$ 还同时受到了差异表达转录因子HMG20A、NFATC3、ZNF792和HNF4G的调节。

表3 肝豆状核变性转录调控网络中的核心转录因子  
Table 3 Core transcription factors in the transcriptional regulatory network of Wilson disease

Gene symbol	Degree of node
U2AF1	238
NFRKB	223
FUS	213
MAX	206
SPSF1	198
CEBPA	192
RXRA	152

### 3 讨论

随着系统生物学方法的发展以及相关实验数据的积累,基于转录调控网络识别复杂疾病的致病基因和基因与靶点之间的调控关系,可以为基础研究提供潜在的候选靶基因。肝豆状核变性的总体发病率约为 1/30 000,在亚洲人群中更为高发<sup>[17]</sup>。该病虽然不是常见疾病,但在亚裔人群中并不罕见。而且肝豆状核变性是为数不多的几种可以治疗的遗传病之一,因此早期诊断和治疗对于患者意义重大。本研究利用转录组测序技术和生物信息学分析方法,分析了野生型 Hep2G 细胞系和经 *ATP7B* 基因敲除的 HepG2 细胞系之间的差异表达基因,以及经铜诱导的野生型 Hep2G 细胞系和经铜诱导的 *ATP7B* 基因敲除的 HepG2 细胞系之间的差异表达基因。保留铜诱导产生的差异表达基因,并对这部分 DEGs 进行 GO 和 KEGG 富集分析。其次,以 Barabási-PPI 网络为背景,进一步分析了肝豆状核变性关键基因和关键功能模块。最后,基于差异的表达基因和转录因子构建了转录调控网络,探讨了核心差异表达转录因子对肝豆状核变性的影响以及相关表型产生的可能原因,具有一定的参考价值。

本研究对关键功能模块中的基因进行了 GO 和 KEGG 富集分析,结果显示上、下调关键功能模块共同富集在某些 GO term 和 KEGG 通路上,如乙型肝炎、MAPK 信号通路、FoxO 信号通路、T 细胞受体信号通路和神经营养信号通路上,以及 mRNA 加工、剪接、组蛋白修饰、蛋白磷酸化和细胞衰老凋亡等与细胞活动和功能相关的 GO term 上。从富集的 KEGG 通路结果来看,WD 和乙型肝炎的症状之间很可能存在某些联系。并且临床证据也表明,WD 的肝脏症状主要包括慢性隐性肝炎和急性暴发

性肝炎<sup>[2]</sup>。此外,Brady 等发现,用于治疗 WD 的铜螯合剂也可用于治疗含有 BRAF 突变的癌症,而 BRAF 激酶磷酸化会刺激 MAPK 信号通路<sup>[18]</sup>。MAPK 信号通路在肝脏炎症的凋亡中具有重要意义,并且该信号通路通过调节肝星状细胞的活化、增殖和凋亡影响肝脏纤维化的形成<sup>[19]</sup>。多项临床研究发现,WD 患者中出现大量以中枢神经系统症状为首发症状的患者<sup>[2-3,20]</sup>,而本研究中关键功能模块内的基因显著富集到了神经营养信号通路和神经变性途径等信号通路上,提示差异基因及其关联的互作用基因的整体变化或许影响了 WD 神经表型的产生。Polishchuk 等<sup>[9]</sup>探讨铜诱导对野生型和敲除型 HepG2 细胞系的影响,经 DAVID 在线工具富集分析发现与野生型细胞相比,铜诱导对 *ATP7B* 缺陷的细胞系有更强的自噬激活作用。本研究侧重于比较铜诱导后的野生型和 *ATP7B* 缺陷型细胞系之间的转录调控差异,而 clusterProfiler 包的富集结果多集中在 mRNA 加工、剪接和翻译等与转录调控相关的 GO term 上。

FUS 是一个 DNA/RNA 结合蛋白,FUS 基因突变会造成神经元细胞死亡进而引发神经退行性疾病的发生<sup>[21]</sup>。多项研究表明 FUS 基因突变与肌萎缩侧索硬化症相关,该病的主要症状表现为肌无力、肌萎缩和震颤痉挛<sup>[21-22]</sup>。临床中,存在很大一部分 WD 患者,以震颤和肌张力障碍等神经症状为最常见的神经学特征。据文献报道有多达 55% 的肝豆状核变性神经系统表型患者在确诊时有震颤症状<sup>[23]</sup>。在本研究结果中,FUS 基因作为核心差异表达转录因子之一,提示 FUS 基因和它调控的基因的正常表达可能与肝豆状核变性神经表型的产生存在某种关联。

U2AF1 和 SRSF1 同属于 U2 核糖核蛋白体



依赖型剪接体相关基因, 主要参与前体 mRNA 的剪接、加工以及调控细胞周期和凋亡。其中, *SRSF1* 基因被认为是一个潜在的原癌基因, 通过选择性剪接能够促进肝癌的发展<sup>[24]</sup>。此外, 还有研究表明 *SRSF1* 不仅可以调控 iNKT 细胞的发育分化, 还能够增强对急性肝损伤的耐受力, 减弱肝损伤程度, 在其功能方面具有至关重要的作用<sup>[25]</sup>。另外, Sanders 等的研究结果表明, 转录因子 MAX 在肝脏的发育、增殖和再生过程中都具有重要的调控作用<sup>[26]</sup>。与 kappaB 结合蛋白 (nuclear factor kappa-B, NFRKB) 相关的核因子在正常肝细胞中低/中表达, 而在肝癌细胞的细胞核中大量积累, NFRKB 的高表达提示不良的预后和较低的生存率, 该基因的功能障碍与肝癌发展密切相关<sup>[27]</sup>。

正常肝脏受到损伤时能够完全再生, 而衰老的肝脏其增殖反应则明显减弱。CEBPA 作为一种肝特异性蛋白, 研究表明它能够通过抑制细胞周期依赖性激酶从而抑制肝细胞的增殖。Iakova 等<sup>[28]</sup>通过实验证实, 衰老能够关闭 CEBPA 抑制 E2F 转录的肝脏特异性途径, 从而导致肝脏增殖反应的丧失。本研究中, 转录因子 CEBPA 的异常很可能影响了 WD 患者肝脏的再生增殖能力。另一方面, 肝星状细胞在肝纤维化中起着重要作用, CCAAT 增强结合蛋白 (CCAAT enhancer binding proteins, CEBPs), 特别是 CEBPA 控制了脂肪细胞的分化。研究表明, 通过 CEBPA 转染激活的肝星状细胞后, 能够同时上调 *RXRA* 的表达, 表明 CEBPA 和 *RXRA* 与肝脏脂肪细胞的分化和肝纤维化存在密切关联<sup>[29]</sup>。此外, *RXRA* 能够与受体基因形成异二聚体, 在配体的作用下, 对靶基因的表达进行调控<sup>[30]</sup>。WD 患者的肝细胞会出现不同程度的脂肪变性<sup>[31]</sup>, 而 Kim 等的研究结果表明, *RXRA* 等相关基因表达的改变影响了肝脂肪代

谢的改变<sup>[32]</sup>。并且, Meacham 等以肝豆状核变性小鼠为研究对象, 探讨了铜-锌平衡对小鼠肝脏的影响, 其研究结果表明, 以 *RXRA* 为代表的一系列核受体基因与 WD 的转录调控密切相关, 这些基因共同参与了细胞的分化、发育, 以及脂质、蛋白质代谢等细胞活动<sup>[8]</sup>。Meacham 等的研究结果与本研究关于 *RXRA* 作为核心转录因子调控疾病的研究结论基本一致。

*ATP7B* 基因作为肝豆状核变性的关键基因, 除了受到 7 个核心差异表达转录因子的调控外, HMG20A、NFATC3、ZNF792 和 HNF4G 也共同调节了该基因的表达变化。其中, Wooton-Kee 等<sup>[33]</sup>报道了肝豆状核变性患者由于受到过量铜的特异性和选择性影响, 导致以 *RXR*、*HNF4* 等为例的一系列核受体基因与启动子元件的结合减少, 使核受体靶基因 mRNA 表达减少的现象, 这与本研究的相关结果一致。

肝豆状核变性作为少数可以有效控制的遗传性疾病, 其病情的严重程度与 *ATP7B* 基因突变类型尚无明确定论, 主要取决于铜在体内的蓄积量和积蓄时间, 所以早期的诊疗对患者尤为重要<sup>[5]</sup>。目前, 对于转录调控如何影响肝豆状核变性的机制研究仍属于初期, 本研究利用生物信息学分析方法筛选出 7 个在肝豆状核变性转录调控网络中位于核心地位的差异表达转录因子, 并初步探讨了它们调控肝豆状核变性神经表型和肝脏表型可能的发生机制。目前, 国内外关于肝豆状核变性的研究多集中在病例报道和突变基因型的确定, 对于该病的转录调控机制和相关基因的功能研究还远远不够。未来将基于本研究结果, 收集并纳入更多的肝豆状核变性就诊患者的血液和组织样本, 进行大规模的高通量测序, 结合生物信息学分析和分子生物学功能实验, 多角度综合、详细地讨论肝豆状核变性患者的致病机制和转录调控机

制。对于本研究发现的核心转录因子在高铜状态下具体如何对 WD 患者产生的影响,仍需要进一步通过实验进行系统地验证。

## REFERENCES

- [1] Gupta A, Das S, Ray K. A glimpse into the regulation of the Wilson disease protein, ATP7B, sheds light on the complexity of mammalian apical trafficking pathways. *Metallomics*, 2018, 10(3): 378-387.
- [2] Shribman S, Poujois A, Bandmann O, et al. Wilson's disease: update on pathogenesis, biomarkers and treatments. *J Neurol Neurosurg Psychiatry*, 2021, 92(10): 1053-1061.
- [3] Chaudhry HS, Anilkumar AC. Wilson Disease [EB/OL]. (2021-08-11) [2021-11-01] <https://www.ncbi.nlm.nih.gov/books/NBK441990/#!po=90.0000>
- [4] Li XJ, Zhang W, Zhou DH, et al. Complex ATP7B mutation patterns in Wilson disease and evaluation of a yeast model for functional analysis of variants. *Hum Mutat*, 2019, 40(5): 552-565.
- [5] Ferenci P, Stremmel W, Czlonkowska A, et al. Age and sex but not ATP7B genotype effectively influence the clinical phenotype of Wilson disease. *Hepatology*, 2019, 69(4): 1464-1476.
- [6] Medici V, LaSalle JM. Genetics and epigenetic factors of Wilson disease. *Ann Transl Med*, 2019, 7(Suppl 2): S58.
- [7] Zhang J, Ma Y, Xie D, et al. Differentially expressed lncRNAs in liver tissues of TX mice with hepatolenticular degeneration. *Sci Reports*, 2021, 11: 1377.
- [8] Meacham KA, Cortés MP, Wiggins EM, et al. Altered zinc balance in the *Atp7b*<sup>-/-</sup> mouse reveals a mechanism of copper toxicity in Wilson disease. *Metallomics*, 2018, 10(11): 1595-1606.
- [9] Polishchuk EV, Merolla A, Lichtmanegger J, et al. Activation of autophagy, observed in liver tissues from patients with Wilson disease and from ATP7B-deficient animals, protects hepatocytes from copper-induced apoptosis. *Gastroenterology*, 2019, 156(4): 1173-1189.e5.
- [10] Ritchie ME, Phipson B, Wu D, et al. Limma Powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*, 2015, 43(7): e47.
- [11] Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 2009, 26(1): 139-140.
- [12] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, 2014, 15(12): 1-21.
- [13] Yu GC, Wang LG, Han YY, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*, 2012, 16(5): 284-287.
- [14] Cheng FX, Kovács IA, Barabási AL. Network-based prediction of drug combination. *Nat Commun*, 2019, 10(1): 1197.
- [15] Zheng RB, Wan CX, Mei SL, et al. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res*, 2018, 47(D1): D729-D735.
- [16] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 2003, 13(11): 2498-2504.
- [17] Huster D. Wilson disease. *Best Pract Res Clin Gastroenterol*, 2010, 24(5): 531-539.
- [18] Brady DC, Crowe MS, Turski ML, et al. Copper is required for oncogenic BRAF signalling and tumorigenesis. *Nature*, 2014, 509(7501): 492-496.
- [19] Chopra P, Kanoje V, Semwal A, et al. Therapeutic potential of inhaled p38 mitogen-activated protein kinase inhibitors for inflammatory pulmonary diseases. *Expert Opin Investig Drugs*, 2008, 17(10): 1411-1425.
- [20] Czlonkowska A, Litwin T, Chabik G. Wilson disease: neurologic features. *Handbook of clinical neurology*, 2017, 142: 101-119.
- [21] Deng JW, Wang P, Chen XP, et al. FUS interacts with ATP synthase beta subunit and induces mitochondrial unfolded protein response in cellular and animal models. *PNAS*, 2018, 115(41): E9678-E9686.
- [22] Zou ZY, Zhou ZR, Che CH, et al. Genetic epidemiology of amyotrophic lateral sclerosis: a systematic review and meta-analysis. *J Neurol Neurosurg Psychiatry*, 2017, 88(7): 540-549.
- [23] Liu ZJ, Lin HX, Wei Q, et al. Genetic spectrum and variability in Chinese patients with amyotrophic lateral sclerosis. *Aging Dis*, 2019, 10(6): 1199-1206.
- [24] Malakar P, Shilo A, Mogilevsky A, et al. Long noncoding RNA MALAT1 promotes hepatocellular carcinoma development by SRSF1 upregulation and mTOR activation. *Cancer Res*, 2017, 77(5): 1155-1167.
- [25] Liu J, You M, Yao Y, et al. SRSF1 plays a critical role

- in invariant natural killer T cell development and function. *Cell Mol Immunol*, 2021, 18: 2502-2515.
- [26] Sanders JA, Gruppuso PA. Coordinated regulation of c-Myc and Max in rat liver development. *Am J Physiol Gastrointest Liver Physiol*, 2006, 290(1): G145-G155.
- [27] Peng QY, Zhou MQ, Zuo SR, et al. Nuclear factor related to kappaB binding protein (NFRKB) is a telomere-associated protein and involved in liver cancer development. *DNA Cell Biol*, 2021, 40(10): 1298-1307.
- [28] Iakova P, Awad SS, Timchenko NA. Aging reduces proliferative capacities of liver by switching pathways of C/EBPalpha growth arrest. *Cell*, 2003, 113(4): 495-506.
- [29] Huang GC, Zhang JS, Tang QQ. Involvement of C/EBP-alpha gene in *in vitro* activation of rat hepatic stellate cells. *Biochem Biophys Res Commun*, 2004, 324(4): 1309-1318.
- [30] Fadel L, Reho B, Volkó J, et al. Agonist binding directs dynamic competition among nuclear receptors for heterodimerization with retinoid X receptor. *J Biol Chem*, 2020, 295(29): 10045-10061.
- [31] Guindi M. Wilson disease. *Semin Diagn Pathol*, 2019, 36(6): 415-422.
- [32] Kim MS, Sweeney TR, Shigenaga JK, et al. Tumor necrosis factor and interleukin 1 decrease RXRalpha, PPARalpha, PPARgamma, LXRalpha, and the coactivators SRC-1, PGC-1alpha, and PGC-1beta in liver cells. *Metabolism*, 2007, 56(2): 267-279.
- [33] Wooton-Kee CR, Jain AK, Wagner M, et al. Elevated copper impairs hepatic nuclear receptor function in Wilson's disease. *J Clin Invest*, 2015, 125(9): 3449-3460.

(本文责编 郝丽芳)