


## Research Article

# Gene Expression Profiling and Biofunction Analysis of HepG2 Cells Targeted by Crocetin

Yi-Ling Wen,<sup>1</sup> Yong Li,<sup>1</sup> Guangcheng Zhu,<sup>1</sup> Zhibing Zheng,<sup>1</sup> Meng Shi,<sup>1</sup> and Si Qin<sup>1,2</sup> 

<sup>1</sup>Lab of Food Function and Nutrigenomics, College of Food Science and Technology, Hunan Agricultural University, Changsha 410128, China

<sup>2</sup>Department of Biochemical Science and Technology, Faculty of Agriculture, Kagoshima University, Korimoto 1-21-24, Kagoshima 890-0065, Japan

Correspondence should be addressed to Si Qin; [qinsiman@hunau.edu.cn](mailto:qinsiman@hunau.edu.cn)

Received 11 February 2021; Revised 3 March 2021; Accepted 17 March 2021; Published 2 April 2021

Academic Editor: Xin Zong

Copyright © 2021 Yi-Ling Wen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Crocetin is a carotenoid extracted from *Gardenia jasminoides*, one of the most popular traditional Chinese medicines, which has been used in the prevention and treatment of various diseases. The present study is aimed at clarifying the effect of crocetin on gene expression profiling of HepG2 cells by RNA-sequence assay and further investigating the molecular mechanism underlying the multiple biofunctions of crocetin based on bioinformatics analysis and molecular evidence. Among a total 23K differential genes identified, crocetin treatment upregulated the signals of 491 genes (2.14% of total gene probes) and downregulated the signals of 283 genes (1.24% of total gene probes) by  $\geq 2$ -fold. The Gene Ontology analysis enriched these genes mainly on cell proliferation and apoptosis (BRD4 and DAXX); lipid formation (EHMT2); cell response to growth factor stimulation (CYP24A1 and GCNT2); and growth factor binding (ABCB1 and ABCG1), metabolism, and signal transduction processes. The KEGG pathway analysis revealed that crocetin has the potential to regulate transcriptional misregulation, ABC transporters, bile secretion, alcoholism, systemic lupus erythematosus (SLE), and other pathways, of which SLE was the most significantly disturbed pathway. The PPI network was constructed by using the STRING online protein interaction database and Cytoscape software, and 21 core proteins were obtained. RT-qPCR datasets serve as the solid evidence that verified the accuracy of transcriptome sequencing results with the same change trend. This study provides first-hand data for comprehensively understanding crocetin targeting on hepatic metabolism and its multiple biofunctions.

## 1. Introduction

Crocetin (*Gardenia jasminoides*) is a carotenoid originally extracted from a medicinal and edible resource and is one of the most popular traditional Chinese medicines. Crocetin has been used in the prevention and treatment of various diseases with health-promoting effects such as antioxidation, antitumor, anticardiovascular illness, anti-inflammation, hepatoprotection, and neuroprotection [1]. These therapeutic effects are achieved with different mechanisms. Accumulated data showed that crocetin could scavenge free radicals by inhibiting lipid peroxidation and increasing the activity of antioxidant enzymes including GST, GSH-Px, catalase, glutathione peroxidase (GPx), and superoxide dismutase

(SOD) [2, 3]. Tumor invasion was suppressed by crocetin through reducing MMP expression and tumor proliferation by inducing cell cycle arrest, and enhancing apoptosis by activating caspase 3 [4, 5]. Cardiovascular disorders such as hypertension, thrombosis formation, and myocardial infarction (MI) could be potentially prevented or treated by crocetin. Moreover, crocetin effectively prevented inflammatory markers such as NF- $\kappa$ B, TNF- $\alpha$ , IL-1 $\beta$ , IL-6, and IL-8 and inhibited iNOS/COX-2 activity and induction of NO/PGE2 [6–8]. The main mechanisms of the protective effects of crocetin in liver and nerve damage are mediated by the restoration of antioxidant enzyme activity and inhibition of the superoxide anion and/or free radical [9–11]. The nutritional value and therapeutic potential of crocetin have been

gradually revealed, and it is therefore necessary to comprehensively and deeply analyze the molecular mechanism of the action of crocetin, and then fully reveal its functional activity.

Nutrigenomics is an emerging research field that studies the genomic changes caused by diet, covering the interaction of health, diet, and genomics. The human diet is composed of a complex mixture of substances and biological activity. It is functionally divided into three aspects: some can directly affect gene expression, some can modulate transcription factor activity after metabolism, while some can induce transcription after stimulating the signal transduction cascade [12]. Transcriptome refers to the sum of RNA transcribed by cells, tissues, or organisms at a certain developmental stage or under physiological conditions, mainly composed of messenger RNA (mRNA), noncoding RNA (ncRNA), and microRNA (miRNA). Transcriptomics is a discipline that systematically studies the global transcription map of the whole genome and reveals the network of molecular mechanisms of complex biological processes and trait regulation. A variety of techniques, including EST sequence construction and study, gene microarray, gene expression series analysis (SAGE), large-scale parallel sequencing (MPSS), and RNA-seq, have been developed to study the transcriptome. RNA-seq technology is established on the basis of EST technology and gene microarray. It is different from the gene expression microarray method in the sense that RNA-seq can not only detect transcripts corresponding to the existing genome sequence, but it can also find and quantify new transcripts, which have more advantages in the study of selective splicing events, new genes and transcripts, and fusion transcripts. RNA-seq has become the most popular method for transcriptome studies due to its low cost, its high throughput, and its applicability to species with no genomic background.

Recently, in the field of nutrition, there have been many studies on organisms, cells, or tissues under different nutritional states based on RNA-seq technology. Studies have reported that dietary composition can affect gene expression, thus affecting biological processes and pathways. Therefore, changing the dietary composition and proportion of farm animals could improve the nutritional value of meat products [13–15]. Suarez-Vega et al. [16] used the RNA-seq technique to study the MED-related mammary responses induced by CLA in breast milk, and compared the transcriptome data of mammary cells in ewes with the CLA-induced MFD group, the non-MED group, and the FO-MFD group, and identified the core genes involved in the downregulation of fatty acid synthesis in CLA-MFD and FO-MFD. Today, RNA-seq remains the preferred method of analysis for transcriptomes, and nutritionists are able to use transcriptome data to develop a framework for improving dietary composition and novel functional foods as a way to promote human health.

To study the effect of crocetin on the whole genome expression profile of human hepatocytes, RNA-seq technology was performed in HepG2 cells with or without crocetin treatment. The GO and KEGG pathway analysis has been applied to investigate the multiple biofunctions of crocetin.

RT-qPCR was further used to verify the gene expression data. This study demonstrated the potential of using of crocetin as a functional ingredient applicable in medicine and food.

## 2. Materials and Methods

**2.1. Materials, Cell Culture, and Cytotoxicity.** Crocetin was provided by Tairui Biotechnology Co., Ltd. (Nanyang, Henan, China). LPS (*Escherichia coli* serotype 055:B5) was from Sigma-Aldrich (St. Louis, USA). Human hepatoblastoma HepG2 cells were purchased from ATCC (Rockville, USA) and cultured at 37°C in a 5% CO<sub>2</sub> atmosphere in Dulbecco's modified Eagle's medium (DMEM; Gibco BRL, Gaithersburg, USA) containing 10% FBS. Fetal bovine serum (FBS) was from Biological Industries (Kibbutz Beit Haemek, Israel). 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) was purchased from Sigma-Aldrich (St. Louis, USA). MTT assay was used to check the cytotoxicity of crocetin. Briefly, HepG2 cells were seeded into a 96-well plate at a density of 10<sup>4</sup> cells per well and preincubated at 37°C in 5% CO<sub>2</sub> for 24 h. The cells were treated with a series of crocetin concentrations (0, 10 μM, 20 μM, and 40 μM) for another 24 h, then MTT (final concentration 0.5 mg/ml) was added to the plate and incubated for 4 h in the dark. Acidic isopropanol (0.04–0.1 M HCl in isopropanol) was added to dissolve the formazan crystals, and the optical density (OD) was measured at 570 nm with a microplate reader (Thermo Fisher Scientific, Waltham, USA). Cell viability was measured by comparing the OD of crocetin-treated cells with that of untreated cells.

**2.2. RNA Extraction and Construction of cDNA Library.** HepG2 cells were precultured for 24 h and then were treated with 20 μM crocetin dissolved in 0.1% DMSO for 8 h. Total RNA was extracted using the Isogen RNA kit (Nippon Gene Co., Tokyo, Japan) according to the manufacturer's protocol. The concentration and quality of RNA were assessed by a NanoDrop Spectrophotometer and an Agilent 2200 Bioanalyzer following the manufacturer's protocol. The cDNA of the purified mRNA was transformed and amplified by PCR technology, and the PCR products were purified by AMPure XP beads to complete the construction of the cDNA library. After the library construction was completed, the library was initially quantified (Qubit 2.0), then the insert fragment size of the library was detected using Agilent 2100, and the qualified cDNA libraries were sequenced on an Illumina HiSeq 2500 platform for high-throughput sequencing.

**2.3. Bioinformatics Analysis.** The R language package DESeq2 was used to screen differentially expressed genes. The significant difference standard for differentially expressed genes in this study is  $|\log_2 \text{fold change}| > 0$  and  $p \text{ value} < 0.05$ , where the fold change value of the gene is the difference fold change, and the higher the value, the greater the difference in the expression of the gene between the two groups; the  $p$  value is the probability of hypothesis testing in the statistical model, and the smaller the  $p$  value, the greater the difference in expression between samples. DAVID is an online, web-based bioinformatics network analysis platform, which can

TABLE 1: The primers used for real-time PCR.

Oligonucleotide	Sequence (5'-3')
FGFBP1-F	CTTCACAGCAAAGTGGTCTCA
FGFBP1-R	GACACAGGAAAATTCATGGTCCA
GCNT2-F	TGTTCTGGCTCTATGCCAAA
GCNT2-R	TTAGCAAACAGGCTTGGTGAAT
GPR35-F	GGGAGGACCGTCTGCACAAA
GPR35-R	CCCAGGTGGCTGAATCTGGTG
TNFRSF21-F	GCCAGTGAGAGGGAGGTTGC
TNFRSF21-R	TCCAGCTGGGTGGTGTCTTC
TGFB1-F	GCGTCTGCTGAGGCTCAAGT
TGFB1-R	GCCGGTAGTGAACCCGTTGAT
TCF19-F	GGGGCGGTGATCTCTACAC
TCF19-R	GGGAGTCGGACATTATTGACCA
PCGF2-F	GCGAGGTCTTGAGCAGGAG
PCGF2-R	GGCGATGTCCATGAGGGTGT
GAPDH-F	CATGGCACCGTCAAGGCTGA
GAPDH-R	ACGTACTCAGCGCCAGCATC
ATG2A-F	CTCGCCTCCTCCAGATCAA
ATG2A-R	GGGCATCCTGGTCCACATTG
FOSL1-F	CAGGCGGAGACTGACAAACTG
FOSL1-R	TCCTTCCGGGATTTTGCAGAT
SOCSS-F	GTGCCACAGAAATCCCTCAAA
SOCSS-R	TCTCTCGTGCAAGTCTTGTC

realize fast, accurate, and comprehensive functional annotation analysis of large-scale gene or protein lists. GO and KEGG analysis methods are used to perform functional analysis on the differentially expressed genes, and to screen out differentially expressed genes enriched in pathways related to crocetin treatment.

**2.4. Construction of Protein Interaction Networks.** Functions and interactions between proteins were analyzed by the protein-protein interaction (PPI) system. The STRING database is the largest repository of its kind, retrieving millions of discoveries from full-text literature and is updated weekly. Using the Cytoscape software to construct the PPI network, setting the protein-protein interaction score greater than 0.7 has statistical significance. CytoHubba and MCODE were used to determine the core proteins in the complex protein interaction network, and clustering was used to construct functional modules.

**2.5. Real-Time PCR.** The genes with obvious differential expression were selected, respectively, and the corresponding sequence files were obtained, and the amplification primers of each gene were designed across exons using the Primer 5 software. The primer sequences were designed and applied for gene expression identification as shown in Table 1. Subsequently, these genes were verified by real-time fluorescent quantitative PCR (RT-qPCR). According to manufacturer's manual, the DyNAmo™ SYBR® Green 2-Step qRT-PCR Kit (Finnzymes Oy, Espoo, Finland) was used for reverse

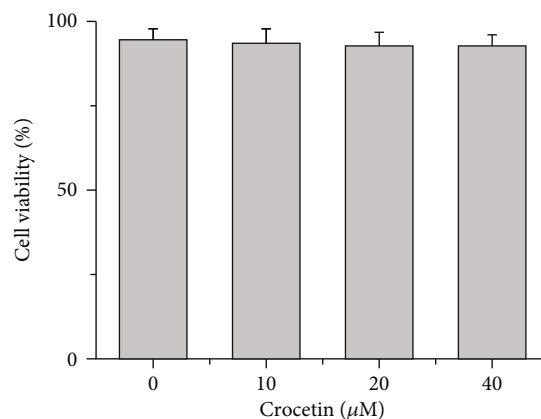


FIGURE 1: The cytotoxicity assay of HepG2 cells treated by crocetin (\* $p < 0.05$ ).

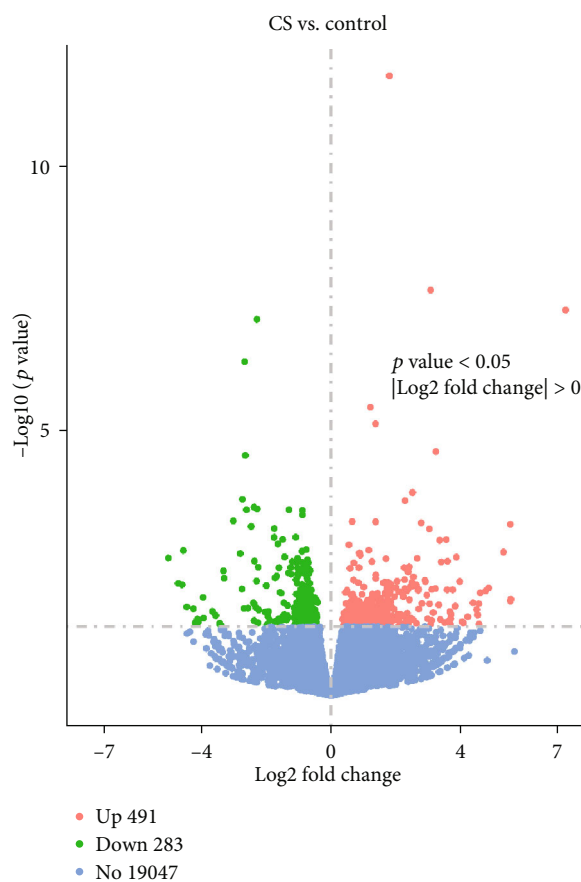


FIGURE 2: Volcano map of differentially expressed genes.

transcription and real-time PCR operations. Specifically, Oligo dT was used to reverse transcribe RNA (200 ng) into cDNA. M-MuLV RNase was applied at 37°C for 30 minutes, and then the reaction was terminated at 85°C for 5 minutes. PCR was determined according to the sequence of each primer with the  $T_m$  value. Each PCR contains 250 ng reverse transcript, 75 ng primer, and 10  $\mu$ l master mix. The thermal cycling conditions are maintained at 95°C for 15 min, and then at 94°C for 55

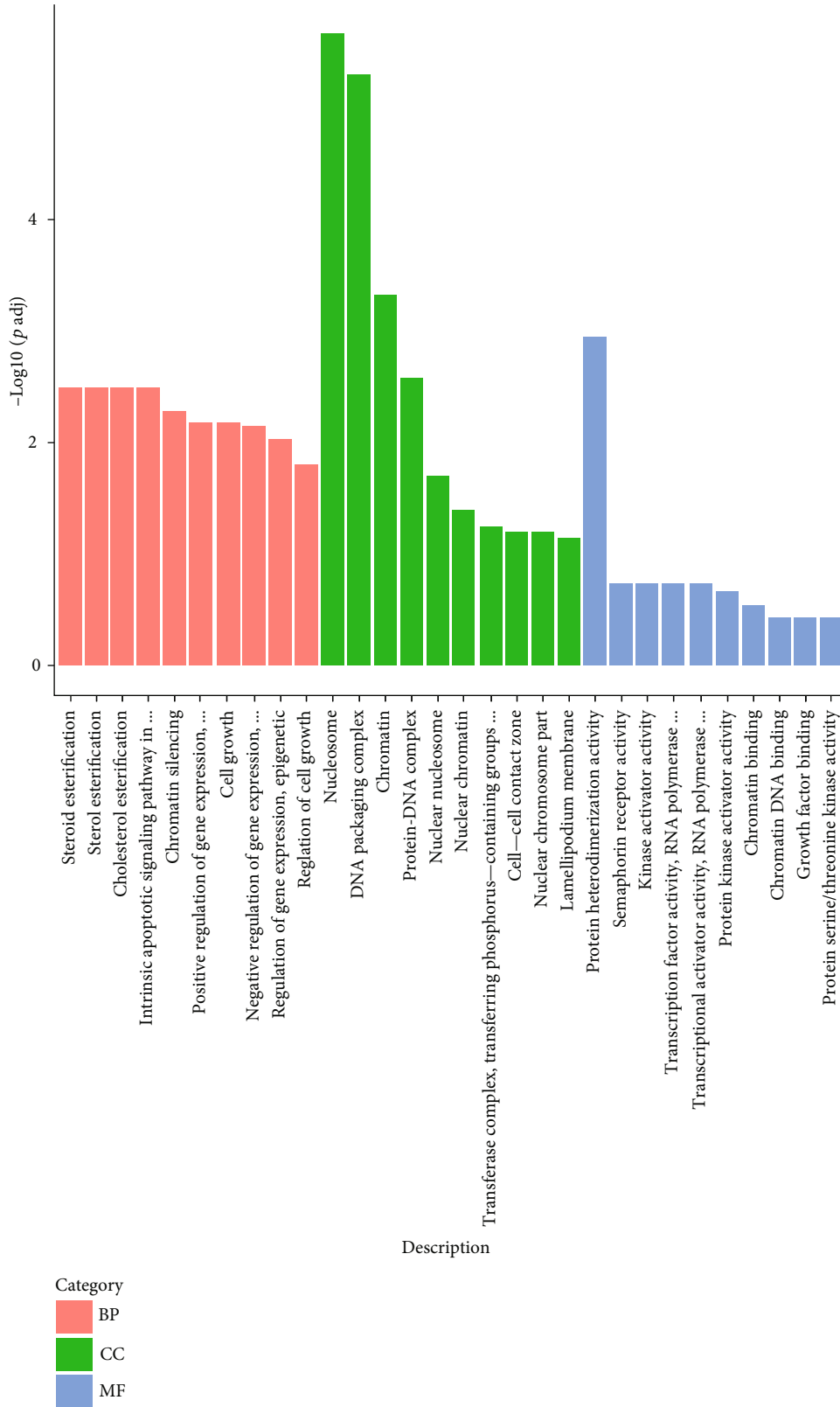


FIGURE 3: GO analysis of differentially expressed genes. GO analysis includes biological process (BP) (red), cell component (CC) (green), and molecular function (MF) (blue).

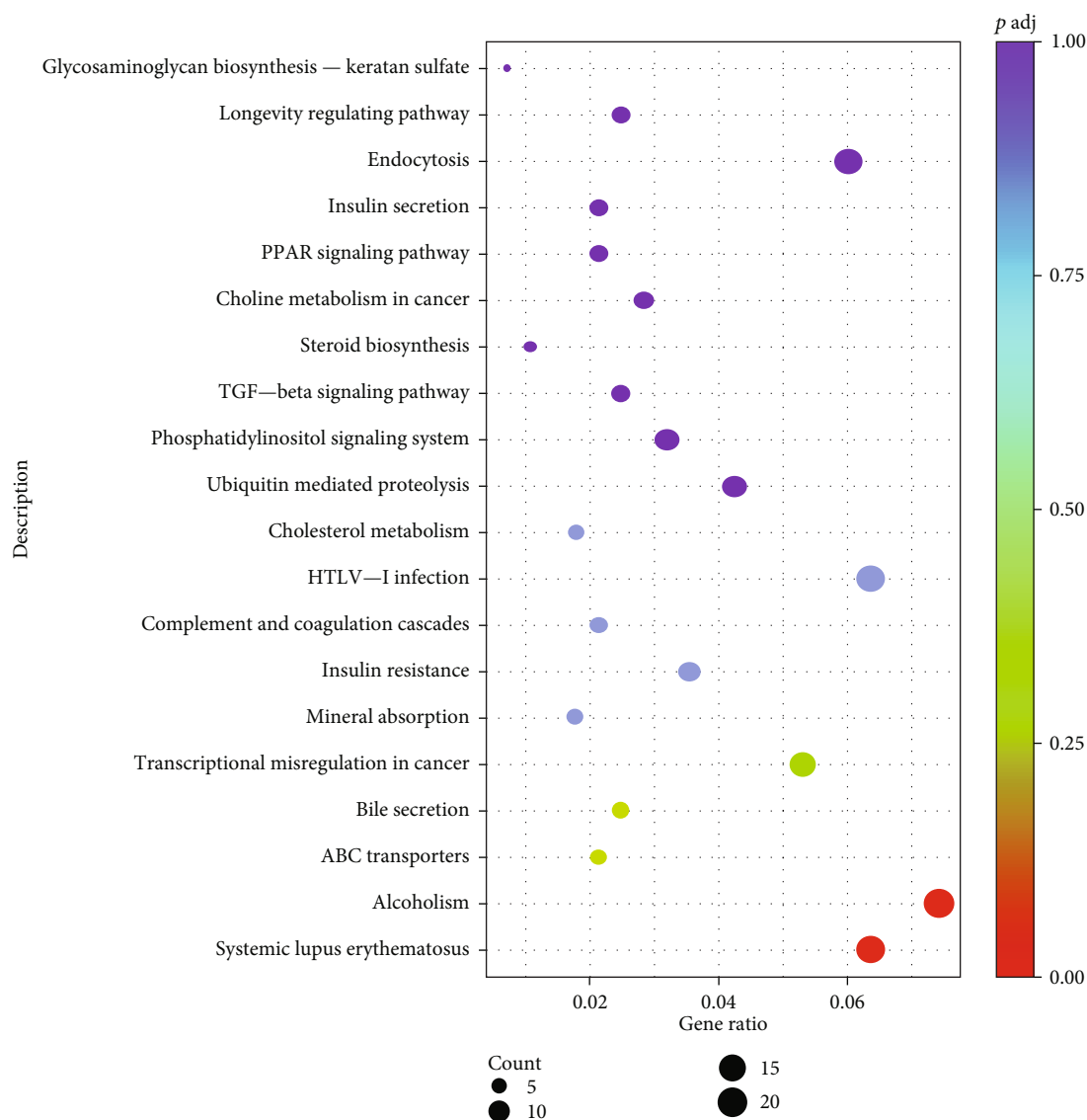


FIGURE 4: KEGG analysis of differentially expressed genes.

cycles of 30 s, with a  $T_m$  value of 30 s (melting temperature) and 30 sec at 72°C in a Rotor-Gene-3000AKAA (Corbett Research Pty., NSW, Australia).

**2.6. Statistical Analysis.** All the experimental data were expressed as mean  $\pm$  SD. Statistical significance was analyzed by Student's  $t$ -test and ANOVA. A statistical probability of  $p < 0.05$  was considered statistically significant.

### 3. Results

**3.1. Gene Expression Profiling.** HepG2 cells were treated with diverse concentrations of crocetin, and the cytotoxicity was detected using a CCK-8 kit according the manufacturer's instructions. As shown in Figure 1, compared with the blank group, when the concentration of crocetin reached 40  $\mu\text{M}$ , the number of living cells had no significant change with the survival rate of cells above 90%, indicating that there

was no toxicity to the cells under the concentration of crocetin in the range of 0–40  $\mu\text{M}$ .

Among the total 45K genes, crocetin can affect the expressions of 19,821 genes among the 22,580 annotated genes in HepG2 cells; the disturbed fold of each gene is listed in Supplementary Table 1. As shown in Figure 2, the volcano map reflects the expression status of all the genes, in which green and red represent the downregulated and upregulated genes, respectively, and blue represents the genes with no differentially expressed genes. The total number of differential genes obtained from the experimental group and the control group was 774, among which 491 were upregulated genes and 283 were downregulated genes. The volcano map comparison showed that our method had sufficient coverage to detect differences in gene expression in the CS group relative to the control group.

Among the differentially expressed genes after crocetin treatment, the differentially expressed multiple of 6 genes was greater than or equal to 5, among which 5 genes were



upregulated and 1 gene was downregulated. The differential expression multiple of 19 genes was greater than or equal to 4, among which 11 genes were upregulated and 8 genes were downregulated. There were 39 genes with differential expression ratios greater than or equal to 3, among which 28 genes were upregulated and 11 genes were downregulated. The differential expression ratio of 71 genes was greater than or equal to 2, among which 49 genes were upregulated and 22 genes were downregulated. In general, from the cells treated with crocetin, there were 135 genes (0.68%, 135 : 19821) with a differential expression ratio higher than or equal to 2 times.

**3.2. Gene Ontology and KEGG Pathway Analysis.** As shown in Figure 3, after CS treatment (control vs. CS), nucleosome and DNA packaging complex-related components were the most significantly disturbed in the cell component group (CC in green), and chromatin and protein-DNA complex-related processes were the most enriched differential genes; the numbers of differential genes were 44 and 22, respectively. From the molecular function group (MF in blue), the most significantly disturbed was the activity of the protein heterodimer, in which the number of differentially enriched genes was 42. The most significantly disturbed in the biological process group (BP in red) was sterol esterification, in which the numbers of differentially enriched genes related to cell growth and gene epigenetic regulation were 40 and 34, respectively.

In order to further understand the functions and pathways of differential genes, the software Clusterprofiler was used to conduct KEGG pathway enrichment analysis of differential gene sets, and the top 20 pathways with the highest enrichment significance were selected for mapping. As shown in Figure 4, after CS treatment, 283 differential genes were compared with 251 KEGG pathways, which were related to transcriptional misregulation, ABC transporters, bile secretion, alcoholism, systemic lupus erythematosus, and other pathways. Among them, systemic lupus erythematosus (SLE) was the most significantly disturbed pathway, which is the first time it was implied that the crocetin has a potent and promising effect on this serious autoimmune disease.

**3.3. PPI Network Analysis and Core Disturbed Gene Screening.** The PPI network belongs to a scale-free network, which is not uniform. Among the PPI networks, most of the nodes have only one or two connections while a few nodes have lots of connections, ensuring that the system is fully connected. Nodes with a high number of connections, in this network, become hubs, which play an irreplaceable role in biological evolution and in maintaining the stability of the interaction network. These nodes normally have very key biological functions and participate in important life activities. The results show that the more interactions a protein interacts with, the more important is its role for the survival of the cell.

In order to understand the function of crocetin-induced differentially expressed genes and screen out Hub genes/proteins, the STRING database was used to analyze the interactions on the PPI network relationship. The list of

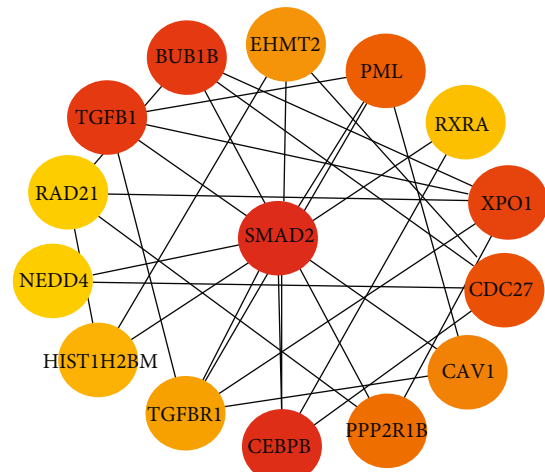


FIGURE 5: The 15 genes with the highest parameter values obtained by calculation are used as hub genes. The color change of the node indicates the increase or decrease of the parameter value.

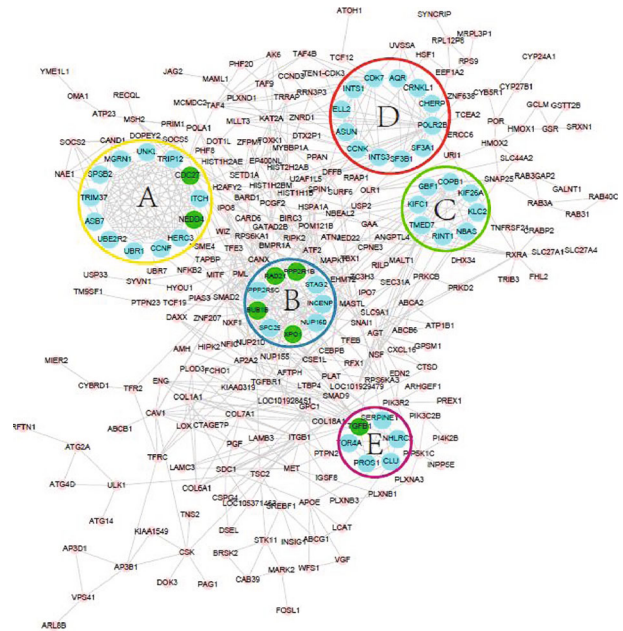


FIGURE 6: The five most important modules in the PPI network. According to the standard with a K-core greater than 6, the five important clusters obtained from formic acid in the PPI network are labeled A, B, C, D, and E, respectively, and are considered to be the topological center of the PPI network. The nodes marked in red represent the upregulated genes participating in the 5 subnetworks (none), and the nodes marked in green represent the downregulated genes.

differentially expressed genes was imported into the STRING database, and the data obtained from experiments, literatures, and high-throughput evidences were combined, and then the protein interaction data set encoded by the differentially expressed genes was downloaded. In the PPI network of differentially expressed genes treated with crocetin, 359 differentially expressed genes that encoded proteins were

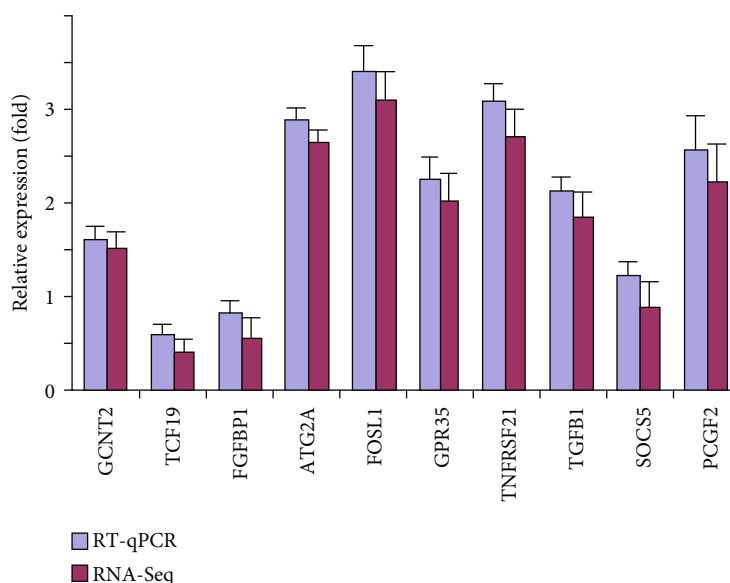


FIGURE 7: Comparison of differential gene expressions obtained by real-time PCR and RNA-seq sequencing results. Error bars show standard deviation.

screened with relatively close interactions, including 359 protein nodes and 761 correlations.

However, it is difficult to visually identify the key nodes in the network due to the scale-free network nature of most biological data networks. CytoHubba plug-ins in the Cytoscape software were then used to calculate the Hub genes/proteins in the network diagram. According to the order of parameter and radiality from high to low as the standard, a total of 15 core proteins were screened, which includes the SMAD family variant 2 (SMAD2), CCAAT enhancer binding protein beta (CEBPB), conversion factor 1 (TGFB1), mitotic checkpoint serine/threonine protein kinase B (BUB1B), output 1 (XPO1), cell division cycle protein 27 (CDC27), early young granulocyte leukemia variant 1 (PML) transcription, protein phosphatase 2 support subunits X5 (PPP2R1B), cellular protein 1 (CAV1), autosomal histone lysine methylation transferase (EHMT2) 2, transforming growth factor beta receptor 1 (TGFBR1), histone (Hist1H2BM), retinol X receptor  $\alpha$  (RXRA), adhesive protein complex (Rad21), and E3 ubiquitin protein ligase (Nedd4) (Figure 5). SMAD proteins encoded by their respective genes are signal transducers and transcriptional modulators that mediate multiple signaling pathways [17, 18]. This protein mediates the signal of the transforming growth factor- (TGF-) beta, and thus regulates multiple cellular processes, such as cell proliferation, apoptosis, and differentiation [19, 20]. SMAD proteins are recruited to the TGF-beta receptors through its interaction with the SMAD anchor for receptor activation (SARA) protein, and further phosphorylated by the TGF-beta receptors [21].

For further understanding the important module of PPI, the Cytoscape MCODE software was used to screen out fine modules with K-core values greater than 6, and mark the above 15 hub genes in each module; we can see that the 15 hub genes were mostly in the five important modules selected as shown in Figure 6.

**3.4. Regulation of the Expressions of Typical Genes by Crocetin Treatment.** Drug metabolic enzymes are composed of phase I and phase II metabolic enzymes and phase III transporters, which play an important role in the metabolism, digestion, and detoxification of foreign substances and drugs. The biological functions of crocetin in the human body, such as anti-cancer, antioxidant, and anti-inflammatory, have been found, but its effect on liver cells, especially on hepatic drug metabolism enzymes and transporters, are still unclear. In the present study, genes with a differential expression ratio greater than or equal to 1.5 were selected as significantly different genes. Among the total 445 metabolic enzyme genes, 28 genes were disturbed by crocetin, with 20 genes upregulated and 8 genes downregulated. For instance, *GCNT2*, *ABCB1*, *ABCG1*, *ABCA2*, *CYP24A1*, *CYP27B1*, and *AHRR* are significantly upregulated genes by crocetin treatment, which implied that crocetin stimulated the expressions of hepatic metabolic enzymes.

In order to know the effect of crocetin on inflammation in the liver cells, several typical inflammatory differentially expressed genes were selected for analysis, and it was found that 199 genes in the total 455 inflammatory genes were significantly disturbed by crocetin treatment, with 158 upregulated genes and 41 downregulated genes. Among them, *GRP35*, *AGT*, *PDE5A*, *MMP28*, *IFIT2*, *TCF19*, *FGFBP1*, *SOCS5*, *TGFBI*, *PCGF2*, *FGFRL1*, *TNFRSF21*, and *VGF* are the most significantly downregulated genes by crocetin treatment, which revealed that crocetin is a potent and promising anti-inflammatory substance. Besides, glycolipid metabolism and other biofunction genes were also selected for analysis, and it was found that many glycolipid metabolism-related genes, 48 from the total of 445 genes, such as *EHMT2*, *AGT*, *RYR3*, *HBPI*, *GCNT2*, *BMT2*, *HSF1*, *NSF*, *TSC2*, *SP2*, *FA2H*, *CAD*, and *APOE*, were the most disturbed genes by crocetin treatment, which indicated that crocetin may exert its multiple biofunction by regulating glycolipid metabolism.

Finally, to verify the accuracy of the transcriptome data, several typical genes are selected by RT-qPCR verification. As shown in Figure 7, the RT-qPCR data showed the same trends with the results of transcriptome, suggesting that RNA-sequencing is believable.

#### 4. Discussion and Conclusion

The present study investigated the effect of crocetin on gene expression profiling of HepG2 cells by RNA-sequence assay and further analyzed the molecular mechanism underlying its multiple biofunction based on bioinformatics analysis and molecular evidence. The 774 differentially expressed genes in the transcription process were annotated by GO and KEGG analysis. In the three categories of the GO analysis catalog, the differentially expressed genes were mainly concentrated in nucleosomes, protein-DNA complexes, transcription factors, lipids, transferases, transcription factors, and other processes in the CC group. When treated with MF, different genes were mainly concentrated in the activities of signal transduction receptors, transcription factors, protein kinases, growth factors, transformation factors, etc. Differential genes were mainly concentrated in the MP group such as esterification, apoptosis, chromatin or gene silencing, and cell response to growth factor stimulation.

Through the PPI network analysis of the interaction between the coding proteins of different genes, the interaction between multiple differentially expressed genes was enriched. Further, 15 core proteins in the network were screened by the Cytoscape software, among which the differentially expressed genes corresponding to the core proteins of PPP2R1b, Rad21, Sub1b, XPO1, and TGFB1 were significantly downregulated. Among them, PPP2R1B is a tumor suppressor gene that encodes the  $\beta$  subtype of the serine/threonine-specific protein phosphatase 2A (PP2A-A $\beta$ ) A subunit, which is inactivated in cancer patients [22]. RAD21 is involved in the regulation of the normal cell cycle, DNA, DSB repair, and apoptosis pathways. In addition, the germline heterozygous or homozygous missense mutations of RAD21 are related to human genetic diseases [23]. XPO1 is an export receptor responsible for the nuclear-cytoplasmic transport of hundreds of proteins and various RNAs. XPO1 is often overexpressed and/or mutated in human cancers, and acts as a carcinogenic driver. Therefore, inhibiting XPO1-mediated nuclear export is a unique treatment strategy [24]. Subsequently, the PPI network of the modules was screened, and 5 important modules were obtained; most of these core proteins were in these 5 important modules.

To be noted, it was found that among the total drug metabolism enzymes (445 genes), 28 genes were significantly disturbed by crocetin treatment, with 20 genes upregulated and 8 genes downregulated, indicating that crocetin could positively influence the processes related to drug metabolism enzymes in human liver cells. Among them, *GCNT2* [25], *ABCBI* [26, 27], *ABCG1* [28, 29], *ABCA2* [30], *CYP24A1* [31, 32], *CYP27B1* [33], and *AHRR* [34] were the core genes of drug metabolism enzymes, which were worthy of further verification. Abnormal cellu-

lar proliferation and apoptosis are typical features of tumor cells, and genetic variations in the study pathway enrichment are mainly involved in transcription, intrinsic apoptotic signaling pathways, gene silencing, miRNA regulation gene silencing, and cell cycle regulation. Of these, 62 and 35 genes directly regulated the cell proliferation and apoptosis processes, respectively, and it is speculated that saffron acid could promote cell proliferation and apoptosis. In particular, *BRD4* [35, 36], *FAM83H* [37, 38], and *FOXK1* [39–42] are reported to be involved in the cell proliferation process of human pancreatic cancer and lung cancer, respectively, while *DAX* [43], *PML* [44], and *ULK1* [45] are reported to be involved in the cell apoptosis process of breast cancer, leukemia, and smooth muscle cells, respectively.

The above results indicate that crocetin is mostly positively associated with drug metabolism enzymes, transporters, inflammatory factors, and glucose and lipid metabolism, as well as cell proliferation and apoptosis processes in HepG2 cells. And the core genes were further detected by an RT-PCR experiment, which was consistent with the results of transcriptome sequencing, indicating that the results of transcriptome sequencing were real and reliable.

To conclude, our transcriptome data revealed gene expression profiles of crocetin in HepG2 cells for the first time. Signaling pathway analysis further demonstrated that systemic lupus erythematosus is involved in crocetin-induced expressions of most hepatic drug metabolizing enzyme genes. These results provide a comprehensive data for understanding the hepatic metabolism, the bioactive role, and the molecular mechanisms of crocetin.

#### Data Availability

The transcriptome data used to support the findings of this study are included within the supplementary information file.

#### Conflicts of Interest

We have no direct or indirect commercial financial incentive associated with publishing this article.

#### Authors' Contributions

Yi-Ling Wen and Yong Li contributed equally to this work.

#### Acknowledgments

This work was partially supported by grants from the National Key Research and Development Program of China (2019YFC1604903), the Natural Science Foundation of Hunan Province (2019JJ40132), and the Double First-Class Construction Project of Hunan Agricultural University (SYL201802025) to Si Qin.



## Supplementary Materials

Supplementary description: raw data of the transcriptome assay in HepG2 cells. (*Supplementary Materials*)

## References

- [1] A. R. Khorasany and H. Hosseinzadeh, "Therapeutic effects of saffron (*Crocus sativus* L.) in digestive disorders: a review," *Iranian Journal of Basic Medical Sciences*, vol. 19, no. 5, pp. 455–469, 2016.
- [2] A. S. Ahmad, M. A. Ansari, M. Ahmad et al., "Neuroprotection by crocetin in a hemi-parkinsonian rat model," *Pharmacology, Biochemistry, and Behavior*, vol. 81, no. 4, pp. 805–813, 2005.
- [3] M. B. Khan, M. N. Hoda, T. Ishrat et al., "Neuroprotective efficacy of *Nardostachys jatamansi* and crocetin in conjunction with selenium in cognitive impairment," *Neurological Sciences*, vol. 33, no. 5, pp. 1011–1020, 2012.
- [4] D. Chryssanthi, P. Dedes, N. Karamanos, P. Cordopatis, and F. Lamari, "Crocetin inhibits invasiveness of MDA-MB-231 breast cancer cells via downregulation of matrix metalloproteinases," *Planta Medica*, vol. 77, no. 2, pp. 146–151, 2011.
- [5] K. He, P. Si, H. Wang et al., "Crocetin induces apoptosis of BGC-823 human gastric cancer cells," *Molecular Medicine Reports*, vol. 9, no. 2, pp. 521–526, 2014.
- [6] J. Yan, Z. Qian, L. Sheng et al., "Effect of crocetin on blood pressure restoration and synthesis of inflammatory mediators in heart after hemorrhagic shock in anesthetized rats," *Shock*, vol. 33, no. 1, pp. 83–87, 2010.
- [7] B. Chen, Z. H. Hou, Z. Dong, and C. D. Li, "Crocetin downregulates the proinflammatory cytokines in methylcholanthrene-induced rodent tumor model and inhibits COX-2 expression in cervical cancer cells," *BioMed Research International*, vol. 2015, Article ID 829513, 5 pages, 2015.
- [8] K. N. Nam, Y. M. Park, H. J. Jung et al., "Anti-inflammatory effects of crocin and crocetin in rat brain microglial cells," *European Journal of Pharmacology*, vol. 648, no. 1–3, pp. 110–116, 2010.
- [9] M. Hashemi and H. Hosseinzadeh, "A comprehensive review on biological activities and toxicology of crocetin," *Food and Chemical Toxicology*, vol. 130, pp. 44–60, 2019.
- [10] S. Higashino, Y. Sasaki, J. C. Giddings et al., "Crocetin, a Carotenoid from *Gardenia jasminoides* Ellis, protects against hypertension and cerebral thrombogenesis in stroke-prone spontaneously hypertensive rats," *Phytotherapy Research*, vol. 28, no. 9, pp. 1315–1319, 2014.
- [11] P. Chen, Y. Chen, Y. Wang et al., "Comparative evaluation of hepatoprotective activities of geniposide, crocins and crocetin by CCl<sub>4</sub>-induced liver injury in mice," *Biomolecules & Therapeutics*, vol. 24, no. 2, pp. 156–162, 2016.
- [12] C. Carlberg, "Nutrigenomics of Vitamin D," *Nutrients*, vol. 11, no. 3, p. 676, 2019.
- [13] M. S. Ferreira, L. A. Tomaz, M. B. Niehues et al., "The inclusion of de-oiled wet distillers grains in feedlot diets reduces the expression of lipogenic genes and fat content in Longissimus muscle from F1 Angus-Nellore cattle," *PeerJ*, vol. 7, article e7699, 2019.
- [14] M. Yu, Z. Li, W. Chen, T. Rong, G. Wang, and X. Ma, "*Hermetia illucens* larvae as a potential dietary protein source altered the microbiota and modulated mucosal immune status in the colon of finishing pigs," *Journal of Animal Science and Biotechnology*, vol. 10, no. 4, article 50, 2019.
- [15] M. Yu, Z. Li, W. Chen, G. Wang, Y. Cui, and X. Ma, "Dietary supplementation with citrus extract altered the intestinal microbiota and microbial metabolite profiles and enhanced the mucosal immune homeostasis in yellow-feathered broilers," *Frontiers in Microbiology*, vol. 10, article 2662, 2019.
- [16] A. Suárez-Vega, B. Gutiérrez-Gil, P. G. Toral, G. Hervás, J. J. Arranz, and P. Frutos, "Conjugated linoleic acid (CLA)-induced milk fat depression: application of RNA-Seq technology to elucidate mammary gene regulation in dairy ewes," *Scientific Reports*, vol. 9, article 4473, 2019.
- [17] S. Zagabathina, R. Ramadoss, A. H. Harini Priya, and R. Krishnan, "Comparative evaluation of SMAD-2 expression in oral submucous fibrosis and reactive oral lesions," *Asian Pacific Journal of Cancer Prevention*, vol. 21, no. 2, pp. 399–403, 2020.
- [18] D. Kamato, M. Burch, Y. Zhou et al., "Individual Smad2 linker region phosphorylation sites determine the expression of proteoglycan and glycosaminoglycan synthesizing genes," *Cellular Signalling*, vol. 53, pp. 365–373, 2018.
- [19] C. R. Gignoux, D. G. Torgerson, M. Pino-Yanes et al., "An admixture mapping meta-analysis implicates genetic variation at 18q21 with asthma susceptibility in Latinos," *Journal of Allergy and Clinical Immunology*, vol. 142, no. 2, pp. 957–969, 2018.
- [20] J. L. Granadillo, W. K. Chung, L. Hecht et al., "Variable cardiovascular phenotypes associated with SMAD2 pathogenic variants," *Human Mutation*, vol. 39, no. 12, pp. 1875–1884, 2018.
- [21] E. Cannaeerts, M. Kempers, A. Maugeri et al., "Novel pathogenic SMAD2 variants in five families with arterial aneurysm and dissection: further delineation of the phenotype," *Journal of Medical Genetics*, vol. 56, no. 4, 2018.
- [22] M. Tamaki, T. Goi, Y. Hirono, K. Katayama, and A. Yamaguchi, "PPP2R1B gene alterations inhibit interaction of PP2A-A $\beta$  and PP2A-C proteins in colorectal cancers," *Oncology Reports*, vol. 11, pp. 655–659, 2004.
- [23] H. Cheng, N. Zhang, and D. J. G. Pati, "Cohesin subunit RAD21: from biology to disease," *Gene*, vol. 758, article 144966, 2020.
- [24] N. G. Azizian and Y. Li, "XPO1-dependent nuclear export as a target for cancer therapy," *Journal of Hematology & Oncology*, vol. 13, no. 1, 2020.
- [25] C. Li and Q. Wu, "Adaptive evolution of multiple-variable exons and structural diversity of drug-metabolizing enzymes," *BMC Evolutionary Biology*, vol. 7, no. 1, p. 69, 2007.
- [26] A. van Helvoort, A. J. Smith, H. Sprong et al., "MDR1 P-glycoprotein is a lipid translocase of broad specificity, while MDR3 P-glycoprotein specifically translocates phosphatidylcholine," *Cell*, vol. 87, no. 3, pp. 507–517, 1996.
- [27] K. Choi, C. J. Chen, M. Krieger, and I. B. Roninson, "An altered pattern of cross-resistance in multidrug-resistant human cells results from spontaneous mutations in the mdr1 (P-glycoprotein) gene," *Cell*, vol. 53, no. 4, pp. 519–529, 1988.
- [28] E. A. Roundhill, S. Jabri, and S. A. Burchill, "ABCG1 and Pgp identify drug resistant, self-renewing osteosarcoma cells," *Cancer Letters*, vol. 453, pp. 142–157, 2019.
- [29] T. Engel, F. Kannenberg, M. Fobker et al., "Expression of ATP binding cassette-transporter ABCG1 prevents cell death by transporting cytotoxic 7 $\beta$ -hydroxycholesterol," *FEBS Letters*, vol. 581, no. 8, pp. 1673–1680, 2007.

- [30] J. Mack, D. Townsend, V. Beljanski, and K. Tew, "The ABCA2 transporter: intracellular roles in trafficking and metabolism of LDL-derived cholesterol and sterol-related compounds," *Current Drug Metabolism*, vol. 8, no. 1, pp. 47–57, 2007.
- [31] K. Yasuda, E. Tohyama, M. Takano et al., "Metabolism of 2 $\alpha$ -[2-(tetrazol-2-yl)ethyl]-1 $\alpha$ ,25-dihydroxyvitamin D(3) by CYP24A1 and biological activity of its 24R-hydroxylated metabolite," *The Journal of Steroid Biochemistry and Molecular Biology*, vol. 178, pp. 333–339, 2018.
- [32] E. W. Tieu, E. K. Y. Tang, and R. C. Tuckey, "Kinetic analysis of human CYP24A1 metabolism of vitamin D via the C24-oxidation pathway," *FEBS Journal*, vol. 281, no. 14, pp. 3280–3296, 2014.
- [33] M. B. Meyer and J. W. Pike, "Mechanistic homeostasis of vitamin D metabolism in the kidney through reciprocal modulation of Cyp27b1 and Cyp24a1 expression," *The Journal of Steroid Biochemistry and Molecular Biology*, vol. 196, p. 105500, 2020.
- [34] T. Haarmann-Stemmann and J. Abel, "The arylhydrocarbon receptor repressor (AhRR): structure, expression, and function," *Biological Chemistry*, vol. 387, no. 9, pp. 1195–1199, 2006.
- [35] Y. F. Tan, M. Wang, Z. Y. Chen, L. Wang, and X. H. Liu, "Inhibition of BRD4 prevents proliferation and epithelial-mesenchymal transition in renal cell carcinoma via NLRP3 inflammasome-induced pyroptosis," *Cell Death & Disease*, vol. 11, no. 4, p. 239, 2020.
- [36] M. Yin, Y. Guo, R. Hu et al., "Potent BRD4 inhibitor suppresses cancer cell-macrophage interaction," *Nature Communications*, vol. 11, no. 1, p. 1833, 2020.
- [37] Y. K. Ma, T. H. Shen, and X. Y. Yang, "Upregulation of lncRNA FAM83H-AS1 in hepatocellular carcinoma promotes cell proliferation, migration and invasion by Wnt/ $\beta$ -catenin pathway," *European Review for Medical and Pharmacological Sciences*, vol. 23, no. 18, pp. 7855–7862, 2019.
- [38] H. Shan, Y. Yang, X. Zhu, X. Han, P. Zhang, and X. Zhang, "FAM83H-AS1 is associated with clinical progression and modulates cell proliferation, migration, and invasion in bladder cancer," *Journal of Cellular Biochemistry*, vol. 120, no. 3, pp. 4687–4693, 2019.
- [39] J. Li, S. Huang, Y. Zhang, W. Zhuo, B. Tong, and F. Cai, "LINC00460 enhances bladder carcinoma cell proliferation and migration by modulating miR-612/FOXK1 axis," *Pharmacology*, vol. 106, pp. 79–90, 2020.
- [40] L. Li, M. Gong, Y. Zhao, X. Zhao, and Q. Li, "FOXK1 facilitates cell proliferation through regulating the expression of p21, and promotes metastasis in ovarian cancer," *Oncotarget*, vol. 8, no. 41, pp. 70441–70451, 2017.
- [41] M. Wencong, W. Jinghan, Y. Yong et al., "FOXK1 promotes proliferation and metastasis of gallbladder cancer by activating AKT/mTOR signaling pathway," *Frontiers in Oncology*, vol. 10, p. 545, 2020.
- [42] H. Xu, S. Huang, X. Zhu, W. Zhang, and X. Zhang, "FOXK1 promotes glioblastoma proliferation and metastasis through activation of Snail transcription," *Experimental and Therapeutic Medicine*, vol. 15, no. 3, pp. 3108–3116, 2018.
- [43] M. Lanzino, P. Maris, R. Sirianni et al., "DAX-1, as an androgen-target gene, inhibits aromatase expression: a novel mechanism blocking estrogen-dependent breast cancer cell proliferation," *Cell Death & Disease*, vol. 4, no. 7, article e724, 2013.
- [44] L. Zhang, Q. S. Chen, P. P. Xu et al., "Catechins induced acute promyelocytic leukemia cell apoptosis and triggered PML-RAR $\alpha$  oncoprotein degradation," *Journal of Hematology & Oncology*, vol. 7, no. 1, p. 75, 2014.
- [45] D. Gui, Z. Cui, L. Zhang et al., "Salidroside attenuates hypoxia-induced pulmonary arterial smooth muscle cell proliferation and apoptosis resistance by upregulating autophagy through the AMPK-mTOR-ULK1 pathway," *BMC Pulmonary Medicine*, vol. 17, no. 1, p. 191, 2017.