

## Article

# Network Pharmacology-Based Approach Combined with Bioinformatic Analytics to Elucidate the Potential of Curcumol against Hepatocellular Carcinoma

Xufeng Huang <sup>1,†</sup>, Hafiz Muzzammel Rehman <sup>2,3,†</sup>, Attila Gábor Szöllösi <sup>4,\*</sup>, Shujing Zhou <sup>5</sup><sup>1</sup> Faculty of Dentistry, University of Debrecen, Debrecen 4032, Hungary; huangxufeng@mailbox.unideb.hu<sup>2</sup> Alnoorians Group of Institutes 55-Elahi Bukhsh Park, Amir Road, Shad Bagh, Lahore 54000, Pakistan; Muz-zammel.phd.ibb@pu.edu.pk<sup>3</sup> School of Biochemistry and Biotechnology, University of the Punjab, Lahore, Punjab 54590, Pakistan; Muz-zammel.phd.ibb@pu.edu.pk<sup>4</sup> Department of Immunology, University of Debrecen, Debrecen 4032, Hungary; szollosi.attila@med.unideb.hu<sup>5</sup> Faculty of Medicine, University of Debrecen, Debrecen 4032, Hungary; shujing.zhou@mailbox.unideb.hu

\* Correspondence: szollosi.attila@med.unideb.hu

† These authors contributed equally to this work

**Abstract:** Purpose: Modern, open-source databases provide an unprecedented wealth of information to help drug development. By combining data available in these databases with the proper bioinformatic tools, we can elucidate the molecular targets of natural compounds. One such molecule is curcumol, a guaiane-type sesquiterpenoid hemiketal isolated from *Rhizoma Curcumae*, which is used for a broad range of diseases in traditional Chinese and Indian medicine. It has been reported to exert anti-tumor activity, but the intrinsic molecular mechanism in hepatocellular carcinoma (HCC) is unclear. Therefore, the present study was designed to reveal the predictive targets and biological mechanisms of curcumol against HCC via a network pharmacology-based approach combined with bioinformatic analytics and to provide proof of concept for further similar investigations. Methods: Data available from open-source databases (Traditional Chinese Medicine Systems Pharmacology, Comparative Toxicogenomic Database, The Cancer Genome Atlas, the Human Protein Atlas project) was processed with the help of a variety of open-source tools (SwissADME, SwissTargetPrediction, JVenn, Gene Ontology, Kyoto Encyclopedia of Genes and Genomes, GENMANIA, Cytoscape). Results: In the present study, the potential of curcumol against HCC was unraveled by network pharmacology-based elucidation. It suggests that curcumol shows exciting druggability with 44 potent homo sapiens biotargets against HCC. The GO terms and KEGG pathways enrichment analyses, curcumol-targets-pathways-HCC network, PPI network, and corresponding in-depth topological analyses, as well as survival analysis, molecular docking simulation indicate that the potential mechanism of curcumol against HCC is complicated, as it may act in various ways, mainly by inducing apoptosis and modulating the inflammatory response, increasing presentation of HCC-specific protein. Conclusion: The present study highlights the potential of curcumol against HCC, giving reference to further experimental study. It also presents a roadmap that can be followed to conduct in silico prescreening of other compounds of interest.

**Keywords:** network modeling; drug discovery; data integration

**Citation:** Lastname, F.; Lastname, F.; Lastname, F. Title. *Genes* **2022**, *13*, x. <https://doi.org/10.3390/xxxxx>

Academic Editor: Yijie Ding and Quan Zou

Received: 01 March 2022

Accepted: 28 March 2022

Published: date

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer that occurs most often in people with long-term liver diseases, especially cirrhosis caused by hepatitis viral infection [1].

Natural products from traditional herbal medicines are precious sources for new drug discovery. Curcumol, as one of the typical examples, is a bioactive extract of essential volatile oils from *Curcuma* Rhizoma [2], which has been used for centuries as an essential ingredient in traditional herbal formulas in the treatments of various diseases (e.g., blood stasis, alleviating pain, and more) [3]. Current studies suggest that curcumol exerts tremendous therapeutic effects against tumors, reducing inflammation, and protecting the liver [4–8]. These findings hint that curcumol may serve as a potential new drug against HCC. However, this possibility is inadequately studied; therefore, the present study attempts to make a critical first step in this field via a network pharmacology-based approach.

We first conducted an “absorption, distribution, metabolism, and excretion” (ADME) evaluation by using the Traditional Chinese Medicine Systems Pharmacology (TCMSP) database and SwissADME analysis platform to investigate the druggability of curcumol at the molecular level.

Then, putative targets of curcumol were mined from Comparative Toxicogenomic Database (CTD) and SwissTargetPrediction server. Meanwhile, we obtained the therapeutic targets of HCC from The Cancer Genome Atlas (TCGA) cohort and Genotype-Tissue Expression (GTEx) portal, GeneCards database, and Online Mendelian Inheritance in Man (OMIM) database. An intersection containing all the common targets of the putative targets of curcumol and the therapeutic targets of HCC was calculated by the JVenn program.

Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analyses were carried out by WEB-based GENE SeT AnaLysis Toolkit (WebGestalt) for these common targets to gain a macroscopic understanding of their functions.

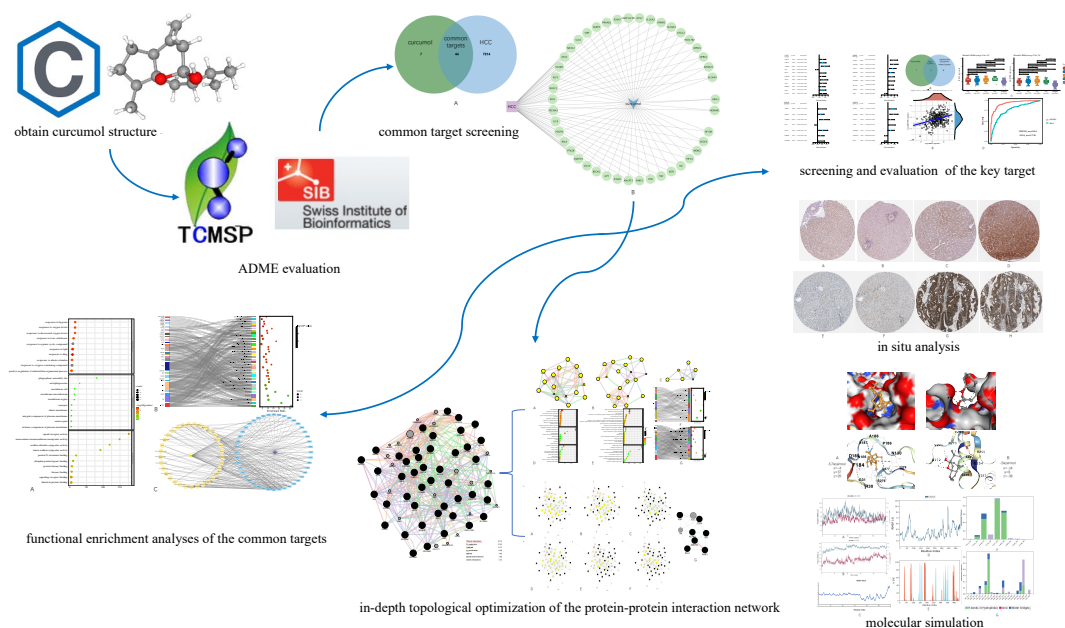
Moreover, a preliminary protein-protein interaction (PPI) network was constructed by the GeneMANIA plug-in of Cytoscape software. To have a deeper insight into the network, a cluster classification of the preliminary PPI network was performed by the MCODE plug-in of Cytoscape software. Each cluster was followed by GO terms and KEGG pathways analyses. In addition, by using the CytoNCA plug-in of Cytoscape software, a core PPI subnetwork with 9 key nodes and 22 edges was extracted.

Moreover, a survival analysis was also performed for the common targets, where we found 17 significantly prognosis-related genes from various types of survival.

By using the JVenn program again, we sorted out two key targets (i.e., RELA and SQSTM1) from the core PPI subnetwork that are significantly prognosis related. We then, from the aspects of cancer grading and Spearman correlation analysis, provided a brief summary of RELA and SQSTM1 to show their importance in HCC. Furthermore, with the help of the Human Protein Atlas project, we clinically verified their importance in HCC patients by comparing the expression in normal liver tissue and pathological liver tissue. Consequently, it was found that they themselves are positively correlated, and their over-expression is associated with HCC carcinogenesis.

Finally, molecular docking and molecular dynamic simulation were performed to investigate their binding to the curcumol molecule.

Figure 1 below demonstrates the workflow of the present study in a graphical manner.



**Figure 1.** Workflow of the present study in a graphical manner.

## 2. Materials and Methods

### 2.1. ADME Evaluation

SwissADME server (<http://www.swissadme.ch/>, 22/11/2021) is an online tool that offers structure based ADME parameter prediction. In the present study, given the three-dimensional structure of the curcuminol molecule, it was first used to roughly assess its druggability. Then, we searched with the keyword “Curcuminol” on the TCMSP database (<https://tcmsp-e.com>, 22/11/2021) to further confirm the pharmacological properties in exact values in a literature-based manner [9–12].

### 2.2. Common Targets Mining

#### 2.2.1. Identification of the Putative Targets of Curcuminol

The Comparative Toxicogenomic Database (CTD; <http://ctdbase.org/>, 22/11/2021) is a publicly available digital ecosystem that relates toxicological information on chemicals, genes, phenotypes, and diseases, to literature data [13]. In the present study, the putative targets of curcuminol were first sorted out of the literature pool and then manually filtered to ensure they were homo sapiens biotargets.

SwissTargetPrediction (<http://www.swisstargetprediction.ch/>, 22/11/2021) is a structure-based target predictor [10–12]. The three-dimensional structure of curcuminol was downloaded from the PubChem database (<https://pubchem.ncbi.nlm.nih.gov/>, 22/11/2021) [14,15] and input into the searching box.

#### 2.2.2. Identification of the Therapeutic Targets of HCC

Raw counts of RNA-sequencing data (level 3) and corresponding clinical information were obtained from the TCGA cohort (<https://portal.gdc.cancer.gov/>, 22/11/2021) [16]. Normal healthy tissue data were obtained from the GTEx portal (<https://gtexportal.org/home/>, 22/11/2021) [17] as part of the control group. All the mentioned data were analyzed by R packages “limma” in R foundation (v.4.0.3) on SciHub (<https://www.aclbi.com>, 22/11/2021) for statistical computing [18]. In addition, the GeneCards database (<https://www.genecards.org/>, 22/11/2021) [19] and OMIM database (<https://www.omim.org/>, 22/11/2021) [20] were also searched.

#### 2.2.3. Common Target Screening

The JVenn program (<http://jvenn.toulouse.inra.fr/app/index.html>, 22/11/2021) is a JavaScript-based program that can exert a Venn diagram of up to 6 data sets [21]. In the present study, it was run to calculate the intersection of the data sets obtained from steps 2.1 and 2.2. Then, we visualized the expression of the distribution of each target in HCC by integrating the data acquired from the TCGA cohort and elucidated the regulation of the expression of common targets in HCC at the gene set level by conducting gene set enrichment analysis (GSEA) in Gene Set Cancer Analysis (GSCA) online platform (<http://bioinfo.life.hust.edu.cn/GSCA/>, 22/11/2021) [22].

### 2.3. Functional Enrichment Analyses and Curcumol-Targets-Hcc-Pathways Network Construction

To gain a macroscopic understanding of the functions of common targets, we carried out GO terms and KEGG pathway enrichment analyses by WebGestalt (<http://www.webgestalt.org/>, 22/11/2021) [23]. The results were then further visualized by the SRPlot online toolkit (<http://www.bioinformatics.com.cn/en>, 22/11/2021). Consequently, by using Cytoscape software, we constructed a curcumol-targets-HCC-pathways network [24].

### 2.4. Ppi Network Construction and In-Depth Structural Optimization

#### 2.4.1. Preliminary PPI Network Construction

The GeneMANIA plug-in of Cytoscape software is a robust application that can analyze and prioritize functional assays [25]. Given query genes, GeneMANIA will list proteins with shared properties and similar functions as a PPI network. In the present study, the common targets were typed into the search bar, and the homo sapiens option was selected in the filter button.

#### 2.4.2. Functional Clustering

After obtaining the preliminary PPI network, MCODE, a Cytoscape plug-in, was used to cluster the protein network [26]. The parameter “K-Core” value was set to be 3, and the other parameters were the same as default (i.e., degree cutoff: 2, cluster finding: haircut, node score cutoff: 0.2, max. depth: 100). Each cluster creates a subnetwork to display separately, followed by GO terms and KEGG pathways enrichment analyses.

#### 2.4.3. Core PPI Subnetwork Extraction

CytoNCA, another Cytoscape plug-in that is capable of network centrality analysis, was used to extract the core PPI subnetwork via in-depth topological analyses [27]. We first set the filtering criteria to be above the average of closeness, betweenness, network, eigenvector, LAC, and degree centralities of the preliminary network to extract the corresponding subnetworks, respectively. Then, we used the intersectional merge function of the Cytoscape software to isolate the core PPI subnetworks from them. Thereafter, we sorted out the key nodes in the preliminary PPI network (i.e., the nodes in the core PPI subnetwork).

### 2.5. Survival Analyses

To identify the targets that were prognosis related in HCC, a total of 44 common targets underwent survival analysis, including overall survival (OS), progression-free survival (PFS), disease-free survival (DFS), and disease-specific survival (DSS) through R packages “ggplot2” and “survival” on Scihub [28,29]. The P values and hazard ratio (HR) with a 95% confidence interval (CI) were generated by log-rank tests and univariate COX proportional hazards regression. The results were plotted as forest plots and Kaplan–Meier (KM) curves.

### 2.6. Screening of the Key Targets and Exploration of Their Expression and Correlation in Hcc

By using JVenn program to calculate the intersection of the 9 key nodes from step 4.3 and the 17 significantly prognosis-related genes from step 5, 2 key targets (i.e., RELA and SQSTM1) were found.

In order to unravel the importance of the key targets, we conducted Kruskal–Wallis test to compare their expression levels in various cancer grades (i.e., G1, G2, G3, G4) and normal tissues. Afterward, we performed a receiver operative characteristic (ROC) analysis to see if RELA and SQSTM1 are eligible to be the biomarker in HCC. Tumoral RNA-seq data were downloaded from the TCGA cohort, while data of normal tissue samples were from the GTEx portal. We also used the Spearman method to describe the correlation between RELA and SQSTM1. All the statistical analyses were performed by R packages “ggstatsplot” and “pROC” on SciHub, where  $p < 0.05$  was considered statistically significant [30,31].

### 2.7. In situ Analysis

At this step, by using the Human Protein Atlas project (<https://www.proteinatlas.org/>, 26/11/2021), we inspected the histological slides of HCC patients and normal individuals [32]. Selected slides that are demonstrated in the present study include 2 normal liver tissue samples and 2 pathological liver tissue samples for both identified markers (HPA068843 and CAB004587 were used to stain SQSTM1, while CAB004264 and CAB005030 were used to stain RELA). Images are available at the following URL: <https://www.proteinatlas.org/ENSG00000173039-RELA> and <https://www.proteinatlas.org/ENSG00000161011-SQSTM1>.

### 2.8. Molecular Docking of Key Targets

To simulate the interaction of the key targets with the curcumol molecule, structures of the key targets from PDB (RELA ID: 1nfi, <https://www.rcsb.org/structure/1NFI>, 28/11/2021) were acquired and prepared in PyMOL (v1.2r3pre) by removing water molecules, adding charge, and parameterizing for molecular docking on AutoDock Vina-based platform, CBdock (<http://clab.labshare.cn/cb-dock/php/>, 28/11/2021) [33–36]. In this way, we obtained their corresponding affinity energy values and specific three-dimensional coordinates. Notably, as there was no entire SQSTM1 structural file available, we used an AlphaFold document (SQSTM ID: AF-Q13501-F1-model, <https://alphafold.ebi.ac.uk/entry/Q13501>, 28/11/2021) for docking [37,38].

### 2.9. Molecular Dynamic Simulation

Molecular dynamics simulations were performed to investigate the binding conformational stability of the protein-ligand complex. The stability of the protein-ligand complex was maintained during 100 ns simulations using Desmond software for compounds based on root mean square deviation (RMSD), root mean square fluctuation (RMSF), and hydrogen bond interactions [39,40].

## 3. Results

### 3.1. ADME Evaluation

During ADME evaluation, multiple pharmacological parameters are discussed (e.g., MW, Hdon, Hacc, RBN, OB, etc.). MW (molecular weight) refers to the mass of a molecule, Hdon and Hacc refer to hydrogen bond donor and receptor, AlogP value represents the partition coefficient between octanol and water [41]. They are the parameters of Lipinski's “rule of five”, which is the most common empirical rule to assess if a chemical is suitable to be developed into orally delivered drugs. RBN is the number of free self-rotationable bonds (i.e., any single bond, not in a ring, bound to a nonterminal heavy atom, excluding the C–N bond) [42]. It reflects the flexibility of the molecule. A molecule with a 10 or even smaller RBN value is predicted to have suitable oral bioavailability. OB (oral bioavailability) is the percentage of the dose of a drug that is orally administered and reaches the

systemic circulation system without changing any of it [43]. A high OB value is often the key to determining whether a bioactive molecule is suitable to serve as an orally administered therapeutic agent. DL (drug-likeness) is a qualitative concept to estimate how “drug-like” a prospective compound is [44]. Caco2 permeability (Caco2) is the passive diffusion of drugs across intestinal epithelium (i.e., to be more exactly speaking, in the case of human intestinal cell line “Caco2”), which represents the efficiency of drug intake in the gut [45]. BBB is an abbreviation of “blood-brain barrier”. Here, its value is a quantified tool for researchers to measure how well a molecule of interest can penetrate the blood-brain barrier. Usually,  $BBB < -0.3$  is non-penetrating (BBB-),  $-0.3 < BBB < +0.3$  is moderate penetrating (BBB±), and  $BBB > 0.3$  is strong penetrating (BBB+) [46]. TPSA is the abbreviation of “the polar surface area” of a molecule. It is defined as the surface sum over all polar atoms, including their attached hydrogen atoms [47].  $TPSA > 140 \text{ \AA}^2$  usually means the molecule is poor at penetrating the cell membrane, while if  $TPSA < 60 \text{ \AA}^2$ , usually it is suitable at penetrating the cell membrane.

As an official recommendation by the TCMSP database, suggested drug screening criteria usually should be  $RBN < 10$ ,  $OB \geq 20\%$ ,  $DL \geq 0.1$ , and  $TPSA < 60 \text{ \AA}^2$ . From this aspect, as the curcumol molecule possesses the following parameters:  $OB = 103.55\%$ ,  $DL = 0.13$ ,  $TPSA = 0.25 \text{ \AA}^2$ ,  $RBN = 1$ , it's deemed to possess exciting druggability (detailed ADME values are available at Table 1).

As a supplementary assessment, the “Lipinski’s rule of five” (i.e., suitable absorption is more likely when  $MW < 500\text{Da}$ ,  $AlogP < 5$ ,  $Hdon < 5$ ,  $Hacc < 10$ ), was applied [48]. According to this rule, as the curcumol molecule possesses the following parameters:  $MW = 236.39$ ,  $AlogP = 2.79$ ,  $Hdon = 1$ ,  $Hacc = 2$ , it is considered as a promising candidate for future drug optimization.

Name	MW	Hdon	Hacc	AlogP	RBN	OB	Caco2	BBB	DL	TPSA
Curcumol	236.39Da	1	2	2.79	1	103.55%	1.12 nm/s	1.23	0.13	0.25Å <sup>2</sup>

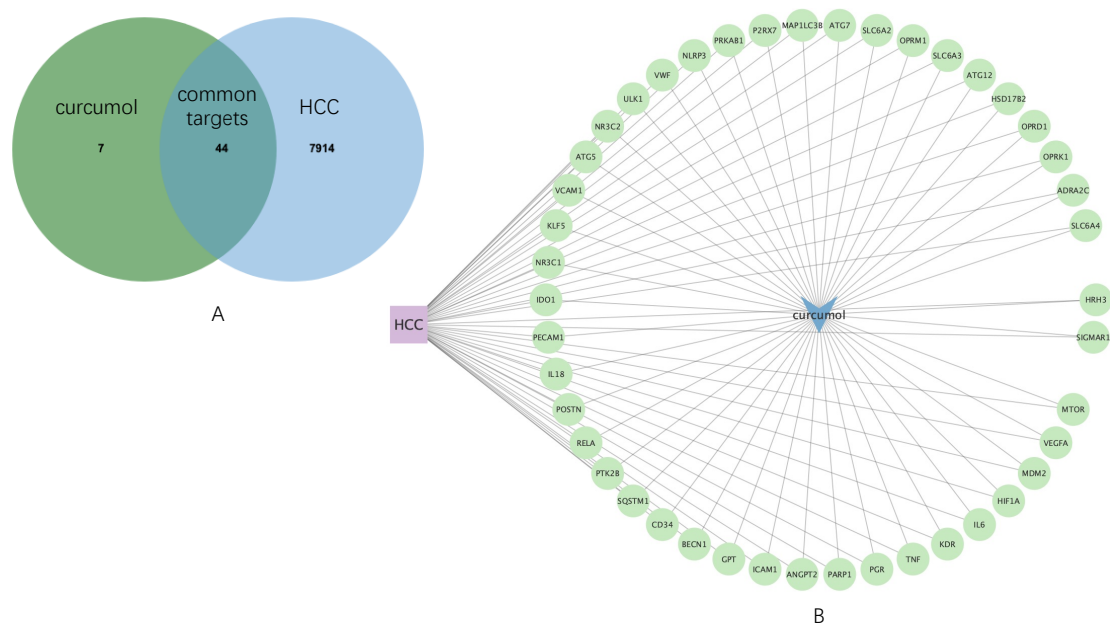
**Table 1.** Pharmacological and molecular properties of curcumol. MW: molecular weight; Hdon: hydrogen donor; Hacc: hydrogen acceptor; AlogP: partition coefficient of octanol to water; RBN: number of bonds that can perform free rotation; OB: oral bioavailability; Caco2: Caco2 permeability; BBB: blood-brain barrier; DL: drug likeliness; TPSA: surface sum over all polar atoms, primarily oxygen and nitrogen, also including their attached hydrogens.

### 3.2. Common Targets Mining

According to the search results, in total, 51 putative targets of curcumol molecules were found. Their gene IDs and official gene names are given in supplementary material T1. For HCC, 7958 targets were identified from multiple open-accessible human genomic platforms (TCGA-GTEx, GeneCards, OMIM, etc.). After deleting the repeated subjects, we created a list containing 7958 potential therapeutic targets of HCC.

Then, by using JVenn, 44 common targets were sorted (Figure 2 (A)), which was used to construct an HCC-curcumol-common targets network by Cytoscape software (Figure 2 (B)). To gain a more comprehensive understanding of the expression of the common targets in HCC, their expression distribution of each gene in HCC patients and healthy individuals as a comparison were screened in supplementary material S1. In addition, we conducted gene set enrichment analysis (GSEA) for the common targets, in which we found the enrichment score is 0.44 and the normalized enrichment score (NES) is 1.05, indicating that the common targets as a gene set, the expression is upregulated in HCC (details can be found in supplementary material S2).

This section may be divided into subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

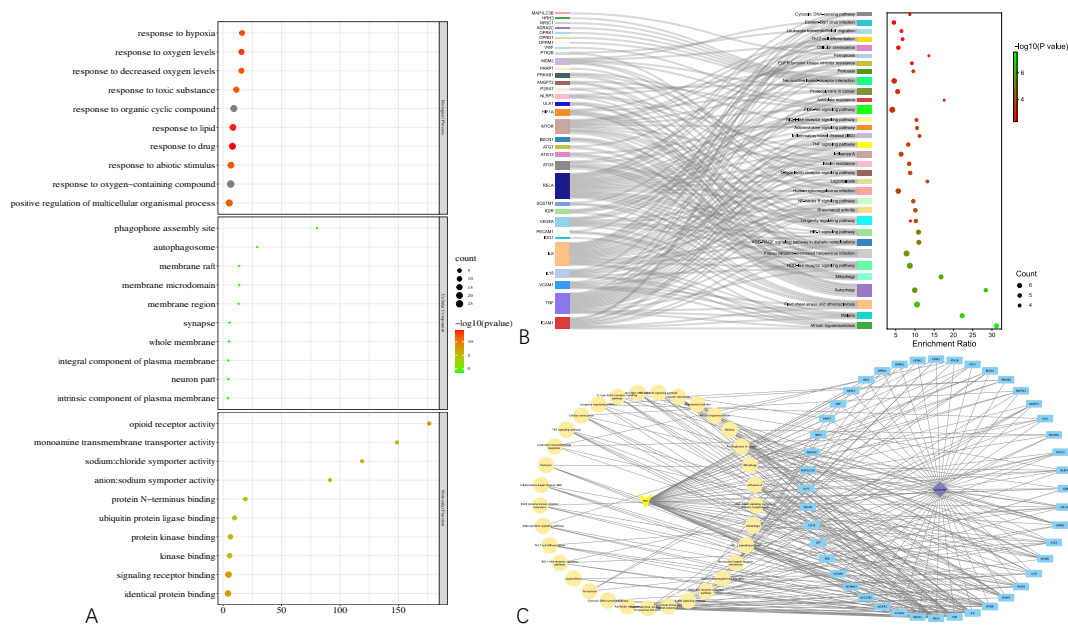


**Figure 2.** Common targets mining and their expression distribution. (A) Venn diagram shows 44 common targets of curcumol and HCC. (B) HCC-curcumol-common targets network.

### 3.3. Functional Enrichment Analyses and Curcumol-Targets-Hcc-Pathways Network Construction

To gain a macroscopic understanding of the functions of the common targets, GO terms and KEGG pathways enrichment analyses were carried out. Consequently, the top 10 most enriched GO terms in biological processes, cellular components, and molecular functions were screened in the form of a bubble plot. As shown in Figure 3 (A), the most enriched GO term in a biological process is “response to oxygen-containing compound”, followed by “response to organic cyclic compound”, and then “response to drug”, while the most enriched GO term in cellular component and molecular function is “whole membrane” and “opioid receptor activity”, respectively. For KEGG pathways analysis, 36 pathways that are statistically significant ( $FDR < 0.05$ ) were screened in the form of a bubble plot as well but combined with a Sankey diagram, which demonstrates the genes within each pathway. As shown in Figure 3 (B), these pathways are mainly related to apoptosis (e.g., hsa04136 and hsa04140 autophagy, hsa04137 mitophagy), inflammation (e.g., hsa04621 Nod-like receptor signaling pathway) and responses to viral infection (e.g., hsa05167 Kaposi sarcoma-associated herpesvirus infection). Notably, through the Sankey diagram, it was found that RELA participates in the largest number of pathways, followed by IL6 and TNF.

Based on the results of the functional enrichment analyses, a curcumol-targets-pathways-HCC network was constructed, as shown in Figure 3 (C), which contains 79 nodes and 265 edges. Among these nodes, RELA, the proto-oncogene that transcribes p65, has the highest degree in the network, followed by TNF, IL6, MTOR, which are responsible for cell growth and inflammation.

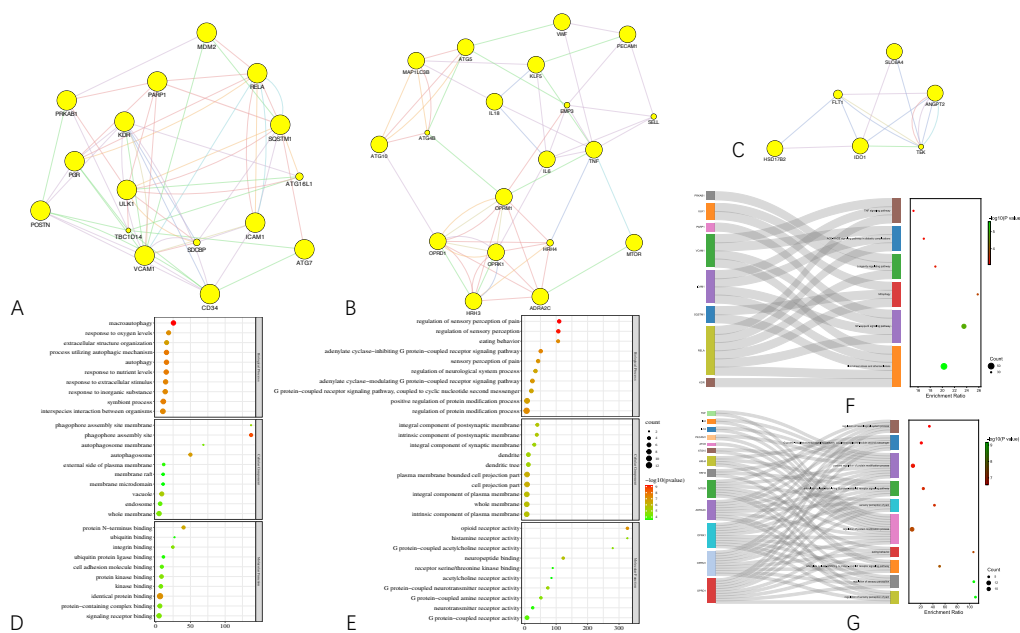


**Figure 3.** Results of functional enrichment analyses of the common targets and curcumol-targets-pathways-HCC network. (A) Bubble plot demonstrating top 10 most enriched GO terms in biological process, cellular component, and molecular function. (B) Bubble plot combined with Sankey diagram demonstrating statistically significant KEGG pathways and the genes within each pathway. (C) Curcumol-targets-pathways-HCC network.

### 3.4. PPI Network Construction and In-Depth Structural Optimization

Attempting to reveal the underlying potential therapeutic mechanism against HCC, we used GeneMANIA to build a preliminary PPI network with 64 nodes and 512 edges, as shown in Figure 4. The network density is 0.191, the heterogeneity is 0.443, and the centrality is 0.212.





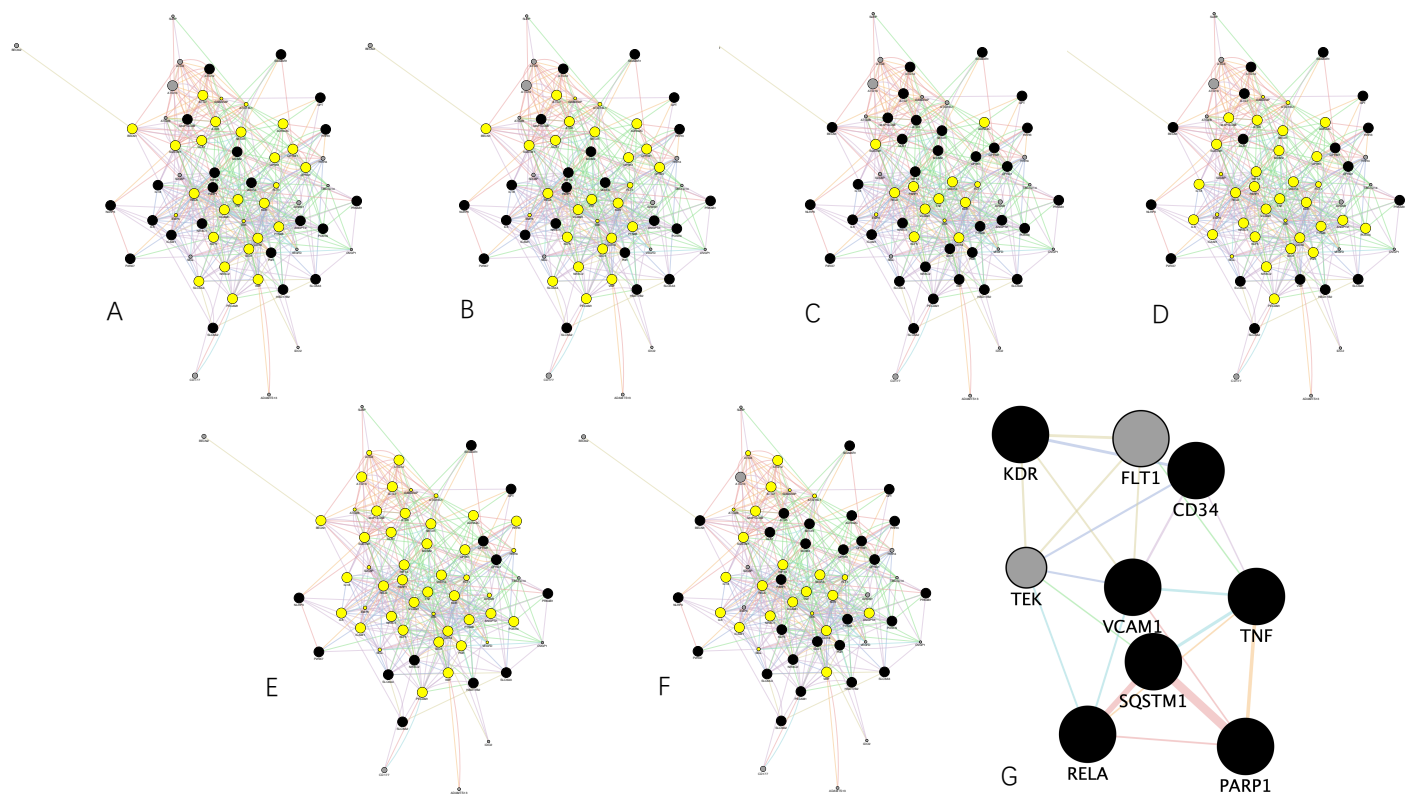
**Figure 5.** Cluster classification of the preliminary PPI network. (A–C) are clusters 1 to 3 divided through the MCODE plug-in, respectively. (D–E) Bubble plots demonstrating the top 10 most enriched GO terms in biological process, cellular component, and molecular function of cluster 1 and cluster 2, respectively. (F–G) Bubble plots combined with Sankey diagrams demonstrating the KEGG pathways with static significance (FDR < 0.05) and the proteins within each pathway of cluster 1 and cluster 2, respectively.

Besides the previous analyses, a topological analysis of the preliminary PPI network is also necessary. Therefore, the CytoNCA plug-in was used to find the core proteins that frame the preliminary network.

First, above all, we calculated the average value of degree, betweenness, closeness, eigenvector, LAC, network centralities of the preliminary network, which is 12, 64.3125, 0.543188007, 0.11156333, 4.789801879, and 6.193579598, respectively.

Then, we extracted the corresponding networks by setting the threshold as these averages. That is to say, in other words, according to the above-mentioned averages, the nodes in the preliminary PPI network greater than these averages were sorted out for corresponding subnetwork construction. Figure 6 (A–F) demonstrates the subnetwork extracted through degree, betweenness, closeness, eigenvector, LAC, and network centrality method, respectively.

Furthermore, we extracted a core PPI subnetwork by using the intersectional merge function of the Cytoscape software. In this way, common nodes showing up within the above-mentioned subnetworks were sorted out for core PPI subnetwork construction. As a result, a core PPI subnetwork in Figure 6 (G) containing 9 key nodes (i.e., KDR, TEK, RELA, SQSTM1, VCAM1, FLT1, CD34, TNF, and PNRP1) and 22 edges was extracted. Through the figure, it can be clearly seen that SQSTM1 shares the strongest and the greatest number of connections with the others, followed by RELA.

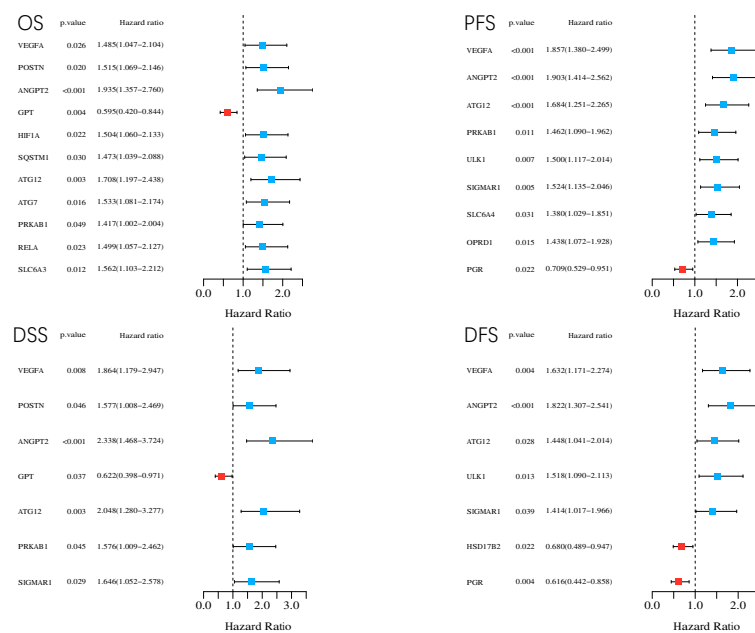


**Figure 6.** The process of core PPI subnetwork generation. (A)–(F) PPI subnetworks demonstrating the subnetworks extracted from the preliminary PPI network according to the degree, betweenness, closeness, eigenvector, LAC, and network centrality method, respectively. (G) Core PPI subnetwork generated by intersectional merge of (A)–(F) PPI subnetworks.

### 3.5. Survival Analysis

By considering  $p < 0.05$  as statistically significant, we conducted log-rank tests and univariate Cox proportional hazards regression for all the 44 common targets to carry out survival analysis, including overall survival (OS), progression-free survival (PFS), disease-free survival (DFS), and disease-specific survival (DSS). In total, 17 genes were identified as prognosis related in HCC in case of different types of survival.

Their corresponding KM curves are available in supplementary material S7. The forest plots in Figure 7 demonstrate statistically significant genes and their P values, hazard ratios (HR) with 95% confidence interval (IC) in different types of survival (i.e., OS, PFS, DFS, DSS). In summary, through the survival analysis, except GPT, PGR, and HSD17B2, the rest (i.e., VEGFA, ANGPT2, ATG12, PRKAB1, ULK1, SIGMAR1, SLC6A4, OPRD1, AGT7, HIF1A, POSTN, SLC6A3, GPT, RELA, and SQSTM1) are related to poor prognosis.



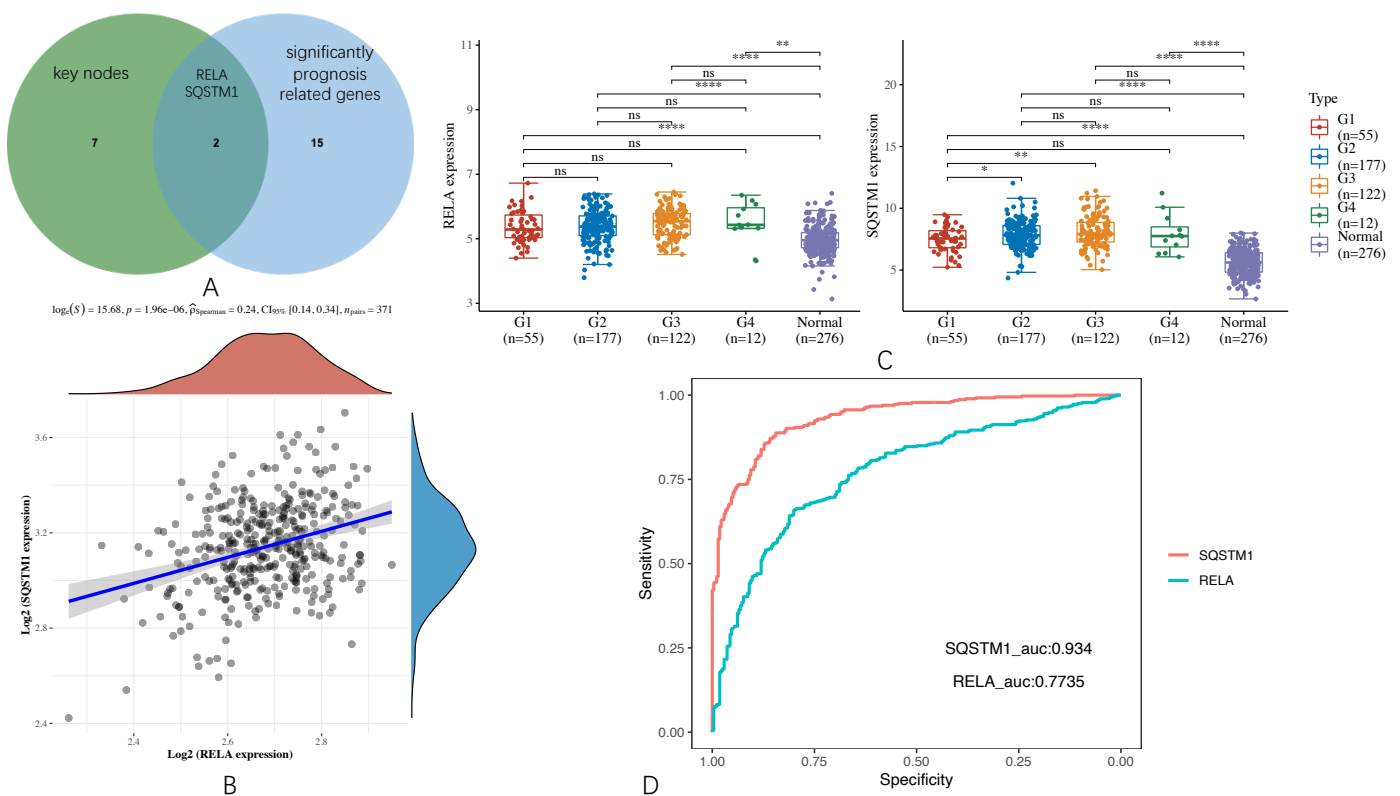
**Figure 7.** Forest plots demonstrating statistically significant genes and their P values, hazard ratios (HR) with 95% confidence interval (IC) in different types of survival (i.e., OS, PFS, DFS, and DSS). OS: overall survival; PFS: progression-free survival; DFS: disease-free survival; DSS: disease-specific survival. HR>1 indicates that the corresponding gene is a protective factor, while HR < 1 indicates the corresponding gene is positively associated with the deterioration of cancer. When  $p < 0.05$ , the line of the gene will not touch the reference vertical line in the middle of the forest plot, meaning that this gene is statistically significant.

### 3.6. Screening for the Key Targets and Exploration of Their Expression and Correlation in HCC

By using the JVenn program, two key targets, RELA and SQSTM1, were sorted out from the intersection of the nine key nodes and the significantly prognosis-related genes, as shown in Figure 8 (A). Then, we explored the correlation between the expression of RELA and SQSTM1 through the 371 samples collected from the TCGA cohort. The result is shown in Figure 8 (B) indicates that RELA and SQSTM1 are positively associated with the Spearman coefficient  $>0$  and the corresponding  $p$ -value  $< 0.001$ . This hints that they have potential synergistic effects.

Moreover, we also explored their expression levels at different HCC grades and normal tissues by Kruskal–Wallis’s test. The result shown in Figure 8 (C) suggests that their expression levels are not significantly related to different cancer grades, but there is a huge difference between cancer and normal tissues.

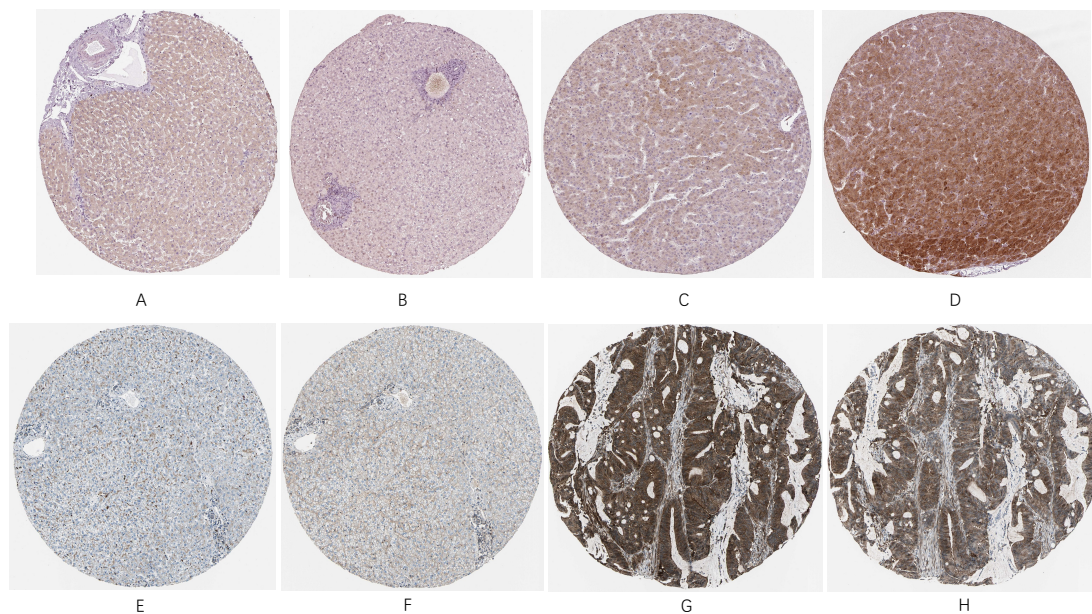
In order to verify whether RELA and SQSTM1 may serve as reliable biomarkers in HCC, we further performed a diagnostic receiver operative characteristic (ROC) analysis, as shown in Figure 8 (D), in which we followed the philosophy of the gold-standard test (i.e., dichotomization of the mentioned samples in TCGA cohort and GTEx portal into disease present, which includes all tumor grades from G1 to G4, or disease absent meaning that the sample is normal) [49,50]. It was found that the area under curve (AUC) of RELA is 0.7735, and that of SQSTM1 is 0.934. As generally speaking, the AUC value was considered excellent if it is from 0.9 to 1, good if it is from 0.8 to 0.9, fair if it is from 0.7 to 0.8; herein, the reliability of RELA should be thought of as fair, and that of SQSTM1 should be thought of as excellent [51].



**Figure 8.** Screening for the key targets and their brief summary. (A) Venn diagram that screens the key targets (i.e., RELA, SQSTM1). (B) The correlations between RELA and SQSTM1 analyzed with the Spearman method. The density curve on the right represents the trend in the distribution of SQSTM1 expression in 371 HCC samples from the TCGA cohort. The upper density curve represents the trend in the distribution of RELA expression in 371 HCC samples from the TCGA cohort. (C) The expression levels of RELA and SQSTM1 in various cancer grades (i.e., G1–G4) and normal tissue as comparison. The number of the stars indicates the statistical significance (\*:  $p < 0.05$ , \*\*:  $p < 0.01$ , \*\*\*\*:  $p < 0.0001$ ). ns means non-significant. (D) Diagnostic ROC curve for RELA and SQSTM1 in HCC.

### 3.7. InSitu Analysis

In order to verify the importance of RELA and SQSTM1 proteins in HCC patients, we compared their expression with that of normal individuals. Both proteins showed relatively low expression in normal tissues, while in the histological slides from HCC patients, they exhibited high expression levels, as shown in Figure 9 (A–H). Therefore, it is verified clinically that they are important poor prognosis-related biomarkers.

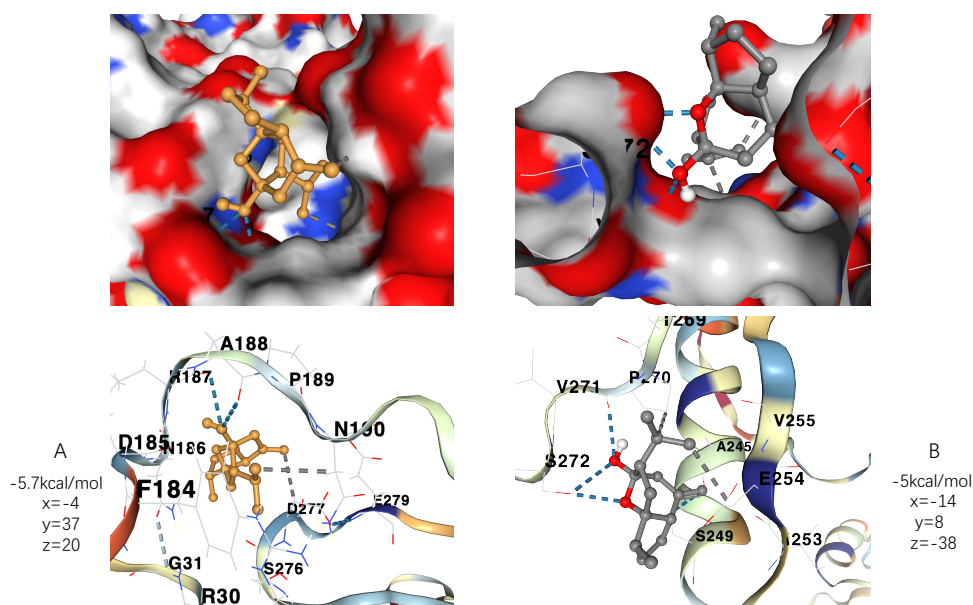


**Figure 9.** (A) and (B) are normal histological slides from liver tissue sample of a female patient at age 50, (C) and (D) are pathological histological slides from liver tissue sample of a male patient at age 67, (E) and (F) are normal histological slides from liver tissue sample of a female patient at age 32, (G) and (H) are pathological histological slides from liver tissue sample of a female patient at age 73.

### 3.8. Molecular Docking of Key Targets

We downloaded the corresponding three-dimensional structures from RCSB PDB to conduct blind docking. The docking site was set in a cubic box in the center of the initial ligand. The results of RELA docking indicates that its binding with curcumol possesses two hydrogen bonds with ARG 187 and ALA 188 and two hydrophobic interactions with ASP 277 and ASN 190, while the results of SQSTM1 docking indicates that its binding with curcumol possesses two hydrogen bonds with ARG 187 and ALA 188 and two hydrophobic interactions with ASP 277 and ASN 190.

Moreover, as the more negative the docking score is, the higher the binding force is between the compound and the protein, we, therefore, found the best run provided the binding affinity of  $-5.7\text{kcal/mol}$  and  $-5\text{kcal/mol}$  to RELA and SQSTM1, respectively, taking together with their specific interactions into account. As generally speaking, affinity energy  $\leq -5\text{kcal/mol}$  is deemed as high affinity. RELA and SQSTM1 thus can be considered as key players that participate in the potential mechanism against HCC of curcumol. The docking results can be viewed in a three-dimensional manner in Figure 10 (A–B).



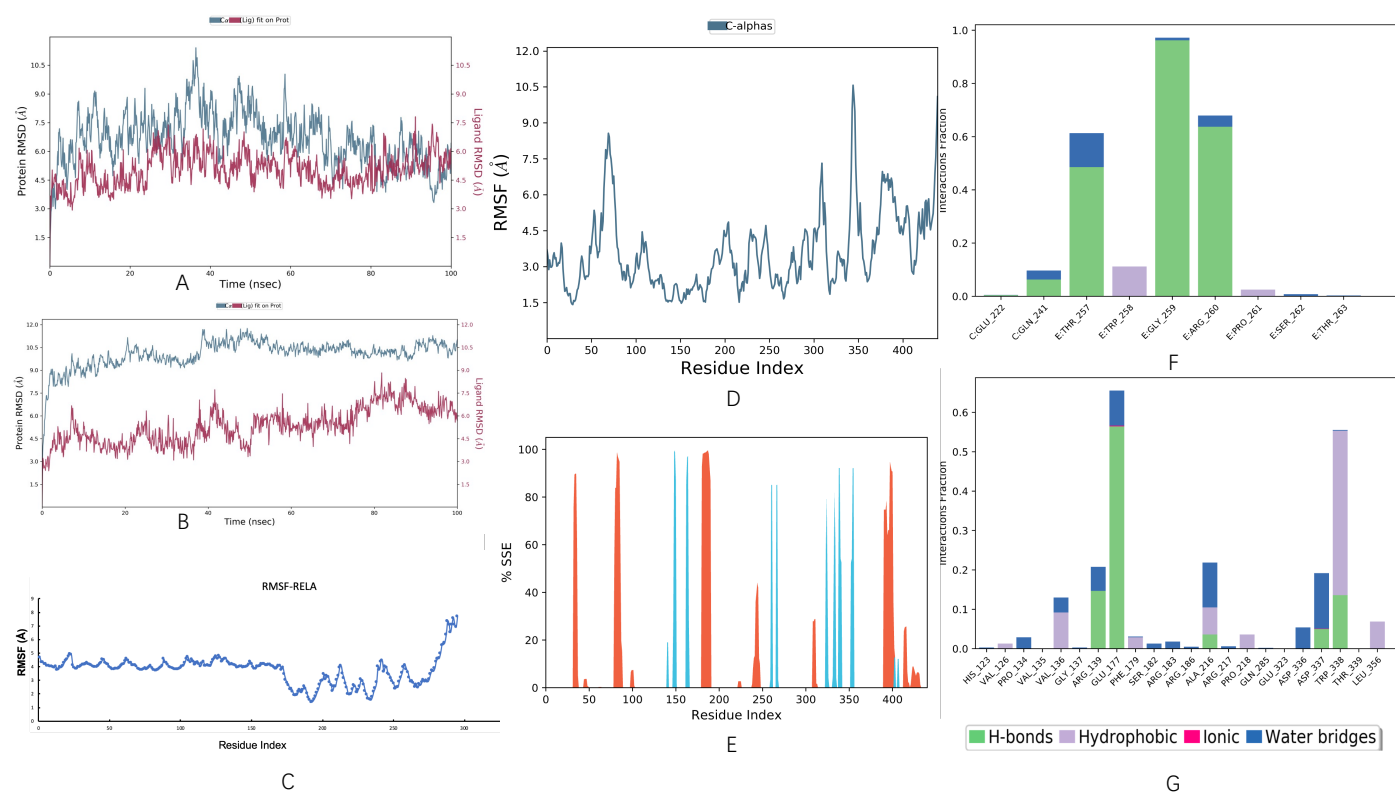
**Figure 10.** Molecular docking results with affinity energy and specific three-dimensional coordinates. (A) demonstrates the three-dimensional conformation of the docking of RELA protein with curcumol molecule and the corresponding binding affinity. (B) Demonstrates the three-dimensional conformation of the docking of SQSTM1 protein with curcumol molecule and the corresponding binding affinity.

### 3.9. Molecular Dynamic Simulation

Molecular dynamics simulations were performed on the top hits containing high binding energies. Over the simulation period, the projected conformational changes from the initial structure were presented in RMSD. Moreover, structural stability, atomic mobility, and residue flexibility at times of interaction of protein-hit were expressed with RMSF values. The peaks of the RMSF graph represent the fluctuation portion of the protein through the simulation. The N- and C-terminals show more changes than any other portion of the protein. Alpha helices and beta strands show less fluctuation, as they are stiffer than the unstructured part of the protein than the loop portion. The RMSD of the curcumol-RELA complex showed deviation at almost 30 ns for 2 Å, and then the system was equilibrated throughout the simulation, as shown in Figure 11 (A). The RMSD of the curcumol-SQSTM1 complex showed little deviation, but it was in the range of 3 Å. In general, the complex was stable throughout the simulation, the ligand remained inside the binding pocket and made important interactions, and the backbone was consistent. The deviation might be due to conformational changes, as shown in Figure 11 (B). The RMSF for RELA showed little fluctuation throughout the simulation, as shown in Figure 11 (C). Similarly, in RMSF of SQSTM1, there was fluctuation near ASP 69 to ALA 76, GLU 309 to LYS 313, and LYS 344 to ASP 347 residues, and the remaining structure was comparatively stabilized, and there was not much fluctuation where ligand made contacts with protein, as shown in Figure 11 (D). The fluctuation in the protein might be due to 83% of the unstructured part of the protein. Furthermore, secondary structure elements (SSE) such as alpha helices and beta strands were monitored throughout the simulation. It turned out to be 8.38% of helix, 4.60% of strands, and 12.98% of total SSE in the SSE analysis, as shown in Figure 11 (E).

The curcumol-RELA complex possessed several different types of interactions, including hydrogen bonds, water bridges, and hydrophobic, as shown in Figure 11 (F), while in the case of the curcumol-SQSTM1 complex, the types of intermolecular interactions during the entire simulation were found to be more diverse, including hydrogen bonds, ionic, water bridges, and hydrophobic. The residues participating in these interactions include VAL 126, PRO 134, VAL 136, ARG 139, GLU 177, PHE 179, SER 182, SER 183,

ALA 216, PRO 218, GLU 323, ASP 336, ASP 337, TRP 338, and LEU 356, as shown in Figure 11 (G).



**Figure 11.** Molecular dynamic simulation performed on the top hits containing high binding energies. (A)–(B) Root mean square deviation (RMSD) plot of curcumol-RELA complex and curcumol-SQSTM1 complex, respectively. (C)–(D) Root mean square fluctuation (RMSF) plot of RELA and SQSTM1, respectively. (E) Protein secondary structure elements (SSE) of SQSTM1. (F)–(G) protein-ligand contacts of curcumol-RELA complex and curcumol-SQSTM1 complex, respectively.

#### 4. Discussion

The broad use of essential volatile oils from *Curcumae Rhizoma* in the treatment of many diseases, including tumors, has a long and storied history. While HCC is a fatal major malignance among liver cancers, previous studies have not yet explored the potential of curcumol against this form of the disease. Therefore, the present study attempted to make the first step to address this issue. As blindly conducting experiments on a large scale is both time and resource consuming, *in silico* method was employed in the present study in order to improve the effectiveness and save labor.

In the present study, out of 371 HCC samples and 276 normal tissue samples, 44 common targets were mined from various authorized databases (e.g., TCGA, GTEx, GeneCards, OMIM, TCD, etc.). With these targets, functional analyses were carried out to gain a macroscopic understanding of their roles against HCC. It appears that curcumol could function by inducing apoptosis and inflammation, disturbing vessel generation of the tumor. Notably, VEGFA appears to be down-regulated so that vascularization of tumor tissue can be suppressed. These findings suggest that curcumol may function against HCC via inducing cell death and immunological reactions, as well as disturbing intra-tumor vascularization [52] and eliminating adverse effects of currently used drugs.

Moreover, a preliminary PPI network was also constructed, trying to gain in-depth insight into their functions. Via topological optimization, we conducted functional clustering for the preliminary PPI network. Each cluster was followed by GO terms and KEGG pathways enrichment analyses through which we found that proteins in cluster 1 were mainly associated with apoptosis, whereas proteins in cluster 2 were mainly associated

with neurological signaling. Moreover, a core PPI subnetwork with nine key nodes was extracted, which was used to screen key targets in the following analyses.

In addition, we conducted a survival analysis for different types of survival (i.e., OS, PFS, DFS, DSS) for the 44 common targets where 17 genes were considered as significantly prognosis related.

Among them, 2 key targets, SQSTM1 and RELA, were sorted out by intersecting between the 9 key targets and these 17 significantly prognosis-related genes. From the aspects of cancer grading and Spearman correlation analysis, we provided a brief summary of RELA and SQSTM1, which emphasized their importance in HCC with ROC curve plotted to evaluate the reliability of our analyses. Then, we conducted *in situ* analysis by using the histological slides from HCC tissues and normal tissues. It was found that both RELA and SQSTM1 showed relatively low expression in normal tissues, while in the histological slides from HCC patients, they exhibited high expression levels, indicating that our computational prediction is in suitable agreement with the observation results. Consequently, it was found that their overexpression is statistically significant in tumor tissues, and they themselves are correlated with poor prognosis.

Via molecular docking, we found RELA and SQSTM1 showed high affinity to the curcumol molecule and are thus thought to be the key players in the potential anti-HCC mechanism of curcumol. Therefore, we further verified the stability of their binding by performing molecular dynamic simulation to ensure the binding was not affected by different docking protocols. As a result, we found that both the RELA-curcumol complex and SQSTM1-curcumol complex were stable during 100 ns-long molecular dynamic simulations. Combined with current literature, it is believed that curcumol decreases the expression level of RELA [53] and has certain indirect interactions with SQSTM1 expression, though whether these interactions may result in up-regulation or down-regulation remains unclear [52].

Although further experiments are needed, the present study elucidated the potential of curcumol against HCC, providing an important basis for *in vivo* studies. The approach described in the current work could also be used to screen for other biologically active compounds used in traditional Eastern medicine.

## 5. Conclusions

In summary, the present study demonstrated the potential underlying mechanism of curcumol in HCC treatment based on a network pharmacology method. We revealed that curcumol seemed to play a critical role in HCC by affecting multiple biotargets. Moreover, our molecular docking and molecular dynamic simulation results also suggested that the RELA and SQSTM1 proteins could bind with curcumol molecule stably, providing an important basis for further investigation, although this study also has certain limitations as pharmacological and clinical research still need to further validate our findings.

**Supplementary Materials:** The following supporting information can be downloaded at: [www.mdpi.com/xxx/s1](http://www.mdpi.com/xxx/s1), File S1: Expression distribution of common targets; File S2: GSEA result of common targets, File S3: Rare data for top 10 most enriched GO terms in biological process, cellular component, and molecular function of common targets, File S4: Rare data for KEGG enrichment result of common targets, File S5: Rare data for cluster classification performed by MCODE, File S6: Rare data for GO terms and KEGG pathways enrichment analyses of cluster 1 and cluster 2, File S7: Prognosis-related genes, Table S1: Putative targets of curcumol.

**Author Contributions:** Conceptualization, Xufeng Huang and Shujing Zhou; methodology, Xufeng Huang; software, Xufeng Huang and Hafiz Muzzammel Rehman; validation, Xufeng Huang, Hafiz Muzzammel Rehman, Shujing Zhou, and Attila Szöllősi; formal analysis, Xufeng Huang and Hafiz Muzzammel Rehman; investigation, Xufeng Huang; resources, Shujing Zhou; data curation, Shujing Zhou; writing—original draft preparation, Xufeng Huang and Shujing Zhou; writing—review and editing, Attila Szöllősi; visualization, Xufeng Huang; supervision, Attila Szöllősi; project administration, Xufeng Huang; funding acquisition, Xufeng Huang. All authors have read and agreed to the published version of the manuscript.”

**Funding:** The present study did not receive any external funding. The article processing fee was covered by the author's stipend issued by Tempus public foundation and the Chinese scholarship council.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

**Acknowledgments:** We sincerely thank the public databases, including Human Protein Atlas project, TCGA, GeneCards, OMIM, CTD, Swiss, and more, for providing open-accessible and high-quality resources for researchers. We also sincerely thank the Tempus public foundation and the Chinese scholarship council for their stipend to the author, which made it possible to cover the article processing fee of the present study.

**Conflicts of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Balogh, J.; Victor, D., 3rd.; Asham, E.H.; Burroughs, S.G.; Boktour, M.; Saharia, A.; Li, X.; Ghobrial, R.M.; Monsour, H.P., Jr. Hepatocellular carcinoma: A review. *J. Hepatocell. Carcinoma* **2016**, *3*, 41–53. <https://doi.org/10.2147/jhc.s61146>.
- Hashem, S.; Nisar, S.; Sageena, G.; Macha, M.A.; Yadav, S.K.; Krishnankutty, R.; Uddin, S.; Haris, M.; Bhat, A.A. Therapeutic Effects of Curcumol in Several Diseases: An Overview. *Nutr. Cancer* **2021**, *73*, 181–195. <https://doi.org/10.1080/01635581.2020.1749676>.
- Xia, Q.; Zhao, K.J.; Huang, Z.G.; Zhang, P.; Dong, T.T.X.; Li, S.P.; Tsim, K.W.K. Molecular genetic and chemical assessment of *Rhizoma Curcumae* in China. *J. Agric. Food Chem.* **2005**, *53*, 6019–6026. <https://doi.org/10.1021/jf0508495>.
- Zhou, Y.; Xie, M.; Song, Y.; Wang, W.; Zhao, H.; Tian, Y.; Wang, Y.; Bai, S.; Zhao, Y.; Chen, X.; et al. Two Traditional Chinese Medicines *Curcumae Radix* and *Curcumae Rhizoma*: An Ethnopharmacology, Phytochemistry, and Pharmacology Review. *Evid. Based Complement. Altern. Med.* **2016**, *2016*, 4973128. <https://doi.org/10.1155/2016/4973128>.
- Huang, Y.; Xue, C.; He, W.; Zhao, X. Inhibition effect of Zedoary turmeric oil on *Listeria monocytogenes* and *Staphylococcus aureus* growth and exotoxin proteins production. *J. Med. Microbiol.* **2019**, *68*, 657–666. <https://doi.org/10.1099/jmm.0.000949>.
- Lee, T.K.; Lee, D.; Lee, S.R.; Ko, Y.-J.; Kang, K.S.; Chung, S.J.; Kim, K.H. Sesquiterpenes from *Curcuma zedoaria* rhizomes and their cytotoxicity against human gastric cancer AGS cells. *Bioorg. Chem.* **2019**, *87*, 117–122. <https://doi.org/10.1016/j.bioorg.2019.03.015>.
- Makabe, H.; Maru, N.; Kuwabara, A.; Kamo, T.; Hirota, M. Anti-inflammatory sesquiterpenes from *Curcuma zedoaria*. *Nat. Prod. Res.* **2006**, *20*, 680–685. <https://doi.org/10.1080/14786410500462900>.
- Matsuda, H.; Morikawa, T.; Ninomiya, K.; Yoshikawa, M. Hepatoprotective constituents from *Zedoariae Rhizoma*: Absolute stereostructures of three new carabran-type sesquiterpenes, curcumenolactones A, B, and C. *Bioorg. Med. Chem.* **2001**, *9*, 909–916. [https://doi.org/10.1016/s0968-0896\(00\)00306-0](https://doi.org/10.1016/s0968-0896(00)00306-0).
- Ru, J.; Li, P.; Wang, J.; Zhou, W.; Li, B.; Huang, C.; Li, P.; Guo, Z.; Tao, W.; Yang, Y.; et al. TCMSP: A database of systems pharmacology for drug discovery from herbal medicines. *J. Cheminform.* **2014**, *6*, 13. <https://doi.org/10.1186/1758-2946-6-13>.
- Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; De Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. <https://doi.org/10.1093/nar/gky427>.
- Guex, N.; Peitsch, M.C.; Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-Pdb-Viewer: A historical perspective. *Electrophoresis* **2009**, *30* (Suppl 1), S162–S173. <https://doi.org/10.1002/elps.200900140>.
- Bienert, S.; Waterhouse, A.; de Beer, T.A.P.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T. The SWISS-MODEL Repository—New features and functionality. *Nucleic Acids Res.* **2017**, *45*, D313–D319. <https://doi.org/10.1093/nar/gkw1132>.
- Davis, A.P.; Grondin, C.J.; Johnson, R.J.; Sciaky, D.; Wiegiers, J.; Wiegiers, T.C.; Mattingly, C.J. Comparative Toxicogenomics Database (CTD): Update 2021. *Nucleic Acids Res.* **2020**, *49*, D1138–D1143. <https://doi.org/10.1093/nar/gkaa891>.
- Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B.; et al. PubChem in 2021: New data content and improved web interfaces. *Nucleic Acids Res.* **2021**, *49*, D1388–D1395. <https://doi.org/10.1093/nar/gkaa971>.
- National Center for Biotechnology Information. PubChem Compound Summary for CID 14240392. Curcumol. Available online: <https://pubchem.ncbi.nlm.nih.gov/compound/Curcumol> (accessed on 9 November 2021).
- Tomczak, K.; Czerwińska, P.; Wizniewowicz, M. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Contemp. Oncol.* **2015**, *19*, 68–77. <https://doi.org/10.5114/wo.2014.47136>.
- GTE Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **2013**, *45*, 580–585. <https://doi.org/10.1038/ng.2653>.

- Ritchie, M.E.; Belinda, P.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. <https://doi.org/10.1093/nar/gkv007>.
- Rebhan, M.; Chalifa-Caspi, V.; Prilusky, J.; Lancet, D. GeneCards: A novel functional genomics compendium with automated data mining and query reformulation support. *Bioinformatics* **1998**, *14*, 656–664. <https://doi.org/10.1093/bioinformatics/14.8.656>.
- Hamosh, A.; Scott, A.F.; Amberger, J.S.; Bocchini, C.A.; Valle, D.; McKusick, V.A. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.* **2002**, *30*, 52–55. <https://doi.org/10.1093/nar/30.1.52>.
- Bardou, P.; Mariette, J.; Escudié, F.; Djemiel, C.; Klopp, C. jvenn: An interactive Venn diagram viewer. *BMC Bioinform.* **2014**, *15*, 293. <https://doi.org/10.1186/1471-2105-15-293>.
- Suk, H.W.; Hwang, H. Functional Generalized Structured Component Analysis. *Psychometrika* **2016**, *81*, 940–968. <https://doi.org/10.1007/s11336-016-9521-1>.
- Liao, Y.; Wang, J.; Jaehnig, E.J.; Shi, Z.; Zhang, B. WebGestalt 2019: Gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* **2019**, *47*, W199–W205. <https://doi.org/10.1093/nar/gkz401>.
- Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13*, 2498–2504. <https://doi.org/10.1101/gr.1239303>.
- Warde-Farley, D.; Donaldson, S.L.; Comes, O.; Zuberi, K.; Badrawi, R.; Chao, P.; Franz, M.; Grouios, C.; Kazi, F.; Lopes, C.T.; et al. The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **2010**, *38*, W214–W220. <https://doi.org/10.1093/nar/gkq537>.
- Bader, G.D.; Hogue, C.W.V. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinform.* **2003**, *4*, 2–27. <https://doi.org/10.1186/1471-2105-4-2>.
- Tang, Y.; Li, M.; Wang, J.; Pan, Y.; Wu, F.-X. CytoNCA: A cytoscape plugin for centrality analysis and evaluation of protein interaction networks. *Biosystems* **2015**, *127*, 67–72. <https://doi.org/10.1016/j.biosystems.2014.11.005>.
- Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016; ISBN 978-3-319-24277-4. Available online: <https://ggplot2.tidyverse.org>.
- Therneau, T.M. A Package for Survival Analysis in R [Internet]. 2020. Available online: <https://CRAN.R-project.org/package=survival>.
- Patil, I. Visualizations with statistical details: The ‘ggstatsplot’ approach. *J. Open Source Softw.* **2021**, *6*, 3167. <https://doi.org/10.21105/joss.03167>.
- Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.-C.; Müller, M. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **2011**, *12*, 77. <https://doi.org/10.1186/1471-2105-12-77>.
- Uhlén, M.; Fagerberg, L.; Hallström, B.M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; et al. Tissue-based map of the human proteome. *Science* **2015**, *347*, 1260419. <https://doi.org/10.1126/science.1260419>. PMID: 25613900.
- Schrödinger, L.; DeLano, W. PyMOL 2020. Available online: <http://www.pymol.org/pymol>.
- Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. <https://doi.org/10.1002/jcc.21334>.
- Eberhardt, J.; Santos-Martins, D.; Tillack, A.F.; Forli, S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J. Chem. Inf. Model* **2021**, *61*, 3891–3898. <https://doi.org/10.1021/acs.jcim.1c00203>.
- Liu, Y.; Grimm, M.; Dai, W.-T.; Hou, M.-C.; Xiao, Z.-X.; Cao, Y. CB-Dock: A web server for cavity detection-guided protein–ligand blind docking. *Acta Pharmacol. Sin.* **2020**, *41*, 138–144. <https://doi.org/10.1038/s41401-019-0228-6>.
- Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
- Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein–sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50*, D439–D444. <https://doi.org/10.1093/nar/gkab1061>.
- Bowers, K.J.; Chow, D.E.; Xu, H.; Dror, R.O.; Eastwood, M.P.; Gregersen, B.A.; Klepeis, J.L.; Kolossvary, I.; Moraes, M.A.; Sacerdoti, F.D.; et al. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In Proceedings of the SC’06: 2006 ACM/IEEE Conference on Supercomputing, Tampa, FL, USA, 11–17 November 2006; pp. 43–43. <https://doi.org/10.1109/sc.2006.54>.
- El Khatabi, K.; El-Mernissi, R.; Moukhliiss, Y.; Hajji, H.; Rehman, H.M.; Yadav, R.; Lakhliifi, T.; Ajana, M.A.; Bouachrine, M. Rational design of novel potential EGFR inhibitors by 3D-QSAR, molecular docking, molecular dynamics simulation, and pharmacokinetics studies. *Chem. Data Collect.* **2022**, *39*, 100851. <https://doi.org/10.1016/j.cdc.2022.100851>.
- Viswanadhan, V.N.; Ghose, A.K.; Revankar, G.R.; Robins, R.K. Atomic physicochemical parameters for three dimensional structure directed quantitative structure–activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 163–172. <https://doi.org/10.1021/ci00063a006>.
- Veber, D.F.; Johnson, S.R.; Cheng, H.-Y.; Smith, B.R.; Ward, K.W.; Kopple, K.D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45*, 2615–2623. <https://doi.org/10.1021/jm020017n>.

- Xu, X.; Zhang, W.; Huang, C.; Li, Y.; Yu, H.; Wang, Y.; Duan, J.; Ling, Y. A Novel Chemometric Method for the Prediction of Human Oral Bioavailability. *Int. J. Mol. Sci.* **2012**, *13*, 6964–6982. <https://doi.org/10.3390/ijms13066964>.
- Shen, M.; Tian, S.; Li, Y.; Li, Q.; Xu, X.; Wang, J.; Hou, T. Drug-likeness analysis of traditional Chinese medicines: 1. property distributions of drug-like compounds, non-drug-like compounds and natural compounds from traditional Chinese medicines. *J. Chem.* **2012**, *4*, 31. <https://doi.org/10.1186/1758-2946-4-31>.
- Lea, T. Caco-2 Cell Line. In *The Impact of Food Bioactives on Health: In Vitro and Ex Vivo Models [Internet]*; Verhoeckx, K., Cotter, P., López-Expósito, I., Kleiveland, C., Lea, T., Mackie, A., Requena, T., Swiatecka, D., Wichers, H., Eds.; Springer: Cham, The Netherlands, 2015; Chapter 10.
- Tattersall MH, Sodergren JE, Dengupta SK, Trites DH, Modest EJ, Frei E 3rd. Pharmacokinetics of actinoymcin D in patients with malignant melanoma. *Clin Pharmacol Ther.* 1975 Jun;17(6):701-8.
- Ertl, P.; Rohde, B.; Selzer, P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J. Med. Chem.* **2000**, *43*, 3714–3717.
- Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3–26. [https://doi.org/10.1016/s0169-409x\(00\)00129-0](https://doi.org/10.1016/s0169-409x(00)00129-0).
- Li, F.; He, H. Assessing the Accuracy of Diagnostic Tests.. *Shanghai Arch. Psychiatry* **2018**, *30*, 207–212. <https://doi.org/10.11919/j.issn.1002-0829.218052>.
- Obuchowski, N.A. An ROC-type measure of diagnostic accuracy when the gold standard is continuous-scale. *Stat. Med.* **2006**, *25*, 481–493. <https://doi.org/10.1002/sim.2228>.
- Mandrekar, J.N. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *J. Thorac. Oncol.* **2010**, *5*, 1315–1316. <https://doi.org/10.1097/jto.0b013e3181ec173d>.
- Gao, L.; Yang, X.; Liang, B.; Jia, Y.; Tan, S.; Chen, A.; Cao, P.; Zhang, Z.; Zheng, S.; Sun, L.; et al. Autophagy-induced p62 accumulation is required for curcumol to regulate KLF5-mediated angiogenesis in liver sinusoidal endothelial cells. *Toxicology* **2021**, *452*, 152707. <https://doi.org/10.1016/j.tox.2021.152707>.
- Jia, Y.; Gao, L.; Yang, X.; Zhang, F.; Chen, A.; Wang, S.; Shao, J.; Tan, S.; Zheng, S. Blockade of periostin-dependent migration and adhesion by curcumol via inhibition of nuclear factor kappa B signaling in hepatic stellate cells. *Toxicology* **2020**, *440*, 152475. <https://doi.org/10.1016/j.tox.2020.152475>.