Check for updates

A network-based pharmacological investigation to identify the mechanistic regulatory pathway of andrographolide against colorectal cancer

Balakarthikeyan Janani[1†], Mayakrishnan Vijayakumar[2†], Kannappan Priya[1*], Jin Hee Kim[2], Ayman Geddawy[3,4], Mohammad Shahid[3], Mahmoud H. El-Bidawy[3,5], Sameer Al-Ghamdi[6], Mohammed Alsaidan[7], Mohammad Hassan Abdelzaher[3,8], Abubucker Peer Mohideen[3] and Thiyagarajan Ramesh [3*]

[1]Department of Biochemistry, PSG College of Arts and Science (Autonomous), Affiliated to Bharathiar University, Coimbatore, Tamil Nadu, India, [2]Department of Integrative Bioscience and Biotechnology, Sejong University, Seoul, South Korea, [3]Department of Basic Medical Sciences, College of Medicine, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia, [4]Department of Pharmacology, Faculty of Medicine, Minia University, Minia, Egypt, [5]Department of Physiology, Faculty of Medicine, Cairo University, Cairo, Egypt, [6]Family and Community Medicine Department, College of Medicine, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia, [7]Internal Medicine Department, College of Medicine, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia, [8]Department of Medical Biochemistry, Faculty of Medicine, Al-Azhar University, Assiut, Egypt

Traditional cancer treatments have posed numerous obstacles, including toxicity, multiple drug resistance, and financial cost. On the contrary, bioactive phytochemicals used in complementary alternative medicine have recently increased attention due to their potential to modulate a wide range of molecular mechanisms with a less toxic effect. Therefore, we investigated the potential regulatory mechanisms of andrographolide to treat colorectal cancer (CRC) using a network pharmacology approach. Target genes of andrographolide were retrieved from public databases (PharmMapper, Swiss target prediction, Targetnet, STITCH, and SuperPred), while targets related to CRC were retrieved from disease databases (Genecards and DisGeNet) and expression datasets (GSE32323 and GSE8671) were retrieved from gene expression omnibus (GEO). Protein-protein interaction networks (PPI) were generated using STRING and Cytoscape, and hub genes were identified by topology analysis and MCODE. Annotation of target proteins was performed using Gene Ontology (GO) database DAVID and signaling pathway enrichment

---

Abbreviations: COX, colorectal cancer; CRC, cyclooxygenase; DEGs, differentially expressed genes; FDR, false discovery rate; GEO, gene expression omnibus; GO, gene ontology; KEGG, kyoto encyclopedia of genes and genomics; MMP, matrix metalloproteinase; PPI, protein-protein interaction

analysis using the Kyoto Encyclopedia and Genome Database (KEGG). Survival and molecular docking analysis for the hub genes revealed three genes (PDGFRA, PTGS2, and MMP9) were involved in the overall survival of CRC patients, and the top three genes with the lowest binding energy include PDGFRA, MET, and MAPK1. MET gene upregulation and PDGFRA and PTGS2 gene downregulation are associated with the survival of CRC patients, as revealed by box plots and correlation analysis. In conclusion, this study has provided the first scientific evidence to support the use of andrographolide to inhibit cellular proliferation, migration, and growth, and induce apoptosis by targeting the hub genes (PDGFRA, PTGS2, MMP9, MAPK1, and MET) involved in CRC migration and invasion.

Introduction

Colorectal cancer accounts for around 10% of all cancer diagnoses and cancer-related deaths globally each year. It is the second most prevalent cancer in women and the third most frequent cancer in men. Incidence and fatality rates are around 25% lower in women than in men (Dekker et al., 2019). Overall, about 4%–5% of the global population suffers from colorectal cancer. Moreover, certain personal characteristics can act as risk factors for polyps' formation, as they can increase the likelihood of developing the disease (Mármol et al., 2017). The likelihood of developing CRC is influenced by many modifiable factors, including obesity, smoking, and eating red meat. Reduced colorectal cancer risks have been associated with physical activity, postmenopausal hormone therapy, nonsteroidal anti-inflammatory drugs, and vegetable and fruit consumption (Johnson et al., 2013). Initiation, promotion, and progression are the stages in the formation of CRC. Epithelial cells in the intestinal mucosa become irreversibly damaged by genetic damage, resulting in subsequent neoplastic transformation (Sawicki et al., 2021). A cancerous growth (cancer) is generated during the promotion phase. In the progression stage, benign carcinoma cells turn malignant and develop aggressive characteristics as well as the capacity to metastasize (Fares et al., 2020). The options for treatment of primary and metastatic cancers (mtCRC) have greatly expanded recently. Among these options are laparoscopic surgery for early-stage CRC, aggressive resection of metastatic CRC (such as lung and liver metastases), radiotherapy for RC, and neoadjuvant and palliative chemotherapy (Dekker et al., 2019; Afshar et al., 2020). The traditional therapies for CRC include surgery and chemotherapy (Woods and Turchi, 2013). Chemotherapy for CRC involves the use of some known drugs with limited clinical applications because of their toxic side effects and resistance to prominent anti-cancer drugs. New treatments for CRC that are less hazardous are therefore required as an alternative or in addition to current methods (Moradi et al., 2019).

Chemotherapeutics can cause cancer cell death by generating DNA damage or activating numerous signaling pathways, such as cell cycle arrest, global translation inhibition, and DNA repair (Woods and Turchi, 2013). The impacts of cytotoxicity, drug resistance, and severe effects are the primary issues linked with chemotherapy (Huang et al., 2019). According to the World Health Organization (WHO), around 80% of the world's population relies on conventional therapies (Aiello et al., 2019).

Phytochemicals, naturally occurring substances in plants, are an important resource for developing new therapeutics and being used as cancer therapies as well. Phytochemicals significantly impact molecular pathways involved in cancer development and progression. Numerous specific strategies are utilized. These include enhancing antioxidant levels, inactivating carcinogens, inhibiting proliferation, causing cell cycle arrest and death, and suppressing immune activity (Choudhari et al., 2019). Due to its numerous medical uses, *Andrographis paniculata*, also known as Kalmegh in India, is largely recognized as a beneficial plant. It is commonly known as the "king of bitters" and is widely distributed throughout India and Sri Lanka. It is a member of the Acanthaceae family (Paul et al., 2019). For centuries, the aerial parts, roots, and entire plant of *A. paniculata* have been used as traditional medicine in Asia to treat a variety of ailments. Traditional healthcare professionals have used it to treat inflammation, pyrexia, stomachaches, and intermittent fevers (Okhuarobo et al., 2014). An important bioactive phytoconstituent in *A. paniculate* is andrographolide, found primarily in the leaves (Jayakumar et al., 2013). It exhibits a variety of biological properties, including immunomodulatory, antiviral, and anti-inflammatory properties, and is often used to treat sore throats, fevers, diarrhea, and inflammatory illnesses such as colitis (Shimura et al., 2021). Additionally, andrographolide has been shown to limit cancer cell motility and invasion by inhibiting the production of proteins or cellular signaling pathways involved in cancer metastasis.

In drug development, the strategy of generating multi-target therapies against complex diseases such as diabetes and cancer is

TABLE 1 Degree, betweenness, closeness, and binding energy values of 19 hub genes.

Gene ID	Probe ID	Name	Degree	Betweenness	Closeness	Binding energy (kcal/mol)
EGFR	5FED	Epidermal growth factor	18	0.013505	1	−6.97
MTOR	4JT6	Mammalian target of rapamycin	18	0.013505	1	−6.97
SRC	3ELD	Proto-oncogene tyrosine-protein kinase	18	0.013505	1	−7.74
HRAS	6ZL3	HRAS proto-oncogene, GTPase	18	0.013505	1	−7.14
IL6	1N26	Interleukin-6	17	0.011354	0.947368	−6.07
ESR1	7MSA	Estrogen receptor1	17	0.010683	0.947368	−8.72
JAK2	6VGL	Janus kinase 2	17	0.012629	0.947368	−8.28
AKT1	3MVH	AKT serine/threonine kinase 1	16	0.011358	0.9	−8.06
KDR	6GQQ	Kinase insert domain receptor	16	0.005546	0.9	−5.99
MAPK1	7PQV	Mitogen-activated protein kinase 1	15	0.004576	0.857143	−7.85
KIT	6GQJ	KIT proto-oncogene, receptor tyrosine kinase	15	0.004987	0.857143	−7.97
CASP3	7SEO	Caspase3	15	0.00394	0.857143	−7.30
PTPN11	3B7O	Protein tyrosine phosphatase non-receptor type 11	15	0.007317	0.857143	−6.32
MMP9	4H1Q	Matrix metallopeptidase 9	15	0.00394	0.857143	−8.85
MET	6UBW	MET proto-oncogene, receptor tyrosine kinase	15	0.00467	0.857143	−6.33
PTGS2	6BL4	Prostaglandin-endoperoxide synthase 2	14	0.00297	0.818182	−8.86
PIK3R1	5ITD	Phosphoinositide-3-kinase regulatory subunit 1	14	0.00685	0.818182	−7.42
TLR4		Toll-like receptor 4	13	0.004612	0.782609	
PDGFRA	6JOL	platelet-derived growth factor receptor alpha	10	8.75E-04	0.692308	−9.01

growing rapidly. In this regard, network pharmacology describes disease processes as networks that are best addressed by a combination of medications that function together. A network is a visual representation of a dataset that highlights the connections between the nodes. These nodes, which stand in for genes, proteins, small molecules, and other entities that can interact in the system being represented, are connected by edges, which stand in for the characteristics of the interactions, to form a graph (Berger and Iyengar, 2009). To gain a better understanding of how herbal medicines operate, network pharmacology is increasingly being used (Noor et al., 2022). Network pharmacology aims to identify genes associated with compounds and diseases, create a network of protein-protein interactions (PPIs), and then evaluate and display the network (Gao et al., 2021; Niu et al., 2021). Through the mechanistic exploration of effects on multiple disease pathways, network pharmacology represents an effective method for selecting and elucidating the synergistic effects of bioactive chemicals (Obaidullah et al., 2022). The hub genes are the highly interconnected nodes in the PPI network that are very likely to play a role in crucial biological regulation. Important biological processes are revealed using hub gene and Protein-Protein Interaction (PPI) network analysis, which offers effective methods for identifying important molecular pathways in cancer biology. Hub genes obtained from network pharmacology either represent the molecular targets of active compound (Lv and Li, 2019) or prognostic genes (Bian et al.,

2019). In this study, the molecular regulatory mechanism of action of andrographolide against colorectal cancer was predicted with the aid of network pharmacology and docking studies.

Materials and methods

Target collection for andrographolide

The potential targets of andrographolide were explored and assessed utilizing public databases such as PharmMapper (http://www.lilab-ecust.cn/pharmmapper/index.html), Swiss target prediction (https://www.swisstargetprediction.ch/), Targetnet (http://targetnet.scbdd.com/home/index/), STITCH (http://stitch.embl.de/) and SuperPred (https://prediction.charite.de/subpages/target_prediction.php). In all these databases species was limited to *Homo Sapiens* once the target has been predicted. The gene list of the targets obtained were standardized using UniProtKB (https://www.uniprot.gov/).

Identification of critical genes in CRC from disease database

Proto-oncogenes and tumour suppressor genes make up the two main classes of genes known as "cancer-critical genes,"

FIGURE 1
(A) GSE32323 volcano plot of differentially upregulated and downregulated genes with significance $p < 0.01$, Upregulated genes are shown in red colour, Downregulated genes are shown in blue colour; **(B)** GSE8671 volcano plot of differentially upregulated and downregulated genes with significance $p < 0.01$, Upregulated genes are shown in red colour, Downregulated genes are shown in blue colour; **(C)** GSE32323 top 100 up and down expressed genes heatmap, Upregulated genes are shown in red colour, Downregulated genes are shown in blue colour. **(D)** GSE8671 top 100 up and down expressed genes heatmap, Upregulated genes are shown in red colour, Downregulated genes are shown in blue colour; **(E,F)** represents network diagram (arranged according to LogFC value) of top 100 DEGs plotted using STRING and Cytoscape.

FIGURE 2
Overlap co-expression Venn map between targets of andrographolide and CRC with a total of 63 genes.

which are responsible for the development of cancer (Alberts et al., 2002). The critical genes related to Colorectal cancer was obtained using the keywords "colorectal cancer," "colorectal adenomas," and "colorectal neoplasms" in five publicly available databases such as DisGeNET (http://www.disgenet.org/) using the cutoff "score_gda >0.1," GeneCards (https://www.genecards.org/) only the genes with the relevance score greater than 30. The target gene information connected to CRC was collected and incorporated for the species "*H. sapiens.*" The integrated gene list was then uploaded in UniProtKB to obtain gene name and ID, then the duplicates were removed.

Identification of critical genes in CRC from expression dataset

GSE32323 and GSE8617 were the two-microarray expression dataset selected for identifying differentially expressed genes in CRC. The datasets were retrieved from NCBI GEO database (http://www.ncbi.nlm.nih.gov/geo/). GSE32323 consists of expression profile of 17 colorectal cancer tissue samples and 17 normal tissue samples (34 samples in total), GSE8671 consists of 32 pair (64 samples in total) of samples corresponding to Colorectal adenoma and normal tissues. GEO2R, an online tool incorporating R/Bioconductor and the Limma package v3.26.8 was used to examine the raw gene expression data. To determine the DEGs among patients with CRC and controls, we used the built-in methods in GEO2R, such as the *t*-test and Benjamini and Hochberg (false discovery rate) and force normalization of the datasets were applied to eliminate redundancy and minimize modification errors of data. Volcano plot was plotted for CRC VS Normal to represent significant up and down regulated genes. The genes with
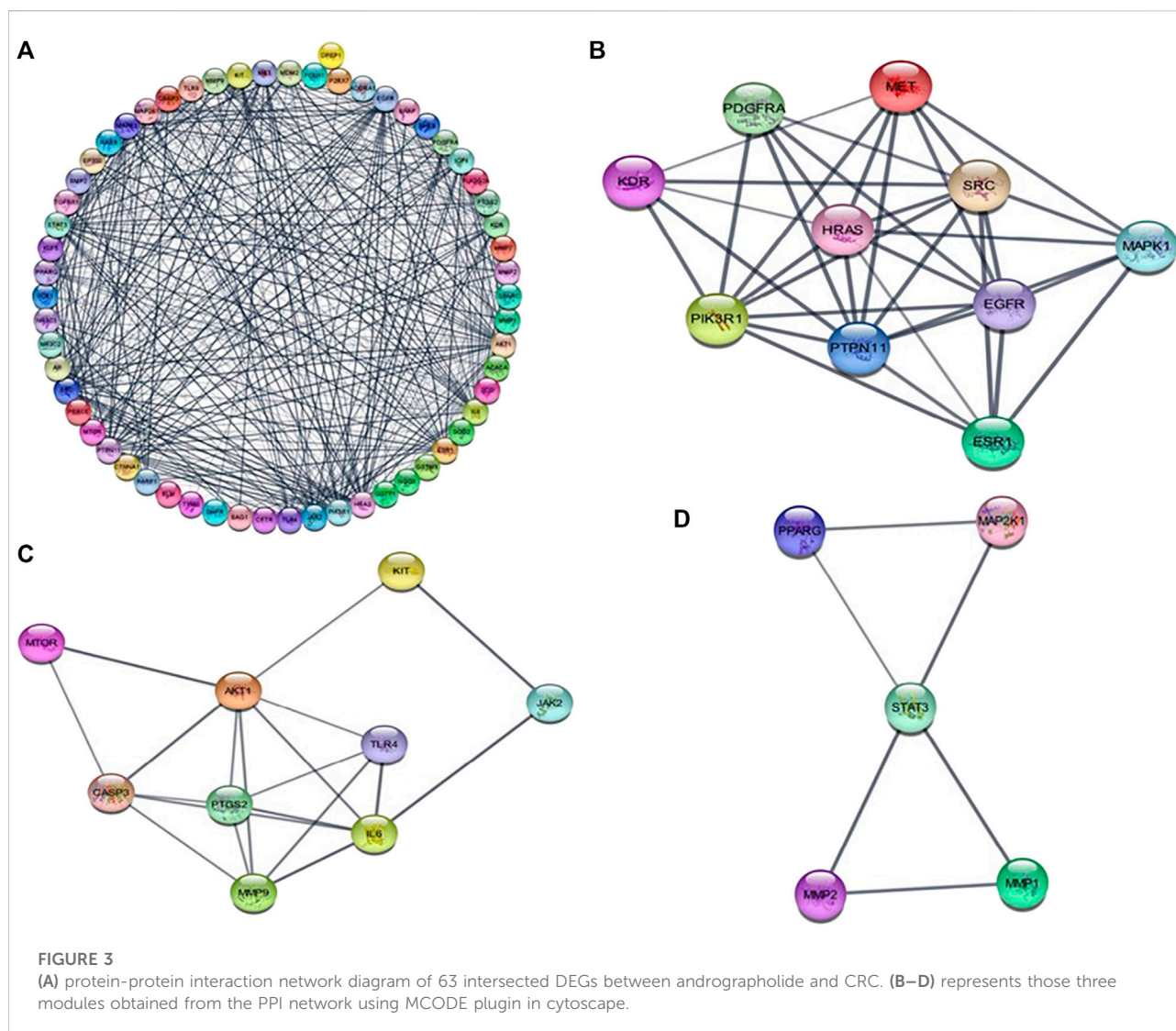
adjusted p value < 0.01 and a $\log_2 FC \geq 2$ were filtered and considered as DEGs. The DEGs obtained from the above process was used for further investigation. The heat map of the two datasets were plotted using SRplot (https://www.bioinformatics.com.cn/) and the network of the top 100 DEGs obtained from two datasets were plotted using Cytoscape. The Targets obtained from disease target database and GEO datasets (Top 100 genes) were integrated and a Venn diagram was created to identify the genes shared by targets of andrographolide and CRC targets.

PPI network construction

Using STRING a web-based tool (http://string-db.org) (v11.5), we established a PPI network to examine the relationships between DEGs acquired from the datasets (Manoochehri et al., 2021). The minimum interaction score was set at 0.09. The protein nodes that do not have any interaction with other nodes were removed from the network. Insights into biological mechanisms can be derived from the analysis of functional interactions between proteins. To identify the hub genes and key modules the PPI network from STRING was imported to Cytoscape (version 3.5.1; http://www.cytoscape.org). Considering the networks were scale-free, the biological importance of genes associated with degree centrality was examined to better understand their functionality. Network parameters such as degree, closeness and betweenness was analyzed using analyze network option in Cytoscape. Degree describes the link between one node with their adjacent nodes, the degree cut off (degree > 10) was set to find the hub genes in the network. From the resulting PPI network, we used the MCODE plugin from Cytoscape to detect intersected clusters using parameters Node cutoff score-0.2, K-score-2 and Max Depth-100.

Pathway and functional enrichment analysis

Gene Ontology (GO) analysis is now a widely used method of analyzing genomic data, especially large-scale transcriptomic data. To determine the biological function of target proteins and cancer-related pathways we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis using The Database for Annotation, Visualization, and Integrated Discovery (DAVID, https://david.ncifcrf.gov/, ver. 6.8) (Manoochehri et al., 2021). The enriched GO terms and pathways with a False Discovery Rate (FDR) of less than 0.01 was selected for further visualization. Bubble graph and of top 20 significant GO terms (BP, CC, and MF) and KEGG pathways were plotted using SRplot (https://www.bioinformatics.com.cn/).

FIGURE 3
(A) protein-protein interaction network diagram of 63 intersected DEGs between andrographolide and CRC. **(B–D)** represents those three modules obtained from the PPI network using MCODE plugin in cytoscape.
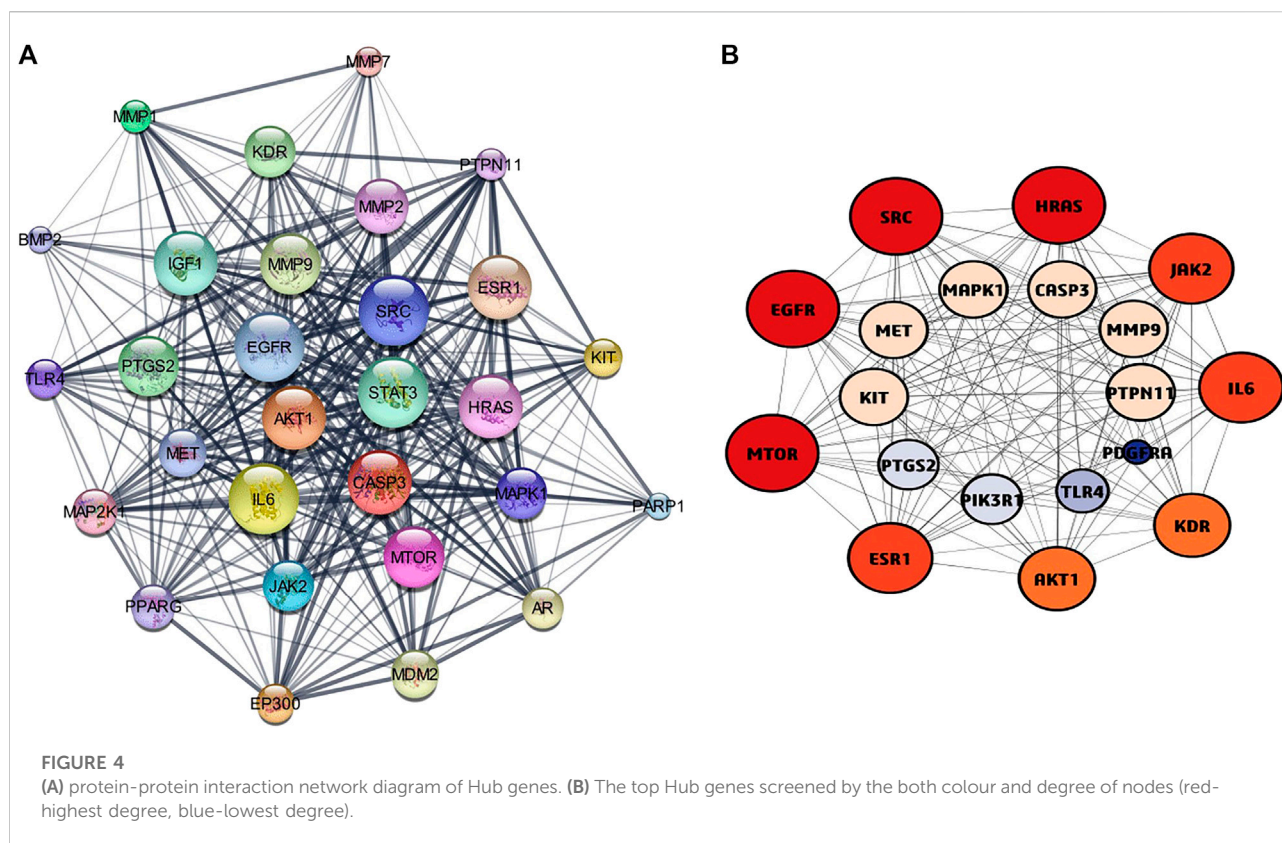
Molecular docking

To predict the interaction between hub genes and andrographolide, molecular docking was performed (Bahmani et al., 2021). PDB structures of the 19 hub genes (PDB IDs of genes in Table 1) were retrieved from RCSB Protein Data Bank (http://www.rcsb.org), whereas the 2D structure of andrographolide was retrieved from PubChem database (https://pubchem.ncbi.nlm.nih.gov/), the SDF format was converted into PDB format using Open Babel. The water molecules and native ligand groups from the protein structure were removed using Discovery Studio Visualizer, which then predicted the active site for most efficient ligand binding. Further processing of proteins and docking of proteins (10 docking runs were performed for each

protein) were performed using AutoDock Tools. A visualization of the 2D and 3D interactions between the top three protein-andrographolide complexes with the lowest binding energy was created using the Discovery Studio Visualizer.

Survival analysis

In GEPIA, a log-rank test is used to analyze overall survival by considering gene expression. In order to analyze the relationship between the hub genes and overall survival of patients with CRC, Kaplan-Meier (K-M) survival curves were plotted using GEPIA2 (http://gepia2.cancer-pku.cn/#index) for COAD (Colon Adenocarcinoma) and READ (Rectum

FIGURE 4
(A) protein-protein interaction network diagram of Hub genes. **(B)** The top Hub genes screened by the both colour and degree of nodes (red-highest degree, blue-lowest degree).

Adenocarcinoma) database. *p* values of <0.05 were considered statistically significant. The top three genes with lowest binding energy from molecular docking and genes involved in survival of CRC patients were integrated and used for further analysis.

Correlation and expression analysis

The correlation analysis between the six genes (three genes from survival analysis and three from docking) was performed using GEPIA after log2 transformation. In terms of significance levels, values less than 0.05 considered (significant-*), 0.01 (highly significant-**), and 0.001 (very highly significant-***). Based on user-defined sample selections and methods, GEPIA dynamically plots expression profiles of a given gene. A comparison of the expression profiles between tumor samples and adjacent normal samples was performed by GEPIA on the Cancer Genome Atlas (TCGA) database for colon adenocarcinoma (COAD and READ) and rectum adenocarcinoma (READ) with statistical significance *p* < 0.01.

Results

Target screening for andrographolide and CRC

After uploading the appropriate format of andrographolide, we were able to retrieve a total of 398 targets from the target database. From the disease database, 985 CRC targets were obtained with the cut-off levels for Genecards and DisGeNet being GDA > 30 and score GDA > 0.3. Following the application of the cut-off adj *p*-value <0.01 and log2FC ≥ 2 to the datasets GSE32323 and GSE8671, the number of DEGs obtained was 441 and 873, respectively. In GSE32323, 145 upregulated and 306 downregulated genes were observed, and in GSE8671, 197 upregulated and 678 downregulated genes were observed. The volcano plots were created from the DEGs in both datasets (Figure 1A). In (Figure 1B), the heatmaps of both datasets are displayed (Figures 1C, D). The network of top 100 DEGs (Figures 1E, F) from each dataset was plotted using Cytoscape. To proceed with the analysis, the top 100 DEGs from each dataset were selected. 63 genes were obtained after finding the overlap between andrographolide targets and CRC targets (Figure 2).
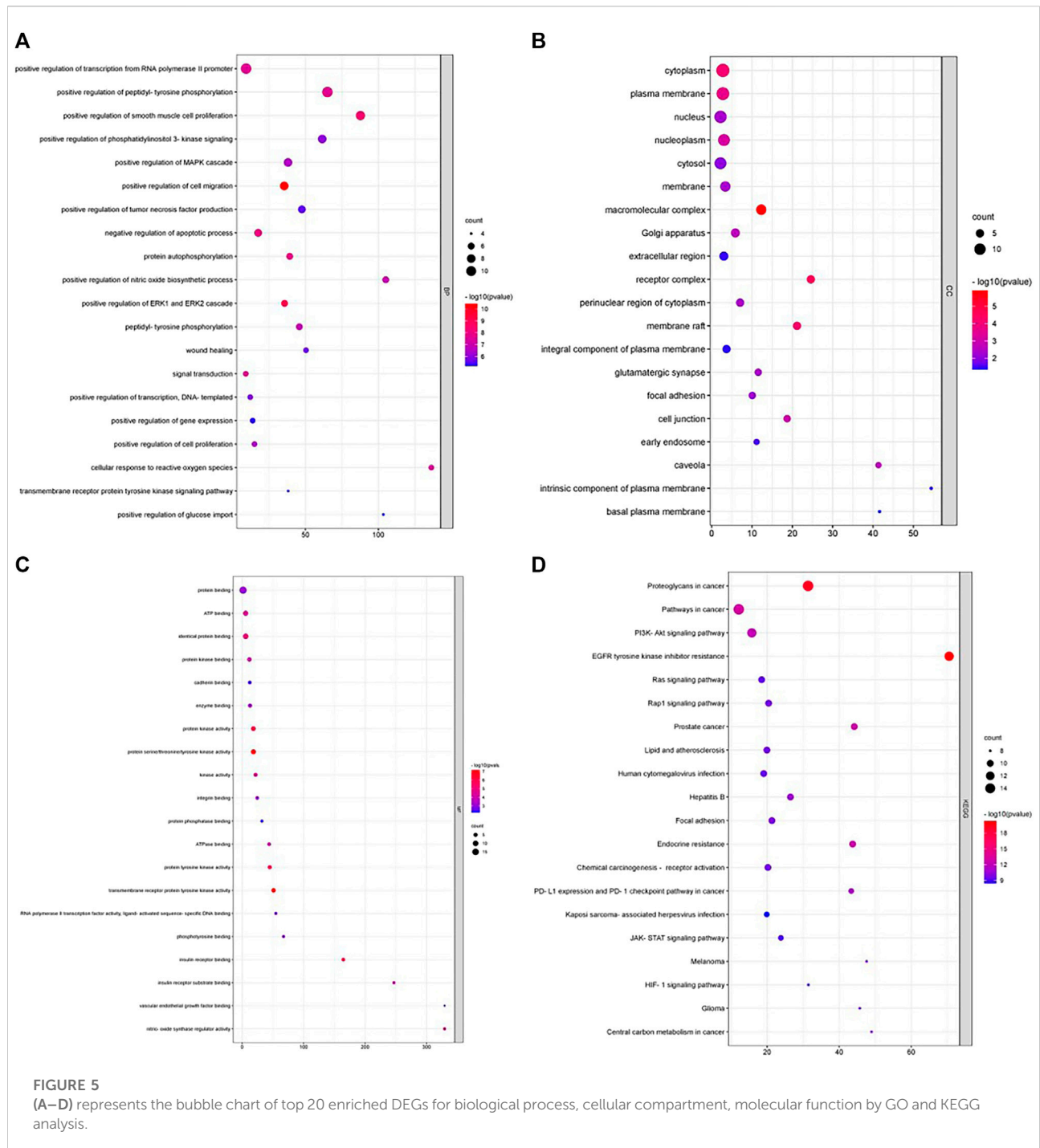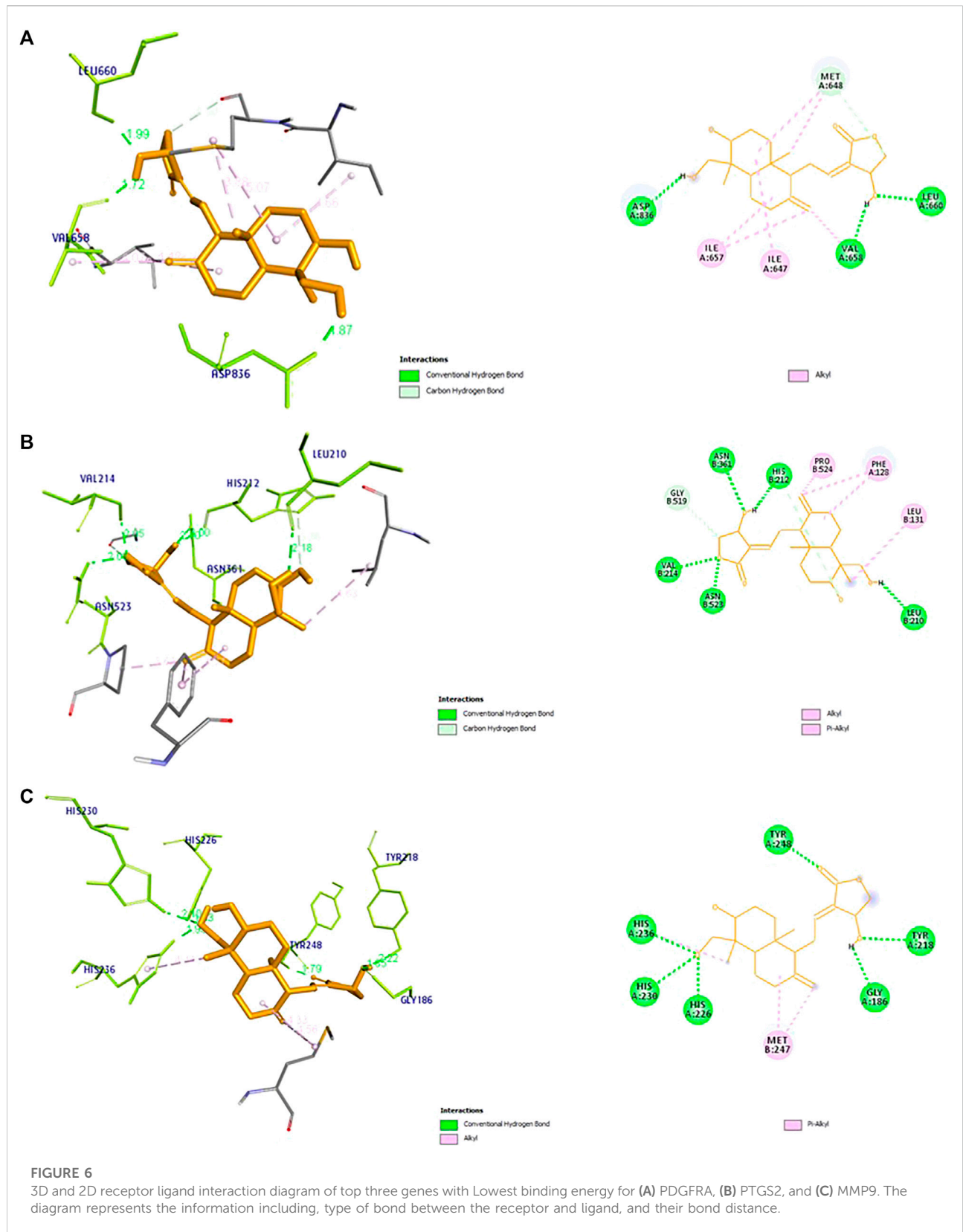
FIGURE 5
(A–D) represents the bubble chart of top 20 enriched DEGs for biological process, cellular compartment, molecular function by GO and KEGG analysis.
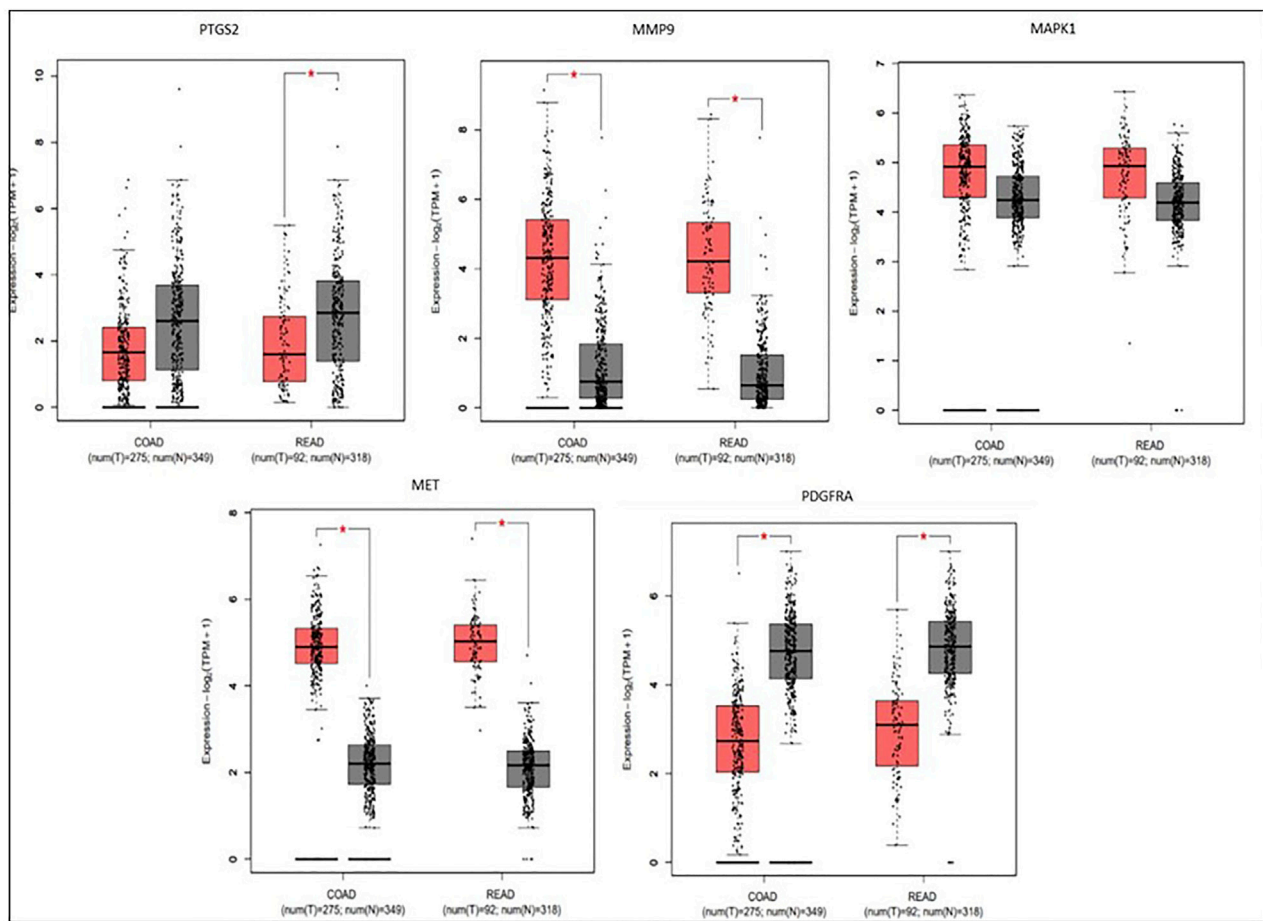
PPI network Analysis

The STRING online tool was used for the analysis of protein-protein interaction network of 63 overlapped genes (Figure 3A) with a minimum interaction score of 0.700 (highest confidence). Cytoscape software was used to construct the PPI network, the network consists of 60 nodes and 240 edges. Three modules (Figures 3B–D) were obtained from the network using MCODE application in Cytoscape software, the modules with score ≥4 and genes with degree ≥10 (Table 1) were selected for further analysis. 19 genes were obtained after filtering, these genes were considered as hub genes. A subnetwork for 19 hub genes were created (Figures 4A,B).

FIGURE 6
3D and 2D receptor ligand interaction diagram of top three genes with Lowest binding energy for **(A)** PDGFRA, **(B)** PTGS2, and **(C)** MMP9. The diagram represents the information including, type of bond between the receptor and ligand, and their bond distance.

FIGURE 7
Survival plots for three genes (MET, PTGS2, and MAPK1) which was found to be involved in the overall survival of CRC patients among 19 hub genes.
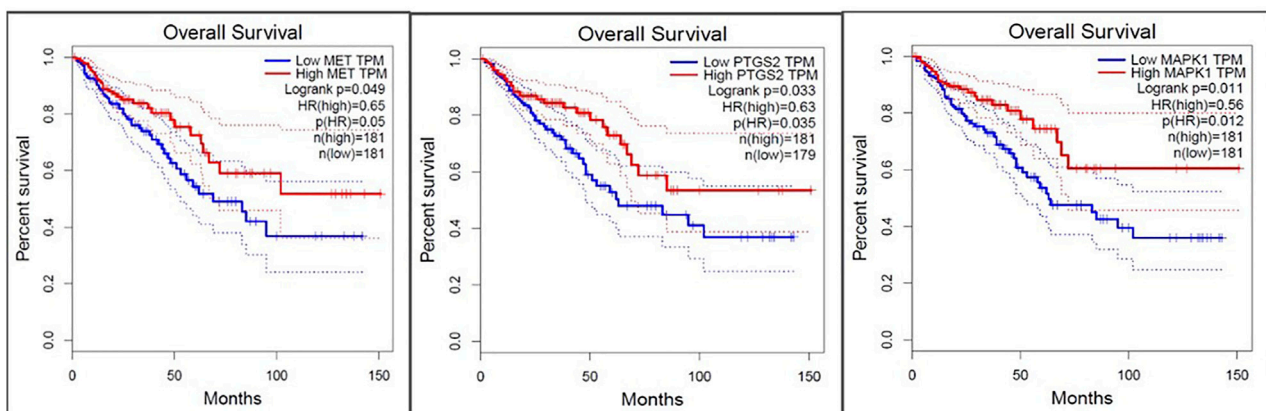
FIGURE 8
Analysis of the correlations between the five genes PTGS2, MMP9, MET, MAPK1, and PDGFRA using Pearson's coefficient analysis at $p < 0.05$. Positive correlation between the genes is shown in green colour, negative correlation between the genes is shown in red color, and genes with no correlation is shown in blue colour *-indicates the correlation is significant and ***-indicates the correlation is highly significant.
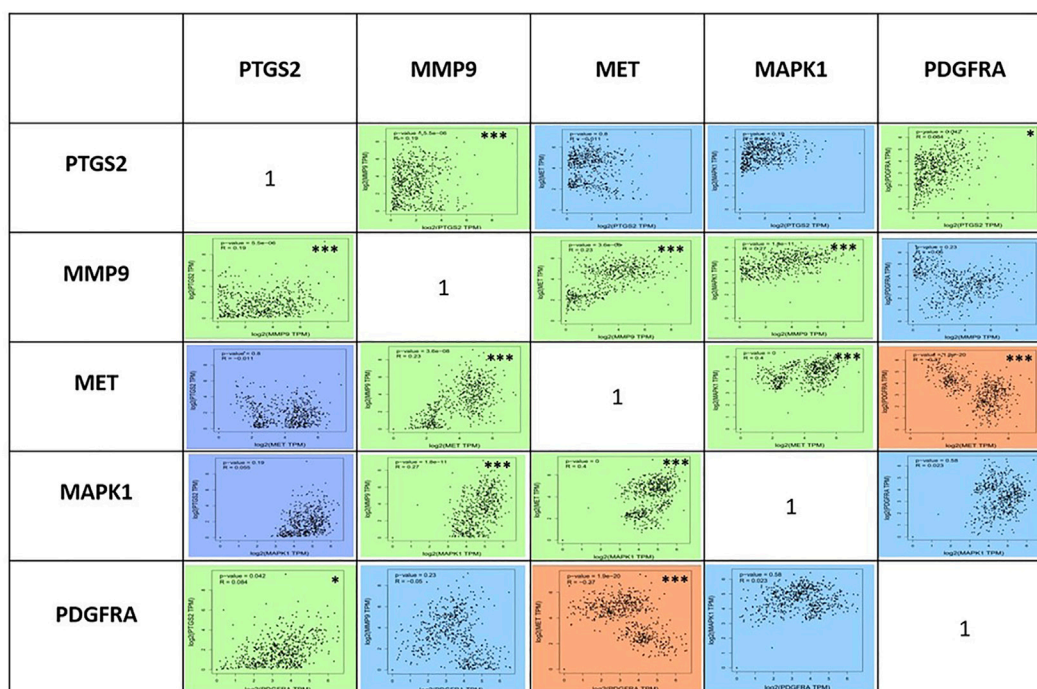
FIGURE 9
Box plots depicting expression levels of five genes in COAD-Colon Adenocarcinoma and READ-Rectum Adenocarcinoma database for tumor group (red) and normal group (grey). MMP9 and MET were over-expressed in tumor groups (COAD and READ), while PTGS2 (READ) and PDGFRA was under-expressed (COAD and READ) in tumor groups.

Functional and pathway enrichment analysis (GO and KEGG analysis)

To understand the functions and pathways of the hub genes in CRC, GO functions and pathways were analyzed with DAVID. A variety of GO enrichment terms were identified, including 204 Biological Processes (BP), 27 Cellular Components (CC) and 43 Molecular Functions (MF). A bubble chart was created displaying the top 20 enriched terms for each GO function (Figures 5A–C) and KEGG pathway (Figure 5D). The hub genes may regulate cell migration, ERK1 and ERK2 cascade, smooth muscle cell proliferation and, negative regulation of apoptotic process through molecular functions like transmembrane receptor protein tyrosine kinase activity, tyrosine kinase activity and, protein kinase activity in cellular components like macromolecular complex, membrane raft, cytoplasm, and plasma membrane. A total of 133 KEGG pathways has been found to be associated with the hub genes, among 133 pathways EGFR tyrosine kinase inhibitor resistance, Proteoglycans in cancer, Pathways in cancer, Prostate cancer, Endocrine resistance, and PI3K-Akt signaling pathway were the most significantly enriched pathways. 14 hub genes were found to be involved with PI3K-Akt signaling pathway. From the pathway study it is evident that andrographolide may exert its

anticancer activity through regulating ERK1 and ERK2 and MAPK cascade *via* PI3K-Akt signaling pathway.

Molecular docking analysis

The docking results of the 19 hub genes is shown in (Table 1). Lower binding energy values indicate more stable structures. Among the target genes PTGS2, ESR1, and PDGFRA found to have lowest binding energy. The 2D and 3D interactions between andrographolide and three target genes is shown in (Figure 6).

Survival, expression, and correlation analysis

To compare the survival rates between the groups with high and low expression, Kaplan-Meier survival curves were plotted (Figure 7) and a log-rank test applied. PTGS2, MAPK1, and MET were among the 19 hub genes with a p value less than 0.05, which indicates that these genes may have a role in colorectal cancer survival. According to the correlation analysis (Figure 8), positive correlation exists between MMP9 & PTGS2, MMP9 & MAPK1, MMP9 & MET, PTGS2 & PDGFRA and, MET & MAPK1,

whereas PGDFRA and MET have a negative correlation and there is no correlation between other gene combinations. Under expression of PTGS2 (READ) and PDGFRA (COAD and READ), over expression of MET & MMP9 (COAD and READ) were observed in the expression analysis using boxplots (Figure 9) of genes between normal and tumor samples.

Discussion

The incidence of colon cancer is increasing, along with its mortality rates (Dong et al., 2014). The cellular mechanisms involved in colon cancer constantly change during its occurrence and progression. Malignant cancer metastasizes and invades in a complicated and polytropic manner, governed by several genes (Zhang et al., 2017). To improve the survival rate of CRC patients and prevent the disease, it is imperative to understand the etiology and mechanisms of CRC progression (Dai et al., 2020). The purpose of this study was to provide insight into the mechanism responsible for the anticancer properties of andrographolide in CRC by using network pharmacology which involves investigating the links between the drug, targets, and its pathway. DEGs from datasets GSE32323 and GSE8671 were obtained using the p value (<0.01) and Log_2FC (≥2) cut off. The DEGs with positive Log_2FC values were considered upregulated and negative Log_2FC values as downregulated genes. The integrated target genes of CRC downloaded from both disease database and expression datasets (DEGs from datasets GSE32323 and GSE8671) were intersected with the targets of andrographolide to find the common genes shared by CRC and andrographolide. 63 genes were found to share targets with both andrographolide and CRC. 19 hub genes were identified from the PPI network with the help of degree cutoff and module score. According to GO functional enrichment analysis the hub genes were associated with GO terms such as regulation of cell migration, ERK1 and ERK2 cascade, smooth cell proliferation, apoptosis, macromolecular complex, receptor complex, membrane raft, cytoplasm, transmembrane receptor protein tyrosine kinase activity, tyrosine kinase activity, and protein kinase activity. While the KEGG pathways associated with the hub genes include EGFR tyrosine kinase inhibitor resistance, Proteoglycans in cancer, Pathways in cancer, endocrine resistance, prostate cancer, and PI3K-Akt signaling pathway. The top three genes with the lowest binding energy in molecular docking study (PDGFRA, PTGS2, and MMP9) and three hub genes (PDGFRA, MET, and MAPK1) were found to be involved in the overall survival of colorectal patients after survival analysis. These five genes (PDGFRA, PTGS2, MMP9, MAPK1, and MET) have been recognized as the target of andrographolide that has an association with CRC, thus making them the most reliable genes that can be used in clinical settings. These five genes are considered as target

genes that are targeted in CRC by andrographolide. The role of these target genes in CRC is explained below.

PDGFRA is a transmembrane protein that works as a normal receptor tyrosine kinase. It has an external ligand-binding domain and an intracellular tyrosine kinase domain. When PDGF binds to PDGFRAs, it dimerizes the receptor, activating its kinase activity. The phosphorylation of tyrosine receptor leads to cellular proliferation, survival, migration, and differentiation which is initiated by the Intracellular signaling cascades, such as Ras-MAPK (mitogen-activated protein kinase), phosphatidylinositol 3-kinase (PI3K), and phospholipase C (Qian et al., 2017). In colorectal carcinogenesis, the alteration of signaling via the PDGFRs family plays a significant role. Most commonly, CRC is associated with overexpression of PDGFRs in tumors. Angiogenesis, invasion, and metastasis are linked to PDGFR overexpression in CRCs, in addition to the poor survival and resistance to targeted therapy (Manzat Saplacan et al., 2017).

Cyclooxygenase-(COX) or Prostaglandin endoperoxide synthases (PTGS2) plays a significant role as a rate limiting enzyme in the conversion of arachidonic acid to prostaglandins. COX exists in two isoforms (COX-1/PTGS1 & COX-2/PTGS2) (Kunzmann et al., 2013). Even though both proteins have the same cyclooxygenase and peroxidase activity, they differ in substrates, cell expression, inhibition, intracellular location, and activation. COX-2 is not generally found in normal tissues, but it is activated by a variety of hormones, cytokines, tumor promoters, and growth factors (Hidalgo-Estévez et al., 2020). PGE2 is a major PG product of COX-2 that may regulate angiogenesis, immunity, and tumorigenesis. COX-2 may be uncontrollably elevated in carcinogenesis, either transcriptionally or post-transcriptionally 55, 56. As a result, increased COX-2 expression is a tumor diagnosis marker that is linked to patient survival. According to growing evidence COX-2 signaling appears to play an important role in the metastasis of colorectal cancer. PGE2 promotes colorectal cancer cell invasion via PI3K 71 and EGFR transactivation by SRC. Besides that, COX-2 overexpression can affect intestinal cell adhesion and promote MMP (matrix metalloproteinase) activity and therefore cancer invasion. COX-2 inhibition can prevent colorectal metastasis in humans (Sinicrope and Gill, 2004).

Matrix metalloproteinases (MMPs) regulate the degradation of extracellular matrix (ECM), which is an important stage of tumor metastasis (Duffy et al., 2000; Stetler-Stevenson and Yu, 2001). The selective proteolytic degradation of Extracellular matrix (ECM) is a critical step in tumor cell migration and invasion (Yang et al., 2014). Mechanisms including transcription, activation, and inhibition control the regulation of MMPs and over 25 types of MMPs have been discovered (Andreas Jonsson, 2018). MMP9 has been detected in substantially higher ratios in the serum of CRC patients compared to normal controls. Overexpression of the p38 gamma MAPK has been demonstrated to enhance cell invasion by increasing MMP9 transcription (Loesch et al., 2010). MMP-9

levels have been proposed as a biological predictor of prognosis in CRC and other cancers such as breast and cervical cancer (Andreas Jonsson, 2018).

MAPK1/ERK1 belongs to the Ser/Thr kinase family, which activates numerous rounds of phosphorylation-activating kinase rings from the cell surface to the nucleus (Koveitypour et al., 2019). Many biological activities, including cell proliferation, differentiation, migration, and death, are regulated by mitogen-activated protein kinase (MAPK) pathways. However, genome-wide association research undertaken in Germany found the MAPK-signaling pathways as one of the most significantly related gene markers with colorectal cancer (CRC) (Slattery et al., 2012). Hepatocyte Growth Factor/Scatter Factor (HGF/SF) belongs to mesenchymal cytokine with pleiotropic effects, including mitogenic, monogenic, and morphogenic features. c-MET (or MET) encodes a receptor tyrosine kinase for HGF/SF (Gayyed et al., 2015). Under normal physiological conditions, the interaction of HGF/cMet powerfully mediates cell division and cellular organization, such as epithelial cell proliferation and migration, angiogenesis, and wound healing. In different types of malignant tumors HGF/cMet pathway plays a key role in tumour proliferation, invasion, and metastasis (Yao et al., 2019). In colorectal cancer, c-Met is associated with tumour invasiveness and aggressiveness, as well as tumor progression and poor prognosis (Liu et al., 2015).

Among the five critical genes previously obtained from survival and molecular docking analysis, MMP9 & MET were the upregulated in tumor samples (COAD and READ), while PDGFRA (COAD and READ) & PTGS2/COX-2 (READ) were downregulated according to the expression analysis by box plots. The overexpression of MET and under expression of PDGFRA and PTGS2/COX-2 was associated with overall survival of CRC patients. The above results were supported by the correlation analysis between genes which infers that there is a positive correlation between MET and MMP9 (both the genes were overexpressed in CRC), and negative correlation between PDGFRA and MET (one gene is under expressed while the other gene is overexpressed).

Conclusion

Overall, findings seem to indicate that andrographolide may inhibit CRC *via* a diverse set of targets and pathways. Both GO annotation and KEGG pathway enrichment suggest that andrographolide may have a therapeutic role in CRC through inhibition of cellular growth, proliferation, and migration and induction of apoptosis by targeting the hub genes (PDGFRA, PTGS2, MMP9, MAPK1, and MET) which is involved in cancer migration and invasion. Based on this multidisciplinary strategy, the current study provided a promising approach for the treatment of disease using colorectal cancer. However, to clarify the relationship between andrographolide and hub genes, along with the specific mechanisms underlying those actions, more pre and clinical experiments are necessary.

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.

Author contributions

BJ, MV, and KP involved in target screening for andrographolide and CRC, and PPI network analysis. AG, MA, and ME-B contributed in functional and pathway enrichment analysis. MS, AM, and SA-G involved in molecular docking analysis, Survival, correlation, and expression analysis. JK, TR, and MA were involved in the interpretation of the scientific data and performed quality checks. All the authors contributed to the manuscript writing.

Acknowledgments

Conflict 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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Afshar, S., Pashaki, A. S., Najafi, R., Nikzad, S., Amini, R., Shabab, N., et al. (2020). Cross-resistance of acquired Radioresistant colorectal Cancer cell line to gefitinib and regorafenib. *Iran. J. Med. Sci.* 45 (1), 50–58. doi:10.30476/ijms.2019.44972

Aiello, P., Sharghi, M., Mansourkhani, S. M., Ardekan, A. P., Jouybari, L., Daraei, N., et al. (2019). Medicinal plants in the prevention and treatment of colon cancer. *Oxid. Med. Cell. Longev.* 2019, 2075614. doi:10.1155/2019/2075614

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). "Molecular biology of the cell," in *Finding the cancer-critical genes*. 4th edition (New York: Garland Science).

Bahmani, A., Tanzadehpanah, H., Moghadam, N. H., and Saidijam, M. (2021). Introducing a pyrazolopyrimidine as a multi-tyrosine kinase inhibitor, using multi-QSAR and docking methods. *Mol. Divers.* 25 (2), 949–965. doi:10.1007/s11030-020-10080-8

Berger, S. I., and Iyengar, R. (2009). Network analyses in systems pharmacology. *Bioinformatics* 25 (19), 2466–2472. doi:10.1093/bioinformatics/btp465

Bian, Q., Chen, J., Qiu, W., Peng, C., Song, M., SunZhang, X. L., et al. (2019). Four targeted genes for predicting the prognosis of colorectal cancer: A bioinformatics analysis case. *Oncol. Lett.* 18, 5043–5054. doi:10.3892/ol.2019.10866

Choudhari, A. S., Mandave, P. C., Deshpande, M., Ranjekar, P., and Prakash, O. (2019). Phytochemicals in cancer treatment: From preclinical studies to clinical practice. *Front. Pharmacol.* 10, 1614. doi:10.3389/fphar.2019.01614

Dai, G. P., Wang, L. P., Wen, Y. Q., Ren, X. Q., and Zuo, S. G. (2020). Identification of key genes for predicting colorectal cancer prognosis by integrated bioinformatics analysis. *Oncol. Lett.* 19, 388–398. doi:10.3892/ol.2019.11068

Dekker, E., Tanis, P. J., Vleugels, J. L. A., Kasi, P. M., and Wallace, M. B. (2019). Colorectal cancer. *Lancet* 394, 1467–1480. doi:10.1016/S0140-6736(19)32319-0

Dong, M., Yang, G., Liu, H., Liu, X., Lin, S., Sun, D., et al. (2014). Aged black garlic extract inhibits HT29 colon cancer cell growth via the PI3K/Akt signaling pathway. *Biomed. Rep.* 2, 250–254. doi:10.3892/br.2014.226

Duffy, M. J., Maguire, T. M., Hill, A., Mcdermott, E., and O'higgins, N. (2000). Metalloproteinases: role in breast carcinogenesis, invasion and metastasis. *Breast Cancer Res.* 2, 252–257. doi:10.1186/bcr65

Fares, J., Fares, M. Y., Khachfe, H. H., Salhab, H. A., and Fares, Y. (2020). Molecular principles of metastasis: a hallmark of cancer revisited. *Signal Transduct. Target. Ther.* 5, 28. doi:10.1038/s41392-020-0134-x

Gao, X., Li, S., Cong, C., Wang, Y., and Xu, L. (2021). A network pharmacology approach to estimate potential targets of the active ingredients of epimedium for alleviating mild cognitive impairment and treating alzheimer's disease. *Evid. Based. Complement. Altern. Med.* 2021, 2302680. doi:10.1155/2021/2302680

Gayyed, M. F., Abd El-Maqsoud, N. M. R., El-Hameed El-Heeny, A. A., and Mohammed, M. F. (2015). c-MET expression in colorectal adenomas and primary carcinomas with its corresponding metastases. *J. Gastrointest. Oncol.* 6, 618–627. doi:10.3978/j.issn.2078-6891.2015.072

Hidalgo-Estévez, A. M., Stamatakis, K., Jiménez-Martínez, M., López-Pérez, R., and Fresno, M. (2020). Cyclooxygenase 2-regulated genes an alternative avenue to the development of new therapeutic drugs for colorectal cancer. *Front. Pharmacol.* 11, 533. doi:10.3389/fphar.2020.00533

Huang, X. M., Yang, Z. J., Xie, Q., Zhang, Z. K., Zhang, H., and Ma, J. Y. (2019). Natural products for treating colorectal cancer: A mechanistic review. *Biomed. Pharmacother.* 117, 109142. doi:10.1016/j.biopha.2019.109142

Jayakumar, T., Hsieh, C. Y., Lee, J. J., and Sheu, J. R. (2013). Experimental and clinical pharmacology of andrographis paniculata and its major bioactive phytoconstituent andrographolide. *Evid. Based. Complement. Altern. Med.* 2013, 846740. doi:10.1155/2013/846740

Johnson, C. M., Wei, C., Ensor, J. E., Smolenski, D. J., Amos, C. I., Levin, B., et al. (2013). Meta-analyses of colorectal cancer risk factors. *Cancer Causes Control* 24, 1207–1222. doi:10.1007/s10552-013-0201-5

Jonsson, A., Hjalmarsson, C., Falk, P., and Ivarsson, M-L. (2018). Stability of matrix metalloproteinase-9 as biological marker in colorectal cancer. *Med Oncol.* 35, 50. doi:10.1007/s12032-018-1109-4

Koveitypour, Z., Panahi, F., Vakilian, M., Peymani, M., Seyed Forootan, F., Nasr Esfahani, M. H., et al. (2019). Signaling pathways involved in colorectal cancer progression. *Cell Biosci.* 9, 97. doi:10.1186/s13578-019-0361-4

Kunzmann, A. T., Murray, L. J., Cardwell, C. R., Mcshane, C. M., Mcmenamin, U. C., and Cantwell, M. M. (2013). PTGS2 (Cyclooxygenase-2) expression and survival among colorectal cancer patients: a systematic review. *Cancer Epidemiol. Biomarkers Prev.* 22, 1490–1497. doi:10.1158/1055-9965.EPI-13-0263

Liu, Y., Yu, X.-F., Zou, J., and Luo, Z.-H. (2015). Prognostic value of c-met in colorectal cancer: a meta-analysis. *World J. Gastroenterol.* 21, 3706–3710. doi:10.3748/wjg.v21.i12.3706

Loesch, M., Zhi, H. Y., Hou, S. W., Qi, X. M., Li, R. S., Basir, Z., et al. (2010). p38gamma MAPK cooperates with c-Jun in trans-activating matrix metalloproteinase 9. *J. Biol. Chem.* 285, 15149–15158. doi:10.1074/jbc.M110.105429

Lv, J., and Li, L. (2019). Hub genes and key pathway identification in colorectal cancer based on bioinformatic analysis. *Biomed. Res. Int.* 2019, 1545680–1546133. doi:10.1155/2019/1545680

Manoochehri, H., Jalali, A., Tanzadehpanah, H., Taherkhani, A., and Saidijam, M. (2021a). Identification of key gene targets for sensitizing colorectal cancer to chemoradiation: an integrative network analysis on multiple transcriptomics data. *J. Gastrointest. Cancer*, 1–20. doi:10.1007/s12029-021-00690-2

Manoochehri, H., Asadi, S., Tanzadehpanah, H., Sheykhhasan, M., and Ghorbani, M. (2021b). CDC25A is strongly associated with colorectal cancer stem cells and poor clinical outcome of patients. *Gene Rep.* 25, 101415. doi:10.1016/j.genrep.2021.101415

Manzat Saplacan, R. M., Balacescu, L., Gherman, C., Chira, R. I., Craiu, A., Mircea, P. A., et al. (2017). The role of PDGFs and PDGFRs in colorectal cancer. *Mediat. Inflamm.* 2017, 4708076. doi:10.1155/2017/4708076

Mármol, I., Sánchez-De-Diego, C., Pradilla Dieste, A., Cerrada, E., and Rodriguez Yoldi, M. J. (2017). Colorectal carcinoma: A general overview and future perspectives in colorectal cancer. *Int. J. Mol. Sci.* 18, E197. doi:10.3390/ijms18010197

Moradi, M., Najafi, R., Amini, R., Solgi, R., Tanzadehpanah, H., Esfahani, A. M., et al. (2019). Remarkable apoptotic pathway of hemiscorpius lepturus scorpion venom on CT26 cell line. *Cell Biol. Toxicol.* 35 (4), 373–385. doi:10.1007/s10565-018-09455-3

Niu, W. H., Wu, F., Cao, W. Y., Wu, Z. G., Chao, Y. C., Liang, C., et al. (2021). Network pharmacology for the identification of phytochemicals in traditional Chinese medicine for COVID-19 that may regulate interleukin-6. *Biosci. Rep.* 41, BSR20202583. doi:10.1042/BSR20202583

Noor, F., Tahir Ul Qamar, M., Ashfaq, U. A., Albutti, A., Alwashmi, A. S. S., and Aljasir, M. A. (2022). Network pharmacology approach for medicinal plants: Review and assessment. *Pharm. (Basel)* 15, 572. doi:10.3390/ph15050572

Obaidullah, A. J., Alanazi, M. M., Alsaif, N. A., Alanazi, A. S., Albassam, H., Az, A., et al. (2022). Network pharmacology- and molecular docking-based identification of potential phytocompounds from *Argyreia capitiformis* in the treatment of inflammation. *Evid. Based Complement. Altern. Med.* 2022, 8037488. doi:10.1155/2022/8037488

Okhuarobo, A., Falodun, J. E., Erharuyi, O., Imieje, V., Falodun, A., and Langer, P. (2014). Harnessing the medicinal properties of *Andrographis paniculata* for diseases and beyond: a review of its phytochemistry and pharmacology. *Asian pac. J. Trop. Dis.* 4 (3), 213–222. doi:10.1016/s2222-1808(14)60509-0

Paul, T., Basu, S., and Saha, N. C. (2019). Anticancer effect of *Andrographis paniculata* by suppression of tumor altered hypoxia signaling cascade in mouse melanoma cells. *J. Cancer Res. Pract.* 6, 117–123. doi:10.4103/jcrp.jcrp_9_19

Qian, C., Wong, C. W. Y., Wu, Z., He, Q., Xia, H., Tam, P. K. H., et al. (2017). Stage specific requirement of platelet-derived growth factor receptor-α in embryonic development. *PLoS One* 12, e0184473. doi:10.1371/journal.pone.0184473

Sawicki, T., Ruszkowska, M., Danielewicz, A., Niedźwiedzka, E., Arłukowicz, T., and Przybyłowicz, K. E. (2021). A review of colorectal cancer in terms of epidemiology, risk factors, development, symptoms and diagnosis. *Cancers (Basel)* 13, 2025. doi:10.3390/cancers13092025

Shimura, T., Sharma, P., Sharma, G. G., Banwait, J. K., and Goel, A. (2021). Enhanced anti-cancer activity of andrographis with oligomeric proanthocyanidins through activation of metabolic and ferroptosis pathways in colorectal cancer. *Sci. Rep.* 11, 7548. doi:10.1038/s41598-021-87283-y

Sinicrope, F. A., and Gill, S. (2004). Role of cyclooxygenase-2 in colorectal cancer. *Cancer Metastasis Rev.* 23, 63–75. doi:10.1023/a:1025863029529

Slattery, M. L., Lundgreen, A., and Wolff, R. K. (2012). MAP kinase genes and colon and rectal cancer. *Carcinogenesis* 33, 2398–2408. doi:10.1093/carcin/bgs305

Stetler-Stevenson, W. G., and Yu, A. E. (2001). Proteases in invasion: matrix metalloproteinases. *Semin. Cancer Biol.* 11, 143–152. doi:10.1006/scbi.2000.0365

Woods, D., and Turchi, J. J. (2013). Chemotherapy induced DNA damage response: convergence of drugs and pathways. *Cancer Biol. Ther.* 14, 379–389. doi:10.4161/cbt.23761

Yang, B., Tang, F., Zhang, B., Zhao, Y., Feng, J., and Rao, Z. (2014). Matrix metalloproteinase-9 overexpression is closely related to poor prognosis in patients with colon cancer. *World J. Surg. Oncol.* 12, 24. doi:10.1186/1477-7819-12-24

Yao, J. F., Li, X. J., Yan, L. K., He, S., Zheng, J. B., Wang, X. R., et al. (2019). Role of HGF/c-Met in the treatment of colorectal cancer with liver metastasis. *J. Biochem. Mol. Toxicol.* 33, e22316. doi:10.1002/jbt.22316

Zhang, R., Zhao, J., Xu, J., Jiao, D.-X., Wang, J., Gong, Z.-Q., et al. (2017). Andrographolide suppresses proliferation of human colon cancer SW620 cells through the TLR4/NF-κB/MMP-9 signaling pathway. *Oncol. Lett.* 14, 4305–4310. doi:10.3892/ol.2017.6669