Identification of Key Genes as Potential Drug Targets for Gastric Cancer

Md. Tofazzal Hossain[†], Md. Selim Reza[†], Yin Peng, Shengzhong Feng, and Yanjie Wei*

Abstract: Gastric cancer (GC) is one of the most common cancers and ranks the third in cancer mortality all over the world. The goal of this study was to identify potential hub-genes, highlighting their functions, signaling pathways, and candidate drugs for the treatment of GC patients. We used publicly available next generation sequencing (NGS) data to identify differentially expressed (DE) genes. The top DE genes were mapped to STRING database to construct the protein-protein interaction (PPI) network and top hub genes were selected for further analysis. We found a total of 1555 DE genes with 870 upregulated and 685 downregulated genes in GC. We selected the top 400 (200 upregulated and 200 downregulated) genes to construct a PPI network and extracted the top 15 hub genes. The gene ontology (GO) term and kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment analyses of the 15 hub genes exposed some important functions and signaling pathways that were significantly associated with GC patients. The survival analysis of the hub genes disclosed that the lower expressions of the three hub genes CDH2, COL4A1, and COL5A2 were associated with better survival of GC patients. These three genes might be the candidate biomarkers for the diagnosis and treatment of GC. Then, we considered 3 key proteins (genomic biomarkers) (COL4A1, CDH2, and CO5A2) as the drug target proteins (receptors), performed their docking analysis with the 102 meta-drug agents, and found Everolimus, Docetaxel, Lanreotide, Venetoclax, Temsirolimus, and Nilotinib as the top ranked 6 candidate drugs with respect to our proposed target proteins for the treatment against GC patients. Therefore, the proposed drugs might play vital role for the treatment against GC patients.

Key words: gastric cancer; hub genes; candidate genes; molecular docking; candidate drugs

1 Introduction

Gastric cancer (GC) is one of the most common malignant tumors and the third leading cause of cancer related mortality globally[1]. The prognosis of GC is still poor despite of having current advanced treatment, and the overall survival rate is not reached higher than 30%[2]. Molecular heterogeneity of GC patients is an obstacle in clinical diagnosis and developing treatments[3]. Previous

● Md. Tofazzal Hossain, Md. Selim Reza, Shengzhong Feng, and Yanjie Wei are with the Center for High Performance Computing, Joint Engineering Research Center for Health Big Data Intelligent Analysis Technology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China. E-mail: tofazzal.stat@gmail.com; selim@siat.ac.cn; sz.feng@siat.ac.cn; yj.wei@siat.ac.cn.

● Md. Tofazzal Hossain and Md. Selim Reza are also with the School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China.

● Md. Tofazzal Hossain is also with the Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj 8100, Bangladesh.

● Yin Peng is with the Guangdong Provincial Key Laboratory for Genome Stability and Disease Prevention and Regional Immunity and Diseases, Department of Pathology, Shenzhen University, Shenzhen 518060, China. E-mail: ypeng@szu.edu.cn.

† Md. Tofazzal Hossain and Md. Selim Reza contribute equally to this paper.

∗ To whom correspondence should be addressed.

Manuscript received: 2022-02-18; revised: 2022-05-31; accepted: 2022-08-29

studies reported that high intake of salt and salty food and inadequate intake of fresh fruit and vegetables are the risk factors for GC[4, 5]. Smoking, alcohol consumption, and inactive or limited physical activity also increase the risk of GC[6]. Currently available GC treatments include surgery[7], radiotherapy[8], neoadjuvant chemotherapy[9], and immunotherapy[10]. Most of the GC patients are diagnosed at advanced stages. Lack of early detection of GC is a leading cause of poor survival rate after diagnosis[11]. Therefore, identification of potential biomarkers for GC diagnosis and finding potential drug targets are of great interest.

Present strategies to identify combinatorial anti-cancer therapies are dependent on large scale experimental data for patient treatment[12] or for suggesting new drug combinations[13–16]. But generating large scale experimental data is a time consuming, expensive, and challenging task. Computational models are also used in predicting drug effects[14, 17] to reduce the experimental obstacles. Drug repurposing (DR) is a promising approach to overcome experimental obstacles in discovering and developing new drugs[18–23]. It is considered as the supporting process to the conventional drug discovery. To explore more suitable repurposable drugs for a new disease, it requires identifying appropriate target proteins associated with the new disease.

In the current study, we used DR approach in finding potential drug targets. We analyzed publicly available next generation sequencing (NGS) data collected from GC tissue and adjacent normal tissue. After finding the differentially expressed (DE) genes, we mapped the top 400 (sorted based on adjusted *p*-value) genes to STRING[24] database to construct a protein-protein interaction (PPI) network. From the PPI network, the top 15 hub genes were selected for further analysis. The gene ontology (GO) term and kyoto encyclopedia of genes and genomes (KEGG) pathway analyses of the hub genes revealed that some hub genes were enriched in several important biological processes, molecular mechanisms, cellular components, and GC related pathways. The expressions of these hub genes were validated by the cancer genome atlas (TCGA) data. The survival analysis of the hub genes using the TCGA data showed that the three genes CDH2, COL4A1, and COL5A2 were associated with the prognosis of GC. Overall, the genes CDH2, COL4A1, and COL5A2 might be the potential candidates for GC prognosis and treatment. Finally, we analyzed the molecular docking of key genes with meta-drug agents, and suggested therapeutic targets with lead small compounds against GC.

2 Method and Material

2.1 RNA-seq data analysis

The RNA-seq data were downloaded from the gene expression omnibus (GEO) database under accession number GSE152309. These data were collected from the 5 GC patient's fresh tumor tissues and paired adjacent non-tumor tissues. After downloading the raw data in fastq format, the quality control analysis was performed. The quality control was done using the NGSQCToolkit[25] and the quality score 20 was used as the cutoff point. Filtering/trimming was carried out if the quality of the reads failed to reach the cutoff point. The fastq reads were aligned to the human reference genome (version GRCh38) using the BWA[26] aligner. Then count data were generated using featureCount[27] with the output of BWA (converted to bam and sorted).

2.2 Collection of meta-drug agents for exploring candidate drugs

We collected host transcriptome-guided 102 meta-drug agents by the literature review against GC patients (Table S1, which is in the Electronic Supplementary Material (ESM) of the online version of this article) for exploring candidate drugs[28–32]. Thus, we considered 102 drug agents to explore candidate drugs by molecular docking with the identified proteins.

2.3 Identification of differentially expressed genes

The raw count data obtained from the tool featureCounts were used for differential expression analysis. The *R* package DESeq2[33] was used to identify DE genes between two groups (normal vs. cancer). |Log2Fold Change| > 2 and padj (adjusted *p*-value) < 0.05 were considered as the cut-off for defining significant DE genes. Among the DE genes, we selected top 400 genes sorted by adjusted *p*-value for performing the PPI network analysis.

2.4 Protein-protein interaction analysis of DE genes

STRING§ is a database for obtaining the PPI between the predicted and experimental interactions of proteins. The top 400 DE genes were mapped to the STRING database and a PPI network was constructed using Cytoscape[34].

---

§ https://string-db.org/

From the PPI network, we constructed a sub-network with the top 15 hub genes using the Cytoscape plugin Cytohuba. The hub genes were then extracted from the sub-network.

2.5 Functional enrichment analysis

Biological processes, cellular components, and molecular functions of the DE genes were investigated by the well-known GO analysis. KEGG pathway analysis was done to evaluate how the DE genes were involved in or influenced by the GC related pathways. We used DAVID[35] to perform the GO and KEGG pathway analyses. The cutoff p-value<0.05 was used for the significance of the functional enrichment analysis.

2.6 Expression and survival analyses of the hub genes

We examined the expression of hub genes in GC tissue samples compared to normal tissue samples using GEPIA[36] database. GEPIA is a new web-based tool that uses gene expressions from TCGA database to compare the expression profiles of genes between normal and cancer samples. In GEPIA, there were 408 GC tumor tissue samples and 36 normal tissue samples from the TCGA database. The default cutoffs |Log2FoldChange| > 1 and p-value<0.01 were considered as statistically significant. The overall survival analysis of the hub genes in GC was also performed by the GEPIA tool. There were survival data of 381 patients with GC in the GEPIA database. Log rank $P < 0.05$ was considered to be statistically significant.

2.7 Association of key genes (KGs) with other risk factors

The Disease-KGs enrichment analysis was performed using the Enrichr web tool[37] with DisGeNET database[38] to explore other disease risk factors for GC patients. The significance level was set to p-value<0.05.

2.8 Drug repurposing by molecular docking study

We performed a molecular docking analysis of our suggested receptor proteins with drug agents to propose in-silico validated efficient candidate drugs for the treatment of GC. As previously mentioned in the data sources (Table S1 in the ESM), we considered our proposed genes based key proteins (KPs) as drug target proteins and 102 meta-drug agents. Both receptor proteins and meta-drug agents require 3-dimensional (3D) structures for molecular docking studies. All of the targeted proteins' 3D structures were downloaded from the protein data bank (PDB)[39] and SWISS-MODEL[40]. All meta-drug agents' 3D structures were downloaded from the PubChem database[41]. Using discovery studio visualizer 2019[42], the 3D structures of the target proteins were displayed, and the target chains that were not part of the genes were deleted. Every protein was defined as a receptor, and the proteins' active sites were found from the receptor cavities using the discovery studio tool. The protonation state of protein was assigned using the PDB2PQR and H++ servers[43, 44]. All the absent hydrogen atoms were properly added as well. The pKa for the receptor amino acids were examined under the physical conditions of pH = 7, salinity = 0.15, external dielectric = 80, and internal dielectric = 10. Then, using AutoDock tools, the receptor was prepared for molecular docking study by eliminating water molecules and ligand heteroatoms and by addition of polar hydrogens[45]. The ligands were prepared for molecular docking study by using AutoDock tools to set the torsion tree and rotatable and nonrotatable bonds in the ligand. AutoDock Vina was used to calculate binding affinities between target proteins and drug agents[46]. The AutoDock Vina scoring functions was as given below.

$$E = \sum_{i}^{ligand} \sum_{j}^{protein} e_{pair}(d_{ij}) \tag{1}$$

Here, d is the surface distance calculated with Eq. (2).

$$d_{ij} = r_{ij} - R_j - R_j \tag{2}$$

where r is the interatomic distance, and R_i and R_j are the radii of the pair's atoms. Every atom pair interacts via a steric interaction described by the 1st component of Eq. (3). In addition, depending on the atom type, hydrophobic and non-directional H-bonding interactions may exist, as indicated by the last two components of Eq. (3).

$$\sum_{i}^{ligand} \sum_{j}^{protein} e_{pair}(d_{ij}) = \sum_{i}^{ligand} \sum_{j}^{protein} w_1 \times gauss_1(d_{ij}) +$$

$$w_2 \times gauss_2(d_{ij}) + w_3 \times repulsion(d_{ij}) +$$

$$\sum_{i}^{ligand} \sum_{j}^{protein} w_4 \times hydrophobic(d_{ij}) +$$

$$\sum_{i}^{ligand} \sum_{j}^{protein} w_5 \times Hbond(d_{ij}) \tag{3}$$

where w_1, w_2, w_3, w_4, and w_5 are weight values equal to -0.0356, -0.00516, 0.840, -0.0351, and -0.587, respectively.

The combination of an attractive Gaussian function with a repulsive parabolic function reproduces the general shape of a typical Lennard-Jones interaction, provided the Gaussian term is negative and the parabolic positive. If both atoms in the pair are hydrophobic, the linear function in Eq. (7) is included. Also, if the pair consists of an H-bond donor and an H-bond acceptor, Eq. (8) is added.

$$gauss_1 = e^{-\left(\frac{d}{0.5}\right)^2} \tag{4}$$

$$gauss_2 = e^{-\left(\frac{d-3}{2}\right)^2} \tag{5}$$

$$repulsion(d) = \begin{cases} d^2, & \text{if } d < 0; \\ 0, & \text{if } d \geqslant 0 \end{cases} \tag{6}$$

$$hydrophobic(d) = \begin{cases} 1.0, & \text{if } d < 0.5; \\ 1.5 - d, & \text{if } 0.5 \leqslant d \leqslant 1.5; \\ 0, & \text{if } d > 1.5 \end{cases} \tag{7}$$

$$Hbond(d) = \begin{cases} 1.0, & \text{if } d < -0.7; \\ \frac{d}{-0.7}, & \text{if } -0.7 \leqslant d \leqslant 0; \\ 0, & \text{if } d > 0 \end{cases} \tag{8}$$

The exhaustiveness parameter was set to 10. PyMol[47] and discovery studio visualizer 2019[42] were used to analyze the docked complexes for surface complexes, types, and distances of non-covalent bonds. Let A_{ij} denote the binding affinity between the i-th target protein $(i = 1, 2, \ldots, m)$ and the j-th drug agent $(j = 1, 2, \ldots, n)$. To select the top-ranked lead compounds as the candidate drugs, we ordered the drug target proteins and agents according to the descending order of row sums $\sum_{j=1}^{n} A_{ij}, i = 1, 2, \ldots, m$ and column sums $\sum_{i=1}^{m} A_{ij}, j = 1, 2, \ldots, n$, respectively. The discovery studio visualizer 2019 and PyMol software[42, 48] were used to examine the hydrogen bonds and hydrophobic interactions between CA compound and the hVDAC protein, as well as α-helix part of hVDAC protein. Besides, the two-dimensional (2D) and 3D structures of the complexes were analyzed using discovery studio visualizer 2019.

3 Result

3.1 Differential expression analysis

We identified the DE genes between two groups (normal vs. cancer). The R package DESeq2 was used to find the DE genes. The volcano plot was used to find the DE genes with the cutoffs p-value<0.05 and |Log2FoldChange| >2 (Fig. 1). A total of 1555
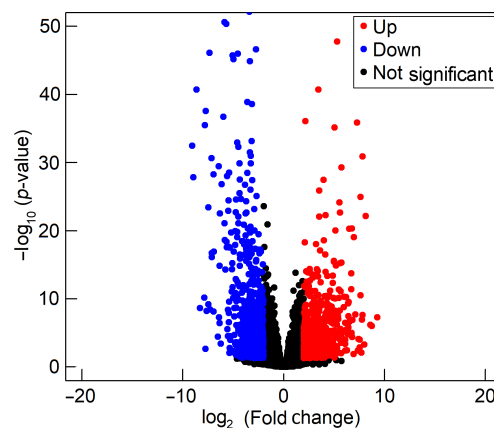

Fig. 1 Volcano plot of the genes. Red colors indicate upregulated genes while blue colors indicate downregulated genes and black colors indicate not significant. |Log2FoldChange|>2 and adjusted p-value <0.05 were considered as statistically significant.

DE genes were identified with 870 upregulated and 685 downregulated genes. The expression patterns of these genes were shown in the heatmap (Fig. 2). The dendrograms showed that the cancer and normal samples were clearly distinguishable. Then, we sorted the DE genes based on adjusted p-value and selected top 400 DE genes (200 upregulated and 200 downregulated) for PPI network analysis.

3.2 PPI networks of DE genes and identification of hub genes

The top 400 DE genes (sorted based on adjusted p-value) were mapped to the STRING database and constructed a PPI network. There were 280 nodes and 1166 edges in the PPI network with an average node degree 8.23. From the PPI network, we constructed a sub-network with the top 15 hub genes. The sub-network was shown in Fig. 3. The 15 hub genes were TIMP1, BGN, SPP1, CFTR, ACAN, CDH2, COL6A3, MMP1, COL4A1, ITGA2, COL2A1, THY1, COL5A2, MMP7, and CDX2. The details of these 15 hub genes were given in Table 1.

3.3 GO and KEGG pathway analysis

To explore the functions of the DE genes, we performed the GO term and KEGG pathway analyses for the top 15 hub genes obtained from the PPI network. The GO term biological process analysis showed that some of the genes were enriched in extracellular matrix organization, collagen catabolic process, cell adhesion, collagen fibril organization, skeletal system development, blood vessel morphogenesis, focal adhesion assembly, etc. The GO term molecular function analysis showed that some host

Fig. 2   Heatmap of the DE genes. The heatmap shows the expression profile of the DE genes in GC compared to the adjacent normal tissues. The color scale indicates the Log2FoldChange of the expression value for each gene in tumor vs. normal tissues. Red colors indicate down-regulation and green colors indicate up-regulation.

Fig. 3   A sub-network with the top 15 hub genes and their interacted genes constructed with the Cytohuba plugin of Cytoscape from the PPI network obtained from STRING database after mapping the top 400 differentially expressed genes (200 upregulated and 200 downregulated). The deepness of the red color indicates higher degree nodes.

Table 1   Top 15 hub genes extracted from the PPI network constructed with 400 top DE genes.

Gene symbol	Degree in PPI network	Log2FoldChange	Adjusted p-value	Expression (up/down)
TIMP	40	3.39	8.64×10^{-15}	Up
BGN	40	3.60	1.41×10^{-6}	Up
SPP1	38	4.89	4.65×10^{-13}	Up
CFTR	36	4.88	3.94×10^{-14}	Up
ACAN	34	3.69	9.38×10^{-8}	Up
CDH2	34	−4.34	1.42×10^{-15}	Down
COL6A3	34	2.83	8.64×10^{-7}	Up
MMP1	32	3.79	8.43×10^{-10}	Up
COL4A1	32	2.39	3.35×10^{-7}	Up
ITGA2	32	2.71	1.81×10^{-14}	Up
COL2A1	30	−5.41	2.42×10^{-18}	Down
THY1	28	3.45	2.73×10^{-11}	Up
COL5A2	28	2.01	2.45×10^{-6}	Up
MMP7	26	4.48	9.30×10^{-10}	Up

genes were enriched in extracellular matrix structural constituent and were able to bind platelet-derived growth factor binding, extracellular matrix binding, etc. The GO term cellular component analysis showed that some genes were enriched in extracellular region, collagen trimer, basement membrane, extracellular exosome, cell surface, extracellular space, apical plasma membrane, etc. The KEGG pathway analysis indicated that some host genes were enriched in several significant pathways including ECM-receptor interaction, PI3K-Akt

signaling pathway, focal adhesion, etc. The results of the significant GO term and KEGG pathway analyses were provided in Fig. 4 and Tables 2 and 3.

3.4   Expression and survival analyses of the hub genes in gastric cancer

The expressions of the 15 hub genes in the normal tissue samples and gastric tissue samples were examined by the GEPIA database. There were 408 gastric tumor samples and 36 normal samples in GEPIA obtained from
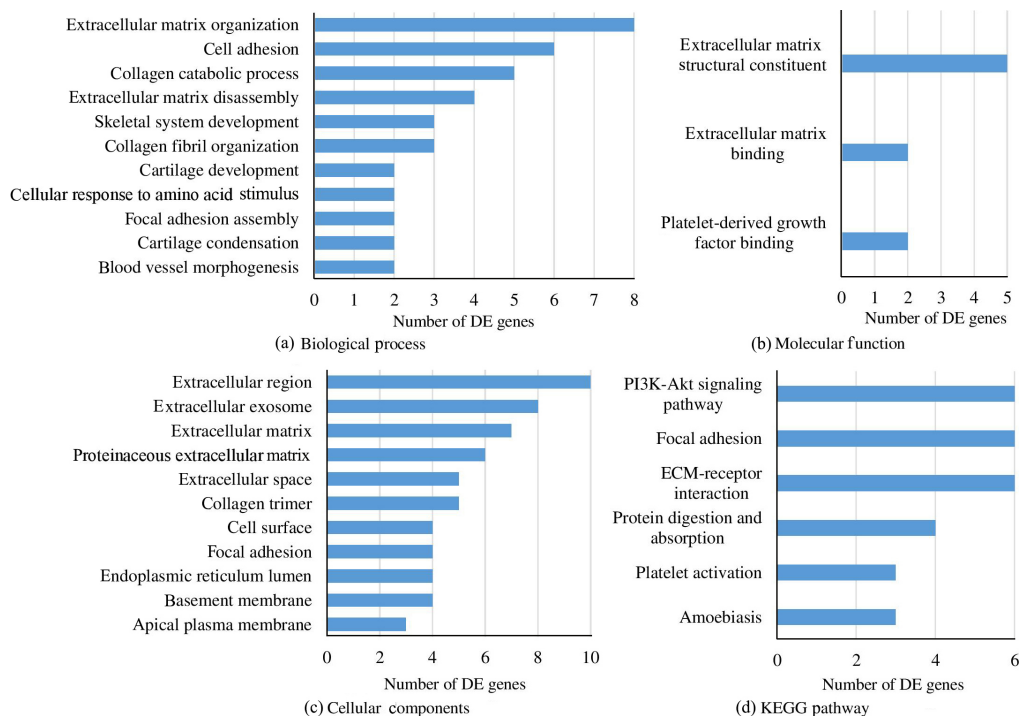

Fig. 4   GO terms and KEGG pathways for the DE genes in GC. The significant GO terms (a) biological processes, (b) molecular functions, (c) cellular components, and the significant (d) KEGG pathways for the 15 hub genes.   The p-value<0.05 was considered as statistically significant for the GO terms and KEGG pathways.

Table 2 Significant GO terms for the top 15 hub genes from the PPI networks of the top 400 DE genes.

GO type	GO term	GO name	Number of genes	*p*-value
Biological process	GO:0030198	Extracellular matrix organization	8	8.49×10^{-11}
	GO:0030574	Collagen catabolic process	5	1.87×10^{-7}
	GO:0007155	Cell adhesion	6	2.44×10^{-5}
	GO:0022617	Extracellular matrix disassembly	4	3.13×10^{-5}
	GO:0030199	Collagen fibril organization	3	4.70×10^{-4}
	GO:0001501	Skeletal system development	3	0.0056
	GO:0048514	Blood vessel morphogenesis	2	0.0141
	GO:0001502	Cartilage condensation	2	0.0182
	GO:0048041	Focal adhesion assembly	2	0.0198
	GO:0071230	Cellular response to amino acid stimulus	2	0.0385
	GO:0051216	Cartilage development	2	0.0481
Molecular function	GO:0005201	Extracellular matrix structural constituent	5	2.20×10^{-7}
	GO:0048407	Platelet-derived growth factor binding	2	0.0091
	GO:0050840	Extracellular matrix binding	2	0.0214
Cellular component	GO:0031012	Extracellular matrix	7	4.70×10^{-8}
	GO:0005576	Extracellular region	10	4.26×10^{-7}
	GO:0005581	Collagen trimer	5	5.86×10^{-7}
	GO:0005578	Proteinaceous extracellular matrix	6	1.19×10^{-6}
	GO:0005604	Basement membrane	4	2.76×10^{-5}
	GO:0005788	Endoplasmic reticulum lumen	4	3.85×10^{-4}
	GO:0070062	Extracellular exosome	8	0.0026
	GO:0005925	Focal adhesion	4	0.0030
	GO:0009986	Cell surface	4	0.0075
	GO:0005615	Extracellular space	5	0.0163
	GO:0016324	Apical plasma membrane	3	0.0204

Table 3 Significant pathways for the top 15 hub genes from the PPI network of the top 400 DE genes.

Pathway ID	Pathway description	Number of genes	*p*-value
hsa04512	ECM-receptor interaction	6	1.25×10^{-7}
hsa04510	Focal adhesion	6	9.15×10^{-6}
hsa04151	PI3K-Akt signaling pathway	6	1.11×10^{-4}
hsa04974	Protein digestion and absorption	4	3.10×10^{-4}
hsa05146	Amoebiasis	3	0.0118
hsa04611	Platelet activation	3	0.0174

TCGA database. Among the 15 hub genes, 13 genes, TIMP1, BGN, SPP1, CFTR, ACAN, COL6A3, MMP1, COL4A1, ITGA2, THY1, COL5A2, MMP7, and CDX2, were upregulated, and 2 genes, CDH2 and COL2A1, were downregulated. We found that 13 genes were upregulated in the GC tissue compared to the normal tissue (Figs. 5a–5e, 5g–5j, and 5l–5o). And 2 genes were downregulated in GC compared to the normal samples (Figs. 5f and 5k). Thus, our findings were validated by the TCGA data.

The correlation analysis between the expression of the hub genes in GC and the overall survival of the GC patients was performed by GEPIA tool using TCGA data. The survival data of 381 patients with GC were given in the GEPIA database. The patients were classified into higher and lower groups based on the median values of the gene expression. As shown in Fig. 6, the lower expression of the genes CDH2, COL4A1, and COL5A2 was associated with better survival of the GC patients.

3.5 Association of KGs with other disease risks

The disease-KGs interaction analysis revealed that top-ranked 10 diseases (axenfeld anomaly (disorder), cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy, tumour budding, retinal hemorrhage, schizencephaly, dilatation of the cerebral artery, congenital porencephaly, cardiomyopathies, diabetic nephropathy, and autism
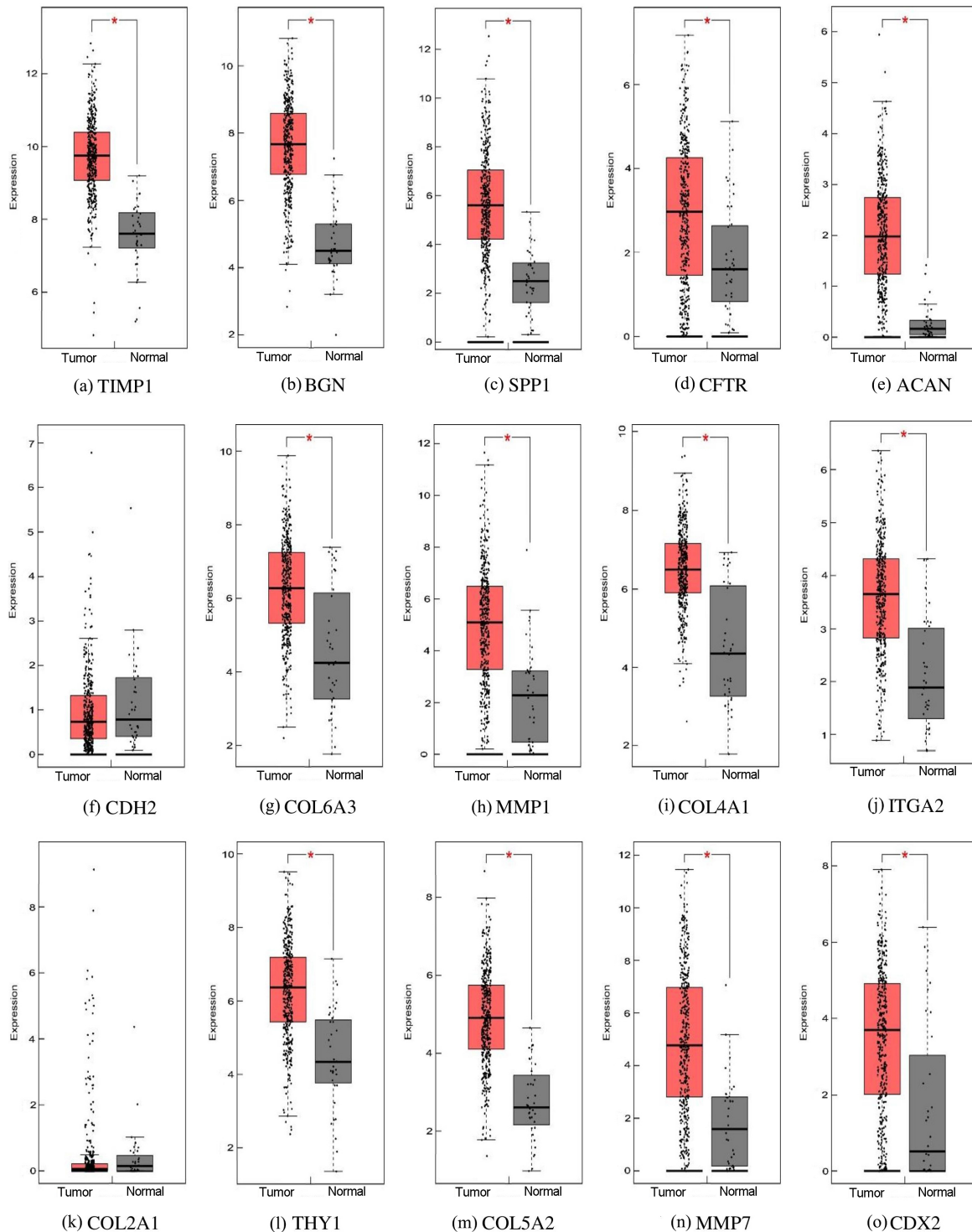
Fig. 5 Validation of the expression of 15 hub genes in GC. The expression of (a) TIMP1, (b) BGN, (c) SPP1, (d) CFTR, (e) ACAN, (g) COL6A3, (h) MMP1, (i) COL4A1, (j) ITGA2, (l) THY1, (m) COL5A2, (n) MMP7, and (o) CDX2 were upregulated, and the expression of (f) CDH2 and (k) COL2A1 were downregulated in gastric tumor tissues compared to normal tissues from the TCGA data through GEPIA. The cutoffs |Log2FoldChange| >1 and *p*-value<0.01 were considered as statistically significant.

spectrum disorders) are the highly significant risk factors for GC patients due to the influence of 3 KGs (Table 4) and the details were provided in Table S2 in the ESM.

3.6 Drug repurposing by molecular docking study

To explore candidate drugs by molecular docking simulation, we considered $m = 3$ drug target proteins

(a) CDH2 overall survival (b) COL4A1 overall survival (c) COL5A2 overall survival
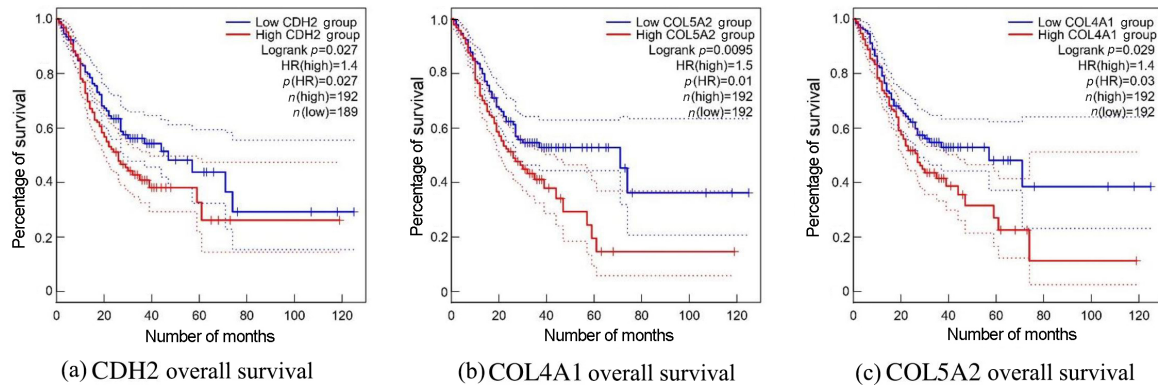
Fig. 6 Overall survival analysis of the candidate hub genes. The association between the expression levels of (a) CDH2, (b) COL4A1, and (c) COL5A2 and the overall survival of the patients with GC was obtained from the TCGA data through GEPIA. The threshold log rank *P*<0.05 was considered as statistically significant.

Table 4 Top-ranked 10 diseases for 3 KGs of the DE genes for GC.

Term	*p*-value	Combined score	Gene
Axenfeld anomaly (disorder)	9.00×10^{-4}	14 021.17	COL4A1
Cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy	9.00×10^{-4}	14 021.17	COL4A1
Tumour budding	9.00×10^{-4}	14 021.17	COL4A1
Retinal hemorrhage	0.001 049 661	11 427.00	COL4A1
Schizencephaly	0.001 049 661	11 427.00	COL4A1
Dilatation of the cerebral artery	0.001 349 432	8255.546	COL4A1
Congenital porencephaly	0.001 349 432	8255.546	COL4A1
Cardiomyopathies	0.001 696 973	520.962	CDH2; COL4A1
Diabetic nephropathy	0.002 312 366	422.106	CDH2; COL4A1
Autism spectrum disorders	0.002 403 121	411.1017	CDH2; COL4A1

(receptors) and $n = 102$ meta-drug agents as mentioned in the data source. We downloaded 3D structure of our COL4A1 from protein data bank (PDB)[39] with source codes 1li1. On the other hand, the 3D structures of CDH2 and COL5A2 proteins were downloaded from UniProt[49] with sources ID of P19022 and P05997, respectively. The 3D structures of 102 meta-drug agents (Table S1 in the ESM) were downloaded from PubChem database[41] as mentioned previously. Then, the molecular docking was carried out between total 3 proteins and 102 meta-drug agents to calculate the binding affinity scores (kcal/mol) for each pair of proteins and drugs. Next, we ordered the proteins in descending order of row sums of the binding affinity matrix and drug agents according to the column sums to select few drug agents as the candidate drugs for GC (Fig. 7 and Table S3 in the ESM). Thus, we selected top-ranked six drug agents (Everolimus, Docetaxel, Lanreotide, Venetoclax, Temsirolimus, and Nilotinib) as candidate drugs with the binding affinity scores –8.2 kcal/mol or less against the 3 proteins (Table S3 in the ESM).

The docked complexes of the top three virtual hits from AutoDock Vina docking were further considered for protein-ligand interaction profiling. As shown in Table 5 and Fig. 8a, COL4A1_Everolimus complex showed three hydrogen bonds with trp192, thr206, and ser208 residues, where trp192 and thr206 were involved in conventional dydrogen bonding with the drug and ser208 was also involved in carbon hydrogen bonding with the drug. The drug formed major hydrophobic interactions with leu210, leu215, met145, and pro205 alkyl bonding while TYR189, PHE191, and TRP192 were responsible for hydrophobic interactions via pi-alkyl bonding. On the other hand, CDH2_Docetaxel (Fig. 8b) formed one hydrogen bonds with ala521 residues while ile532 was responsible for hydrophobic interactions via alkyl bonding. Furthermore, docetaxel also formed electrostatic with asp523 residue. In Fig. 8c, CO5A2_Lanreotide complex showed three hydrogen bonds with gln77, pro94, and glu75 residues and the major hydrophobic interactions with val79, cys93, val79, and val95 residues.
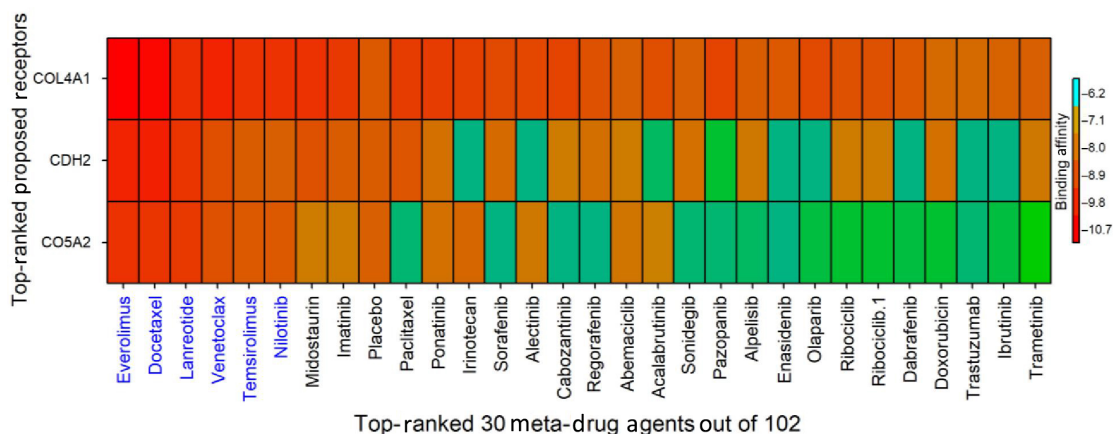
Fig. 7   Image of binding affinity scores based on the top-ordered 30 meta-drug agents out of 102 against the ordered 3 receptors, where red colors indicate the strong binding affinities between target proteins and drug agents, and green colors indicate their weak bindings.

Table 5   Detailed molecular interactions analysis for best hits obtained from Autodock Vina docking.

Complex name	Conventional hydrogen bond	Carbon hydrogen bond	Hydrophobic (alkyl)	Hydrophobic (pi-alkyl)	Electrostatic
COL4A1_Everolimus	TRP192, THR206	SER208	LEU210, LEU215, MET145, PRO205	TYR189, PHE191, TRP192	−
CDH2_Docetaxel	ALA521	−	ILE532	−	ASP523
CO5A2_Lanreotide	GLN77, PRO94, GLU75	−	VAL79, CYS93, VAL79, VAL95	−	−

4   Discussion

Gastric cancer is one of the most commonly diagnosed tumors all over the world. Although, a lot of advancements have been made in GC diagnosis and treatment, the survival rate is still poor due to lack of early diagnosis and proper treatments at advanced level of GC. In this study, we investigated some potential candidate genes for GC using the paired cancer-normal sequencing data obtained from the GEO datasets. Primarily, we identified 15 hub genes TIMP1, BGN, SPP1, CFTR, ACAN, CDH2, COL6A3, MMP1, COL4A1, ITGA2, COL2A1, THY1, COL5A2, MMP7, and CDX2 from the PPI network of top 400 DE genes. Several of these genes were reported as important biomarkers of GC by previous studies. For example, overexpression of TIMP1 promoted GC cell proliferation[50]. BGN was identified as a candidate biomarker for GC prognosis and tumor immune infiltration[51]. Overexpression of SPP1 was closely correlated with GC occurrence[52]. THY1 was identified as a potential novel biomarker for GC[53]. ITGA2 was reported as a potential therapeutic target for GC[54]. MMP7 was identified as a prognostic biomarker to predict the outcome of GC patients[55].

We performed the GO term and KEGG pathway analyses for the top 15 hub genes. The GO term analyses showed that some hub genes were enriched in several important biological processes, molecular functions, and cellular components (Table 2). Previous studies also reported that the biological processes extracellular matrix organization, collagen catabolic process, cell adhesion, extracellular matrix disassembly, collagen fibril organization and skeletal system development, the molecular functions extracellular matrix structural constituent and extracellular matrix binding, and the cellular components extracellular region, collagen trimer, extracellular exosome, cell surface, and basement membrane were associated with GC related genes[56–59]. The KEGG pathway analysis showed that some hub genes were enriched in GC related signaling pathways including ECM-receptor interaction, PI3K-Akt signaling pathway, focal adhesion, etc. ECM-receptor interaction was reported as an important pathway linked with the progression of GC[60]. Focal adhesion was identified as a key player in regulating cell survival and proliferation, migration, and invasion of GC cells[61]. The PI3K-Akt signaling pathway was known to play an important role in the development and progression of GC[62].
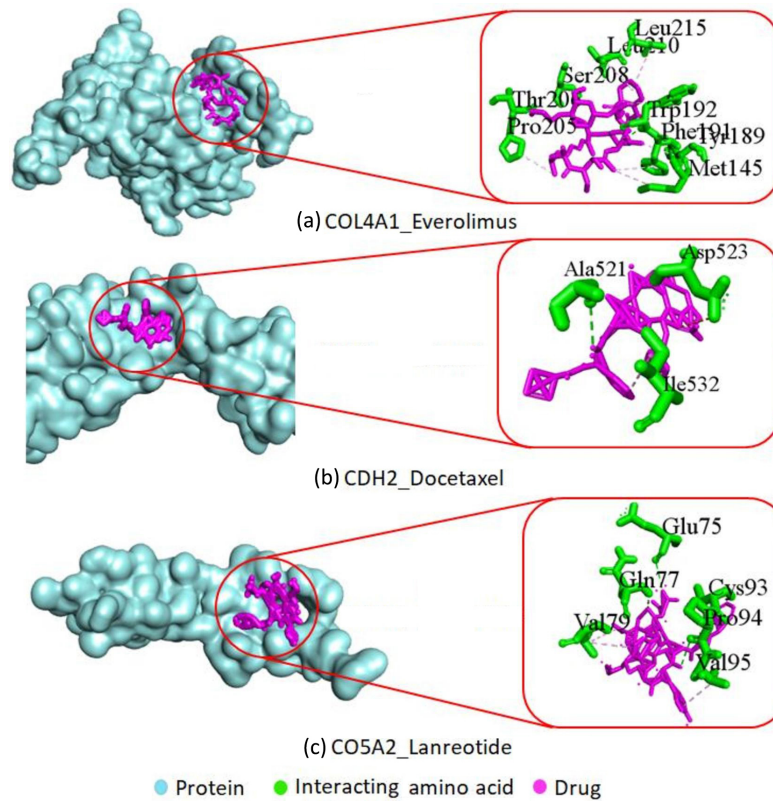
Fig. 8   Top three potential targets and top three lead drugs based on docking results. Lead three drugs Everolimus, Docetaxel, and Lanreotide were selected by investigating the binding affinity score. The 3D structure of the key protein with candidate drugs is shown in left side. The 2D Schematic diagram of key protein with candidate drugs interaction and the neighbor residues (within 4×10^{-10} m of the drug) are given in right side.

The survival analysis showed that the lower expression of the genes CDH2, COL4A1, and COL5A2 was associated with better survival of the GC patients (Fig. 6). CDH2 can be used as a potential biomarker for the prognosis of GC[63]. COL4A1 might be able to confer trastuzumab resistance in GC[64]. COL5A2 is strongly associated with the prognosis of GC[65]. Thus, these three genes CDH2, COL4A1, and COL5A2 might serve as potential biomarkers for GC.

To explore our proposed genomic biomarkers guided repurposable drugs, we proposed 3 key proteins (CDH2, COL4A1, and COL5A2) as the drug target receptors and performed their docking simulation with the 102 meta-drug agents. Then, we selected top ranked six drugs (Everolimus, Docetaxel, Lanreotide, Venetoclax, Temsirolimus, and Nilotinib) as the most probable repurposable candidate drugs for GC infections based on their strong binding affinity scores (kcal/mol) with all the target proteins (Fig. 7). Among the identified candidate drugs, Everolimus was an mTOR inhibitor with antitumor activity. Everolimus was recommended as a potential drug against GC[32, 66, 67]. In a phase I

clinical trial, Everolimus was used in combination with capecitabine in patients with refractory GC where the clinical benefits were modest[68]. In phase I clinical trials NCT01049620 and NCT01042782, Everolimus was used in combination with capecitabine and oxaliplatin, and mitomycin C, respectively, in patients with advanced GC but the results of these trials were unknown[69]. In a multicenter phase II study (NCT00519324), Everolimus was used in patients with metastatic GC with previous chemotherapy failure[70]. Docetaxel was a standard chemotherapy regimen for GC patients[30, 66, 67, 71]. Perioperative chemotherapy with docetaxel, oxaliplatin, fluorouracil, and leucovorin significantly improved the progression-free survival and overall survival among patients with resectable GC compared with epirubicin, cisplatin, and fluorouracil or capecitabine (ECF/ECX)[72]. Patients with advanced GC benefited more with the combination of apatinib and docetaxel than with apatinib monotherapy[72]. Lanreotide was approved in the USA and Europe to improve progression-free survival (PFS) in patients with unresectable gastroenteropancreatic

neuroendocrine tumors (GEP-NETs)[73]. Lanreotide acetate inhibitor was considered as a potential therapeutic against GC patients[32, 67]. Venetoclax was used as a potential treatment against GC patients[32]. Venetoclax was a potent, clinically approved BH3-mimetic that induced apoptosis by targeting BCL-2[74]. Several authors recommended the mTOR inhibitor temsirolimus as a potentially useful drug by screening drugs that reduced the proliferation of diffuse-type GC-initiating cells[32, 75]. Several articles reported that mTOR played an important role in the growth regulation of GCs[76–78]. Temsirolimus was used for the treatment of patients with renal cell carcinoma, multiple myeloma, and mantle cell lymphoma[75]. Nilotinib was considered a potential treatment for GC patients[32, 79, 80]. Two cases of GC were reported with nilotinib therapy in a post-marketing clinical use survey in Japan and one case was reported in a global phase III multicenter trial[80].

The literature review also supported our proposed drugs for the treatment against GC infections. Therefore, the proposed six candidate drugs might play the vital role for the treatment against GC patients with comorbidities as our proposed target proteins are also associated with several comorbidities. Further wet lab experimental validation is needed for both the proposed target proteins and candidate drugs to confirm the role of the candidate drugs for the treatment of GC.

5   Conclusion

We identified a total of 1555 DE genes with 870 upregulated and 685 downregulated genes. Through the PPI network analysis, we screened 15 hub genes. The GO term analysis showed that some hub genes were enriched in several significant biological processes, molecular mechanisms, and cellular components. Again, the KEGG pathway analysis showed that some hub genes were associated with several GC related pathways. The expressions of the hub genes were validated by the TCGA data. Among the 15 hub genes, three genes, CDH2, COL4A1, and COL5A2, were significantly associated with the GC patients found by survival analysis. Therefore, these three genes, CDH2, COL4A1, and COL5A2, might be considered as potential biomarkers for GC diagnosis and treatment. In this paper, we also attempted to suggest effective supporting drugs for the treatment against GC patients. For this purpose, we identified 3 host receptor proteins (CDH2, COL4A1, and COL5A2) guided top ranked 6 repurposable

drugs (Everolimus, Docetaxel, Lanreotide, Venetoclax, Temsirolimus, and Nilotinib) for the treatment against GC patients by molecular docking simulation. The literature review also supported our proposed drugs for the treatment against GC patients. Therefore, our findings might be effective therapeutic resource for the better treatment against GC patients.

Electronic Supplementary Material

Supplementary materials including
- Table S1: 102 meta-drug agents for the treatment against GC.
- Table S2: Significant diseases name for 3 KGs of the DE genes for GC.
- Table S3: Docking score of target proteins with meta-drug agent against GC.

All the tables are available in the online version of this article at http:// doi.org/10.26599/TST.2022.9010035.

References

[1]   B. Freddie, F. Jacques, S. Isabelle, R. L. Siegel, L. A. Torre, and J. Ahmedin, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA. Cancer J. Clin.*, vol. 68, no. 6, pp. 394–424, 2018.

[2]   M. Rugge, R. M. Genta, F. D. Mario, E. M. El-Omar, H. B. El-Serag, M. Fassan, R. H. Hunt, E. J. Kuipers, P. Malfertheiner, K. Sugano, et al., Gastric cancer as preventable disease, *Clin. Gastroenterol. Hepatol.*, doi: 10.1016/j.cgh.2017.05.023.

[3]   Y. Shi, L. Qi, H. Chen, J. Zhang, Q. Guan, J. He, M. Li, Z. Guo, H. Yan, and P. Li, Identification of genes universally

differentially expressed in gastric cancer, *Biomed Res. Int.*, doi: 10.1155/2021/7326853.

[4] E. C. Smyth, M. Nilsson, H. I. Grabsch, N. C. V. Grieken, and F. Lordick, Gastric cancer, *The Lancet*, doi: 10.1016/S0140-6736(20)31288-5.

[5] J. Machlowska, J. Baj, M. Sitarz, R. Maciejewski, and R. Sitarz, Gastric cancer: Epidemiology, risk factors, classification, genomic characteristics and treatment strategies, *International Journal of Molecular Sciences*, doi: 10.3390/ijms21114012.

[6] A. R. Yusefi, K. B. Lankarani, P. Bastani, M. Radinmanesh, and Z. Kavosi, Risk factors for gastric cancer: A systematic review, *Asian Pacific Journal of Cancer Prevention*, doi: 10.22034/APJCP.2018.19.3.591.

[7] G. Herrera-Almario and V. E. Strong, Minimally invasive gastric surgery, *Ann. Surg. Oncol.*, doi: 10.1245/s10434-016-5429-3.

[8] J. Tey, Y. Y. Soon, W. Y. Koh, C. N. Leong, B. Vellayappan, K. Lim, and I. W. K. Ivan, Palliative radiotherapy for gastric cancer: A systematic review and meta-analysis, *Oncotarget*, doi: 10.18632/oncotarget.15554.

[9] O. Zhandossov, G. Kaussova, and A. Koten, Combined treatment for gastric cancer: Immunological approach, *Turkish J. Gastroenterol.*, doi: 10.5152/tjg.2018.17398.

[10] E. Niccolai, A. Taddei, D. Prisco, and A. Amedei, Gastric cancer and the epoch of immunotherapy approaches, *World J. Gastroenterol.*, doi: 10.3748/wjg.v21.i19.5778.

[11] L. Necula, L. Matei, D. Dragu, A. I. Neagu, C. Mambet, S. Nedeianu, C. Bleotu, C. C. Diaconu, and M. Chivu-Economescu, Recent advances in gastric cancer early diagnosis, *World J. Gastroenterol.*, doi: 10.3748/wjg.v25.i17.2029.

[12] A. S. Crystal, A. T. Shaw, L. V. Sequist, L. Friboulet, M. J. Niederst, E. L. Lockerman, R. L. Frias, J. F. Gainor, A. Amzallag, P. Greninger, et al., Patient-derived models of acquired resistance can identify effective drug combinations for cancer, *Science*, doi: 10.1126/science.1254721.

[13] A. A. Borisy, P. J. Elliott, N. W. Hurst, M. S. Lee, J. Lehar, E. R. Price, G. Serbedzija, G. R. Zimmermann, M. A. Foley, B. R. Stockwell, et al., Systematic discovery of multicomponent therapeutics, *Proc. Natl. Acad. Sci.*, doi: 10.1073/pnas.1337088100.

[14] B. Klinger, A. Sieber, R. Fritsche-Guenther, F. Witzel, L. Berry, D. Schumacher, Y. Yan, P. Durek, M. Merchant, R. Schafer, et al., Network quantification of EGFR signaling unveils potential for targeted combination therapy, *Mol. Syst. Biol.*, doi: 10.1038/msb.2013.29.

[15] M. Bansal, J. Yang, C. Karan, M. P. Menden, J. C. Costello, H. Tang, G. Xiao, Y. Li, J. Allen, R. Zhong, et al., A community computational challenge to predict the activity of pairs of compounds, *Nat. Biotechnol.*, doi: 10.1038/nbt.3052.

[16] L. A. M. Griner, R. Guha, P. Shinn, R. M. Young, J. M. Keller, D. Liu, I. S. Goldlust, A. Yasgar, C. MaKnight, M. B. Boxer, et al., High-throughput combinatorial screening identifies drugs that cooperate with ibrutinib to kill activated B-cell-like diffuse large B-cell lymphoma cells, *Proc. Natl. Acad. Sci.*, doi: 10.1073/pnas.1311846111.

[17] M. L. Miller, E. J. Molinelli, J. S. Nair, T. Sheikh, R. Samy, X. Jing, Q. He, A. Korkut, A. M. Crago, S. Singer, et al., Drug synergy screen and network modeling in dedifferentiated liposarcoma identifies CDK4 and IGF1R as synergistic drug targets, *Sci. Signal.*, doi: 10.1126/scisignal.2004014.

[18] Y. Tu, Q. Wu, J. He, J. Xu, S. Yu, Q. Wang, Y. Cheng, Q. Yang, S. Xu, and Y. Cao, Exploring the potential molecular mechanism of scutellaria baicalensis georgi in the treatment of gastric cancer based on network pharmacological analysis and molecular docking technology, *Front. Pharmacol.*, doi: 10.3389/fphar.2021.697704.

[19] L. Guo, H. Shi, and L. Zhu, Siteng fang reverses multidrug resistance in gastric cancer: A network pharmacology and molecular docking study, *Front. Oncol.*, doi: 10.3389/fonc.2021.671382.

[20] O. A. Kolawole, F. Olatomide A., and S. Banjo, Anti-gastric cancer activity of 1,2,3-triazolo[4,5-d]pyrimidine hybrids (1,2,3-TPH): QSAR and molecular docking approaches, *Heliyon*, doi: 10.1016/j.heliyon.2020.e03561.

[21] K. Li, Y. Du, L. Li, and D. -Q. Wei, Bioinformatics approaches for anti-cancer drug discovery, *Curr. Drug Targets*, vol. 21, no. 1, pp. 3–17, 2019.

[22] A. C. Kaushik, Y. J. Wang, X. Wang, and D. -Q. Wei, Irinotecan and vandetanib create synergies for treatment of pancreatic cancer patients with concomitant TP53 and KRAS mutations, *Brief. Bioinform.*, doi: 10.1093/bib/bbaa149.

[23] M. Ö. Cingiz, G. Biricik, and B. Diri, The performance comparison of gene co-expression networks of breast and prostate cancer using different selection criteria, *Interdiscip. Sci.: Comput. Life Sci.*, vol. 13, no. 3, pp. 500–510, 2021.

[24] D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, P. Bork, et al., STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, *Nucleic Acids Res.*, doi: 10.1093/nar/gky1131.

[25] R. K. Patel and M. Jain, NGS QC toolkit: A toolkit for quality control of next generation sequencing data, *PLoS One*, doi: 10.1371/journal.pone.0030619.

[26] H. Li and R. Durbin, Fast and accurate short read alignment with burrows-wheeler transform, *Bioinformatics*, doi: 10.1093/bioinformatics/btp324.

[27] Y. Liao, G. K. Smyth, and W. Shi, Featurecounts: An efficient general purpose program for assigning sequence reads to genomic features, *Bioinformatics*, doi: 10.1093/bioinformatics/btt656.

[28] Z. Li, P. Hou, D. Fan, M. Dong, M. Ma, H. Li, R. Yao, B.

Huang, G. Wang, P. Geng, et al., The degradation of EZH2 mediated by lncRNA ANCR attenuated the invasion and metastasis of breast cancer, *Cell Death Differ.*, vol. 24, no. 1, pp. 59–71, 2017.

[29] E. D. Tetzlaff, P. Cen, and J. A. Ajani, Emerging drugs in the treatment of advanced gastric cancer, *Expert Opinion on Emerging Drugs*, doi: 10.1517/14728214.13.1.135.

[30] L. Vecchione, M. Orditura, F. Ciardiello, and F. D. Vita, Novel investigational drugs for gastric cancer, *Expert Opinion on Investigational Drugs*, doi: 10.1517/13543780902969455.

[31] R. Geng, L. Song, J. Li, and L. Zhao, The safety of apatinib for the treatment of gastric cancer, *Expert Opin. Drug Saf.*, doi: 10.1080/14740338.2018.1535592.

[32] D. A. Hescheler, P. S. Plum, T. Zander, A. Quaas, M. Korenkov, A. Gassa, M. Michel, C. J. Bruns, and H. Alakus, Identification of targeted therapy options for gastric adenocarcinoma by comprehensive analysis of genomic data, *Gastric Cancer*, doi: 10.1007/s10120-020-01045-9.

[33] M. I. Love, W. Huber, and S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.*, doi: 10.1186/s13059-014-0550-8.

[34] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker, Cytoscape: A software environment for integrated models of biomolecular interaction networks, *Genome Res.*, doi: 10.1101/gr.1239303.

[35] D. W. Huang, B. T. Sherman, and R. A. Lempicki, Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists, *Nucleic Acids Res.*, doi: 10.1093/nar/gkn923.

[36] Z. Tang, B. Kang, C. Li, T. Chen, and Z. Zhang, GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis, *Nucleic Acids Res.*, doi: 10.1093/nar/gkz430.

[37] M. V. Kuleshov, M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. L. Jenkins, K. M. Jagodnik, A. Lachmann, et al., Enrichr: A comprehensive gene set enrichment analysis web server 2016 update, *Nucleic Acids Res.*, doi: 10.1093/nar/gkw377.

[38] J. Pinero, J. M. Ramírez-Anguita, J. Saüch-Pitarch, F. Ronzano, E. Centeno, F. Sanz, and L. I. Furlong, The DisGeNET knowledge platform for disease genomics: 2019 update, *Nucleic Acids Res.*, doi: 10.1093/nar/gkz1021.

[39] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, The protein data bank, *Nucleic Acids Research*, doi: 10.1093/nar/28.1.235.

[40] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. D. Beer, C. Rempfer, L. Bordoli, et al., SWISS-MODEL: Homology modelling of protein structures and complexes, *Nucleic Acids Res.*, doi: 10.1093/nar/gky427.

[41] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, et al., PubChem 2019 update: Improved access to chemical data, *Nucleic Acids Res.*, doi: 10.1093/nar/gky1033.

[42] D. S. v4. 0. 100. 13345 Visualizer, Accelrys Software Inc, https://discover.3ds.com/discovery-studio-visualizer-download, 2005.

[43] T. J. Dolinsky, P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe, and N. A. Baker, PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations, *Nucleic Acids Res.*, doi: 10.1093/nar/gkm276.

[44] J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath, and A. Onufriev, H++: A server for estimating pKas and adding missing hydrogens to macromolecules, *Nucleic Acids Res.*, doi: 10.1093/nar/gki464.

[45] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, and A. J. Olson, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, *J. Comput. Chem.*, vol. 30, no. 16, pp. 2785–2791, 2009.

[46] O. Trott and A. J. Olson, AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.*, vol. 31, no. 2, pp. 455–461, 2010.

[47] W. L. Delano and S. Bromberg, PyMOL user's guide, https://pymol.sourceforge.net/newman/userman.pdf, 2004.

[48] W. L. DeLano, The PyMOL molecular graphics system, version 2.3, Schrödinger LLC, https://sourceforge.net/projects/pymol/, 2020.

[49] D506-D515, UniProt: A worldwide hub of protein knowledge, The UniProt Consortium, *Nucleic Acids Res.*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6323992/, 2019.

[50] O. M. Omar, M. Soutto, N. S. Bhat, A. A. Bhat, H. Lu, Z. Chen, and W. El-Rifai, TFF1 antagonizes TIMP-1 mediated proliferative functions in gastric cancer, *Mol. Carcinog.*, vol. 57, no. 11, pp. 1577–1587, 2018.

[51] W. Chen and Z. Yang, Identification of differentially expressed genes reveals BGN predicting overall survival and tumor immune infiltration of gastric cancer, *Comput. Math. Methods Med.*, doi: 10.1155/2021/5494840.

[52] T. Chen, L. Sun, C. He, Y. Gong, Q. Xu, and Y. Yuan, Serum OPN expression for identification of gastric cancer and atrophic gastritis and its influencing factors, *PLoS One*, doi: 10.1371/journal.pone.0114005.

[53] D. Oikonomou, K. Hassan, J. T. Kaifi, H. C. Fiegel, P. G. Schurr, U. Reichelt, K. Aridome, E. F. Yekebas, O. Mann, D. Kluth, et al., Thy-1 as a potential novel diagnostic marker for gastrointestinal stromal tumors, *J. Cancer Res. Clin. Oncol.*, vol. 133, no. 12, pp. 951–955, 2007.

[54] Y. C. Chuang, H. -Y. Wu, Y. -L. Lin, S. -C. Tzou, C. -H. Chuang, T. -Y. Jian, P. -R. Chen, Y. -C. Chang, C. -H. Lin, T. -H. Huang, et al., Blockade of ITGA2 induces apoptosis and inhibits cell migration in gastric cancer, *Biol. Proced. Online*, doi: 10.1186/s12575-018-0073-x.

[55] S. Soleyman-Jahi, S. Nedjat, A. Abdirad, N. Hoorshad, R. Heidari, and K. Zendehdel, Prognostic significance of matrix metalloproteinase-7 in gastric cancer survival: A meta-analysis, *PLoS One*, doi: 10.1371/journal.pone.0122316.

[56] J. Qiu, M. Sun, Y. Wang, and B. Chen, Identification of hub genes and pathways in gastric adenocarcinoma based on bioinformatics analysis, *Med. Sci. Monit.*, doi: 10.12659/MSM.920261.

[57] B. Jiang, S. Li, Z. Jiang, and P. Shao, Gastric cancer associated genes identified by an integrative analysis of gene expression data, *Biomed Res. Int.*, doi: 10.1155/2017/7259097.

[58] X. Chong, R. Peng, Y. Sun, L. Zhang, and Z. Zhang, Identification of key genes in gastric cancer by bioinformatics analysis, *Biomed Res. Int.*, doi: 10.1155/2020/7658230.

[59] W. Wang, Y. He, Q. Zhao, X. Zhao, and Z. Li, Identification of potential key genes in gastric cancer using bioinformatics analysis, *Biomed. Reports*, vol. 12, no. 4, pp. 178–192, 2020.

[60] K. Hu and F. Chen, Identification of significant pathways in gastric cancer based on protein-protein interaction networks and cluster analysis, *Genet. Mol. Biol.*, vol. 35, no. 3, pp. 701–708, 2012.

[61] D. Mao, R. Xu, H. Chen, X. Chen, D. Li, S. Song, Y. He, Z. Wei, and C. Zhang, Cross-talk of focal adhesion-related gene defines prognosis and the immune microenvironment in gastric cancer, *Front. Cell Dev. Biol.*, doi: 10.3389/fcell.2021.716461.

[62] T. Matsuoka and M. Yashiro, The role of PI3K/Akt/mTOR signaling in gastric carcinoma, *Cancers*, vol. 6, no. 3, pp. 1441–1463, 2014.

[63] S. Luo, R. Lin, X. Liao, D. Li, and Y. Qin, Identification and verification of the molecular mechanisms and prognostic values of the cadherin gene family in gastric cancer, *Sci. Rep.*, vol. 11, no. 1, p. 23674, 2021.

[64] R. Huang, W. Gu, B. Sun, and L. Gao, Identification of COL4A1 as a potential gene conferring trastuzumab resistance in gastric cancer based on bioinformatics analysis, *Mol. Med. Rep.*, vol. 17, no. 5, pp. 6387–6396, 2018.

[65] Y. L. Ding, S. F. Sun, and G. L. Zhao, COL5A2 as a potential clinical biomarker for gastric cancer and renal metastasis, *Medicine* (*Baltimore*)., vol. 100, no. 7, p. e24561, 2021.

[66] X. Liu and S. J. Meltzer, Gastric cancer in the era of precision medicine, *CMGH*, vol. 3, no. 3. pp. 348–358, 2017.

[67] G. A. Colditz, National cancer institute, *SAGE Encycl. Cancer Soc.*, doi: 10.4135/9781483345758.n393.

[68] T. Lim, J. Lee, D. J. Lee, H. Y. Lee, B. Han, K. K. Baek, H. K. Ahn, S. J. Lee, S. H. Park, J. O. Park, et al., Phase I trial of capecitabine plus everolimus (RAD001) in patients with previously treated metastatic gastric cancer, *Cancer Chemother. Pharmacol.*, vol. 68, no. 1, pp. 255–262, 2011.

[69] C. Pauligk, S. Janowski, K. Steinmetz, A. Atmaca, A. Pustowka, E. Jaeger, and S. Al-Batran, Daily RAD001 plus mitomycin C, every three weeks in previously treated patients with advanced gastric cancer or cancer of the esophagogastric junction: Preliminary results of a phase

I study, *J. Clin. Oncol.*, vol. 28, no. 15_suppl, p. e14531, 2010.

[70] T. Doi, K. Muro, N. Boku, Y. Yamada, T. Nishina, H. Takiuchi, Y. Komatsu, Y. Hamamoto, N. Ohno, and Y. Fujita, Multicenter phase II study of everolimus in patients with previously treated metastatic gastric cancer, *J. Clin. Oncol.*, vol. 28, no. 11, pp. 1904–1910, 2010.

[71] E. V. Cutsem, C. Boni, J. Tabernero, B. Massuti, G. Middleton, F. Dane, P. Reichardt, F. L. Pimentel, A. Cohn, P. Follana, et al., Docetaxel plus oxaliplatin with or without fluorouracil or capecitabine in metastatic or locally recurrent gastric cancer: A randomized phase II study, *Ann. Oncol.*, vol. 26, no. 1, pp. 149–156, 2015.

[72] H. Lin, D. Han, G. Fu, C. Liu, L. Wang, S. Han, B. Liu, and J. Yu, Concurrent apatinib and docetaxel vs apatinib monotherapy as third-or subsequent-line therapy for advanced gastric adenocarcinoma: A retrospective study, *Onco. Targets. Ther.*, vol. 12, pp. 1681–1689, 2019.

[73] E. M. Wolin, A. Manon, C. Chassaing, A. Lewis, L. Bertocchi, J. Richard, and A. T. Phan, Lanreotide depot: An antineoplastic treatment of carcinoid or neuroendocrine tumors, *Journal of Gastrointestinal Cancer*, vol. 47, no. 4. pp. 366–374, 2016.

[74] A. J. Souers, J. D. Leverson, E. R. Boghaert, S. L. Ackler, N. D. Catron, J. Chen, B. D. Dayton, H. Ding, S. H. Enschede, W. J. Fairbrother, et al., ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets, *Nat. Med.*, vol. 19, no. 2, pp. 202–208, 2013.

[75] H. Fukamachi, S. -K. Kim, J. Koh, H. S. Lee, Y. Sasaki, K. Yamashita, T. Nishikawaji, S. Shimada, Y. Akiyama, S. J. Byeon, et al., A subset of diffuse-type gastric cancer is susceptible to mTOR inhibitors and checkpoint inhibitors, *J. Exp. Clin. Cancer Res.*, doi: 10.1186/s13046-019-1121-3.

[76] S. J. Byeon, N. Han, J. Choi, M. A. Kim, and W. H. Kim, Prognostic implication of TSC1 and mTOR expression in gastric carcinoma, *J. Surg. Oncol.*, vol. 109, no. 8, pp. 812–817, 2014.

[77] S. E. Al-Batran, M. Ducreux, and A. Ohtsu, MTOR as a therapeutic target in patients with gastric cancer, *International Journal of Cancer*, vol. 130, no. 3. pp. 491–496, 2012.

[78] T. Murayama, M. Inokuchi, Y. Takagi, H. Yamada, K. Kojima, J. Kumagai, T. Kawano, and K. Sugihara, Relation between outcomes and localisation of p-mTOR expression in gastric cancer, *Br. J. Cancer*, vol. 100, no. 5, pp. 782–788, 2009.

[79] T. Huang, F. Zhou, X. Yuan, T. Yang, X. Liang, Y. Wang, H. Tu, J. Chang, K. Nan, and Y. Wei, Reactive oxygen species are involved in the development of gastric cancer and gastric cancer-related depression through ABL1-mediated inflammation signaling pathway, *Oxid. Med. Cell. Longev.*, doi: 10.1155/2019/5813985.

[80] N. Kassem, O. M. Ismail, H. Elomri, and M. A. Yassin, Nilotinib induced recurrent gastric polyps: Case report and review of literature, *American Journal of Case Reports*, vol. 18, pp. 794–798, 2017.

Yin Peng received the BS degree in biology from Sichuan University in 2004, Chengdu, China and the PhD degree in molecular and cellular biochemistry from Loyola University Chicago in 2011. She is currently an assistant professor in pathology at Shenzhen University, China. She also worked as a postdoctoral researcher at Shenzhen University from 2014 to 2017. Her long-standing research interests lie in the molecular mechanisms of cell signaling transduction and tumorigenesis. Her lab is actively exploring how oncogenic circRNAs regulate critical signaling transduction pathways, such as Wnt/β-catenin pathway and NF-κB signaling pathway to promote cell survival, proliferation, and migration in cancer cells. She has published over 20 first author/corresponding author papers in professional and peer-reviewed journals.

Md. Tofazzal Hossain received the BS and MS degrees from University of Rajshahi, Bangladesh in 2009 and 2010, respectively. Currently, he is pursuing the PhD degree at Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. He is also an assistant professor from 2016 at the Department of Statistics, Bangbandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj, Bangladesh. His research interest includes bioinformatics and computational biology.

Md. Selim Reza received the BS and MS degrees from University of Rajshahi, Bangladesh in 2014 and 2015, respectively. Currently, he is pursuing the PhD degree at Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. His research interest includes bioinformatics and computational biology.

Shenzhong Feng is a professor and the director assistant of Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences. He received the bachelor degree from University of Science and Technology of China in 1991, and the PhD degree from Beijing Institute of Technology in 1997. Before joining SIAT in 2009, he has worked as an associate professor in the Institute of Computing technology and a visiting professor in University of Toronto. His research interest includes high performance computing, cloud computing and bioinformatics, etc. He has published more than 60 research papers which are indexed by SCI/EI, and applied more than 20 patents in the resent 5 years.

Yanjie Wei is a professor and the executive director in Center for High Performance Computing, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences. He received the PhD degree from Michigan Tech University in the field of computational biophysics in 2007. From 2008 to 2011, he worked as a postdoctoral research associate at Princeton University. His research focuses on high performance computing and computational biology/bioinformatics. He has published more than 100 peer reviewed journal/conference papers.