



NATURE'S SYMPHONY

ISSN-E: 3007-2034

ISSN-P: 3007-2026

Website link: www.tsfn.com

Exploring Prognostically Relevant Genes in Gastric Cancer via Gene Expression Profiling and Pathway Analysis

Momna Irfan, Shahzadi Noreen, Rizwan Ashraf and Saba Nawaz

To cite: Irfan, M., Noreen, S., Ashraf, R., and Nawaz, S. 2025. Exploring Prognostically Relevant Genes in Gastric Cancer via Gene Expression Profiling and Pathway Analysis. *Nature's Symphony*, 3(2):01-25.

Article QR



Published online: 13 August 2025

Submit your article to this journal: <https://tsfjb.com/ojs/index.php/tsfn/about/submissions>

View archive: <https://tsfjb.com/ojs/index.php/tsfn/issue/archive/>



NATURE'S SYMPHONY
Volume 3 Issue 2, Fall 2025

Exploring Prognostically Relevant Genes in Gastric Cancer via Gene Expression Profiling and Pathway Analysis

Momna Irfan¹ | Shahzadi Noreen^{1*} | Rizwan Ashraf¹ | Saba Nawaz¹

¹Department of Biotechnology,
Faculty of Sciences, International
Institute of Science, Arts and
Technology, Gujranwala, Pakistan

*Correspondence

Shahzadi Noreen

Email:

shahzadi.noreen@iisat.edu.pk

Keywords:

Gastric cancer;
Genomics; Serum- and
glucocorticoid-
regulated kinases;
Peroxisomal membrane
protein 2; Orphan
receptors linked to G
proteins; Tissue
inhibitor of
metalloproteinase 1

ABSTRACT

Gastric cancer (GC), the fifth most common malignancy worldwide, continues to pose diagnostic and prognostic challenges owing to the lack of reliable non-invasive biomarkers. This study aims to identify and validate prognostic biomarkers in gastric cancer through integrated bioinformatics analysis. Three publicly available gene expression profiles (GSE54129, GSE79973, and GSE118916), comprising 182 gastric cancer and 175 normal tissue samples, were retrieved from the NCBI Gene Expression Omnibus (GEO) database. A total of 179 common differentially regulated genes (DRGs) were identified through Venn diagram analysis. Functional enrichment, Gene Ontology (GO), and KEGG pathway analyses were conducted to explore the biological significance of these genes. Protein-protein interaction networks and survival analysis were performed using STRING, Cytoscape, and Kaplan-Meier plotter tools. Among the identified genes, SGK1 and PXMP2 were associated with favorable overall survival (log-rank $p < 0.05$; HR < 1), suggesting a protective role, while COL10A1, GPRC5C, AKR1C1, and TIMP1 were significantly linked to poor prognosis (log-rank $p < 0.05$; HR > 1). These findings enhance our understanding of gastric cancer pathogenesis and offer potential candidates for the development of prognostic biomarkers and potential therapeutic targets for future management of GC.

1. INTRODUCTION

Gastric cancer (GC) ranking as the fifth most common cancer and the third leading cause of cancer-related deaths worldwide poses significant health concerns (Herrera-Kok *et al.*, 2023). This is because GC is mostly diagnosed at an advanced stage after it has spread to other bodily areas. Early detection is crucial for improving patient outcomes, yet current diagnostic methods—such as endoscopy, biopsy, and imaging—are invasive, costly, and often uncomfortable for patients. These conventional approaches also have showed

limited sensitivity in diagnosing early-stage tumors (Yang *et al.*, 2020; Guan *et al.*, 2023).

This underscores the urgent need for the development of precise, non-invasive diagnostic methods, such as liquid biopsy approaches, which can detect tumor-related genetic and epigenetic changes in body fluids. Recent advancements in omics technologies and systems biology approaches have provided new opportunities to explore the complex molecular mechanisms underlying gastric cancer development and to discover novel biomarkers and therapeutic targets (Puliga *et al.*, 2021; Wang *et al.*, 2024). The field of healthcare and diagnostics has undergone a

revolution thanks to advancements in genomic research and the availability of large-scale datasets (Behl *et al.*, 2021). The introduction of powerful genomic, proteomic, and epigenomic technologies has revolutionized our understanding of the genetic basis of cancer including GC (Dai and Shen, 2022).

Multinational projects, such as the International Human Cancer Consortium and the Cancer Genome Atlas, have provided insights on the complicated genomic alterations underlying various cancer types, including GC (Giunta, 2021). Additionally, a vast collection of genomic datasets from multiple research studies may be found in the National Center for Biotechnology Information's (NCBI) Gene Expression Omnibus (GEO), a comprehensive public database. These datasets offer invaluable insights into genetic variants, gene expression patterns, and molecular markers associated with diseases, including cancer (Clough *et al.*, 2024). By utilizing these datasets researchers can create reliable and non-invasive diagnostic methods which offer improved early detection capabilities with better accuracy and sensitivity (Clark and Lillard, 2024).

Importantly, epigenetic alterations in oncogenes are more common, stable, and easily detectable in blood and other body fluids, making them ideal targets for least invasive liquid biopsy practices (Prabhu *et al.*, 2024). Unusual methylation patterns like hyper methylation of tumor suppressor genes (local) (Wen *et al.*, 2023) and hypo methylation of oncogenes (global) as significant role players in the biogenesis of various cancers including GC offer vast therapeutic potential (Hosseini *et al.*, 2023).

This study aimed to identify potential GC biomarkers through the integration and bioinformatics analysis of genomic datasets from the GEO NCBI repository. We utilized a systems biology approach to integrate

genomics data and expose the main molecular pathways and networks linked with gastric cancer progression. By analyzing three gastric cancer datasets GSE79973, GSE13911, and GSE29272 from the NCBI GEO database we have identified 179 DRGs in GC. After the functional and pathway analyses of the identified DRGs, interactive regulatory PPI network was constructed and analyzed for seed genes identification via module analyses. The identified seed genes were subjected to comparative expression analysis in TCGA GC and GTEx normal samples, epigenetic methylation and survival analysis. In addition to exploring novel biomarkers, this study also reports previously implicated genes in GC. Remarkably, PXMP2, TIMP1 and COL10A1, have been previously found correlated with GC progression and bad prognosis, signifying their translational significance and clinical pertinence (Necula *et al.*, 2020; Liu *et al.*, 2021b; Li *et al.*, 2024; Zheng *et al.*, 2024). The study findings contribute to the understanding of GC pathogenesis and propose potential biomarkers for non-invasive diagnosis, prognosis, and targeted therapy of the disease.

2. METHODS

2.1 Obtaining Data and Identifying Differentially Regulated Genes (DRGs)

Present study began with an inquiry of the GEO database using the search term "Gastric cancer" and particular filters customized to GC data from humans. This resulted in the selection of three distinct datasets: GSE79973, GSE13911, and GSE29272 from the GEO database. These datasets varied in numbers of samples analyzed and the data collection platforms, i.e., GSE79973 analyzed 10 normal adjacent and 10 GC tissue samples (Jin *et al.*, 2017), while GSE13911 contained 38 tumors and 31 normal tissues. Additionally, GSE29272 included 134 samples each of GC and normal adjacent mucosa totaling at 268 samples (Wang *et al.*,



2013; Li *et al.*, 2014). The identification of DRGs between GC and adjacent normal tissues was performed by the GEO2R tool from NCBI with set criteria of $P\text{-value} \leq 0.05$ and $|\log_2FC| \geq 1.5$. The probes without corresponding IDs were discarded and those representing the same genes were averaged. To identify common DRGs across the three datasets, Venn diagram online tool accessible at <https://bioinformatics.psb.ugent.be/webtools/Venn/> was used. The identified common DRGs were used for subsequent analyses.

2.2 Functional Annotation and Pathway Enrichment

Differentially Regulated Genes identified across the three datasets were subjected to Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) analysis with the help of tools like Metascape (Zhou *et al.*, 2019); (v3.5.20240101; <https://metascape.org/>) and DAVID (Database for Annotation, Visualization, and Integrated Discovery; <https://david.ncifcrf.gov/>; v2023q4) (Wu *et al.*, 2006) with significance threshold set at $p\text{-value} \leq 0.05$. Metascape using a kappa score of 0.3 reveals insights into enriched ontology terms and pathways, that are further grouped based on gene similarity index into networked clusters using MCODE (1.5+) algorithm of cytoscape. In the network, node sizes represent gene associations, colors indicate clusters, and edges denote similarities. Metascape also examines tissue enrichments of genes via PaGenBase (Piñero *et al.*, 2016) and their relationships with various diseases by the gene-disease database DisGeNET (Pan *et al.*, 2013) that hosts a vast repository of genes and variants associated with human disorders.

2.3 Protein-protein Interaction (PPI) Network and Hub Genes Identification

As genes and proteins sharing similar functions tend to appear as interconnected, protein-protein interaction (PPI) analysis is crucial for

thorough understanding of the functional roles of proteins in biological systems and predicting their therapeutic targets. PPI network construction was performed by using data from the GeneMania database (Franz *et al.*, 2018), which takes into account information from interactions between proteins, proteins and DNA including pathways and functionality. In current study analysis the gene identifiers of the GC DRGs were submitted to GeneMANIA using the default parameters. The resultant network was then imported into Cytoscape 3.8 software (Smoot *et al.*, 2011) and topological analysis performed using Cytoscape's Network Analyzer plugin (Assenov *et al.*, 2007). In addition, the Molecular Complex Detection (MCODE) module of the Cytoscape (Bader and Hogue, 2003) was used to identify clustered modules in the GC network with default parameters (degree cutoff= 2, node score cutoff= 0.2), K-score =2, maximum depth= 100).

2.4 Profiling Expression of Key Genes and Survival Outcomes Analysis

Hub genes identified in the GC DRGs were validated using Genotype-Tissue Expression (GTEx) and TCGA data-based gene expression profile interactive analysis (GEPIA) tool (Tang *et al.*, 2017). cBioPortal investigated relationships between methylation patterns and hub gene expression profiles (Cerami *et al.*, 2012). Furthermore, the Kaplan Meier-plotter was used to evaluate the effect of a gene's regulation (up/-down expression) on a patient's overall survival using data from The European Genome-phenome Archive (EGA), GEO, and The Cancer Genome Atlas TCGA. The hazard ratio (HR) at a 95% confidence interval and the log rank P value were calculated and presented on the graph.



3. RESULTS

3.1 Differentially Regulated Genes (DRGs) Identification

This study used the NCBI GEO2R tool to analyze three genome profiles (GSE13911, GSE29272, and GSE79973) to identify the genes that are differentially expressed in gastric cancer (GC). The analysis identified a total 1214, 225, and 671 upregulated DRGs (adjusted P value ≤ 0.05 ; $|\log_2FC| \geq 1.5$) from

GSE13911, GSE29272, and GSE79973 datasets, respectively. While downregulated DRGs from these datasets were 3603, 391, and 2903 (adjusted P value ≤ 0.05 and $|\log_2FC| \leq -1.5$), respectively (Figure 1a). Venn analysis identified 179 DRGs that were common to all datasets. Among them, 76 were upregulated ($\log_2FC \geq 1.5$) and 103 were downregulated ($\log_2FC \leq -1.5$) (Figure 1b, 1c; Table 1; Table S1).

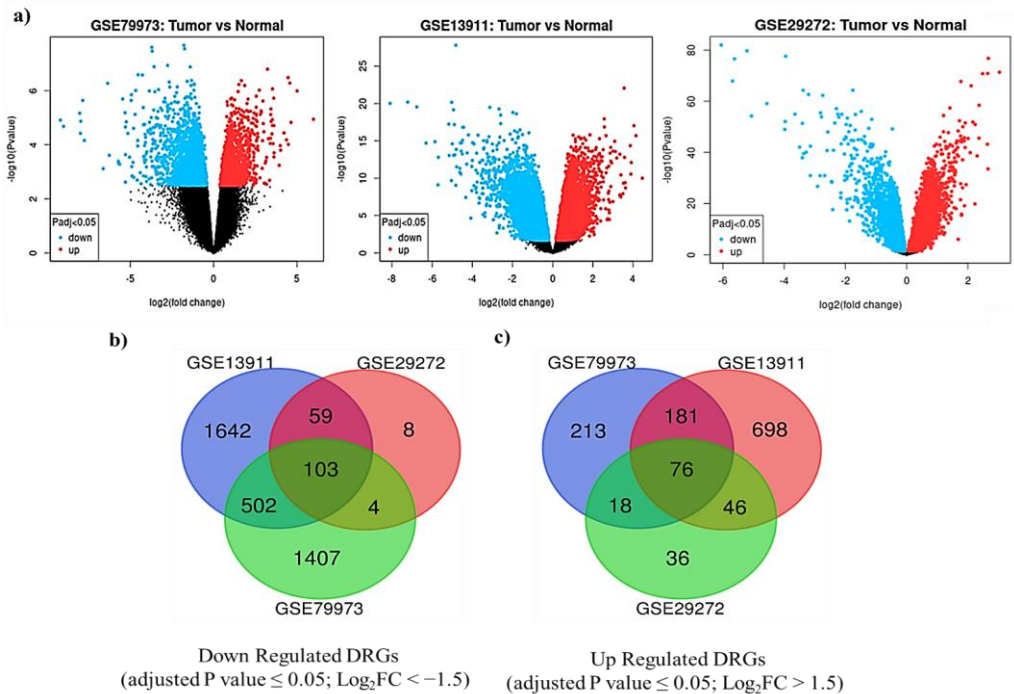


Figure 1. Differentially Regulated Genes identified in GC. a) Volcano Plots showing DRGs in three GC genomic datasets (GSE79973, GSE13911, and GSE29272, respectively). b) Among the three datasets, 103 DRGs were found downregulated ($\log_2FC < -1.5$), c) while 76 genes were found upregulated ($\log_2FC > 1.5$). The corresponding datasets are shown by different colors in b) and c)

Table 1. Differentially Regulated Genes in GC as identified from 3 genomic profiles

Expression	No.	DRG IDs
Down regulated	103	EPB41L4B HMGS2 XK IGH MAOA LOC101930400///AKR1C2 KIAA1324 GKN1 SULT1C2 FCGBP GPRC5C AKR7A3 PBLD PXMP2 DGKD KLF4 MT1G CYP3A5 GC ALDH6A1 NEDD4L PDGFD SGK1 CA9 ADH1C CYP2C9 PIK3C2G XYLT2 PGA4///PGA3///PGA5 MT1E ADAM28 KRT20 HRASLS2 ADGRG2 MYOC GIF MT1HL1 ADTRP RNASE1 MT1F AKR1C1 SOSTDC1 SH3GL2 CKMT2 KCNJ16 CKM GSTA1 S100P FBP2 AKR1B10 AADAC SLC16A7 ATP4A UGT2B15 AZGP1P1///AZGP1 FOLR1 MAL CHGA CAPN9 MT1M ESRRG VILL CPA2 FA2H CCKBR HPGD MT1X ADA TMPRSS2 LIPF ETNPPL SLC28A2 MYRF UBL3 NQO1 CYP2C18 ALDH3A1 AZGP1 COL2A1 ALDH1A1 FMO5 CA2 LTF ATP4B SIDT2 CHIA METTL7A PLLP PGC SCNN1A KCNE2 RAB27A GPX3 MUC5AC MT1H AKR1C3 ALDOB KCNJ15 C1orf116 IGHA2///IGHA1///IGH HDC SELENBP1 JCHAIN
Up Regulated	76	PLAU IGF2BP3 HMGB3 LOC100129518///SOD2 GPNMB CDC25B KPNA2 COL1A1 LGALS1 APOE EDNRA IGFBP7 CST1 CXCL8 FN1 MAD2L1 CKS1B TNFAIP6 LEF1 CTSB COL5A2 THY1 SPP1 SFRP4 CSE1L CDH11 CAMK2N1 GREM1 FBN1 TIMP1 RAB31 BGN MIR1292///SNORD110///SNORD86///SNORD57///NOP56 ECT2 ASPN COL3A1 COL4A2 ASPM SERPINH1 FAP INHBA S100A4 UBE2C PRC1 CEP55 HOXC10 CEMIP TOP2A THBS1 NT5DC2 THBS2 NID2 MFAP2 OLFML2B S100A10 COL4A1 VCAN DTL SULF1 MEST LY6E RARRES1 COL6A3 COL10A1 FCER1G SPARC COL5A1 LOC101928916///NNMT PMEPA1 COL11A1 HOXC6 FLNA COL1A2 APOC1 NEK2 SPOCK1

3.2 Functional Enrichment and Pathway Analysis of GC DRGs

The Gene Ontology (GO) enrichment of the DRGs identified in gastric cancer (GC) was conducted using the DAVID and Metascape software. The enriched biological processes for upregulated DEGs (P-value ≤ 0.05) included cell adhesion, extracellular matrix organization, positive regulation of cell migration, angiogenesis, positive regulation of cell proliferation, signal transduction, and cell-matrix adhesion (Table 2). The identified molecular functions of the upregulated DRGs were platelet-derived growth factor binding,

extracellular matrix binding, protein binding, identical protein binding, and extracellular matrix structural constituent that confers tensile strength. Aldo-keto reductase (NADP), D-threo-aldose 1-dehydrogenase, retinal dehydrogenase, oxidoreductase, and alditol: NADP+ 1-oxidoreductase activity, on the other hand, are the molecular functions of downregulated DRGs (Table 2). Drug metabolism-cytochrome P450, gastric acid secretion, chemical carcinogenesis-DNA adducts, glycolysis/gluconeogenesis, tyrosine, arginine proline, and glycerolipid metabolism are among the key KEGG pathways of

downregulated GC DRGs (P-value ≤ 0.05). On the other hand, KEGG pathways in which upregulated DRGs found enriched were protein digestion and absorption, ECM-receptor interaction, focal adhesion, PI3K-Akt signaling pathway, proteoglycans in cancer, TGF-beta signaling pathway, and pathways in cancer (Table 2). The biological pathways and enriched functions linked to GC genes were also revealed by Metascape. Metascape analysis hierarchically clustered statistically significant terms or pathways in the form of network and enriched terms identified were digestion, hormone response, gastric acid secretion, and negative regulation of cell migration (Figure 2). The networked

significant terms or pathways further subjected to module analysis by MCODE algorithm of cytoscape that resulted in identification of five clusters each with a hub gene. Tissue enrichment analysis by Pattern Gene Database revealed abundant presence of GC DRGs in the stomach, smooth muscle, liver, placenta, small intestine, kidney, trachea, uterus, and prostate (Figure 3a) Additionally, DRGs were found to

have strong associations with meningioma, lupus nephritis, adenoma of the large intestine, and Barrett esophagus by DisGeNET analysis, a database of the gene-disease search (Figure 3b).

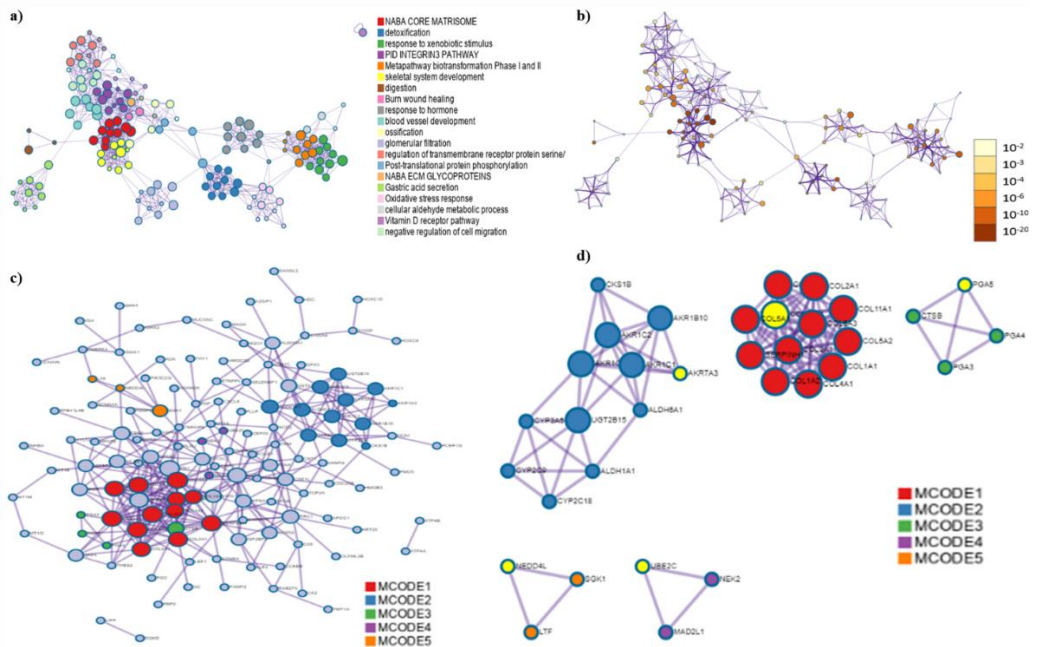


Figure 2. Metascape identified Enriched Terms (GO/KEGG/pathways) of GC DRGs. Representative terms (GO/KEGG/pathways) were hierarchically clustered based on gene similarity, with a threshold of 0.3 Kappa (a) and displayed as circles with size proportional to gene counts and colors denoting individual ontology clusters. Terms with Similarity > 0.3 are linked by edges. Enrichment network colored by P-value (b). Enriched ontology clustered network further subjected to analysis by the MCODE Algorithm which identified five clusters (c) with a hub gene in each (d).

Table 2. Ontology Enrichment analysis of GC DRGs.

Regulation	Category	Term	Count	%	P-Value
Downregulated DRGs	DISGENET	Carcinoma, Cribriform and Granular Cell	6	6.3	1.2E-3
	DISGENET	Adenocarcinoma, Oxyphilic, Tubular and Basal Cell	6	6.3	1.2E-3
	DISGENET	Malignant neoplasm of prostate	11	11.6	9.2E-3
	DISGENET	Malignant neoplasm of stomach	7	7.4	1.7E-2
	DISGENET	Malignant tumor of colon	4	4.2	9.7E-2
	GAD_DISEASE	Breast Neoplasms Neoplasms, Germ Cell and Embryonal Ovarian Neoplasms	5	5.3	1.7E-4
	GAD_DISEASE	drug-related genes	10	10.5	2.0E-3
	GAD_DISEASE	prostate cancer	11	11.6	8.1E-2
	GAD_DISEASE	Esophageal Neoplasms Head and Neck Neoplasms Laryngeal neoplasm Laryngeal Neoplasms Mouth neoplasms Oesophageal neoplasm Pharyngeal Neoplasms	3	3.2	2.7E-2
	GAD_DISEASE	Breast Neoplasms Neoplasms, Germ Cell and Embryonal Ovarian Neoplasms	5	5.3	2.4E-3
	KEGG_PATHWAY	Drug metabolism - cytochrome P450	8	8.4	5.90E-07
	KEGG_PATHWAY	Gastric acid secretion	7	7.4	1.30E-05
	KEGG_PATHWAY	Chemical carcinogenesis - DNA adducts	5	5.3	1.40E-03
	KEGG_PATHWAY	Glycolysis / Gluconeogenesis	4	4.2	1.10E-02
	KEGG_PATHWAY	Tyrosine, Arginine proline and Glycerolipid metabolism	3	3.2	2.60E-02
	GOTERM_MF_DIRECT	aldo-keto reductase (NADP) activity	4	4.2	4.80E-05
	GOTERM_MF_DIRECT	D-threo-aldose 1-dehydrogenase activity	4	4.2	8.50E-05
	GOTERM_MF_DIRECT	retinal dehydrogenase activity	3	3.2	6.40E-04



Regulation	Category	Term	Count	%	P-Value
	GOTERM_MF_DIRECT	oxidoreductase activity	7	7.4	9.50E-04
	GOTERM_MF_DIRECT	alditol: NADP+ 1-oxidoreductase activity	3	3.2	1.80E-03
Upregulated DRGs	DISGENET	Liver carcinoma	13	17.8	1.40E-04
	DISGENET	Breast Carcinoma	11	15.1	3.20E-03
	DISGENET	Malignant neoplasm of salivary gland	4	5.5	3.50E-03
	DISGENET	Hereditary Diffuse Gastric Cancer	6	8.2	4.90E-02
	DISGENET	Malignant neoplasm of stomach	6	8.2	5.30E-02
	GAD_DISEASE	prostate cancer	7	9.6	4.80E-02
	GAD_DISEASE	gallstones	2	2.7	5.10E-02
	GAD_DISEASE	metabolism disorders	2	2.7	7.50E-02
	GAD_DISEASE	Duodenal Ulcer Helicobacter Infections	2	2.7	8.50E-02
	GAD_DISEASE	lung cancer	7	9.6	8.60E-02
	KEGG_PATHWAY	Protein digestion and absorption	10	13.7	5.70E-10
	KEGG_PATHWAY	ECM-receptor interaction	9	12.3	4.60E-09
	KEGG_PATHWAY	Focal adhesion	10	13.7	2.20E-07
	KEGG_PATHWAY	PI3K-Akt signaling pathway	9	12.3	1.70E-04
	KEGG_PATHWAY	Proteoglycans in cancer	6	8.2	2.20E-03
	KEGG_PATHWAY	TGF-beta signaling pathway	4	5.5	1.30E-02
	KEGG_PATHWAY	Pathways in cancer	7	9.6	3.10E-02
	GOTERM_MF_DIRECT	extracellular matrix structural constituent	18	24.7	1.70E-21
	GOTERM_MF_DIRECT	extracellular matrix structural constituent conferring tensile strength	10	13.7	7.70E-14
	GOTERM_MF_DIRECT	platelet-derived growth factor binding	5	6.8	6.20E-08
	GOTERM_MF_DIRECT	extracellular matrix binding	6	8.2	1.30E-07
	GOTERM_MF_DIRECT	protein binding	62	84.9	4.70E-04
	GOTERM_MF_DIRECT	identical protein binding	14	19.2	1.30E-02

Regulation	Category	Term	Count	%	P-Value
	GOTERM_BP_DIRECT	cell adhesion	16	21.9	1.6E-9
	GOTERM_BP_DIRECT	extracellular matrix organization	10	13.7	2.9E-8
	GOTERM_BP_DIRECT	positive regulation of cell migration	7	9.6	4.3E-4
	GOTERM_BP_DIRECT	angiogenesis	6	8.2	2.4E-3
	GOTERM_BP_DIRECT	positive regulation of cell proliferation	8	11.0	3.7E-3
	GOTERM_BP_DIRECT	signal transduction	10	13.7	4.8E-2
	GOTERM_BP_DIRECT	cell-matrix adhesion	3	4.1	6.0E-2

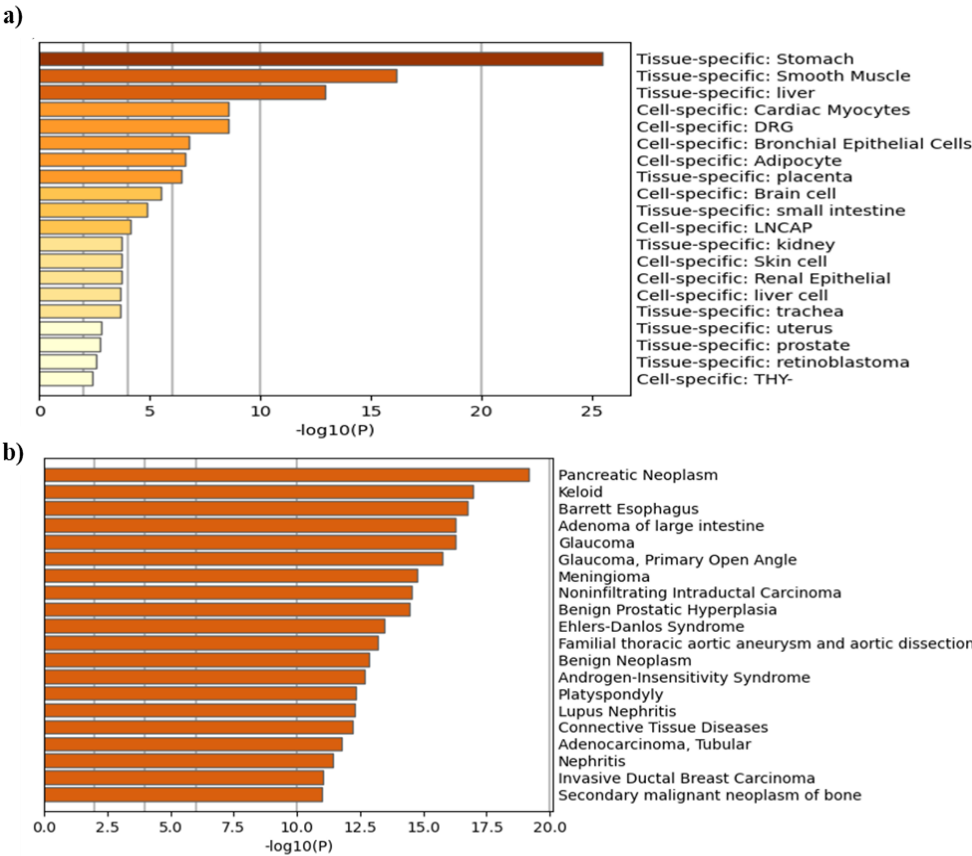


Figure 3. Metascape enrichment of GC DRGs a) Tissue enrichment of GC genes from Pattern Gene Database (b) and relations with various disease/illnesses in body by DisGeNET

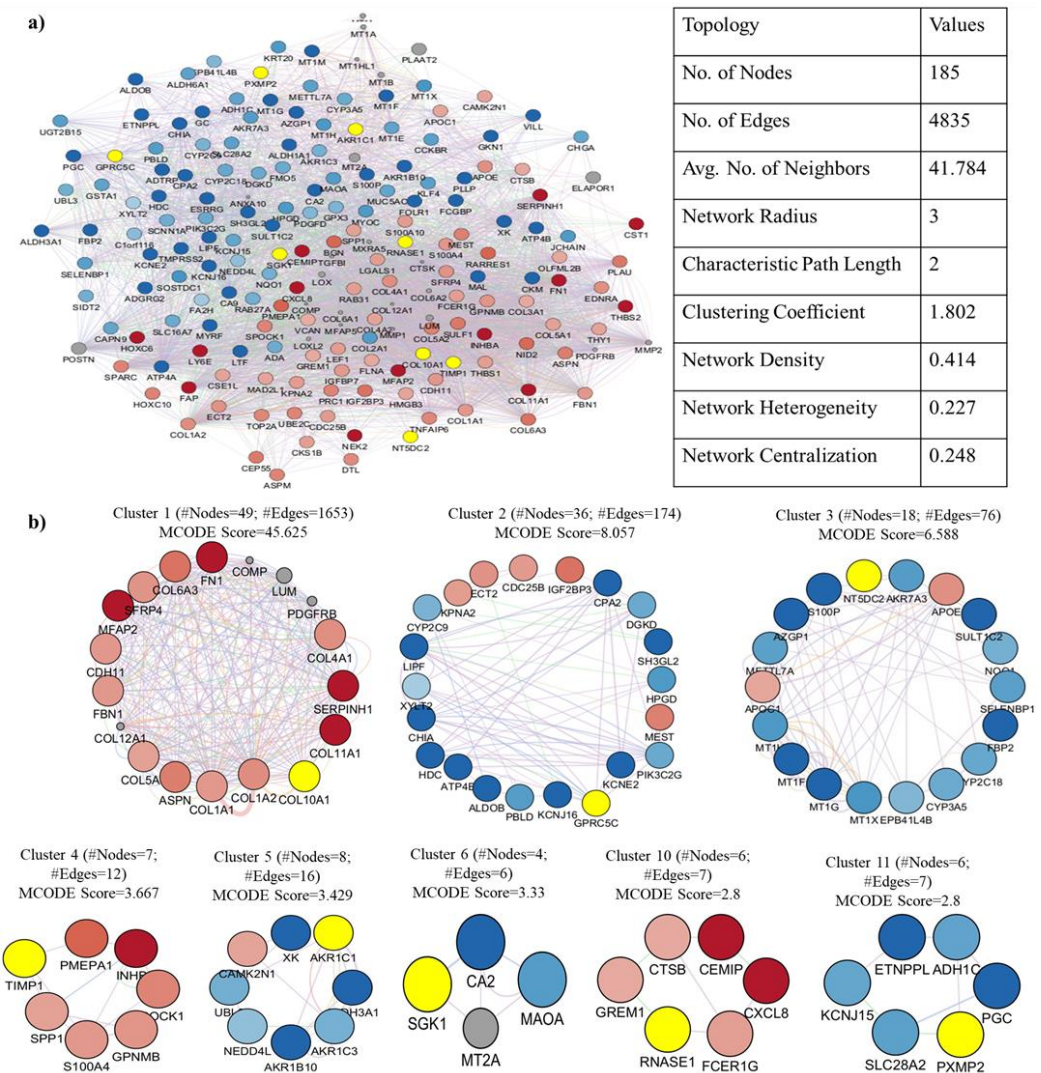


Figure 4. Protein-protein interaction (PPI) network of GC DRGs a) the network generated by GeneMANIA and further visualized by Cytoscape. In this presentation nodes represent proteins with blue color denoting down- and red up-regulation of a protein while edges signify interactions. The topological parameters of the network are shown in table next to the network. b) MCODE identified GC network distinct modules with the Seed/Hub genes shown by yellow nodes in each identified clustered module

3.3 GC Network and Module Analysis

In order to observe the functional roles of the GC DRGs, interactions between the 179 common DRGs identified among three datasets were extracted using the GeneMANIA database after entering their IDs into the server. A total of 4835 interactions were retrieved that were found free of self-loops. This interactive data then saved into a TSV file for further analysis of the network in the Cytoscape. Using the Cytoscape Network Analyzer tool, various topologies of the network were detected that involved the number of nodes (185), number of edges (4835), average number of neighbors (41.784), network radius (3), characteristic path length (2), clustering coefficient (1.802), network density (0.414), network heterogeneity (0.227), and network centralization (0.248) (Figure 4a). The MCODE app in Cytoscape was then used to perform cluster analysis in the GC network (Bader & Hogue 2003). Following analysis, eight clusters were identified within the GC network, each centered around a distinct seed gene (Figure 4b). As presented by the yellow nodes in Figure 4b, the hub/seed genes within these clusters were COL10A1 (MCODE score: 45.625), GPRC5C (MCODE score: 8.057), NT5DC2 (MCODE score: 6.588), TIMP1 (MCODE score: 3.667), AKR1C1 (MCODE score: 3.429), SGK1 (MCODE score: 3.33), RNASE1 (MCODE score: 2.8), and PXMP2 (MCODE score: 2.8), respectively.

3.4 Analysis of Gene Expression Patterns and Assessment of Survival Outcomes

Using the Cytoscape tool's MCODE plugin, the study found eight seed genes (GPRC5C, AKR1C1, SGK1, PXMP2, RNASE1, COL10A1, NT5DC2, and TIMP1) that were further examined for comparative expression profiling in normal and GC samples, methylation and survival analysis using the UALCAN and GEPIA tools. While the five genes, GPRC5C, AKR1C1, SGK1, PXMP2,

and RNASE1, were upregulated in normal samples, the results showed that three genes, COL10A1, NT5DC2, and TIMP1, had higher expression levels in tumor samples compared to GTEx normal samples (Figure 5a & 5b).

The impact of epigenetics regulation on hub gene's expression was evaluated by methylation profiling with the help of cBioPortal and UALCAN analyses. RNASE1, and COL10A1 were found to be hypermethylated (Beta value ≥ 0.5), whereas five hub genes NT5DC2, GPRC5C, PXMP2, SGK1, and TIMP1 were found hypomethylated (Beta value ≤ 0.5). Further the influence of methylation on gene expression was investigated. Except for SGK1, which showed a positive correlation (Beta value > 0.1 ; P-value < 0.05 ; Table 3; Figure 6a), all other hub genes showed negative correlations between DNA methylation and mRNA intensities (Table 3).

The effect of hub genes' up- or down-regulation on the overall survival of GC patients was evaluated by Kaplan–Meier survival analysis, and logrank P-values and hazard ratio (HR) was used for comparison. SGK1 and PXMP2 found to have HR values less than 1 (Logrank p-value < 0.05), signifying high expression levels association with favorable prognostic outcomes in GC hence referred as defensive genes (Figure 6b). However, six genes had HR > 1 (COL10A1, GPRC5C, AKR1C1, NT5DC2, TIMP1, and RNASE1) indicating that high gene levels linked to a poor prognosis in GC thus classified as risk genes (Figure 6b). Moreover, the survival outcomes were found insignificant for RNASE1 and GPRC5C (Logrank p-value > 0.05) (Figure 6b; Table S2).

The Human Protein Atlas (HPA) protein expression analysis revealed that PXMP2 and COL10A1 lacked expression data. HPA protein—expression analysis revealed that regulation/expression data for all other hub



genes, however, matched with the transcriptomic data from the NCBI GEO and TCGA datasets (as analyzed by GEPIA, UALCAN, and cBioPortal) exposing low levels of GPRC5C, AKR1C1, SGK1, RNASE1 and up regulated levels of NT5DC2 and TIMP1 in GC (Figure 7).

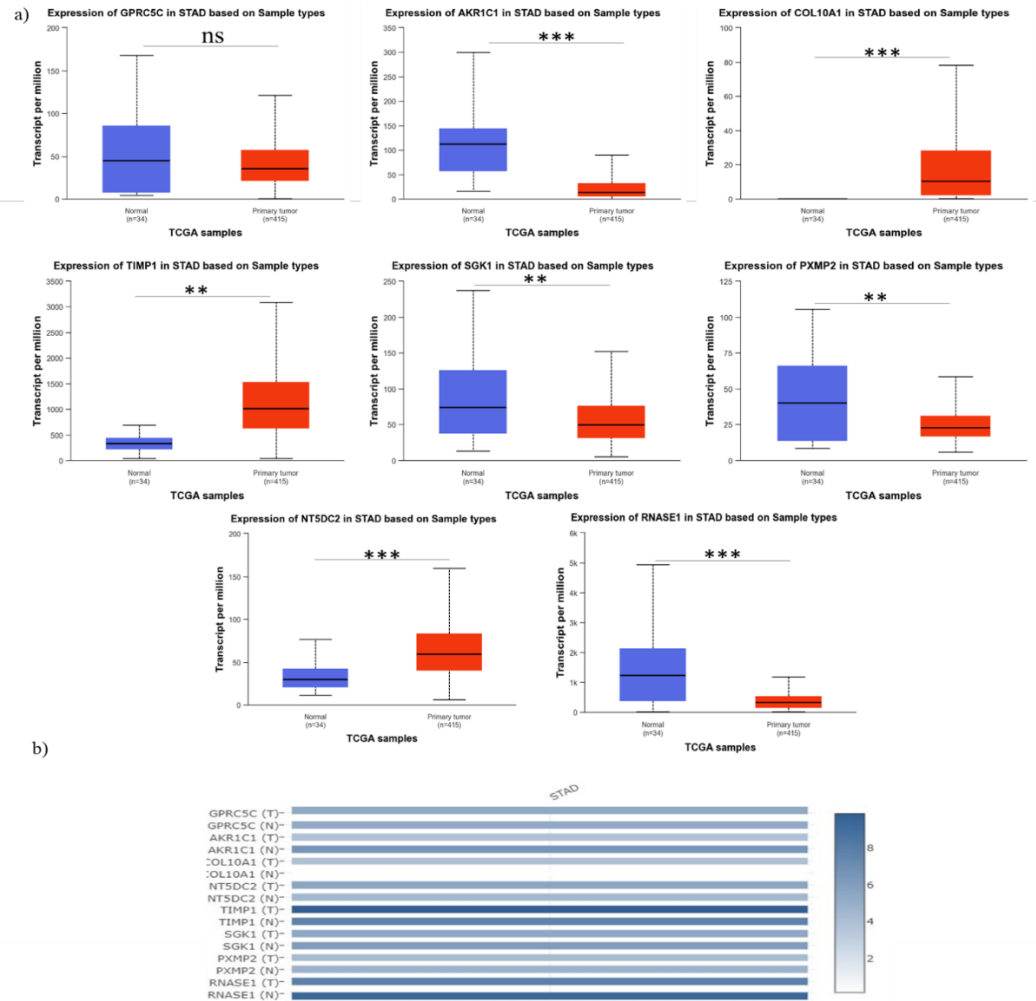


Figure 5. Comparative expression profiling of GC DRGs a) Comparative seed genes expression in TCGA stomach adenocarcinoma (STAD) and normal GTEx samples [$**p<0.01$, $***p<0.001$]. b) Heatmap portraying comparative profile of differentially represented genes in GC and normal samples

Table 3. The relationship between methylation and mRNA levels of hub genes

Hub genes	Methylation and mRNA Correlation
GPRC5C	Spearman: -0.09(p = 0.0983) Pearson: -0.01(p = 0.822)
AKR1C1	Spearman: -0.09(p = 0.0983) Pearson: -0.01(p = 0.812)
COL10A1	Spearman: -0.11(p = 0.0411) Pearson: -0.00(p = 0.950)
NT5DC2	Spearman: -0.24(p = 5.192e-6) Pearson: -0.16(p = 2.661e-3)
SGK1	Spearman: 0.19(p = 4.004e-4) Pearson: 0.09(p = 0.0919)
PXMP2	Spearman: -0.27(p = 4.01e-7) Pearson: -0.12(p = 0.0291)
RNASE1	Spearman: -0.06(p = 0.274) Pearson: 0.04(p = 0.407)

4. DISCUSSION

Gastric cancer is a serious health concern standing as fifth most prevalent and the fourth most common cause of cancer-related deaths globally (Smyth *et al.*, 2020). With nearly one million new cases per annum GC is predicted to reach up to 62% global burden by 2040 (Sexton *et al.*, 2020). Although the incidence rate of GC has declined in the United States over the past 50 years, the number of young people (under 50) diagnosed with advanced-stage GC is increasing (Thrift *et al.*, 2023).

Due to late-stage diagnosis of GC, the 5-year survival rate continues to be 28% or even lower (Rugge *et al.*, 2018) which signifies need for discovering novel precise biomarkers to facilitate early detection and improve GC therapeutics. Targeted approaches in GC management are desperately needed to get over these challenges. Limitations in molecular biology studies have prompted bioinformatics research to discover potential markers for cancer development and progression (Ma *et al.*, 2024). This study employs integrated

bioinformatics and systems biology approach to identify potential molecular signatures associated with GC. Through the analysis of multiple gene expression datasets from the GEO database, followed by functional enrichment and network analyses, a set of differentially regulated genes was identified. Notably, 08 hub genes including TIMP1, COL10A1, SGK1, PXMP2, GPRC5C, AKR1C1 and LOX were found to be significantly associated with GC prognosis.

Gene ontology and pathway enrichment analyses with various bioinformatics tools showed that GC DRGs are significantly enriched in multiple biological processes related to tumorigenesis, including cell adhesion, extracellular matrix organization, positive regulation of cell migration, angiogenesis, positive regulation of cell proliferation, signal transduction, and cell-matrix adhesion (Table 2). Furthermore, enriched molecular functions of the GC DRGs were aldo-keto reductase (NADP) activity, D-threo-aldo-1-dehydrogenase activity, retinal

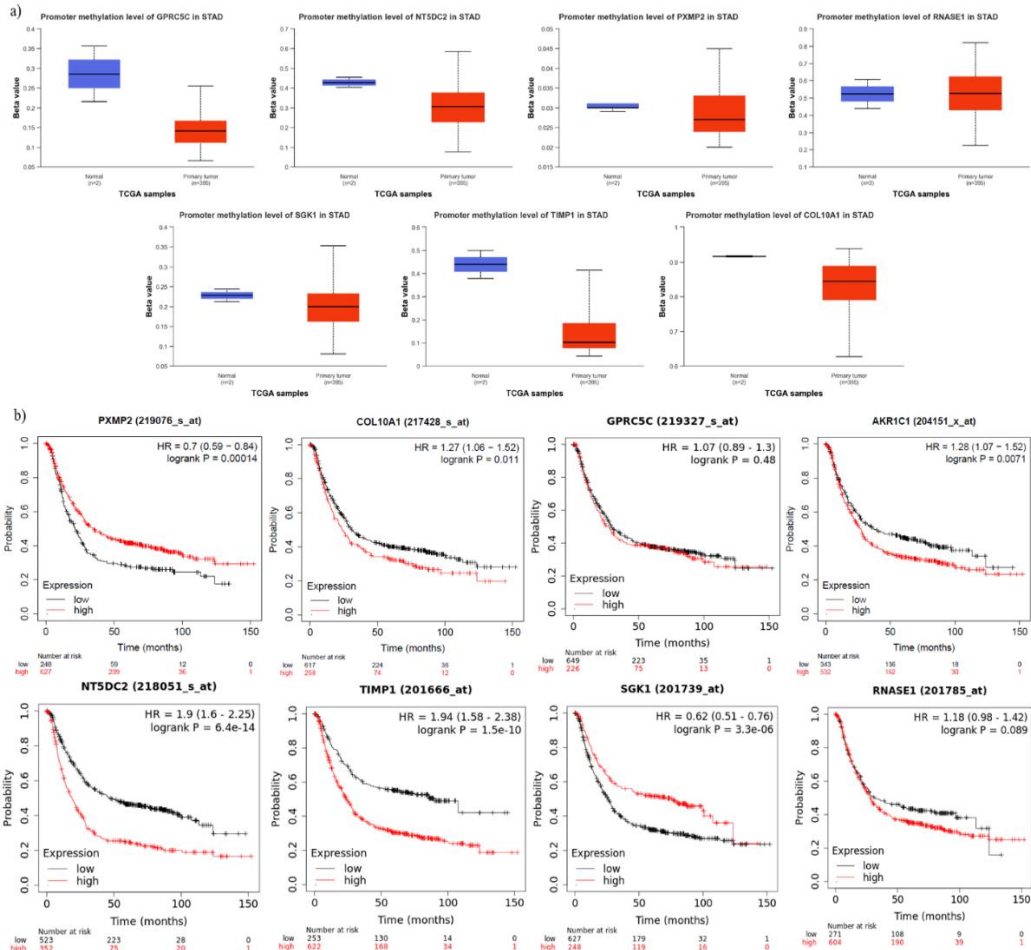


Figure 6. a) Methylation profile of GC hub genes. The analysis identified RNASE1 and COL10A1 as hyper-methylated genes (Beta value >0.5). b) Kaplan Meier Survival Plots of the GC hub genes. Among the 8 analyzed genes, two genes GPRC5C and RNASE1 had no significant impact on survival of GC patients ($P > 0.05$)

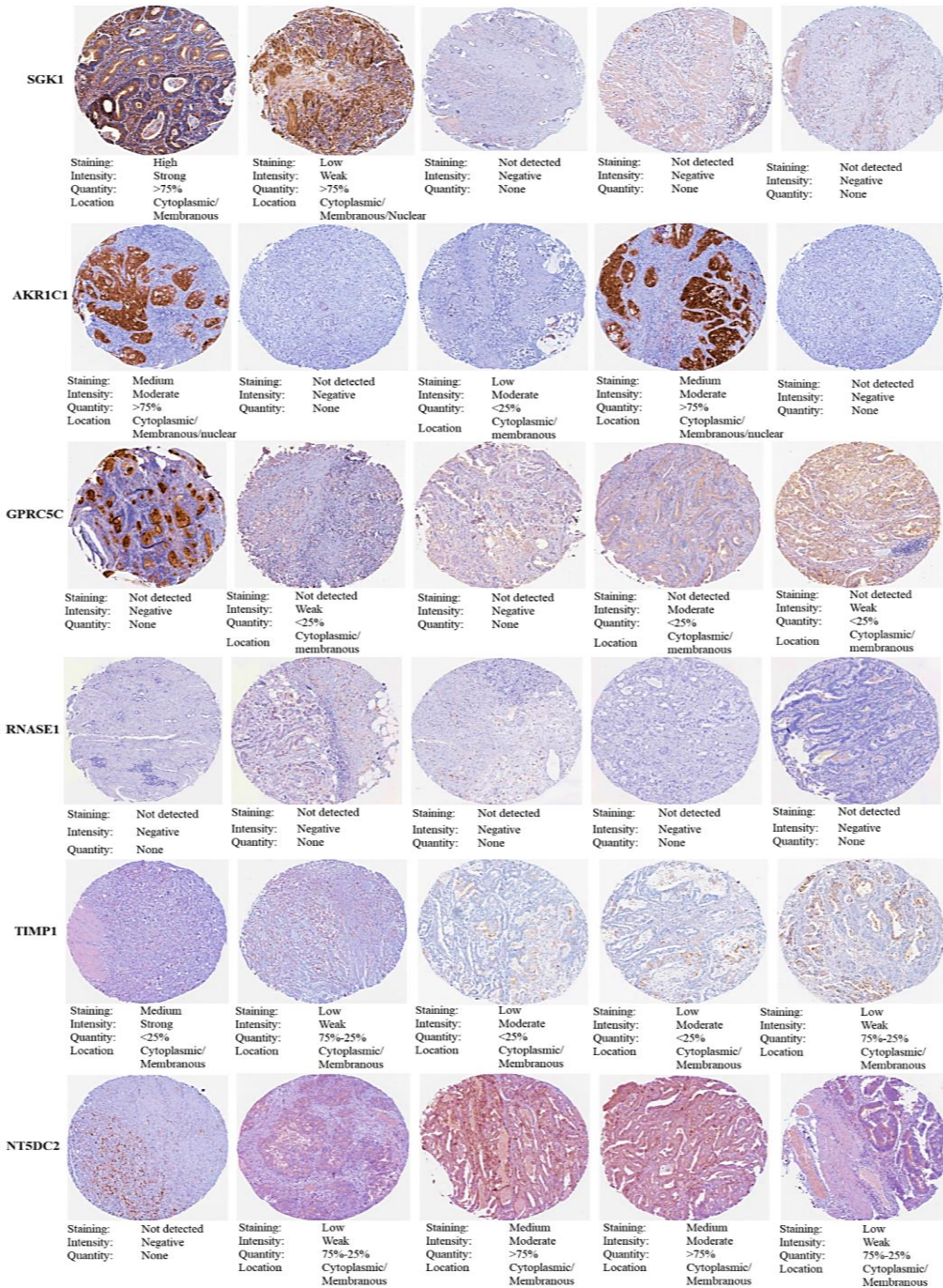


Figure 7. Protein expression analysis by Human Protein Atlas (HPA)

dehydrogenase activity, oxidoreductase activity, and alditol: NADP+ 1-oxidoreductase activity, extracellular matrix structural constituent, platelet-derived growth factor binding, extracellular matrix binding, protein binding, and identical protein binding (Table 2).

According to current study DAVID analysis, thirteen DRGs (FCGBP, AKR1C3, ALDOB, CA2, GKN1, KRT20, SELENBP1, CXXL8, CST1, GREM1, IGFBP7, PLAU, and RARRES1) revealed potential roles in the development of stomach malignant neoplasms. While HMGCS2, CKM, FOLR1, and LTF identified with established roles in malignant tumors of the colon (Table 2). GC DRGs identified as potential drug-related genes were UGT2B15, ADH1C, ALDH1A1, ALDH3A1, ALDH6A1, AADAC, CYP2C9, CYP3A5, GSTA1, and SULT1C2 (Table 2).

KEGG pathways in which the present study up regulated DRGs were found enriched included focal adhesion, the PI3K-Akt signaling pathway, protein digestion and absorption, the interaction among extracellular matrix and receptors, proteoglycans in cancer, and the TGF-beta signaling pathway. While chemical carcinogenesis, DNA adducts, drug metabolism, cytochrome P450, gastric acid secretion, glycolysis/gluconeogenesis, tyrosine, arginine proline, and glycerolipid metabolism were found enriched KEGG pathways of down-regulated DRGs (Table 2).

Focal adhesion pathway has previously been reported to linked with the metastatic spread of gastric cancer (Zhang *et al.*, 2022) like others i.e. breast (Ramírez-Ricardo *et al.*, 2020), and lung cancer (Wang *et al.*, 2020). The pathway by activating further signaling pathways through phosphorylation and protein-protein interactions, is crucial for the migration and invasion of cancer cells (Niland and Eble, 2019).

According to earlier research, PI3K/Akt/mTOR pathway displayed essential part in GC development as reported by Le Rhun *et al.* (2017) abnormal PI3K/Akt/mTOR pathway activation increases cell survival and metastasis by causing an imbalance between pro- and anti-apoptotic genes. By blocking genes associated with apoptosis, namely BCL2, this pathway is crucial for improving cell survival. Given its potential connection with GC, the pathway has been proposed as a promising therapeutic target for GC (Duan *et al.*, 2020).

Previous research has also highlighted links between interaction of extracellular matrix with receptors and GC development via direct promotion of cell division, migration, and survival. Also reported were key findings about the application of ECM-associated molecules as drug targets of the disease (Moreira *et al.*, 2022). Proteoglycans via regulating angiogenesis, cell migration, and survival play potential in GC progression (Pinto *et al.*, 2021; Yang *et al.*, 2022) just as in breast and ovarian cancer (Kizhakkeppurath Kumaran *et al.*, 2023). Cytochrome P450 plays a crucial role in carcinogenesis, influencing lung, breast, and stomach cancers, among others (Elfaki *et al.*, 2018).

PPI network analysis of 179 GC DRGs identified eight seed genes (GPRC5C, AKR1C1, SGK1, PXMP2, RNASE1, COL10A1, NT5DC2, and TIMP1) as major players in gastric cancer. These seed genes' expressions, epigenetic regulation, survival, and protein expression were further verified through various tools. The results exposed three seed genes—COL10A1, NT5DC2, and TIMP1—elevated expression in tumor samples, while five genes—GPRC5C, AKR1C1, SGK1, PXMP2, and RNASE1—were down regulated (Fig 5a & 5b). According to Kaplan–Meier analysis, two genes (PXMP2 and SGK1) with HR < 1 were found linked to good prognostic outcomes in GC, while six



genes (COL10A1, GPRC5C, AKR1C1, NT5DC2, TIMP1, and RNASE1) returned with a worse prognostic significance in GC (HR > 1) (Figure 6b).

Shao *et al.* (2020) using GSE79973 GEO genomic dataset along with GSE54129, and GSE19826 identified, IGF2BP3 and FOLR1 genes with strong relationship with H19 and PVT1 long non-coding RNAs, respectively in GC. A previous study reported reduced levels of *ATP4A*, *CA9*, *FGA*, *ALDH1A1*, and *GHRL* and enhanced expression of *TIMP1*, *SPPI1*, *CXCL8*, *THY1*, and *COL1A1* in GC while hunting GSE79973, GSE118916, and GSE29998 genome profiles (Liu *et al.*, 2021a). Zhao *et al.* (2023) reported high expression of COL5A2 and COL4A1 genes associated with worse survival in GC by profiling GSE profiles (GSE29272, GSE54129, GSE13911, GSE79973, GSE19826). CASR, CXCL12, and SST had been reported as key genes with prognostic significance in GC through bioinformatics profiling of GSE13911 and GSE29272 datasets (Li *et al.*, 2017).

Present study identified down regulated expression of GPRC5C in GC which is the third member of a new class of orphan receptors linked to G proteins (GPCRs) along with GPRC5A, GPRC5B, and GPRC5D that may play a part in carcinogenesis (Jiang *et al.*, 2018). Although no previous study has linked GPRC5C to GC, other members of this family, like GPRC5A, are found more abundant in cell membranes of GC tissues. Increased GPRC5A levels are linked with aggressive tumor characteristics and a lower overall survival (OS), influencing pathways including cAMP, NF- κ B, STAT3, and FAK/Src signaling (Iglesias González *et al.*, 2023).

Aldo-keto reductase family member C1 (AKR1C1) is an NADPH-dependent enzyme involved in hormone metabolism. AKR1C1 dysregulation is reported in GC tissues, while its overexpression encourages metastasis in

non-small-cell lung cancer and cisplatin resistance in bladder cancer (Zhou *et al.*, 2023). Serum- and glucocorticoid-regulated kinases (SGKs) within AGC family of serine/threonine kinases, have three isoforms: SGK1, SGK2, and SGK3. SGK is identified upregulated in certain malignancies while downregulated in others (Sang *et al.*, 2020). By stimulating K⁺ and Ca²⁺ channels, the Na⁺/H⁺ exchanger, amino acid and glucose transporters, downregulating Foxo3a and p53, and upregulating β -catenin and NF κ B, SGK1 encourages the epithelial to mesenchymal transition, adherence, invasion, motility, and survival of cancer cells (Cicenas *et al.*, 2022). In contrast to the current study findings a previous study reported enhanced levels of SGK1 in GC (Yao *et al.*, 2016).

Peroxisomal membrane protein 2 (PXMP2) is a channel-forming protein in the membrane of peroxisomes in mammals (Rokka *et al.*, 2009). PXMP2 along with PXMP 4 plays significant part in the regulation of self-renewal and chemo-resistance in GC. Additionally, PXMP2 gene expression reported as an independent predictor of disease-free and overall survival in breast cancer (Tong *et al.*, 2010).

RNASE1 is a member of RNA binding proteins (RBPs) with potential roles in human malignancies. In line with the current study analysis earlier studies have also revealed correlation between worse overall survival in GC and enhanced expressions of RNASE1 (Wei *et al.*, 2022). Collagen type X alpha 1 (COL10A1), a member of the collagen family is the main matrix component in the stroma. Studies have established the dynamic impacts of extracellular matrix on tumor cells growth, differentiation, progression, and metastasis. A significant rise in COL10A1 levels in GC tissue compared to adjacent normal tissue ($P < 0.05$) already stated by peers (Shen *et al.*, 2022).



The NT5DC family is an evolutionarily conserved family of 5'-nucleosidases that catalyze the hydrolytic cleavage of nucleotides. NT5DC family dysregulation has been linked to the development and progress of several cancers (Yu *et al.*, 2023). In present study analysis NT5DC2 upregulated expression observed in GC adenocarcinoma compared to the normal samples (Figure 5a & 5b). Further high levels of this gene correlated with worst prognosis in GC (HR =1.9, logrank P-value 6.4e-14). Parallel to current study findings upregulated expression of NT5DC2 had been reported linked with GC progression and poor previously (Sha *et al.*, 2023). Enhanced levels of COL10A1 are also found associated with progressive tumor stage in GC (Necula *et al.*, 2020).

Tissue inhibitor of metalloproteinase 1 (TIMP1) inhibits proteins with proteolytic activity and is linked to tumor differentiation, metastatic spread, and poor prognosis in GC (Zheng *et al.*, 2024). In alignment with present study report, previous research has described higher TIMP1 expression in GC, with advanced levels linking with worse survival outcomes in GC patients (Liu *et al.*, 2021b), supporting its potential as a non-invasive diagnostic and prognostic marker.

4.1 Limitations

In spite of the promising findings, the current study has some limitations. Firstly, the results are derived from in silico analyses and lack experimental validation via techniques i.e. qRT-PCR or immunohistochemistry. Secondly, while the normalization of the data was applied, integrating datasets from different sources (GSE79973, GSE13911, GSE29272) may still introduce batch effects that could influence gene expression profiles. Hence, while our analysis highlights non-invasive biomarker potential, further clinical validation using patient-derived liquid biopsy samples is

essential to approve their diagnostic and prognostic efficacy.

5. CONCLUSION

By integrating transcriptomic data with survival and methylation analyses, the study identified several key genes, including TIMP1, COL10A1, SGK1, and PXMP2, with potential clinical relevance. These findings offer a basis for further experimental and translational studies, predominantly in the area of liquid biopsy development, which could enable non-invasive, sensitive, and early detection of GC. Future investigations should aim to validate these biomarkers in large patient cohorts to support their potential implementation in the clinical management of GC.

Conflict of interest

The authors declare no conflict of interest.

Author Contributions

MI and SN carried out the data analysis and drafted the manuscript. RA and SN interpreted the data and finalized the manuscript.

Ethics Approval and Informed Consent

Not applicable

Availability of Data

Not applicable

Funding

Not applicable

REFERENCES

- Assenov, Y., Ramírez, F., Schelhorn, S.-E., Lengauer, T. and Albrecht, M. 2007. Computing topological parameters of biological networks. *Bioinformatics*, 24:282-284.
- Bader, G. D. and Hogue, C. W. 2003. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*, 4:2.



- Behl, T., Kaur, I., Sehgal, A., Singh, S., Bhatia, S., Al-Harrasi, A., Zengin, G., Babes, E. E., Brisc, C., Stoicescu, M., Toma, M. M., Sava, C. and Bungau, S. G. 2021. Bioinformatics accelerates the major tetrad: A aeal boost for the pharmaceutical industry. *International Journal of Molecular Sciences*, 22.
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E., Antipin, Y., Reva, B., Goldberg, A. P., Sander, C. and Schultz, N. 2012. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discovery*, 2:401-404.
- Cicenas, J., Meskinyte-Kausiliene, E., Jukna, V., Rimkus, A., Simkus, J. and Soderholm, D. 2022. SGK1 in Cancer: Biomarker and Drug Target. *Cancers (Basel)*, 14.
- Clark, A. J. and Lillard, J. W., Jr. 2024. A comprehensive review of bioinformatics tools for genomic biomarker discovery driving precision oncology. *Genes (Basel)*, 15.
- Clough, E., Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Lee, H., Zhang, N., Serova, N., Wagner, L., Zalunin, V., Kochergin, A. and Soboleva, A. 2024. NCBI GEO: archive for gene expression and epigenomics data sets: 23-year update. *Nucleic Acids Research*, 52:D138-d144.
- Dai, X. and Shen, L. 2022. Advances and trends in omics technology development. *Frontiers in Medicine*, Volume 9 - 2022.
- Duan, Y., Haybaeck, J. and Yang, Z. 2020. Therapeutic potential of PI3K/AKT/mTOR pathway in gastrointestinal stromal tumors: Rationale and progress. *Cancers (Basel)*, 12.
- Elfaki, I., Mir, R., Almutairi, F. M. and Duhier, F. M. A. 2018. Cytochrome P450: Polymorphisms and roles in cancer, diabetes and atherosclerosis. *Asian Pacific Journal of Cancer Prevention*, 19:2057-2070.
- Franz, M., Rodriguez, H., Lopes, C., Zuberi, K., Montojo, J., Bader, G. D. and Morris, Q. 2018. GeneMANIA update 2018. *Nucleic Acids Research*, 46:W60-w64.
- Giunta, S. 2021. Decoding human cancer with whole genome sequencing: a review of PCAWG Project studies published in February 2020. *Cancer and Metastasis Reviews*, 40:909-924.
- Guan, W. L., He, Y. and Xu, R. H. 2023. Gastric cancer treatment: recent progress and future perspectives. *Journal of Hematology & Oncology*, 16:57.
- Herrera-Kok, J. H., Parmar, C., Bangash, A. H., Samadov, E., Demirli Atici, S., Cheruvu, C. V., Abouelazayem, M., Yang, W., Galanis, M., Di Maggio, F., Isik, A., Bandopyadaya, S. and Viswanath, Y. K. 2023. Global impact of COVID-19 pandemic on gastric cancer patients. *European Journal of Surgical Oncology*, 49:876-877.
- Hosseini, M., Lotfi-Shahreza, M. and Nikpour, P. 2023. Integrative analysis of DNA methylation and gene expression through machine



- learning identifies stomach cancer diagnostic and prognostic biomarkers. *Journal of Cellular and Molecular Medicine*, 27:714-726.
- Iglesias González, P. A., Valdivieso, A. G. and Santa-Coloma, T. A. 2023. The G protein-coupled receptor GPRC5A-a phorbol ester and retinoic acid-induced orphan receptor with roles in cancer, inflammation, and immunity. *Biochemistry and Cell Biology*, 101:465-480.
- Jiang, X., Xu, X., Wu, M., Guan, Z., Su, X., Chen, S., Wang, H. and Teng, L. 2018. GPRC5A: An Emerging Biomarker in Human Cancer. *Biomed Research International*, 2018:1823726.
- Jin, Y., He, J., Du, J., Zhang, R. X., Yao, H. B. and Shao, Q. S. 2017. Overexpression of HS6ST2 is associated with poor prognosis in patients with gastric cancer. *Oncology Letters*, 14:6191-6197.
- Kizhakkeppurath Kumaran, A., Sahu, A., Singh, A., Aynikkattil Ravindran, N., Sekhar Chatterjee, N., Mathew, S. and Verma, S. 2023. Proteoglycans in breast cancer, identification and characterization by LC-MS/MS assisted proteomics approach: A review. *Proteomics Clinical Applications*, 17:e2200046.
- Le Rhun, E., Bertrand, N., Dumont, A., Tresch, E., Le Deley, M. C., Mailliez, A., Preusser, M., Weller, M., Revillion, F. and Bonnetterre, J. 2017. Identification of single nucleotide polymorphisms of the PI3K-AKT-mTOR pathway as a risk factor of central nervous system metastasis in metastatic breast cancer. *European Journal of Cancer*, 87:189-198.
- Li, H., Liu, J. W., Liu, S., Yuan, Y. and Sun, L. P. 2017. Bioinformatics-based identification of methylated-differentially expressed genes and related pathways in gastric cancer. *Digestive Diseases and Sciences* 62:3029-3039.
- Li, W., Dong, X., Wan, Z., Wang, W., Zhang, J., Mi, Y., Li, R., Xu, Z., Wang, B., Li, N. and He, G. 2024. PXMP4 promotes gastric cancer cell epithelial-mesenchymal transition via the PI3K/AKT signaling pathway. *Molecular Biology Reports* 51:350.
- Li, W. Q., Hu, N., Burton, V. H., Yang, H. H., Su, H., Conway, C. M., Wang, L., Wang, C., Ding, T., Xu, Y., Giffen, C., Abnet, C. C., Goldstein, A. M., Hewitt, S. M. and Taylor, P. R. 2014. PLCE1 mRNA and protein expression and survival of patients with esophageal squamous cell carcinoma and gastric adenocarcinoma. *Cancer Epidemiology, Biomarkers & Prevention*, 23:1579-1588.
- Liu, H., Qu, Y., Zhou, H., Zheng, Z., Zhao, J. and Zhang, J. 2021a. Bioinformatic analysis of potential hub genes in gastric adenocarcinoma. *Science Progress*, 104:368504211004260.
- Liu, H., Xiang, Y., Zong, Q. B., Zhang, X. Y., Wang, Z. W., Fang, S. Q., Zhang, T. C. and Liao, X. H. 2021b. miR-6745-TIMP1 axis inhibits cell growth and metastasis in gastric cancer. *Aging (Albany NY)*, 13:24402-24416.
- Ma, X., Huang, T., Li, X., Zhou, X., Pan, H., Du, A., Zeng, Y., Yuan, K. and



- Wang, Z. 2024. Exploration of the link between COVID-19 and gastric cancer from the perspective of bioinformatics and systems biology. *Frontiers in Medicine*, 11:1428973.
- Moreira, A. M., Ferreira, R. M., Carneiro, P., Figueiredo, J., Osório, H., Barbosa, J., Preto, J., Pinto-do-Ó, P., Carneiro, F. and Seruca, R. 2022. Proteomic identification of a gastric tumor ECM signature associated with cancer progression. *Frontiers in Molecular Biosciences*, 9:818552.
- Necula, L., Matei, L., Dragu, D., Pitica, I., Neagu, A. I., Bleotu, C., Dima, S., Popescu, I., Diaconu, C. C. and Chivu-Economescu, M. 2020. High plasma levels of COL10A1 are associated with advanced tumor stage in gastric cancer patients. *World Journal of Gastroenterology*, 26:3024-3033.
- Niland, S. and Eble, J. A. 2019. Neuropilins in the context of tumor vasculature. *International Journal of Molecular Sciences*, 20:639.
- Pan, J.-B., Hu, S.-C., Shi, D., Cai, M.-C., Li, Y.-B., Zou, Q. and Ji, Z.-L. 2013. PaGenBase: A pattern gene database for the global and dynamic understanding of gene function. *PLoS One*, 8:e80747.
- Piñero, J., Bravo, À., Queralt-Rosinach, N., Gutiérrez-Sacristán, A., Deu-Pons, J., Centeno, E., García-García, J., Sanz, F. and Furlong, L. I. 2016. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Research*, 45:D833-D839.
- Pinto, F., Santos-Ferreira, L., Pinto, M. T., Gomes, C. and Reis, C. A. 2021. The extracellular small leucine-rich proteoglycan biglycan is a key player in gastric cancer aggressiveness. *Cancers (Basel)*, 13.
- Prabhu, K. S., Sadida, H. Q., Kuttikrishnan, S., Junejo, K., Bhat, A. A. and Uddin, S. 2024. Beyond genetics: Exploring the role of epigenetic alterations in breast cancer. *Pathology - Research and Practice*, 254:155174.
- Puliga, E., Corso, S., Pietrantonio, F. and Giordano, S. 2021. Microsatellite instability in gastric cancer: Between lights and shadows. *Cancer Treat Reviews*, 95:102175.
- Ramírez-Ricardo, J., Leal-Orta, E., Martínez-Baeza, E., Ortiz-Mendoza, C., Breton-Mora, F., Herrera-Torres, A., Elizalde-Acosta, I., Cortes-Reynosa, P., Thompson-Bonilla, R. and Perez Salazar, E. 2020. Circulating extracellular vesicles from patients with breast cancer enhance migration and invasion via a Src-dependent pathway in MDA-MB-231 breast cancer cells. *Molecular Medicine Reports*, 22:1932-1948.
- Rokka, A., Antonenkov, V. D., Soininen, R., Immonen, H. L., Pirilä, P. L., Bergmann, U., Sormunen, R. T., Weckström, M., Benz, R. and Hiltunen, J. K. 2009. Pxm2 is a channel-forming protein in Mammalian peroxisomal membrane. *PLoS One*, 4:e5090.
- Rugge, M., Genta, R. M., Fassan, M., Valentini, E., Coati, I., Guzzinati, S., Savarino, E., Zorzi, M., Farinati, F. and Malfertheiner, P. 2018. OLGA gastritis staging for the



- prediction of gastric cancer risk: A long-term follow-up study of 7436 patients. *The American Journal of Gastroenterology*, 113:1621-1628.
- Sang, Y., Kong, P., Zhang, S., Zhang, L., Cao, Y., Duan, X., Sun, T., Tao, Z. and Liu, W. 2020. SGK1 in human cancer: Emerging roles and mechanisms. *Frontiers in Oncology*, 10:608722.
- Sexton, R. E., Al Hallak, M. N., Diab, M. and Azmi, A. S. 2020. Gastric cancer: a comprehensive review of current and future treatment strategies. *Cancer and Metastasis Reviews*, 39:1179-1203.
- Sha, R., Zhang, J., Meng, F. and Zhaori, G. 2023. Gastric cancer metastasis-related NT5DC2 indicates unfavorable prognosis of patients. *Medicine (Baltimore)*, 102:e35030.
- Shao, L., Ming, Y. and Zhou, Z. 2020. Identification of dysregulated long noncoding RNA and associated mechanism in gastric cancer. *Research Square*:PPR170470
- Shen, N., Zhu, S., Zhang, Z. and Yong, X. 2022. High expression of COL10A1 is an independent predictive poor prognostic biomarker and associated with immune infiltration in advanced gastric cancer microenvironment. *Journal of Oncology*, 2022:1463316.
- Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. and Ideker, T. 2011. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, 27:431-432.
- Smyth, E. C., Nilsson, M., Grabsch, H. I., van Grieken, N. C. T. and Lordick, F. 2020. Gastric cancer. *The Lancet*, 396:635-648.
- Tang, Z., Li, C., Kang, B., Gao, G., Li, C. and Zhang, Z. 2017. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Research*, 45:W98-w102.
- Thrift, A. P., Wenker, T. N. and El-Serag, H. B. 2023. Global burden of gastric cancer: epidemiological trends, risk factors, screening and prevention. *Nature Reviews Clinical Oncology*, 20:338-349.
- Tong, D., Heinze, G., Pils, D., Wolf, A., Singer, C. F., Concin, N., Hofstetter, G., Schiebel, I., Rudas, M. and Zeillinger, R. 2010. Gene expression of PMP22 is an independent prognostic factor for disease-free and overall survival in breast cancer patients. *BMC Cancer*, 10:682.
- Wang, G., Hu, N., Yang, H. H., Wang, L., Su, H., Wang, C., Clifford, R., Dawsey, E. M., Li, J. M., Ding, T., Han, X. Y., Giffen, C., Goldstein, A. M., Taylor, P. R. and Lee, M. P. 2013. Comparison of global gene expression of gastric cardia and noncardia cancers from a high-risk population in china. *PLoS One*, 8:e63826.
- Wang, M. R., Chen, R. J., Zhao, F., Zhang, H. H., Bi, Q. Y., Zhang, Y. N., Zhang, Y. Q., Wu, Z. C. and Ji, X. M. 2020. Effect of Wenxia Changfu formula combined with cisplatin reversing non-small cell lung cancer cell adhesion-mediated drug resistance. *Fronts in Pharmacology*, 11:500137.
- Wang, X., Zhao, J., Marostica, E., Yuan, W., Jin, J., Zhang, J., Li, R., Tang, H., Wang, K., Li, Y., Wang, F., Peng, Y., Zhu, J., Zhang, J., Jackson, C.



- R., Zhang, J., Dillon, D., Lin, N. U., Sholl, L., Denize, T., Meredith, D., Ligon, K. L., Signoretti, S., Ogino, S., Golden, J. A., Nasrallah, M. P., Han, X., Yang, S. and Yu, K.-H. 2024. A pathology foundation model for cancer diagnosis and prognosis prediction. *Nature*, 634:970-978.
- Wei, X., Liu, J., Hong, Z., Chen, X., Wang, K. and Cai, J. 2022. Identification of novel tumor microenvironment-associated genes in gastric cancer based on single-cell RNA-sequencing datasets. *Frontiers in Genetics*, Volume 13 - 2022.
- Wen, J., Deng, S., Bi, Y., Qiao, L. and Xu, H. 2023. Association between multiple gene promoter hypermethylation and the risk of gastric cancer: A systematic review and meta-analysis. *Digestive and Liver Disease*, 55:40-45.
- Wu, J., Mao, X., Cai, T., Luo, J. and Wei, L. 2006. KOBAS server: a web-based platform for automated annotation and pathway identification. *Nucleic Acids Research*, 34:W720-724.
- Yang, J., Liao, Q., Price, M., Moriarity, B., Wolf, N., Felices, M., Miller, J. S., Geller, M. A., Bendzick, L., Hopps, R., Starr, T. K., O'Connor, C. H., Tarullo, S., Nelson, A. C., Turley, E., Wang, J. and McCarthy, J. B. 2022. Chondroitin sulfate proteoglycan 4, a targetable oncoantigen that promotes ovarian cancer growth, invasion, cisplatin resistance and spheroid formation. *Translational Oncology*, 16:101318.
- Yang, L., Ying, X., Liu, S., Lyu, G., Xu, Z., Zhang, X., Li, H., Li, Q., Wang, N. and Ji, J. 2020. Gastric cancer: Epidemiology, risk factors and prevention strategies. *Chinese Journal of Cancer Research*, 32:695-704.
- Yao, Y., Jiang, Q., Jiang, L., Wu, J., Zhang, Q., Wang, J., Feng, H. and Zang, P. 2016. Lnc-SGK1 induced by *Helicobacter pylori* infection and highsalt diet promote Th2 and Th17 differentiation in human gastric cancer by SGK1/Jun B signaling. *Oncotarget*, 7:20549-20560.
- Yu, X., Sun, R., Yang, X., He, X., Guo, H. and Ou, C. 2023. The NT5DC family: expression profile and prognostic value in pancreatic adenocarcinoma. *Journal of Cancer*, 14:2274-2288.
- Zhang, P., Cao, X., Guan, M., Li, D., Xiang, H., Peng, Q., Zhou, Y., Weng, C., Fang, X., Liu, X., Mao, H., Li, Q., Liu, G. and Lu, L. 2022. CPNE8 promotes gastric cancer metastasis by modulating focal adhesion pathway and tumor microenvironment. *International Journal of Biological Sciences*, 18:4932-4949.
- Zhao, T., Chen, Z., Liu, W., Ju, H. and Li, F. 2023. Identification of hub genes associated with gastric cancer via bioinformatics analysis and validation studies. *International Journal of General Medicine*, 16:4835-4848.
- Zheng, M., Wang, P., Wang, Y., Jia, Z., Gao, J., Tan, X., Chen, H. and Zu, G. 2024. Clinicopathological and prognostic significance of TIMP1 expression in gastric cancer: a systematic review and meta-analysis. *Expert Review of Anticancer Therapy*, 24:1169-1176.



- Zhou, Y., Lin, Y., Li, W., Liu, Q., Gong, H., Li, Y. and Luo, D. 2023. Expression of AKRs superfamily and prognostic in human gastric cancer. *Medicine (Baltimore)*, 102:e33041.
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., Benner, C. and Chanda, S. K. 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications*, 10:1523.

