RESEARCH





Network and functional analyses of differentially expressed genes in gastric cancer provide new biomarkers associated with disease pathogenesis

Mousa Fadaei¹, Maryam Kohansal^{2,3}, Omidreza Akbarpour⁴, Mahsa Sami¹ and Ali Ghanbariasad^{1,2*}

Abstract

Background Gastric cancer is a dominant source of cancer-related death around the globe and a serious threat to human health. However, there are very few practical diagnostic approaches and biomarkers for the treatment of this complex disease.

Methods This study aimed to evaluate the association between differentially expressed genes (DEGs), which may function as potential biomarkers, and the diagnosis and treatment of gastric cancer (GC). We constructed a protein-protein interaction network from DEGs followed by network clustering. Members of the two most extensive modules went under the enrichment analysis. We introduced a number of hub genes and gene families playing essential roles in oncogenic pathways and the pathogenesis of gastric cancer. Enriched terms for Biological Process were obtained from the "GO" repository.

Results A total of 307 DEGs were identified between GC and their corresponding normal adjacent tissue samples in GSE63089 datasets, including 261 upregulated and 261 downregulated genes. The top five hub genes in the PPI network were CDK1, CCNB1, CCNA2, CDC20, and PBK. They are involved in focal adhesion formation, extracellular matrix remodeling, cell migration, survival signals, and cell proliferation. No significant survival result was found for these hub genes.

Conclusions Using comprehensive analysis and bioinformatics methods, important key pathways and pivotal genes related to GC progression were identified, potentially informing further studies and new therapeutic targets for GC treatment.

Keywords Gastric cancer, Gene expression, Signaling networks, Biomarkers

*Correspondence:

Ali Ghanbariasad

alighanbariasad@gmail.com

¹ Noncommunicable Diseases Research Center, Fasa University of Medical Sciences, Fasa, Iran

² Department of Medical Biotechnology, Fasa University of Medical Sciences, Fasa, Iran

³ Department of Biology, Payame Noor University, Tehran, Iran

⁴ Genetic Laboratory, Fasa University of Medical Sciences, Fasa, Iran

Background

Gastric cancer is a common form of cancer with the second-highest cancer-related mortality rate [1]. More than 90% of gastric tumors are adenocarcinoma, and there is a poor prognosis for that. Early stages of the disease are often silent, so late diagnosis results in a low survival rate [2]. In spite of the improvement in the diagnosis of gastric tumors and the development of new molecular targeted drugs, there is still a lack of diagnostic biomarkers and effective treatments [3, 4].



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.



Fig. 1 Illustration of outlier samples in the PCA plot. PC1 is the eigenvector one, and PC2 is the eigenvector two. Normal samples are blue with the suffix N, while tumor samples are red with the suffix C. one of the apparent outlier samples is GSM1540941_N in the top right corner

In recent years, the development of bioinformatics methods and tools has made significant progress at determining the molecular pathogenesis of many carcinomas and adenocarcinomas such as breast and gastric cancer [5, 6]. One way to identify biomarkers in a biological context is by analyzing gene expression transcriptomic data [7]. Expression of many genes is deregulated once cells start to transform toward a cancerous phenotype, and this is different from cell to cell and tissue to tissue [8, 9]. As a result, analyzing the differentially expressed genes (DEGs) in a specific biological context like gastric cancer enables us to find potential diagnostic biomarkers and therapeutic genetic targets. For instance, CXCL1, SPARC, SPP1, and SULF1 are the genes overexpressed together in gastric cancer [10]. In another study, they proposed seventeen genes differentially expressed between gastric cancer samples and paired normal samples responsible for tumorigenesis [11]. Upregulation of CCNE1 and downregulation of NR3C1 is an indicator of primary GC tumor, while downregulation of NR4A2 and upregulation of HSP90AA1 are promising markers of liver metastasis [12].

In the present study, a gene expression microarray dataset with GSE63089 accession ID was downloaded from Gene Expression Omnibus (GEO) database. There were two groups of normal and gastric cancer samples in the dataset statistically compared to identify genes different in expression between the two groups. The goal was to recognize new biomarkers among a great extent of DEGs with desired thresholds by the tools of network analysis and gene set enrichment analysis. Several hub genes and gene families engaged in biological processes and signaling pathways related to cancer progression were introduced, and their molecular mechanism leading to GC progression was explained.

Results

Data preprocessing

The gastric cancer gene expression dataset was imported into R using "getGEO" function in "GEOquery" R package. Data were visualized using PCA and boxplots to recognize biased samples. Figure 1 illustrates the PCA plot for all samples before outlier sample removal. Two clusters appeared in the PCA plane segregated well based on the group definition. A number of samples were located at a distance from their cluster set regarded as the outlier or biased samples. Figure 2A depicts the sample boxplots before outlier removal. Samples with extreme IQRs demonstrated the presence of batch effects in the dataset, and they were removed as well. As a result, outlier samples could no longer impact the downstream processing steps. Thirteen outliers were detected, presented in



Fig. 2 Boxplots before (A) and after normalization (B). Sample names with the suffix N are normal, and the ones with the suffix C are cancerous. Values were log2 transformed. Quantile normalization equalized the percentiles of all samples

Supplementary file 1. Next, data were normalized using "normalize.quantiles" function in "preprocessCore" R package. Figure 2B depicts sample boxplots after data normalization.

Identifying differentially expressed genes

The cancer group was compared to the normal group in the dataset. There were 42 samples in the cancer group and 35 samples in the normal group. Genes with absolute log fold change (LogFC) larger than one and adjusted *p*-value less than 0.05 were regarded as DEGs. 261 upregulated genes and 216 downregulated genes were recognized. The mentioned metrics for all DEGs are presented in Supplementary file 2. SFRP4, CDH17, FAP, CLDN1, and OLFM4 were of the highest LogFC values among the upregulated genes, while GIF, PGA4, GKN1, ATP4B, and CPA2 had the lowest LogFC values among the downregulated genes.

Undirected protein-protein interaction network

All DEGs were given to the STRING database to construct the protein-protein interaction (PPI) network. Three sources of evidence were used to predict more valid interactions between genes, namely Experiment, Database, and Co-expression. 307 DEGs were identified to have interaction with at least one gene (protein) that participated in the network configuration. The giant component of this network with 268 nodes and 3582 edges is illustrated in Fig. 3. At one glance, there is a cluster of genes on the right-hand side of the network that may be responsible for one or some specific biological functions. Therefore, we decided to apply cluster analysis



Fig. 3 The whole network giant component. Labels are protein/gene symbols. This is a scale-free network [13] that follows a power-law distribution (most network nodes have a low degree while there are few nodes with a high degree)

on the network and separately carry out the enrichment analysis on each cluster.

Network descriptive and hub genes

The network diameter was Eleven containing AKR1C1, AKR1C3, CYP3A5, GSTA, GGH, CXCL,1AGT, CFTR, HKDC1, ALDOB, SULT1C2, and ACSM3 genes. Transitivity was around 80%, edge density was about 10% and the mean distance was 4.2. Two important centralities, degree and betweenness, and regulation status for some genes with high centrality are presented in Table 1. CDK1 had the highest value for both degree and betweenness centralities upregulated in our analysis. HGF had a high betweenness but a low degree. Contrary, CCNB1 and CCNA2 had a high degree but low betweenness. Other descriptives such as edge-betweenness for edges and closeness and average distances for each node are provided in Supplementary file 3.

Network clustering and enrichment analysis

Network giant component was clustered using the "Fast Greedy" algorithm in "igraph" R package. Six clusters emerged and Gene Set Enrichment Analysis (GSEA) was performed on the two largest groups. Gene sets were given to the "Enrichr" online tool. Figure 4A and B depict the enrichment results for the first module. Enrichment results for cluster 1 are presented in Fig. 4 while results for cluster 2 are presented in Supplementary file 4. In Fig. 4A, the first two genes, MMP1 and MMP3, were related to the important terms for extracellular matrix (ECM) organization and degradation. Moreover, the CXCL gene family owned a number of terms associated with cytokine- and chemokine-mediated leukocyte migration. Collagen gene family was associated with ECM collagen fibers organization, but MMP1 and MMP3 genes were enriched for ECM remodeling proposing that extracellular matrix in gastric tumor samples has been altered probably in favor of the tumor. The last genes were bolding some Biological processes related to

Table 1 The network hub genes. DEGs were sorted based on the highest degree. Status column exhibits whether a gene is upregulated or downregulated in the analysis

De genes	Degree	Betweenness	Status
cdk1	102	6366.23	Up
CCNB1	94	324.32	Up
CCNA2	93	476.87	Up
CDC20	89	1735.14	Up
PBK	87	2489.09	Up
CHEK1	80	1230.53	Up
ZWINT	79	2548.27	Up
CDKN3	72	1791.96	Up
HIST1H2BB	25	747.8	Up
HIST1H2BM	25	747.8	Up
COL1A2	21	2108.26	Up
COL1A1	19	1422	Up
COL3A1	18	1345.35	Up
CXCL1	18	4174.93	Up
ORM2	15	3172.16	Down
EPHB2	14	2414.36	Up
AGT	14	1555.01	Up
FGA	12	1101.23	Down
SPARC	12	4279.68	Up
GSTA1	12	2290.52	Up
GSTA3	12	2290.52	Down
ITGA2	11	2797.82	Up
HGF	11	6349.3	Down
ZBTB16	7	3430.08	Down
RCN1	5	2436.24	Up
CFTR	5	3687.92	Up

neutrophils activation and immune system. In cluster 2, "tubulin-binding" had the highest *p*-value among the biological processes. In addition, Kinesins (KIF) gene family was enriched for several terms associated with microtubules activity and organization as well as Kinase and ATPase activities. Furthermore, "DNA-dependent ATPase activity" and "DNA binding" terms were related to the same genes.

In B, some crucial terms for focal adhesion arrangement and PI3K-AKT signaling pathways were associated with many genes such as collagens, integrins, HGF, and so on. Moreover, matrix metalloproteinase-linked enriched pathways were associated with MMP1 and MMP3, similar to GO biological process terms. As a result, the role of obtained terms in gastric cancer should be interrogated. Some genes were involved in retinoblastoma protein signaling in cluster 2. A number of enriched genes encompassed terms related to DNA damage and repair, proposing that DNA damage might have occurred in tumor cells. RFC3, TOP2A, and FANCI were the genes linked to the signaling pathways engaged in Gastric Cancer Network 1 and 2.

Gene expression validation

Survival analysis was performed on the enriched genes in Fig. 5. To verify the expression and impact of the genes on survival rate. Collagen gene families were the dominant protein family in the enrichment analysis, some of which had a significant *p*-value, < 0.05. Patients with higher expression levels of collagens in Fig. 4, had lower daily survival rates presented in Supplementary file 5. However, the results for the hub genes were not significant therefore, we only validated them using expression profiles in TCGA and GTEx datasets. Hence, upregulation of all the top five hub genes in Table 1 was verified in Fig. 5.

Discussion

Gastric cancer is still the fourth most common cancer globally and ranks as the second leading cause of cancer death alongside lung cancer [14]. GC is a complex disease affected by many environmental and genetic factors. Despite the increase in knowledge and advances in drug development, treatment still performs poorly due to late diagnosis and extremely high heterogeneity within the tumor and among patients. Therefore, there is an urgent need to identify more specific and sensitive biomarkers to clarify this complex disease's pathogenesis. These include the telomerase reverse transcriptase promoter region (TP53, BRAF, and RAS), DNA hypermethylation, and other gene mutations that can be used to explore the pathogenesis of gastric cancer.

In our study, several hub genes as well as outcomes for functional analysis went under investigation. As a result of tumor heterogeneity, gene expression profiles (GSE) emerged from multiple modules from protein-protein interaction information. The largest (first) module enriched mostly for the immune system, tumor occurrence, and progression. Components of this cluster are engaged in focal adhesion formation, ECM remodeling, and cell migration. Some common cancer signaling pathways such as PI3K-Akt, chemokine-mediated signaling pathway, ECM-receptor interaction, and so on emerged [15–17]. The ECM-receptor interaction pathway also plays a vital role in the proliferation, differentiation, and metastasis of cancer cells [18]. ECM can promote cancer metastasis by inducing epithelial-mesenchymal transition (EMT) of tumor cells [19]. integrins, collagens, and matrix metalloproteinases were among the important DEGs responsible for de novo remodeling of ECM and the stiffness pertinent for EMT mechanism and cell migration [20-22]. Integrin signals bridge between cell ECM and cell differentiation [23]. Survival analysis shows



Fig. 4 Gene Set Enrichment Analysis for cluster 1 of the PPI network. Part A illustrates the enriched terms for biological process from GO repository. Part B shows the enriched pathways from WikiPathways signaling datasets. Red cells reveal the involvement of the genes in the enriched elements. Enriched terms were sorted based on the highest *p*-value

that all collagens enriched in our study play a pivotal role in integrin signaling since they are against the survival rate of gastric cancer patients. The majority of them even had a significant *p*-value demonstrating the significance of collagens upregulation in GC progression. These collagens were enriched in the PI3K-Akt signaling pathway engaged in focal adhesion formation. Furthermore, The PI3K/Akt survival signals regulate gene expression and cell metabolism, and the reaction products catalyzed by PI3K essential for adhesion, survival, cytoskeleton rearrangement, and vesicle transport [24]. The control of cell growth by the PI3K/Akt pathway via regulating cell proliferation, cell cycle progression, and apoptosis implicates a crucial role of this pathway in carcinogenesis and cancer development [24, 25]. Therefore, regulating ECM and PI3K1/AKT signaling pathways appear to be promising treatment strategies [24]. There were some CXC chemokine ligands (CXCL) in cluster 1 as well. They have important roles in lymphocyte trafficking, particularly the attraction of leukocytes to tumor sites, induction of apoptosis, regulation of cell growth, and mediation of angiogenesis [26, 27]. All the CXCL genes were upregulated in our analysis increasing ligands for chemokine receptors that triggers cell migration of tumor cells [27]. CXCL1 was among the hub genes and possessed

a relatively significant p-value in survival analysis. The higher this gene is expressed, the survival rates is reduced in GC patients.

The top five hub genes in the PPI network CDK1, CCNB1, CCNA2, CDC20, and PBK, exhibited unique expression Patterns. These DEGs and related functions would be related to the progression of gastric cancer. However, no significant survival result was found for these hub genes. We validated them by expression profiling of these genes in TCGA and GTEx genomic repositories rather than the former analysis.

Cdk1 has a catalytic subunit that promotes the M-phase process essential for the G1/S and G2/M phase transitions during cell proliferation [28, 29]. Besides, cyclin B-Cdk1 is involved in cell survival at mitotic checkpoints called spindle checkpoints [30, 31]. In the past ten years, a large number of studies have shown that the disorder of CDK1 not only leads to the rapid growth of tumors but also leads to the spontaneous proliferation of cancer cells [32, 33]. Prior studies have indicated that aberrant activation of CDKs and their modulators exist in many tumors [34]. Neganova et al. reported that CDK1 is a critical element of cell cycle regulation, especially mitosis, and plays a vital role in maintaining the pluripotency and genome stability of human pluripotent stem





Fig. 5 Expression profiling of the hub genes. Data were firstly TPM normalized and then log2 transformed. The median for all the hub genes in gastric tumor samples was larger than in normal samples

cells [35]. In addition, CDK1 can activate the JAK/STAT3 signaling pathway through the phosphorylation of JAK1, thereby promoting the progression of colorectal cancer (CRC) [36].

Cancer is characterized by cell cycle dysregulation. CCNB1, Cycline B1, is one of the central genes in the first module of the PPI network whose product promotes the transition of cells from the G2 to M phase [37]. CCNB1 depletion or stable gene silencing of CCNB1 can inhibit human tumor cell proliferation and induce apoptosis [38]. Related studies have shown that the overexpression of CCNB1 is connected to the clinical stage, lymph node metastasis, and low survival rate of GC patients [39]. Furthermore, CDK1 and CCNB1 are highly expressed in neuroblastoma (NB) cell lines enhancing their proliferation [40]. The important role of the CDK1/CCNB1 complex in tumor cell survival was confirmed [34]. Finally, A recent study reported that cyclin B1 and cyclin B2 are the most important candidate biomarkers in GC [41].

Like CCNB1, Cyclin A2 (CCNA2) is another hub gene a member of the cyclin family that functions as a regulator of cyclin-dependent kinases (CDKs) affected by KRAS gene mutations. It is also a predictive biomarker of gastric cancer associated with Polo-like kinase 1 (PLK1), a mitotic serine/threonine-protein kinase [42]. Cyclin B1 and Cyclin A2 form a complex with CDK1 to trigger G2/M transition [43]. Upregulation of CCNA2 is found in numerous types of cancer, including pancreatic ductal adenocarcinoma and colorectal cancers [44, 45]. CCNA2 has the potential to be a new diagnostic biomarker and cancer therapy target that aims to monitor the efficacy of breast cancer chemotherapy [46]. Furthermore, these two genes were present in cluster 2 of the PPI network enriched alongside Kinesin gene family (KIF). KIF proteins are the hub proteins in the intracellular transport system by transport of cellular cargo. Mitotic spindle kinesins are essential for cell division. Therefore, genes in cluster 2 were engaged mostly in cellular proliferation, such as TOP2A and RFC3. Furthermore, many non-mitotic kinesins are associated with tumorigenesis and anticancer drug resistance [47].

Cell division cycle 20 homolog (CDC20) is an oncogenic gene that has long been recognized as one of the significant regulatory components of the cell cycle [48, 49]. Its product forms a complex with Anaphase Promoting Complex (APC) necessary for spindle assembly and chromosome segregation [50]. Overexpression of CDC20 has been reported in various malignant tumors [51]. It has been proposed to be a promising therapeutic target for cancer treatment as well [50]. The high expression of CDC20 is associated with increased tumor grade and stage in the majority of common carcinoma and adenocarcinoma [52]. P53 negatively regulates CDC20 expression, and silencing of CDC20 significantly inhibits cell growth in vitro [53]. Inactivation of p53 has been observed in various cancer tissues, including acute myeloid leukemia and lung cancer [54– 56]. This effect might be attributed to CDC20 upregulation in gastric cancer. APC/CDC20 complex can suppress apoptosis by targeting the apoptotic protein BIM for destruction and ubiquitination [57, 58].

PBK is a serine/threonine-protein kinase related to the mitogen-activated protein kinase kinase (MAPKK) family. Overexpression of this gene has been implicated in tumorigenesis [59, 60]. PBK gene is upregulated in various types of cancers and tumors such as bladder cancer, brain tumor, breast cancer, hepatocarcinoma, lung cancer, and gastric cancer [61–65]. PBK is also located in the nucleus and is involved in the phosphorylation of histone H3 and the inhibition of p53 in colorectal and breast cancer cells in different cancers [66]. PBK is connected to geranylgeranylation signaling, most likely in advanced-stage cancers [67], which is essential for cancer cell proliferation, confirming that PBK is an important molecular target for cancer therapy.

Conclusion

In conclusion, chemokine ligands, integrins, collagens, and matrix metalloproteinases, and hub genes including CDK1, CCNB1, CCNA2, CDC20, and PBK have been identified to be associated with GC progression. Through GO and pathway enrichment of the two most extensive modules, we identified the functions and pathways of the hub genes as well as genes responsible for cell differentiation and migration in cluster 1 and the ones responsible for cell cycle progression in cluster 2. In vitro studies are further required to test the functional results that pave a prospective way towards gastric cancer treatment.

Methods

Database searching to find a suitable experiment

Gene Expression Omnibus (http://www.ncbi.nlm.nih. gov/geo/) database was searched to detect an experiment containing high-quality transcriptomic samples in concordance to the study design. Searches were filtered for Homo sapiens, while gastric cancer and metastasis were the search keywords. Microarray raw data with accession numbers GSE63089 was selected and the gene expression matrix was downloaded from GEO using "GEOquery" R package version 2.5 [68]. The dataset contained 45 normal samples and 45 samples from gastric cancer patients. All the sample IDs are listed in Supplementary file 1.

Identifying differentially expressed genes

Outlier samples were identified and removed using the PCA method. Next, data were normalized using the quantile normalization method [69]. GPL5175 annotation dataset was downloaded from GEO, and annotation was carried out by mapping probesets to the gene symbols. "Limma" R package, which applies linear models on the expression matrix, was utilized to discover DEGs between three groups of samples [70]. Genes with absolute log fold change larger than 1, and Benjamini Hochberg adjusted *p*-value [71] less than 0.05 were selected as the DEGs.

Network construction

STRING database was used to generate the Interactions between all DEGs according to five sources of evidence, namely Experiments, Databases, Co-expression, Gene fusion, and Co-occurrence. Using "igraph" package version 1.2.4 in R software [72], the giant component of the network was extracted from the whole network. Next, different network descriptive and centralities were computed employing the same package.

Enrichment analysis

Enrichment analysis was performed using the "Enrichr" online tool [73]. Enriched terms for Biological Process were obtained from the "GO" repository. For pathway enrichment analysis, the "wikiPathways" signaling repository version 2019 for humans was used. Enriched terms with the top score and a *p*-value less than 0.05 were selected.

Survival analysis and expression profiling

Genes were given to the "GEPIA2" web server, and TCGA and GTEx gene expression data from stomach adenocarcinoma were employed for survival analysis and expression profiling [74, 75]. Kaplan-Meier estimate was used to perform survival analysis [76]. Overall survival was measured based on days with a 95% confidence interval, and the median was used as the cutoff for grouping. For expression profiling, the LogFC cut-off was set on one and *q*-value on 0.01. Data were Log2 transformed, and the "LIMMA" method was used for the statistical inference.

Abbreviations

DEGs	Differentially expressed genes
GC	Gastric cancer
PPI	Protein-protein interaction
GSEA	Gene Set Enrichment Analysis
ECM	Extracellular matrix
GSE	Gene expression profiles
EMT	Epithelial-mesenchymal transition
CXCL	CXC chemokine ligands
CCNA2	Cyclin A2
PLK1	Polo-like kinase 1
KIF	Kinesin gene family
CDC20	Cell division cycle 20 homolog
MAPKK	Mitogen-activated protein kinase kinase

Supplementary Information

The online version contains supplementary material available at https://doi. org/10.1186/s43046-023-00164-5.

Additional file 1.

Acknowledgements

Not applicable.

Authors' contributions

MF: conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing – original draft preparation, writing – review and editing; MK: writing – original draft preparation, writing – review and editing; OA: writing – original draft preparation, writing – review and editing; MS: review and editing; AG: supervision, validation, writing – review and editing, project administration. All authors have read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

Anybody can access the GEO database that uses in this study with the below link: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE63089.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 27 February 2022 Accepted: 13 February 2023 Published online: 10 April 2023

References

- 1. Hartgrink HH, et al. Gastric cancer. Lancet. 2009;374(9688):477-90.
- 2. Correa P. Gastric cancer: overview. Gastroenterol Clin N Am. 2013;42(2):211.

- Wu H-H, Lin W-C, Tsai K-W. Advances in molecular biomarkers for gastric cancer: miRNAs as emerging novel cancer markers. Expert Rev Mol Med. 2014;16:e1.
- 4. Ryu J-W, et al. The proteomics approach to find biomarkers in gastric cancer. J Korean Med Sci. 2003;18(4):505.
- Liu X, et al. Identification of potential key genes associated with the pathogenesis and prognosis of gastric cancer based on integrated bioinformatics analysis. Front Genet. 2018;9:265.
- Sun C, et al. Identification of core genes and outcome in gastric cancer using bioinformatics analysis. Oncotarget. 2017;8(41):70271.
- 7. Vecchi M, et al. Gene expression analysis of early and advanced gastric cancers. Oncogene. 2007;26(29):4284–94.
- Andor N, et al. Joint single cell DNA-seq and RNA-seq of gastric cancer cell lines reveals rules of in vitro evolution. NAR Genom Bioinform. 2020;2(2):lqaa016.
- Sathe A, et al. Single-Cell Genomic Characterization Reveals the Cellular Reprogramming of the Gastric Tumor Microenvironment. Clin Cancer Res. 2020;26(11):2640–53.
- Junnila S, et al. Gene expression analysis identifies over-expression of CXCL1, SPARC, SPP1, and SULF1 in gastric cancer. Genes Chromosom Cancer. 2010;49(1):28–39.
- 11. Rajkumar T, et al. Identification and validation of genes involved in gastric tumorigenesis. Cancer Cell Int. 2010;10(1):45.
- 12. Chang W, et al. Identification of novel hub genes associated with liver metastasis of gastric cancer. Int J Cancer. 2009;125(12):2844–53.
- Khanin R, Wit E. How scale-free are biological networks. J Comput Biol. 2006;13(3):810–8.
- 14. Bray F, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.
- Duan S, et al. Novel prognostic biomarkers of gastric cancer based on gene expression microarray: COL12A1, GSTA3, FGA and FGG. Mol Med Rep. 2018;18(4):3727–36.
- 16. Lindsey S, Langhans SA. Crosstalk of oncogenic signaling pathways during epithelial-mesenchymal transition. Front Oncol. 2014;4:358.
- 17. Sanchez-Vega F, et al. Oncogenic signaling pathways in the cancer genome atlas. Cell. 2018;173(2):321–37 e10.
- Bao Y, et al. Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer. Cell Mol Biol Lett. 2019;24(1):38.
- 19. Peixoto P, et al. EMT is associated with an epigenetic signature of ECM remodeling genes. Cell Death Dis. 2019;10(3):1–17.
- Frantz C, Stewart KM, Weaver VM. The extracellular matrix at a glance. J Cell Sci. 2010;123(24):4195–200.
- 21. Ehrbar M, et al. Elucidating the role of matrix stiffness in 3D cell migration and remodeling. Biophys J. 2011;100(2):284–93.
- 22. Harjanto D, Maffei JS, Zaman MH. Quantitative analysis of the effect of cancer invasiveness and collagen concentration on 3D matrix remodeling. PLoS One. 2011;6(9):e24891.
- 23. Canel M, et al. E-cadherin–integrin crosstalk in cancer invasion and metastasis. J Cell Sci. 2013;126(2):393–401.
- 24. Ye B, et al. Exression and PI3K/AKT pathway in gastric cancer and its blockade suppresses tumor growth and metastasis. Int J Immunopathol Pharmacol. 2012;25(3):627–36.
- Roy SK, Srivastava RK, Shankar S. Inhibition of PI3K/AKT and MAPK/ ERK pathways causes activation of FOXO transcription factor, leading to cell cycle arrest and apoptosis in pancreatic cancer. J Mol Signal. 2010;5(1):10.
- Strieter RM, et al. Cancer CXC chemokine networks and tumour angiogenesis. Eur J Cancer. 2006;42(6):768–78.
- 27. Vandercappellen J, Van Damme J, Struyf S. The role of CXC chemokines and their receptors in cancer. Cancer Lett. 2008;267(2):226–44.
- Herman-Antosiewicz A, et al. Diallyl trisulfide-induced G2/M phase cell cycle arrest in DU145 cells is associated with delayed nuclear translocation of cyclin-dependent kinase 1. Pharm Res. 2010;27(6):1072–9.
- Suryadinata R, Sadowski M, Sarcevic B. Control of cell cycle progression by phosphorylation of cyclin-dependent kinase (CDK) substrates. Biosci Rep. 2010;30(4):243–55.
- Hayward D, Alfonso-Pérez T, Gruneberg U. Orchestration of the spindle assembly checkpoint by CDK1-cyclin B1. FEBS Lett. 2019;593(20):2889–907.

- Serpico AF, Grieco D. Recent advances in understanding the role of Cdk1 in the Spindle Assembly Checkpoint. F1000Research. 2020;9:F1000 Faculty Rev-57.
- Liu P, Kao T, Huang H. CDK1 promotes cell proliferation and survival via phosphorylation and inhibition of FOXO1 transcription factor. Oncogene. 2008;27(34):4733–44.
- Linares JF, et al. Phosphorylation of p62 by cdk1 controls the timely transit of cells through mitosis and tumor cell proliferation. Mol Cell Biol. 2011;31(1):105–17.
- Zhang L, et al. LncRNA CASC11 promoted gastric cancer cell proliferation, migration and invasion in vitro by regulating cell cycle pathway. Cell Cycle. 2018;17(15):1886–900.
- Neganova I, et al. CDK1 plays an important role in the maintenance of pluripotency and genomic stability in human pluripotent stem cells. Cell Death Dis. 2014;5(11):e1508.
- 36. Kuang Y, et al. Iron-dependent CDK1 activity promotes lung carcinogenesis via activation of the GP130/STAT3 signaling pathway. Cell Death Dis. 2019;10(4):1–12.
- Huang Y, Sramkoski RM, Jacobberger JW. The kinetics of G2 and M transitions regulated by B cyclins. PLoS One. 2013;8(12):e80861.
- Akimov I, Chernolovskaya E. Silencing of the CCNB1, Her2, and PKC genes by small interfering RNA differently retards the division of different human cancer cell lines. Mol Biol. 2010;44(1):89–96.
- 39. Zhang H-P, et al. Clinical significance and biological roles of cyclins in gastric cancer. OncoTargets Ther. 2018;11:6673.
- Schwermer M, et al. Sensitivity to cdk1-inhibition is modulated by p53 status in preclinical models of embryonal tumors. Oncotarget. 2015;6(17):15425.
- Shi Q, et al. ISL1, a novel regulator of CCNB1, CCNB2 and c-MYC genes, promotes gastric cancer cell proliferation and tumor growth. Oncotarget. 2016;7(24):36489.
- 42. Lee Y, et al. Pharmacogenomic Analysis Reveals CCNA2 as a Predictive Biomarker of Sensitivity to Polo-Like Kinase I Inhibitor in Gastric Cancer. Cancers. 2020;12(6):1418.
- 43. Doan P, et al. Alkylaminophenol induces G1/S phase cell cycle arrest in glioblastoma cells through p53 and cyclin-dependent kinase signaling pathway. Front Pharmacol. 2019;10:330.
- 44. Gan Y, et al. CCNA2 acts as a novel biomarker in regulating the growth and apoptosis of colorectal cancer. Cancer Manag Res. 2018;10:5113.
- Dong S, et al. Overexpression of BUB1B, CCNA2, CDC20, and CDK1 in tumor tissues predicts poor survival in pancreatic ductal adenocarcinoma. Biosci Rep. 2019;39(2):BSR20182306.
- Gao T, et al. CCNA2 is a prognostic biomarker for ER+ breast cancer and tamoxifen resistance. PLoS One. 2014;9(3):e91771.
- 47. Rath O, Kozielski F. Kinesins and cancer. Nat Rev Cancer. 2012;12(8):527–39.
- Kidokoro T, et al. CDC20, a potential cancer therapeutic target, is negatively regulated by p53. Oncogene. 2008;27(11):1562–71.
- Ding Z-Y, et al. Expression characteristics of CDC20 in gastric cancer and its correlation with poor prognosis. Int J Clin Exp Pathol. 2014;7(2):722.
- Wang L, et al. Targeting Cdc20 as a novel cancer therapeutic strategy. Pharmacol Ther. 2015;151:141–51.
 Kapapidau M. Cuttic NI. Balapage Cascia VM. Cdc20, at the superconduct
- Kapanidou M, Curtis NL, Bolanos-Garcia VM. Cdc20: at the crossroads between chromosome segregation and mitotic exit. Trends Biochem Sci. 2017;42(3):193–205.
- 52. Gayyed MF, et al. A comprehensive analysis of CDC20 overexpression in common malignant tumors from multiple organs: its correlation with tumor grade and stage. Tumor Biol. 2016;37(1):749–62.
- Banerjee T, Nath S, Roychoudhury S. DNA damage induced p53 downregulates Cdc20 by direct binding to its promoter causing chromatin remodeling. Nucleic Acids Res. 2009;37(8):2688–98.
- 54. Herrero AB, et al. Molecular mechanisms of p53 deregulation in cancer: an overview in multiple myeloma. Int J Mol Sci. 2016;17(12):2003.
- 55. Dragoj M, et al. Anti-invasive effects of CXCR4 and FAK inhibitors in non-small cell lung carcinomas with mutually inactivated p53 and PTEN tumor suppressors. Investig New Drugs. 2017;35(6):718–32.
- 56. Yang D, et al. HUWE1 controls the development of non-small cell lung cancer through down-regulation of p53. Theranostics. 2018;8(13):3517.
- 57. Gao Y, et al. Cdc20 inhibitor apcin inhibits the growth and invasion of osteosarcoma cells. Oncol Rep. 2018;40(2):841–8.

- Lub S, et al. Inhibiting the anaphase promoting complex/cyclosome induces a metaphase arrest and cell death in multiple myeloma cells. Oncotarget. 2016;7(4):4062.
- Fukukawa C, et al. Critical roles of LGN/GPSM2 phosphorylation by PBK/ TOPK in cell division of breast cancer cells. Genes Chromosom Cancer. 2010;49(10):861–72.
- Shih M, et al. TOPK/PBK promotes cell migration via modulation of the PI3K/PTEN/AKT pathway and is associated with poor prognosis in lung cancer. Oncogene. 2012;31(19):2389–400.
- Ohashi T, et al. Overexpression of PBK/TOPK contributes to tumor development and poor outcome of esophageal squamous cell carcinoma. Anticancer Res. 2016;36(12):6457–66.
- 62. Dong C, Fan W, Fang S. PBK as a Potential Biomarker Associated with Prognosis of Glioblastoma. J Mol Neurosci. 2020;70(1):56–64.
- Ohashi T, et al. Overexpression of PBK/TOPK relates to tumour malignant potential and poor outcome of gastric carcinoma. Br J Cancer. 2017;116(2):218–26.
- 64. Yang Q-X, et al. PBK overexpression promotes metastasis of hepatocellular carcinoma via activating ETV4-uPAR signaling pathway. Cancer Lett. 2019;452:90–102.
- Park JH, et al. PBK attenuates paclitaxel-induced autophagic cell death by suppressing p53 in H460 non-small-cell lung cancer cells. FEBS Open Bio. 2020;10(5):937–50.
- 66. Majeed F, et al. Ginseng phytochemicals as therapeutics in oncology: Recent perspectives. Biomed Pharmacother. 2018;100:52–63.
- 67. Dou X, et al. PBK/TOPK mediates geranylgeranylation signaling for breast cancer cell proliferation. Cancer Cell Int. 2015;15(1):1–9.
- Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007;23(14):1846–7.
- 69. Do JH, Choi D-K. Normalization of microarray data: single-labeled and dual-labeled arrays. Mol Cell. 2006;22(3):254–61.
- Smyth GK. In: Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S, editors. Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Springer; 2005. p. 397–420.
- 71. Ferreira J, Zwinderman A. On the Benjamini–Hochberg method. Ann Stat. 2006;34(4):1827–49.
- Luke D. A User's Guide to Network Analysis in R. London; 2015. https:// doi.org/10.1007/978-3-319-23883-8.
- Kuleshov MV, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44(W1):W90–7.
- 74. Tang Z, et al. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res. 2019;47(W1):W556–60.
- Weinstein JN, et al. The cancer genome atlas pan-cancer analysis project. Nat Genet. 2013;45(10):1113.
- Goel MK, Khanna P, Kishore J. Understanding survival analysis: Kaplan-Meier estimate. Int J Ayurveda Res. 2010;1(4):274.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- ► Rigorous peer review
- Open access: articles freely available online
- ► High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com