



Identification of the hub genes in gastric cancer through weighted gene co-expression network analysis

Chunyang Li^{1,2}, Haopeng Yu^{1,2}, Yajing Sun^{1,2}, Xiaoxi Zeng^{1,2} and Wei Zhang^{1,2}

¹ West China Biomedical Big Data Center, West China Hospital, Sichuan University, Cheng, China

² Medical Big Data Center, Sichuan University, Chengdu, China

ABSTRACT

Background. Gastric cancer is one of the most lethal tumors and is characterized by poor prognosis and lack of effective diagnostic or therapeutic biomarkers. The aim of this study was to find hub genes serving as biomarkers in gastric cancer diagnosis and therapy.

Methods. GSE66229 from Gene Expression Omnibus (GEO) was used as training set. Genes bearing the top 25% standard deviations among all the samples in training set were performed to systematic weighted gene co-expression network analysis (WGCNA) to find candidate genes. Then, hub genes were further screened by using the “least absolute shrinkage and selection operator” (LASSO) logistic regression. Finally, hub genes were validated in the GSE54129 dataset from GEO by supervised learning method artificial neural network (ANN) algorithm.

Results. Twelve modules with strong preservation were identified by using WGCNA methods in training set. Of which, five modules significantly related to gastric cancer were selected as clinically significant modules, and 713 candidate genes were identified from these five modules. Then, *ADIPOQ*, *ARHGAP39*, *ATAD3A*, *C1orf95*, *CWH43*, *GRIK3*, *INHBA*, *RDH12*, *SCNN1G*, *SIGLEC11* and *LYVE1* were screened as the hub genes. These hub genes successfully differentiated the tumor samples from the healthy tissues in an independent testing set through artificial neural network algorithm with the area under the receiver operating characteristic curve at 0.946.

Conclusions. These hub genes bearing diagnostic and therapeutic values, and our results may provide a novel prospect for the diagnosis and treatment of gastric cancer in the future.

Submitted 11 August 2020

Accepted 9 December 2020

Published 5 March 2021

Corresponding author

Xiaoxi Zeng, zengxiaoxi@wchscu.cn

Academic editor

Alberto Davalos

Additional Information and
Declarations can be found on
page 15

DOI 10.7717/peerj.10682

© Copyright
2021 Li et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Bioinformatics, Gastroenterology and Hepatology, Oncology, Medical Genetics

Keywords Gastric cancer, Weighted gene co-expression network analysis, WGCNA, LASSO regression, Supervised machine learning

BACKGROUND

Gastric carcinoma remains the fifth most frequently diagnosed cancer and the third leading cause of cancer-related deaths, with an estimated 1,033,701 new cases and 782,685 deaths worldwide in 2018 (Bray et al., 2018; Pormohammad et al., 2018). Gastric cancer is also one of the most common malignancies and the third leading cause of death in China, where 427,100 cases with 301,200 deaths were observed in 2013 (Chen et al., 2017b). Despite the

several existing treatments including chemo-, radio-, or targeted therapy, the overall 5-year survival rate of stomach cancer patients is still <20% ([Raimondi et al., 2018](#)).

There are two types of gastric cancer, diffuse and intestinal types, which differ in their histological manifestations, epidemiological features and etiologic pathogenesis ([Huang et al., 2019](#)). Histopathology is the gold standard approach for diagnosing gastric cancer; however, this approach is not suitable for everyone due to the invasive nature of the biopsy ([Yoon & Kim, 2015](#)). Although there are several commonly used serum biomarkers such as alpha-fetoprotein (AFP), carcinoembryonic antigen (CEA), cancer antigen 125 (CA125), and cancer antigen 19–9 (CA19-9) ([He et al., 2013](#)) for gastric cancer diagnosis, none of them are sensitive or gastric-cancer-specific ([Smyth et al., 2016](#)). Moreover, effective and specific targeted therapies for gastric cancer remain to be identified. Presently, the major treatment strategies for gastric cancer are anti-human epidermal growth factor receptor 2 (HER2) and anti-vascular therapies ([Raimondi et al., 2018](#)). However, resistance to the targeted agents is common in some gastric tumor types. Therefore, novel practical approaches are needed for specific diagnosis and effective treatment of gastric cancer. Accordingly, identification of the key genes and biomarkers that are involved in the pathogenesis of gastric cancer is of paramount significance.

With recent advancements in bioinformatics methods, comprehensive identification of potential biomarkers through large-scale screening of expression profiles has been proposed ([Li et al., 2018a](#); [Takeno et al., 2008](#); [Wang et al., 2014](#); [Zeng et al., 2019](#)). A weighted gene co-expression network analysis (WGCNA) approach provides a systematic analysis to investigate the functional clustering of expression profiles, based on the theory that genes with similar expression profiles may have closely functional linkages and/or pathways ([Carlson et al., 2006](#); [Carter et al., 2004](#); [Zhou et al., 2018](#)). This approach groups highly co-expressed genes into the same module. Modules bearing high correlation with certain clinical traits are identified as clinically significant modules ([Zhou et al., 2018](#)).

By using this systematic bioinformatic method, followed by the “least absolute shrinkage and selection operator” (LASSO) logistic regression, a suitable method for high-dimensional gene data analysis ([Friedman, Hastie & Tibshirani, 2010](#); [Zeng et al., 2019](#)), candidate variables were selected from clinically significant modules. Finally, supervised artificial neural network (ANN) method was performed to test the reliability of the results in an independent dataset. ANN approach has been widely used in the prediction of cancer diagnosis, staging and recurrence since the mid-1990s ([Hu et al., 2013](#)), which is an useful method to incorporate and analyze large amounts of omics and health-care data ([Ngiam & Khor, 2019](#)).

Consequently, we attempt to construct a co-expression network by using systematic WGCNA method followed by LASSO regression to identify hub genes, which could effectively discriminate cancer samples from normal tissue. These findings may provide potential diagnostic and therapeutic targets in future research and clinical intervention of gastric cancer.

MATERIAL AND METHODS

Data collection and preprocessing

The workflow of this study is shown in Fig. 1. Raw expression datasets were downloaded from the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>) by using the keywords “stomach/gastric cancer/tumor/carcinoma”, “normal”, “GPL570”, and “*Homo sapiens*”. Our inclusion criteria for the training set were that: (1) datasets based on the Affymetrix Human Genome U133 Plus 2.0 Array Platform (Affymetrix, Santa Clara, CA, USA) and (2) datasets derived from human case-control studies, with gastric tumor patients as the case group, regardless of the histopathological types and stages, and non-tumor individuals as the control group.

Therefore, only two datasets (GSE66229 (Cristescu *et al.*, 2015; Oh *et al.*, 2018) and GSE54129) as of October 10, 2020 met the screening criteria. GSE66229 contained 300 tumor and 100 normal samples, and was used as the training set to screen for the hub genes. GSE54129 (contained 111 tumor and 21 normal samples) served as an independent testing set to validate the hub genes.

All the analyses in this study were conducted using R software (version 3.5.1). FitPLM weight, Relative Log Expression (RLE), Normalized Unscaled Standard Errors (NUSE), and RNA degradation images were analyzed to evaluate the quality of each dataset. Then, the “rma” function with the default parameters of the “affy” package was used to perform background correction and normalization (Gautier *et al.*, 2004). Missing values in each dataset were imputed by using the function “impute.knn” with the default parameters of the “impute” package (Hastie & Narasimhan, 2001). Platform annotations were downloaded from the GEO database, and finally, the gene symbol expression matrices were acquired from each dataset for further analyses.

Weighted gene co-expression network construction

Weighted gene co-expression network in the training set was constructed using the “WGCNA” package (Langfelder & Horvath, 2008; Zhang & Horvath, 2005). The genes with the top 25% SD among all the 400 samples in the expression matrix of the training set were selected as the input genes (5,115 genes in total).

In brief, first, the appropriate soft-thresholding power (β) was selected by using the “pickSoftThreshold” function with the default parameters (herein, $\beta = 4$). Subsequently, the Pearson’s correlation matrix was calculated to evaluate the similarity among all the pair-wise genes by using the “cor” function with the default parameters. Then, the adjacency was calculated based on β and the Pearson’s correlation matrix by using the “TOMsimilarity” function with the default parameters, and the corresponding dissimilarity (dissTOM) was also calculated. Finally, average linkage hierarchical clustering was conducted according to the dissTOM value with a minimum size of 30 for each gene dendrogram.

Module eigengenes (MEs), considered the first principal component (PC) of gene expression patterns of a corresponding module, were obtained for each module. To further strengthen the reliability of the modules, a cut line was set at 0.25 so that modules bearing <0.25 would be merged (Chen *et al.*, 2017a).

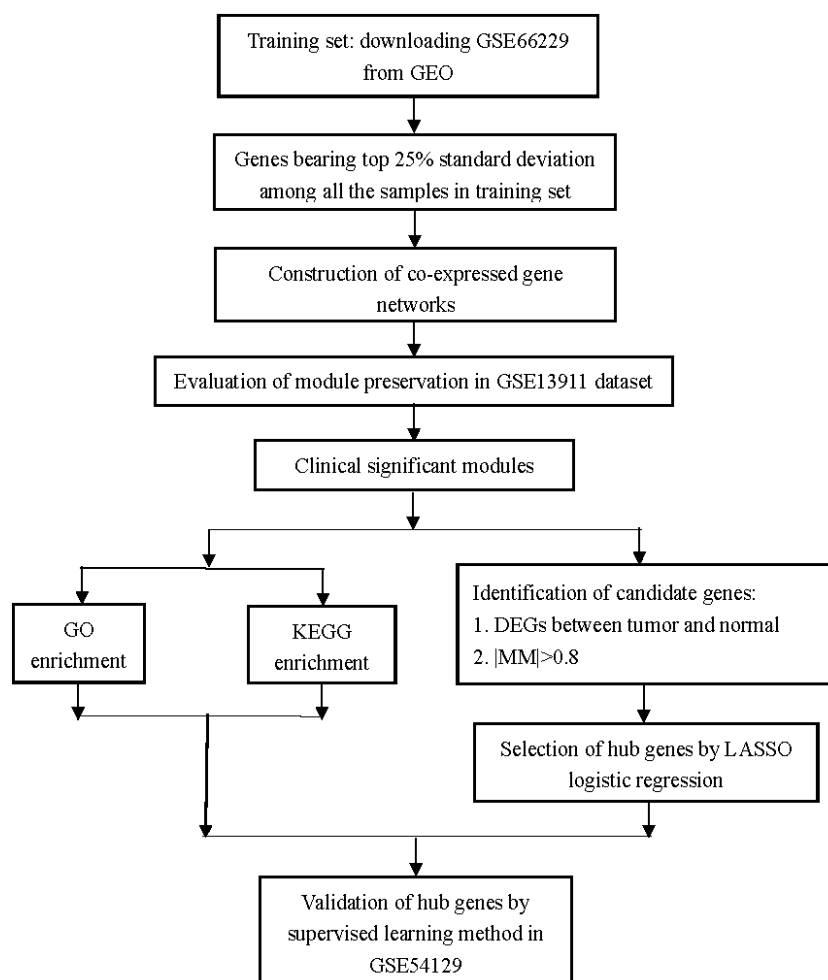


Figure 1 Flow diagram of the study. Data collection, analysis, and hub gene selection and validation.

Full-size  DOI: [10.7717/peerj.10682/fig-1](https://doi.org/10.7717/peerj.10682/fig-1)

Module preservation analysis

To evaluate the stability of the modules in the training set, GSE13911 was used to validate the module preservation of the training set (Chen *et al.*, 2018; Neidlin, Dimitrakopoulou & Alexopoulos, 2019; Obeidat *et al.*, 2017). Preservation analysis for GSE13911 was performed using the “modulePreservation” function by setting referenceNetworks = 2, nPermutations = 200, randomSeed = 1, and verbose = 3, maxModuleSize = 3000 and maxGoldModuleSize = 3000 in “WGCNA” package. The Z summary scores of each module were calculated to indicate the module preservations. Z summary scores <2, [2–10], and >10 indicated that the modules had no, moderate, and strong preservation, respectively (Liu *et al.*, 2019b; Lou *et al.*, 2017). The grey module contained the genes that did not belong to any of the other modules, and a gold module was generated for statistical purposes. Therefore, these two modules were not shown in the preservation analysis results (Langfelder *et al.*, 2011b).

Identification of clinically significant modules

Herein, the most interesting clinical trait was the tissue type, which was designated as tumor or normal samples. We calculated \log_{10} transformation of the P -value in the logistic regression between the MEs and clinical trait. Modules with \log_{10} transformation of the P -value greater than 10 were considered to be closely correlated with tissue types.

Functional enrichment and pathway analyses of significant modules

To determine whether the clinically significant modules were closely correlated with gastric cancer, GO functional annotation and KEGG pathway analyses were performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (version 6.8) (<https://david.ncifcrf.gov/home.jsp>) (Ashburner *et al.*, 2000; Dennis *et al.*, 2003). The visualization of the functional enrichment and pathway analyses was performed by the “GOplot” (Robin *et al.*, 2011) and “ggplot2” (Wickham, 2016) packages of R, respectively.

Candidate gene selection

Candidate genes among clinically significant modules were selected according to the following criteria: (1) differentially expressed genes (DEGs) between gastric cancer samples and normal samples with $|\log_2\text{FC (Fold Change)}| > 1$ and adjusted P -value < 0.05 based on the “limma” package (Ritchie *et al.*, 2015), (2) high module membership (defined as the correlation between the expression of each genes and MEs) $|\text{MM}(\text{Module Membership})| > 0.8$.

Selection of hub genes by LASSO logistic regression analysis

Candidate genes were subjected to LASSO regression, which was performed using the “glmnet” package by setting $\alpha = 1$, and ten-fold cross-validation for tuning parameter selection. Lambda was defined as the minimum partial likelihood deviance (Friedman, 2010).

Validation of hub genes

The machine learning method of ANN (by using the “neuralnet” function with hidden = 2) (Fritsch *et al.*, 2019) was performed to determine whether the hub genes could correctly distinguish the gastric cancer samples from the normal samples in a testing set. Moreover, in order to demonstrate whether these hub genes could specifically distinguish between gastric tumor and normal samples, we also evaluated the predictive effects of 11-gene model in pancreatic cancer (GSE15471) (Badea *et al.*, 2008) and colorectal cancer (GSE37364 excluding adenoma samples) (Galamb *et al.*, 2012) by using ANN algorithm.

Areas under the receiver operating characteristic (ROC) curve was calculated to show the predictive effect of supervised machine learning model, and then the ROC curves were plotted using the “pROC” package (Robin *et al.*, 2011). An area under the curve (AUC) value between 0.8 and 0.9 is considered an excellent classification, while greater than 0.9 is considered as outstanding discrimination (Lemeshow, 2000).

RESULTS

Construction of co-expression networks

After the quality check of the input data, no sample was removed (Fig. S1); herein, two clinical traits (tissue type and stage) are presented. According to different tissue types (tumor or normal), the 400 samples could be mainly divided into two clusters.

As shown in Fig. 2, the soft thresholding was set at 4, while the scale-free topology fit index reached 0.89, indicating approximate scale-free topology. Co-expressed gene modules were identified with the dynamic tree cut method (Fig. 3A) (Chen et al., 2017a). In total, 12 modules were found, and each color represented one module (Fig. 3B). The biggest module was the turquoise module, which contained 2,120 genes, followed by the blue module, bearing 1,503 genes. The grey module comprised 2 genes, which did not have a similar expression pattern and did not belong to any other module.

Module preservation analysis

5,115 genes in GSE13911 clustered into 11 colored modules (grey module only contained 2 genes, and did not show in Fig. 4), as determined in the training set. All gene modules were found to bear strong conservation, as the Z summary scores were all >10 (Fig. 4B).

Selection of clinically significant modules

After the assessment of the relationship by using regression analysis between the MEs and clinical traits, the \log_{10} transformation of the P -value was shown in Fig. 5. According to the screening criteria, there were 5 modules closely related to tissue types, which were black, turquoise, greenyellow, salmon and blue modules. These modules were selected for further analysis.

Functional enrichment of clinically significant modules

Gene Ontology (GO) enrichment results showed that 1503 genes in the blue module, 137 genes in the black module, 2,120 genes in the turquoise module, 61 in the salmon module and 194 genes in the greenyellow module mainly participated in 141, 15, 221, 54 and 146 different significant biological processes, respectively (Table S1 to Table S5). The top three most significantly enriched biological processes were cell division ($P = 2.29e-29$), G1/S transition of the mitotic cell cycle ($P = 6.32e-21$), mitotic nuclear division ($P = 3.49e-19$) in the blue module (Fig. 6A), and potassium ion import ($P = 3.39e-05$), digestion ($P = 8.25e-04$), multicellular organismal water homeostasis ($P = 0.0012$) in the black module (Fig. 6B), and extracellular matrix organization ($P = 6.30e-06$), positive regulation of cell migration ($P = 7.57e-06$), axon guidance in the turquoise module ($P = 9.18e-06$) (Fig. 6C), and inflammatory response ($P = 2.80e-21$), immune response ($P = 5.70e-17$), neutrophil chemotaxis ($P = 5.15e-15$) in the greenyellow module (Fig. 6D), and defense response to virus ($P = 3.33e-25$), type I interferon signaling pathway ($P = 8.52e-23$), response to virus ($P = 9.02e-16$) in salmon module (Fig. 6E).

KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis showed that genes in the blue, black, turquoise, salmon and greenyellow modules were mainly significantly enriched in 15, 7, 35, 14 and 23 pathways, respectively (from Table S6 to Table S10).

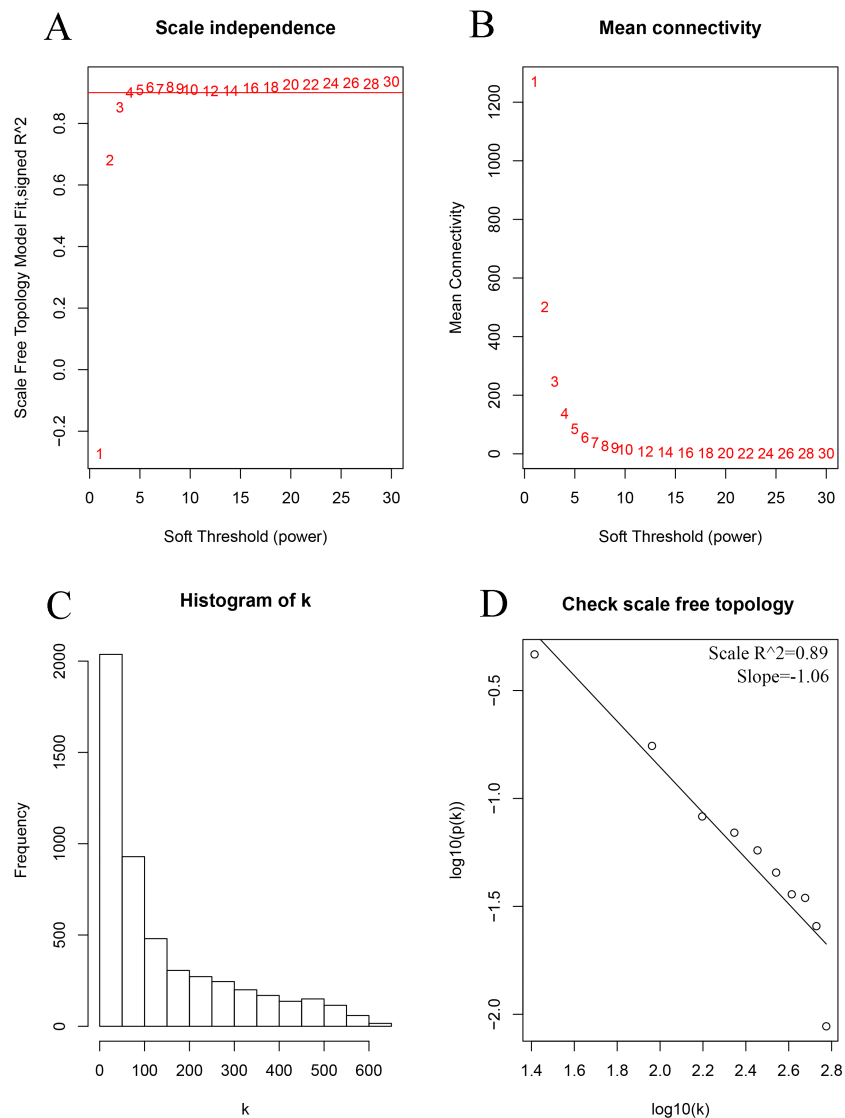


Figure 2 Determination of the soft-thresholding power in the weighted gene co-expression network analysis in the training set. (A) Screening soft-thresholding powers. (B) Analysis of the mean connectivity for various soft-thresholding powers. (C) Histogram of the connectivity distribution with the soft-thresholding powers set at 4. (D) Checking the scale-free topology with the soft-thresholding powers set at 4.

Full-size DOI: [10.7717/peerj.10682/fig-2](https://doi.org/10.7717/peerj.10682/fig-2)

The top three most significantly enriched pathways were cell cycle ($P = 1.21e-15$), DNA replication ($P = 6.21e-15$), the p53 signaling pathway ($P = 2.03e-06$) in blue module (Fig. 7A), and gastric acid secretion ($P = 4.46e-05$), protein digestion and absorption ($P = 1.28e-04$), drug metabolism-cytochrome P450 ($P = 0.0032$) in black moule (Fig. 7B), and focal adhesion ($P = 1.49e-07$), arrhythmogenic right ventricular cardiomyopathy ($P = 2.27e-05$), complement and coagulation cascades ($P = 3.58e-05$) in turquoise module (Fig. 7C), and cytokine-cytokine receptor interaction ($P = 2.35e-08$), amoebiasis ($P = 3.30e-05$), and rheumatoid arthritis ($P = 5.65e-05$) in greenyellow

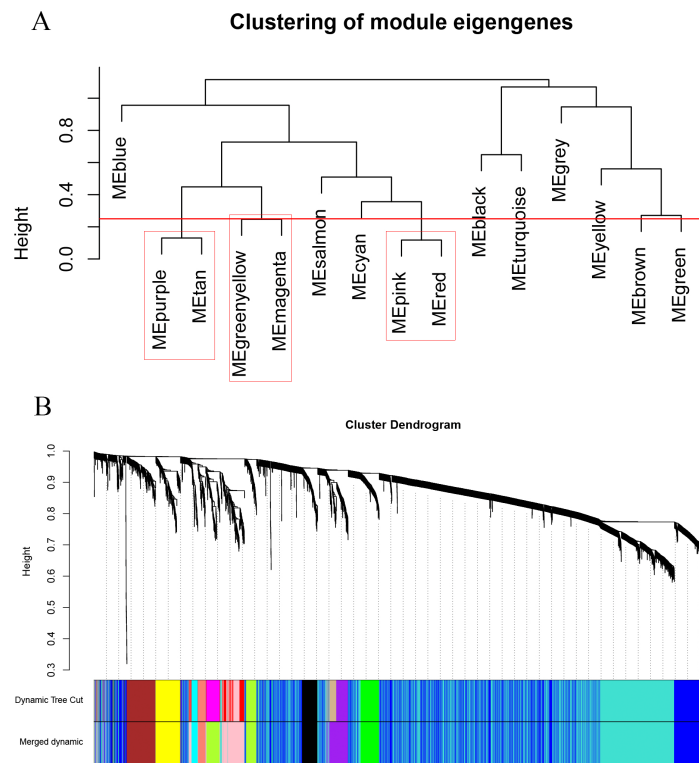


Figure 3 Clustering dendrograms of the 5,115 genes in the training set. (A) Clustering of the module eigengenes to identify the merged modules. Upon setting the threshold at 0.25, 15 modules were merged into 12 modules. (B) Co-expression module of the training set.

Full-size DOI: 10.7717/peerj.10682/fig-3

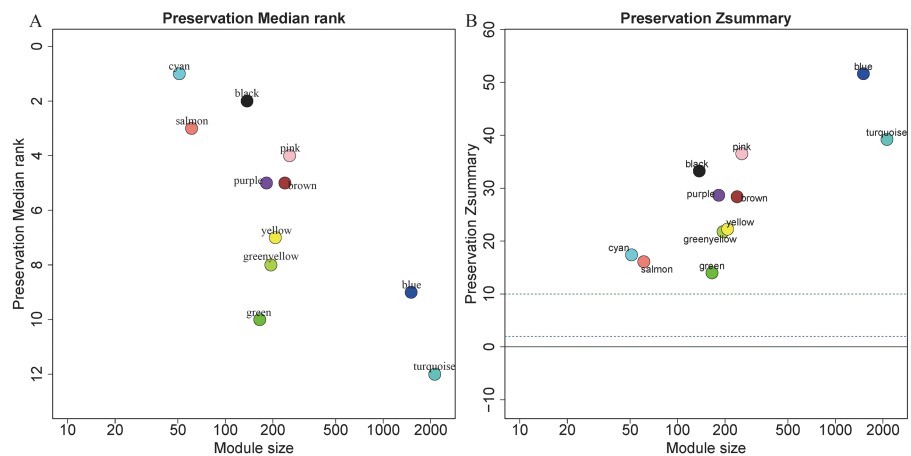


Figure 4 Evaluation of module preservation. The x- and y-axes present module size and preservation median rank (A) as well as preservation Z summary (B), respectively.

Full-size DOI: 10.7717/peerj.10682/fig-4

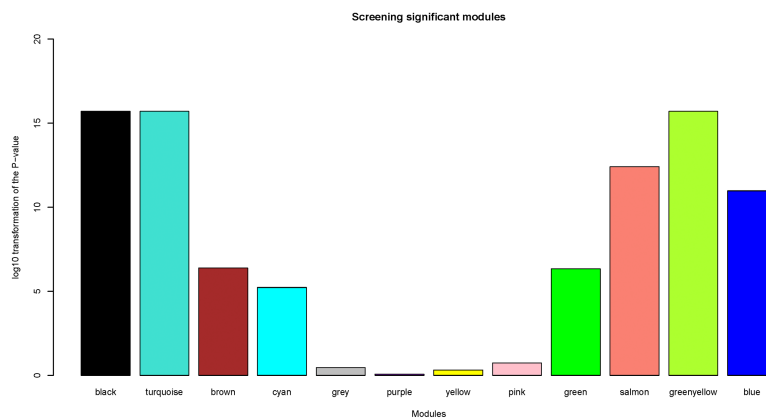


Figure 5 Identification of clinically significant modules. log₁₀ transformation of the *P*-value in the logistic regression between the MEs and clinical trait. The height of bars represents the log₁₀ transformation of the *P*-value, and modules with log₁₀ transformation of the *P*-value greater than 10 were considered to be closely correlated with tissue types.

Full-size DOI: 10.7717/peerj.10682/fig-5

module (Fig. 7D). While in the salmon module, influenza A ($P = 1.09e-09$), measles ($P = 9.44e-07$), and herpes simplex infection ($P = 7.99e-06$) (Fig. 7E) were the most significantly enriched pathways.

Identification of hub genes

Using the screening criteria of $|MM| > 0.8$, 926 genes were identified from 5 significant modules, of which 713 genes were differentially expressed between the normal and tumor samples with $|\logFC| > 1$ and adjusted *P*-value < 0.05 (all the 713 candidate genes were listed in Table S11). Finally, 11 genes [Adiponectin (*ADIPOQ*); Rho GTPase activating protein 39 (*ARHGAP39*); ATPase family AAA-domain containing protein 3A (*ATAD3A*); *C1orf95* (also known as *STUM* gene); Cell wall biogenesis 43 C-terminal homolog (*CWH43*); Glutamate receptor, ionotropic kainate 3 (*GRIK3*); Inhibin subunit beta A (*INHBA*); sodium channel epithelial 1 subunit gamma (*SCNN1G*); Sialic acid-binding immunoglobulin-like lectin-11 (*SIGLEC11*); Retinol dehydrogenase 12 (*RDH12*) and lymphatic vessel endothelial hyaluronan receptor 1 (*LYVE1*)] were identified as the hub genes by using LASSO logistic regression (Table 1). The heatmap of these 11 hub genes were shown in Fig. 8, indicating that these 11 hub genes differentially expressed between tumor and normal samples.

Validation of the hub genes

GSE54129 was utilized as the testing set to validate the 11-gene model. The AUC value of this classifier upon using artificial neural network was 0.946 indicating the excellent classification effects of the model (Fig. 9A). Furthermore, both the AUC values of this 11-gene model were around 0.5 in colorectal cancer (Fig. 9B) and pancreatic cancer (Fig. 9C), indicating specifically predictive effect in gastric cancer.



Figure 6 The bubble plot of gene ontology terms in the (A) blue, (B) black, (C) turquoise, (D) greenyellow and (E) salmon modules. The z-score is assigned to the x-axis and the negative logarithm of the p -value to the y-axis, as in the barplot (the higher the more significant). The area of the displayed circles is proportional to the number of genes assigned to the terms. Herein, only the meaningful enriched GO terms were presented, and the most three significant GO terms' labels were displayed.

Full-size [DOI: 10.7717/peerj.10682/fig-6](https://doi.org/10.7717/peerj.10682/fig-6)

DISCUSSION

In this study, five modules were identified as clinically significant and preserved modules by using WGCNA. The GO and KEGG analyses revealed that the genes in these four modules were significantly enriched in the biological processes of the cell cycle, cell division, and stomach-related functions. All these biological functions are closely related to gastric cancer (Cao *et al.*, 2018; Waldum, Sagatun & Mjones, 2017). Eventually, 11 hub genes including *ADIPOQ*, *ARHGAP39*, *ATAD3A*, *C1orf95*, *CWH43*, *GRIK3*, *INHBA*, *RDH12*, *SCNN1G*, *SIGLEC11* and *LYVE1* were screened by using WGCNA method followed by LASSO regression. Then, artificial neural network algorithms were performed, and demonstrated that this 11-gene model could effectively discriminate between gastric cancer and normal tissues.

In preservation analysis, Z summary is used to assess the significance of observed statistics and is defined as the mean of Z scores computed for density and connectivity measures (Lou *et al.*, 2017). When density and connectivity based preservation statistics are important

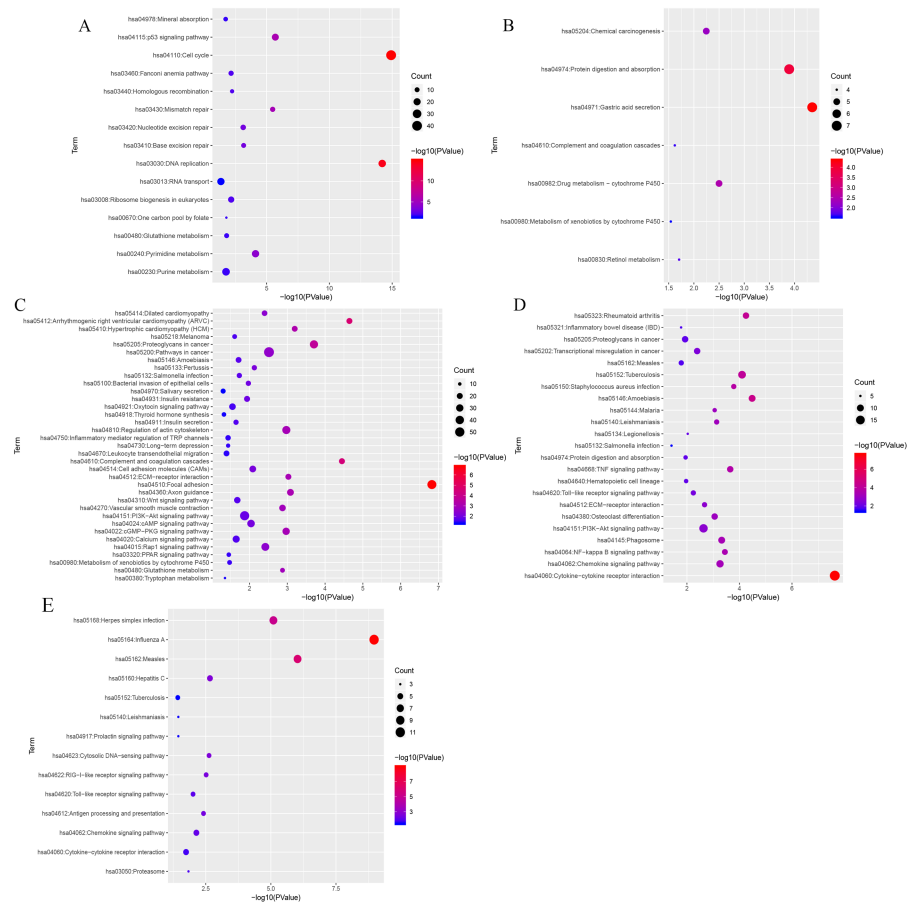


Figure 7 KEGG pathway enrichment analyses of the (A) blue, (B) black, (C) turquoise, (D) greenyellow and (E) salmon modules. The negative logarithm of the p -value is assigned to the x -axis and the term of each pathway to the y -axis. The size of the bubble shows the numbers of the genes enriched in each pathway, while the colors indicate the enrichment significance (from blue to red designated as less to high significance).

Full-size [DOI: 10.7717/peerj.10682/fig-7](https://doi.org/10.7717/peerj.10682/fig-7)

factors for judging the preservation of a network, Z summary score was preferentially selected to evaluate the preservation (Langfelder et al., 2011b). Although turquoise, blue and greenyellow had low preservation according to the median rank results, they are still considered to be conserved since the Z summary scores of these modules were 39, 52 and 22, respectively. Black module has relatively high Z summary score and low median rank, indicating high preservation. In the current study, all the four gene modules were considered to be conserved and selected for further analysis.

INHBA, Inhibin- β A (INHBA), a ligand belonging to the transforming growth factor- β superfamily (Oshima et al., 2014), is associated with cell proliferation in various tumor types including colon adenocarcinoma (Lin et al., 2020a; Miao et al., 2020; Miyamoto et al., 2020), pancreatic cancer (Liu et al., 2020), gastric cancer (Chen et al., 2019) as well as oral squamous cell carcinoma (Lin et al., 2020b). Many studies have demonstrated the prognostic role of INHBA in colon adenocarcinoma (Chen et al., 2020; Li et al., 2020;

Table 1 LASSO regression results. Genes selected by the LASSO logistic regression, with the estimated coefficients and odds ratio.

Gene	Coefficient	Odds ratio
ADIPOQ	-0.16241431554	0.8500889265
ARHGAP39	1.14238882470	3.134246595
ATAD3A	0.85913784838	2.361124169
C1orf95	-1.99447221881	0.1360854586
CWH43	-0.05325279770	0.9481402945
GRIK3	-4.02900881511	0.01779195633
INHBA	0.19110414846	1.210585526
LYVE1	-0.01065353474	0.9894030132
RDH12	-0.09270793497	0.9114596669
SCNN1G	-0.02709362847	0.9732701115
SIGLEC11	-0.17337880264	0.84081905

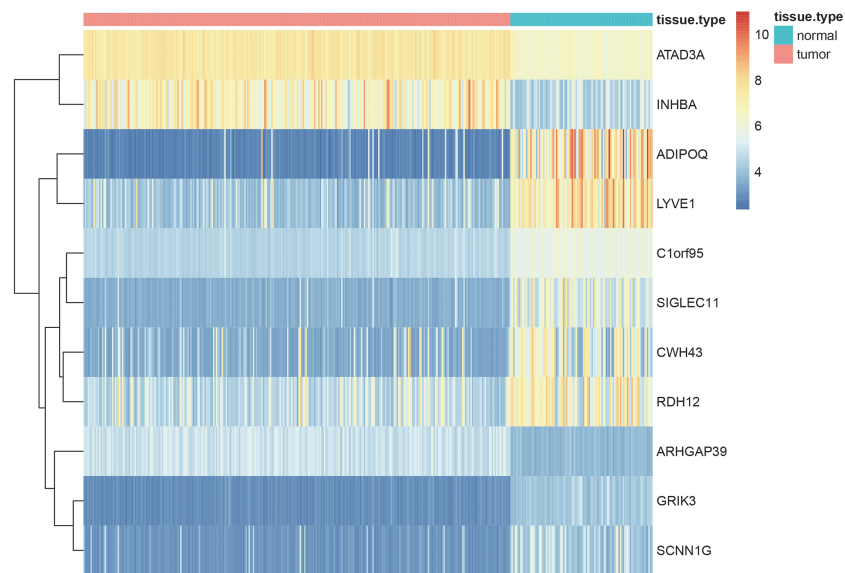


Figure 8 Heatmaps showing the expression profiles of 11 hub genes between the gastric cancer patients and normal controls in training set. The x - and y -axes present the samples and genes, respectively. In the x -axis, pink and green represent the gastric cancer and normal samples, respectively. The scale bar from blue to red represented low to high expressions of each gene in each sample.

Full-size [DOI: 10.7717/peerj.10682/fig-8](https://doi.org/10.7717/peerj.10682/fig-8)

Miao et al., 2020; Sun et al., 2020), and the role of INHBA in gastric cancer has been widely reported also. INHBA was highly expressed (*Kaneda et al., 2011; Seeruttun et al., 2019; Zhang et al., 2010*) and aberrantly methylated (*Zhang et al., 2019*) in gastric tumor samples, and high INHBA expression was associated with significantly poorer 5-year survival than low expression group (*Katayama et al., 2017; Wang et al., 2012*). One study has demonstrated that INHBA gene silencing could inhibit gastric cancer cell migration and invasion by impeding TGF- β signaling pathway (*Chen et al., 2019*).

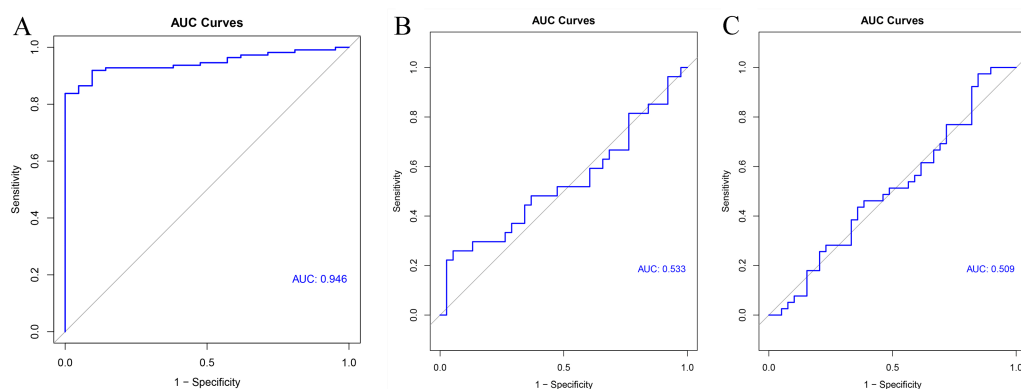


Figure 9 Validation results. ROC curve of the classifier predicted by these 11 hub genes in the testing set of (A) gastric cancer, (B) colorectal cancer and (C) pancreatic cancer.

Full-size DOI: [10.7717/peerj.10682/fig-9](https://doi.org/10.7717/peerj.10682/fig-9)

ADIPOQ is one of the most important adipocytokines secreted by adipocytes (*Parida, Siddharth & Sharma, 2019*), and the polymorphisms of *ADIPOQ* have been reported to correlate with several types of cancer, including colorectal (*Nimptsch et al., 2017; Tan et al., 2017*) and breast (*Mendez-Hernandez et al., 2017*) cancers. A study focusing on the molecular mechanisms *ADIPOQ* participated in has revealed that *ADIPOQ* induces cytotoxic autophagy in breast cancer cells through *STK11/LKB1*-mediated activation of the *AMPK-ULK1* axis (*Chung et al., 2017*). Another study has reported that *miR-370* inhibits the proliferation, invasion, and epithelial-mesenchymal transition of gastric cells by directly downregulating receptor 4 of *ADIPOQ* (*Feng et al., 2018*). Overexpression of another microRNA, *miR-15b-5p*, promotes the metastasis of gastric cancer by regulating *ADIPOQ* receptor 3 (*Zhao et al., 2017*).

ATAD3A is a nuclear-encoded mitochondrial enzyme, involving in mitochondrial dynamics, cell death, and cholesterol metabolism (*Teng, Lang & Shay, 2019*). It has been reported to correlate with hepatocellular carcinoma (*Liu et al., 2019a*) and breast cancer (*Daniel et al., 2019*), and it might be an effective therapeutic target in cancer treatment (*Teng et al., 2016*). *ATAD3A* is differentially expressed between paclitaxel-resistant and -sensitive MCF7 breast cancer cells (*Daniel et al., 2019*). A study has revealed that *ATAD3A* is upregulated in hepatocellular carcinoma and *ATAD3A* upregulation is correlated with poor prognosis (*Liu et al., 2019a*).

LYVE1 acts as a receptor and binds to both soluble and immobilized hyaluronan (*Banerji et al., 2016*), may function in lymphatic hyaluronan transport and tumor metastasis (*Wu et al., 2019*). The dysregulation of *LYVE1* closely correlated with many types of tumor, like gastric cancer (*Ozmen et al., 2011*), colorectal cancer (*Gao et al., 2006*), breast cancer (*Kato et al., 2005*), lung cancer (*Koukourakis et al., 2005*) and liver cancer (*Mouta Carreira et al., 2001*). *LYVE1* has been studied extensively for its possible role in cancer diagnosis and prognosis in cancer. One study has demonstrated that *LYVE1* was upregulated in gastric cancer, and overexpression of *LYVE1* positively correlated with perineural invasion and lymph node in gastric cancers (*Ozmen et al., 2011*). And the expression of *LYVE1* might

be a biomarker to predict the existence of regional lymph node metastasis in early gastric cancer (*Fujimoto et al., 2007*).

RDH12, an NADPH-dependent retinal reductase, catalyzes the reduction of all-trans retinal to all-trans retinol (*Belyaeva et al., 2005*). It was significantly decreased in gastric tumor samples (*Kropotova et al., 2013*) and cervical squamous cell carcinoma samples (*Peng et al., 2015*). RDH12 was also one of the differentially expressed metabolism-related genes, and correlated with the prognosis of gastric cancer patients (*Wen et al., 2020*).

GRIK3 mainly participates in the neuroactive ligand–receptor interaction pathway, and GRIK3 upregulation is associated with poor survival in gastric cancer (*Gong et al., 2017*). GRIK3 promotes epithelial-mesenchymal transition by regulating the SPDEF/CDH1 signaling in breast cancer cells (*Xiao et al., 2019*).

There have been few studies focusing on the relationship between gastric cancer and *SCNN1G*, *ARHGAP39*, *Clorf95*, *CWH43* or *SIGLEC11*. *SCNN1G* is one of the genes significantly upregulated in Ewing's sarcoma and fibromatosis samples (*Sarver et al., 2015*). *ARHGAP39* mutations or variations in copy number or expression level were found in several types of tumor-like tissues from the central nervous system, skin, prostate, and gastrointestinal tract (*Nowak, 2018*). *ARHGAP39* interacts with p53 and BAX, and decreased expression of *ARHGAP39* increases cell proliferation, leading to tumorigenesis (*Jones, 2017*). *Clorf95* is one of the uncharacterized proteins correlated with diverse human cancers (*Delgado et al., 2014*). Another study focusing on scleroderma patients demonstrated the involvement of *Clorf95* in cancer incidence (*Xu et al., 2016*). Sialic acid-binding immunoglobulin-like lectin-11 (*SIGLEC11*) is a primate-lineage-specific receptor of human tissue macrophages, and it is also expressed in brain microglia (*Angata et al., 2002*; *Shahraz et al., 2015*). A missense mutation of *SIGLEC11* has been detected in pancreatic cancer patients (*Jones et al., 2008*), and *SIGLEC11* was significantly upregulated in the poor prognostic group of pancreatic cancer patients (*Stratford et al., 2010*). *CWH43* was correlated with tumorigenesis in thyrotropin-secreting pituitary adenomas (*Sapkota et al., 2017*). One meta-analysis also showed that *CWH43* was differentially expressed between colorectal cancer and normal (*Chu et al., 2014*) samples.

All these results from the previous studies demonstrate that the hub genes identified in our study are closely correlated with gastric cancer and play important roles in cancer development, progression, or proliferation.

The significant module and hub genes identified in this study are biologically rational. First, the clinically significant module identified in our study bears strong preservation, implying that this clinically significant module is conservative and could also be reproduced in other datasets. Further, it suggests that that modules constructed by WGCNA are reliable. Second, most of the genes in the significant module were enriched for specific GO terms and KEGG pathways closely relating to stomach or cancer physiology. For instance, GO analysis demonstrated that most of the genes in the clinically significant modules were closely related to digestion, carbohydrate metabolic process, and gastric acid secretion, as well as cell division and cell cycle. KEGG enrichment analysis also indicated that most of the genes in the clinically significant module were implicated in gastric acid secretion, protein digestion and absorption, as well as glycerolipid metabolism and the p53 signaling pathway.

Third, all the hub genes identified in our study had previously been reported to relate to cancer. Moreover, several hub genes are implicated in metabolic processes, influencing the development and progression of gastric cancer. A previous study has demonstrated the association between metabolic syndrome and gastric cancer (*Li et al., 2018b*). A study has detected increased fatty acid oxidation in gastric cancer (*Lee et al., 2019*), and adipocytes fuel gastric cancer by mediating fatty acid metabolism (*Tan et al., 2018*). It may thus be inferred that these genes are genuinely the hub genes in charge of the key processes in gastric cancer, and they deserve a deeper analysis and validation. Finally, by using machine learning methods, the hub genes were demonstrated to effectively discriminate the gastric tumor samples from normal samples. In our study, the predictive effects of ANN method was evaluated by AUC values (*Huang & Ling, 2005*). Herein, the AUC value was >0.8 , indicating the excellent predictive results. Furthermore, these 11-gene model might be the specific predictors for gastric cancer, since the AUC values of this predictive model were less than 0.8 in other tumor types including, colorectal cancer and pancreatic cancer. All the results indicated that the expression profiles of these 11 hub genes have excellent predictive effects when discriminating gastric cancer samples from normal samples.

However, our study has limitations. First, all the hub genes were identified and validated only through bioinformatics, and further exploration of the biological functions and molecular mechanisms of these hub genes both in vitro and in vivo is required. Second, due to the limited availability of the data, we did not differentiate between intestinal-type and diffuse-type gastric cancers. More data are needed to analyze and identify the hub genes between these two types of gastric cancer and normal samples.

In summary, through WGCNA, we identified 11 hub genes, which might serve as potential diagnostic and/or therapeutic biomarkers for gastric cancer. Profile data mining by bioinformatics analysis is an available method to find potential diagnostic or therapeutic biomarkers systematically. Nevertheless, further investigations about the molecular mechanisms in which these hub genes are involved are still needed to verify the involvement of these genes in gastric cancer. Our findings provide a better understanding of the molecular mechanisms and putative diagnostic or therapeutic biomarkers for gastric cancer.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the Science & Technology Department of Sichuan Province funding project (No. 2016FZ0108, 2017FZ0104), the 1.3.5 project for disciplines of excellence, West China Hospital, Sichuan University (ZYJC18010). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

Science & Technology Department of Sichuan Province: 2016FZ0108, 2017FZ0104.

1.3.5 project for disciplines of excellence, West China Hospital, Sichuan University: ZYJC18010.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Chunyang Li conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Haopeng Yu performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Yajing Sun performed the experiments, analyzed the data, prepared figures and/or tables, performed the double check of the results, and approved the final draft.
- Xiaoxi Zeng conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Wei Zhang conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

R code and training set, pre-derivation analysis, and testing set raw data are available in the [Supplemental Files](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.10682#supplemental-information>.

REFERENCES

- Angata T, Kerr SC, Greaves DR, Varki NM, Crocker PR, Varki A. 2002. Cloning and characterization of human Siglec-11. A recently evolved signaling molecule that can interact with SHP-1 and SHP-2 and is expressed by tissue macrophages, including brain microglia. *Journal of Biological Chemistry* 277:24466–24474 DOI 10.1074/jbc.M202833200.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nature Genetics* 25:25–29 DOI 10.1038/75556.
- Badea L, Herlea V, Dima SO, Dumitrascu T, Popescu I. 2008. Combined gene expression analysis of whole-tissue and microdissected pancreatic ductal adenocarcinoma identifies genes specifically overexpressed in tumor epithelia. *Hepatogastroenterology* 55:2016–2027.

- Banerji S, Lawrance W, Metcalfe C, Briggs DC, Yamauchi A, Dushek O, Van der Merwe PA, Day AJ, Jackson DG. 2016.** Homodimerization of the lymph vessel endothelial receptor LYVE-1 through a redox-labile disulfide is critical for hyaluronan binding in lymphatic endothelium. *Journal of Biological Chemistry* **291**:25004–25018 DOI [10.1074/jbc.M116.736926](https://doi.org/10.1074/jbc.M116.736926).
- Belyaeva OV, Korkina OV, Stetsenko AV, Kim T, Nelson PS, Kedishvili NY. 2005.** Biochemical properties of purified human retinol dehydrogenase 12 (RDH12): catalytic efficiency toward retinoids and C9 aldehydes and effects of cellular retinol-binding protein type I (CRBPI) and cellular retinaldehyde-binding protein (CRALBP) on the oxidation and reduction of retinoids. *Biochemistry* **44**:7035–7047 DOI [10.1021/bi050226k](https://doi.org/10.1021/bi050226k).
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 2018.** Global Cancer Statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *A Cancer Journal for Clinicians* **68**(6):394–424 DOI [10.3322/caac.21492](https://doi.org/10.3322/caac.21492).
- Cao J, Wu N, Han Y, Hou Q, Zhao Y, Pan Y, Xie X, Chen F. 2018.** DDX21 promotes gastric cancer proliferation by regulating cell cycle. *Biochemical and Biophysical Research Communications* **505**:1189–1194 DOI [10.1016/j.bbrc.2018.10.060](https://doi.org/10.1016/j.bbrc.2018.10.060).
- Carlson MR, Zhang B, Fang Z, Mischel PS, Horvath S, Nelson SF. 2006.** Gene connectivity, function, and sequence conservation: predictions from modular yeast co-expression networks. *BMC Genomics* **7**:40 DOI [10.1186/1471-2164-7-40](https://doi.org/10.1186/1471-2164-7-40).
- Carter SL, Brechbuhler CM, Griffin M, Bond AT. 2004.** Gene co-expression network topology provides a framework for molecular characterization of cellular state. *Bioinformatics* **20**:2242–2250 DOI [10.1093/bioinformatics/bth234](https://doi.org/10.1093/bioinformatics/bth234).
- Chen L, Yuan L, Qian K, Qian G, Zhu Y, Wu CL, Dan HC, Xiao Y, Wang X. 2018.** Identification of biomarkers associated with pathological stage and prognosis of clear cell renal cell carcinoma by co-expression network analysis. *Frontiers in Physiology* **9**:399 DOI [10.3389/fphys.2018.00399](https://doi.org/10.3389/fphys.2018.00399).
- Chen L, Yuan L, Wang Y, Wang G, Zhu Y, Cao R, Qian G, Xie C, Liu X, Xiao Y, Wang X. 2017a.** Co-expression network analysis identified FCER1G in association with progression and prognosis in human clear cell renal cell carcinoma. *International Journal of Biological Sciences* **13**:1361–1372 DOI [10.7150/ijbs.21657](https://doi.org/10.7150/ijbs.21657).
- Chen S, Cao GD, Wei W, Yida L, Xiaobo H, Lei Y, Ke C, Chen B, Xiong MM. 2020.** Prediction and identification of immune genes related to the prognosis of patients with colon adenocarcinoma and its mechanisms. *World Journal of Surgical Oncology* **18**:146 DOI [10.1186/s12957-020-01921-9](https://doi.org/10.1186/s12957-020-01921-9).
- Chen W, Zheng R, Zhang S, Zeng H, Xia C, Zuo T, Yang Z, Zou X, He J. 2017b.** Cancer incidence and mortality in China, 2013. *Cancer Letters* **401**:63–71 DOI [10.1016/j.canlet.2017.04.024](https://doi.org/10.1016/j.canlet.2017.04.024).
- Chen ZL, Qin L, Peng XB, Hu Y, Liu B. 2019.** INHBA gene silencing inhibits gastric cancer cell migration and invasion by impeding activation of the TGF-beta signaling pathway. *Journal of Cellular Physiology* **234**:18065–18074 DOI [10.1002/jcp.28439](https://doi.org/10.1002/jcp.28439).

- Chu CM, Yao CT, Chang YT, Chou HL, Chou YC, Chen KH, Terng HJ, Huang CS, Lee CC, Su SL, Liu YC, Lin FG, Wetter T, Chang CW. 2014.** Gene expression profiling of colorectal tumors and normal mucosa by microarrays meta-analysis using prediction analysis of microarray, artificial neural network, classification, and regression trees. *Disease Markers* **2014**:634123 DOI [10.1155/2014/634123](https://doi.org/10.1155/2014/634123).
- Chung SJ, Nagaraju GP, Nagalingam A, Muniraj N, Kuppusamy P, Walker A, Woo J, Gyorffy B, Gabrielson E, Saxena NK, Sharma D. 2017.** ADIPOQ/adiponectin induces cytotoxic autophagy in breast cancer cells through STK11/LKB1-mediated activation of the AMPK-ULK1 axis. *Autophagy* **13**:1386–1403 DOI [10.1080/15548627.2017.1332565](https://doi.org/10.1080/15548627.2017.1332565).
- Cristescu R, Lee J, Nebozhyn M, Kim KM, Ting JC, Wong SS, Liu J, Yue YG, Wang J, Yu K, Ye XS, Do IG, Liu S, Gong L, Fu J, Jin JG, Choi MG, Sohn TS, Lee JH, Bae JM, Kim ST, Park SH, Sohn I, Jung SH, Tan P, Chen R, Hardwick J, Kang WK, Ayers M, Hongyue D, Reinhard C, Loboda A, Kim S, Aggarwal A. 2015.** Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nature Medicine* **21**:449–456 DOI [10.1038/nm.3850](https://doi.org/10.1038/nm.3850).
- Daniel P, Halada P, Jelinek M, Balusikova K, Kovar J. 2019.** Differentially expressed mitochondrial proteins in human MCF7 breast cancer cells resistant to paclitaxel. *International Journal of Molecular Sciences* **20**(12):2986 DOI [10.3390/ijms20122986](https://doi.org/10.3390/ijms20122986).
- Delgado AP, Brandao P, Chaoado MJ, Hamid S, Narayanan R. 2014.** Open reading frames associated with cancer in the dark matter of the human genome. *Cancer Genomics and Proteomics* **11**(4):201–213.
- Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. 2003.** DAVID: database for annotation, visualization, and integrated discovery. *Genome Biology* **4**:R60 DOI [10.1186/gb-2003-4-9-r60](https://doi.org/10.1186/gb-2003-4-9-r60).
- Feng Y, Sun T, Yu Y, Gao Y, Wang X, Chen Z. 2018.** MicroRNA-370 inhibits the proliferation, invasion and EMT of gastric cancer cells by directly targeting PAQR4. *Journal of Pharmacological Sciences* **138**:96–106 DOI [10.1016/j.jphs.2018.08.004](https://doi.org/10.1016/j.jphs.2018.08.004).
- Friedman J. 2010.** Package ‘glmnet’. Available at <https://cran.r-project.org/web/packages/glmnet/index.html>.
- Friedman J, Hastie T, Tibshirani R. 2010.** Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software* **33**:1–22.
- Fritsch S, Guenther F, Wright MN, Suling M, Mueller SM, Wright MN. 2019.** neuralnet: training of neural networks. Available at <https://journal.r-project.org/archive/2010/RJ-2010-006/index.html>.
- Fujimoto A, Ishikawa Y, Akishima-Fukasawa Y, Ito K, Akasaka Y, Tamai S, Maehara T, Kiguchi H, Ogata K, Nishimura C, Miki K, Ishii T. 2007.** Significance of lymphatic invasion on regional lymph node metastasis in early gastric cancer using LYVE-1 immunohistochemical analysis. *American Journal of Clinical Pathology* **127**:82–88 DOI [10.1309/LJQ9G0X8KP17QXP3](https://doi.org/10.1309/LJQ9G0X8KP17QXP3).
- Galamb O, Wichmann B, Sipos F, Spisak S, Krenacs T, Toth K, Leiszter K, Kalmar A, Tulassay Z, Molnar B. 2012.** Dysplasia-carcinoma transition specific transcripts in

- colonic biopsy samples. *PLOS ONE* 7:e48547
DOI 10.1371/journal.pone.0048547.
- Gao F, Lu YM, Cao ML, Liu YW, He YQ, Wang Y. 2006. Expression and quantification of LYVE-1 in human colorectal cancer. *Clinical and Experimental Medicine* 6:65–71
DOI 10.1007/s10238-006-0097-4.
- Gautier L, Cope L, Bolstad BM, Irizarry RA. 2004. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20:307–315
DOI 10.1093/bioinformatics/btg405.
- Gong B, Li Y, Cheng Z, Wang P, Luo L, Huang H, Duan S, Liu F. 2017. GRIK3: a novel oncogenic protein related to tumor TNM stage, lymph node metastasis, and poor prognosis of GC. *Tumour Biology* 39:1010428317704364
DOI 10.1177/1010428317704364.
- Hastie TTR, Narasimhan B. 2001. impute: imputation for microarray data. *Bioinformatics* 17:6.
- He CZ, Zhang KH, Li Q, Liu XH, Hong Y, Lv NH. 2013. Combined use of AFP, CEA, CA125 and CA19-9 improves the sensitivity for the diagnosis of gastric cancer. *BMC Gastroenterology* 13:87 DOI 10.1186/1471-230X-13-87.
- Hu X, Cammann H, Meyer HA, Miller K, Jung K, Stephan C. 2013. Artificial neural networks and prostate cancer—tools for diagnosis and management. *Nature Reviews Urology* 10:174–182 DOI 10.1038/nrurol.2013.9.
- Huang J, Ling CX. 2005. Using AUC and accuracy in evaluating learning algorithms. *Ieee Transactions on Knowledge and Data Engineering* 17:299–310
DOI 10.1109/Tkde.2005.50.
- Huang SC, Ng KF, Yeh TS, Cheng CT, Lin JS, Liu YJ, Chuang HC, Chen TC. 2019. Subtraction of Epstein-Barr virus and microsatellite instability genotypes from the Lauren histotypes: combined molecular and histologic subtyping with clinicopathological and prognostic significance validated in a cohort of 1,248 cases. *International Journal of Cancer* 145:3218–3230 DOI 10.1002/ijc.32215.
- Jones R. 2017. Identification of novel risk variants for sarcoma and other cancers by whole exome sequencing in cancer cluster families. PhD Thesis, The University of Western Australia.
- Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, Mankoo P, Carter H, Kamiyama H, Jimeno A, Hong SM, Fu B, Lin MT, Calhoun ES, Kamiyama M, Walter K, Nikolskaya T, Nikolsky Y, Hartigan J, Smith DR, Hidalgo M, Leach SD, Klein AP, Jaffee EM, Goggins M, Maitra A, Iacobuzio-Donahue C, Eshleman JR, Kern SE, Hruban RH, Karchin R, Papadopoulos N, Parmigiani G, Vogelstein B, Velculescu VE, Kinzler KW. 2008. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* 321:1801–1806
DOI 10.1126/science.1164368.
- Kaneda H, Arao T, Matsumoto K, De Velasco MA, Tamura D, Aomatsu K, Kudo K, Sakai K, Nagai T, Fujita Y, Tanaka K, Yanagihara K, Yamada Y, Okamoto I, Nakagawa K, Nishio K. 2011. Activin A inhibits vascular endothelial cell growth

- and suppresses tumour angiogenesis in gastric cancer. *British Journal of Cancer* **105**:1210–1217 DOI [10.1038/bjc.2011.348](https://doi.org/10.1038/bjc.2011.348).
- Katayama Y, Oshima T, Sakamaki K, Aoyama T, Sato T, Masudo K, Shiozawa M, Yoshikawa T, Rino Y, Imada T, Masuda M. 2017.** Clinical significance of INHBA gene expression in patients with gastric cancer who receive curative resection followed by adjuvant S-1 chemotherapy. *In Vivo* **31**:565–571 DOI [10.21873/invivo.11095](https://doi.org/10.21873/invivo.11095).
- Kato T, Prevo R, Steers G, Roberts H, Leek RD, Kimura T, Kameoka S, Nishikawa T, Kobayashi M, Jackson DG, Harris AL, Gatter KC, Pezzella F. 2005.** A quantitative analysis of lymphatic vessels in human breast cancer, based on LYVE-1 immunoreactivity. *British Journal of Cancer* **93**:1168–1174 DOI [10.1038/sj.bjc.6602844](https://doi.org/10.1038/sj.bjc.6602844).
- Koukourakis MI, Giatromanolaki A, Sivridis E, Simopoulos C, Gatter KC, Harris AL, Jackson DG. 2005.** LYVE-1 immunohistochemical assessment of lymphangiogenesis in endometrial and lung cancer. *Journal of Clinical Pathology* **58**:202–206 DOI [10.1136/jcp.2004.019174](https://doi.org/10.1136/jcp.2004.019174).
- Kropotova ES, Zinov'eva OL, Zyrianova AF, Choinzonov EL, Afanas'ev SG, Cherdyntseva NV, Beresten SF, Oparina N, Mashkova TD. 2013.** Expression of genes involved in retinoic acid biosynthesis in human gastric cancer. *Molekuliarnaia Biologiya* **47**:317–330 DOI [10.7868/s0026898413020079](https://doi.org/10.7868/s0026898413020079).
- Langfelder P, Horvath S. 2008.** WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**:559 DOI [10.1186/1471-2105-9-559](https://doi.org/10.1186/1471-2105-9-559).
- Langfelder P, Luo R, Oldham MC, Horvath S. 2011.** Is my network module preserved and reproducible? *PLOS Computational Biology* **7**:e1001057 DOI [10.1371/journal.pcbi.1001057](https://doi.org/10.1371/journal.pcbi.1001057).
- Lee JS, Kim SH, Lee S, Kang JH, Lee SH, Cheong JH, Kim SY. 2019.** Gastric cancer depends on aldehyde dehydrogenase 3A1 for fatty acid oxidation. *Scientific Reports* **9**:16313 DOI [10.1038/s41598-019-52814-1](https://doi.org/10.1038/s41598-019-52814-1).
- Lemeshow DWHS. 2000.** *Applied logistic regression*. 2nd Ed. Hoboken: Wiley.
- Li C, Zeng X, Yu H, Gu Y, Zhang W. 2018a.** Identification of hub genes with diagnostic values in pancreatic cancer by bioinformatics analyses and supervised learning methods. *World Journal of Surgical Oncology* **16**:223 DOI [10.1186/s12957-018-1519-y](https://doi.org/10.1186/s12957-018-1519-y).
- Li F, Du H, Li S, Liu J. 2018b.** The association between metabolic syndrome and gastric cancer in Chinese. *Frontiers in Oncology* **8**:326 DOI [10.3389/fonc.2018.00326](https://doi.org/10.3389/fonc.2018.00326).
- Li X, Yu W, Liang C, Xu Y, Zhang M, Ding X, Cai X. 2020.** INHBA is a prognostic predictor for patients with colon adenocarcinoma. *BMC Cancer* **20**:305 DOI [10.1186/s12885-020-06743-2](https://doi.org/10.1186/s12885-020-06743-2).
- Lin H, Hong YG, Zhou JD, Gao XH, Yuan PH, Xin C, Huang ZP, Zhang W, Hao LQ, Hou KZ. 2020a.** LncRNA INHBA-AS1 promotes colorectal cancer cell proliferation by sponging miR-422a to increase AKT1 axis. *European Review for Medical and Pharmacological Sciences* **24**:9940–9948 DOI [10.26355/eurrev_202010_23206](https://doi.org/10.26355/eurrev_202010_23206).

- Lin TA, Wu TS, Li YJ, Yang CN, Illescas Ralda MM, Chang HH. 2020b.** Role and mechanism of LIF in oral squamous cell carcinoma progression. *Journal of Clinical Medicine* **9**:295 DOI [10.3390/jcm9020295](https://doi.org/10.3390/jcm9020295).
- Liu D, Zhou D, Sun Y, Zhu J, Ghoneim D, Wu C, Yao Q, Gamazon ER, Cox NJ, Wu L. 2020.** A transcriptome-wide association study identifies candidate susceptibility genes for pancreatic cancer risk. *Cancer Research* **80**:4346–4354 DOI [10.1158/0008-5472.CAN-20-1353](https://doi.org/10.1158/0008-5472.CAN-20-1353).
- Liu X, Li G, Ai L, Ye Q, Yu T, Yang B. 2019a.** Prognostic value of ATAD3 gene cluster expression in hepatocellular carcinoma. *Oncology Letters* **18**:1304–1310 DOI [10.3892/ol.2019.10454](https://doi.org/10.3892/ol.2019.10454).
- Liu Z, Li M, Hua Q, Li Y, Wang G. 2019b.** Identification of an eight-lncRNA prognostic model for breast cancer using WGCNA network analysis and a Cox-proportional hazards model based on L1-penalized estimation. *International Journal of Molecular Medicine* **44**:1333–1343.
- Lou Y, Tian GY, Song Y, Liu YL, Chen YD, Shi JP, Yang J. 2017.** Characterization of transcriptional modules related to fibrosing-NAFLD progression. *Scientific Reports* **7**:4748 DOI [10.1038/s41598-017-05044-2](https://doi.org/10.1038/s41598-017-05044-2).
- Mendez-Hernandez A, Gallegos-Arreola MP, Moreno-Macias H, Espinosa Fematt J, Perez-Morales R. 2017.** LEP rs7799039, LEPR rs1137101, and ADIPOQ rs2241766 and 1501299 Polymorphisms are associated with obesity and chemotherapy response in mexican women with breast cancer. *Clinical Breast Cancer* **17**:453–462 DOI [10.1016/j.clbc.2017.03.010](https://doi.org/10.1016/j.clbc.2017.03.010).
- Miao YD, Wang JT, Ma XP, Yang Y, Mi D. 2020.** Identification prognosis-associated immune genes in colon adenocarcinoma. *Bioscience Reports* **40**(11):BSR20201734 DOI [10.1042/BSR20201734](https://doi.org/10.1042/BSR20201734).
- Miyamoto Y, Schirripa M, Suenaga M, Cao S, Zhang W, Okazaki S, Berger MD, Matsusaka S, Yang D, Ning Y, Baba H, Loupakis F, Lonardi S, Pietrantonio F, Borelli B, Cremolini C, Yamaguchi T, Lenz HJ. 2020.** A polymorphism in the cachexia-associated gene INHBA predicts efficacy of regorafenib in patients with refractory metastatic colorectal cancer. *PLOS ONE* **15**:e0239439 DOI [10.1371/journal.pone.0239439](https://doi.org/10.1371/journal.pone.0239439).
- Mouta Carreira C, Nasser SM, Di Tomaso E, Padera TP, Boucher Y, Tomarev SI, Jain RK. 2001.** LYVE-1 is not restricted to the lymph vessels: expression in normal liver blood sinusoids and down-regulation in human liver cancer and cirrhosis. *Cancer Research* **61**:8079–8084.
- Neidlin M, Dimitrakopoulou S, Alexopoulos LG. 2019.** Multi-tissue network analysis for drug prioritization in knee osteoarthritis. *Scientific Reports* **9**:15176 DOI [10.1038/s41598-019-51627-6](https://doi.org/10.1038/s41598-019-51627-6).
- Ngiam KY, Khor IW. 2019.** Big data and machine learning algorithms for health-care delivery. *The Lancet Oncology* **20**:e262–e273 DOI [10.1016/S1470-2045\(19\)30149-4](https://doi.org/10.1016/S1470-2045(19)30149-4).
- Nimptsch K, Song M, Aleksandrova K, Katsoulis M, Freisling H, Jenab M, Gunter MJ, Tsilidis KK, Weiderpass E, Bueno-De-Mesquita HB, Chong DQ, Jensen**

- MK, Wu C, Overvad K, Kuhn T, Barrdahl M, Melander O, Jirstrom K, Peeters PH, Sieri S, Panico S, Cross AJ, Riboli E, Van Guelpen B, Myte R, Huerta JM, Rodriguez-Barranco M, Quiros JR, Dorronsoro M, Tjonneland A, Olsen A, Travis R, Boutron-Ruault MC, Carbonnel F, Severi G, Bonet C, Palli D, Janke J, Lee YA, Boeing H, Giovannucci EL, Ogino S, Fuchs CS, Rimm E, Wu K, Chan AT, Pischon T. 2017. Genetic variation in the ADIPOQ gene, adiponectin concentrations and risk of colorectal cancer: a mendelian randomization analysis using data from three large cohort studies. *European Journal of Epidemiology* **32**:419–430 DOI [10.1007/s10654-017-0262-y](https://doi.org/10.1007/s10654-017-0262-y).
- Nowak FV. 2018. Porf-2 = Arhgap39 = Vilse: a pivotal role in neurodevelopment, learning and memory. *eNeuro* **5**(5):ENEURO.0082-18.2018 DOI [10.1523/ENEURO.0082-18.2018](https://doi.org/10.1523/ENEURO.0082-18.2018).
- Obeidat M, Nie Y, Chen V, Shannon CP, Andiappan AK, Lee B, Rotzschke O, Castaldi PJ, Hersh CP, Fishbane N, Ng RT, McManus B, Miller BE, Rennard S, Pare PD, Sin DD. 2017. Network-based analysis reveals novel gene signatures in peripheral blood of patients with chronic obstructive pulmonary disease. *Respiratory Research* **18**:72 DOI [10.1186/s12931-017-0558-1](https://doi.org/10.1186/s12931-017-0558-1).
- Oh SC, Sohn BH, Cheong JH, Kim SB, Lee JE, Park KC, Lee SH, Park JL, Park YY, Lee HS, Jang HJ, Park ES, Kim SC, Heo J, Chu IS, Jang YJ, Mok YJ, Jung W, Kim BH, Kim A, Cho JY, Lim JY, Hayashi Y, Song S, Elimova E, Estralla JS, Lee JH, Bhutani MS, Lu Y, Liu W, Lee J, Kang WK, Kim S, Noh SH, Mills GB, Kim SY, Ajani JA, Lee JS. 2018. Clinical and genomic landscape of gastric cancer with a mesenchymal phenotype. *Nature Communications* **9**:1777 DOI [10.1038/s41467-018-04179-8](https://doi.org/10.1038/s41467-018-04179-8).
- Oshima T, Yoshihara K, Aoyama T, Hasegawa S, Sato T, Yamamoto N, Akito N, Shiozawa M, Yoshikawa T, Numata K, Rino Y, Kunisaki C, Tanaka K, Akaike M, Imada T, Masuda M. 2014. Relation of INHBA gene expression to outcomes in gastric cancer after curative surgery. *Anticancer Research* **34**:2303–2309.
- Ozmen F, Ozmen MM, Ozdemir E, Moran M, Seckin S, Guc D, Karaagaoglu E, Kansu E. 2011. Relationship between LYVE-1, VEGFR-3 and CD44 gene expressions and lymphatic metastasis in gastric cancer. *World Journal of Gastroenterology* **17**:3220–3228 DOI [10.3748/wjg.v17.i27.3220](https://doi.org/10.3748/wjg.v17.i27.3220).
- Parida S, Siddharth S, Sharma D. 2019. Adiponectin, obesity, and cancer: clash of the bigwigs in health and disease. *International Journal of Molecular Sciences* **20**:2519 DOI [10.3390/ijms20102519](https://doi.org/10.3390/ijms20102519).
- Peng G, Dan W, Jun W, Junjun Y, Tong R, Baoli Z, Yang X. 2015. Transcriptome profiling of the cancer and adjacent nontumor tissues from cervical squamous cell carcinoma patients by RNA sequencing. *Tumour Biology* **36**:3309–3317 DOI [10.1007/s13277-014-2963-0](https://doi.org/10.1007/s13277-014-2963-0).
- Pormohammad A, Mohtavinejad N, Gholizadeh P, Dabiri H, Salimi Chirani A, Hashemi A, Nasiri MJ. 2018. Global estimate of gastric cancer in Helicobacter pylori-infected population: a systematic review and meta-analysis. *Journal of Cellular Physiology* **234**(2):1208–1218 DOI [10.1002/jcp.27114](https://doi.org/10.1002/jcp.27114).

- Raimondi A, Nichetti F, Peverelli G, Di Bartolomeo M, Braud FDe, Pietrantonio F. 2018.** Genomic markers of resistance to targeted treatments in gastric cancer: potential new treatment strategies. *Pharmacogenomics* **19**:1047–1068 DOI [10.2217/pgs-2018-0077](https://doi.org/10.2217/pgs-2018-0077).
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015.** limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**:e47 DOI [10.1093/nar/gkv007](https://doi.org/10.1093/nar/gkv007).
- Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. 2011.** pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**:77 DOI [10.1186/1471-2105-12-77](https://doi.org/10.1186/1471-2105-12-77).
- Sapkota S, Horiguchi K, Tosaka M, Yamada S, Yamada M. 2017.** Whole-exome sequencing study of thyrotropin-secreting pituitary adenomas. *Journal of Clinical Endocrinology and Metabolism* **102**:566–575 DOI [10.1210/jc.2016-2261](https://doi.org/10.1210/jc.2016-2261).
- Sarver AE, Sarver AL, Thayanithy V, Subramanian S. 2015.** Identification, by systematic RNA sequencing, of novel candidate biomarkers and therapeutic targets in human soft tissue tumors. *Laboratory Investigation* **95**:1077–1088 DOI [10.1038/labinvest.2015.80](https://doi.org/10.1038/labinvest.2015.80).
- Seeruttun SR, Cheung WY, Wang W, Fang C, Liu ZM, Li JQ, Wu T, Wang J, Liang C, Zhou ZW. 2019.** Identification of molecular biomarkers for the diagnosis of gastric cancer and lymph-node metastasis. *Gastroenterology Report* **7**:57–66 DOI [10.1093/gastro/goy023](https://doi.org/10.1093/gastro/goy023).
- Shahraz A, Kopatz J, Mathy R, Kappler J, Winter D, Kapoor S, Schutza V, Scheper T, Gieselmann V, Neumann H. 2015.** Anti-inflammatory activity of low molecular weight polysialic acid on human macrophages. *Scientific Reports* **5**:16800 DOI [10.1038/srep16800](https://doi.org/10.1038/srep16800).
- Smyth EC, Verheij M, Allum W, Cunningham D, Cervantes A, Arnold D, Committee EG. 2016.** Gastric cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* **27**:v38–v49 DOI [10.1093/annonc/mdw350](https://doi.org/10.1093/annonc/mdw350).
- Stratford JK, Bentrem DJ, Anderson JM, Fan C, Volmar KA, Marron JS, Routh ED, Caskey LS, Samuel JC, Der CJ, Thorne LB, Calvo BF, Kim HJ, Talamonti MS, Iacobuzio-Donahue CA, Hollingsworth MA, Perou CM, Yeh JJ. 2010.** A six-gene signature predicts survival of patients with localized pancreatic ductal adenocarcinoma. *PLOS Medicine* **7**:e1000307 DOI [10.1371/journal.pmed.1000307](https://doi.org/10.1371/journal.pmed.1000307).
- Sun YL, Zhang Y, Guo YC, Yang ZH, Xu YC. 2020.** A prognostic model based on the immune-related genes in colon adenocarcinoma. *International Journal of Medical Sciences* **17**:1879–1896 DOI [10.7150/ijms.45813](https://doi.org/10.7150/ijms.45813).
- Takeno A, Takemasa I, Doki Y, Yamasaki M, Miyata H, Takiguchi S, Fujiwara Y, Matsubara K, Monden M. 2008.** Integrative approach for differentially overexpressed genes in gastric cancer by combining large-scale gene expression profiling and network analysis. *British Journal of Cancer* **99**:1307–1315 DOI [10.1038/sj.bjc.6604682](https://doi.org/10.1038/sj.bjc.6604682).

- Tan X, Wang GB, Tang Y, Bai J, Ye L. 2017.** Association of ADIPOQ and ADIPOR variants with risk of colorectal cancer: a meta-analysis. *Journal of Huazhong University of Science and Technology* 37:161–171 DOI 10.1007/s11596-017-1710-3.
- Tan Y, Lin K, Zhao Y, Wu Q, Chen D, Wang J, Liang Y, Li J, Hu J, Wang H, Liu Y, Zhang S, He W, Huang Q, Hu X, Yao Z, Liang B, Liao W, Shi M. 2018.** Adipocytes fuel gastric cancer omental metastasis via PITPNC1-mediated fatty acid metabolic reprogramming. *Theranostics* 8:5452–5468 DOI 10.7150/thno.28219.
- Teng Y, Lang L, Shay C. 2019.** ATAD3A on the path to cancer. *Advances in Experimental Medicine and Biology* 1134:259–269 DOI 10.1007/978-3-030-12668-1_14.
- Teng Y, Ren X, Li H, Shull A, Kim J, Cowell JK. 2016.** Mitochondrial ATAD3A combines with GRP78 to regulate the WASF3 metastasis-promoting protein. *Oncogene* 35:333–343 DOI 10.1038/onc.2015.86.
- Waldum HL, Sagatun L, Mjones P. 2017.** Gastrin and gastric cancer. *Frontiers in Endocrinology* 8:1 DOI 10.3389/fendo.2017.00001.
- Wang J, Ni Z, Duan Z, Wang G, Li F. 2014.** Altered expression of hypoxia-inducible factor-1alpha (HIF-1alpha) and its regulatory genes in gastric cancer tissues. *PLOS ONE* 9:e99835 DOI 10.1371/journal.pone.0099835.
- Wang Q, Wen YG, Li DP, Xia J, Zhou CZ, Yan DW, Tang HM, Peng ZH. 2012.** Upregulated INHBA expression is associated with poor survival in gastric cancer. *Medical Oncology* 29:77–83 DOI 10.1007/s12032-010-9766-y.
- Wen F, Huang J, Lu X, Huang W, Wang Y, Bai Y, Ruan S, Gu S, Chen X, Shu P. 2020.** Identification and prognostic value of metabolism-related genes in gastric cancer. *Aging* 12:17647–17661 DOI 10.18632/aging.103838.
- Wickham H. 2016.** *ggplot2: elegant graphics for data analysis*. New York: Springer-Verlag.
- Wu S, Liu X, He J, Wang H, Luo Y, Gong W, Li Y, Huang Y, Zhong L, Zhao Y. 2019.** A dual targeting magnetic nanoparticle for human cancer detection. *Nanoscale Research Letters* 14:228 DOI 10.1186/s11671-019-3049-0.
- Xiao B, Kuang Z, Zhang W, Hang J, Chen L, Lei T, He Y, Deng C, Li W, Lu J, Qu J, Zhou Q, Hao W, Sun Z, Li L. 2019.** Glutamate Ionotropic Receptor Kainate Type Subunit 3 (GRIK3) promotes epithelial-mesenchymal transition in breast cancer cells by regulating SPDEF/CDH1 signaling. *Molecular Carcinogenesis* 58:1314–1323 DOI 10.1002/mc.23014.
- Xu GJ, Shah AA, Li MZ, Xu Q, Rosen A, Casciola-Rosen L, Elledge SJ. 2016.** Systematic autoantigen analysis identifies a distinct subtype of scleroderma with coincident cancer. *Proceedings of the National Academy of Sciences of the United States of America* 113:E7526–E7534 DOI 10.1073/pnas.1615990113.
- Yoon H, Kim N. 2015.** Diagnosis and management of high risk group for gastric cancer. *Gut Liver* 9:5–17 DOI 10.5009/gnl14118.
- Zeng X, Li C, Li Y, Yu H, Fu P, Hong HG, Zhang W. 2019.** A network-based variable selection approach for identification of modules and biomarker genes associated with end-stage kidney disease. *Nephrology* 25:775–784 DOI 10.1111/nep.13655.

- Zhang B, Horvath S. 2005.** A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology* **4**:Article17 DOI [10.2202/1544-6115.1128](https://doi.org/10.2202/1544-6115.1128).
- Zhang C, Liang Y, Ma MH, Wu KZ, Dai DQ. 2019.** KRT15, INHBA, MATN3, and AGT are aberrantly methylated and differentially expressed in gastric cancer and associated with prognosis. *Pathology, Research and Practice* **215**:893–899 DOI [10.1016/j.prp.2019.01.034](https://doi.org/10.1016/j.prp.2019.01.034).
- Zhang X, Yang JJ, Kim YS, Kim KY, Ahn WS, Yang S. 2010.** An 8-gene signature, including methylated and down-regulated glutathione peroxidase 3, of gastric cancer. *International Journal of Oncology* **36**:405–414.
- Zhao C, Li Y, Chen G, Wang F, Shen Z, Zhou R. 2017.** Overexpression of miR-15b-5p promotes gastric cancer metastasis by regulating PAQR3. *Oncology Reports* **38**:352–358 DOI [10.3892/or.2017.5673](https://doi.org/10.3892/or.2017.5673).
- Zhou XG, Huang XL, Liang SY, Tang SM, Wu SK, Huang TT, Mo ZN, Wang QY. 2018.** Identifying miRNA and gene modules of colon cancer associated with pathological stage by weighted gene co-expression network analysis. *OncoTargets and Therapy* **11**:2815–2830 DOI [10.2147/OTT.S163891](https://doi.org/10.2147/OTT.S163891).