Clinical Trial/Experimental Study

Medicine®

OPEN

PRRX1 promotes lymph node metastasis of gastric cancer by regulating epithelial-mesenchymal transition

Jibin Yao, MD[a], Yongbin Zhang, MD[a], Yu Xia, MD[b], Chenglou Zhu, MD[a], Xiaoxiong Wen, MD[c], Tianxiang Liu, MD[a], Mingxu Da, PhD[a,*]

Abstract

Background: Gastric cancer has multiple metastasis pathways, of which lymph node metastasis plays a dominant role. However, the specific mechanism of lymph node metastasis is still not unclear.

Methods: The bioinformatics technology was utilized to mine gene chip data related to gastric cancer and Epithelial-Mesenchymal Transition (EMT) in a high-throughput gene expression database (Gene Expression Omnibus, GEO), we screened out all genes that have differential expression levels in gastric cancer tissues and in adjacent normal gastric mucosa tissues. The corresponding function package of R language software were performed for gene annotation and cluster analysis, then enrichment analysis of genes with differential expression and protein interaction network diagram for correlation analysis were performed, we finally screened out the paired related homeobox 1 gene (PRRX1) related to EMT. Next, we collected 65 metastatic lymph node samples and 93 gastric cancer tissue samples. The expression levels of PRRX1 and EMT-related protein E-cadherin (E-ca) and vimentin (Vim) in gastric cancer tissues and metastatic lymph node tissues were determined by immunohistochemistry (IHC) staining of streptavidin-peroxidase (SP). The expression differences of PRRX1, E-ca and Vim in gastric cancer tissues and metastatic lymph node tissues as well as the correlation were analyzed by the experimental data, and the clinical significance was analyzed in combination with the clinicopathological data.

Results: The PRRX1 expression levels in gastric cancer tissues are significantly higher than that in adjacent normal gastric mucosa tissues. The positive expression rates of PRRX1, Vim and E-ca in gastric cancer and in metastatic lymph node tissues were significantly different. Comparing with that in gastric cancer, expression of PRRX1 and Vim was significantly down-regulated, and E-ca expression was significantly up-regulated in metastatic lymph nodes.

Conclusion: PRRX1 may promote lymph node metastasis of gastric cancer by regulating EMT, and then affect the prognosis of patients. PRRX1 may be used as a new biological indicator to predict or prevent lymph node metastasis in gastric cancer.

Abbreviations: E-ca = E-cadherin, EMT = epithelial-mesenchymal transition, GEO = Gene Expression Omnibus, LNM = lymph node metastasis, PRRX1 = paired related homeobox 1, Vim = vimentin.

Keywords: E-cadherin, epithelial-mesenchymal transition, gastric cancer, lymph node metastasis, PRRX1, vimentin

1. Introduction

As the fourth most common cancer worldwide, gastric cancer is still incurable and induces about 865,000 mortality each year.[1] What is worse, its poor prognosis is closely associated with lymph node metastasis of gastric cancer, whereas the effect of systemic chemotherapy is minimal. Thus, novel biomarkers are urgently needed to elucidate the specific mechanism of lymph node metastasis in gastric cancer so as to improve its prognosis.

Editor: Peeyush Goel.

JY, YZ, and YX contributed equally to this study.

This study was funded by the National Natural Science Foundation of China (no. 81560391).

The authors have no conflicts of interests to disclose.

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

[a] Department of Surgical Oncology, Gansu Provincial Hospital, Lanzhou, [b] The First Clinical Medical College, Lanzhou University, [c] Day Clinic, Gansu Provincial Maternal and Child-care Hospital, Lanzhou, China.

[*] Correspondence: Mingxu Da, Department of Surgical Oncology, Gansu Provincial Hospital, The First Clinical Medical College, Lanzhou University, Lanzhou 730000, Gansu Province, China (e-mail: hxdamingxu@hotmail.com).

How to cite this article: Yao J, Zhang Y, Xia Y, Zhu C, Wen X, Liu T, Da M. PRRX1 promotes lymph node metastasis of gastric cancer by regulating epithelial-mesenchymal transition. Medicine 2021;100:6(e24674).

Received: 8 February 2020 / Received in final form: 11 January 2021 / Accepted: 17 January 2021

http://dx.doi.org/10.1097/MD.0000000000024674

Many studies reported that the occurrence, development, metastasis and prognosis of various tumors are closely correlated with the EMT process.[2–5] Throughout EMT process, the morphology and characteristics of epithelial cells gradually transform into that of mesenchyme cells.[4] More specifically, it includes loss of cell polarity and tight junctions between epithelial cells, as well as the acquisition of adhesive junctions between mesenchyme cells and the ability to invade and migrate to other tissues.[5] The sign of the EMT process occurrence is marked by down-regulation of epithelial markers expression levels and up-regulation of mesenchyme markers expression levels, in which the main characteristic molecules are epithelial cadherin (E-ca) and vimentin (Vim).[6] PRRX1 has been reported as a new EMT inducer.[7] PRRX1 overexpression in breast cancer and colorectal cancer can induce EMT; however, the functional role of PRRX1 in different tumors is different. In colorectal cancer,[8] the high expression level of PRRX1 is closely related to increased metastasis and poor prognosis. The opposite relationship is observed in breast cancer.[7] The expression of PRRX1 in gastric cancer and its possible role in the occurrence of lymph node metastasis in gastric cancer is still unclear.

2. Materials and methods

2.1. Source of gene chip data

Gastric cancer and EMT-related chip datasets were retrieved on the National Center for Biotechnology Information (NCBI) website. Gene chip data were downloaded from the link (https://www.ncbi.nlm.nih.gov/gds/). Inclusion criteria for chip data: aiming at whole-genome sequencing data of gastric cancer tissues tissue samples and adjacent normal gastric mucosal tissue samples. Exclusion criteria: lack of control or too small sample size, poor chip quality or drug or other interventions. After the screening, the number GSE65801 (platform: GPL14550 Agilent-028004 SurePrint G3 Human GE 8x60K Microarray; Updated: October 11, 2016) chip data was selected for relevant data analysis in this study, including 32 gastric cancer tissue samples and 32 adjacent normal gastric mucosa tissue samples. All data are available online for free, and no author has ever experimented with animals or humans.

2.2. Screening of differential genes

Raw data of the GSE65801 chip from the NCBI website was downloaded, and the GEO2R online analysis platform (http://www.ncbi.nlm.nih.gov/geo/geo2r/) was utilized to draw a boxplot for chip quality evaluation. The "affy" package of R language was utilized to read the chip data, the "limma" package was utilized to extract the differential genes in the chip, and the "annotate" package was utilized to annotate the corresponding genes. According to the change value of expression multiple (logFC), the differentially extracted genes were divided into 3 types: $\log FC > 1.5$, $adj.p.Val < 0.01$ was up-regulated gene; $LogFC < -1.5$, $adj.p.Val < 0.01$ were down-regulated genes; others were normal.

2.3. GO enrichment analysis

The "clusterProfiler" package of R language was utilized to perform differential gene GO enrichment analysis. Gene ontology GO is widely utilized in the field of bioinformatics, which involves 3 aspects of biology: cellular component (CC), molecular function (MF), and biological process (BP). The differentially up-regulated genes and differentially down-regulated genes were screened for enrichment analysis in CC, MF, and BP. A "tiff" mapping device was utilized to map a network of up-regulated genes and GO relationships.

2.4. KEGG enrichment analysis

The "clusterProfiler" function package performs KEGG enrichment analysis on the differentially expressed genes obtained by screening, and annotates the pathways of the differential genes based on the *P* value. KEGG is a database for the systematic analysis of metabolic pathways of gene products in cells, focusing on intermolecular interactions and intermolecular networks. Biological metabolic pathways are divided into the following 6 categories: metabolism, genetic information processing, environmental information processing, cellular processes, biological systems, human diseases, and drug development. In the current study, we analyzed the integrated microarray gastric cancer data of significantly up-regulated and down-regulated differentially expressed genes, and considered that $P < .05$ was statistically significant.

2.5. Differential gene protein interaction analysis

The differential genes obtained from the above screening were selected according to *P* size. And the genes included in the study were imported into the STRING online tool (https://string-db.org/) for protein–protein interaction analysis (PPI). The STRING database contains a vast network of interactions between 9.6 million proteins and 13.8 million proteins in 2,031 species, enabling researchers to search for known and predicted protein interactions. Indirect functional correlation and direct physical interaction are 2 aspects of protein interactions, and analysis of protein interaction networks can help to study the molecular mechanism of disease pathogenesis from a systematic perspective and discover new targets for drug action. Then the top 100 genes of up-regulated genes and down-regulated genes were imported into the STRING database and a protein interaction network map was drawn. When the comprehensive score is greater than 0.4, the PPI was extracted, repeated interactions were removed, and then Cytoscape (http://www.cytoscape.org/) was utilized to analyze the PPI network topology. CytoHubba was used as a plug-in tool for Cytoscape to calculate the degree of each protein node, and a gene network with 10 or more genes in PPI is considered a pivot gene.

2.6. Patient selection and tissue preparation

From May 2018 to May 2019, 93 cases of primary tumor tissue of gastric cancer diagnosed at the Department of Oncology, Gansu Provincial People's Hospital (Lan Zhou, China), along with 65 cases of the corresponding metastatic lymph node tissue, were used in this study. The patients received neither chemotherapy nor radiotherapy before surgery. Cancer tissues and metastatic lymph node tissue were excised, fixed in 10% neutral-buffered formalin. The average age of the included cases was 43 years old, including 32 females and 61 males. Other clinicopathological characteristics are provided in Table 1. The maximum diameter of the tumor was used as the criterion to determine the tumor size, including 56 cases with ≥ 5 cm and 37 cases with < 5 cm. According to the depth of invasion of the

Table 1

Patients characteristics.

Patients characteristics	Frequency(n)	Percentage(%)
Age(years)		
>50	57	61.3
≤50	36	38.7
Gender		
Male	61	65.6
Female	32	34.4
Tumor size(cm)		
<5	37	39.8
≥5	56	60.2
Lauren's classification		
Mixture	31	33.3
Intestinal type	38	40.9
Diffuse type	24	25.8
Differentiation		
Poorly	43	46.2
Well/Moderately	50	53.8
Lymph node metastasis		
Yes	65	69.9
No	28	30.1
LNM stage		
N0	28	30.1
N1	21	22.6
N2	23	24.7
N3	21	22.6
TNM stage		
I	28	30.1
II	21	22.6
III	23	24.7
IV	21	22.6

tumor: 32 cases of T1-T2 and 61 cases of T3-T4. There were 43 cases of moderate-high differentiation, and 50 cases of low-undifferentiated. TNM staging: 39 cases of I+II and 54 cases of III +IV. Informed consent for tissue samples derived from patients, and the study was approved by the Research Ethics Committee of Gansu Provincial People's Hospital.

2.7. Immunohistochemistry

The expression of PRRX1, E-ca, and Vim proteins in gastric cancer tissues and metastatic lymph node tissues was determined by immunohistochemical streptavidin-peroxidase (SP) staining technique. The staining results of targeted proteins were observed under the microscope, and 5 cell fields were randomly selected at high magnification (×400), and the positive rate of staining cell was used as the scoring standard. The positive cell number score (0 points for 0%–5%; 1 point for 6%–25%; 2 points for 26%–50%; 3 points for >50%) and staining intensity score(0 points = unstained, 1 point = light yellow staining, 2 points = brownish yellow staining, 3 points = tan staining).

When the product of positive cell number score and staining intensity score is greater than 3, the section is considered to be a positive expression, and less than or equal to 3 is considered to be a negative expression. The product of PRRX1 positive cell score and staining degree score is less than or equal to 4, 6, and 9, which are considered to be low expression, moderate expression, high expression respectively.[9]

2.8. Statistical analysis

All statistical analyses were performed by SPSS v17.0 software.

The expression of PRRX1, E-ca, Vim, and their relationship with clinical pathological parameters was analyzed by χ^2 test in gastric cancer tissues and metastatic lymph nodes. Correlation of expression in metastatic lymph nodes was analyzed using spearman correlation analysis. The relationship between PRRX1 score and lymph node metastasis of gastric cancer was analyzed by rank sum test; independent factors of lymph node metastasis of gastric cancer were analyzed by multivariate logistic regression analysis. P values lower than .05 were considered statistically significant.

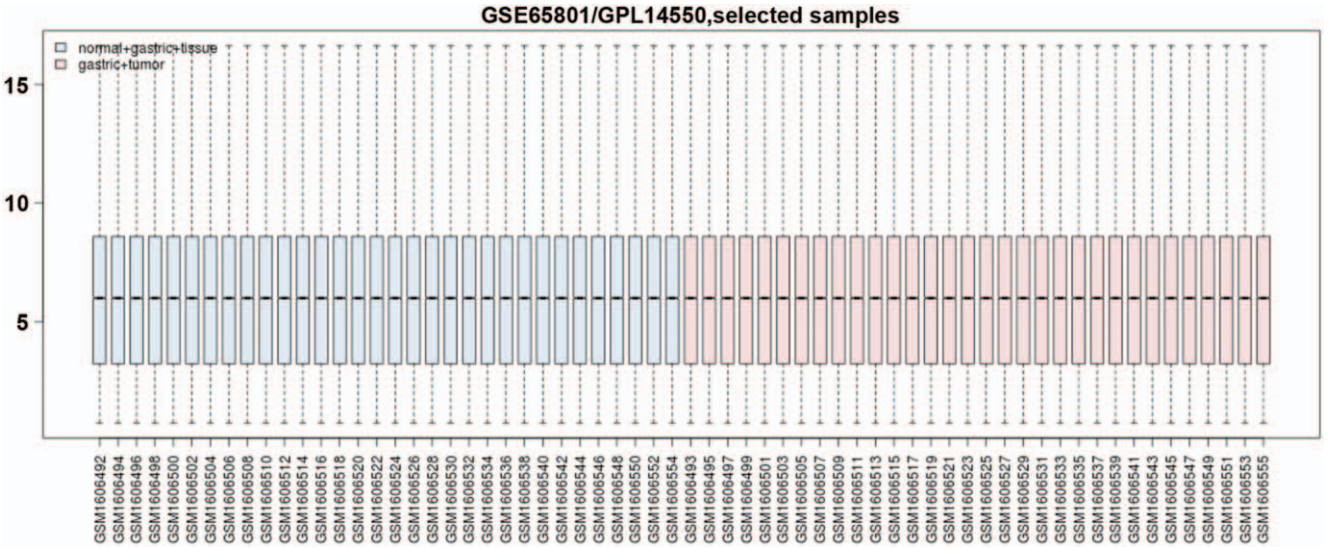

Figure 1. GSE65801 chip quality evaluation.

3. Results

3.1. Chip quality evaluation

According to the established inclusion and exclusion criteria, the GSE65801 chip data contained a total of 64 tissue samples, of which 32 were gastric cancer tissue samples and 32 adjacent normal gastric mucosa tissue samples. The online quality evaluation of GEO2R showed that the data of gastric cancer tissue samples and adjacent normal gastric mucosa samples were not significantly dispersed in the distribution, and the R language quality analysis report showed that no obvious plaques were seen in the chip scanned image (grayscale). After simple statistical analysis of the gray value, no obvious bias was found, nor was there any significant abnormality by RNA degradation analysis, as shown in Figure 1.

3.2. Differential gene screening

Based on logFC and *P* values, we finally screened out a total of 1143 differentially expressed genes, including 468 up-regulated

Table 2

Differential gene screening results.

Symbols	logFC	*t*	*P*.Value	adj.*P*.Val	type
CHSY3	2.660424425	9.780018216	2.35E-14	1.52E-10	up
MDFI	2.651448963	9.777464627	2.37E-14	1.52E-10	up
MFAP2	2.538537628	9.72756826	2.90E-14	1.52E-10	up
SLC39A10	1.557173366	9.669278805	3.65E-14	1.52E-10	up
SEMA5B	1.823538131	9.651708531	3.92E-14	1.52E-10	up
CHST1	2.093027359	9.304855365	1.57E-13	3.91E-10	up
BGN	2.704431125	9.257857772	1.89E-13	3.91E-10	up
INHBA	4.001728481	9.252987936	1.93E-13	3.91E-10	up
TIMP1	2.459945922	9.16402666	2.76E-13	4.70E-10	up
GABRD	1.790820744	9.128905463	3.18E-13	4.82E-10	up
HOXC9	4.805896082	9.121085142	3.28E-13	4.82E-10	up
KIAA1199	3.994276658	9.085412384	3.79E-13	5.07E-10	up
CCL26	2.179716388	9.052295409	4.33E-13	5.59E-10	up
NOTCH3	1.706174563	9.02776497	4.78E-13	5.74E-10	up
FOXS1	2.796109584	9.023900442	4.86E-13	5.74E-10	up
CLDN1	3.87136765	8.997389288	5.41E-13	6.20E-10	up
F2R	1.540058141	8.99139666	5.54E-13	6.20E-10	up
ANGPT2	2.016910659	8.939585082	6.83E-13	6.52E-10	up
RARRES1	3.238919031	8.931686414	7.05E-13	6.52E-10	up
BMP1	1.707134909	8.809842097	1.15E-12	1.00E-09	up
SULF1	3.101675388	8.799290745	1.20E-12	1.02E-09	up
MFI2	3.151526121	8.769134524	1.36E-12	1.09E-09	up
LZTS1	2.059319713	8.73384192	1.57E-12	1.21E-09	up
SERPINH1	1.909906594	8.663488226	2.09E-12	1.48E-09	up
PRRX1	2.659404784	7.476752539	2.60E-10	3.72E-08	up
CCKBR	−6.562700803	−10.87044817	3.26E-16	6.94E-12	down
GKN1	−9.312196287	−9.749416733	2.65E-14	1.52E-10	down
KCNE2	−7.473439434	−9.732908619	2.84E-14	1.52E-10	down
APOBEC2	−4.744765636	−9.715362624	3.04E-14	1.52E-10	down
PTGDR2	−4.150922469	−9.670936606	3.63E-14	1.52E-10	down
GRIA4	−4.664249023	−9.499514412	7.19E-14	2.55E-10	down
UBE2QL1	−2.761570278	−9.42691192	9.62E-14	3.15E-10	down
ALDH6A1	−1.528451672	−9.22501529	2.16E-13	4.11E-10	down
GCNT4	−2.291412156	−9.218424691	2.22E-13	4.11E-10	down
GKN2	−9.52385975	−9.163997783	2.76E-13	4.70E-10	down
SH3GL2	−4.255459516	−9.135285655	3.10E-13	4.82E-10	down
ADRB3	−1.541952941	−8.961590926	6.25E-13	6.52E-10	down
DNER	−4.086198063	−8.951458687	6.51E-13	6.52E-10	down
CCDC160	−2.554555506	−8.941980156	6.76E-13	6.52E-10	down
LIPF	−10.30206283	−8.821391724	1.10E-12	9.75E-10	down
ABCC8	−2.831949719	−8.780232144	1.30E-12	1.08E-09	down
ATP4A	−9.580547197	−8.740805885	1.52E-12	1.20E-09	down
PAIP2B	−2.434546688	−8.509667734	3.89E-12	2.30E-09	down
SLC1A2	−3.492330696	−8.450777855	4.94E-12	2.70E-09	down
DRD5	−5.54833483	−8.447069337	5.02E-12	2.70E-09	down
SLC7A8	−1.821403584	−8.357656813	7.21E-12	3.30E-09	down
PGA3	−8.595481369	−8.293372672	9.37E-12	3.90E-09	down
CA14	−2.466272575	−8.291253179	9.45E-12	3.90E-09	down
MICALL1	−1.733218081	−8.243873762	1.15E-11	4.43E-09	down
FAM99A	−1.780408316	−8.215526607	1.29E-11	4.84E-09	down

genes and 675 down-regulated genes. These up-regulated and down-regulated genes were sorted according to the expression levels and the adjusted *P* value, and the top 25 genes which expression were up-regulated were as follows: CHSY3, MDFI, MFAP2, SLC39A10, SEMA5B, CHST1, BGN, INHBA, TIMP1, GABRD, HOXC9, KIAA1199, CCL26, NOTCH3, FOXS1, CLDN1, F2R, ANGPT2, RARRES1, BMP1, SULF1, MFI2, LZTS1, SERPINH1, PRRX1; the top 25 genes which expression were down-regulated were as follows: CCKBR, GKN1, KCNE2, APOBEC2, PTGDR2, GRIA4, UBE2QL1, ALDH6A1, GCNT4, GKN2, SH3GL2, ADRB3, DNER, CCDC160, DNER, ABCC8, ATP4A, PAIP2B, SLC1A2, DRD5, SLC7A8, PGA3, CA14, MICALL1, FAM99A. The differential gene screening results are shown in Table 2. The heat map for the cluster analysis of differentially expressed genes is shown in Figure 2. The volcano map for the cluster analysis of differentially expressed genes is shown in Figure 3.

3.3. GO enrichment analysis of differential genes

Enrichment analysis of differentially expressed genes was performed in 3 aspects: cellular component (CC), molecular

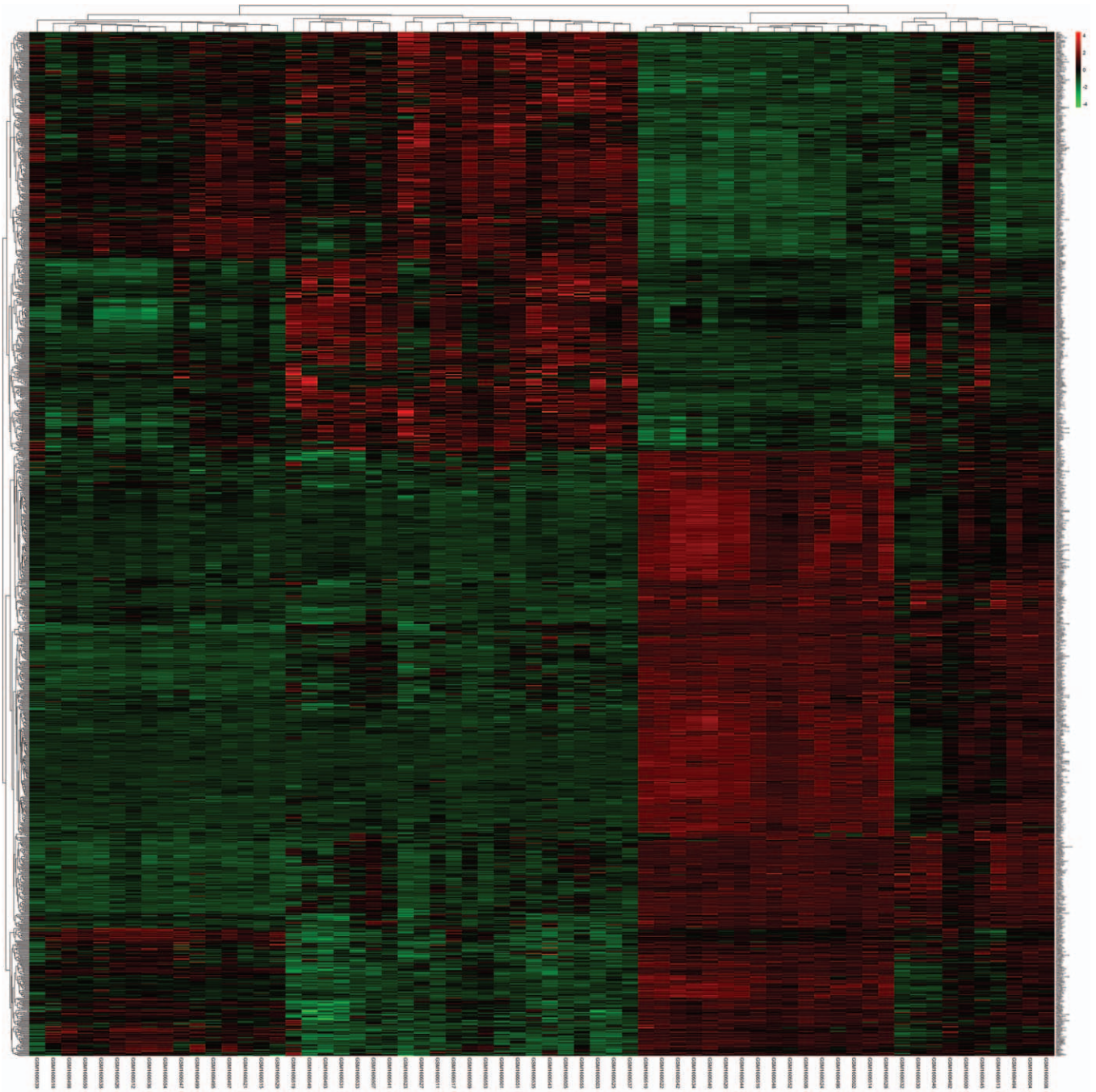

Figure 2. cluster analysis of differentially expressed genes.

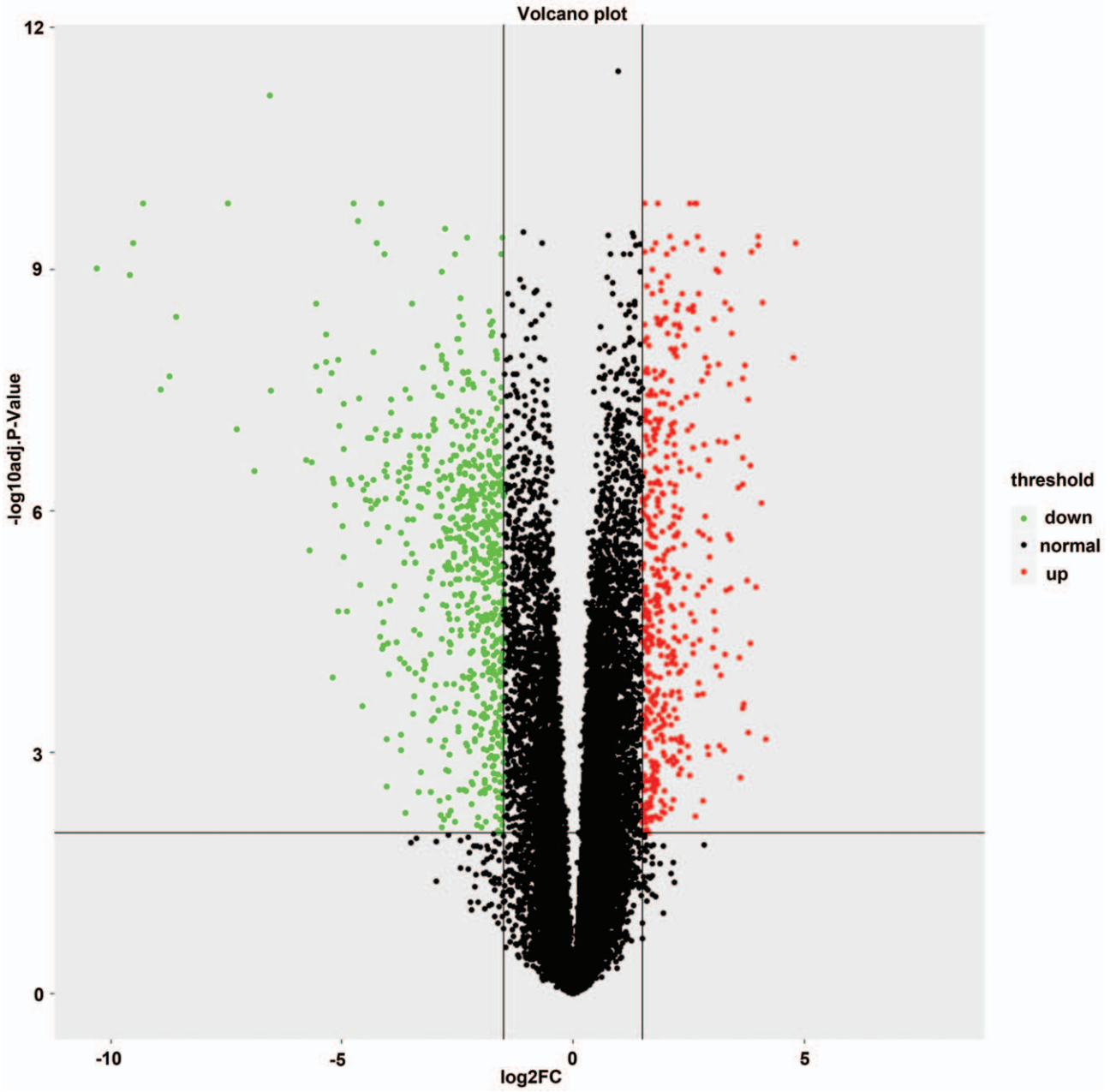

Figure 3. Volcano map of differentially expressed genes.

function (MF), and biological process (BP). The GO analysis results showed that differential genes were mainly enriched in biological processes, including extracellular matrix (ECM) tissue, endothelial cell differentiation, endothelial formation, wound healing, epidermal cell diffusion, and negative regulation of chemotaxis. GO enrichment analysis results are shown in Table 3 and the relationship between differentially expressed up-regulated genes and GO is shown in Figure 4. Molecular functional enrichment analysis showed that the differentially expressed genes were significantly enriched in collagen binding, peptidase activity, platelet-derived growth factor binding, binding between low-density lipoprotein particles, and serine peptidase activity. In terms of cell composition, differential genes were rich in endoplasmic reticulum lumen, platelet granule lumen, platelet granule, frill membrane, and filamentous pseudopodia. These above results suggested that most of the differential genes are involved in cell proliferation, cell cycle regulation and transcription.

3.4. Enrichment analysis of key KEGG pathway of differential genes

Through KEGG pathway enrichment analysis, we observed that the main biological pathways, in which the differential genes are involved included protein digestion and absorption, gastric acid secretion, calcium ion signaling pathway, pi3k-akt signaling

Table 3

GO enrichment analysis results.

ID	*P*.Value	*P*.adjust	q.value	Gene ID	type
GO_CC:0031012	1.99E-26	6.90E-24	5.96E-24	MFAP2, BGN, TIMP1, BMP1	up
GO_CC:0005578	7.92E-25	1.37E-22	1.19E-22	MFAP2, BGN, TIMP1, BMP1	up
GO_CC:0005581	1.40E-13	1.62E-11	1.40E-11	CTHRC1, MSR1, COL18A1, COL1A1	up
GO_CC:0044420	2.93E-11	2.54E-09	2.20E-09	MFAP2, TIMP1, COL18A1, COL1A1	up
GO_CC:0005788	7.53E-11	5.23E-09	4.52E-09	TIMP1, MELTF, SERPINH1, P4HA3	up
GO_MF:0005201	1.02E-13	6.11E-11	5.51E-11	BGN, ACAN, COL1A1, VCAN	up
GO_MF:0030545	2.32E-12	6.92E-10	6.23E-10	INHBA, TIMP1, CCL26, BMP1	up
GO_MF:0048018	6.21E-12	1.24E-09	1.11E-09	INHBA, TIMP1, CCL26, BMP1	up
GO_MF:0005125	3.96E-09	4.73E-07	4.26E-07	INHBA, TIMP1, CCL26, BMP1	up
GO_MF:0000982	0.000248242	0.006736376	0.006069449	FOXS1, PRRX1, DLX5, LEF1	up
GO_BP:0040017	4.24E-10	1.54E-07	1.22E-07	CEMIP, CCL26, F2R, THY1	up
GO_BP:0048736	1.08E-09	3.32E-07	2.64E-07	CRABP2, PRRX1, GJA5, DLX5	up
GO_BP:0060173	1.08E-09	3.32E-07	2.64E-07	CRABP2, PRRX1, GJA5	up
GO_BP:0051216	2.23E-09	6.35E-07	5.04E-07	TIMP1, BMP1, SULF1, PRRX1	up
GO_BP:0061448	2.47E-09	6.57E-07	5.22E-07	TIMP1, BMP1, SULF1, PRRX1	up
GO_CC:0016324	4.49E-07	0.000115073	0.000106914	GIF, MAL, P2RX2, SCNN1B	down
GO_CC:0045177	8.23E-07	0.000115073	0.000106914	HOMER2, GIF, MAL, P2RX2	down
GO_CC:0016323	1.11E-06	0.000115073	0.000106914	CHRM3, PROM2, SLC4A4, RAB17	down
GO_CC:0034702	1.13E-06	0.000115073	0.000106914	CLCNKA, GRIA2, SCNN1B, KCNMB2	down
GO_CC:1902495	1.74E-06	0.000142625	0.000132512	KCNE2, GRIA4, ABCC8, CLCNKA	down
GO_MF:0022838	9.20E-12	6.65E-09	5.63E-09	GRIA4, ABCC8, TMEM38A, CLCNKA	down
GO_MF:0022839	3.30E-11	9.92E-09	8.39E-09	HTR3E, KCNJ15, GABRB3, KCNK10	down
GO_MF:0022836	4.30E-11	9.92E-09	8.39E-09	HPN, KCNJ11, KCNJ15, GABRB3	down
GO_MF:0015267	6.39E-11	9.92E-09	8.39E-09	ABCC8, TMEM38A, CLCNKA, AQP4	down
GO_MF:0022803	6.86E-11	9.92E-09	8.39E-09	ABCC8, TMEM38A, CLCNKA, AQP4	down
GO_BP:0007586	6.57E-10	2.45E-06	2.37E-06	CCKBR, GKN1, PGA3, CHIA	down
GO_BP:0015698	2.77E-07	0.00035125	0.000338571	ABCC8, CLCNKA, SLC5A5, GABRG2	down
GO_BP:0071804	3.76E-07	0.00035125	0.000338571	KCNE2, ABCC8, ATP4A, TMEM38A	down
GO_BP:0071805	3.76E-07	0.00035125	0.000338571	ATP4A, TMEM38A, NEDD4L, ATP4B	down
GO_BP:0042403	1.74E-06	0.001301496	0.001254516	GCNT4, SLC5A5, DUOXA1, DUOX1	down

pathway, complement, and coagulation cascade reaction. The key pathway enrichment results of KEGG differential genes are shown in Table 4.

3.5. PPI protein interaction network analysis

Top 100 up-regulated and down-regulated genes were imported into the online website of STRING for PPI protein interaction network analysis. The results of PPI analysis of up-regulated genes showed that the key genes which play a pivotal regulatory role in all up-regulated genes are as follows: HOC8, PRRX1, DLX5, FOXC1, PLAU, MFAP2, ADAM12, TNFSF11, BGN, CLDN1, INHBA, NOTCH3, LOX, ANGPT2, SULF1, THBS2, TIMP1, COL1A1, WNT2, VCAN, EPHB2, COL18A1, ACAN, BMP1, PGF, SFRP4. The results of PPI analysis of down-regulated genes showed that the key genes which play a pivotal role in all down-regulated genes are as follows: SLC26A7, SLC7A8, HEPH, ATP8A2, AGXT2L1, PGA3, PTGDR2, ATP4B, SLC1A2, ABCC8, EGFR, RET, GIF, SLC5A5, AQP4, CLCNKA, GABRG2, GRIA2, NEDD4L, KCNE2, SLC2A12, GPER, ESRRG. As shown in Figures 5 and 6.

3.6. The expression of PRRX1, E-ca, and Vim

The results suggested that the chromogenic expression of PRRX1 protein appears in the nucleus, and the chromogenic expression of E-ca and Vim both appear in the cell membrane and/or cytoplasm; The positive expression rates of PRRX1 and Vim both in gastric cancer tissues and metastatic lymph node tissues are significantly different, and the positive expression rates of PRRX1 and Vim are respectively 69.89% and 63.44% in gastric cancer tissues, respectively 43.08% and 36.92% in metastatic lymph nodes. The positive expression rates of PRRX1 and Vim in gastric cancer tissues are significantly higher than that in metastatic lymph node tissues; however the positive expression rate of E-ca in gastric cancer tissues is significantly lower than that in metastatic lymph nodes; the positive expression rate of E-ca in gastric cancer tissues and metastatic lymph nodes is 44.09% and 61.54%, respectively. ($P < .05$), as shown in Figure 7.

3.7. Correlation of expression of PRRX1, E-ca, and Vim in metastatic lymph node tissues

There were 17 cases of PRRX1 and Vim positive expression, 11 cases of PRRX1 positive expression and Vim negative expression, 30 cases of PRRX1 and Vim negative expression, and 7 cases of PRRX1 negative expression and Vim positive expression in metastatic lymph node tissues samples. Statistical analysis suggested that PRRX1 expression was positively correlated with Vim expression in metastatic lymph node tissue samples ($r = 0.429$, $P < .001$). There were 12 cases of PRRX1 and E-ca both positive expression, 16 cases of PRRX1 positive expression and E-ca negative expression, 9 cases of PRRX1 and E-ca both negative expression, and 28 cases of PRRX1 negative expression and E-ca positive expression in metastatic lymph node tissues

samples. Statistical analysis suggested that PRRX1 expression is negatively correlated with E-ca expression in metastatic lymph node tissue samples ($r = -0.334$, $P = .007$).

3.8. Relationship between clinical pathological parameters and expression of PRRX1, E-ca, and Vim

The expressions of PRRX1, E-ca and vim in metastatic lymph node tissue are closely correlated with the depth of tumor invasion, tumor differentiation, tumor pathological stage, and neurovascular invasion ($P < .05$), as shown in Table 5.

3.9. Relationship between PRRX1 protein expression and lymph node metastasis in gastric cancer

According to the separate score of PRRX1, all the included samples were divided into low expression group, moderate expression group, and high expression group. The final results showed that the positive rates in the low expression group, moderate expression group, and high expression group of metastatic lymph nodes were 40.7%, 66.7%, and 90.5%, respectively. This data suggested that when PRRX1 expression level is up-regulated, the positive rate of lymph nodes also suggests an increasing trend, as shown in Table 6. Furthermore,

Figure 4. Network diagram of up-regulated genes and GO.

Table 4

KEGG pathway enrichment analysis results.

ID	Description	*P*.Value	*P*.adjust	Count
hsa04974	Protein digestion and absorption	7.40E-09	2.03E-06	22
hsa04971	Gastric acid secretion	0.000106665	0.006059708	14
hsa04020	Calcium signaling pathway	0.000110177	0.006059708	25
hsa04151	PI3K-Akt signaling pathway	0.000233556	0.009175396	38
hsa04610	Complement and coagulation cascades	0.000674746	0.020617253	13

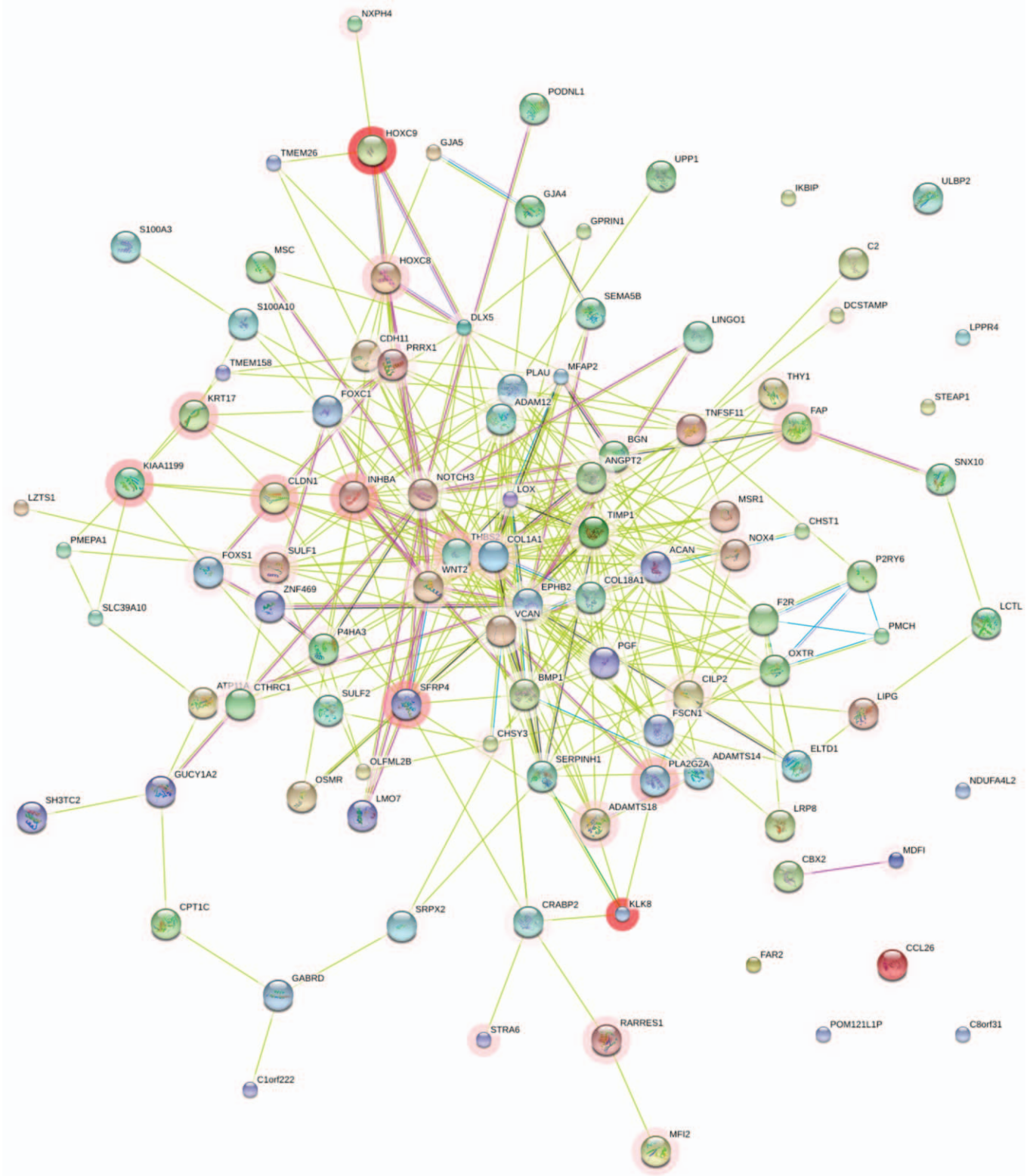

Figure 5. Topological diagram of PPI network for up-regulated genes.

multivariate logistic regression analysis revealed independent risk factors for distant lymph node metastasis in gastric cancer, including PRRX1 independent score, tumor sizes, and depth of invasion, as shown in Table 7.

4. Discussion

Lymph node metastasis plays a dominant role in the metastasis of gastric cancer, and most patients with a poor prognosis are mostly associated with early lymph node metastasis.[10] There-

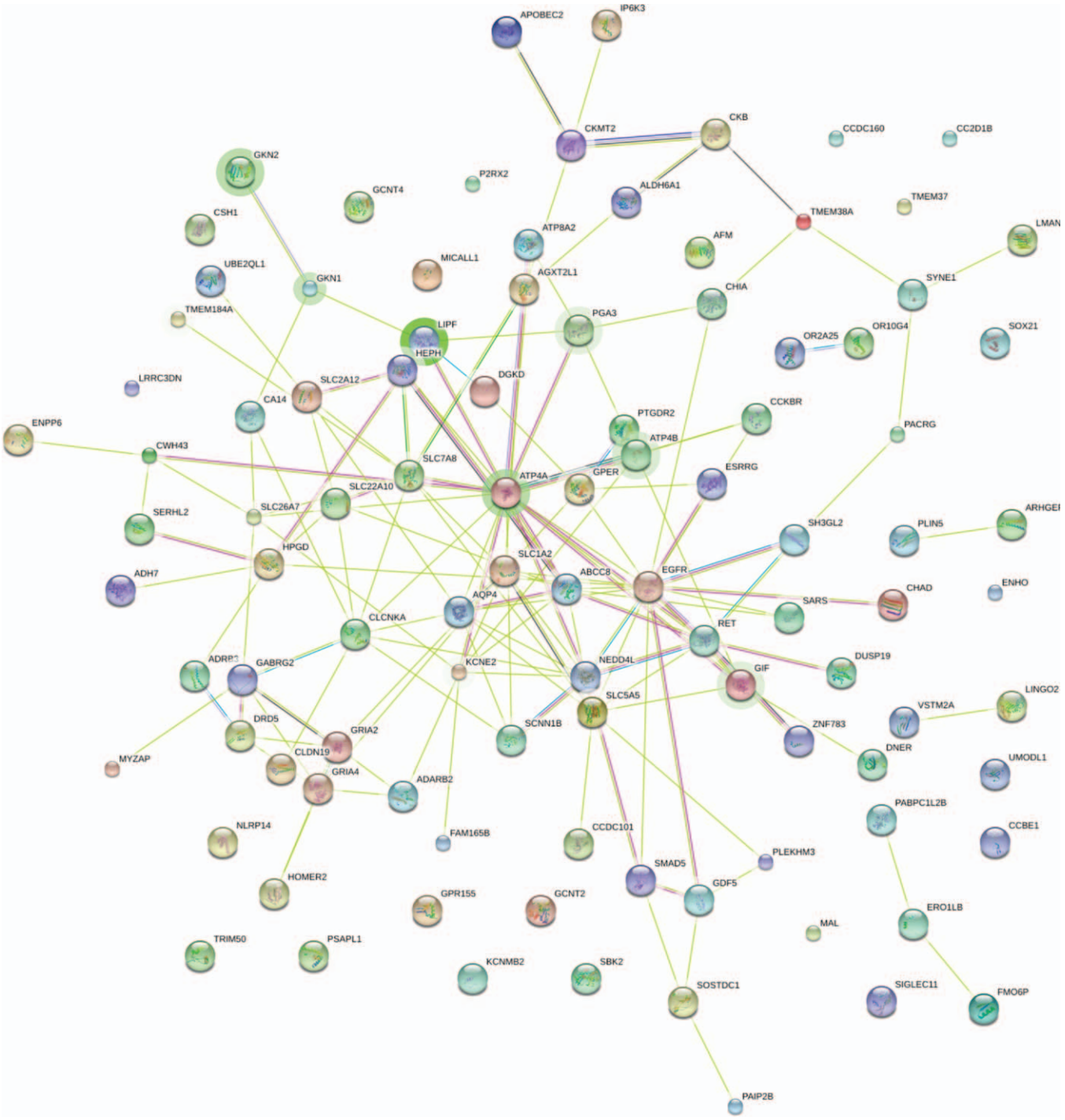

Figure 6. Topological diagram of PPI network for down-regulated genes.

fore, elucidating the mechanism of lymph node metastasis in gastric cancer will provide new ideas for the study of specific diagnosis and treatment of gastric cancer, so as to improve its survival rate and prognosis. Although increasing studies have been focused on the mechanism of lymph node metastasis in gastric cancer, it is still largely unclear.

In this study, we conducted differential screening of genome-wide expression profile data of 32 gastric cancer samples and 32 normal gastric mucosal tissue samples from GEO database. Then GO enrichment analysis was performed on the differential genes, which were divided into MF group, BP group and CC group. We observed that genes with significantly increased expression in gastric cancer patients were mainly enriched in the biological processes and signaling pathways. The former mainly included endothelial cell differentiation, endothelial formation, wound healing, epidermal cell diffusion and negative regulation chemotaxis; and the latter included protein digestion and absorption, gastric acid secretion, the calcium ion signaling,

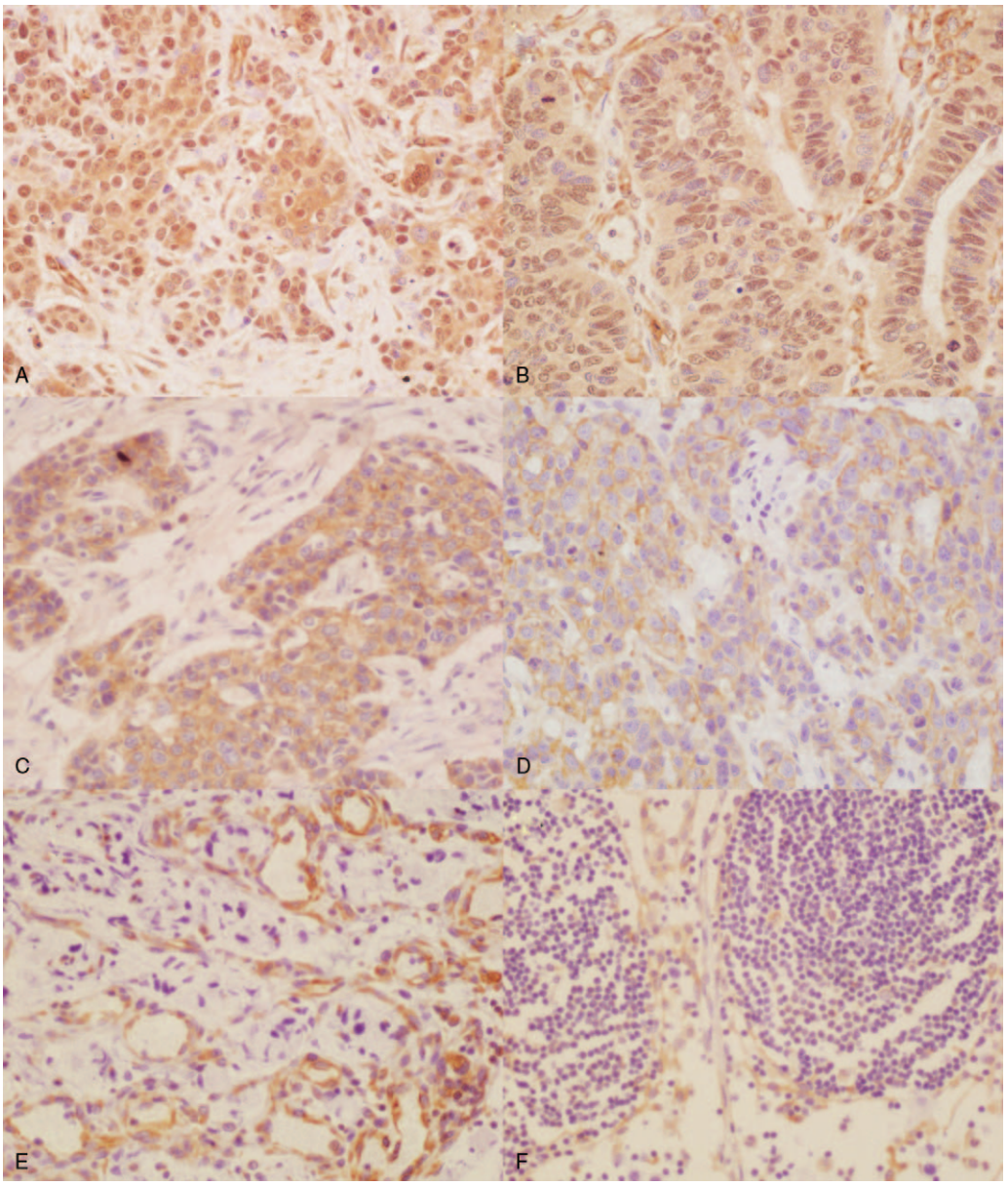

Figure 7. expression levels of PRRX1, E-ca, Vim by immunohistochemistry (×400). A, expression of PRRX1 in gastric cancer tissues. B, expression of PRRX1 in metastatic lymph node tissues. C, expression of E-ca in gastric cancer tissues. D, expression of E-ca in metastatic lymph node tissues. E, expression of Vim in gastric cancer tissues. F, expression of Vim in metastatic lymph node tissues.

pi3k-akt signaling, complement and coagulation cascade. By constructing a PPI network encoding differentially expressed genes, the targeted gene PRRX1 related to EMT was screened out according to the correlation. The differential screening results suggested that PRRX1 expression level in gastric cancer tissue is 2.65 times higher than that in adjacent normal gastric mucosa tissues, and the enrichment analysis results suggested that the gene has transcription factor activity and participates in the process of RNA polymerase II proximal promoter sequence binding to specific DNA. The pair-related homeobox 1 (PRRX1) is a member of the homeobox family of proteins, located in the 1q24 region of the chromosome, which encodes a protein that is linked to DNA and located in the nucleus. It has been reported that PRRX1 is a novel transcription factor, which can induce EMT.[7,8] PRRX1 may initiate EMT through Wnt/β-catenin pathway in gastric cancer, and the expression levels of epithelial-mesenchymal markers may change due to PRRX1 expression.[11] Relevant studies have reported that EMT occurs in breast cancer cells when PRRX1 expression level is up-regulated, and at the same time, the dryness of breast cancer cells is down-regulated.[12] Furthermore, it is reported that PRRX1 expression upregulation in triple negative breast cancer corresponds to the enhancement of cancer cell invasion, proliferation and metastasis, which are mainly manifested as the invasion of nerve, blood vessels and lymphatic vessels. However, the ability of cancer cells to invade, proliferate and migrate is significantly reduced when PRRX1 expression level is down-regulated.[13]

E-ca is a calcium-dependent transmembrane glycoprotein molecule that exists on the surface of epithelial cells. Its main function is to maintain the morphological stability of epithelial

Table 5

Relationship between clinical pathological parameters and expression of PRRX1, E-ca, and Vim.

	PRRX1			E-ca			Vim		
	+	−	*P*	+	−	*P*	+	−	*P*
Patients	52	13		17	48		48	17	
Age (years)									
>50	35	9	.894	12	32	.766	33	11	.759
≤50	17	4		5	16		15	6	
Gender									
Male	35	8	.694	12	31	.653	32	11	.883
Female	17	5		5	17		16	6	
Tumor size(cm)									
<5	22	6	.802	12	32	.766	20	8	.700
≥5	30	7		5	16		28	9	
Tumor location									
Upper third	14	3	.954	4	13	.924	13	4	.924
Middle third	16	4		5	15		15	5	
Lower third	22	6		8	20		20	8	
Lauren's classification									
Mixture	20	5	.986	7	18	.899	19	6	.654
Intestinal type	21	5		6	20		20	6	
Diffuse type	11	3		4	10		9	5	
Differentiation									
Poorly	38	5	.018	10	33	.457	31	12	.653
Well/Moderately	14	8		7	15		17	5	
Depth of tumor invasion									
T1/T2	13	8	.012	12	9	<.001	9	12	<.001
T3/T4	39	5		5	39		39	5	
TNM stage									
I/II	15	9	.007	10	14	.029	14	10	.029
III/IV	37	4		7	34		34	7	
Lymphatic invasion									
Negative	16	9	.011	14	11	<.001	15	10	.045
Positive	36	4		3	37		33	7	
Vascular invasion									
Negative	9	8	.001	10	7	<.001	9	8	.022
Positive	43	5		7	41		39	9	
Nerve invasion									
Negative	13	9	.003	12	10	<.001	11	11	.002
Positive	39	4		38			37	6	

cells, to mediate cellular adhesion, cell polarity and tissue structural integrity of homogenous cells. In the EMT process, E-ca expression levels change so as to mediate tumor metastasis and colonization.

Increasing evidence suggested that E-ca expression levels in lung adenocarcinoma,[14] colorectal cancer,[15] breast cancer[16] and other tumor tissues are significantly down-regulated, leading to the enhancement of metastatic ability of tumor cells.

Vim is a type III intermediate filament that maintains cell integrity and involved in cell migration, motility, and adhesion, which overexpression is known to be associated with aggressive disease and worse outcomes in solid cancers such breast, gastrointestinal, and prostate cancers.[17–19]

Given the relevance of EMT-related protein E-ca and Vim to cancer metastasis, we examined the expressions of PRRX1, E-ca, and Vim by immunohistochemistry in gastric cancer tissues and

Table 6

Relationship between PRRX1 protein expression and lymph node metastasis in gastric cancer.

PRRX1	All samples	Low expression	Moderate expression	High expression	*P*
All samples	93	27	24	42	<.001
Positive	65 (69.9%)	11 (40.7%)	16 (66.7%)	38 (90.5%)	
N1	21 (22.58%)	6 (22.22%)	7 (29.17%)	8 (19.05%)	
N2	23 (24.73%)	4 (14.81%)	7 (29.17%)	12 (28.57%)	
N3	21 (22.58%)	1 (3.70%)	2 (8.33%)	18 (42.86%)	
N3a	20 (21.51%)	1 (3.70%)	2 (8.33%)	17 (40.48%)	
N3b	1 (1.08%)	0	0	1 (2.38%)	
Negative	28 (30.10%)	16 (59.3%)	8 (33.3%)	4 (9.5%)	

N1: 1–2 metastasis lymph nodes. N2: 3–6 metastasis lymph nodes. N3: ≥7 metastasis lymph nodes N3a: 7–15 metastasis lymph nodes. N3b: ≥16 metastasis lymph nodes.

Table 7

Multivariate analysis of independent risk factors for lymph node metastasis in gastric cancer.

	Total	OR	95%CI	P
Age(years)				
≤50	36	1 (reference)	–	–
<50	57	1.147	[0.938–1.267]	.059
Gender				
Male	61	1 (reference)	–	–
Female	32	0.926	[0.687–1.247]	.178
Tumor size(cm)				
<5	37	1 (reference)	–	–
≥5	56	1.223	[1.119–1.335]	.009
Tumor invasion depth				
T1/T2	32	1 (reference)	–	–
T3/T4	61	1.674	[1.029–2.721]	.038
Lauren classification				
Mixture	31	1 (reference)	–	–
Intestinal type	38	1.164	[0.80–2.459]	.209
Diffuse type	24	1.406	[0.826–1.640]	.145
PRRX1 expression levels				
Low	27	1 (reference)	–	–
Moderate	24	1.317	[1.010–1.715]	.047
High	42	2.807	[1.874–4.203]	<.001

metastatic lymph node tissues, and then analyzed the relationship between clinical pathological parameters and expression of PRRX1, E-ca and Vim, and the relationship between PRRX1 expression and lymph node metastasis in gastric cancer.

Our results demonstrate that the expression levels of PRRX1, E-ca, and Vim are closely correlated with the degree of tumor differentiation, the depth of infiltration, TNM stage, lymphatic invasion, vascular invasion, and nerve invasion ($P < .05$). Furthermore, we found by data analysis that PRRX1 is positively correlated with Vim expression, and PRRX1 expression is negatively correlated with E-ca expression in metastatic lymph node tissues. This study further confirms that PRRX1, as an independent risk factor, is significantly correlated with lymph node metastasis status in gastric cancer. Differential expression levels of PRRX1, E-ca, and Vim in gastric cancer tissues and metastatic lymph node tissues suggested that PRRX1, E-ca, and Vim may be involved in lymph node metastasis of gastric cancer. Thus, we can speculate that the dynamic transformation of EMT-MET may be involved in the whole process of gastric cancer from the primary tumor to lymph node metastasis. Furthermore, this study reveals that there is a significant correlation between lymph node metastasis of gastric cancer and PRRX1 expression.

In conclusion, this study reports that PRRX1 expression may promote lymph node metastasis of gastric cancer by regulating EMT. PRRX1 stimulates gastric cancer invasion and metastasis, therefore contributing to poor prognosis.[20] PRRX1 may be used as a new biological indicator to predict the possibility of lymph node metastasis of gastric cancer. However, the possible regulatory relationship of EMT in lymph node metastasis of gastric cancer and the specific mechanism of action remain unclear, and the further experimental researches are needed.

Author contributions

Conceptualization: Chenglou Zhu, Mingxu Da.
Data curation: Yongbin Zhang, Mingxu Da.
Formal analysis: Tianxiang Liu.
Methodology: Mingxu Da.
Resources: Yongbin Zhang, Yu Xia, Chenglou Zhu, Xiaoxiong Wen.
Software: Jibin Yao, Yongbin Zhang, Yu Xia, Chenglou Zhu, Xiaoxiong Wen, Tianxiang Liu.
Supervision: Chenglou Zhu, Mingxu Da.
Visualization: Yu Xia, Chenglou Zhu, Tianxiang Liu, Mingxu Da.
Writing – original draft: Yongbin Zhang, Yu Xia, Mingxu Da.
Writing – review & editing: Yu Xia, Chenglou Zhu, Xiaoxiong Wen.

References

[1] Petrillo A, Smyth EC. 27 years of stomach cancer: painting a global picture. Lancet Gastroenterol Hepatol 2020;5:5–6.
[2] Savagner P. Epithelial-mesenchymal transitions: from cell plasticity to concept elasticity. Curr Top Dev Biol 2015;112:273–300.
[3] Murai T, Yamada S, Fuchs BC, et al. Epithelial-to-mesenchymal transition predicts prognosis in clinical gastric cancer. J Surg Oncol 2014;109:684–9.
[4] Heerboth S, Housman G, Leary M, et al. EMT and tumor metastasis. Clin Transl Med 2015;4:1–13.
[5] Jolly MK, Ware KE, Gilja S, et al. EMT and MET: necessary or permissive for metastasis? Mol Oncol 2017;11:755–69.
[6] Busch EL, McGraw KA, Sandler RS. The potential for markers of epithelial-mesenchymal transition to improve colorectal cancer outcomes: a systematic review. Cancer Epidemiol Biomarkers Prev 2014;23:1164–75.
[7] Ocana OH, Corcoles R, Fabra A, et al. Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. Cancer Cell 2012;22:709–24.
[8] Takahashi Y, Sawada G, Kurashige J, et al. Paired related homoeobox 1, a new EMT inducer, is involved in metastasis and poor prognosis in colorectal cancer. Br J Cancer 2013;109:307–11.
[9] Lee HH, Lee SH, Song KY, et al. Evaluation of Slug expression is useful for predicting lymph node metastasis and survival in patients with gastric cancer. BMC Cancer 2017;17:670.
[10] Wu L, Liang Y, Zhang C, et al. Prognostic significance of lymphovascular infiltration in overall survival of gastric cancer patients after surgery with curative intent. Chin J Cancer Res 2019;31:785–96.
[11] Guo J, Fu Z, Wei J, et al. PRRX1 promotes epithelial-mesenchymal transition through the Wnt/beta-catenin pathway in gastric cancer. Med Oncol 2015;32:393.
[12] Zhu H, Sun G, Dong J, et al. The role of PRRX1 in the apoptosis of A549 cells induced by cisplatin. Am J Transl Res 2017;9:396–402.
[13] Lv ZD, Kong B, Liu XP, et al. miR-655 suppresses epithelial-to-mesenchymal transition by targeting Prrx1 in triple-negative breast cancer. J Cell Mol Med 2016;20:864–73.
[14] Zhang Y, Sun L, Gao X, et al. RNF43 ubiquitinates and degrades phosphorylated E-cadherin by c-Src to facilitate epithelial-mesenchymal transition in lung adenocarcinoma. BMC Cancer 2019;19:1–18.
[15] Palaghia M, Mihai C, Lozneanu L, et al. E-cadherin expression in primary colorectal cancer and metastatic lymph nodes. Rom J Morphol Embryol 2016;57:205–9.
[16] Tavakolian S, Goudarzi H, Faghihloo E. E-cadherin, Snail, ZEB-1, DNMT1, DNMT3A and DNMT3B expression in normal and breast cancer tissues. Acta Biochim Pol 2019;66:409–14.
[17] Zhu QS, Rosenblatt K, Huang KL, et al. Vimentin is a novel AKT1 target mediating motility and invasion. Oncogene 2011;30:457–70.
[18] Kidd ME, Shumaker DK, Ridge KM. The role of vimentin intermediate filaments in the progression of lung cancer. Am J Respir Cell Mol Biol 2014;50:1–6.
[19] Sun BO, Fang Y, Li Z, et al. Role of cellular cytoskeleton in epithelial-mesenchymal transition process during cancer progression. Biomed Rep 2015;3:603–10.
[20] Otsuki S, Inokuchi M, Enjoji M, et al. Vimentin expression is associated with decreased survival in gastric cancer. Oncol Rep 2011;25:1235–42.