

The Comprehensive Analysis of Competitive Endogenous RNA Networks and Tumor-Infiltrating Immune Cells in Gastric Cancer With Lymph Node Metastasis

Fangxiao Gong

Department of Intensive Care Unit, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou, Zhejiang, China.

Tianyu Liang

Department of Intensive Care Unit, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou, Zhejiang, China.

Zongbin Lin

Department of Intensive Care Unit, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou, Zhejiang, China.

Renyang Liu

Department of Intensive Care Unit, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou, Zhejiang, China.

Jun Hong

Department of Intensive Care Unit, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou, Zhejiang, China.

Minhua Chen (✉ zjpacmh@126.com)

Department of Intensive Care Unit, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou, Zhejiang, China.

Research Article

Keywords: gastric cancer, lymph node metastasis, competing endogenous RNA, tumor-infiltrating immune cell, predicting

Posted Date: December 8th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1133390/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Gastric cancer is a kind of tumor with strong heterogeneity. Long non-coding RNAs (lncRNAs) acting as competing endogenous RNAs (ceRNAs) play significant roles in the development of tumors.

Methods: In this study, we divided all TCGA gastric cancer patients into the whole, intestinal and diffuse cohorts for further analysis, and constructed competitive endogenous RNA network and evaluated immune cells using CIBERSORTx. The support vector machines recursive feature elimination (SVM-RFE) was used for screening significant signatures and the support vector machines (SVM) for establishing model predicting the lymph node metastasis.

Results: In this study, we divided all TCGA gastric cancer patients into the whole, intestinal and diffuse cohorts for further analysis, and constructed competitive endogenous RNA network and evaluated immune cells using CIBERSORTx. The support vector machines recursive feature elimination (SVM-RFE) was used for screening significant signatures and the support vector machines (SVM) for establishing model predicting the lymph node metastasis. The performance of SVM model was good in the intestinal and diffuse cohort, while the model in the whole cohort was relatively poor. Some important co-expression patterns between immune cells and ceRNAs network indicated significant correlation CD70 with dendritic cells and so on.

Conclusion: Our research inferred competing endogenous RNA network of lymph node metastasis and built an excellent predicting model.

Introduction:

Gastric cancer (GC) is the sixth most common cancer and the second leading cause of cancer-related deaths worldwide¹. Worldwide mortality rates for GC have declined in the past 10 years, however the survival rate remains low². Many clinical, molecular, and pathologic data suggest that GC is a heterogeneous disease³. It's urgent to investigate the underlying mechanism by analyzing different groups of patients. The lymph node metastasis is the most common metastasis pattern of GC. According to the Lauren classification, there are the intestinal-type GC that is associated with lymphatic or vascular invasion and the lesions are scattered in distant positions and the diffuse-type GC that have the characteristic of non-cohesive, scattered tumor cells. So, the Lauren classification is an excellent classification method for our research.

The competitive endogenous RNA (ceRNA) hypothesis was first proposed by Salmena L, Poliseno L, et al in 2011⁴. In the ceRNA network which includes long noncoding RNA (lncRNA), microRNA (miRNA), and messenger RNA (mRNA), lncRNA can act as endogenous molecular sponges that competitively bind miRNA via shared microRNA response elements (MREs) with reverse complementary binding seed regions to indirectly regulate mRNA expression levels. In recent years, many studies have validated the

ceRNA regulatory mechanism is involved in many carcinoma^{5 6 7}. Not only tumor cells but also tumor-infiltrating immune cells participate in lymph node metastasis⁸.

In this study, based on the whole cohort, the intestinal- and the diffuse cohort, we constructed the ceRNA network respectively and evaluate the immune cells fraction using the CIBERSORTx algorithm⁹. The support vector machines recursive feature elimination (SVM-RFE) algorithm was applying for selected lymph node metastasis associated RNA and tumor-infiltrating immune cells. Moreover, we built a predicted model for lymph node metastasis using the support vector machines (SVM). The flowchart of this research is shown in Figure 1.

Result:

Data set acquisition and identification of differentially expressed RNAs: We had 14,358 lncRNAs, 2,177 miRNAs and 19,581 mRNAs from the TCGA database and 357 GC patients, 157 intestinal GC patients and 63 diffuse GC patients after removing the patients without N stage. The baseline characteristics of three patient cohorts available from the TCGA are described in Table 1. We identified 68 DEmRNAs and 80 DElncRNAs in the whole patient cohort, 247 DEmRNAs and 58 DElncRNAs in the intestinal patient cohort and 46 DEmRNAs and 16 DElncRNAs in the diffuse patient cohort.

Table 1. Basic characteristics of the datasets.

Variables	the whole cohort	the intestinal cohort	the diffuse cohort
	n=357	n=157	n=63
Age (Mean ± SD)	65.88±10.67	66.01±10.10	61.65±11.09
Gender			
Male	229(64.15%)	101(64.33%)	38(60.32%)
Female	128(35.85%)	56(35.67%)	25(39.68%)
Stage			
I	53(14.85%)	23(14.65%)	7(11.11%)
II	107(29.97%)	34(21.66%)	19(30.16%)
III	148(41.46%)	75(47.77%)	29(46.03%)
IV	38(10.64%)	24(15.29%)	5(7.94%)
Unknown	11(3.08%)	1(0.63%)	3(4.76%)
T category			
T1	19(5.32%)	10(6.37%)	0(0.00%)
T2	76(21.29%)	30(19.11%)	17(26.98%)
T3	164(45.94%)	75(47.77%)	27(42.86%)
T4	95(26.61%)	42(26.75%)	19(30.16%)
Unknown	3(0.84%)	0(0.00%)	0(0.00%)
N category			
N0	111(31.09%)	42(26.75%)	14(22.22%)
N1	97(27.17%)	40(25.48%)	21(33.33%)
N2	75(21.01%)	43(27.39%)	13(20.63%)
N3	74(20.73%)	32(20.38%)	15(23.80%)
Unknown	0(0.00%)	0(0.00%)	0(0.00%)
Grade			
G1	7(1.96%)	5(3.18%)	1(1.59%)
G2	128(35.85%)	81(51.59%)	1(1.59%)
G3	214(59.94%)	67(42.68%)	58(92.06%)
G4	0(0.00%)	0(0.00%)	0(0.00%)

Construction of the ceRNA network and screening the best optimal RNAs: CeRNA is said to act as a sponge for miRNAs and regulate mRNA from miRNA binding. To better comprehend the role of differentially expressed lncRNAs and mRNAs in GC, we constructed a ceRNA regulatory network to elucidate the interaction Between lymph node metastasis or not. In the whole patient cohort, we identified 65 lncRNA-miRNA interaction pairs and 65 miRNA-mRNA interaction pairs (Table S1 and S2) and constructed ceRNA network with 22 lncRNA, 43 miRNAs and 22 mRNAs (Figure S1). In the intestinal-type patient cohort, we identified 233 lncRNA-miRNA interaction pairs and 305 miRNA-mRNA interaction pairs (Table S3 and S4) and constructed ceRNA network with 39 lncRNAs, 157 miRNAs and 66 mRNAs (Figure S2). In the diffuse-type patient cohort, we identified 3 lncRNA-miRNA interaction pairs and 3 miRNA-mRNA interaction pairs (Table S5 and S6) and constructed ceRNA network with 3 lncRNAs, 3 miRNAs and 2 mRNAs (Figure S3). One lncRNA and 3 mRNAs were screened out in the whole cohort applying SVM-RFE. 6 lncRNAs and 17 mRNAs were screened out in the intestinal cohort applying SVM-RFE. 3 lncRNAs and one mRNA were screened out in the diffuse-type cohort applying SVM-RFE.

Functional enrichment analysis: To further clarify the potential biological functions of mRNAs in gastric cancer, KOBAS 3.0 online database was used to perform functional enrichment analysis. Gene ontology analysis showed that there were significantly 3 GO terms in the whole cohort (Figure 2A), 20 GO terms in the intestinal cohort (Figure 2B) and 18 GO terms in the diffuse cohort (Figure 2C). The number of KEGG pathways enriched by mRNAs of ceRNA network was relatively less. There was just biosynthesis of amino acids in the whole cohort. Neuroactive ligand-receptor interaction and nicotine addiction were significantly enriched in the intestinal cohort and pyrimidine metabolism, thyroid hormone synthesis and purine metabolism were significantly associated with the diffuse cohort.

Composition of immune cells in GC and screening the best optimal immune cell types: Immune cells estimated by CIBERSORT algorithm are displayed in Table S7, S8 and S9. The violin plot depicted results of the Wilcoxon rank-sum test (Figure 3). There were significantly different Mast cells resting between lymph node metastasis and non- lymph node metastasis in the whole cohort ($p=0.012$) (Figure 3A) and intestinal-type cohort ($p=0.05$) (Figure 3B) and significantly different Plasma cells ($p=0.049$) and T cells CD8 ($p=0.035$) in the diffuse-type cohort (Figure 3C).

In addition, we applied SVM-RFE to screen and have 5 immune cells related with lymph node metastasis in the whole cohort, including: NK cells activated, macrophages M0, macrophages M2, mast cells resting and neutrophils. In the intestinal-type cohort, we had NK cells resting, dendritic cells resting and mast cells resting. Finally, we identified 4 immune cells, including: B cells memory, plasma cells, T cells CD8 and macrophages M0 in the diffuse-type cohort.

The co-expression analysis: We performed co-expression analysis of immune cells and RNAs significantly associated with lymph node metastasis. In the whole cohort and diffuse cohort, there was

no evident co-expression interaction between RNAs and immune cells (Figure 4A and 4C). In the intestinal cohort, FGF13-AS1 and KCNJ2 were positively associated with NK cells resting and CD70 was positively related with dendritic cells resting (Figure 4B).

Construction of SVM classifier: lncRNA dataset, mRNA dataset and immune cells dataset associated with the lymph node metastasis were merged into a single dataset for improving the performance and reliability of model. The three categories of signature labels were utilized to build an SVM classification model, where the datasets were split into 10 folds and used for training dataset and testing dataset with a 7/3 ratio. We assessed the performance of prediction with accuracy, sensitivity, specificity, positive prediction value (PPV) and negative prediction value (NPV) and the detailed results are in Table 2. The considerable good prediction effects were observed in term of 5 metrics of models. Compared to the whole cohort, SVM classifier in the other two cohort were having a better performance.

Table 2. Performance of SVM model of three cohort in training set and testing set.

Datasets	Accuracy	sensitivity	specificity	PPV	NPV
train					
the whole(polynomial)	0.7835	0.9943	0.3165	0.7632	0.9615
intestinal(polynomial)	0.9643	1	0.875	0.9524	1
diffuse(radial)	0.8913	1	0.5	0.878	1
test					
the whole	0.699	0.97183	0.09375	0.70408	0.6
intestinal	0.8222	0.8286	0.8	0.9355	0.5714
diffuse	0.8235	1	0.25	0.8125	1

Note: the whole cohort for polynomial kernel functions, the intestinal cohort for polynomial kernel functions and the diffuse cohort for radial kernel functions.

Discussion:

With the development of technology, the diagnosis and treatment of GC have made great progress. But the recurrence and metastasis of tumor are important factors affecting prognosis. Then, Lymph node metastasis is the most common form of metastasis in GC. To investigate the mechanism of lymph node metastasis and to estimate the status of its are urgent. Intestinal-type GC is preceded by premalignant lesions, including chronic atrophic gastritis and intestinal metaplasia. Diffuse type GC is more common in young patients, in whom there is a female preponderance and behaves more aggressively than the intestinal type^{17,18}. So, we thought that investigating GC as the whole could lose the specific subtype related discoverable points. Previous studies showed that ceRNA network¹⁹⁻²⁵ and immune cells fraction

^{26 27 28} existing in or influencing the many carcinomas' tumorigenesis, progress and therapy. We supposed that these may also have an effect on lymph node metastasis. In present study, we screened differentially expressed mRNAs and lncRNAs and constructed ceRNA network for GC's lymph node metastasis to reveal its potential functions and mechanisms. The immune cell fractions were evaluated via CIBERSORTx. The lncRNAs, mRNAs and immune cell fractions selected by means of SVM-RFE were used to establish the SVM model for predicting the status of lymph node metastasis in the whole, intestinal and diffuse cohort, respectively. The results showed that the predicting effect of model in the intestinal and diffuse cohort were superior to that in the whole cohort, which proved their excellent clinical application.

In this study, we systematically integrated gene expression profiles and identified lncRNAs and mRNAs in GC. And we didn't identify differentially expressed miRNAs, the reason is that we thought miRNA work as intermediate in the ceRNA network and an extreme example proving our ideas is that when the miRNA's expression is constant, the ceRNA network still works. So, using differentially expressed miRNAs for constructing ceRNA network could lead to the neglect of important lncRNA-miRNA-mRNA triples. We investigated the DElncRNAs, miRNAs and DEGs in the ceRNA network and took the intersection to RNAs of three cohort, respectively. There are not only common RNAs, but also specific subtype's RNAs, which proved our hypothesis that the behavior of lymph node metastasis is different in Lauren subtype. Of course, the detailed mechanism needs us to further explore.

The performance of prediction model of three cohort using significant signatures after SVM-RFE perfectly illustrated different Lauren subtypes had different mechanism of lymph node metastasis. Probably because of fewer mRNAs, the few KEGG enrichment pathways were significant ($p < 0.05$). The KEGG enrichment pathways and GO terms of the diffuse cohort didn't need to be paid more attention, which the reason was that the number of selected mRNAs of the diffuse cohort was small. We found the intrinsic component of membrane of GO terms were common in the whole and intestinal cohort, which meant the intrinsic component of membrane maybe related with lymph node metastasis. In addition, the cell-cell signaling by wnt was specific for the whole cohort and vesicle lumen, catalytic activity, positive regulation of nitrogen compound metabolic process, cellular anatomical entity and so on were specific for the intestinal cohort.

We observed the differences in the components of immune cells between lymph node metastasis and non-metastasis and found significant NK cells activated, Macrophages M0, Macrophages M2, Mast cells resting and Neutrophils for the whole cohort, NK cells resting, dendritic cells resting and mast cells resting for the intestinal cohort and B cells memory, Plasma cells, T cells CD8 and Macrophages M0 for the diffuse cohort. The tumor microenvironment contains innate and adaptive immune cells, which display pro or anti-tumor functions ²⁹. Evidence accumulated from many cancer models suggested that macrophages ^{30 31 32}, Dendritic cells ^{33 34}, Mast cells ^{35 36} and NK cells ³⁷ contributed to the lymph node metastasis of tumors. Significantly different immune cells just supported their association with lymph node metastasis.

In the co-expression of the intestinal cohort, our study suggested that dendritic cells were significantly associated with CD70. CD70 is not only implicated in tumor cell and regulatory T cell survival through interaction with its ligand, CD27³⁸. Moreover, CD70 was reported that it was related with dendritic cells and NK cells^{39 40}. Thus,

There were several limitations of our study that should be acknowledged. The public data we use are all data on the population of Western countries, and this conclusion should be applied cautiously to Asian countries. Because we're analyzing Lauren subgroup population data, the number of cases in the subgroup of diffuse is relatively small, which may lead to less reliable results. Moreover, many inferences need to be proved by further experiments.

Although there are some inadequacies in our research, it still has many bright spots. First of all, our research established the ceRNA network and combined immune cells using CIBERSORTx, and these are based on two Lauren subgroups, which can reduce the effect of tumor heterogeneity. Then, the prediction results of SVM model also demonstrated that it was correct for us to classify and analyze patients with GC. The good performance of SVM model meant it could be used in clinical diagnosis.

Conclusion:

We speculated that CD70 might also play a role in the lymph node metastasis of GC and it maybe a potential therapy target. Our research inferred competing endogenous RNA network of lymph node metastasis and built an excellent predicting model.

Materials And Methods:

Data collection and differential gene expression analysis: The RNA-seq data with fragments per kilobase of transcript per million mapped reads (FPKM) and count and clinical data of the 375 GC samples from The Cancer Genome Atlas (TCGA; <https://cancergenome.nih.gov/>) database were downloaded by using the "TCGAbiolinks" package in R¹⁰. We matched and selected LncRNA and mRNA using GENCODE Release 22 (https://www.encodegenes.org/human/release_22.html)¹¹ as gene annotation, which consisted with the official pipeline of TCGA data portal. Because the data were extracted from TCGA and because this study strictly followed the publication guidelines approved by TCGA (<https://cancergenome.nih.gov/publications/publicationguidelines>), there was no requirement for ethics committee approval. The RNA-seq data was loaded with FPKM, which were converted to transcripts per million (TPM) after removing duplicated genes and zero expression genes. Based on the clinical data, we screened the intestinal patients and diffuse patients and removed the patients without N stage information. The miRNA data transferred by $\log_2(\text{Reads per million mapped reads} + 1)$ was downloaded in UCSC Xena (<http://xena.ucsc.edu/>)¹².

According to node stage (N stage) of tumor node metastasis stage (TNM stage), we divided all patients from TCGA into the whole cohort (all patients with N stage), the intestinal cohort (patients with intestinal-

type classification and N stage) and the diffuse cohort (patients with diffuse-type classification and N stage). We obtained the differentially expressed mRNAs and lncRNAs by using the “DESeq2” packages in R software with $|\log_2 \text{fold change}| > 1$ with an adjusted false discovery rate (FDR) of $P < 0.05$ ¹³.

Construction of the ceRNA Network: First, we predicted the lncRNA–miRNA interaction pairs using differentially expressed lncRNAs (DELncRNAs) by [LncBase Predicted v.2](#) with the threshold of 0.9¹⁴ in all three groups. We thought the miRNA with zero expression in many samples can't work in the ceRNA network. So, the miRNAs with zero expression in more than 20% patients were removed in three groups, respectively. Then, we had the intersections of predicted lncRNA–miRNA and remaining 80% miRNA from above. The miRTarBase¹⁵ were used to predict for the miRNA–mRNA interaction pairs. The miRNAs in the miRNA–mRNA interaction pairs were again took the intersection with the two kinds of interaction pairs. Finally, we established matched lncRNA– miRNA–mRNA triples.

Functional enrichment analysis: To annotate the underlying biological processes of dysregulated mRNAs in the ceRNA network, Gene Ontology (GO) Biological Process enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was performed using KOBAS 3.0 online database¹⁶. The cut-off criteria of significant GO terms and KEGG pathways was $\text{FDR} < 0.05$.

CIBERSORTx Estimation: In order to further explore the cytological causes of lymph node metastasis to some extent, the CIBERSORT algorithm⁹, with the B-mode of batch correction mode and 1000 permutations, was used to estimate the fraction of 22 immune cell types in all patients. Only cases with CIBERSORT $P < 0.05$ were considered eligible for subsequent analysis.

SVM-RFE and SVM: SVM-RFE is a machine learning method based on support vector machine, which is used to find the best optimal variables by deleting SVM-generated eigenvectors. The variables closely related with lymph node metastasis status were selected from mRNA of ceRNA and lncRNA of ceRNA and immune cells, respectively. Then, we merged three data matrices into single matrix with lymph node metastasis status. In order to avoid overfitting of predicting model, the patient cohorts were divided two groups on the basis of 7:3, one in front was named as the train cohort and one in back was named as the test cohort. We built SVM classifiers in the train cohort using screened lncRNAs, mRNAs and immune cells above and examined the performance of classifiers in the test cohort. The above analysis was implemented using the e1071 package in R available from <https://CRAN.Rproject.org/package=e1071>.

Abbreviations:

lncRNAs: Long non-coding RNAs; ceRNAs: competing endogenous RNAs; SVM-RFE: support vector machines recursive feature elimination; SVM: support vector machines; GC: Gastric cancer; DELncRNAs : differentially expressed lncRNAs; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; FPKM: transcript per million mapped reads; TCGA: The Cancer Genome Atlas; TPM : transcripts per million; FDR: false discovery rate.

gastric cancer=GC, competitive endogenous RNA=ceRNA, long noncoding RNA=lncRNA, microRNA=miRNA, messenger RNA=mRNA, microRNA response elements=MREs, support vector machines recursive feature elimination=SVM-RFE, support vector machines=SVM, fragments per kilobase of transcript per million mapped reads=FPKM, The Cancer Genome Atlas=TCGA, transcripts per million=TPM, node stage=N stage, tumor node metastasis stage=TNM stage, Kyoto Encyclopedia of Genes and Genomes =FDR, Gene Ontology=GO, Kyoto Encyclopedia of Genes and Genomes=KEGG, positive prediction value=PPV, negative prediction value=NPV

Declarations:

Acknowledgments: The authors thank the Zhejiang Provincial Research Projects of Medical and Healthy Industries and the General Project Funds from the Health Department of Zhejiang Province.

Author Contributions: Fangxiao Gong and Tianyu Liang designed experiments; The preliminary preparation work was carried out by Fangxiao Gong; Minhua Chen contributed to statistical analysis using R; Minhua Chen edited the manuscript.; Renyang Liu and Jun Hong helped perform the analysis with constructive discussions; all authors read and approved the manuscript; all authors read and approved the manuscript.

Consent for publication: All authors have read the manuscript and approved for publication.

Competing interest: The authors declare no potential conflicts of interest.

Funding: This work was supported by the Zhejiang Provincial Research Projects of Medical and Healthy Industries [Grant NO:2020383843 and 2021448043] and the General Project Funds from the Health Department of Zhejiang Province [Grant No. 2018KY269].

Availability of data and materials' statement: The data that was used in our research is available at <https://portal.gdc.cancer.gov/> and <http://xena.ucsc.edu/>

References:

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394-424.
2. Patru CL, Surlin V, Georgescu I, Patru E. Current issues in gastric cancer epidemiology. *Rev Med Chir Soc Med Nat Iasi.* 2013;117(1):199-204.
3. Yan L. The journey of personalizing gastric cancer treatment. *Chinese journal of cancer.* 2016;35(1):84.
4. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell.* 2011;146(3):353-358.

5. Yang L, Duff MO, Graveley BR, Carmichael GG, Chen LL. Genomewide characterization of non-polyadenylated RNAs. *Genome Biol.* 2011;12(2):R16.
6. Ergun S, Oztuzcu S. Oncocers: ceRNA-mediated cross-talk by sponging miRNAs in oncogenic pathways. *Tumour Biol.* 2015;36(5):3129-3136.
7. Qi X, Zhang DH, Wu N, Xiao JH, Wang X, Ma W. ceRNA in cancer: possible functions and clinical implications. *J Med Genet.* 2015;52(10):710-718.
8. Schlößer HA, Thelen M, Lechner A, Wennhold K, Garcia-Marquez MA, Rothschild SI, Staib E, Zander T, Beutner D, Gathof B, et al. B cells in esophago-gastric adenocarcinoma are highly differentiated, organize in tertiary lymphoid structures and produce tumor-specific antibodies. *Oncoimmunology.* 2019;8(1):e1512458.
9. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, Khodadoust MS, Esfahani MS, Luca BA, Steiner D, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nature Biotechnology.* 2019;37(7):773-782.
10. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, et al. TCGAblinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 2016;44(8):e71.
11. Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, Mudge JM, Sisu C, Wright J, Armstrong J, et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* 2019;47(D1):D766-d773.
12. Goldman MJ, Craft B, Hastie M, Repčeka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN, et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nature biotechnology.* 2020;38(6):675-678.
13. Varet H, Brillet-Guéguen L, Coppée JY, Dillies MA. SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. *PLoS One.* 2016;11(6):e0157022.
14. Paraskevopoulou MD, Vlachos IS, Karagkouni D, Georgakilas G, Kanellos I, Vergoulis T, Zagganas K, Tsanakas P, Floros E, Dalamagas T, et al. DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. *Nucleic Acids Res.* 2016;44(D1):D231-238.
15. Chou CH, Shrestha S, Yang CD, Chang NW, Lin YL, Liao KW, Huang WC, Sun TH, Tu SJ, Lee WH, et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* 2018;46(D1):D296-d302.
16. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li CY, Wei L. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 2011;39(Web Server issue):W316-322.
17. Kaneko S, Yoshimura T. Time trend analysis of gastric cancer incidence in Japan by histological types, 1975-1989. *Br J Cancer.* 2001;84(3):400-405.
18. Adachi Y, Yasuda K, Inomata M, Sato K, Shiraishi N, Kitano S. Pathology and prognosis of gastric carcinoma: well versus poorly differentiated type. *Cancer.* 2000;89(7):1418-1424.

19. Bai Y, Long J, Liu Z, Lin J, Huang H, Wang D, Yang X, Miao F, Mao Y, Sang X, et al. Comprehensive analysis of a ceRNA network reveals potential prognostic cytoplasmic lncRNAs involved in HCC progression. *J Cell Physiol.* 2019;234(10):18837-18848.
20. Liu H, Zhang Z, Wu N, Guo H, Zhang H, Fan D, Nie Y, Liu Y. Integrative Analysis of Dysregulated lncRNA-Associated ceRNA Network Reveals Functional lncRNAs in Gastric Cancer. *Genes (Basel).* 2018;9(6).
21. Zhang Y, Han T, Li J, Cai H, Xu J, Chen L, Zhan X. Comprehensive analysis of the regulatory network of differentially expressed mRNAs, lncRNAs and circRNAs in gastric cancer. *Biomed Pharmacother.* 2020;122:109686.
22. Gao C, Li H, Zhuang J, Zhang H, Wang K, Yang J, Liu C, Liu L, Zhou C, Sun C. The construction and analysis of ceRNA networks in invasive breast cancer: a study based on The Cancer Genome Atlas. *Cancer Manag Res.* 2019;11:1-11.
23. Li MJ, Zhang J, Liang Q, Xuan C, Wu J, Jiang P, Li W, Zhu Y, Wang P, Fernandez D, et al. Exploring genetic associations with ceRNA regulation in the human genome. *Nucleic Acids Res.* 2017;45(10):5653-5665.
24. Yan Y, Lu Y, Mao K, Zhang M, Liu H, Zhou Q, Lin J, Zhang J, Wang J, Xiao Z. Identification and validation of a prognostic four-genes signature for hepatocellular carcinoma: integrated ceRNA network analysis. *Hepatol Int.* 2019;13(5):618-630.
25. Yao Y, Zhang T, Qi L, Zhou C, Wei J, Feng F, Liu R, Sun C. Integrated analysis of co-expression and ceRNA network identifies five lncRNAs as prognostic markers for breast cancer. *J Cell Mol Med.* 2019;23(12):8410-8419.
26. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, Bin J, Liao Y, Rao J, Liao W. Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures. *Cancer Immunol Res.* 2019;7(5):737-750.
27. Zhou R, Zhang J, Zeng D, Sun H, Rong X, Shi M, Bin J, Liao Y, Liao W. Immune cell infiltration as a biomarker for the diagnosis and prognosis of stage III colon cancer. *Cancer Immunol Immunother.* 2019;68(3):433-442.
28. Zhang B, Wu Q, Li B, Wang D, Wang L, Zhou YL. m(6)A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. *Mol Cancer.* 2020;19(1):53.
29. Sag D, Cekic C, Wu R, Linden J, Hedrick CC. The cholesterol transporter ABCG1 links cholesterol homeostasis and tumour immunity. *Nat Commun.* 2015;6:6354.
30. Weichand B, Popp R, Dziumbila S, Mora J, Strack E, Elwakeel E, Frank AC, Scholich K, Pierre S, Syed SN, et al. S1PR1 on tumor-associated macrophages promotes lymphangiogenesis and metastasis via NLRP3/IL-1 β . *J Exp Med.* 2017;214(9):2695-2713.
31. Bieniasz-Krzywiec P, Martín-Pérez R, Ehling M, García-Caballero M, Pinioti S, Pretto S, Kroes R, Aldeni C, Di Matteo M, Prenen H, et al. Podoplanin-Expressing Macrophages Promote Lymphangiogenesis and Lymphoinvasion in Breast Cancer. *Cell Metab.* 2019;30(5):917-936.e910.

32. Wang N, Liu W, Zheng Y, Wang S, Yang B, Li M, Song J, Zhang F, Zhang X, Wang Q, et al. CXCL1 derived from tumor-associated macrophages promotes breast cancer metastasis via activating NF- κ B/SOX4 signaling. *Cell Death Dis.* 2018;9(9):880.
33. Gadalla R, Hassan H, Ibrahim SA, Abdullah MS, Gaballah A, Greve B, El-Deeb S, El-Shinawi M, Mohamed MM. Tumor microenvironmental plasmacytoid dendritic cells contribute to breast cancer lymph node metastasis via CXCR4/SDF-1 axis. *Breast Cancer Res Treat.* 2019;174(3):679-691.
34. Mansfield AS, Heikkila P, von Smitten K, Vakkila J, Leidenius M. Metastasis to sentinel lymph nodes in breast cancer is associated with maturation arrest of dendritic cells and poor co-localization of dendritic cells and CD8+ T cells. *Virchows Arch.* 2011;459(4):391-398.
35. Ammendola M, Sacco R, Zuccalà V, Luposella M, Patrino R, Gadaleta P, Zizzo N, Gadaleta CD, De Sarro G, Sammarco G, et al. Mast Cells Density Positive to Tryptase Correlate with Microvascular Density in both Primary Gastric Cancer Tissue and Loco-Regional Lymph Node Metastases from Patients That Have Undergone Radical Surgery. *Int J Mol Sci.* 2016;17(11).
36. Jain S, Phulari RG, Rathore R, Shah AK, Sancheti S. Quantitative assessment of tumor-associated tissue eosinophilia and mast cells in tumor proper and lymph nodes of oral squamous cell carcinoma. *J Oral Maxillofac Pathol.* 2018;22(2):227-233.
37. Zhang H, Zhu Z, Meadows GG. Chronic alcohol consumption decreases the percentage and number of NK cells in the peripheral lymph nodes and exacerbates B16BL6 melanoma metastasis into the draining lymph nodes. *Cell Immunol.* 2011;266(2):172-179.
38. Jacobs J, Deschoolmeester V, Zwaenepoel K, Rolfo C, Silence K, Rottey S, Lardon F, Smits E, Pauwels P. CD70: An emerging target in cancer immunotherapy. *Pharmacol Ther.* 2015;155:1-10.
39. Shaw J, Wang YH, Ito T, Arima K, Liu YJ. Plasmacytoid dendritic cells regulate B-cell growth and differentiation via CD70. *Blood.* 2010;115(15):3051-3057.
40. Al Sayed MF, Ruckstuhl CA, Hilmenyuk T, Claus C, Bourquin JP, Bornhauser BC, Radpour R, Riether C, Ochsenbein AF. CD70 reverse signaling enhances NK cell function and immunosurveillance in CD27-expressing B-cell malignancies. *Blood.* 2017;130(3):297-309.

Figures

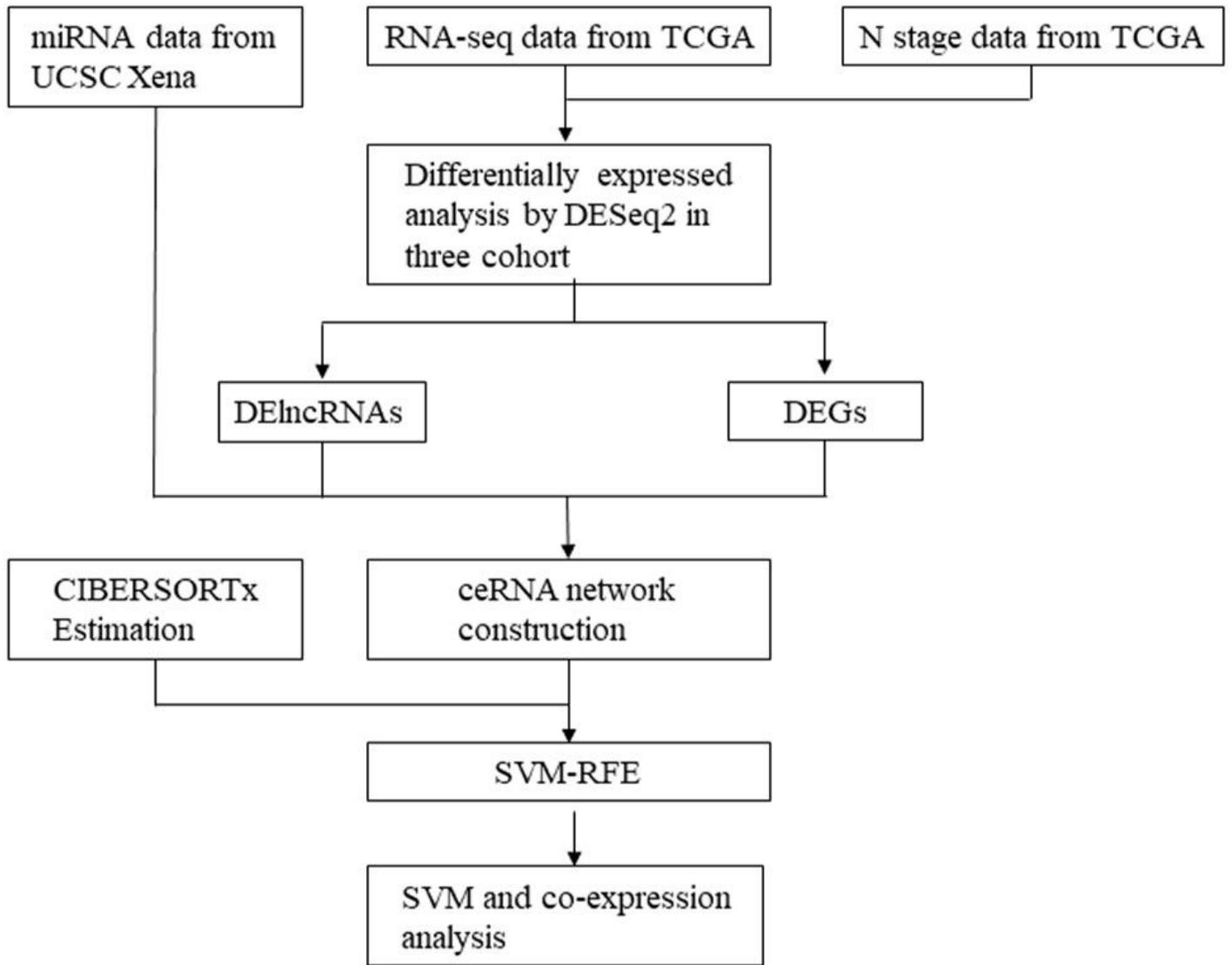


Figure 1

Workflow of our bioinformatics analyses.

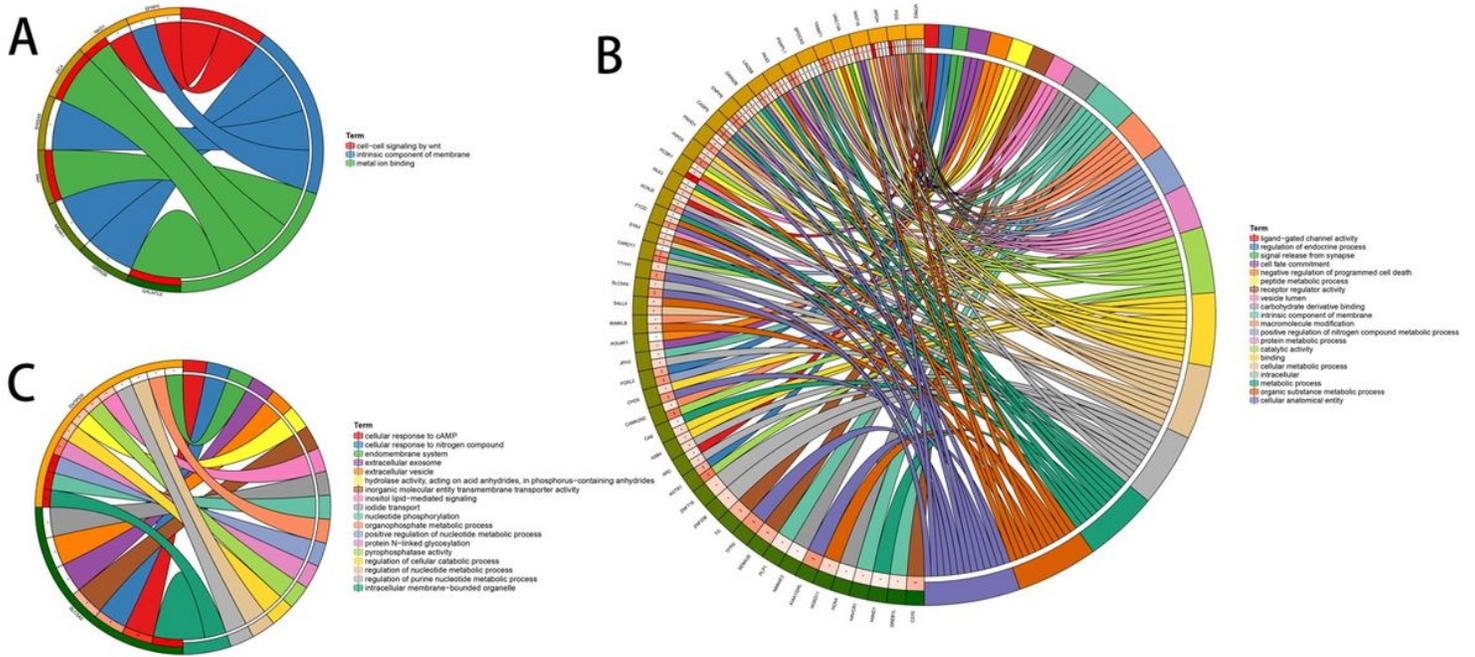


Figure 2

GO terms enrichment analyses of DEGs. (A) in the whole cohort. (B) in the intestinal cohort. (C) in the diffuse cohort.

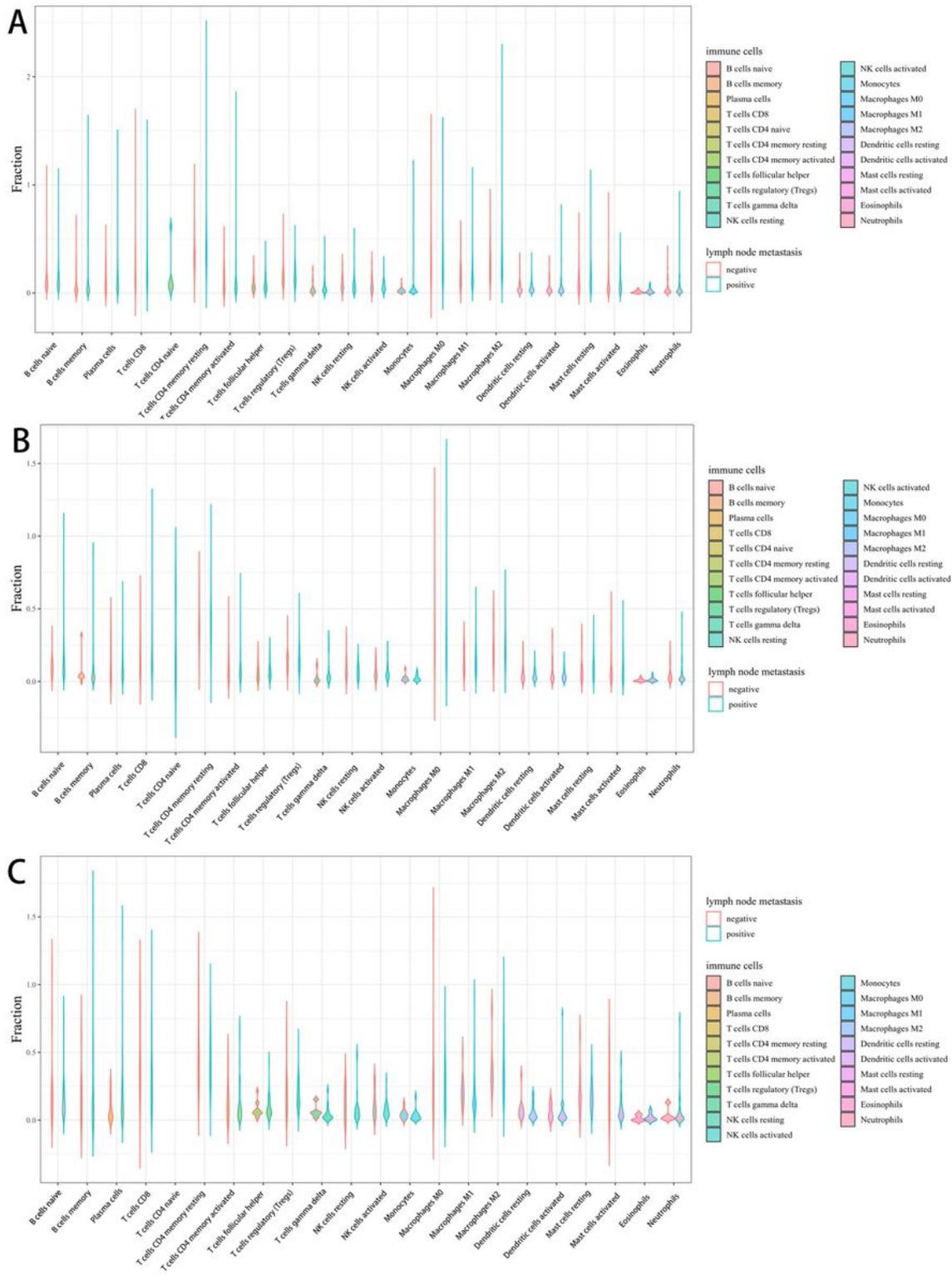


Figure 3

The violin plot of immune cells and the blue and red bar represent lymph node metastasis and non-lymph node metastasis, respectively. (A) in the whole cohort. (B) in the intestinal cohort. (C) in the diffuse cohort.

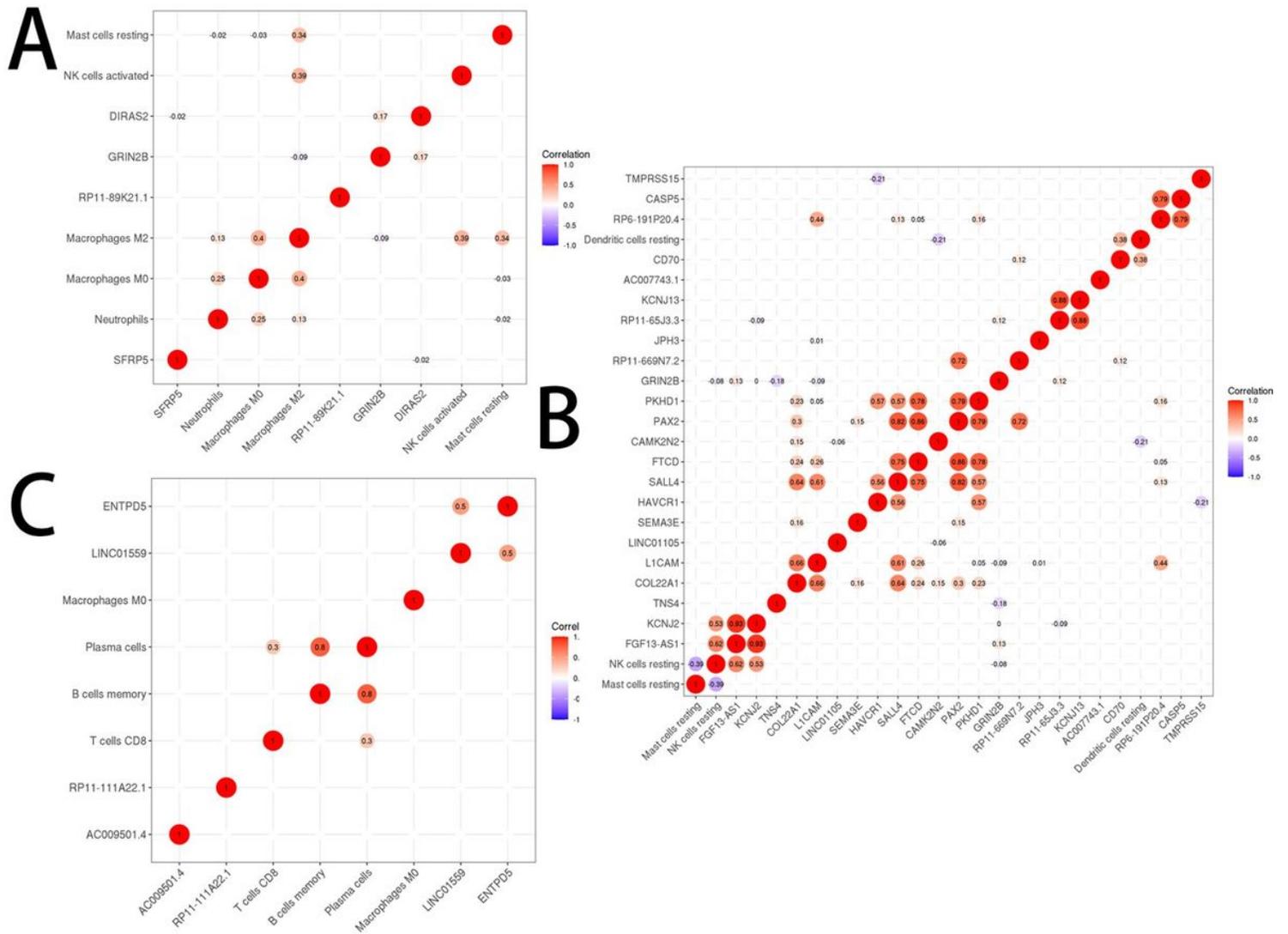


Figure 4

The co-expression patterns among fractions of immune cells, lncRNAs and mRNAs screened by SVM-RFE (the relationships were calculated using Pearson correlation coefficients, the co-expression interactions with coefficient were significant with $p \leq 0.05$), (A) in the whole cohort. (B) in the intestinal cohort. (C) in the diffuse cohort.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.docx](#)
- [TableS2.docx](#)
- [TableS3.docx](#)
- [TableS4.docx](#)
- [TableS5.docx](#)

- TableS6.docx
- TableS7.docx
- TableS8.docx
- TableS9.docx
- FigureS1.tif
- FigureS2.tif
- FigureS3.tif