Identification of CXCL8 as an Immune Microenvironment Regulator of Colorectal Cancer: A Bioinformatics Study Based on Cancer Genome Atlas and Gene Expression Omnibus Datasets

Ying Shao (✉ shying0610@126.com)

Shanxi Medical University <https://orcid.org/0000-0003-4211-1166>

Shuhua Gao

Shanxi Medical University

Minrong Cheng

Shanxi Medical University

Jinxu Zheng

Shanxi Medical University

Ting Yang

Shanxi Medical University

Tie Li

NO.985 Hosipital of the Chinese People's Liberation Army

Lijun Yang

Shanxi Medical University

Tao Yang

Shanxi Medical University

Primary research

Keywords: CXCL8, Colorectal cancer, Immune microenvironment, Immune escape, CD8 T cell, M2 macrophage

Posted Date: October 15th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-91602/v1>

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

[Read Full License](#)

Abstract

Background

At present, a lack of knowledge of the molecular mechanisms underlying the genesis and progression of colorectal cancer (CRC) exists. This is one of the reasons for poor prognosis of this increasingly occurring disease. The purpose of our research is to identify molecular candidates for therapeutic intervention that would help stop the malignant progression of CRC.

Methods

In this study, we performed bioinformatic analysis of the combined microarray data GSE40967 and GSE8671 in Gene Expression Omnibus (GEO) and the CRC RNA-seq dataset in The Cancer Genome Atlas (TCGA). This analysis consisted of 1067 CRC samples and 92 normal mucosae in total. We identified the common differentially expressed genes (DEGs) using R and Venn software. Protein–protein interaction (PPI) network information was obtained using the Search Tool for the Retrieval of Interacting Genes (STRING) Database. The Cytoscape plug-ins MCODE and cytoHubba were used to screen hub modules. *CXCL8*, which was the selected hub gene, was subjected to functional analysis and Gene Set Enrichment Analysis (GSEA). The algorithm “Cell type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORT)” was employed to estimate the relative abundance of tumor infiltrating lymphocyte cells (TILs) with *CXCL8* expression. Finally, correlation analysis was carried out to explore the underlying mechanisms.

Results

We identified 135 common DEGs, which consisted of 38 upregulated genes and 97 downregulated genes. The upregulated DEGs were strikingly enriched in regulation of neutrophil chemotaxis and cytokine-cytokine receptor interaction. The downregulated DEGs were enriched in bicarbonate transport. We identified the candidate molecule *CXCL8* as a product of the hub gene. Functional enrichment analysis and GSEA indicated that *CXCL8* may be closely related to a regulator of the CRC immune microenvironment. Finally, TILs analysis and correlation analysis showed that *CXCL8* may inhibit CD8⁺ T cell immune infiltration through the HIF-1 α /PD-Ls axis. In addition, *CXCL8* could induce the polarization of M2 macrophages through the PI3K/AKT3 pathway.

Conclusions

We determined that hub gene *CXCL8* may promote immune escape of CRC by inhibiting CD8⁺ T cell immune infiltration or promoting polarization of M2 macrophages. *CXCL8* is a potential therapeutic target for the treatment of CRC patients.

Background

CRC is one of the most common malignancies of the digestive system. Most CRC is related to old age, gene aberrations, environmental influence, and lifestyles [1]. However, accumulating evidence has demonstrated that the incidence of CRC in young and middle-aged people has been increasing gradually in recent years [2]. There is a lack of knowledge of the molecular mechanisms underlying CRC progression and uncertainty about the increasing occurrence of the disease and its poor prognosis. Therefore, clarifying the roles of the key molecules and the pathogenesis involved in the progression of CRC is extremely important for the development of effective diagnostic and therapeutic strategies.

In recent years, the high-throughput data from the analysis of gene expression have been archived and deposited in public databases. This includes for example data from platforms such as microarrays. Two databases—GEO and TCGA—serve to collect data from different microarray platforms. More and more researchers have begun to use bioinformatics to discover the precise molecular mechanisms of disease by data mining the many gene expression profiles [3-5]. These comprehensive analyses helped reveal the precise molecular mechanism of tumor pathogenesis and progression. They also provided a theoretical basis for effective diagnostic and therapeutic strategies [6].

In the present study, bioinformatics methods and techniques were used to analyze and integrate the mRNA expression data of CRC from GEO and TCGA. We determined *CXCL8* as the hub molecule that might regulate the microenvironment of CRC using PPI information, hub module analysis, and functional enrichment analysis. In addition, we explored the underlying mechanisms of CXCL8 to regulate the immune microenvironment of CRC based on the algorithm CIBERSORT and correlation analysis. This study provides a reliable way to investigate the mechanism of inflammatory factors mediating immune escape of CRC and to search for potential therapeutic targets for clinical diagnosis or treatment.

Methods

Microarray and RNA-seq Data

The National Center for Biotechnology Information (NCBI)-GEO (www.ncbi.nlm.nih.gov/geo/) is a free public database of microarray profiles. We used GEO to obtain the gene expression profiles of GSE40967 and GSE8671 in CRC samples and normal tissue. The microarray data of GSE40967 and GSE8671 were all compiled on GPL570 Platforms ([HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array). The GSE40967 dataset contained 585 samples, including 566 colorectal adenocarcinoma and 19 non-tumoral colorectal mucosae. The GSE8671 dataset contained 32 colorectal adenoma samples and matched normal mucosae.

The RNA-seq dataset for CRC was downloaded from TCGA database (portal.gdc.cancer.gov). The data contained 478 colorectal cancer samples and 41 normal tissue samples.

Data Preprocessing of DEGs

After CRC data were downloaded, probe identification numbers were transformed into gene symbols using the Perl software. If multiple probes corresponded to one gene, the average of multiple expression was

taken as the gene expression level. Gene expression values were subsequently normalized using the R software.

The aberrantly expressed genes in normal and cancer specimens were analyzed by R software. DEGs were screened by using the criterion $|\log_2(\text{Fold Change})| > 1.5$ and adjusted $P < 0.05$. The raw data were checked online in Venn software (bioinfogp.cnb.csic.es/tools/venny/) to detect common DEGs among the three datasets. The DEGs with $\log_2(\text{Fold Change}) < 0$ were considered as down-regulated genes, while the DEGs with $\log_2(\text{Fold Change}) > 0$ were considered as up-regulated genes.

A heatmap of DEGs between *CXCL8* high-expression group and *CXCL8* low-expression expression samples in the TCGA CRC dataset was completed using the “heatmap” package in R software [7].

Gene Ontology and Pathway Enrichment

Gene ontology (GO) analysis is a useful method for annotating genes and identifying characteristic biological properties of genome and transcriptome data [8]. Kyoto Encyclopedia of Genes and Genomes (KEGG) is a collection of databases dealing with genomes, diseases, biological pathways, and chemical materials. The functional annotation tool of Visualization and Integrated Discovery (DAVID) Bioinformatics Resources 6.8 (david.ncifcrf.gov/) was used to verify the remarkable enrichment of gene function and pathways in the resulting dataset. A Q value < 0.05 was set as significant.

Integration of PPI Network and Module Analysis

PPI information can be evaluated online using the STRING tool (string-db.org/) [9]. We mapped the DEGs using the STRING website. A minimum required interaction score > 0.4 was selected as the significant threshold. PPI networks were constructed using the Cytoscape software [10]. MCODE plug-in was used to find highly interconnected subgraphs as molecular complexes or clusters of PPI networks. The parameter criteria were set as: MCODE scores > 3 and number of nodes > 4 . Finally, the cytoHubba plug-in of Cytoscape was utilized to explore PPI network hub genes. This provides a user-friendly interface to explore important nodes in biological networks and computes using eleven methods. Out of these methods, Maximal Clique Centrality (MCC) had a better performance in the PPI network [11].

Gene set enrichment analysis

GSEA (software.broadinstitute.org/gsea/index.jsp) is a computational evaluation of whether the gene expression differences across biological samples among certain gene sets reach statistical significance [12]. 478 samples of CRC from TCGA were divided into two groups according to the median values of the expression of the *CXCL8* gene: *CXCL8* high-expression group vs *CXCL8* low-expression group. GSEA was used to screen the biological processes and pathways relevant to *CXCL8*. A False Discovery Rate (FDR) < 0.05 and a Q value < 0.05 were set to be significant.

Tumor infiltrating lymphocyte cell analysis

The CIBERSORT algorithm is highly sensitive and specifically discriminates 22 human immune cell phenotypes using support vector regression [13]. It has already been used for the construction of immunoscore models in several cancers [14-16]. The *CXCL8* high-expression group (239 samples) and *CXCL8* low-expression group (239 samples) divided as described above were used for the following analysis. The CIBERSORT algorithm was used to estimate the relative abundance of 21 TILs with *CXCL8* expression using R software. $P < 0.05$ was considered to be statistically significant.

Correlation analysis

The corrplot package was used to calculate the correlation between different genes. The darkest red means the maximum positive correlation coefficient, and the darkest blue means the maximum negative correlation coefficient.

Statistical analysis of Scatter diagram was completed by SPSS 19.0 software package and Prism 8.0. $P < 0.05$ was considered statically significant.

Prediction of transcription factor binding site

The promoter region of PD-L2 was identified using resources available at the NCBI website (www.ncbi.nlm.nih.gov/). First, we searched for the promotor region of PD-L2 at NCBI-GENE (www.ncbi.nlm.nih.gov/gene/). Then we predicted the transcription factor HIF-1 α binding site in promotor region of PD-L2 using Jaspar (jaspar.genereg.net/).

Results

Identification of DEGs in CRC

There were 1076 CRC samples and 92 normal colorectal mucosae in our present study. We extracted 434, 618, and 1395 DEGs from the GSE40967, GSE8671, and TCGA CRC samples, respectively, using the limma package and Student's *t*-test. Venn diagram software was used to identify the common DEGs in the three datasets. 135 DEGs were detected, which included 38 up-regulated genes and 97 down-regulated genes in the CRC tissues and normal colorectal mucosae (Table 1, Fig.1a and b).

Gene ontology and KEGG pathway analysis of DEGs

All 135 of the DEGs were analyzed by DAVID 6.8 to identify GO categories and KEGG pathways. The results of the GO analysis indicated the following: 1) for the biological process (BP), up-regulated DEGs were particularly enriched in "positive regulation of neutrophil chemotaxis", "chemokine-mediated signaling pathway", "hair follicle development", "inflammatory response", and down-regulated genes were enriched in "bicarbonate transport"; 2) for molecular function (MF), up-regulated DEGs were enriched in "chemokine activity", "CXCR chemokine receptor binding", and down-regulated genes were enriched in "carbonate dehydratase activity", and "chloride channel activity"; 3) for the GO cell component (CC), up-regulated DEGs were particularly enriched in "extracellular space", "extracellular region", and down-

regulated genes were enriched in “apical plasma membrane”, and “brush border membrane” (Table 2, Fig.1c and d). Fig.1e shows the most significantly enriched pathways of the up-regulated DEGs and down-regulated DEGs analyzed by KEGG analysis. The up-regulated DEGs were particularly enriched in regulation of “cytokine-cytokine receptor interaction”, “Tumor Necrosis Factor (TNF) signaling pathway”, “NOD-like receptor signaling pathway”, and “chemokine signaling pathway”, and down-regulated genes were enriched in “retinol metabolism”, “bile secretion”, “nitrogen metabolism”, and “mineral absorption”.

PPI network and modular analysis

We used the STRING online database and Cytoscape software to look at the PPI network. All 135 of the DEGs were imported into the PPI network complex, which contained 88 nodes and 152 edges (Fig.2a), 49 genes did not fall into the PPI network. We applied MCODE for further analysis and it showed that the significant module (8 nodes, 25 edges, Fig.2b) from the PPI network was selected. Moreover, a total of 135 nodes and 152 edges were analyzed using the cytoHubba plug-in of Cytoscape. We used the top ten genes for the analysis in each of the eleven methods. The results of five methods, namely, MCC, Maximum Neighborhood Component (MNC), degree, Edge Percolated Component (EPC), and closeness, indicated that their first ranked gene was *CXCL8* (Fig.2c-g).

Enrichment analysis of CXCL8 in the TCGA CRC dataset

We detected the DEGs between the *CXCL8* high-expression and *CXCL8* low- expression patients (Fig.3a). 1229 DEGs were enriched using the DAVID database for KEGG pathway and GO functional enrichment analysis to make a first pass at understanding the biological relevance of *CXCL8*. The genes were mainly enriched in “cytokine secretion”, “cytokine-cytokine receptor interaction”, “immune response”, in addition to accumulating “neutrophil chemotaxis” (Fig.3b and c). Furthermore, GSEA was also applied to look at how the biological process was enriched in *CXCL8* highly expressed samples. Similarly, we analyzed several gene sets that may be associated with immune microenvironment. GSEA analysis not only enriched the gene sets related to “cytokine-cytokine receptor interaction”, but also enriched many immune cell receptor signaling pathways (Fig.3d and Fig.4). Therefore, it can be speculated that the function of *CXCL8* in CRC tissues may be closely related to its immune microenvironment.

***CXCL8* expression negatively correlates with the immune infiltration of CD8⁺ T cell and positively correlates with the immune infiltration of M2 macrophages**

Analysis was performed to explore the mechanism by which *CXCL8* regulates the immune microenvironment of CRC. We estimated the relative abundance of 21 TILs between *CXCL8* high- and low-expression samples of TCGA using the CIBERSORT algorithm. The results showed that CD8⁺ T cell exhibited low infiltration in the *CXCL8* high-expression group, while M2 macrophages were enriched in the *CXCL8* high-expression group (Fig.5b). We also observed that the *CXCL8* high-expression samples of GSE40967 had lower abundance of CD8⁺ T cell and higher abundance of M2 macrophages (Fig.5c).

***CXCL8* may through HIF-1 α /PD-Ls axis inhibited immune infiltration of CD8⁺ T cell**

Next, we analyzed the mechanism by which CXCL8 regulates the immune microenvironment by inhibiting CD8⁺ T cell infiltration. Immunotherapy currently relies on the activation of the anti-tumor effect of T lymphocytes to improve the therapeutic effect. Many researchers have focused on the function of immune inhibitory signaling molecules, including PD-L1 and PD-L2, which are highly expressed on tumor cells and one of the reasons for evasion of the immune system [17, 18]. In this study, Spearman analysis demonstrated that *CXCL8* was positively correlated with PD-L1 (*CD274*) and PD-L2 (*PDCD1LG2*) (Fig.6a-c). This indicated that *CXCL8* inhibited T cell proliferation and that activation may proceed through effects on the expression of PD-L1 and PD-L2. Therefore, we analyzed the transcriptome data of TCGA to explore the upstream transcription factors that regulated PD-L1 and PD-L2 expression. The correlation analysis showed that HIF-1 α was the highest positively correlating entity with *CXCL8* (Fig.6d and e). In addition, HIF-1 α also positively correlated with PD-L1 and PD-L2 (Fig.6f and g). HIF-1 α can regulate the expression of PD-L1 by binding directly to a hypoxia response element (HRE) [19]. The promoter region CCACATGCCT of PD-L2 may be the HIF-1 α binding site according to analysis performed using Jasper (Table 4). These predictions provide clues that HIF-1 α may regulate PD-L1 and PD-L2 expression. In summary, we preliminarily speculate that CXCL8 may inhibit CD8⁺ T cell infiltration in the immune microenvironment through the HIF-1 α /PD-Ls axis.

CXCL8 may act through the PI3K/AKT3 pathway to induce the polarization of M2 macrophages

It has been reported that CXCL8 induces the trafficking of M2 macrophages [20] and that the PI3K/AKT pathway could regulate the polarization of M2 macrophages [21]. We found that the high expression of *CXCL8* was positively related to the activation of the PI3K pathway and AKT3 expression (Fig.6h and i). CXCL8 may therefore promote the polarization of M2 macrophages through the PI3K/AKT3 pathway to exert its immunosuppressive role.

Discussion

CRC is the result of accumulated genetic, epigenetic, and endocrine aberrations [11, 22]. The limitations of previous studies are that they focused on a single cohort study or results from a single database [23-25]. In the present study, we analyzed two microarray datasets GEO40967 and GEO8671 from GEO and RNA sequencing data from TCGA. This included 1067 CRC specimens and 92 normal specimens and compensated for the limitation of a single database.

Although several studies indicated that CXCL8 can promote the malignant progression of a variety of tumors including CRC [26-30], the function of CXCL8 in the immune microenvironment still needs further research. Recently, researchers found that CXCL8 could reshape the tumor microenvironment and accelerate the malignant progression of tumor through recruiting myeloid-derived suppressor cells [31], stimulating angiogenesis [32], and promoting tumor cell epithelial-mesenchymal transition [27]. However, the influence of CXCL8 on the CRC microenvironment is still poorly understood and needs further exploration. Further studies are needed if suitable therapies are to be developed for CRC.

Previously, Oliver and colleagues found that CXCL8-derived macrophages can exert immunosuppressive function by inducing macrophages to express PD-L1 in gastric cancer [33]. Furthermore, more recent research indicated that high expression of PD-L1 can lead to CD8⁺ T cell exhaustion, which may be a key mechanism of tumor immune escape [34]. In this study, we first indicated that CXCL8 upregulation is negatively associated with CD8⁺ T cell immune infiltration in CRC. Moreover, evidence showed that a series of molecular features influenced immunotherapy efficacy, including expression of immune checkpoint molecules [33], infiltration of CD8⁺ T cell [35, 36], oncogenic pathways [37], and tumor-specific neoantigens [38]. In this study, we found that *CXCL8* up-regulation is negatively associated with PD-L1 (*CD274*) and PD-L2 (*PDCD1LG2*). Meanwhile, we established that the high expression of *CXCL8* was positively correlated with the *HIF-1α* expression. Furthermore, evidence showed that HIF-1α was positively correlated with the expression of PD-L1 and PD-L2 [39, 40]. Therefore, we speculated that CXCL8 may play an immunosuppressive function through the HIF-1α/PD-Ls axis. At present, immunotherapy targeting immunosuppressive checkpoints (such as PD-1/PD-L1) has been applied in the clinic and can improve the treatment of various malignant tumors. Only CRC patients with high microsatellite instability and DNA mismatch repair defects can benefit from this type of treatment [41]. Our research helps provide a better understanding of the mechanism of PD-L1 expression and could improve therapies targeting PD-1/PD-L1.

In addition, we found that high expression of CXCL8 is positively correlated with the degree of M2 macrophage infiltration. The imbalance of M1/M2 polarization of macrophages is usually associated with various inflammatory diseases. M1 macrophages are mainly involved in initiating and maintaining the inflammatory response; M2 macrophages are mainly involved in the regression of inflammation [42]. We can preliminarily speculate that CXCL8 can accelerate the immune escape of CRC cells by promoting M2 macrophages polarization. The polarization of macrophages is regulated by a variety of molecules and pathways, including the PI3K/AKT pathway, Notch pathway, IRF-STAT signaling, TGF-β pathway and TLR4/NF-κB pathway [43]. Our results showed that high expression of *CXCL8* was positively related to the activation of the PI3K pathway and *AKT3* expression. Therefore, the results indicated that CXCL8 may promote the polarization of M2 macrophages through the PI3K/AKT3 pathway, and induce its immunosuppressive effect on T cells.

We integrated gene expression profiles and sequencing data from datasets of TCGA and GEO. Bioinformatics analysis identified *CXCL8* as key genes involved in pathways that may play important roles in CRC immune escape. Our results are expected to complement the mechanism of inflammatory factors mediating immune escape of CRC. This is conducive to the search for potential therapeutic targets of CRC and lays a theoretical basis for promoting the application of targets in the clinic. However, there are also limitations to the study. First, our results were only an analysis of microarray or RNA-seq data. The effect of the key molecule CXCL8 on T cells and macrophages needs to be verified in CRC samples and tumor cells. In addition, the mechanisms of CXCL8-mediated immune escape of CRC remain unclear. Further studies are needed to determine if CXCL8 can become a potential target of immune therapy. This will require further validation by subsequent experiments.

Conclusion

This study explored the key role of CXCL8 in the microenvironment of CRC from the perspective of bioinformatics. CXCL8 could regulate the CRC microenvironment by preventing CD8⁺ T cell infiltration or inducing M2 macrophages polarization by different pathways. CXCL8 may therefore serve as a potential immunotherapy target for CRC.

Abbreviations

CRC: Colorectal cancer

GEO: Gene Expression Omnibus

TCGA: The Cancer Genome Atlas

DEGs: Differentially expressed genes

PPI: Protein–protein interaction

STRING: Search Tool for the Retrieval of Interacting Genes Database

GSEA: Gene Set Enrichment Analysis

CIBERSORT: Cell type Identification By Estimating Relative Subsets Of RNA Transcripts

TILs: Tumor infiltrating lymphocyte cells

KEGG: Kyoto Encyclopedia of Genes and Genomes

NCBI: National Center for Biotechnology Information

GO: Gene ontology

DAVID: Visualization and Integrated Discovery

BP: Biological process

MF: Molecular function

CC: Cell component

Declarations

Availability of data and materials

All GEO datasets (GSE40967, GSE8671) and the TCGA CRC dataset are available from the GEO and TCGA databases.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

The work was supported by the grant from Shanxi Provincial Department Finance in China (SD1910), the grant from Shanxi Medical University of Shanxi Province in China (XD1910) and the grant from the National Natural Science Foundation of China (81972325).

Authors' contributions

SG and MC performed the data collecting and analysis. JZ, TY and TL participated in the data interpretation and further analysis. YS designed the study and drafted the manuscript. LJ and TY were responsible for supervised the study and critical revision of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

References

1. Islami F, Goding Sauer A, Miller KD, Siegel RL, Fedewa SA, Jacobs EJ, McCullough ML, Patel AV, Ma J, Soerjomataram I et al: Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. *CA Cancer J Clin.* 2018; 68(1):31-54.
2. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, Cercek A, Smith RA, Jemal A: Colorectal cancer statistics, 2020. *CA Cancer J Clin.* 2020; 70(3):145-164.
3. Li S, Yang R, Sun X, Miao S, Lu T, Wang Y, Wo Y, Jiao W: Identification of SPP1 as a promising biomarker to predict clinical outcome of lung adenocarcinoma individuals. *Gene.* 2018; 679:398-404.
4. Peng Q, Shen Y, Zhao P, Cheng M, Zhu Y, Xu B: Biomarker roles identification of miR-106 family for predicting the risk and poor survival of colorectal cancer. *BMC cancer.* 2020; 20(1):506.
5. Gong L, Zhang D, Dong Y, Lei Y, Qian Y, Tan X, Han S, Wang J: Integrated Bioinformatics Analysis for Identifying the Therapeutic Targets of Aspirin in Small Cell Lung Cancer. *Journal of biomedical informatics.* 2018; 88:20-28.
6. Yuan L, Zeng G, Chen L, Wang G, Wang X, Cao X, Lu M, Liu X, Qian G, Xiao Y et al: Identification of key genes and pathways in human clear cell renal cell carcinoma (ccRCC) by co-expression analysis.

- International journal of biological sciences. 2018; 14(3):266-279.
7. Galili T, O'Callaghan A, Sidi J, Sievert C: heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics* (Oxford, England). 2018; 34(9):1600-1602.
 8. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT et al: Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics*. 2000; 25(1):25-29.
 9. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP et al: STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic acids research*. 2015; 43(Database issue):D447-452.
 10. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T: Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* (Oxford, England). 2011; 27(3):431-432.
 11. Cao L, Chen Y, Zhang M, Xu DQ, Liu Y, Liu T, Liu SX, Wang P: Identification of hub genes and potential molecular mechanisms in gastric cancer by integrated bioinformatics analysis. *PeerJ*. 2018; 6:e5180.
 12. Wang J, Vasaikar S, Shi Z, Greer M, Zhang B: WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic acids research*. 2017; 45(W1):W130-w137.
 13. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA: Robust enumeration of cell subsets from tissue expression profiles. *Nature methods*. 2015; 12(5):453-457.
 14. Zhou L, Xu L, Chen L, Fu Q, Liu Z, Chang Y, Lin Z, Xu J: Tumor-infiltrating neutrophils predict benefit from adjuvant chemotherapy in patients with muscle invasive bladder cancer. *Oncoimmunology*. 2017; 6(4):e1293211.
 15. Fu H, Zhu Y, Wang Y, Liu Z, Zhang J, Xie H, Fu Q, Dai B, Ye D, Xu J: Identification and Validation of Stromal Immunity Predict Survival and Benefit from Adjuvant Chemotherapy in Patients with Muscle-Invasive Bladder Cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2018; 24(13):3069-3078.
 16. Zeng D, Zhou R, Yu Y, Luo Y, Zhang J, Sun H, Bin J, Liao Y, Rao J, Zhang Y et al: Gene expression profiles for a prognostic immunoscore in gastric cancer. *The British journal of surgery*. 2018; 105(10):1338-1348.
 17. Kitano A, Ono M, Yoshida M, Noguchi E, Shimomura A, Shimoi T, Kodaira M, Yunokawa M, Yonemori K, Shimizu C et al: Tumour-infiltrating lymphocytes are correlated with higher expression levels of PD-1 and PD-L1 in early breast cancer. *ESMO open*. 2017; 2(2):e000150.
 18. Rozali EN, Hato SV, Robinson BW, Lake RA, Lesterhuis WJ: Programmed death ligand 2 in cancer-induced immune suppression. *Clinical & developmental immunology*. 2012; 2012:656340.
 19. Noman MZ, Chouaib S: Targeting hypoxia at the forefront of anticancer immune responses. *Oncoimmunology*. 2014; 3(12):e954463.
 20. Zhang M, Huang L, Ding G, Huang H, Cao G, Sun X, Lou N, Wei Q, Shen T, Xu X et al: Interferon gamma inhibits CXCL8-CXCR2 axis mediated tumor-associated macrophages tumor trafficking and enhances anti-PD1 efficacy in pancreatic cancer. *Journal for immunotherapy of cancer*. 2020; 8(1).

21. Vergadi E, Ieronymaki E, Lyroni K, Vaporidi K, Tsatsanis C: Akt Signaling Pathway in Macrophage Activation and M1/M2 Polarization. *Journal of immunology* (Baltimore, Md : 1950). 2017; 198(3):1006-1014.
22. Esteller M, Risques RA, Toyota M, Capella G, Moreno V, Peinado MA, Baylin SB, Herman JG: Promoter hypermethylation of the DNA repair gene O(6)-methylguanine-DNA methyltransferase is associated with the presence of G:C to A:T transition mutations in p53 in human colorectal tumorigenesis. *Cancer research*. 2001; 61(12):4689-4692.
23. Chen ZH, Qi JJ, Wu QN, Lu JH, Liu ZX, Wang Y, Hu PS, Li T, Lin JF, Wu XY et al: Eukaryotic initiation factor 4A2 promotes experimental metastasis and oxaliplatin resistance in colorectal cancer. *Journal of experimental & clinical cancer research : CR*. 2019; 38(1):196.
24. Soldevilla B, Carretero-Puche C, Gomez-Lopez G, Al-Shahrour F, Riesco MC, Gil-Calderon B, Alvarez-Vallina L, Espinosa-Olarte P, Gomez-Esteves G, Rubio-Cuesta B et al: The correlation between immune subtypes and consensus molecular subtypes in colorectal cancer identifies novel tumour microenvironment profiles, with prognostic and therapeutic implications. *European journal of cancer* (Oxford, England : 1990). 2019; 123:118-129.
25. Chen F, Li Z, Zhou H: Identification of prognostic miRNA biomarkers for predicting overall survival of colon adenocarcinoma and bioinformatics analysis: A study based on The Cancer Genome Atlas database. *Journal of cellular biochemistry*. 2019; 120(6):9839-9849.
26. Liu Q, Li A, Yu S, Qin S, Han N, Pestell RG, Han X, Wu K: DACH1 antagonizes CXCL8 to repress tumorigenesis of lung adenocarcinoma and improve prognosis. *Journal of hematology & oncology*. 2018; 11(1):53.
27. Cheng XS, Li YF, Tan J, Sun B, Xiao YC, Fang XB, Zhang XF, Li Q, Dong JH, Li M et al: CCL20 and CXCL8 synergize to promote progression and poor survival outcome in patients with colorectal cancer by collaborative induction of the epithelial-mesenchymal transition. *Cancer letters*. 2014; 348(1-2):77-87.
28. Ogawa R, Yamamoto T, Hirai H, Hanada K, Kiyasu Y, Nishikawa G, Mizuno R, Inamoto S, Itatani Y, Sakai Y et al: Loss of SMAD4 Promotes Colorectal Cancer Progression by Recruiting Tumor-Associated Neutrophils via the CXCL1/8-CXCR2 Axis. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2019; 25(9):2887-2899.
29. Sun L, Wang Q, Chen B, Zhao Y, Shen B, Wang H, Xu J, Zhu M, Zhao X, Xu C et al: Gastric cancer mesenchymal stem cells derived IL-8 induces PD-L1 expression in gastric cancer cells via STAT3/mTOR-c-Myc signal axis. *Cell death & disease*. 2018; 9(9):928.
30. Cheng J, Li Y, Liu S, Jiang Y, Ma J, Wan L, Li Q, Pang T: CXCL8 derived from mesenchymal stromal cells supports survival and proliferation of acute myeloid leukemia cells through the PI3K/AKT pathway. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*. 2019; 33(4):4755-4764.
31. Alfaro C, Teixeira A, Oñate C, Pérez G, Sanmamed MF, Andueza MP, Alignani D, Labiano S, Azpilikueta A, Rodriguez-Paulete A et al: Tumor-Produced Interleukin-8 Attracts Human Myeloid-Derived Suppressor Cells and Elicits Extrusion of Neutrophil Extracellular Traps (NETs). *Clinical cancer*

- research : an official journal of the American Association for Cancer Research. 2016; 22(15):3924-3936.
32. Song YS, Kim MJ, Sun HJ, Kim HH, Shin HS, Kim YA, Oh BC, Cho SW, Park YJ: Aberrant Thyroid-Stimulating Hormone Receptor Signaling Increases VEGF-A and CXCL8 Secretion of Thyroid Cancer Cells, Contributing to Angiogenesis and Tumor Growth. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2019; 25(1):414-425.
 33. Lin C, He H, Liu H, Li R, Chen Y, Qi Y, Jiang Q, Chen L, Zhang P, Zhang H et al: Tumour-associated macrophages-derived CXCL8 determines immune evasion through autonomous PD-L1 expression in gastric cancer. *Gut*. 2019; 68(10):1764-1773.
 34. Jansen CS, Prokhnevskaya N, Master VA, Sanda MG, Carlisle JW, Bilen MA, Cardenas M, Wilkinson S, Lake R, Sowalsky AG et al: An intra-tumoral niche maintains and differentiates stem-like CD8 T cells. *Nature*. 2019; 576(7787):465-470.
 35. Oliver AJ, Davey AS, Keam SP, Mardiana S, Chan JD, von Scheidt B, Beavis PA, House IG, Van Audenaerde JR, Darcy PK et al: Tissue-specific tumor microenvironments influence responses to immunotherapies. *Clinical & translational immunology*. 2019; 8(11):e1094.
 36. Li J, Byrne KT, Yan F, Yamazoe T, Chen Z, Baslan T, Richman LP, Lin JH, Sun YH, Rech AJ et al: Tumor Cell-Intrinsic Factors Underlie Heterogeneity of Immune Cell Infiltration and Response to Immunotherapy. *Immunity*. 2018; 49(1):178-193.e177.
 37. Liao W, Overman MJ, Boutin AT, Shang X, Zhao D, Dey P, Li J, Wang G, Lan Z, Li J et al: KRAS-IRF2 Axis Drives Immune Suppression and Immune Therapy Resistance in Colorectal Cancer. *Cancer cell*. 2019; 35(4):559-572.e557.
 38. Gubin MM, Zhang X, Schuster H, Caron E, Ward JP, Noguchi T, Ivanova Y, Hundal J, Arthur CD, Krebber WJ et al: Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature*. 2014; 515(7528):577-581.
 39. Noman MZ, Desantis G, Janji B, Hasmim M, Karray S, Dessen P, Bronte V, Chouaib S: PD-L1 is a novel direct target of HIF-1 α , and its blockade under hypoxia enhanced MDSC-mediated T cell activation. *The Journal of experimental medicine*. 2014; 211(5):781-790.
 40. Pinato DJ, Black JR, Trousil S, Dina RE, Trivedi P, Mauri FA, Sharma R: Programmed cell death ligands expression in pheochromocytomas and paragangliomas: Relationship with the hypoxic response, immune evasion and malignant behavior. *Oncoimmunology*. 2017; 6(11):e1358332.
 41. Segal NH, Saltz LB: Translational Considerations on the Outlook of Immunotherapy for Colorectal Cancer. *Current Colorectal Cancer Reports*. 2015; 11(2):92-97.
 42. Kong X, Gao J: Macrophage polarization: a key event in the secondary phase of acute spinal cord injury. *Journal of cellular and molecular medicine*. 2017; 21(5):941-954.
 43. Wang N, Liang H, Zen K: Molecular mechanisms that influence the macrophage m1-m2 polarization balance. *Frontiers in immunology*. 2014; 5:614.

Tables

Table 1 All 135 common DEGs were detected from three dataset

DEGs	Gene names
Up-regulated	CDH3 CXCL3 MMP7 SERPINB5 FOXQ1 MMP3 INHBA ANLN SLC7A5 CXCL2 SRPX2 PSAT1 LRP8 NFE2L3 CEMIP SLC6A6 TRIM29 ACSTD2 CLDN2 PHLDA1 S100A2 CXCL1 CHI3L1 CST1 CXCL8 PCDD1 KLK10 ASCL2 LGR5 AJUBA ETV4 CLDN1 EPHX4 PTP4A3 GFBI MMP1 KRT23 TEAD4
Down-regulated	CPNE8 SPIB SD17B2 HRDL1 BCG2 MMP28 ZG16 TMCC3 UCA2B GT2B17 CHP2 OGN SCARA5 CLCA4 PHLPP2 DHRS11 BTNL8 CA4 RPM6 ADAMDEC1 CES2 GHR PYY LIFR LYVE1 B3GNT7 UGT2B15 SLC26A3 CD36 CXCL12 THRB SLC30A10 NR5A2 B3GALT5 TSPAN7 KLB HHLA2 VIP CA2 SEMA6D MFAP5 IL6R CHGA MT1M TP53INP2 SLC26A2 PKIB CDHR5 CEACAM7 PADI2 MT1G TMEM171 ABCA8 MT1F BEST2 CLDN23 SLC51B ADH1B EDN3 AQP8 BRINP3 LYPD8 TNFRSF17 CD177 MS4A12 SECTM1 PCK1 CPM ADH1C ABI3BP SLC17A4 UGT2A3 GCNT2 RUNDC3B SCNN1B C2orf88 EPB41L3 CDKN2B HPGD LAMA1 WH43 AMDC2 BEST4 CA1 STMN2 MEP1A MALL TEX11 SLC4A4 SFRP1 PDK4 TMEM37 SLC51A GUCA2A TMEM100 DHRS9 CA7

Table 2 Gene ontology analysis of differentially expressed genes associated with CRC

Expression	Category	Term	Count	%	PValue
Up-regulated	GOTERM_BP_DIRECT	GO:0090023~positive regulation of neutrophil chemotaxis	4	0.080305	1.24E-05
	GOTERM_BP_DIRECT	GO:0070098~chemokine-mediated signaling pathway	4	0.080305	0.00043
	GOTERM_BP_DIRECT	GO:0001942~hair follicle development	3	0.060229	0.002831
	GOTERM_BP_DIRECT	GO:0006954~inflammatory response	5	0.100381	0.007698
	GOTERM_MF_DIRECT	GO:0008009~chemokine activity	4	0.080305	1.67E-04
	GOTERM_MF_DIRECT	GO:0045236~CXCR chemokine receptor binding	3	0.060229	1.67E-04
	GOTERM_CC_DIRECT	GO:0005615~extracellular space	12	0.240915	3.69E-05
	GOTERM_CC_DIRECT	GO:0005576~extracellular region	12	0.240915	0.000186
Down-regulated	GOTERM_BP_DIRECT	GO:0015701~bicarbonate transport	7	0.053603	1.05E-07
	GOTERM_MF_DIRECT	GO:0004089~carbonate dehydratase activity	4	0.03063	4.31E-05
	GOTERM_MF_DIRECT	GO:0005254~chloride channel activity	5	0.038288	1.56E-04
	GOTERM_CC_DIRECT	GO:0016324~apical plasma membrane	9	0.068918	1.40E-04
	GOTERM_CC_DIRECT	GO:0031526~brush border membrane	5	0.038288	1.44E-04

Table 3 KEGG pathway analysis of DEGs associated with CRC

Expression	Category	Term	Count	%	PValue	Genes
Up-regulated	KEGG_PATHWAY	hsa05134: Legionellosis	4	0.080305	0.000288	CXCL1, CXCL3, CXCL2, CXCL8
	KEGG_PATHWAY	hsa05132: Salmonella infection	4	0.080305	0.001019	CXCL1, CXCL3, CXCL2, CXCL8
	KEGG_PATHWAY	hsa04668: TNF signaling pathway	4	0.080305	0.002123	CXCL1, CXCL3, CXCL2, MMP3
	KEGG_PATHWAY	hsa04060: Cytokine-cytokine receptor interaction	5	0.100381	0.002514	CXCL1, INHBA, CXCL3, CXCL2, CXCL8
	KEGG_PATHWAY	hsa04621: NOD-like receptor signaling pathway	3	0.060229	0.008185	CXCL1, CXCL2, CXCL8
	KEGG_PATHWAY	hsa04062: Chemokine signaling pathway	4	0.080305	0.010001	CXCL1, CXCL3, CXCL2, CXCL8
Down-regulated	KEGG_PATHWAY	hsa00830: Retinol metabolism	6	0.045945	6.08E-05	UGT2B17, ADH1C, DHRS9, ADH1B, UGT2A3, UGT2B15
	KEGG_PATHWAY	hsa04976: Bile secretion	6	0.045945	8.74E-05	AQP8, SLC51B, CA2, SLC51A, SLC4A4, ABCG2
	KEGG_PATHWAY	hsa00910: Nitrogen metabolism	4	0.03063	1.78E-04	CA7, CA4, CA2, CA1
	KEGG_PATHWAY	hsa04978: Mineral absorption	5	0.038288	1.95E-04	SLC26A3, TRPM6, MT1M, MT1G, MT1F

Table 4 The HIF-1α binding sites in PD-L2 promoter region

Matrix ID	Name	Score	Relative score	Sequence ID	Start	End	Strand
MA1106.1	HIF-1 α	4.66475	0.801956146724	5'	441	450	+
Predicted sequence: CCACATGCCT							

Figures

Figure 1

Identification of DEGs and Enrichment analysis in CRC dataset. a Authentication of 36 common up-regulated DEGs in the three datasets (TCGA, GSE40967 and GSE8671) through Venn diagrams software. b 97 DEGs were down-regulated in three datasets. c, d Go pathway analysis of the upregulated DEGs (c) and downregulated DEGs (d) by DAVID. e KEGG pathway analysis of the common DEGs by DAVID.

Figure 1

Identification of DEGs and Enrichment analysis in CRC dataset. a Authentication of 36 common up-regulated DEGs in the three datasets (TCGA, GSE40967 and GSE8671) through Venn diagrams software. b 97 DEGs were down-regulated in three datasets. c, d Go pathway analysis of the upregulated DEGs (c) and downregulated DEGs (d) by DAVID. e KEGG pathway analysis of the common DEGs by DAVID.

Figure 1

Identification of DEGs and Enrichment analysis in CRC dataset. a Authentication of 36 common up-regulated DEGs in the three datasets (TCGA, GSE40967 and GSE8671) through Venn diagrams software. b 97 DEGs were down-regulated in three datasets. c, d Go pathway analysis of the upregulated DEGs (c) and downregulated DEGs (d) by DAVID. e KEGG pathway analysis of the common DEGs by DAVID.

Figure 2

PPI network of common DEGs constructed by STRING database and Module analysis. a The PPI network contains 88 nodes and 152 edges. The nodes meant DEGs, the edges meant the interaction of proteins. b Module analysis via MCODE of Cytoscape software. c-g Module analysis via cytoHubba of Cytoscape software. c: MCC, d: MNC, e: Degree, f: EPC, g: Closeness.

Figure 3

Enrichment analysis of CXCL8 in TCGA CRC dataset. a Heatmap of DEGs between CXCL8 high- and low-expression patients in TCGA CRC dataset. Green meant downregulated DEGs, red meant upregulated DEGs. b, c 1229 DEGs were enriched by using DAVID database for KEGG pathway (b) and GO functional enrichment analysis (c). d Go functional annotation genes in the module obtained by GSEA. The y-axis showed the GO terms and x-axis showed the $-\log_{10}(Q\text{ value})$ of each term.

Figure 3

Enrichment analysis of CXCL8 in TCGA CRC dataset. a Heatmap of DEGs between CXCL8 high- and low-expression patients in TCGA CRC dataset. Green meant downregulated DEGs, red meant upregulated DEGs. b, c 1229 DEGs were enriched by using DAVID database for KEGG pathway (b) and GO functional enrichment analysis (c). d Go functional annotation genes in the module obtained by GSEA. The y-axis showed the GO terms and x-axis showed the $-\log_{10}(Q\text{value})$ of each term.

Figure 3

Enrichment analysis of CXCL8 in TCGA CRC dataset. a Heatmap of DEGs between CXCL8 high- and low-expression patients in TCGA CRC dataset. Green meant downregulated DEGs, red meant upregulated DEGs. b, c 1229 DEGs were enriched by using DAVID database for KEGG pathway (b) and GO functional enrichment analysis (c). d Go functional annotation genes in the module obtained by GSEA. The y-axis showed the GO terms and x-axis showed the $-\log_{10}(Q\text{value})$ of each term.

Figure 4

GSEA reveals potential downstream signaling of CXCL8. GSEA results showed that cytokine-cytokine receptor interaction (a), chemokine signaling pathway (b), B cell receptor signaling pathway (c), Toll like receptor signaling pathway (d) were enriched in the CXCL8 high-expression group. Top panels indicated the enrichment scores of each gene. Bottom panels showed the ranking metrics of each gene. NES: normalized enrichment score.

Figure 4

GSEA reveals potential downstream signaling of CXCL8. GSEA results showed that cytokine-cytokine receptor interaction (a), chemokine signaling pathway (b), B cell receptor signaling pathway (c), Toll like receptor signaling pathway (d) were enriched in the CXCL8 high-expression group. Top panels indicated the enrichment scores of each gene. Bottom panels showed the ranking metrics of each gene. NES: normalized enrichment score.

Figure 4

GSEA reveals potential downstream signaling of CXCL8. GSEA results showed that cytokine-cytokine receptor interaction (a), chemokine signaling pathway (b), B cell receptor signaling pathway (c), Toll like receptor signaling pathway (d) were enriched in the CXCL8 high-expression group. Top panels indicated the enrichment scores of each gene. Bottom panels showed the ranking metrics of each gene. NES: normalized enrichment score.

Figure 5

The relative abundance of 21 TILs between CXCL8 high- and low-expression samples of CRC. a Group of high- and low-expression of CXCL8 in TCGA CRC dataset. b Violin diagram showed the relative abundance of 21 TILs of CXCL8 in TCGA CRC dataset. Red color represented the CXCL8 high-expression samples, blue color represented the CXCL8 low-expression samples. c Violin diagram showed the relative abundance of 21 TILs of CXCL8 in GSE40967 dataset. Red color represented the CXCL8 high-expression samples, blue color represented the CXCL8 low-expression samples.

Figure 5

The relative abundance of 21 TILs between CXCL8 high-and low-expression samples of CRC. a Group of high- and low-expression of CXCL8 in TCGA CRC dataset. b Violin diagram showed the relative abundance of 21 TILs of CXCL8 in TCGA CRC dataset. Red color represented the CXCL8 high-expression samples, blue color represented the CXCL8 low-expression samples. c Violin diagram showed the relative abundance of 21 TILs of CXCL8 in GSE40967 dataset. Red color represented the CXCL8 high-expression samples, blue color represented the CXCL8 low-expression samples.

Figure 5

The relative abundance of 21 TILs between CXCL8 high- and low-expression samples of CRC. a Group of high- and low-expression of CXCL8 in TCGA CRC dataset. b Violin diagram showed the relative abundance of 21 TILs of CXCL8 in TCGA CRC dataset. Red color represented the CXCL8 high-expression samples, blue color represented the CXCL8 low-expression samples. c Violin diagram showed the relative abundance of 21 TILs of CXCL8 in GSE40967 dataset. Red color represented the CXCL8 high-expression samples, blue color represented the CXCL8 low-expression samples.

Figure 6

Correlation analysis of CXCL8 with a series of genes. a Correlation heatmap of CXCL8 with T cell activation related genes. b Scatter diagram showed the correlation of CXCL8 with CD274 expression. c Scatter diagram showed the correlation of CXCL8 with PDCD1LG2 expression. d Correlation heatmap of CXCL8 with the transcription factors of CD274, PDCD1LG2. e Scatter diagram showed the correlation of CXCL8 with HIF-1α expression. f Scatter diagram showed the correlation of HIF-1α with CD274 expression. g Scatter diagram showed the correlation of HIF-1α with PDCD1LG2 expression. h Correlation heatmap of CXCL8 with various genes. i Scatter diagram showed the correlation of AKT3 with CXCL8 expression.

g Scatter diagram showed the correlation of HIF-1α with PDCD1LG2 expression. h Correlation heatmap of CXCL8 with PI3K/AKT pathway and ERK1/2 pathway. i Scatter diagram showed the correlation of CXCL8 with AKT3 expression.

Figure 6

Correlation analysis of CXCL8 with a series of genes. a Correlation heatmap of CXCL8 with T cell activation related genes. b Scatter diagram showed the correlation of CXCL8 with CD274 expression. c

Scatter diagram showed the correlation of CXCL8 with PDCD1LG2 expression. d Correlation heatmap of CXCL8 with the transcription factors of CD274, PDCD1LG2. e Scatter diagram showed the correlation of CXCL8 with HIF-1α expression. f Scatter diagram showed the correlation of HIF-1α with CD274 expression. g Scatter diagram showed the correlation of HIF-1α with PDCD1LG2 expression. h Correlation heatmap of CXCL8 with PI3K/AKT pathway and ERK1/2 pathway. i Scatter diagram showed the correlation of CXCL8 with AKT3 expression.

Figure 6

Correlation analysis of CXCL8 with a series of genes. a Correlation heatmap of CXCL8 with T cell activation related genes. b Scatter diagram showed the correlation of CXCL8 with CD274 expression. c Scatter diagram showed the correlation of CXCL8 with PDCD1LG2 expression. d Correlation heatmap of CXCL8 with the transcription factors of CD274, PDCD1LG2. e Scatter diagram showed the correlation of CXCL8 with HIF-1 α expression. f Scatter diagram showed the correlation of HIF-1 α with CD274 expression. g Scatter diagram showed the correlation of HIF-1 α with PDCD1LG2 expression. h Correlation heatmap of CXCL8 with PI3K/AKT pathway and ERK1/2 pathway. i Scatter diagram showed the correlation of CXCL8 with AKT3 expression.