Heliyon 9 (2023) e22433

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Prediction of immune infiltration and prognosis for patients with gastric cancer based on the immune-related genes signature

Xianghui Li [a,b,1], Yuanyuan Chen [a,1], Yuxiang Dong [c,*,1], Zhongjin Ma [a], Wenjun Zheng [a,**], Youkun Lin [a,***]

[a] Nanjing Drum Tower Hospital, the Affiliated Hospital of Nanjing University Medical School, Nanjing, 210008, China
[b] Department of Dermatology, The First Affiliated Hospital of Guangxi Medical University, Nanning 530021, China
[c] Second Clinical College of Nanjing Medical University, Nanjing, 210029, China

ARTICLE INFO

Keywords:
Gastric cancer
Immune-related genes
Immune microenvironment
Prognostic signature
Immune cell types

ABSTRACT

Objective: The immune microenvironment influenced clinical outcomes and treatment response of gastric cancer (GC) patients. Though thousands of immune-related genes (IRGs) have been identified, their effects on GC are not fully understood. The objective of the study is to analyze the correlations between the expression and effect of IRGs and clinical outcomes. Moreover, we evaluate the efficacy and value of utilizing the immune-related genes signature as a prognosis prediction model for GC patients.

Methods: We identified the differentially expressed IRGs and systematically analyzed their functions. We constructed a novel GC prognostic signature and a new nomogram, Moreover, we explored the infiltrated immune cell types in the immune microenvironment and discussed the genetic variation in GC IRGs.

Results: Eight IRGs, including CCL15, MSR1, GNAI1, NR3C1, ITGAV, NMB, AEN, and TGFBR1 were identified. Based on the prognostic signature, GC patients were distinguished into two subtype groups. As verified in multiple datasets, the prognostic signature exhibited good performance in predicting the prognosis (AUC = 0.803, *P*-value <0.001) and revealed the different clinical features and infiltrated immune cell types in the immune microenvironment.

Conclusions: In summary, we found that IRGs contributed to GC prognosis prediction and constructed an IRGs-based GC prognostic signature, which could serve as an effective prognostic stratification tool.

1. Introduction

Gastric cancer (GC) is one of the most common malignancies worldwide, with the incidence ranking 6th and the mortality ranking 3rd. Most patients are diagnosed at an advanced stage and succumb to a worse prognosis and living quality [1]. Though fluoropyrimidine and platinum-based chemotherapy as the first-line therapy are standard treatments, the prognosis for the advanced-stage

* Corresponding author. 300 Zhongshan Rd, Nanjing, 210008, China.
** Corresponding author. 6 Shuangyong Rd, Nanning, 530000, China.
*** Corresponding author. 6 Shuangyong Rd, Nanning, 530000, China.
E-mail addresses: yuxiangdong@njmu.edu.cn (Y. Dong), gxmuzwj@163.com (W. Zheng), linyoukun7@aliyun.com (Y. Lin).
[1] Xianghui Li, Yuanyuan Chen, and Yuxiang Dong contributed equally.

https://doi.org/10.1016/j.heliyon.2023.e22433
Received 6 May 2023; Received in revised form 10 November 2023; Accepted 13 November 2023
Available online 23 November 2023
2405-8440/© 2023 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

patient remains unfavorable [2]. Existing treatments for GC are insufficient. Tumor recurrence and metastasis affect clinical management and lead to pessimistic survival.

The immune microenvironment has been educated by the tumor, and meanwhile supports tumor growth. Human genes with immune system-related functions, named immune-related genes (IRGs), play a critical role in the development of the immune microenvironment [3]. However, the function and regulatory mechanism of these immune-related genes remain unclear and have not been systematically explored. Moreover, as the immune microenvironment influences tumor growth and response to therapeutic interventions, IRGs may serve as an effective predictor for GC prognosis [4,5].

In this study, we identified the differential expression IRGs in GC and found that expression differences of IRGs contribute to GC prognosis. Thus, we conducted a multilevel analysis, and a novel immune-related prognosis signature was established. The prognosis signature exhibited high accuracy in the training cohort, internally validated cohort, and external cohort. We further explored the infiltration cells in the immune microenvironment and discuss the possible regulatory mechanism of IRGs. We found that the prognosis of GC patients could be predicted and stratified effectively, and IRGs demonstrated the potential for offering more valuable information for GC patients.

2. Methods

2.1. Data acquisition and processing

We obtained 2498 immune genes from the ImmPort database (https://immport.niaid.nih.gov). The transcriptome data of GC patients with complete data were obtained from the TCGA database (https://portal.gdc.cancer.gov/), which contained 270 cases of tumor tissue and 28 cases of adjacent normal tissue. To validate the novel prognosis model, an external validation cohort was obtained from Gene Expression OMNIBUS (GEO) database (GSE84437, https://www.ncbi.nlm.nih.gov/geo/). As shown in Table 1, clinical parameters of the 270 GC patients including age, gender, pathological stage, histological grade, T stage, M stage, N stage, survival status, and survival duration were collected. Samples with missing information were excluded. The transcriptome data was corrected through the "Limma" package in R software. Immune genes with a *P*-value <0.05 and $|\log_2 FC \text{ (fold change)}| > 1$ were considered as different expression immune genes (DEIGs). The data was shared voluntarily and could be obtained from the above-mentioned websites freely with no requirement of ethics committee approval.

2.2. Gene function enrichment analysis

To realize the major functions of the above-mentioned DEIGs, we conducted the gene function enrichment analysis. Kyoto Encyclopedia of Genes and Genomes (KEGG) and gene ontology (GO) analysis were also applied using the R ClusterProfiler package [6,

Table 1
Clinical parameters of included GC patients from TCGA in the training cohort and the validation cohort.

Variables	Total (n = 270)	Training cohort (n = 136)	Validation cohort (n = 134)
Age (year)			
<60	91	45	46
≥60	179	91	88
Gender			
FEMALE	109	52	57
MALE	161	84	77
Stage			
I	38	16	22
II	81	43	38
III	122	59	63
IV	29	18	11
T stage			
T1	13	9	4
T2	55	21	34
T3	132	68	64
T4	70	38	32
N stage			
N0	84	37	47
N1	69	45	24
N2	61	30	31
N3	59	25	34
M stage			
M0	251	125	126
M1	19	11	8
Grade			
G1	5	2	3
G2	95	55	40
G3	170	78	92

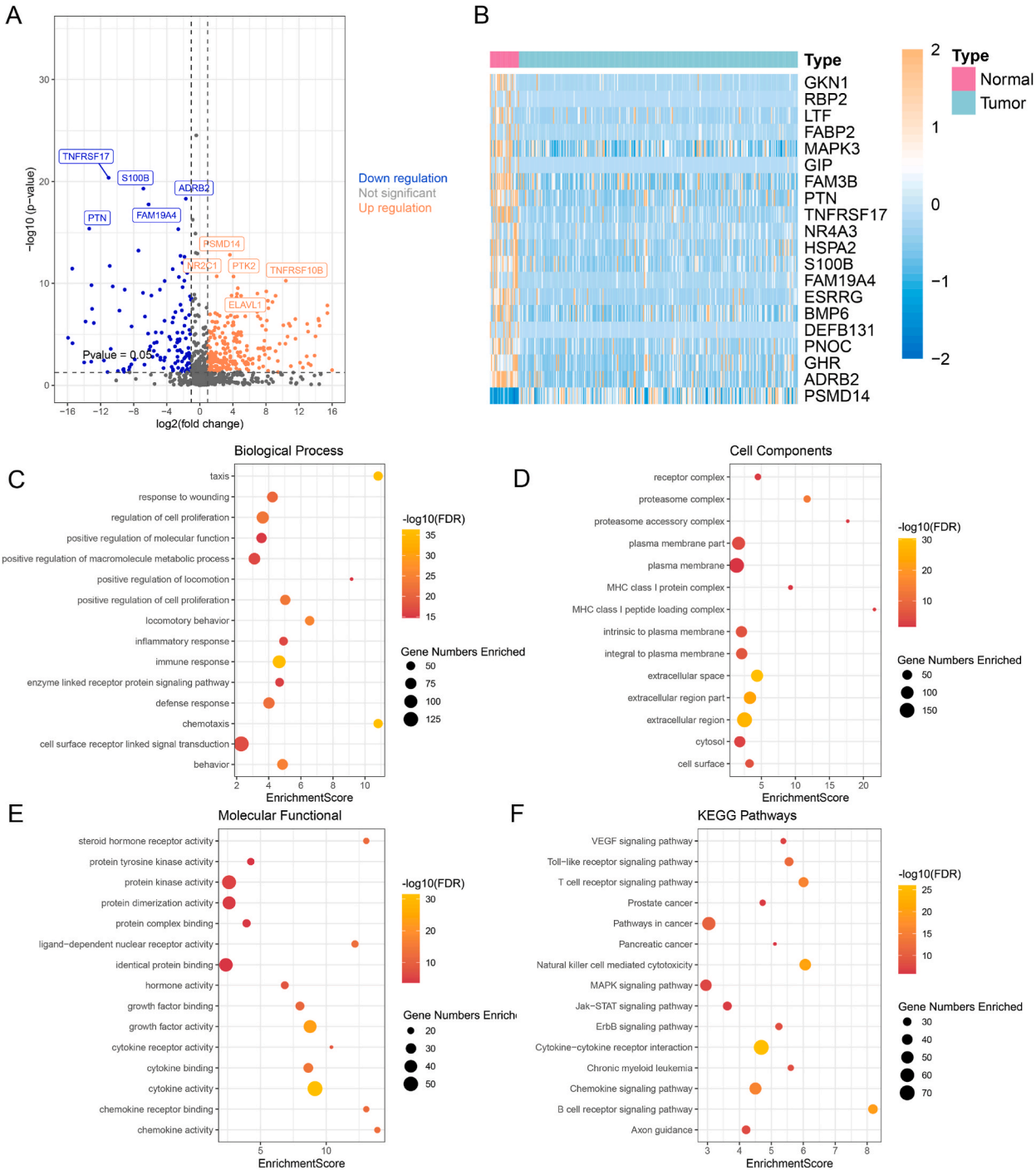


Fig. 1. Identification of differential expressed immune genes. (A) Volcano plot of differently expressed immune genes from TCGA database with a standard of $|\log_2 FC| > 1$ and *P*-value <0.05. The orange dots in the volcano plots represent up-regulation, the blue dots represent down-regulation and the grey dots represent genes without differential expression. The top 5 up-regulated and top 5 down-regulated IRGs were listed. (B) The heatmap illustrated the top 20 differently expressed IRGs in comparison of normal tissues and GC tissues from the TCGA database. The colors in the heatmap from blue to yellow represent the fold change of the expression levels from low to high. (C ~ F) GO and KEGG analysis results of gene function enrichment analysis. The top 10 function annotations of each part were listed. DEIGs were mostly enriched in cell surface receptor-linked signal transduction, plasma membrane, cytokine activity, and cytokine-cytokine receptor interaction. (C) GO analysis results for biological process. (D) GO analysis results for cell components. (E) GO analysis results for molecular function. (F) KEGG pathways analysis results.

7]. Database for Annotation, Visualization, and Integrated Discovery (DAVID, https://david.ncifcrf.gov/) was employed to identify enriched KEGG and GO terms [8,9].

2.3. Survival analysis and expression comparison of DEIGs

Univariate Cox regression analysis was conducted to explore the influence of each gene on overall survival (OS). We didn't perform WGCNA combined with the PPI network for that protein translation was outside the scope of this article and not ignoring prognosis-related IRGs, The survival-related IRGs with a *P*-value<0.05 were identified and integrated into further multivariate Cox regression. Eight hub genes were screened from the multivariate Cox regression (*P*-value<0.05). The risk score of each patient was calculated according to the formula in Supplementary Formula 1. The patients in the cohort were divided into a high-risk group and a low-risk group according to the median risk score value of all patients. Survival information including survival status and survival time was obtained from the TCGA database. The "Survival" R package was utilized in the survival analysis. A Log-rank test was conducted to detect significant differences (*P*-value <0.05) in OS rates. The results were demonstrated in Kaplan-Meier (KM) survival curves.

2.4. Functional analysis

To understand the internal biological mechanisms of the immune genes scoring signature, the gene set enrichment analysis (GSEA, http://software.broadinstitute.org/gsea/downloads.jsp, version 3.0, the broad institute of MIT and Harvard) was performed between the high-risk group and the low-risk group. In detail, the number of marks was set to 1000, the 'permutation type' was as phenotype, the 'collapse data set to gene symbols was set as false, and the 'enrichment statistic' was set to weight. The high-risk group was set as the experimental group while the low-risk group was set as the reference group. Moreover, Gene size >500 and < 15, *P*-value <0.05, and FDR <0.25 were set as the cut-off criteria. The immune microenvironment score was obtained by ESTIMATE, which is an algorithmic tool for inferring the abundance of infiltrating immune cells. The bubble plot was drawn by R software. Based on TCGA GC data, the copy number variation (CNV) analysis was performed by the "R-Circos" package and the "R-ggplot2" package. Additionally, the online tool website Cbioportal (http://www.cbioportal.org/) was used to analyze gene mutation status, in which the threshold was set as *P*-value <0.05. The EPIC database was used for the quantitative analysis of immune cells. Eight IRGs sets were acquired from the MSigDB database, and ssGSEA (single-sample GSEA) used the 'GSVA' R package to calculate the signal intensity of a single sample.

2.5. Statistical analysis

All analyses were performed using R (version 3.6.1; https://www.r-project.org/). All statistical tests were two-tailed, and a *P*-value <0.05 was considered to be significant. Continuous variables with normal distribution were compared by independent *t*-test, while those with skewed distribution were compared by Mann-Whitney *U* test. Based on Pearson Correlation Coefficient, the correlation matrix was constructed. The ROC curves were performed by the "SurvivalROC" R package and used to explore the sensitivity and specificity of the signature. To explore the effects of individual variables on survival, the univariate regression model was applied. To further confirmed the independent impact factors associated with survival, the multivariate Cox regression model was used. With the regression coefficients based on the above Cox analysis, the nomogram was constructed.

3. Results

3.1. Identification of differentially expressed IRGs (DIRGs) and functional annotation

In total, 2498 IRGs were obtained from the ImmPort database. The transcriptome data of 270 gastric cancer cases and 28 adjacent normal tissue cases gained from the TCGA database were conducted for differential analysis. Ultimately, 446 IRGs emerged with significant expression differences between GC and normal tissue (*P*-value<0.05, Fig. 1A). There were 283 up-regulated IRGs and 163 down-regulated IRGs were identified (Fig. 1A, Table S1). Moreover, the expression profile of the top 10 up-regulated and the top 10 down-regulated genes was displayed in the heat map (Fig. 1B).

We mainly focus on the regulatory function of the immune microenvironment of these DEIRGs. Thus, we conducted the Kyoto Encyclopedia of Genes and Genomes (KEGG) and gene ontology (GO) analyses to further explore their biological connotation. As shown in Fig. 1C, the top 3 enriched biological processes were cell surface receptor-linked signal transduction, immune response, and regulation of cell proliferation; the top 3 enriched cell components were extracellular region, plasma membrane, and plasma membrane part (Fig. 1D); the top 3 molecular functions were cytokine activity, protein kinase activity, and growth factor activity (Fig. 1E); top 3 pathway were cytokine-cytokine receptor interaction, pathways in cancer and chemokine signaling pathway (Fig. 1F). Most of the above-mentioned biological behaviors were associated with immune regulation and tumor progression. We could conclude that these DEIRGs contribute to an abnormal immune microenvironment in GC tissues.

3.2. Identification of the prognosis-related DEIRGs and construction of the prognostic signature

To further reveal which of the DEIGs were related to overall survival (OS), univariate Cox regression analysis was applied. 28 of the 446 DEIRGs were identified significantly associated with OS (Fig. S1). TCGA GC data were randomly divided into two cohorts (training cohort = 134, validation cohort = 136, Table 1). Further regression analysis was further performed (Fig. 2A). Finally, 8 hub-DEIRGs

were selected as the independent risk factors of OS.

Based on the expression coefficient of the 8 hub-DEIRGs, the prognostic signature was established (Table S3, Formula S1). Given the risk score calculated by the prognostic signature, patients were arranged into the high-risk group and low-risk group (Fig. 2B). The survival profile of different groups was demonstrated in Fig. 2C, and a heat map was implemented to depict the heterogeneity expression of the 8 DEIGs between groups (Fig. 2D).

Next, we evaluated the efficacy of the prognostic signature in multiple datasets. The Kaplan–Meier analysis revealed that the high-risk group showed a significantly lower overall survival compared with the low-risk group in the training cohort (P-value<0.001, Fig. 2E). As shown in Fig. 2F, the ROC curve indicated the amplified sensitivity and specificity of the prognostic signature (1-year OS AUC = 0.803). The signature was further verified in the validation cohort (P-value = 0.013, Figs. S2 and S3). Similarly, the prognostic

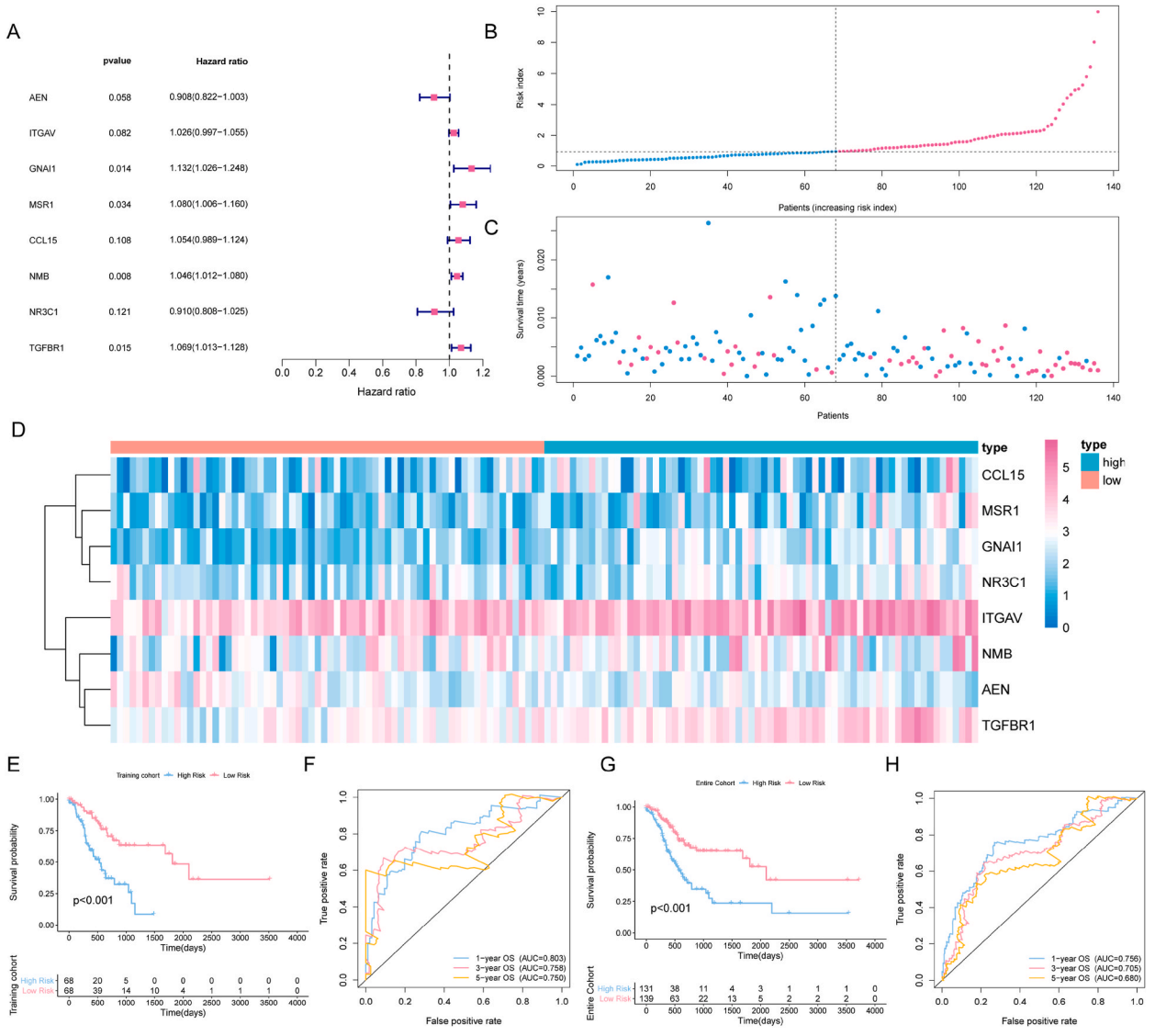
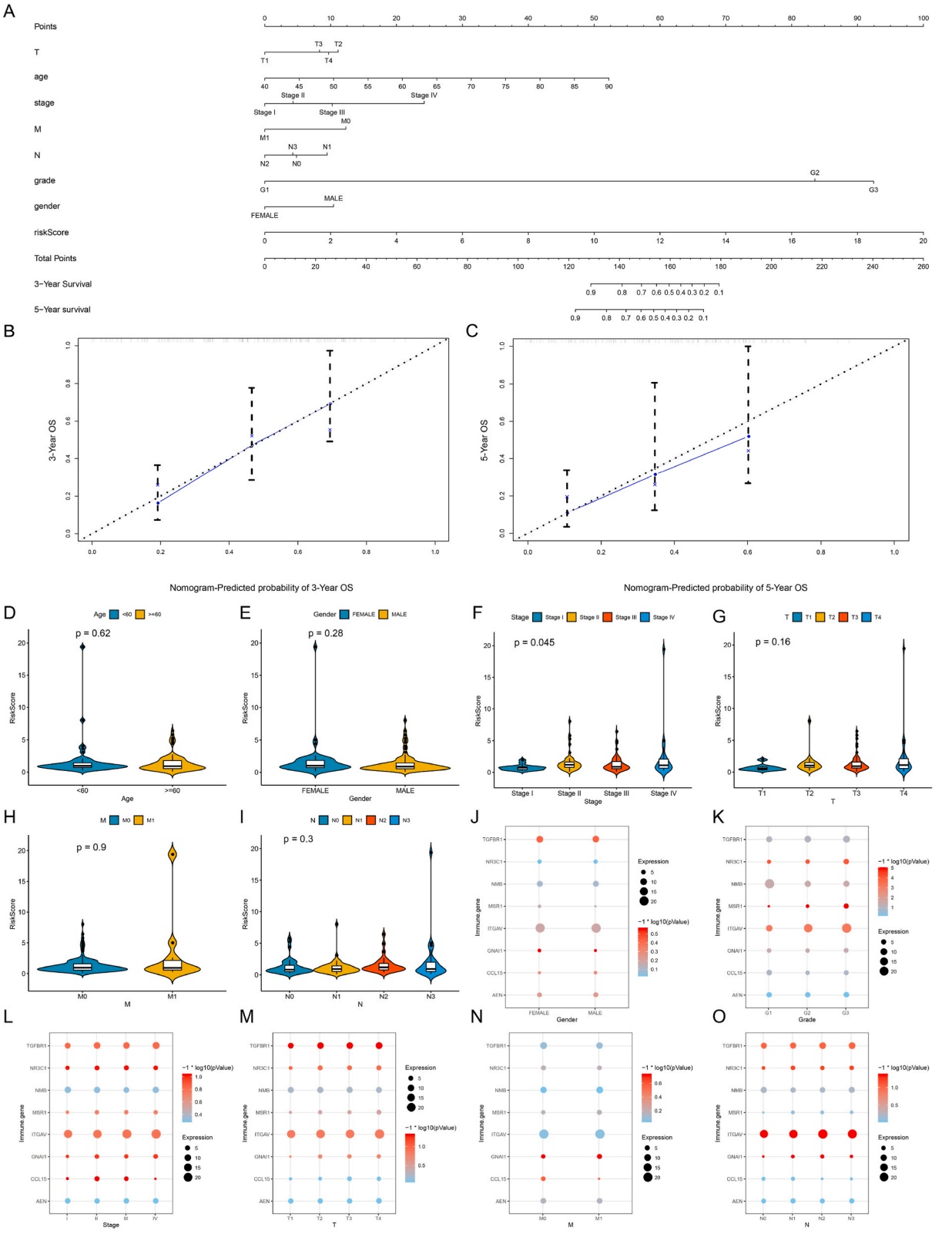


Fig. 2. Construction and validation of the novel prognostic signature. (A) Multivariate Cox regression analysis of the 8 hub IRGs in the novel prognostic score. (B) According to the risk scores of the patients in the training cohort, the patients were divided into a high-risk group (pink) and a low-risk group (blue) by the mean value of the risk score of all patients. (C) Survival status of patients in a high-risk group (pink) and low-risk group (blue). The survival time of the high-risk group was shorter than the low-risk group. (D) Heatmap of the expression of the 8 hub IRGs of the novel prognostic signature in comparison between high-risk group and low-risk group. E-H. Kaplan–Meier survival analyses and time-dependent ROC curves in the training and total cohorts. ROC, receiver operator characteristic. AUC is the area under the curve. (E) Kaplan–Meier survival analysis between the high-risk group and low-risk group in the validation cohort with a P-value <0.001. (F) AUC in ROC analysis for signature at the 1-, 3-, and 5-year survival time in the validation cohort, with an AUC of 0.803, 0.758, and 0.750. (G) Kaplan–Meier survival analysis between the high-risk group and low-risk group in both cohorts with a P-value <0.001. (H) AUC in ROC analysis for signature at the 1-, 3-, and 5-year survival time in both cohorts with an AUC of 0.756, 0.705, and 0.680.



(caption on next page)

Fig. 3. Nomogram construction for OS prediction in GC patients and the validation of clinical features. (A) Based on the prognostic signature, the nomogram was established to predict the OS of patients with GC. (B) Plots displaying the calibration of the nomogram measuring the consistency between the prediction and actual 3−year OS in all patients. (C) Plots displaying the calibration of the nomogram measuring the consistency between the prediction and actual 5−year OS in all patients. (D–I) Correlation between the risk score of IRGs and different clinical features. (D) Age. (E) Gender. (F) Stage. (G) T stage. (H) N stage. (I) M stage. (J–O) Correlation analysis between TNM & Stage and 8 model genes in gastric cancer cases. (J) Correlation analysis between pathologic stage and 8 model genes expression in gastric cancer cases. (K) Correlation analysis between tumor stage and 8 model genes expression in gastric cancer cases. (L) Correlation analysis between metastasis stage and 8 model genes expression in gastric cancer cases. (M) Correlation analysis between node stage and 8 model genes in gastric cancer cases.

signature exhibited as an effective predictor in the entire cohort, which was verified in the Kaplan–Meier analysis ($P < 0.001$, Fig. 2G) and demonstrated a 1-year OS AUC = 0.756 (Fig. 2H).

3.3. Clinical validation of the prognostic signature

As the efficacy of the prognostic signature was verified, we constructed a nomogram for the convenience of clinical application (Fig. 3A). Following the above results, the high-risk group demonstrated a shorter OS than the low-risk group. Moreover, the calibration curve of 3-year OS (Fig. 3B) and 5-year OS (Fig. 3C) suggested that the nomogram exhibited a high accuracy in OS prediction.

Furthermore, univariate and multivariate Cox regression analyses about age, gender, stage, grade, TNM stage, and risk score were conducted to verify the independence. The Univariate Cox model revealed that age, pathological stage, T stage, M stage, N stage, and risk score were correlated with OS (Fig. S2). While in the multivariate Cox model, only age and risk score were proved to be the independent predicted factor (Fig. S3).

We also found an obvious correlation between risk score and pathological stage (Fig. 3D–I). Moreover, the proportion of 8 hub genes was researched in different gender, different histological grades, and different pathological stages (Fig. 3J-O). It was presented that NR3C1 and GNAI1 were significantly relevant to the development of gastric cancer.

3.4. Functional analysis of risk scores and external validation

Additionally, we investigated the bio-correlation between risk scores and the biological behavior of gastric cancer. Gene set enrichment analysis of risk score was performed based on the TCGA gastric cancer cohort. The result indicated that the high-risk score was related to angiogenesis, epithelial-mesenchymal transition, hypoxia, and UV response DNA pathways (Fig. 4A–D). Based on the TCGA GC dataset, the copy number variation (CNV) analysis of 8 model genes was performed by utilizing the R-Circos package and R-ggplot2, and the resultant frequency of copy number variation was displayed (Fig. 4F), in which GNAI1 and MSR1 had the most variable frequency. Additionally, the single nucleotide polymorphism composition (SNP) analysis was also conducted (Fig. 4G). As a result, ITGAV exhibited most SNPs, consistent with the missense mutation, frameshift delete, and so on.

As shown in Fig. 4E, the expression of immune genes is correlated with tumor microenvironment immune cell infiltration. Tumor purity is mainly associated with MSR1, NR3C1, and ITGAV. Up-regulated MSR1 contributes to macrophage M2 infiltration. GNAI1 contributes to CD4 T cells' memory resting and inhibits T cell CD4 memory activation. In summary, the expression of the 8 hub immune genes in the high-risk group demonstrated an immunosuppressive environment, mediated by M2 macrophage infiltration, dendritic cell (DC) resting promotion, and activation inhibited. These results showed that immune genes play immense roles in tumor-suppressive microenvironment regulation.

Moreover, an external validation cohort was set ($P = 0.003$) (Fig. 4H). The results exhibited similar significance and served as an effective prediction tool (Fig. 4I).

4. Discussion

Notwithstanding numerous efforts that have been made to improve the diagnosis and therapeutic effects, gastric cancer remains one of the most threats to human health, with more than 1,000,000 new cases worldwide per year [10]. However, owing to gene heterogeneity, not everyone could benefit from surgery or chemotherapy. It is well-documented that the role of the immune system as the key defense pathway against cancer contributes to the development of the new era of cancer immunotherapies [11]. Moreover, tumor cells exacerbate immunosuppressive pathways and foster a tolerant microenvironment [12]. Suppressive immune cells and secreted chemokines cytokines modulate the tumor microenvironment and favor the growth and expansion of cancer cells [13]. Thus, exploring molecular biomarkers and histopathology is greatly meaningful. Establishing an IRG-based signature to identify GC patients quickly and appropriately may contribute to purposive and personalized clinical treatment decisions.

In this study, we developed an individualized IRGs-based signature and validated its effectiveness in predicting the prognosis for GC patients. Specifically, we identified 446 DEIGs and constructed an 8 hub-IRGs (AEN, ITGAV, GNAI1, MSR1, CCL15, NMB, NR3C1, and TGFBR1) risk score signature. We conducted the ROC curve to investigate the prognostic value of the signature with a *P*-value$<$0.001 and an AUC equal to 0.803 [14]. We further explore the association between the signature and clinical features. As expected, the high-risk group possessed worse OS, higher histological grade, and higher pathological stage, which may correlate to poor outcomes. Furthermore, the effectiveness of the signature was examined in the validation cohort and total cohort. These results demonstrated the efficacy of the IRGs-based signature in prognosis prediction and may serve as an effective prognostic stratification tool.

Moreover, we explore the function of the immune genes. In the KEGG and GO analysis, the DEIGs showed significant relevance with

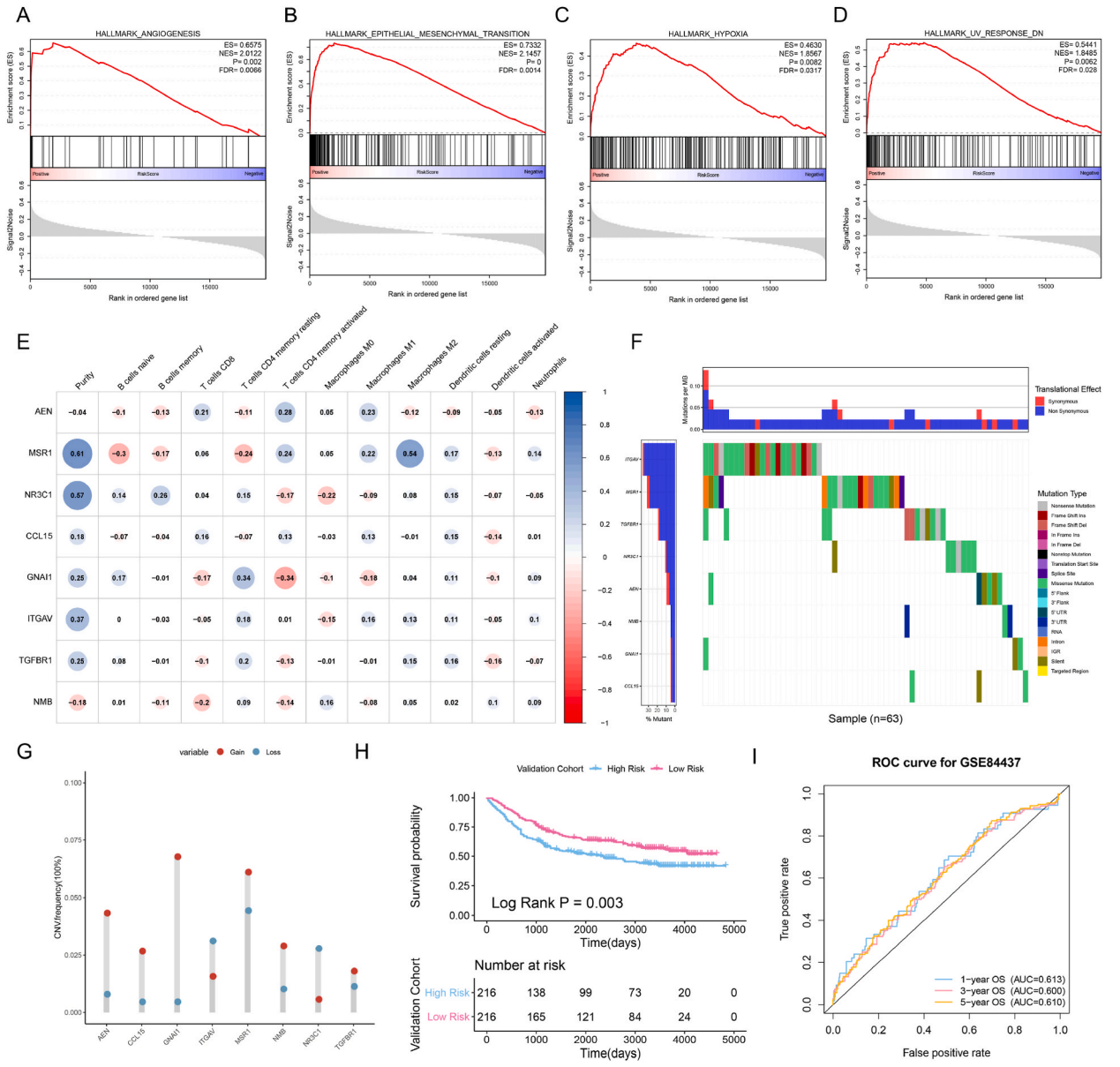


Fig. 4. GESA analysis, immune cell infiltration analysis, and copy number variation analysis of the IRGs in GC. (A–D) Gene set enrichment analysis (GSEA). (A) Angiogenesis. (B) Epithelial-mesenchymal transition. (C) Hypoxia. (D) UV response DN. (E) Correlation between 8 hub IRGs and immune cell infiltration. The fold change of the high-risk group compared with the low-risk group was shown as the number in the circle. Immune cell infiltration analysis indicates an immunosuppressive microenvironment, mediated by increased tumor M2 macrophage infiltration and decreased dendritic cell activation in the high-risk group. (F–G) Copy number variation analysis and gene mutation analysis. (F) Stem plot of copy number variation analysis. (G) Distribution of highly variant mutated genes correlated with signature. (H–I) Kaplan–Meier survival analysis and Time-dependent ROC curve in the external validation cohorts (GSE84437), from the GEO (https://www.ncbi.nlm.nih.gov/geo/) database. (H) Kaplan–Meier survival analysis between the high-risk group and low-risk group, with a *P*-value = 0.003. (I) AUC in ROC analysis for signature at the 1-, 3-, and 5-year survival time, with an AUC = 0.613, 0.600, and 0.610, respectively.

the immune microenvironment. The most enriched biological processes including cell surface receptor-linked signal transduction and immune response exhibited the immense role of the DEIGs in the regulation of the immune microenvironment. Enriched cell components, molecular functions, and pathways also attached the importance of immune regulators in tumor progression. GSEA analysis further confirmed that 8 hub immune genes contributed to tumorigenesis and tumor progression. Immune cell infiltration analysis indicates increased macrophage M2 and decreased dendric cell infiltration in the high-risk group, suggesting the suppressive immune microenvironment may lead to a worse prognosis.

The hub IRGs were proven to regulate tumor immune microenvironment in literature. Integrin αv (ITGAV) belongs to the integrin

family, one of the transmembrane receptors that take part in cellular interaction with the extracellular matrix, which is related to the growth, migration, and invasion of GC cells [15]. Integrin has been reported upregulated in GC mediated by M2 macrophages [16], which contributes to the immunosuppressive microenvironment and serves as the potential target for immunotherapy. MSR1 (macrophage scavenger receptor1) is another target of tumor-associated macrophages. Monoclonal antibodies targeting MSR1 enhance NK cell activation, NK cell-mediated killing, and reverse immunosuppressive tumor microenvironment [17]. GNAI1 (G protein subunit αi1) is relevant to GTP binding and signal transduction activity [18]. GNAI1 regulates CXCL12-CXCR4/CXCR7 chemokine axis and leads to immune resistance in gastrointestinal malignancies [19]. NMB (Neuromedin B) is a negative regulator for Type 2 innate lymphocytes and inhibits immune response [20]. CCL15 recruits myeloid-derived suppressive cells (MDSCs) to the tumor tissue, resulting in an immunosuppressive microenvironment. The immunosuppressive microenvironment accounts for shorter relapse-free survival (RFS) and poor prognosis [21]. AEN (apoptosis-enhancing nuclease) is a mediator of p53-dependent apoptosis [22]. Increased NR3C1 (Nuclear Receptor Subfamily 3 Group C Member 1, Glucocorticoid Receptor) expression is found in $CD8^+$ tumor-infiltration lymphocytes [23]. TGFBR1 (TGFβ receptor1) takes part in TGFβ signaling, which drives immune evasion that promotes T cell exclusion [24]. Overall, most genes are relevant to immune system regulation [25]. On the one hand, abnormal gene expression of tumors leads to an immunosuppressive microenvironment; on the other hand, the repurposed microenvironment supports tumor progression in return [26].

Though previous studies have proposed several signatures for prognosis [27–29], no genome-wide profiling studies of GC have been performed. Due to small datasets and the absence of sufficient validation, none of them was utilized in clinical practice [30]. In our work, differentially expressed IRGs in GC were investigated. We constructed and validated a novel immune-related predictable signature for GC patients. We utilized univariate Cox regression analysis and multivariate Cox regression analysis for DEIGs screening rather than WGCNA combined with the PPI network. The analysis mode of WGCNA and PPI is to classify genes through data processing and classify genes with similar expression patterns or interacting genes into one module for further classification [31,32]. However, such a screening procedure may ignore some hub genes that do not exist in modules or have no interaction but are equally important. In this study, we performed expressional differences analysis, univariate Cox regression analysis, and multivariate Cox regression analysis for targeted gene screening. These genes were further validated via clinical correlation. Therefore, we confirmed the genes that predict the prognosis of patients. Mendelian randomization analysis is another method for causal relationship analysis [33,34]. It provides insight into the causal relationship between exposure factors and epidemiology outcomes on gene differences level [35]. However, in limited sample size, such a strategy may lead to unfavorable bias [36].

The limitations of our study still existed. Firstly, datasets with sequencing information would be better. Secondly, our attention was mainly paid to gene expression and gene mutation level, while other genetic behaviors such as epigenetic regulation and protein translation were outside the scope of this article. For a similar reason, we didn't perform WGCNA combined with the PPI network. Moreover, due to the limitation of sample size, mendelian randomization analysis was not utilized. In addition, the effectiveness of the signature needs to be proven in clinical practice and prospective studies should be conducted to test the utility in the individualized management of GC. Nevertheless, our analysis provides valuable information about GC survival outcomes.

Collectively, we established and validated the prognostic signature in the training cohort and validated cohort respectively, which effectively divides GC patients into different risk groups. Therefore, individuals with worse predicted prognoses can be identified for personalized therapeutic interventions before the disease advances. With a high performance (P-value<0.001, AUC $= 0.803$) and validated prediction, the nomogram could serve as a prognostic stratification tool for GC patients to assist in decision-making and individual treatment scheduling. Moreover, all genes in the signature were proved to contribute to the immunosuppressive microenvironment. Based on these results we may conclude that IRGs themself may be a manipulated variable in the tumor microenvironment. And the immune-related predictable instrument for prognosis will be helpful to clinical guidance and personalized immunotherapy development.

5. Conclusions

We constructed an eight-gene signature and a new nomogram, which can be used as prognostic stratification tools for GC patients to assist in decision-making and individual treatment scheduling. In addition, the relationship between genes and tumor immune microenvironment was also taken into account in the research, which may facilitate immunotherapy progression.

Ethics approval and consent to participate

All data were obtained from public databases and used following the Helsinki Declaration II. **We confirmed all methods were carried out following relevant guidelines and regulations.**

Data availability statement

The TCGA data were obtained from GDC data portal (https://portal.gdc.cancer.gov/). The immune-related genes (IRGs) data were downloaded from the ImmPort database (https://immport.niaid.nih.gov). The R software (https://www.r-project.org/) was used for all statistical analyses. The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request without breaching participant confidentiality.

Funding

This work was supported by the National Natural Science Foundation of China (82172645), Natural Science Foundation of Jiangsu Province for Outstanding Young Scholars (BK20200052), Key Project of Nanjing Health Commission (ZKX21013), Bethune Charitable Foundation (05002), Projects of the Guangxi Health Commission (Z-A20230464), and Clinical Trials from the Affiliated Drum Tower Hospital Medical School of Nanjing University (2021-LCYJ-MS-09 and 2021-LCYJ-PY-17).

CRediT authorship contribution statement

Xianghui Li: Conceptualization, Writing – original draft. **Yuanyuan Chen:** Data curation. **Yuxiang Dong:** Formal analysis. **Zhongjin Ma:** Methodology. **Wenjun Zheng:** Visualization. **Youkun Lin:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e22433.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA A Cancer J. Clin. 68 (6) (2018) 394–424.
[2] W. Koizumi, H. Narahara, T. Hara, A. Takagane, T. Akiya, M. Takagi, K. Miyashita, T. Nishizaki, O. Kobayashi, W. Takiyama, Y. Toh, T. Nagaie, S. Takagi, Y. Yamamura, K. Yanaoka, H. Orita, M. Takeuchi, S-1 plus cisplatin versus S-1 alone for first-line treatment of advanced gastric cancer (SPIRITS trial): a phase III trial, Lancet Oncol. 9 (3) (2008) 215–221.
[3] S. Bhattacharya, P. Dunn, C. Thomas, B. Smith, H. Schaefer, J. Chen, Z. Hu, K. Zalocusky, R. Shankar, S. Shen-Orr, E. Thomson, J. Wiser, A. Butte, ImmPort, toward repurposing of open access immunological assay data for translational and clinical research, #N/A 5 (2018), 180015.
[4] K. Jin, S. Qiu, D. Jin, X. Zhou, X. Zheng, J. Li, X. Liao, L. Yang, Q. Wei, Development of prognostic signature based on immune-related genes in muscle-invasive bladder cancer: bioinformatics analysis of TCGA database, Aging 12 (2021).
[5] S. Duan, P. Wang, F. Liu, H. Huang, W. An, S. Pan, X. Wang, Novel immune-risk score of gastric cancer: a molecular prediction model combining the value of immune-risk status and chemosensitivity, #N/A 8 (5) (2019) 2675–2685.
[6] G. Yu, L. Wang, Y. Han, Q. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS A J. Integr. Biol. 16 (5) (2012) 284–287.
[7] T. Wu, E. Hu, S. Xu, M. Chen, P. Guo, Z. Dai, T. Feng, L. Zhou, W. Tang, L. Zhan, X. Fu, S. Liu, X. Bo, G. Yu, clusterProfiler 4.0: a universal enrichment tool for interpreting omics data, Innovation 2 (3) (2021), 100141.
[8] Z. Zhu, J. Xu, L. Li, W. Ye, B. Chen, J. Zeng, Z. Huang, #N/A, CTHRC1Comprehensive Analysis Reveals , , and as the Novel Prognostic Markers in Gastric Cancer, vol. 9, 2020, pp. 4093–4110 (7).
[9] X. Gu, Y. Jiang, W. Wang, J. Zhang, D. Shang, C. Sun, J. Tian, J. Tian, B. Yu, Y. Zhang, Comprehensive circRNA expression profile and construction of circRNA-related ceRNA network in cardiac fibrosis, Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie 125 (2020), 109944.
[10] R.L. Siegel, K.D. Miller, A. Jemal, Cancer Statistics, 2019, CA Cancer J. Clin, 2019.
[11] I. Mellman, G. Coukos, G. Dranoff, Cancer immunotherapy comes of age, Nature 480 (7378) (2011) 480–489.
[12] R. Wang, M. Dang, K. Harada, G. Han, F. Wang, M. Pool Pizzi, M. Zhao, G. Tatlonghari, S. Zhang, D. Hao, Y. Lu, S. Zhao, B.D. Badgwell, M. Blum Murphy, N. Shanbhag, J.S. Estrella, S. Roy-Chowdhuri, A.A.F. Abdelhakeem, Y. Wang, G. Peng, S. Hanash, G.A. Calin, X. Song, Y. Chu, J. Zhang, M. Li, K. Chen, A. J. Lazar, A. Futreal, S. Song, J.A. Ajani, L. Wang, Single-cell dissection of intratumoral heterogeneity and lineage diversity in metastatic gastric adenocarcinoma, Nat. Med. 27 (1) (2021), 141-+.
[13] Y. Oya, Y. Hayakawa, K. Koike, Tumor microenvironment in gastric cancers, Cancer Sci. 12 (2021).
[14] H. Yu, R. Pan, Y. Qi, Z. Zheng, J. Li, H. Li, J. Ying, M. Xu, S. Duan, LEPR hypomethylation is significantly associated with gastric cancer in males, Exp. Mol. Pathol. 116 (2020).
[15] H.S. Wang, H.L. Chen, Z.P. Jang, Y.J. Lin, X.Y. Wang, J. Xiang, J.S. Peng, Integrin subunit alpha V promotes growth, migration, and invasion of gastric cancer cells, #N/A 215 (9) (2019) 6.
[16] D. Ngabire, I. Niyonizigiye, M.P. Patil, Y.-A. Seong, Y.B. Seo, G.-D. Kim, M2 Macrophages Mediate the Resistance of Gastric Adenocarcinoma Cells to 5-Fluorouracil through the Expression of Integrin Beta 3, Focal Adhesion Kinase, and Cofilin, 2020. #N/A 2020.
[17] S. Eisinger, D. Sarhan, V.F. Boura, I. Ibarlucea-Benitez, S. Tyystjarvi, G. Oliynyk, M. Arsenian-Henriksson, D. Lane, S.L. Wikstrom, R. Kiessling, T. Virgilio, S. F. Gonzalez, D. Kaczynska, S. Kanatani, E. Daskalaki, C.E. Wheelock, S. Sedimbi, B.J. Chambers, J.V. Ravetch, M.C.I. Karlsson, Targeting a scavenger receptor on tumor-associated macrophages activates tumor cell killing by natural killer cells, Proc. Natl. Acad. Sci. U. S. A. 117 (50) (2020) 32005–32016.
[18] Y.J. Shen, S.K. Dong, J.H. Liu, L.Q. Zhang, J.C. Zhang, H. Zhou, W.D. Dong, Identification of Potential Biomarkers for Thyroid Cancer Using Bioinformatics Strategy: A Study Based on GEO Datasets, 2020, p. 21. #N/A 2020.
[19] S.K. Daniel, Y.D. Seo, V.G. Pillarisetty, The CXCL12-CXCR4/CXCR7 axis as a mechanism of immune resistance in gastrointestinal malignancies, Semin. Cancer Biol. 65 (2020) 176–188.
[20] J.M. Inclan-Rico, J.J. Ponessa, N. Valero-Pacheco, C.M. Hernandez, C.B. Sy, A.D. Lemenze, A.M. Beaulieu, M.C. Siracusa, Basophils prime group 2 innate lymphoid cells for neuropeptide-mediated inhibition, Nat. Immunol. 21 (10) (2020) 1181–1193.

[21] B.H. Li, M.A. Garstka, Z.F. Li, Chemokines and their receptors promoting the recruitment of myeloid-derived suppressor cells into the tumor, Mol. Immunol. 117 (2020) 201–215.
[22] T. Kawase, H. Ichikawa, T. Ohta, N. Nozaki, F. Tashiro, R. Ohki, Y. Taya, p53 target gene AEN is a nuclear exonuclease required for p53-dependent apoptosis, Oncogene 27 (27) (2008) 3797–3810.
[23] N. Acharya, A. Madi, H. Zhang, M. Klapholz, G. Escobar, S. Dulberg, E. Christian, M. Ferreira, K.O. Dixon, G. Fell, K. Tooley, D. Mangani, J. Xia, M. Singer, M. Bosenberg, D. Neuberg, O. Rozenblatt-Rosen, A. Regev, V.K. Kuchroo, A.C. Anderson, Endogenous glucocorticoid signaling regulates $CD8^{+}$ T cell differentiation and development of dysfunction in the tumor microenvironment, Immunity 53 (3) (2020), 658-671.e6.
[24] D.V.F. Tauriello, S. Palomo-Ponce, D. Stork, A. Berenguer-Llergo, J. Badia-Ramentol, M. Iglesias, M. Sevillano, S. Ibiza, A. Canellas, X.H. Momblona, D. Byrom, J.A. Matarin, A. Calon, E.I. Rivas, A.R. Nebreda, A. Riera, C.S.O. Attolini, E. Batlle, TGF beta drives immune evasion in genetically reconstituted colon cancer metastasis, Nature 554 (7693) (2018) 538–+.
[25] K.H. Pak, K.C. Park, J.H. Cheong, VEGF-C induced by TGF- beta1 signaling in gastric cancer enhances tumor-induced lymphangiogenesis, BMC Cancer 19 (1) (2019) 799.
[26] C. Lyssiotis, A. Kimmelman, Metabolic interactions in the tumor microenvironment, Trends Cell Biol. 27 (11) (2017) 863–875.
[27] C. Zhang, L. Jing, Z. Li, Z. Chang, H. Liu, Q. Zhang, Q. Zhang, Identification of a prognostic 28-gene expression signature for gastric cancer with lymphatic metastasis, Biosci. Rep. 39 (5) (2019).
[28] K.C. Nie, Z.T. Deng, Z.H. Zheng, Y. Wen, J.L. Pan, X.T. Jiang, Y.H. Yan, P. Liu, F.B. Liu, P.W. Li, Identification of a 14-lncRNA signature and construction of a prognostic nomogram predicting overall survival of gastric cancer, DNA Cell Biol. 13 (2021).
[29] E. Zhao, C. Zhou, S. Chen, A signature of 14 immune-related gene pairs predicts overall survival in gastric cancer, Clin. Transl. Oncol. 10 (2021).
[30] B. Li, Y. Cui, M. Diehn, R. Li, Development and validation of an individualized immune prognostic signature in early-stage nonsquamous non-small cell lung cancer, #N/A 3 (2017) 1529–1537.
[31] J. Chen, X. Zhao, L. Cui, G. He, X. Wang, F. Wang, S. Duan, L. He, Q. Li, X. Yu, F. Zhang, M. Xu, #N/A, Genetic Regulatory Subnetworks and Key Regulating Genes in Rat hippocampus Perturbed by Prenatal Malnutrition: Implications for Major Brain Disorders, vol. 12, 2020, pp. 8434–8458.
[32] H. Li, X. Wang, X. Lu, H. Zhu, S. Li, S. Duan, X. Zhao, F. Zhang, G. Alterovitz, F. Wang, Q. Li, X.-L. Tian, M. Xu, #N/A, Co-expression Network Analysis Identified Hub Genes Critical to Triglyceride and Free Fatty Acid Metabolism as Key Regulators of Age-Related Vascular Dysfunction in Mice, vol. 11, 2019, pp. 7620–7638.
[33] X. Wang, X. Fang, W. Zheng, J. Zhou, Z. Song, M. Xu, J. Min, F. Wang, Genetic support of A causal relationship between iron status and type 2 diabetes: a mendelian randomization study, #N/A 106 (11) (2021) E4641–E4651.
[34] L. Hou, M. Xu, Y. Yu, X. Sun, X. Liu, L. Liu, Y. Li, T. Yuan, W. Li, H. Li, F. Xue, Exploring the causal pathway from ischemic stroke to atrial fibrillation: a network Mendelian randomization study, Mol. Med. 26 (1) (2020).
[35] F. Zhang, A. Baranova, C. Zhou, H. Cao, J. Chen, X. Zhang, M. Xu, Causal influences of neuroticism on mental health and cardiovascular disease, Hum. Genet. 140 (9) (2021) 1267–1281.
[36] D. Adam, The causation detector, Nature 576 (7786) (2019) 196–199.