Diagnostic Pathology: Open Access
ISSN: 2476-2024
Wei, et al., Diagn Pathol Open 2025, 10:1
Research Article
Open Access

Identification of Lysosome-Related Genes in Gastric Cancer and Development of a Prognostic Model

Zihao Wei, Dongmei Luo*, Zhengrui Wang and Chen Zi
Department of Microelectronics and Data Science, Anhui University of Technology, Anhui, China

Abstract

Background: Despite significant advancements in the diagnosis and treatment of Gastric Cancer (GC), it remains a major contributor to global cancer-related mortality, characterized by high rates of distant recurrence and late-stage fatality.

Methods: We obtained gene expression data for the GC cohort (TCGA-STAD) from the Genomic Data Public Portal (GDC) and lysosome-related genomic data from AMIGO. We used the GSE66259 dataset from the Gene Expression Omnibus (GEO) database for external validation. Initially, we screened GC tissues for Differentially Expressed Genes (DEGs) compared to normal tissues and overlapped DEGs with lysosomal genes to obtain Lysosome-Related Genes (LRGs). Key LRGs were then selected for prognostic modeling through univariate Cox regression, Least Absolute Shrinkage and Selection Operator (LASSO) regression, and multivariate Cox stepwise regression. The prognostic models were evaluated using Kaplan-Meyer (K-M) analysis and Receiver Operating Characteristic (ROC) curves. Independent prognostic factors were also identified by univariate Cox regression and multivariate Cox regression analyses. Additionally, we investigated biological functions, response to Immune Checkpoint Inhibitors (ICIs), as well as Tumor Mutational Load (TMB) mutations in high and low-risk groups.

Results: The univariate Cox analysis identified a total of 48 lysosomal genes with significant differential expression. Subsequently, LASSO regression was employed to select 11 prognostic genes. Finally, a multivariate Cox stepwise regression yielded a set of 7 genes (*TRIM29, EGF, GPC3, RETN, RNASE3, GRP, and PSAL1*) for constructing the prognostic models. The validated risk model demonstrated accurate prediction of clinical outcomes. Furthermore, independent prognostic analyses revealed that the risk score along with stage and age were independent prognostic factors. Notably, significant differences in biological function, immune microenvironment characteristics, as well as immunotherapy response, were observed between the high and low-risk groups.

Conclusions: The identification of 7 key prognostic LRGs associated with GC patients facilitates accurate prognosis prediction and presents a novel avenue for enhancing clinical management and prognostic outcomes in the GC patient population.

Keywords: Gastric Cancer; Lysosome-Related Genes (LRGs); Prognostic model; Immune infiltration

Abbreviations: GC: Gastric cancer; OS: Overall survival; LDCD: Lysosome-Dependent Cell Death; TCGA-STAD: The Cancer Genome Atlas-Stomach Adenocarcinoma; GDC: Genomic Data Commons; GEO: Gene Expression Omnibus; DEGs: Differential Expressed Genes; LRGs: Lysosome-Related genes; LASSO: Least Absolute Shrinkage and Selection Operator; K-M: Kaplan-Meyer; ROC: Receiver Operating Characteristic; AUC: Area under the Curve; ICIs: Immune Checkpoint Inhibitors; TMB: Tumor Mutational; EGF: Epidermal Growth Factor; RNA: Glypican-3; GPC3: Ribonuclease 3; GRP: Gastrin Releasing Peptide; RETN: Resistin; PSAL1: Prostate Cancer Susceptibility Protein 1; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; CC: Cellular Components; MF: Molecular Functions; BP: Biological Processes; GSEA: Gene Set Enrichment Analysis; TRIM29: Tripartite Motif Containing 29; SNPs: Single Nucleotide Polymorphisms

Introduction

Gastric Cancer (GC) ranks as the fifth most prevalent malignancy worldwide and stands as the third leading cause of cancer-related mortality [1]. The incidence of GC remains alarmingly high, accompanied by a dismal prognosis. As per global statistics in 2020, over a million new cases of GC were reported, with approximately 769,000 patients succumbing to this disease [2,3]. Despite notable advancements in screening and treatment modalities for GC, its fatality rates and distant recurrence remain elevated, and the median Overall Survival (OS) for individuals with advanced GC merely reaches 14.2 months [4,5]. Consequently, it becomes imperative to identify more precise and effective cellular molecular markers that can enhance patient outcomes while alleviating their burden.

***Corresponding author:** Dongmei Luo, Department of Microelectronics and Data Science, Anhui University of Technology, Anhui, China; E-mail: luodamhut@126.com

Received: 23-January-2024, Manuscript No. DPO-24-125664; **Editor assigned:** 25-January-2024, PreQC No. DPO-24-125664 (PQ); **Reviewed:** 08-February-2024, QC No. DPO-24-125664; **Revised:** 15-January-2025, Manuscript No. DPO-24-125664 (R); **Published:** 22-January-2025, DOI: 10.4172/2476-2024.10.1.246

Citation: Wei Z, Luo D, Wang Z, Zi C (2025) Identification of Lysosome-Related Genes in Gastric Cancer and Development of a Prognostic Model. Diagn Pathol Open 10: 246.

Lysosomes, membrane-enclosed organelles, play a crucial role in intracellular macromolecule degradation, leading to their previous designation as "rubbish disposals" [6-9]. Further investigations have revealed that lysosomes also serve essential functions in cellular adaptation to various signaling stimuli and act as pivotal signaling hubs for nutrient responses and signal transduction [10]. Dysregulation of pH and dysfunction within lysosomes have been associated with human longevity, cellular senescence, Parkinson's disease, and Alzheimer's disease [11,12]. Additionally, lysosomes modulate the proliferation of tumor cells through growth factor signaling regulation, hence malignant cell phenotypes in certain tumor-related diseases are linked to altered lysosomal activity [13]. Notably, researchers have proposed utilizing Lysosome-Dependent Cell Death (LDCD) as a potential strategy for eliminating cancer cells due to observed alterations in the lysosomal structure in some cancers. Furthermore, lysosomes play a critical role in the development of resistance to radiation and chemotherapy in tumor cells [14]. Unfortunately, there is a dearth of existing literature specifically investigating the correlation between GC cells and Lysosome-Related Genes (LRGs).

The primary objective of this study is to analyze the expression patterns of LRGs in GC cells and establish a prognostic model for LRGs. Additionally, we also investigate the correlation between the prognostic model and immune response as well as mutational status, thereby exploring the feasibility of our proposed model. Overall, the prognostic LRGs model developed in this research can effectively predict the prognosis of GC patients and contribute to relevant clinical treatment strategies. The analytical workflow employed in this study is illustrated in Figure 1.

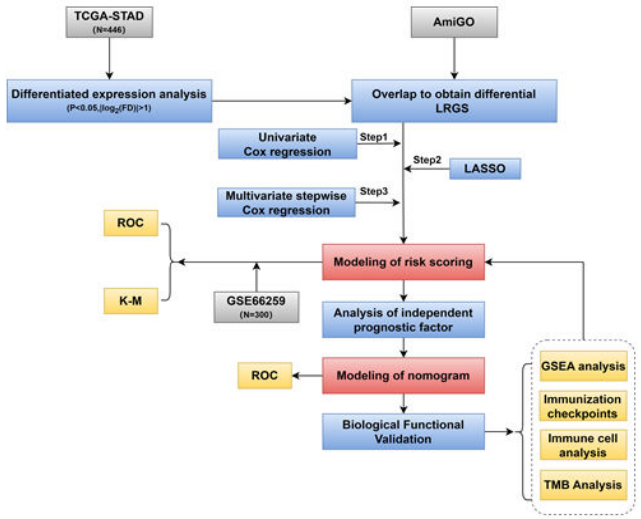

Figure 1: The workflow of the current study.

Materials and Methods

Database of GC patients

The gene expression data of GC and normal samples, as well as the clinical information data (TCGA-STAD), were obtained from the TCGA database (https://portal.gdc.cancer.gov/). The TCGA-STAD dataset comprised 410 GC tissues and 36 normal tissues. Additionally, the gene expression data of the GSE66259 gastric cancer cell dataset was acquired from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), which included 300 GC tissues and 100 normal tissues. The clinical information data for the GSE66259 dataset was retrieved from the literature [15-16]. A total of 875 LRGs were obtained from AmiGO2 (http://amigo.geneontology.org/amigo) and analyzed using R 4.2.2 software.

Identification of differential prognostic LRGs in STAD

We initially normalized the gene expression data using the "limma" package, followed by conducting differential expression analysis on 36 normal tissues and 410 GC tissues utilizing the "edgeR" package. The Differentially Expressed Genes (DEGs) in TCGA-STAD were identified based on a cut-off criteria of p-value<0.05, and |log2(fold change)|>1. Subsequently, we obtained the differential prognostic LRGs by overlapping the DEGs with lysosomal genes. Visualization of DEG volcanoes was performed using the "ggplot2" package, while Venn diagram were employed to illustrate overlapped LRGs. Correlation coefficients between genes were depicted using both the "corrplot" and "heatmap" packages, respectively, with clustering spectra also being generated.

To investigate the relevant biological functions of LRGs, we initially employed the "org.hs.egg.db" package to convert gene symbols into EntrezIDs. Subsequently, we utilized the enrichment function from the "ClusterProfiler" package for conducting Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis and the GO enrichment analysis included Cellular Components (CC), Molecular Functions (MF), and Biological Processes (BP).

Construction and evaluation of the risk model

We initially filtered the sample data in TCGA-STAD (Filtering criteria: Exclude the samples with incomplete survival information or survival time of less than 30 days). And finally obtained 415 TCGA-STAD samples, and performed log2 transformation on the gene expression data. These 415 samples were used as the training set for TCGA-STAD analysis. Similarly, we applied the same process to the GSE66259 dataset and identified 300 samples after screening as the external validation set.

To identify differential prognostic LRGs, univariate Cox regression analysis was conducted using the "survival" package, resulting in a selection of 48 genes (p-value<0.05). Further screening was performed using least absolute shrinkage and selection operator regression (LASSO) with parameters set as follows: family="Cox", alpha="1", leading to the identification of 11 genes (p-value<0.05). Multivariate Cox analysis was then employed on these 11 genes, followed by stepwise method-based screening to identify critical prognostic genes and to obtain risk coefficients for each of them. Ultimately, 7 key LRGs were identified from this process. A risk score model was established based on these 7 LRGs using the formula:

$$\text{Risk score}=\sum_{i=1}^{n} coei \times xi$$

Where, coei denotes the multivariate Cox regression coefficient of each prognostic LRGs, and xi denotes its corresponding expression level.

This formula was utilized to calculate individual risk scores for both training and validation sets.

To evaluate the prognostic value of the risk model, we initially categorized the samples into high-risk and low-risk groups based on the median risk score. Subsequently, we employed the "survival", "timeROC", and "ggplot2" packages to generate Receiver Operating Characteristic (ROC) curves for subjects at 1, 3, and 5-year survival time. The Area under the ROC Curve (AUC) was utilized to assess the predictive performance of our risk model. Additionally, Kaplan-Meier survival curve analysis was conducted to compare the survival patterns between high-risk and low-risk groups predicted by our model. Furthermore, the external dataset GSE66259 was employed for validation purposes.

Analysis of independent prognosis factors and construction of the nomogram model

To assess whether our constructed risk scores and other clinical characteristics of TCGA-STAD patients (gender, age, T-stage (T1, T2, T3, T4), N-stage (N0, N1, N2, N3), M-stage (M0, M1, M2), and staging) were independent prognostic factors, we initially identified significant clinicopathological factor for predicting GC patients' prognosis using univariate Cox regression analysis. Subsequently, clinical factors with p-value<0.05 were further selected through stepwise Cox regression employing the "cpxh" function ("rms" package). Finally obtained clinical factors were utilized to construct a prognostic model and develop a nomogram for survival prediction using multivariate Cox regression. The performance of the nomogram model was validated by ROC curves at 1-year, 3-year, and 5-year time.

Gene enrichment analysis of high-risk and low-risk groups

To investigate the biological differences between the high-risk and low-risk groups, we conducted a Gene Set Enrichment Analysis (GSEA) on the TCGA high-risk and low-risk groups. Subsequently, all DEGs in these risk groups were subjected to GO enrichment and KEGG pathway analysis using the "clusterProfiler" software package. The results were visualized using "path view" and "enrichplot" software packages.

Immunotherapy analysis and mutation analysis in risk groups

To investigate the distribution of immune cells surrounding tumor cells, we utilized function "subsea" in "GSVA" package to quantify the abundance of 28 immune cells in GC samples. Additionally, leveraging nucleotide mutation data for STAD patients, we employed the "TMB" function ("mdftools" package) to calculate the Tumor Mutation Load (TMB) for each STAD sample. Subsequently, we visualized these mutations to discern differences in TMB between low-risk and high-risk groups.

Results

Identification of differential prognostic LRGs in GC cells

The idea of our study is illustrated in Figure 1. To identify differentially expressed LRGs in GC, we initially obtained 410 GC and 36 normal tissues from the TCGA-STAD dataset, followed by conducting differential expression analysis using the "edgeR" package (p-value<0.05, |log2(fold change)|>1). This analysis revealed a total of 4172 genes (2051 up-regulated and 2121 down-regulated, as shown in Figure 2A) that exhibited significant differential expression between GC and normal tissues. Subsequently, we overlapped these identified DEGs with the set of 875 known lysosomal genes available on the AmiGO2 database, resulting in the identification of a subset comprising 145 differentially prognostic LRGs (as depicted in Figure 2B). Furthermore, to facilitate further investigation, we ranked these differentially prognostic LRGs based on their p-values in ascending order and generated correlation coefficients along with heat maps for the top 10 differentially prognostic LRGs (Figures 2C and D). The ACAN and ACE genes are strongly correlated, with mutations in the ACAN gene associated with spinal dysplasia and the ACE gene associated with hypertension and kidney disease (Figure 2C) [17]. There is a gap in gene expression patterns in the two states of tumor and normal (Figure 2D) [18].

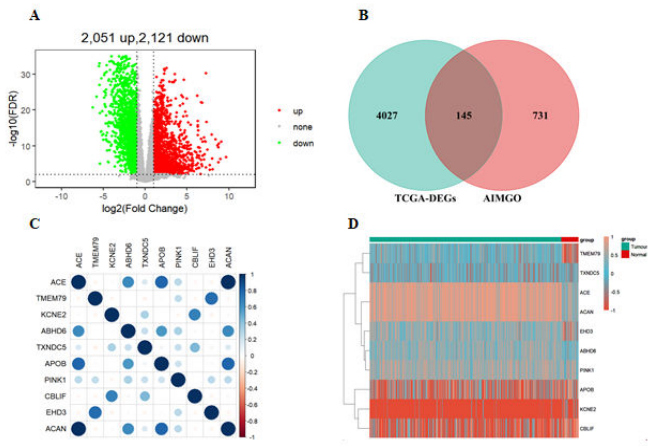

Figure 2: Expression of lysosome-related genes in gastric cancer and normal tissues. (A) Volcano plots of differentially expressed genes. (B) Wayne plots of differentially expressed genes in TCGA-STAD associated with lysosomal genes in AMIGO2. (C) Correlation coefficient plots of the top 10 differential bases. (D) Hierarchical clustering of the top 10 differentially expressed lysosomal genes.

Functional analysis of LRGs

To investigate the relevant pathways associated with 145 differentially prognostic LRGs in tumor development and progression, we conducted enrichment analysis of LRGs using GO and KEGG (Figures 3A and B). A total of 733 GO-enriched terms were identified. Under Biological Process (BP), the pathways primarily enriched included transition metal ion transport, positive regulation of protein catabolic process, and lysosomal transport. Under Cellular Component (CC), the enriched terms consisted of azurophil granule lumen, late endosome, lysosomal lumen, and secretory granule lumen, among others. Under Molecular Function (MF), the main enrichments were observed for proteinase binding and glycosaminoglycan binding (Figure 3A). KEGG enrichment analysis revealed significant involvement of neuroactive ligand-receptor interaction pathway as well as protein digestion and absorption pathway along with other 34 pathways in GC tumorigenesis (Figure 3B). Furthermore, a network diagram was constructed to visualize the associations within this pathway set where lysosomal transport, transition metal ion transport, regulation of trans-synaptic signaling, myeloid leukocyte activation, vacuolar transport emerged as highly interconnected components

(Figure 3C). In summary, we have identified a panel of 145 differentially prognostic LRGs in GC that are significantly enriched in key pathways involved in tumor development. These findings provide valuable insights into understanding the role of LRGs.

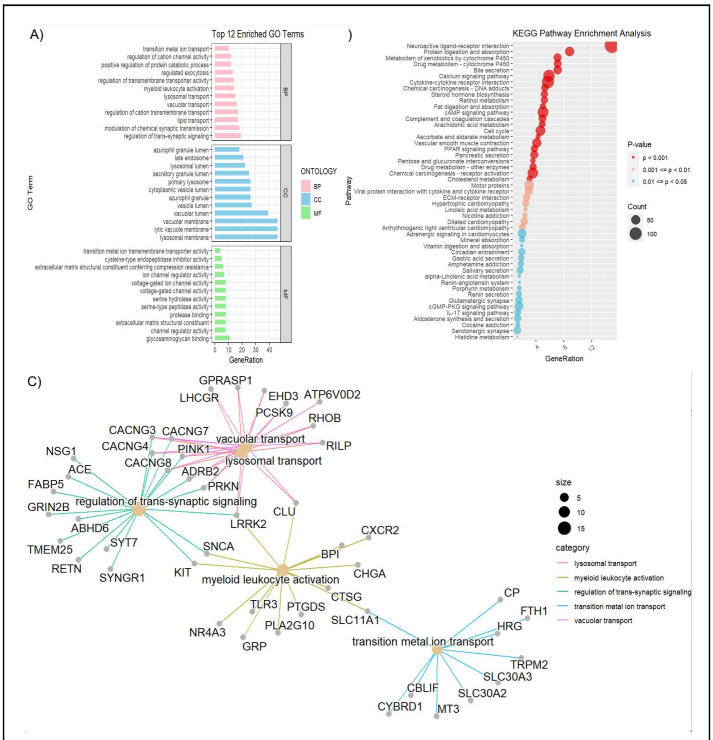

Figure 3: KEGG and GO enrichment analyses of differentially prognosticated lysosome-related genes and PPI networks. (A) GO enrichment analysis of 145 differentially prognostic LRGs. BP: Biological Progress; CC: Cellular Component; MF: Molecular Function. (B) KEGG pathway analysis of the 145 differentially prognostic LRGs. (C) Association network diagram of pathway sets.

Construction of risk model

We initially performed univariate Cox regression analysis on the 145 differential prognostic LRGs, identifying 48 genes with prognostic significance (p-value<0.05). Subsequently, we employed LASSO regression to further screen these 48 prognostic genes and obtained a final set of 11 genes that exhibited significant prognostic significance based on the smallest cross-validation error (Figures 4A and B). Multivariate stepwise Cox regression analysis revealed that a predictive model comprising of 7 genes (*TRIM29, EGF, GPC3, RETN, RNASE3, GRP*, and *PSAL1*) demonstrated the highest prognostic value in determining patient outcomes, and risk score was calculated using the formula: Risk score=0.05964 × TRIM29+0.08971 × EGF +0.10998GPC3+0.08204×RETN+0.08986×RNASE3+0.06407×GRP +0.04003×PSAPL1. After constructing the risk model, we divided our sample into high-risk and low-risk groups by utilizing the median risk score as a threshold. Risk curves based on these risk scores predicted a poorer prognosis for patients in high-risk group (Figure 4C). The AUC values of the ROC curves at 1, 3, and 5-years in our training set were all greater than 0.64, notably, the 3-year AUC value exceeded 0.7 (Figure 4D). Additionally, Kaplan-Meier curves demonstrated higher survival rates among patients in the low-risk group compared to those in the high-risk group (Figure 4E).

In our validation set, the survival outcomes of patients in the high-risk group remained inferior to those of patients in the low-risk group (Figure 4F). The AUC values of the ROC curves in the validation set were consistently above 0.6 (Figure 4H), indicating that our risk scores exhibited excellent predictive performance. Furthermore, significant differences were observed in the K-M survival curves between high-risk and low-risk groups (Figure 4I).

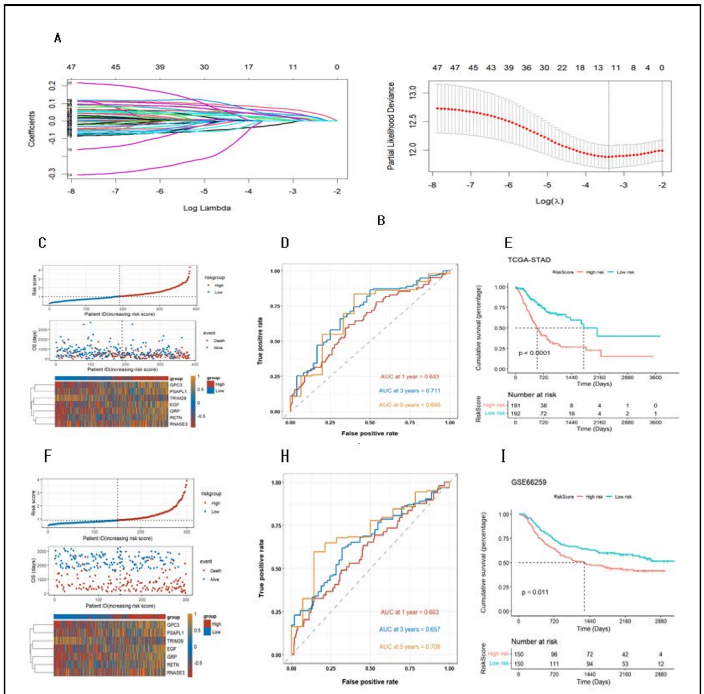

Figure 4: Construction of the risk model. (A,B) Distribution plots of LASSO coefficients for differential prognostic genes. (C) Risk curves in the TCGA training set, scatter plots of survival and death, and heat maps of model gene expression in the high and low-risk groups. (D) ROC curves at 1, 3, and 5 years in the training set. (E) K-M survival analysis curves for the high-risk and low-risk groups in the training set (F) Risk curves, scatter plots of survival and death, and heat maps of model gene expression in the high-risk and low-risk groups in the GSE66259 external validation set. (H) ROC curves at 1, 3, and 5 years for the GSE66259 external validation set. (I) K-M survival analysis curves for the high-risk and low-risk groups of the GSE66259 external validation set.

Independent prognosis factors and accurate prediction of the nomogram

The risk score was subjected to univariate Cox regression analysis along with other clinical factors in the GC samples, revealing a significant association between the risk score, stage, and age with GC (Figure 5A). Subsequently, these significantly associated factors were further analyzed using multivariate Cox regression analysis. The findings revealed that the risk score, stage, and age could serve as independent prognostic factors for patients with GC (p-value<0.05) (Figure 5B). Notably, the significance of stage may be attributed to its representation of N-stage indicating lymph node involvement. Furthermore, a nomogram model incorporating risk score, stage, and age was constructed (Figure 5C), exhibiting a consistency index of 0.681 on the nomogram. Additionally, the AUC values of ROC curves

in predicting 1, 3 and 5-year survival time exceeded 0.69 (Figure 5D), thereby confirming high accuracy in survival prediction.

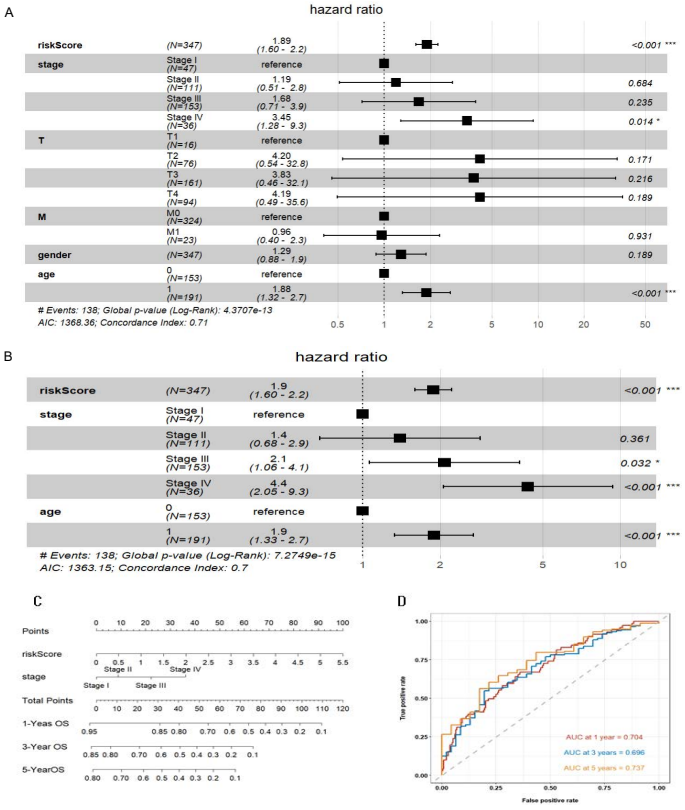

Figure 5: Relationship between risk scores and clinicopathological characteristics. (A) Forest plot of univariate Cox regression analysis of risk score and clinicopathological characteristics. (B) Forest plot of multivariate Cox regression analysis for significantly correlated characteristics in univariate Cox regression analysis. (C) Column line plots of constructed independent prognostic models. (D) ROC curves for predicting 1, 3, and 5-year patient survival using column-line plots. *p-value<0.05, **p-value<0.01, ***p-value<0.001.

Functional and enrichment analyses in high and low-risk groups

To further investigate the difference in gene function and pathways between the high-risk and low-risk groups, we analyzed the impact of risk scores on tumor development using GSEA. GSEA analysis revealed that significant enrichment of epidermal cell differentiation, epidermis development, keratinization, and keratinocyte differentiation in the high-risk group for BP term in GO analysis (Figure 6A), as well as cell cycle, cytokine-cytokine receptor interaction, fat digestion and absorption, neuroactive ligand-receptor interaction among KEGG pathways (Figure 6B). Conversely, GSEA analysis revealed that significant enrichment of chromosome organization, chromosome segregation, DNA-templated DNA replication, sister chromatid segregation in the low-risk group for BP terms in GO analysis (Figure 6C), as well as cell cycle, fanconl anemia pathway, fat digestion and absorption, and neuroactive ligand-receptor interaction among KEGG pathways (Fig. 6D). We also investigated the difference of gene function for MF and CC terms in GO enrichment analyses between the high-risk and low-risk groups.

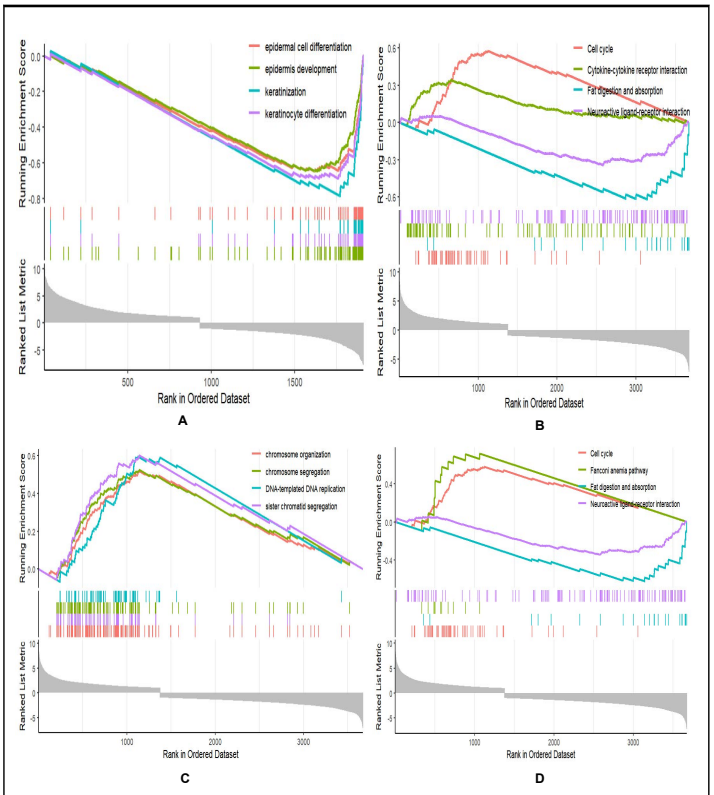

Figure 6: Functional enrichment analysis between high-risk and low-risk groups (A) GO-enriched BP terms in GSEA analysis of the high-risk group. (B) KEGG pathway analysis in GSEA analysis of the high-risk group. (C) GO-enriched BP terms in the GSEA analysis of the low-risk group. (D) KEGG pathway analysis in GSEA analysis of the low-risk group.

Relationship between the immune microenvironment and immunotherapy

Relevant studies have demonstrated that the involvement of lysosomes in regulating immune cell function. To further investigate this, we utilized function "subsea" in "GSVA" package to calculate the abundance of each immune cell in GC samples. The resulting box-and-line plot revealed significant differences in immune cells between high-risk and low-risk groups (Figure 7A). Specifically, CD4 T cells, B cells, memory B cells, and T cells were expressed at significantly higher levels in the low-risk group compared to the high-risk group, indicating functional differences between these two groups. Additionally, our analysis identified several differentially expressed immune checkpoint genes including *TNFSF4, CD276, NRP1, TNFRSF4, LAIR1, TNFRSF9*, and *CD28* that were all up-regulated in the high-risk group relative to their expression levels in the low-risk group (Figure 7B).

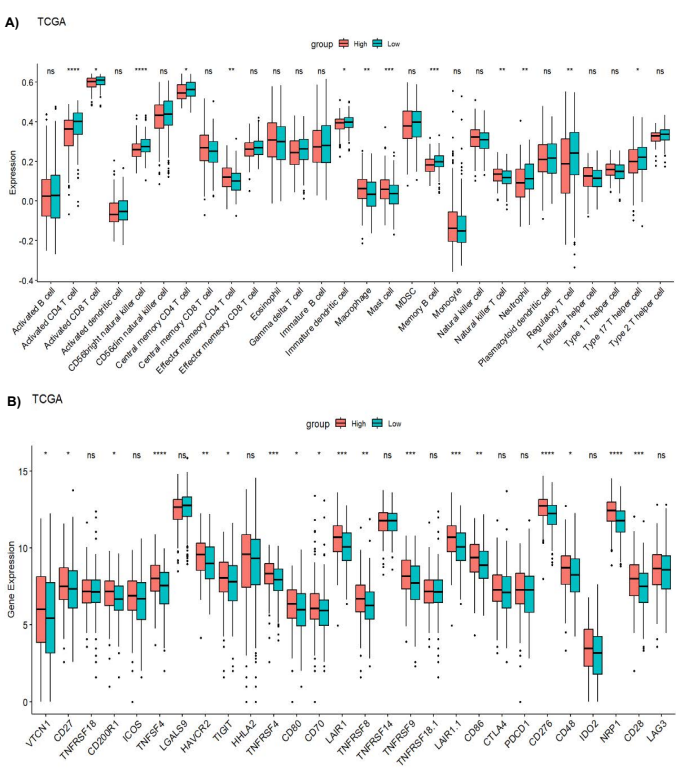

Figure 7: Correlation between risk score and immune cells (A) GSVA analysis of the differences in immune cells in the high-risk and low-risk groups in TCGA. (B) Differences in immune checkpoint genes between high-risk and low-risk groups. *p-value<0.05, **p-value<0.01, ***p-value<0.001, ****p-value<0.0001. ns: no significance.

TMB analyses in the high-risk and low-risk groups

Somatic mutation data were downloaded using the "TCGAbiolinks" package and visualized for TCGA-STAD patients (Figure 8A). Mistranslated mutations were the most predominant mutation classification, and Single Nucleotide Polymorphisms (SNPs) were the most predominant type of mutation, with a median mutation value of 62 for each sample. *MUC16, LRP1B, SYNE1, CSMD3*, and *FLG* are the top 10 mutated genes in both the high-risk and low-risk groups. Classification information for both groups was combined with mutation data to extract samples from each group. A waterfall map was used to display the top 15 genes with the highest frequency of mutations (Figures 8B and C). In the high-risk group, mutated genes were present in 154 out of 171 samples (90.06%). The most frequently mutated genes in this group included *TTN* (54%), *TP53* (47%), *MUC16* (35%), *LRP1B* (30%), *SYNE1* (27%), *ARID1A* (27%), *FAT4* (26%), and *CSMD3* (23%) (Figure 8B). In contrast, within the low-risk group, mutated genes were present in 148 out of 167 samples (88.62%). The highly-mutated genes within this cohort included *TTN* (50%), *TP53* (43%), *ARID1A* (28%), *MUC15* (26%), *LRP1B* (26%), *CSMD3* (23%), *SYNE1* (23%), and *SPTA1* (20%) (Figure 8C).

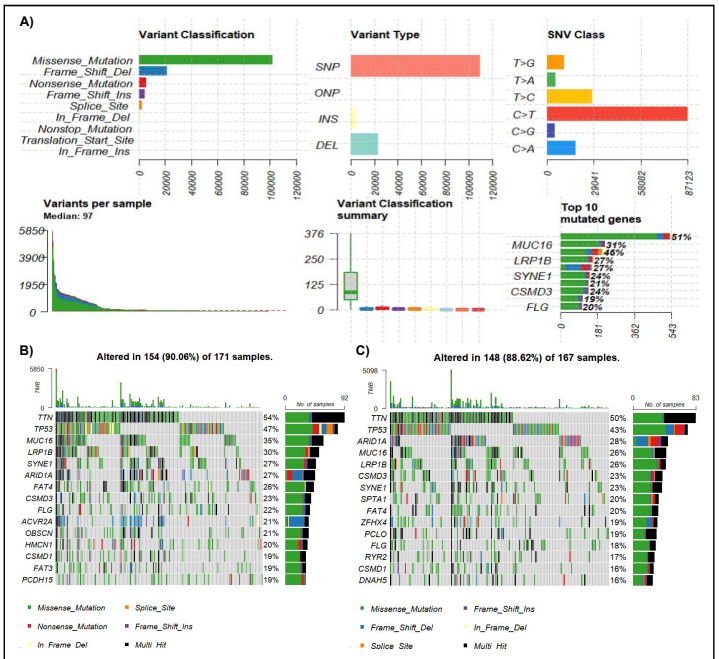

Figure 8: Mutation analysis in the high-risk and low-risk groups (A) General description of the mutation landscape in TCGA-STAD patients. (B) Mutated loci in the high-risk group. (C) Mutation loci in the low-risk group.

Discussion

GC is a rapidly emerging and highly malignant form of cancer in terms of human cancer incidence, but the current conventional clinical therapies for GC remain significantly limited [19]. Therefore, it is imperative to explore novel and effective molecular markers to enhance the clinical outcomes of GC treatment. Immunotherapy has emerged as a pivotal approach in improving therapeutic efficacy against lung cancer, with mounting evidence suggesting an association between alterations in lysosomal biological functions and immune cells as well as cancer development [20]. Nevertheless, there exists a dearth of systematic investigation into differential LRGs among patients with STAD at present. Hence, this study aims to investigate the prognostic significance of LRGs in GC patients.

We first screened 7 key LRGs-*TRIM29, EGF, GPC3, RETN, RNASE3, GRP*, and *PSAPL1*, for the construction of prognostic risk model by univariate Cox, LASSO regression, and multivariate stepwise Cox regression. There is accumulating evidence supporting the relevance of these identified LRGs to malignant tumor prognosis. Studies have demonstrated that the *TRIM29* gene acts as a crucial negative regulator of DNA viral and cytoplasmic DNA immune responses by targeting STING degradation, and there present an association between over-expression of the *TRIM29* gene and squamous cell carcinomas of the skin as well as ovarian cancer.

EGF, the earliest discovered growth factor, plays a pivotal role in cell growth, differentiation, and proliferation. Previous study has reported a positive correlation between EGF presence in GC and infiltration as well as lymph node metastasis. Moreover, the detection of EGF in human GC may indicate an elevated level of cancer malignancy.

Although the precise function of the *GPC3* gene remains elusive, an increasing body of evidence suggests that GPC3 serves as a

promising target molecule for early diagnosis of hepatocellular carcinoma. Moreover, a novel imaging strategy utilizing GPC3-targeted immune positron emission tomography has been developed to facilitate early diagnosis of hepatocellular carcinoma. Consequently, further investigation into the role of GPC3 in GC is warranted, with the expectation that it may unveil a new target molecule for improving early detection of this malignancy.

The *RETN* gene is capable of encoding resistin and adipokines in the human body, exhibiting risk associations between resistin and RETN with susceptibility to breast cancer and type 2 diabetes mellitus. Furthermore, it has been suggested that polymorphisms in RETN may contribute to an increased susceptibility to colon cancer disease.

RNASE3, a member of the RNASEA superfamily involved in host immunity, is expressed by leukocytes and possesses direct antimicrobial and immunomodulatory properties, prognostic models for idiopathic pulmonary fibrosis have been developed by researchers utilizing five types of immune cells, including RNASE3, resulting in improved outcomes.

GRP belongs to the belladonna peptide family of gastrin-releasing peptides, and it acts as an autocrine growth factor that stimulates the proliferation of various cancer cells and regulates numerous functions within the gastrointestinal and central nervous systems. Furthermore, immune responses triggered by novel chimeric proteins targeting GRP have demonstrated inhibition of mouse mammary tumor cells EMT-6. Additionally, PSAPL1 has been identified as a valuable biodiagnostic marker for GC.

The risk score model was evaluated using training and validation sets. Firstly, the samples were divided into two groups based on the median risk score threshold, and the low-risk group exhibited a significantly longer survival time compared to the high-risk group. Additionally, the validity of the risk score's predictive performance was confirmed by K-M curve and ROC curve analysis. Univariate and multivariate Cox analyses further demonstrated that risk score served as an independent prognostic factors for GC, and a nomogram model incorporating risk score, stage, and age confirmed high accuracy in predicting 1, 3, and 5-year survival time. Finally, we verified the rationality of our risk model from biological functions. Immunoassays were conducted to explore functional differences between the high-risk and low-risk groups. The GO enrichment analyses of the BP term revealed differential results in the first four biological processes between the high-risk and low-risk groups, while only one of the first four pathways showed a significant difference in the KEGG pathway analyses between these two risk groups.

With advancements in science and medicine, immunotherapy has emerged as a groundbreaking approach to cancer treatment, with Immune Checkpoint Inhibitors (ICIs) playing a pivotal role in altering the treatment and prognosis of gastric cancer. Hence, we also investigated the association between the risk score model and immune checkpoints and revealed significant differential expression of immune checkpoint genes such as *TNFSF4, CD276, NRP1, TNFRSF4, LAIR1, TNFRSF9, LAIR1*, and *CD28* between the high-risk and low-high groups. Infiltration of immune cells into tumor tissues along with their modulation of cytokine signaling significantly influences the biological function of cancer cells. Our findings demonstrate a significantly higher expression level of CD4 T cells, B cells, memory B cells, and T cells in low-risk groups compared to high-risk groups. These results are consistent with previous research.

We observed that TTN and TP53 exhibited the highest mutation frequencies in both high-risk and low-risk groups, which are closely associated with immunotherapy. TTN/TP53 co-mutations may be a potent predictor of OS and chemotherapeutic response in patients with lung cancer. High mutation of TTN is positively correlated with the survival rate of GC patients, and the *TTN* gene is important in improving the level of immunity. TP53 oncogene mutations are common in 50% of human cancers, and TP53 act as a transcription factor capable of directly regulating the expression of approximately 500 genes.

However, our study still has certain limitations. Firstly, we employ the traditional statistical model to identify the LRGs. Although our model demonstrates excellent performance in prognostic prediction for GC patients, leveraging machine learning or deep learning algorithms may yield more accurate predictive outcomes. Secondly, further experimental exploration is required to elucidate the role of these 7 LRGs we identified in GC pathologic function. Addressing these aforementioned shortcomings will be the primary focus of our future work.

Conclusion

In this study, we have identified 7 key prognostic LRGs associated with GC and developed a risk score model based on these genes. Our model has demonstrated high accuracy in predicting the OS of patients with GC. Furthermore, significant differences were observed in the immune microenvironment, immunotherapy response, as well as TTN and TP53 mutations between the high and low-risk groups. These findings provide valuable insights for future studies aiming to improve the prognosis of GC patients through lysosomal-related mechanisms.

Ethics Approval and Consent to Participate

All experiments involving human tissues complied with the principles of the declaration of Helsinki.

Consent for Publication

Not applicable.

Availability of Data and Materials

The datasets used in this study can be found in the GEO database (https://www.ncbi.nlm.nih.gov/geo/), TCGA database (https://portal.gdc.cancer.gov/) and AmiGO2 database (http://amigo.geneontology.org/amigo). Further inquiries can be directed to the corresponding author.

Funding

The present study was funded by the provincial natural science foundation of Anhui (grant no. 1808085MG220), the research projects of Anhui teaching (grant no. 2020jyxm0238); the national innovation and entrepreneurship training project (grant no. 202110360094, 202210360086); and the research preparation project of Anhui Province (grant no. 2022AH050328).

Competing Interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors' Contributions

Zihao Wei conceptualized and designed the study, Zhengrui Wang and Zi Chen created the graphs and charts, and Dongmei Luo guided the research ideas and revisions. All members read and approved the manuscript.

Acknowledgement

The authors acknowledge the TCGA database, the GEO database, the AmiGO2 database, and other groups for providing invaluable datasets for statistical analyses.

References

1. Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F (2020) Gastric cancer. Lancet 396: 635-648.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, et al. (2021) Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 71: 209-249.
3. Zhao Q, Cao L, Guan L, Bie L, Wang S, et al. (2019) Immunotherapy for gastric cancer: dilemmas and prospect. Brief Funct Genomics 18: 107-112.
4. Hu HM, Tsai HJ, Ku HY, Lo SS, Shan YS, et al. (2021) Survival outcomes of management in metastatic gastric adenocarcinoma patients. Sci Rep 11: 23142.
5. Catalano V, Labianca R, Beretta GD, Gatta G, de Braud F, et al. (2005) Gastric cancer. Crit Rev Oncol Hematol 54: 209-241.
6. Wang F, Gomez-Sintes R, Boya P (2018) Lysosomal membrane permeabilization and cell death. Traffic 19: 918-931.
7. Ballabio A (2016) The awesome lysosome. EMBO Mol Med 8: 73–76.
8. Kallunki T, Olsen OD, Jaattela M (2013) Cancer-associated lysosomal change7s: friends or foes?. Oncogene 32: 1995–2004.
9. Saftig P, Klumperman J (2009) Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. Nat Rev Mol Cell Biol 10: 623–635.
10. Yang C, Wang X (2021) Lysosome biogenesis: Regulation and functions. J Cell Biol 220: e202102001.
11. Colacurcio DJ, Nixon RA (2016) Disorders of lysosomal acidification emerging role of v-ATPase in aging and neurodegenerative disease. Ageing Res Rev 32: 75-88.
12. Whyte LS, Lau AA, Hemsley KM, Hopwood JJ, Sargeant TJ (2017) Endo-lysosomal and autophagic dysfunction: a driving factor in Alzheimer's disease?. J Neurochem 140: 703-717.
13. Zhang Z, Yue P, Lu T, Wang Y, Wei Y, et al. (2021) Role of lysosomes in physiological activities, diseases, and therapy. J Hematol Oncol 14: 79-118.
14. Serrano-Puebla A, Boya P (2018) Lysosomal membrane permeabilization as a cell death mechanism in cancer cells. Biochem Soc Trans 46: 207-215.
15. Cristescu R, Lee J, Nebozhyn M, Kim KM, Ting JC, et al. (2015) Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. Nat Med 21: 449-456.
16. Oh SC, Sohn BH, Cheong JH, Kim SB, Lee JE, et al. (2018) Clinical and genomic landscape of gastric cancer with a mesenchymal phenotype. Nat Commun 9: 1-15.
17. Masunaga Y, Ohkubo Y, Nishimura G, Ueno T, Fujisawa Y, et al. (2022) ACAN biallelic variants in a girl with severe idiopathic short stature. J Hum Genet 67: 481-486.
18. Rahimi Z (2012) ACE insertion/deletion (I/D) polymorphism and diabetic nephropathy. J Nephropathol 1: 143-151.
19. Li K, Zhang A, Li X, Zhang H, Zhao L (2021) Advances in clinical immunotherapy for gastric cancer. Biochim Biophys Acta Rev Cancer 1876: 188615.
20. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144: 646-674.