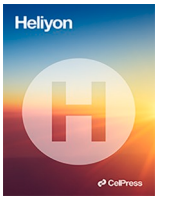
Heliyon 10 (2024) e33795

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Identification of a novel apoptosis-related genes signature to improve gastric cancer prognosis prediction

Xiaopeng Li [a,b], Xiaolei Yin [a], Lili Mi [a], Ning Li [a], Shumei Li [b], Fei Yin [a,*]

[a] Department of Gastroenterology, The Fourth Hospital of Hebei Medical University, Shijiazhuang, 050035, Hebei, China
[b] Medical Record Room, The Fourth Hospital of Hebei Medical University, Shijiazhuang, 050035, Hebei, China

ARTICLE INFO

Keywords:
Apoptosis genes
Gastric cancer
Prognostic model
Molecular characteristics
Chemotherapies sensitivity

ABSTRACT

Dysregulation of apoptosis occurs in different types of malignant tumors and is likely to influence the tumor evolution, as well as clinical prognosis. However, the limited number of studies investigating the predictive power of apoptosis-related genes (ARGs) in gastric cancer indicates a gap in the current research. 174 ARGs who differentially expressed were screened using public databases, including the Gene Expression Omnibus and the Molecular Signatures Database. Univariate and LASSO regression analyses were rigorous approaches to recognize the 12 optimal genes (*CTHRC1, PDGFRL, VCAN, GJA1, LOX, UPP1, ANGPT2, CRIM1, HIF1A, APOD*, *RNase1*, and *ID1*) that make up the prognostic risk model. Molecular mutations, related signaling pathways, and immune system characteristics in different subgroups defined by the risk model were analyzed using different R packages. Moreover, based on the database of Genomics of Drug Sensitivity in Cancer, chemotherapy sensitivity was predicted among the risk subgroups. As a result, there were differences in mutation profiles, signaling pathways, and infiltrated immune cells between patients in various risk groups. Moreover, the low-risk group displayed greater sensitivity to chemotherapy than the high-risk group. Risk model provided a better prognostic value than the T, N, and M stages, according to the receiver operating characteristic curve. Finally, in a nomogram, the risk model and clinical factors were combined to visualize the survival rates of patients with GC. In response to the differential expression of apoptosis-related genes, a novel model for predicting the prognosis of GC patients was developed. This model may be highly valuable for guiding doctors to deliver treatment plans tailored to the need of patients with GC.

1. Introduction

Gastric cancer (GC) is one of the most prevalent gastrointestinal malignancies in the world, ranking fifth in cancer-related mortality [1]. Radical surgery remains the cornerstone of treatment for potentially resectable GC [2]. Although substantial advancements have been made in anticancer therapies, including chemotherapy and immunotherapy, patients with advanced GC still have unsatisfactory prognoses [3]. Consequently, it is imperative to understand the mechanisms underlying disease progression and identify novel and effective prognostic biomarkers. By doing so, the survival rates and overall well-being of GC patients can be improved.

Apoptosis is a dynamic and evolutionarily conserved phenomenon in the development of various organisms. It maintains the

* Corresponding author.
E-mail address: xiaoyin2627@126.com (F. Yin).

https://doi.org/10.1016/j.heliyon.2024.e33795
Received 5 September 2023; Received in revised form 24 June 2024; Accepted 26 June 2024
Available online 27 June 2024

balance and harmony of the intracellular environment, ensuring the proper functioning and homeostasis of cells. Intrinsic or acquired resistance to apoptosis is a major hallmark of human cancers [4]. Escaping apoptosis leads to uncontrolled proliferation of cancer cells, resulting in tumorigenesis and therapeutic resistance [5]. Tumors employ various mechanisms to evade apoptosis, which allows them to survive and persist. Cancer cells often exhibit alterations in the expression or activity of key apoptotic regulators such as Bcl-2 family proteins, p53, and caspases. Anti-apoptotic molecules, such as Bcl-2, are frequently upregulated in cancer cells that are resistant to drug treatments [6]. The expression of BAK, a member of the Bcl-2 family, can predict chemotherapeutic responses to docetaxel in patients with GC and is an effective prognostic biomarker [7]. Recently, a growing number of genes are being recognized as biomarkers for GC, which contribute to improving the outcomes of patients with this disease. For instance, the discovery of a gene signature that can speculate the prognosis and effectiveness of immunotherapy in GC represents a substantial advancement in personalized medicine and development of immunotherapeutic strategies [8]. Additional research has indicated that genes associated with angiogenesis can serve as prognostic indicators for patients with GC [9]. However, the molecular basis of apoptosis and the predictive implications of apoptosis-related genes (ARGs) in GC remain poorly understood. Therefore, such research could have profound implications for personalized treatment, particularly for high-risk patients.

Our study involved an extensive bioinformatics examination using the Gene Expression Omnibus (GEO) database to assess the prognostic significance of ARGs in GC patients. The purpose of this research was to formulate a prognosis risk model on account of 12 apoptosis-associated genes and to investigate the relationship between gene expression profiles and GC prognosis. Through our analysis, we generated a novel prognostic risk prediction model and examined its association with various factors in different risk subgroups, including tumor mutation burden, gene set enrichment analysis, immune cell invasion, and chemosensitivity. Additionally, a nomogram was created to evaluate the clinical prognosis of these GC patients. This comprehensive evaluation allowed us to gain insight into the potential underlying mechanisms and clinical implications of our prognostic model. In conclusion, a novel prognostic risk prediction model was established that provides more advantageous information that indirectly improves the prognosis of these patients.

2. Materials and methods

2.1. Datasets and preprocessing

The clinical data and gene expression profiles were downloaded from the GEO and The Cancer Genome Atlas (TCGA) databases. Four GEO datasets (GSE118916, GSE54129, GSE79973, and GSE81948) were selected to identify the differentially expressed genes in GC. Molecular signatures database (MSigDB) was used to screen genes associated with apoptosis. Data from 315 GC patients at TCGA, comprising RNA sequences and clinical data, served as an internal training dataset. We incorporated an external testing dataset, GSE62254, into our study. This dataset consisted of information from 300 patients with GC who had complete follow-up data. Supplementary Table S1 presents the clinical data for the training and testing datasets used in this research. Principal component analysis (PCA), as well as hierarchical cluster analysis were employed to discern different group information based on the distance matrix using the R statistical language (v3.6.1). For background adjustment, the raw files of the five GEO datasets were downloaded and processed using a robust multichip average (RMA) algorithm after quantile normalization and log transformation. Moreover, TCGA gene expression profiles were transformed to base-2 logarithms for further analysis.

2.2. Identification of apoptosis-related hub genes

We utilized the R programming language to process the downloaded matrix files. The probe names in the matrix files have been converted to their international standard names utilizing the annotation package in R. We screened for differentially expressed genes (DEGs) between GC and normal gastric samples using four microarray datasets obtained from the GEO database. For this analysis, the limma and robust rank aggregation (RRA) packages in R were employed. The significance levels were set at $P < 0.05$ and $|\log_2 FC| > 0.5$ [10]. ARGs were screened using MSigDB as described previously. We then compared the list of ARGs with the identified DEGs, resulting in a final set of 174 apoptosis-related DEGs. We performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses to investigate the probable molecular mechanisms linked to these apoptosis-related DEGs using the ClusterProfiler package.

2.3. Construction and validation of the prognostic model

In TCGA training dataset, a univariate Cox analysis was performed to assess the connection among 174 candidate genes and the overall survival (OS) of patients with GC. LASSO regression allows the identification of a subset of genes with the highest predictive value for patient prognosis [11]. Optimal genes and their coefficients were screened by applying the best penalty parameter from the smallest cross-validation of 10 folds. We obtained 12 genes and calculated the coefficients for each prognostic gene. The risk score is calculated by multiplying the expression of each gene by its corresponding coefficient and summing the results. Patients were divided into low- and high-risk groups based on their median risk score. The score's predictive ability was assessed with Kaplan-Meier survival analysis.

2.4. Molecular characteristics in different risk subgroups

Using Maftools software, we conducted an in-depth analysis of somatic variants in different risk subgroups with information retrieved from TCGA database [12]. Microarray data were analyzed through Gene Set Enrichment Analysis (GSEA). ClusterProfiler in R was utilized to investigate the crucial signal pathways enriched in both subgroups using a significance threshold of $P < 0.05$, and a false discovery rate <0.25.

2.5. Immune characteristics in different risk subgroups

The immune cell subtypes were assessed utilizing CIBERSORT4, which utilizes linear support vector regression. This method allows for the evaluation of the percentage of immune cells that have infiltrated into different risk categories due to the gene expression profile of patients with GC. The distribution of 22 immune cell types infiltrating in each risk subgroup was estimated [13].

2.6. Chemotherapies sensitivity in different risk subgroups

We investigated the chemotherapeutic response of each sample in the TCGA dataset using the Genomics of Drug Sensitivity in Cancer database and the pRRophetic R package. The package used ridge regression analysis to determine the half-maximal inhibitory concentration (IC50) of certain chemotherapeutics drugs. The prediction accuracy of the regressions was assessed using 10-fold cross-validation.

2.7. Construction of nomogram

According to the risk scores and clinicopathological features, a nomogram was created to estimate the risk stratification of patients with GC. The nomogram aimed to predict the 1-, 3-, and 5-year OS rates in these patients. The nomogram efficiency was estimated using a calibration curve.

2.8. Statistical analysis

R software was utilized to perform the statistical analyses and graphical visual representations in this study. To determine the statistical significance of the variables, Student's t-test and the chi-squared test were employed to analyze the continuous and categorical variables separately. Kaplan–Meier analysis estimated the OS. The independent prognostic significance of the risk score and clinical characteristics was evaluated by both univariate and multivariate Cox regression analyses. Statistical significance was defined

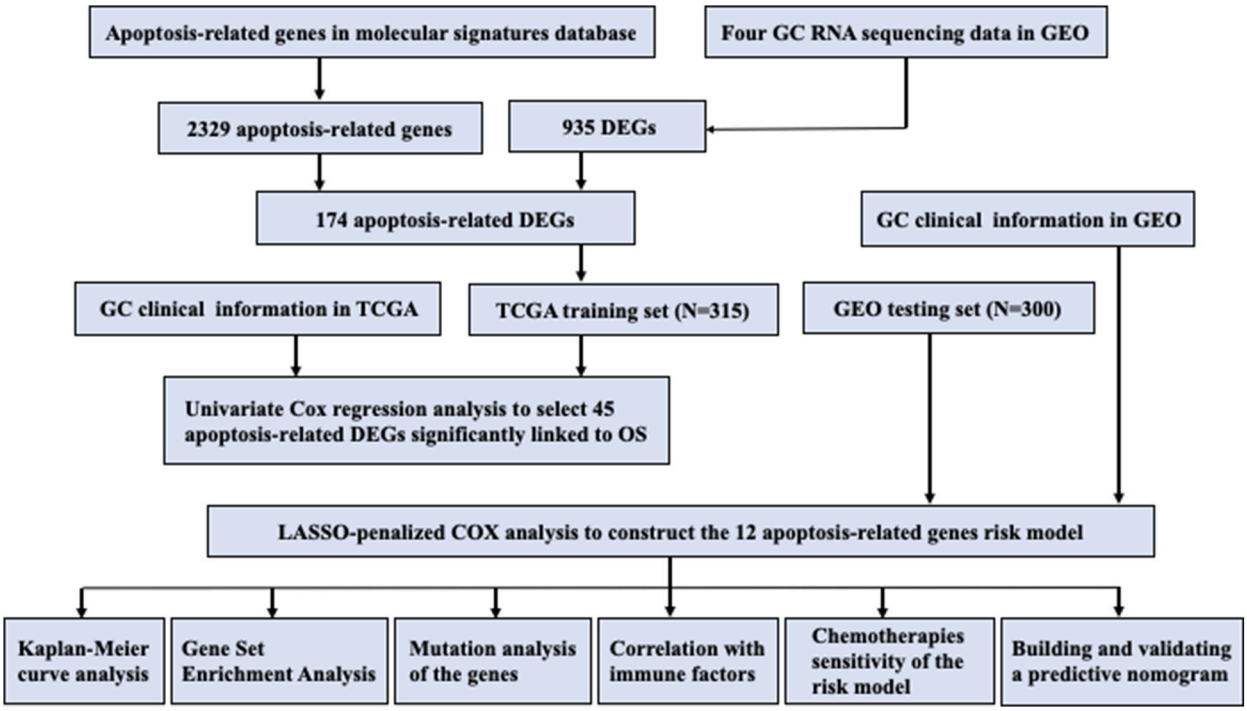
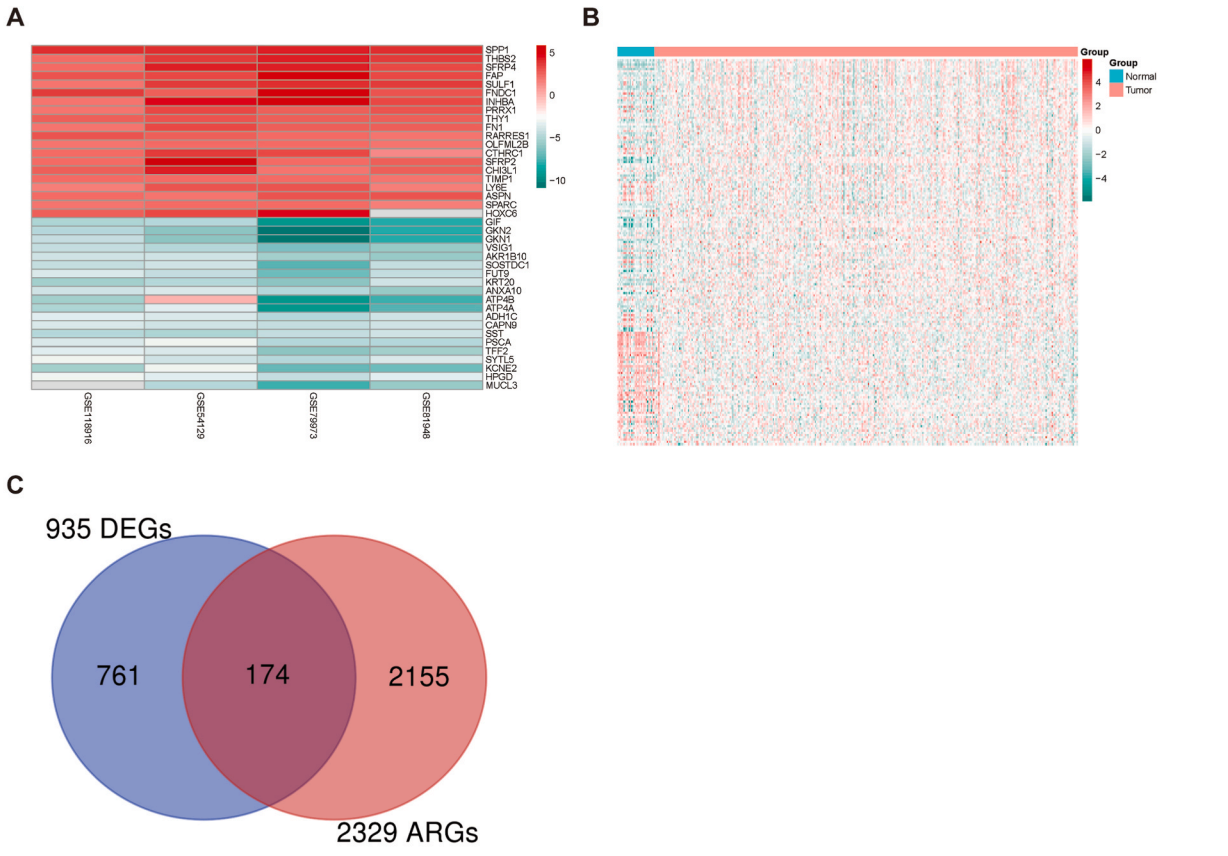


Fig. 1. Flowchart of this present study.



(caption on next page)

Fig. 2. Identification and functional enrichment analyses of apoptosis-related differentially expressed genes. (A) The top 20 up and downregulated differentially expressed genes in these four GEO datasets. (B) Heatmap of apoptosis-related genes in molecular signatures database. (C) The intersection of the 935 DEGs and 2329 ARGs by Venn diagram. (D) Gene Ontology (GO) enrichment analysis of the apoptosis-related DEGs. (E) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of the apoptosis-related DEGs.

as $P < 0.05$.

3. Results

The detailed flowchart for this research is shown in Fig. 1.

3.1. Identification of apoptosis related hub genes

Datasets involved in GC expression profiles (GSE118916, GSE79973, GSE81948, and GSE54129) were standardized. The results of the PCA (Supplementary Fig. S1A) indicated a distinct separation or discernible dimensions between the normal and tumor groups, suggesting notable distinctions in the gene expression profiles among these two groups. We screened 2266, 1650, 1507 and 2802 upregulated genes in GSE118916, GSE79973, GSE81948 and GSE54129, while the number of the downregulated genes in these datasets respectively are 2014, 1553, 1339 and 3094. For each dataset, our cut-off criteria for DEGs screening were corrected $P < 0.05$ and $|\log_2 FC| \geq 0.5$ (Supplementary Fig. S1B).

The obtained DEGs were analyzed utilizing the RRA method with a significance cut-off of $P < 0.05$, after correction, which assumed

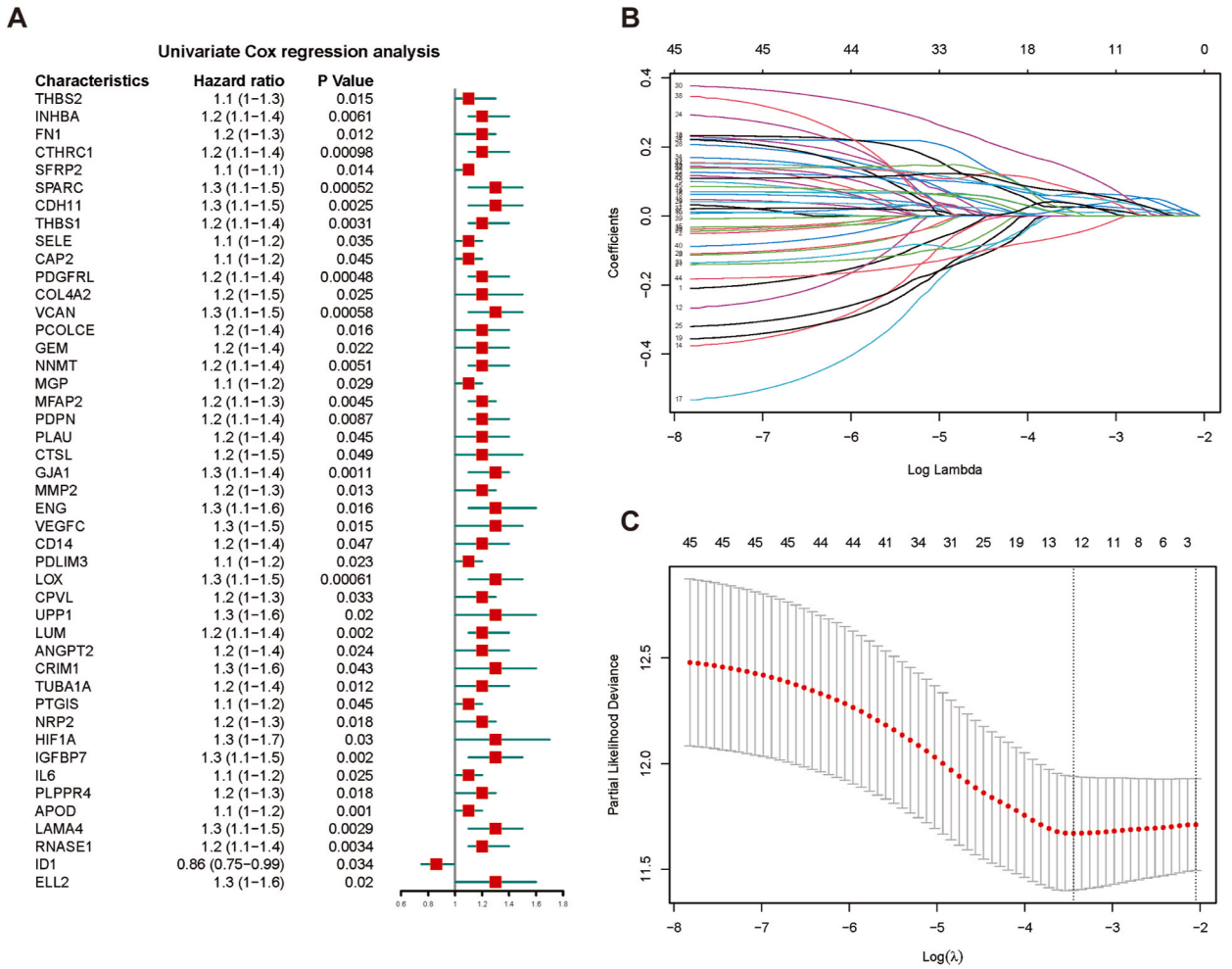


Fig. 3. Construction and risk score analysis of a prognosis model. (A) Univariate Cox regression analysis revealed that 45 apoptosis-related DEGs significantly correlated with clinical prognosis. (B) LASSO coefficient curves for 45 prognostic apoptotic genes in the training dataset. (C) The coefficient profile plot was generated with log (λ).

that genes in each experiment were arranged randomly. Genes with higher ranks were more likely to be DEGs, indicating a smaller p-value after correction. Using this approach, we identified 935 DEGs consisting of 412 upregulated and 523 downregulated genes. Fig. 2A illustrates the top 20 most upregulated and downregulated genes among DEGs. Additionally, Fig. 2B depicts the extraction of 2329 ARGs, whereas Fig. 2C shows the overlap between the 935 DEGs and 2329 apoptotic genes, resulting in the identification of 174 apoptosis-related DEGs. For gaining insight into the biological processes and pathways enriched in these 174 apoptosis-related DEGs, GO terms and KEGG pathway were analyzed. Analyses of enrichment results helped us to better understand the functional implications and potential molecular mechanisms underlying apoptosis in GC (Fig. 2D and E).

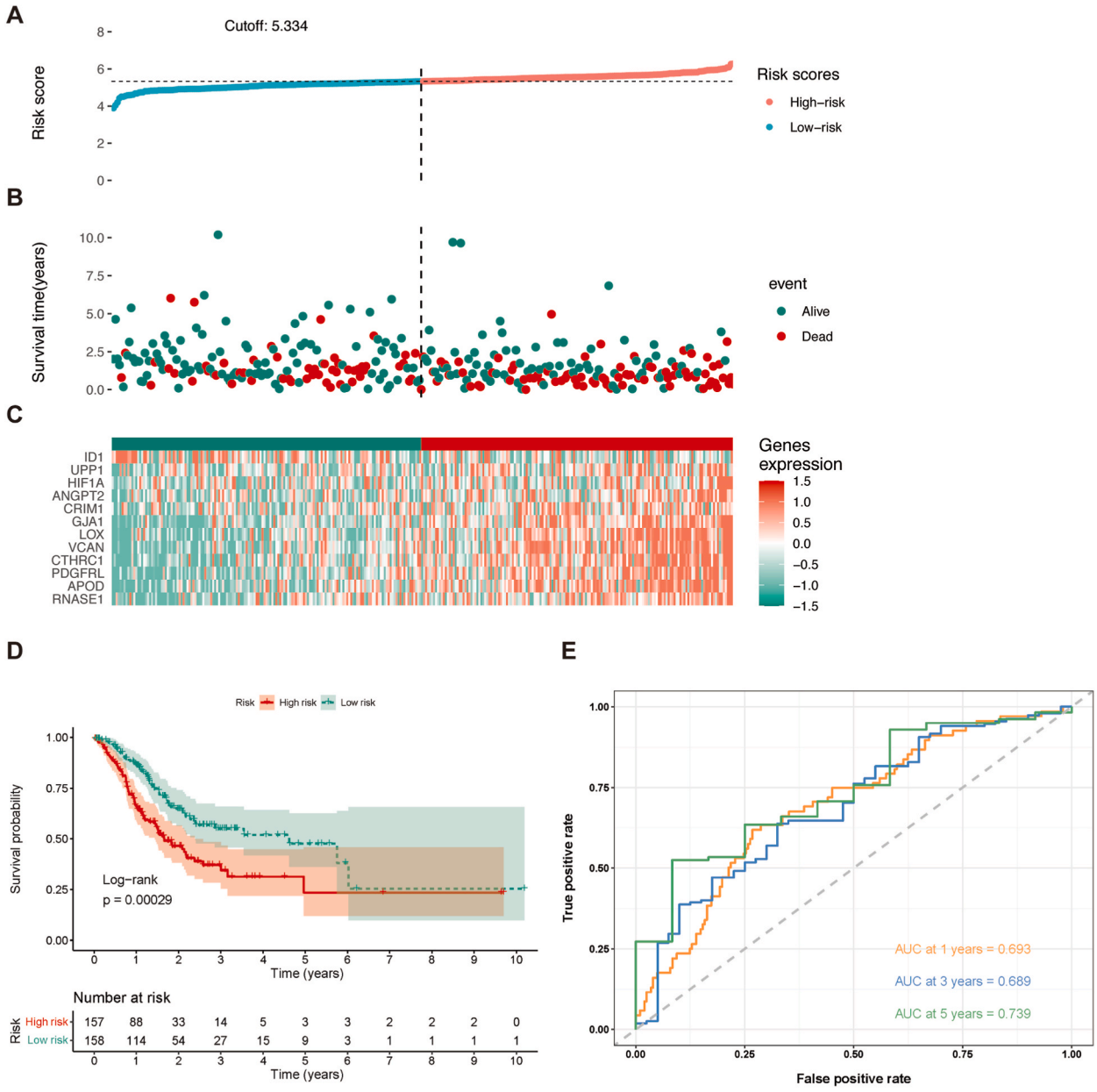
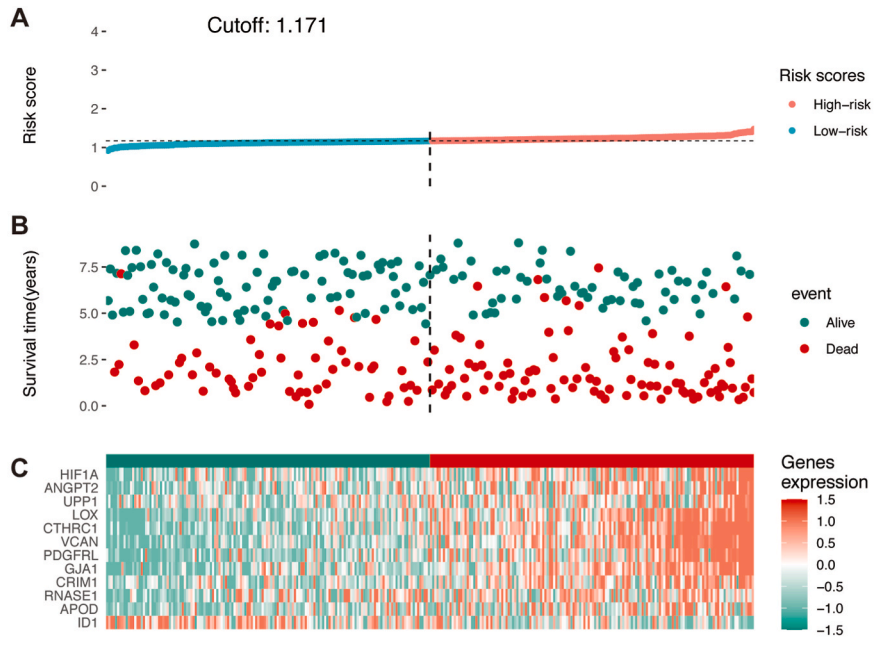
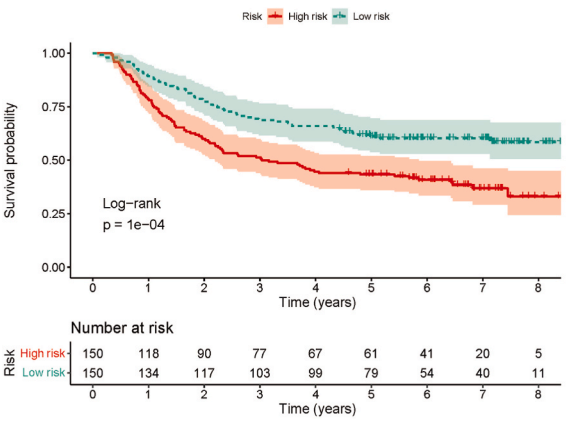
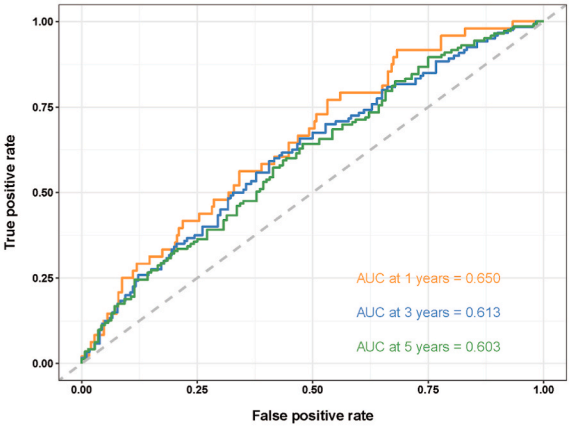
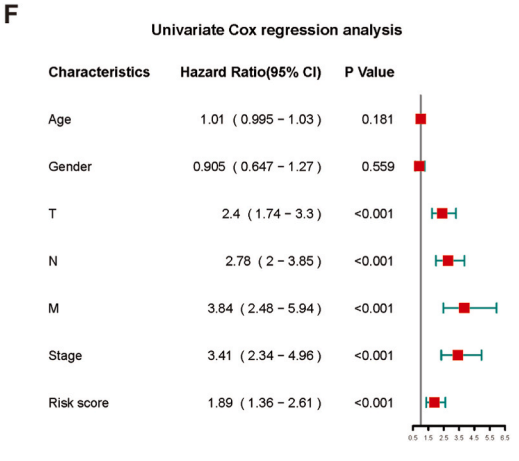


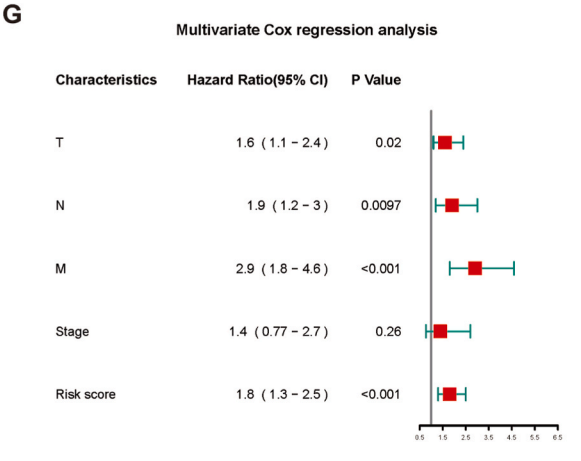
Fig. 4. Prognostic value of the risk model in the TCGA training dataset. (A) The median risk score allocated GC patients in TCGA into two groups. From left to right, the risk scores were ascending and each dot represents an individual. (B) Patients with different survival times and statuses were arranged with the increasing risk score from left to right. (C) The heatmap showed expression profile of the 12 ARGs. The expression level from high to low was manifested with the colors from red to green. (D) Kaplan–Meier analysis showed different prognosis between the two groups. (E) Validation of ROC curves for the risk score in the training dataset from TCGA.











(caption on next page)

Fig. 5. External validation of the apoptosis-related gene prognostic risk model. (A) Patient with different risks were separated into two groups via the same median risk score in the training dataset. (B) Patients in the testing dataset with different survival times and statuses were arranged by the increasing risk score from left to right. (C) The heatmap of the 12 ARGs in the testing dataset showed the same expression pattern. (D) Kaplan–Meier analysis showed different prognoses between the two groups as well. (E) The ROC curves of the models for survival rate in the testing dataset. (F) Forest plot of the univariate Cox regression analysis in the testing dataset. (G) Forest plot of multivariate Cox regression analysis evaluating the independent prognostic value of risk score.

3.2. Construction and risk score analysis of a prognosis model

Univariate analysis was performed on the 174 apoptosis-related DEGs using the TCGA training dataset consisting of 315 samples, and 45 genes showed significant associations with OS ($P < 0.05$) (Fig. 3A). We obtained correlation coefficients for each of the 45 genes using LASSO Cox regression, which allowed us to assess the strength and direction of their association with OS. This analysis aided in selecting the most informative genes that contributed significantly to the prognostic risk model. Finally, 12 candidate genes (CTHRC1, PDGFRL, VCAN, GJA1, LOX, UPP1, ANGPT2, CRIM1, HIF1A, APOD, RNASE1, and ID1) were screened from the training array (Fig. 3B and C). The formula of risk score = the expression of CTHRC1 · (0.0349) + PDGFRL · (0.0110) + VCAN · (0.0259) + GJA1 · (0.0644) + LOX · (0.0262) + UPP1 · (0.1300) + ANGPT2 · (0.0736) + CRIM1 · (0.0096) + HIF1A · (0.0348) + APOD · (0.0545) + RNASE1 · (0.0690) + ID1 · (−0.0490).

TCGA training cohort patients were split into low- and high-risk groups on the basis of the median apoptosis risk score. This categorization allowed the stratification of patients according to their predicted risk of adverse outcomes (Fig. 4A). The survival rate of each patient is illustrated in Fig. 4B. To gain a more particular knowledge of the expression of 12 apoptosis-related DEGs used in the risk model, Fig. 4C shows the individual expression profiles of these genes for each patient. This graphical representation highlights the variability between the two groups in gene expression patterns, and the different outcomes were assessed using Kaplan–Meier curves. Compared to the high-risk group, the low-risk group survived longer (Fig. 4D; $P < 0.001$). Moreover, a receiver operating characteristic analysis (ROC) assessed the accuracy of prognostic risk scores. The AUC values were 0.693, 0.689, and 0.739 at 1, 3, and 5 years survival time, respectively, indicating that it accurately predicted patient survival at these specific time points (Fig. 4E). Overall, the risk model demonstrated high accuracy and stability for the assessment of GC patients' prognoses.

3.3. External validation of the risk model

The model's prognostic value was estimated using the external testing dataset, GSE62254 (n = 300), obtained from the GEO database. In the testing dataset, we employed a uniform formula similar to that used in the training database to calculate the risk scores for each individual. Risk groups were also assigned to patients in the validation dataset, as shown in Fig. 5A. We plotted the patient status, risk scores and survival times in the testing dataset using a scattergram (Fig. 5B). This visualization allowed a clear understanding of the relevance of the risk levels to survival outcomes in the testing data. Testing and training datasets showed similar expression profiles for the 12 apoptosis-related DEGs in the heatmap, indicating consistency in the biological relevance of these genes across different datasets (Fig. 5C). As shown in Fig. 5D, testing data showed differences in survival times between risk groups, which matched the training data results ($P < 0.001$). Furthermore, the accuracy of the predictive value was evaluated using ROC curves. There was a 0.650, 0.613, and 0.603 AUC for 1-, 3-, and 5-year OS, respectively, in the testing dataset, indicating a reasonable predictive capability of the model (Fig. 5E). Univariate regression analysis indicated that the risk score, as well as the T, N, M, and clinical stages, were related to the survival time of patients (Fig. 5F; all $P < 0.001$). Finally, multivariate regression analysis confirmed the risk score as a substantial independent prognostic index, even after considering other clinical factors ($P < 0.001$; Fig. 5G).

3.4. Molecular characteristics in different risk subgroups

To gain a deeper understanding of subgroup differences, we analyzed the gene mutation profiles of the two different risk groups. Results demonstrated that mutation rates were higher in the low-risk group compared to the high-risk group. Specifically, a considerably higher mutation rate for TTN was observed in the low-risk group. We conducted additional research into the various types of mutation in these subgroups and found that missense was the most commonly observed mutation type, then followed by nonsense and frameshift deletions. These different mutation types can lead to alterations in protein structure and function, potentially affecting tumor development and progression. As shown in Fig. 6A and B, both risk groups shared 20 genes exhibiting the most elevated mutation frequencies. These figures provide a visual representation of the mutation rates and highlight the genes with frequent mutations in each subgroup. Among these genes, TTN, TP53, CSMD3, LRP1B, MUC16, SYNE1, ARID1A, and FLG consistently showed mutation rates exceeding 20 %.

Next, GSEA analysis was performed to gain insights into the possible molecular mechanisms underlying the progression of tumor. Mitogen-activated protein kinase (MAPK) pathways and cytokine-cytokine receptor interactions were extensively enriched in the high-risk group (Fig. 6C). The enrichment of these pathways suggests their potential role in promoting tumor progression. In contrast, in the low-risk group, pathways related to DNA replication and homologous recombination were enriched. These pathways suggest a potential association with more controlled and stable cell growth, potentially contributing to a better prognosis (Fig. 6D).

The tumor microenvironment (TME) is primarily made up of immune cells, tumor-associated fibroblasts, multiple inflammatory factors, extracellular matrix and growth factors. Understanding the TME and immune cell infiltration patterns can provide valuable

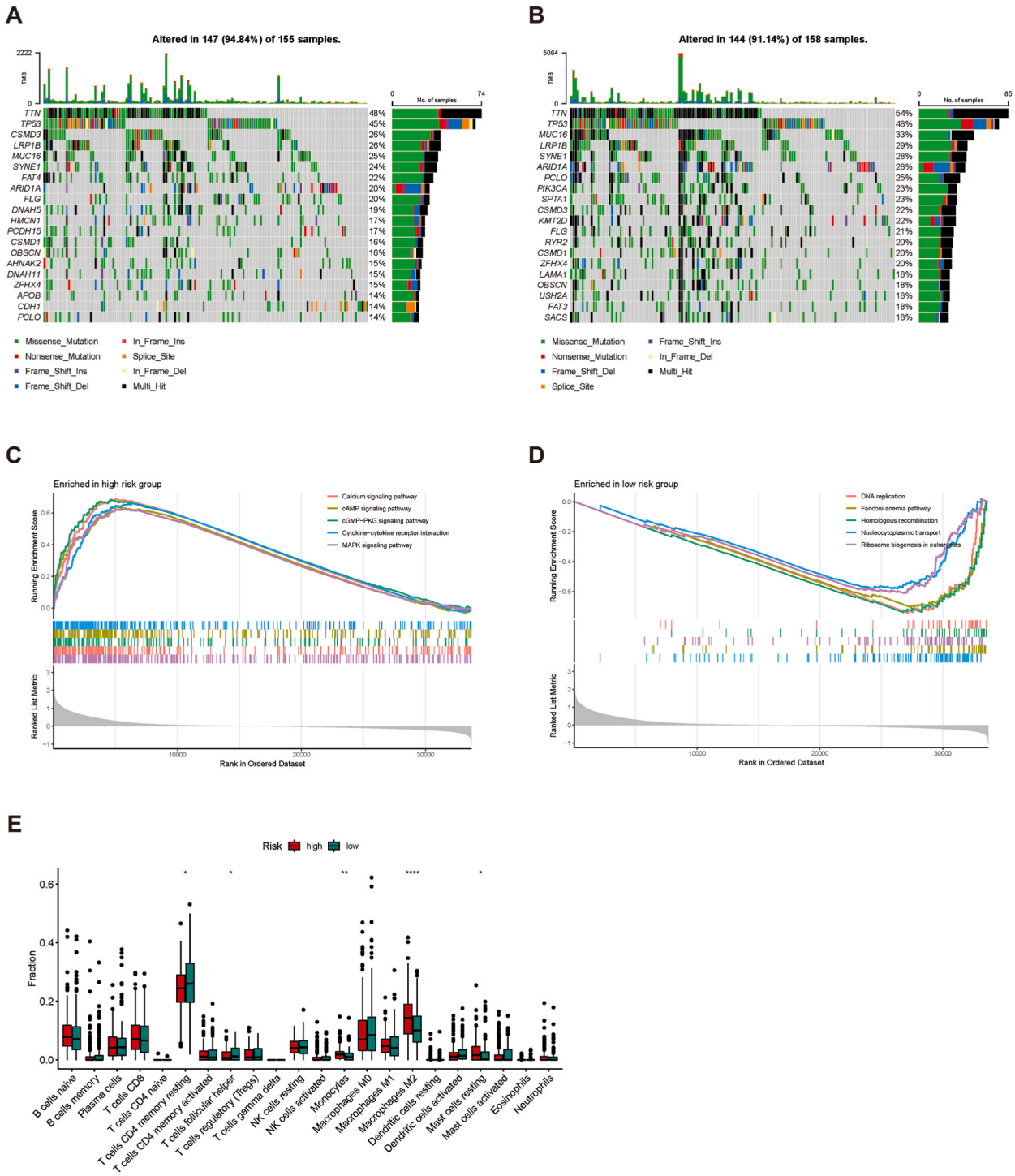


Fig. 6. Molecular characteristics in different risk subgroups. (A) Waterfall plot displays significantly mutated genes in high-risk subgroup. (B) Waterfall plot displays significantly mutated genes in low-risk subgroup. (C) Gene sets enriched in high-risk subgroup. (D) Gene sets enriched in low-risk subgroup. (E) The proportions of the 22 immune cells in different risk subgroups.

information regarding the clinical response to cancer treatment and patient survival. The CIBERSORT algorithm findings described that there was a significant increase of the distribution of monocytes, resting mast cells, and M2 macrophages in the high-risk subgroup. In contrast, the low-risk subgroup exhibited a lower percentage of resting CD4 + memory cells as well as follicular helper T cells than the high-risk subgroup (Fig. 6E). As we know, a subset of T cells referred to as $CD4^+$ memory resting cells plays a critical part in

long-term immunity and antigen reactivity. B cells can produce antibodies and exert adaptive immunity with the help of follicular helper T cells.

3.5. Chemotherapies sensitivity prediction in different risk subgroups

Chemotherapy remains the primary therapeutic schedule for GC. To assess the sensitivity of patients to chemotherapy in relation to the apoptosis risk score, we selected six commonly used drugs: oxaliplatin, paclitaxel, irinotecan, cisplatin, docetaxel, and fluorouracil. It was suggested that apoptosis risk score could substantially influence the sensitivity of patients to specific chemotherapy drugs, namely oxaliplatin, paclitaxel, docetaxel and fluorouracil. We observed lower IC50 values in the low-risk group, indicating that this segment of the population has a better response to these chemotherapeutic drugs (Fig. 7).

3.6. Construction of the prognostic nomogram

First, we assessed the prognostic precision of the risk score by using additional clinical and pathological parameters, like age, T, N, M and clinical stages. It was evident from the ROC curve that the risk score performed better in forecasting the outcomes of patients

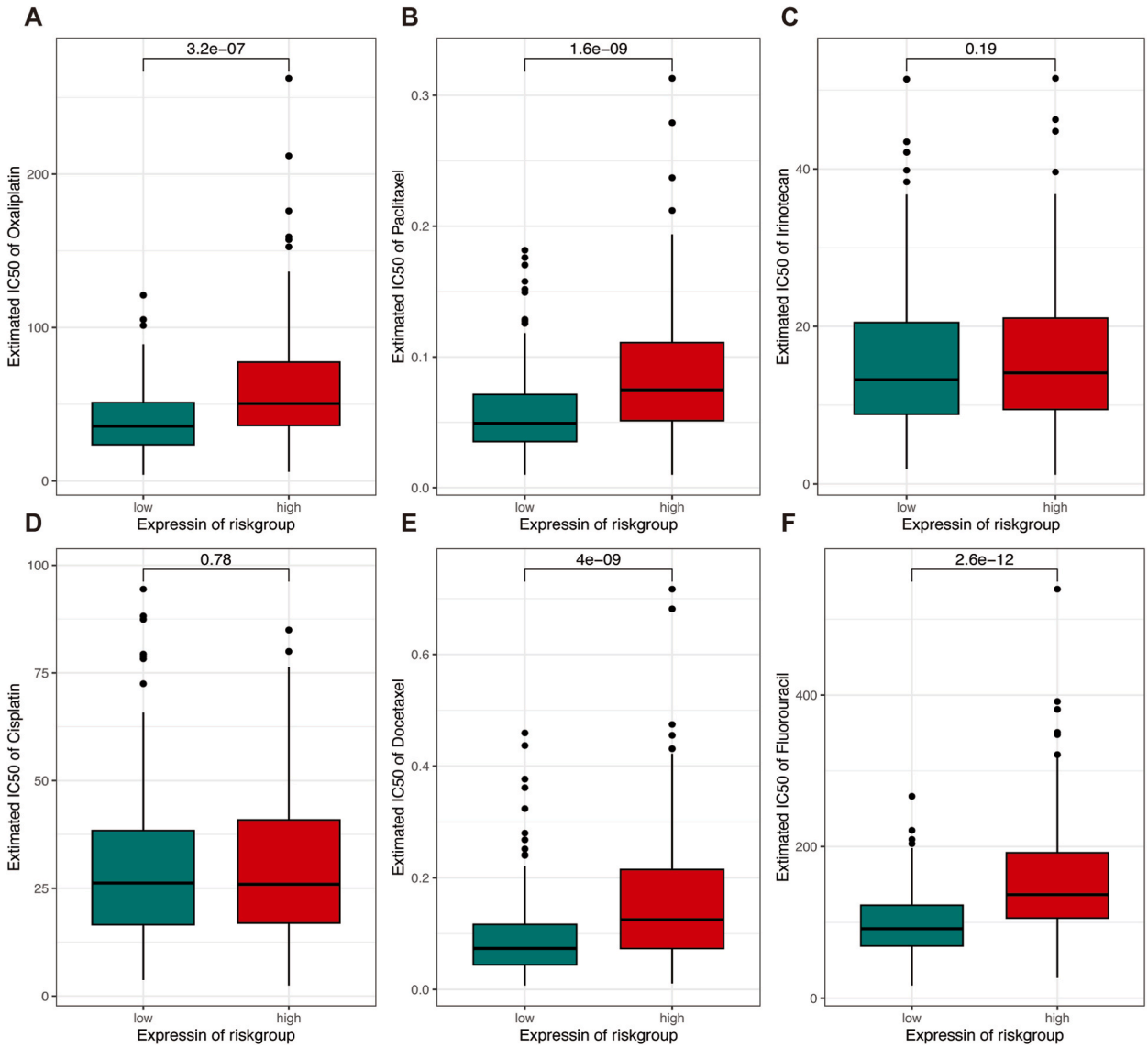


Fig. 7. Relationship between apoptosis risk score and chemotherapy sensitivity. (A) Risk score and the IC 50 of oxaliplatin. (B) Risk score and the IC 50 of paclitaxel. (C) Risk score and the IC 50 of irinotecan. (D) Risk score and the IC 50 of cisplatin. (E) Risk score and the IC 50 of docetaxel. (F) Risk score and the IC 50 of fluorouracil.

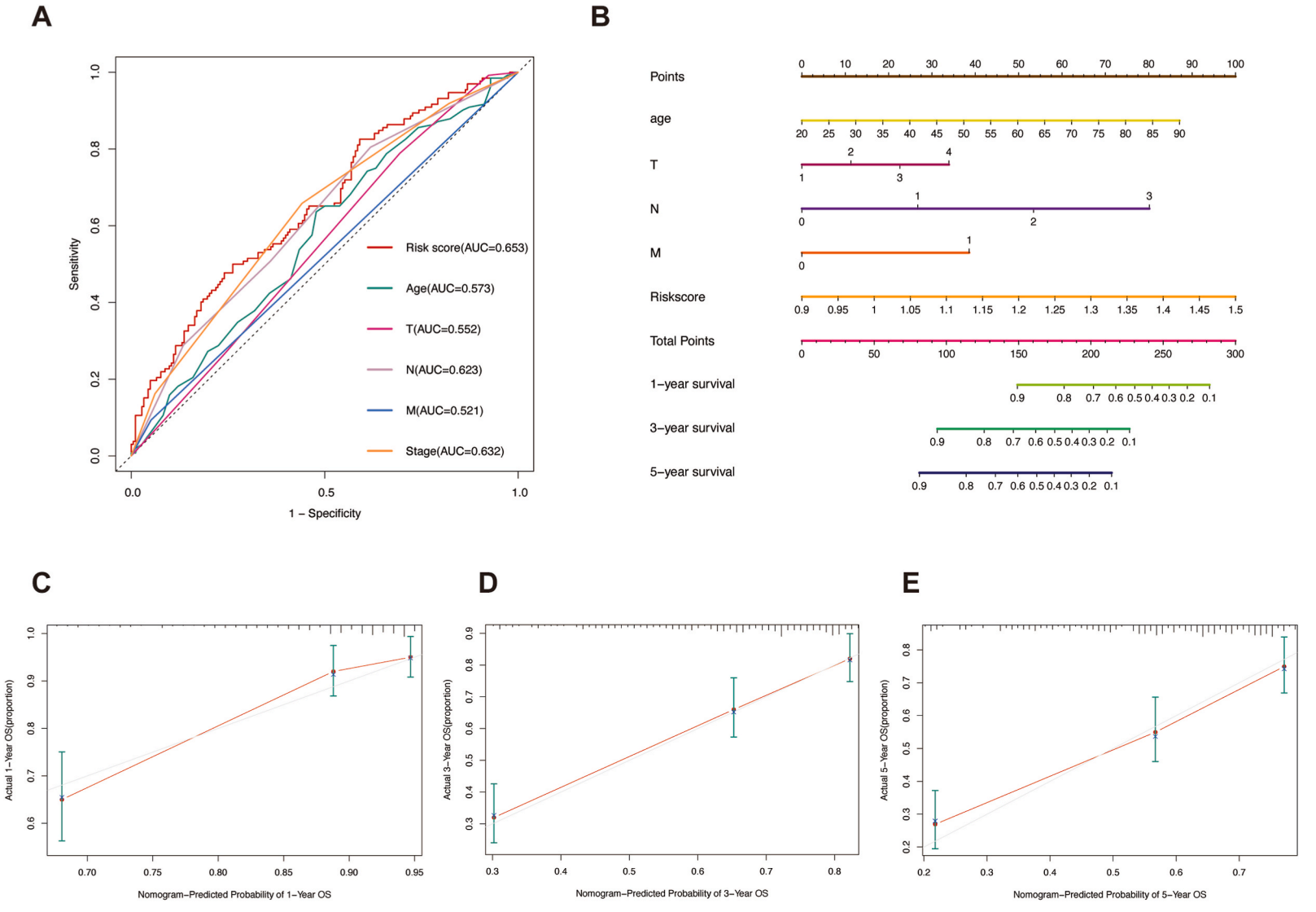


Fig. 8. Construction and evaluation of the prognostic nomogram. (A) ROC curves of the clinical characteristics and risk score. (B) The nomogram predicts the probability of the 1-, 3-, and 5-year OS in GC patients by combining risk score and four clinicopathological features. (C–E) The calibration plot is utilized to evaluate the accuracy of the 1-, 3-, and 5-year progress forecasts of the nomogram.

than the other parameters (Fig. 8A). Based on these findings, we generated a nomogram model that incorporated both the risk grade and clinical risk characteristics, which made it possible to estimate the probabilities of patients' survival for 1, 3, and 5 years. When comparing the predictive abilities of the clinical factors and risk score in the nomogram, the latter displayed a superior predictive ability (Fig. 8B). The correlation charts visually depict the relationship between the observed and the corresponding predicted survival rates of 1-, 3-, and 5-year (Fig. 8C–E). The high consistency between these two rates indicated the nomogram's excellently predictive ability.

4. Discussion

As we know, despite the improvements in multidisciplinary anticancer therapies, patients with GC continue to have an unsatisfactory 5-year survival rate, especially in the advanced stage of the disease [14]. Currently, the prognosis of these patients has mostly relied on the pathological judgment of biopsy basing on the extent of tumor penetration, the number of regional lymph nodes invaded as well as the stage of distant metastasis. Due to the invasive and complex nature of this technology, its use is limited. As next-generation sequencing continues to develop, bioinformatic analysis has emerged as a crucial tool for identifying molecular biomarkers capable of predicting the prognosis of numerous cancers. The process of apoptosis is an active, orderly, self-destructive event that does not induce inflammatory responses [15]. One of the major hallmarks of human cancers is that cells resist to apoptosis and proliferate indefinitely. Recently, increasing attention has been paid to ARGs, which may help assess the prognosis of cancers, including colon cancer, oral cancer, hepatocellular carcinoma, and osteosarcoma [16–19]. However, few prognostic signatures related to apoptosis are available to predict personalized survival in GC. Therefore, discovery of novel apoptosis-related prognostic signatures may provide valuable prognostic information as well as therapeutic targets for these patients.

According to this study, 174 apoptosis-related DEGs were screened in GC datasets, 45 of which were significantly related to prognosis. Finally, we constructed a risk model consisting of 12 optimal ARGs (*CTHRC1, PDGFRL, VCAN, GJA1, LOX, UPP1, ANGPT2, CRIM1, HIF1A, APOD, RNASE1,* and *ID1*) that can effectively predicte the prognosis for patients with GC. Some of these 12 genes are reportedly involved in gastric cancer. For example, CTHRC1 is an extracellular matrix protein, which has a high level of expression in GC. It was demonstrated that CTHRC1 could contribute to tumor invasion and metastasis by inhibiting tumor cell apoptosis [20]. In addition, Song et al. reported that the high level of VCAN in GC indicated a poor prognosis of these patients. Moreover, patients expressing low levels of VCAN benefited more from adjuvant chemotherapy, radiotherapy and immunotherapy [21]. Other genes among the 12 ARGs need to be investigated for their roles in the apoptosis of GC cells. According to our research, a risk model including 12 ARGs was constructed. Furthermore, the low-risk group of the model had a higher chance of survival than the high-risk group. For GC, it functioned as a standalone prognostic marker which provided additional evidence of its accuracy in gauging patient outcomes. Finally, the nomogram predicted GC prognosis at 1-, 3-, and 5-years. With the help of this model, high-risk patients can be screened and more reasonable treatment options can be devised.

In addition, we analyzed TMB, signaling pathways, TME, as well as the sensitivity to chemotherapeutics in different subgroups of GC according to the apoptosis risk score. TMB is commonly characterized as the cumulative count of non-synonymous somatic mutations affected by the emergence of neoantigens [22]. With a higher TMB, neoantigens are more likely to be recognized by the immune system. TMB is a significant biomarker in metastatic colorectal cancer with high microsatellite instability and an indicator of a patient's likelihood of responding to immunotherapy [23]. In our study, a higher TMB was identified in the low-risk subgroup. Among the two groups, the maximal mutation difference was noted in PCLO, the prevalence of which was higher in low-risk samples than in high-risk samples (25 % vs. 14 %). There were reports that mutation in TTN was linked to a higher TMB and better outcomes in several tumors, including GC and lung squamous cell carcinoma [24,25]. These findings are consistent with our results.

Immune cells constitute a substantial portion of the TME in solid tumors. Tumor growth may be promoted or inhibited by the infiltration of various innate and adaptive immune cells [26]. Apoptotic genes play important roles in the TME. For instance, BCL-2, an inhibitory gene in apoptosis, can induce immunosuppression by enhancing regulatory T cells abundance and cytotoxicity T lymphocyte depletion [27]. The present study found that monocytes, resting mast cells, as well as M2 macrophages were more abundant in high-risk subgroups. In cancer, elevated levels of resting mast cells, M2 macrophages, and monocytes have been linked to tumor progression, angiogenesis, and metastasis. They can contribute to tumor-promoting effects by facilitating immunosuppression, promoting inflammation, and supporting tumor cell invasion and migration [28–30]. However, patients in the low-risk subgroup exhibited elevated proportions of follicular helper T cells and resting CD4 + memory cells. A large number $CD4^{+}$ memory T cells infiltrating the TME indicate a favorable prognosis [31,32]. Based on the above findings, we propose that apoptosis may play a role in tumor immunity and this model may offer a suggestion to find reliable immune biomarkers for GC treatment.

Chemotherapeutic agents commonly target apoptosis. Various chemotherapeutic drugs induce apoptosis in cancer cells via different signaling pathways. For example, studies have demonstrated that paclitaxel can promote apoptosis via ROS/HIF-1α signaling pathway in prostate cancer cells [33], while anlotinib inhibited breast cancer cell growth by promoting cell apoptosis [34]. As there is a wide range of responses to antineoplastic drugs among patients, precision therapy is necessary to recommend suitable drugs for each patient. In this context, the apoptosis risk scores identified in our study significantly affected the sensitivity of patients with GC to oxaliplatin, paclitaxel, docetaxel, and fluorouracil, which are commonly used chemotherapeutic agents. Personalized chemotherapy according to the risk score may improve the sensitivity to chemotherapy and the clinical outcome of patients with GC.

In past studies, there have been several researches on the features about the prognosis of GC, including mitochondrial-related gene, microRNA, cancer-associated fibroblast related gene and tumor immunophenotyping-derived signatures [35–38]. As far as we know, there have been no correlational researches in prognostic studies on GC apoptosis-related signatures. Then, we developed an apoptosis model which was established with internal and external verification. Nevertheless, there are several limitations that may affect the

robustness and generalizability of our study. Only public databases were utilized to build and validate the prediction model. Therefore, larger sample sizes are necessary to testify the clinical applicability. Furthermore, the potential molecular mechanisms of the 12 genes in the signature of GC remain unclear and need to be notarized by in vivo or in vitro experiments using cell lines or clinical samples from patients with cancer.

5. Conclusion

We successfully established a novel prognostic risk model on the strength of 12 ARGs for patients with GC. The apoptosis risk score derived from this model stratified patients with GC into two distinct risk subgroups, thereby providing valuable prognostic information. Additionally, correlations were evaluated between the risk model and TMB, signaling pathways, tumor immunity, and chemotherapy sensitivity. Our findings highlighted the significance of ARGs as potential prognostic biomarkers. A risk model based on these genes is a promising tool for predicting patient outcomes and guiding treatment decisions for patients with GC.

Funding

This work was supported by the Fourth Hospital of Hebei Medical University.

Data availability statement

The datasets used in this study are available online. The raw data can be found in the article/supplementary files.

Ethics approval and consent to participate

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication

Not applicable.

CRediT authorship contribution statement

Xiaopeng Li: Writing – original draft, Conceptualization. **Xiaolei Yin:** Software, Formal analysis. **Lili Mi:** Supervision, Methodology. **Ning Li:** Software, Methodology. **Shumei Li:** Project administration, Funding acquisition, Conceptualization. **Fei Yin:** Writing – review & editing, Funding acquisition.

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.

Acknowledgements

The authors are thankful to the TCGA and GEO databases.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e33795.

References

[1] F. Bray, M. Laversanne, H. Sung, J. Ferlay, R.L. Siegel, I. Soerjomataram, et al., Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, Ca - Cancer J. Clin. 74 (3) (2024 May-Jun) 229–263. https://doi:10.3322/caac.21834.

[2] E.C. Smyth, M. Nilsson, H.I. Grabsch, N.C. Van Grieken, F. Lordick, Gastric cancer, Lancet 396 (10251) (2020 Aug 29) 635–648, https://doi.org/10.1016/S0140-6736(20)31288-5.

[3] W.J. Yang, H.P. Zhao, Y. Yu, J.H. Wang, L. Guo, J.Y. Liu, et al., Updates on global epidemiology, risk and prognostic factors of gastric cancer, World J. Gastroenterol. 29 (16) (2023 Apr 28) 2452–2468, https://doi.org/10.3748/wjg.v29.i16.2452.

[4] A. Jassim, E.P. Rahrmann, B.D. Simons, R.J. Gilbertson, Cancers make their own luck: theories of cancer origins, Nat. Rev. Cancer 23 (10) (2023 Oct) 710–724, https://doi.org/10.1038/s41568-023-00602-5.

[5] A. Strasser, D.L. Vaux, Cell death in the origin and treatment of cancer, Mol. Cell 78 (6) (2020 Jun 18) 1045–1054, https://doi.org/10.1016/j.molcel.2020.05.014.

[6] J.W. Kwon, J.S. Oh, S.H. Seok, H.W. An, Y.J. Lee, N.Y. Lee, et al., Combined inhibition of bcl-2 family members and yap induces synthetic lethality in metastatic gastric cancer with rasa1 and nf2 deficiency, Mol. Cancer 22 (1) (2023 Sep 20) 156, https://doi.org/10.1186/s12943-023-01857-0.

[7] Y. Kuang, Z. He, L. Li, C. Wang, X. Cheng, Q. Shi, et al., The developmental regulator hand1 inhibits gastric carcinogenesis through enhancing er stress apoptosis via targeting chop and bak which is augmented by cisplatin, Int. J. Biol. Sci. 19 (1) (2023 Jan 1) 120–136, https://doi.org/10.7150/ijbs.76345.

[8] D. Liu, Y. Xu, Y. Fang, K. Hu, Development of a novel immune-related gene signature to predict prognosis and immunotherapeutic efficiency in gastric cancer, Front. Genet. 13 (2022 May 27) 885553, https://doi.org/10.3389/fgene.2022.885553.

[9] N. Ma, J. Li, L. Lv, C. Li, K. Li, B. Wang, Bioinformatics evaluation of a novel angiogenesis related genes-based signature for predicting prognosis and therapeutic efficacy in patients with gastric cancer, Am J Transl Res 14 (7) (2022 Jul 15) 4532–4548.

[10] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, et al., Limma powers differential expression analyses for rna-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015 Apr 20) e47, https://doi.org/10.1093/nar/gkv007.

[11] B. Zhang, L. Xie, J. Liu, A. Liu, M. He, Construction and validation of a cuproptosis-related prognostic model for glioblastoma, Front. Immunol. 14 (2023 Feb 6) 1082974. https://doi:10.3389/fimmu.2023.1082974.

[12] A. Mayakonda, D.C. Lin, Y. Assenov, C. Plass, H.P. Koeffler Maftools, Efficient and comprehensive analysis of somatic variants in cancer, Genome Res. 28 (11) (2018 Nov) 1747–1756, https://doi.org/10.1101/gr.239244.118.

[13] A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, et al., Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (5) (2015 May) 453–457, https://doi.org/10.1038/nmeth.3337.

[14] T.H. Patel, M. Cecchini, Targeted therapies in advanced gastric cancer, Curr. Treat. Options Oncol. 21 (9) (2020 Jul 28) 70, https://doi.org/10.1007/s11864-020-00774-4.

[15] D. Thiyagarajan, H.L. Pedersen, N. Seredkina, K.D. Horvei, L. Arranz, R. Sonneveld, et al., Il-1β promotes a new function of dnase i as a transcription factor for the fas receptor gene, Front. Cell Dev. Biol. 6 (2018 Feb 6) 7, https://doi.org/10.3389/fcell.2018.00007.

[16] H. Tang, J. Wang, X. Luo, Q. Wang, J. Chen, X. Zhang, et al., An apoptosis-related gene prognostic index for colon cancer, Front. Cell Dev. Biol. 9 (2021 Dec 8) 790878, https://doi.org/10.3389/fcell.2021.790878.

[17] S. Wang, S. Zhang, Z. Lin, J. Ma, L. Zhu, G. Liao, Identification and validation of an apoptosis-related gene prognostic signature for oral squamous cell carcinoma, Front. Oncol. 13 (2023 Jan 10) 1100417, https://doi.org/10.3389/fonc.2022.889049.

[18] J. Yan, J. Cao, Z. Chen, Mining prognostic markers of asian hepatocellular carcinoma patients based on the apoptosis-related genes, BMC Cancer 21 (1) (2021 Feb 18) 175, https://doi.org/10.1186/s12885-021-07886-6.

[19] F. Yang, Y. Zhang, Apoptosis-related genes-based prognostic signature for osteosarcoma, Aging 14 (9) (2022 May 3) 3813–3825, https://doi.org/10.18632/aging.204042.

[20] M. Chivu-Economescu, L.G. Necula, L. Matei, D. Dragu, C. Bleotu, A. Sorop, et al., Collagen family and other matrix remodeling proteins identified by bioinformatics analysis as Hub genes involved in gastric cancer progression and prognosis, Int. J. Mol. Sci. 23 (6) (2022 Mar 16) 3214. https://doi:10.3390/ijms23063214.

[21] J. Song, R. Wei, S. Huo, C. Liu, X. Liu, Versican enrichment predicts poor prognosis and response to adjuvant therapy and immunotherapy in gastric cancer, Front. Immunol. 13 (2022 Sep 20) 960570. https://doi:10.3389/fimmu.2022.960570.

[22] D. Sha, Z. Jin, J. Budczies, K. Kluck, A. Stenzinger, F.A. Sinicrope, Tumor mutational burden as a predictive biomarker in solid tumors, Cancer Discov. 10 (12) (2020 Dec) 1808–1825, https://doi.org/10.1158/2159-8290.CD-20-0522.

[23] V. Martelli, A. Pastorino, A.F. Sobrero, Prognostic and predictive molecular biomarkers in advanced colorectal cancer, Pharmacol. Ther. 236 (2022 Aug) 108239, https://doi.org/10.1016/j.pharmthera.2022.108239.

[24] Y. Yang, J. Zhang, Y. Chen, R. Xu, Q. Zhao, W. Guo, Muc4, muc16, and ttn genes mutation correlated with prognosis, and predicted tumor mutation burden and immunotherapy efficacy in gastric cancer and pan-cancer, Clin. Transl. Med. 10 (4) (2020 Aug) e155, https://doi.org/10.1002/ctm2.155.

[25] X. Xie, Y. Tang, J. Sheng, P. Shu, X. Zhu, X. Cai, et al., Titin mutation is associated with tumor mutation burden and promotes antitumor immunity in lung squamous cell carcinoma, Front. Cell Dev. Biol. 9 (2021 Oct 21) 761758, https://doi.org/10.3389/fcell.2021.761758.

[26] J. Yang, J. Xu, W. Wang, B. Zhang, X. Yu, S. Shi, Epigenetic regulation in the tumor microenvironment: molecular mechanisms and therapeutic targets, Signal Transduct. Targeted Ther. 8 (1) (2023 May 22) 210, https://doi.org/10.1038/s41392-023-01480-x.

[27] L. Liu, X. Cheng, H. Yang, S. Lian, Y. Jiang, J. Liang, et al., Bcl-2 expression promotes immunosuppression in chronic lymphocytic leukemia by enhancing regulatory t cell differentiation and cytotoxic t cell exhaustion, Mol. Cancer 21 (1) (2022 Feb 22) 59, https://doi.org/10.1186/s12943-022-01516-w.

[28] T. Yamamoto, K. Kawada, K. Obama, Inflammation-related biomarkers for the prediction of prognosis in colorectal cancer patients, Int. J. Mol. Sci. 22 (15) (2021 Jul 27) 8002, https://doi.org/10.3390/ijms22158002.

[29] M. Nowak, M. Klink, The role of tumor-associated macrophages in the progression and chemoresistance of ovarian cancer, Cells 9 (5) (2020 May 22) 1299, https://doi.org/10.3390/cells9051299.

[30] R. Bahri, O. Kiss, I. Prise, K.M. Garcia-Rodriguez, H. Atmoko, J.M. Martínez-Gómez, et al., Human melanoma-associated mast cells display a distinct transcriptional signature characterized by an upregulation of the complement component 3 that correlates with poor prognosis, Front. Immunol. 13 (2022 May 20) 861545, https://doi.org/10.3389/fimmu.2022.861545.

[31] J. Wu, T. Zhang, H. Xiong, L. Zeng, Z. Wang, Y. Peng, et al., Tumor-infiltrating cd4 central memory t cells correlated with favorable prognosis in oral squamous cell carcinoma, J. Inflamm. Res. 15 (2022 Jan 8) 141–152, https://doi.org/10.2147/JIR.S343432.

[32] S. Ida, H. Takahashi, R. Kawabata-Iwakawa, I. Mito, H. Tada, K. Chikamatsu, Tissue-resident memory t cells correlate with the inflammatory tumor microenvironment and improved prognosis in head and neck squamous cell carcinoma, Oral Oncol. 122 (2021 Nov) 105508, https://doi.org/10.1016/j.oraloncology.2021.105508.

[33] Y. Zhang, Y. Tang, X. Tang, Y. Wang, Z. Zhang, H. Yang, Paclitaxel induces the apoptosis of prostate cancer cells via ros-mediated hif-1α expression, Molecules 27 (21) (2022 Oct 24) 7183, https://doi.org/10.3390/molecules27217183.

[34] S. Chen, Y. Gao, P. Zhu, X. Wang, L. Zeng, Y. Jin, et al., Anti-cancer drug anlotinib promotes autophagy and apoptosis in breast cancer, Front. Biosci. 27 (4) (2022 Apr 7) 125, https://doi.org/10.31083/j.fbl2704125.

[35] J. Chang, H. Wu, J. Wu, M. Liu, W. Zhang, Y. Hu, et al., Constructing a novel mitochondrial-related gene signature for evaluating the tumor immune microenvironment and predicting survival in stomach adenocarcinoma, J. Transl. Med. 21 (1) (2023 Mar 13) 191. https://doi:10.1186/s12967-023-04033-6.

[36] J. Xu, J. Song, X. Chen, Y. Huang, T. You, C. Zhu, et al., Genomic instability-related twelve-microRNA signatures for predicting the prognosis of gastric cancer, Comput. Biol. Med. 155 (2023 Mar) 106598. https://doi:10.1016/j.compbiomed.2023.106598.

[37] R. Xu, L. Yang, Z. Zhang, Y. Liao, Y. Yu, D. Zhou, et al., Cancer-associated fibroblast related gene signature in Helicobacter pylori-based subtypes of gastric carcinoma for prognosis and tumor microenvironment estimation in silico analysis, Front. Med. 10 (2023 Jan 18) 1079470. https://doi:10.3389/fmed.2023.1079470.

[38] J.B. Wang, Q.Z. Qiu, Q.L. Zheng, Y.J. Zhao, Y. Xu, T. Zhang, et al., Tumor immunophenotyping-derived signature identifies prognosis and neoadjuvant immunotherapeutic responsiveness in gastric cancer, Adv. Sci. 10 (15) (2023 May) e2207417. https://doi:10.1002/advs.202207417.