

A Risk Assessment Model for Prognosis Prediction: The Ferroptosis-related Gene Signature Based on Thyroid Cancer Mechanism Exploration from Multiple Perspectives

Xin Fan

Department of Otolaryngology Head and Neck Surgery, The First Affiliated Hospital of Nanchang University

Yun Zhong

The First Clinical Medical College of Nanchang University

Yating Fan

The First Clinical Medical College of Nanchang University

Lingling Zhang

School of Stomatology, Nanchang University

Fei Xie

Department of Otolaryngology Head and Neck Surgery, The First Affiliated Hospital of Nanchang University

Zhi-yuan Zhang (✉ zzyent@126.com)

Department of Otolaryngology Head and Neck Surgery, The First Affiliated Hospital of Nanchang University

Research Article

Keywords: Thyroid cancer, Ferroptosis, Prognosis, Gene signature, tumor microenvironment, The Cancer Genome Atlas database

Posted Date: March 22nd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1464127/v1>

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

[Read Full License](#)

Abstract

Background: Ferroptosis is an iron-dependent, new type of programmed cell death different from apoptosis, necrosis, and autophagy. At present, ferroptosis has been confirmed to be closely related to the prognosis and treatment of cancer. However, the relationship between ferroptosis and the progression and prognosis of different types of thyroid cancer (THCA) is unclear.

Methods: First, we performed differential expression analysis on the data of the two databases to obtain differentially expressed ferroptosis-related genes (DE-FRGs). Through differential expression, univariate Cox and lasso regression analysis, we identified 14 prognostic-related differently expressed ferroptosis-related genes (PR-DE-FRGs) for building risk assessment models. Subsequently, various validation methods are used to test the performance of the model. Then, we explored the mechanism of ferroptosis in the development and prognosis of THCA from the aspects of gene set enrichment analysis (GSEA), tumor microenvironment (TME), high frequency gene mutation, cell stemness and AL928654.4/miR-1287-5p/GPX4 regulatory axis. Finally, we verified the promising clinical application of the model.

Results: ANGPTL7, DRD4, SRXN1, TXNRD1, CDKN2A, MIOX, PGD and TFRC (HR>1) were identified as prognostic risk factors, whereas CAPG, GPX4, ARNTL, ISCU, BID and DPP4 (HR<1) were the opposite. Immunohistochemical (IHC) images and Gene Expression Omnibus (GEO) data validated differential expression of 7 and 12 PR-DE-FRGs, respectively. Through the study of the tumor microenvironment, we obtained the immune landscape of THCA and revealed to us that ferroptosis may influence the prognosis of THCA patients by affecting anti-tumor immunity. Subsequently, we found that the predicted and validated AL928654.4/miR-1287-5p/GPX4 regulatory axis may play an important role in regulating cancer cell apoptosis and ferroptosis. Finally, immunophenoscore, chemotherapeutic drug sensitivity analysis, ICIs related genes, and Nomogram were performed to prove the good clinical application prospects of the risk assessment model.

Conclusion: The risk assessment model can effectively predict the treatment effect and prognosis of THCA patients, and achieve the effect that can guide clinical treatment. Ferroptosis may be involved.

Introduction

Thyroid cancer (THCA) is the most frequent malignancy of the thyroid. THCA accounts for 3.0% (586,202 cases) of newly diagnosed cancer cases worldwide in 2020 and ranks ninth in cancer incidence in 2020. The incidence of females is relatively high, and males to females ratio of 1:3.3[1]. The national mortality is meager (< 2%) in both men and women with THCA [2]. However, both incidence and mortality of THCA have continued to rise in recent years [3]. The etiologies of THCA are unclear, and the only well-established risk factor is ionizing radiation [4]. The increase in ionizing radiation due to the rise in diagnostic imaging methods contributes to the rapid growth in THCA in recent years [5, 6].

Currently, surgery is the mainstay treatment for THCA [7]. However, extensive studies have shown that the clinical outcomes in the thyroid lobectomy and the total thyroidectomy groups were similar, and there

was no significant difference in patient survival [8–10]. After surgery, some patients need to be supplemented with endocrine therapy according to the patient's prognostic status, and comprehensive treatment including radiotherapy and chemotherapy should be selected if necessary. Considering the limitations of THCA treatment strategies, finding new therapeutic interventions for THCA is essential. This study confirmed that the risk score is significantly related to the immunophenoscore (IPS), the expression level of immune checkpoint inhibitors (ICIs)-related genes, and the half-maximal inhibitory concentrations (IC50) of chemotherapy drugs. The risk assessment model can effectively predict chemotherapy agents' sensitivity and effect. Stem cell characteristics can reveal novel targets for anti-cancer therapy[11][12]. Hence, evaluating the relationship between the risk assessment model and the features of tumor stem cells is beneficial for developing THCA treatment.

In recent years, targeting the ferroptosis of tumor cells has become a hot topic. Ferroptosis is an iron-dependent, new type of programmed cell death different from apoptosis, necrosis, and autophagy [13]. The primary mechanism of ferroptosis is severe lipid peroxidation, which depends on the generation of reactive oxygen species (ROS) and the participation of intracellular iron [13, 14]. Ferroptosis is gradually accepted as an adaptive feature that eliminates malignant cells. It plays a pivotal role in suppressing tumor formation by removing insufficient cells in critical nutrients in the environment or damaged by infection or environmental stressors [15]. Ferroptosis is expected to become a novel treatment for inducing cell death for resistant tumors to traditional therapies[16, 17]. Ferroptosis involves complex biochemical reactions involving different genes' expression, regulation, and signal systems. For example, Erastin[13, 18] RSL3 [19], Lanperisone [16, 20], SAS [13, 16, 21], glutamate [13, 16], and so on can induce ferroptosis in tumor cells. In recent years, although some studies have aimed to explore the mechanism of ferroptosis-related genes (FRGs) in THCA, there is a shortage of whole-genome analysis of ferroptosis-related miRNA and lncRNA, especially analysis based on high-throughput sequencing. In the first few years of discovering ferroptosis, several studies focused on glutathione metabolism to investigate the mechanism of ferroptosis. Therefore, the ability of GPX4 to prevent the accumulation of lipid peroxides has also received extensive attention [14, 22–24].

For this reason, we hope to explore the regulation mode of GPX4-related ceRNAs, one of the prognostic-related differently expressed ferroptosis-related genes (PR-DE-FRGs), in the occurrence, development, and prognosis of THCA to contribute a unique perspective for transcriptome research and resolve some academic phenomena more comprehensively and in-depth. To gain insight into the biochemical reactions related to ferroptosis from thyroid cancer, we used gene set enrichment analysis (GSEA) to investigate the biological processes and pathways of PR-DE-FRGs. Gene Ontology (GO) analysis found that immune-related biological processes and pathways are enriched, so we further reveal the relationship between the risk assessment model and the tumor microenvironment (TME). Numerous studies show that ferroptosis is associated with RAS gene mutations [13, 25]. Studying the relationship between risk score and RAS mutations in thyroid cancer is crucial in studying ferroptosis. At the same time, we analyzed the mutation characteristics of PR-DE-FRGs. And through immunohistochemical (IHC) staining to verify the difference of PR-DE-FRGs protein expression between thyroid cancerous and normal tissues. Clinically, there is a need for decision support tools that can easily predict patients' 1-, 3-, and 5-year overall survival. A single

independent prognostic factor has low accuracy in predicting the prognosis of patients. Nomogram has been widely employed to predict cancer's overall survival at 1-, 3-, and 5-years [26]. Therefore, we used factors (age, stage, gender, and risk score) that may affect the prognosis to build a Nomogram for predicting thyroid cancer's 1-, 3-, and 5-year overall survival.

Collectively, in this study, we established a risk assessment model related to ferroptosis, deeply analyzed the mechanism of ferroptosis in THCA, and finally explored the possibility of this model being widely used in clinical practice.

Methods

Data collection and DE-FRGs extraction

We follow the methods of our previous research[27] in writing the methodology section of this article. On April 1, 2021, we downloaded the RNA sequence data and relevant clinical data of 510 THCA samples and 58 adjacent normal samples from the TCGA database. The list of FRGs used in this article came from the FerrDb database[27]. Differently expressed ferroptosis-related genes (DE-FRGs) between adjacent non-tumor tissue and tumor tissue were determined by the R package "limma" with an FDR < 0.05.

Establishment of the regulatory network and risk assessment model

The data in the TCGA database was screened, and we retained 501 samples with complete overall survival (OS) and RNA-seq data. 15 PR-DE-FRGs screened out by univariate Cox analysis. The 501 samples were randomly allocated to the training set (n = 301) and the test set (n = 200) at a ratio of 6:4[27]. To screen out highly relevant genes and minimize the adverse effects of overfitting of selected features, we performed the lasso regression analysis on the expression values of 15 PR-DE-FRGs in the training group to predict the clinical efficacy of THCA cases accurately. A risk assessment model based on 14 PR-DE-FRGs was constructed by choosing the best Lasso penalization parameter (λ) determined by the smallest k-fold cross-validation with K = 10. (Supplementary Figure S1 C-D). The sample's risk score was calculated according to the expression level of each PR-DE-FRGs, and the coefficient obtained in the lasso regression algorithm. Risk score equation:

$$\text{Risk score} = \sum(\text{Gene expression value} \times \text{Gene corresponding coefficient})$$

We used the STRING database to construct a PPI network based on 14 PR-DE-FRGs. The R package "reshape2", "psych", "RColorBrewer", and "igraph" were used for correlation analysis of 14 genes and visualization of related networks[27]. The establishment of the correlation network diagram was based

on 13 associated genes. Similarly, the establishment of PPI was established on 12 genes that interact with other genes. In addition, the Kaplan-Meier survival curve was used to analyze the relationship between the expression of 14 genes and OS in the model.

Evaluation of Risk assessment model

By calculating the risk scores of the training/test/whole sets separately, each group of samples was divided into a high-risk group and a low-risk group based on the median value of the risk score. We performed the R package "survival" and "survminer" to perform the Kaplan-Meier test to identify the difference in survival between the high-risk and the low-risk groups[28]. The multivariate receiver operating characteristic (ROC) curve was used to verify the model's ability to predict prognosis and assess whether it has more predictive value than other single prognostic factors. According to the expression of genes in the model, the "prcomp" function of the R package "stats" was used to perform principal component analysis (PCA) to explore the distribution of different risk populations[29]. Finally, we performed univariate and multivariate Cox regression analysis to test whether the risk assessment model can be an independent predictor of prognosis.

Stratified analysis

To assess whether each clinical feature is related to the risk score, we compared the differences in risk scores between the different subtypes of each clinical feature and visualized the corresponding results in boxplots. Kaplan-Meier test was used to evaluate the model's capability to foretell OS in each subgroup of different clinical characteristics.

Gene set enrichment analysis

Based on the PR-DE-FRGs between the high-risk and low-risk groups, GSEA was performed using Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathways and GO functions[30]. Select "c2.cp.kegg.v7.4.symbols.gmt" and "c5.go.v7.4.symbols.gmt" as the gene set database. Normalized enrichment score (NES) was calculated by setting the permutation value to 1000, and significant enrichment pathways and functions were screened using a P-value < 0.05 or an FDR Q value < 0.05 [31]. The R package "GSVA" was used for GSEA.

The relationship between risk assessment model and tumor microenvironment

We used the R package "estimate" to calculate each sample's immune cell score and stromal cell score and analyze the difference in immune cell score and stromal cell score between the high-risk and low-risk groups. Spearman correlation analysis was used to determine the correlation between risk score and the immune cell score or stromal cell score. The single-sample gene set enrichment analysis (ssGSEA) based on the R packages "GSEABase" and "GSVA" was used to quantify the scores[32] of 16 immune cells and 13 immune functions. Similarly, Spearman correlation analysis evaluated the correlation between each immune cell, immune function, and risk score. Based on the risk assessment model, the difference

analysis of different risk groups' immune cell scores and immune function scores was carried out, and boxplots were drawn.

Stemness Features of THCA

We downloaded the accurately calculated tumor cell stemness index (mDNAsi) from related research. The stemness index in the study was calculated according to the one-class logistic regression (OCLR) algorithm, which was trained on stem cell types (ESC: embryonic stem cells; iPSC: induced pluripotent stem cells) their differentiated ectoderm, mesoderm, and endoderm progenitor cells [1]. Apply OCLR-based methylation data to TCGA data set to calculate mDNAsi. Each mDNAsi ranges from low (zero) to high (one)[33]. To explore the correlation between the predictive model and mDNAsi, we used Spearman correlation analysis to determine the correlation between risk score and mDNAsi. We used the Wilcoxon rank-sum test to compare mDNAsi differences between high-risk and low-risk groups. In addition, the Wilcoxon rank-sum test was used to compare further the differences in mDNAsi among different subgroups of clinicopathological characteristics.

Gene mutation and tumor mutational burden analysis

Obtain the genetic mutation data of THCA samples and the relevant clinical data from TCGA database. The integrated data was used to analyze the association between gene mutations and risk score models. Use the R package "GenVisR" to visualize the 20 most frequently mutated genes[34] in the high-risk and low-risk groups, respectively. The R package "maftools" was used to calculate TMB, defined as the number of somatic non-synonymous mutations in a specific genomic region[35]. It was usually expressed as the number of modifications per megabase (mut/Mb)[35]. To explore the predictive effect of TMB on the prognosis of patients, we divided the samples into high TMB and low TMB groups. We performed Kaplan-Meier survival analyses on the two groups. According to the mutation status of NRAS or HRAS, THCA samples were divided into the wild and mutant groups. To explore the association between ferroptosis and RAS (NRAS and HRAS) mutations, we compared the difference in risk scores between the RAS mutant group and the RAS wild group. Not only that, but Kaplan-Meier survival analysis was used to compare differences in OS.

The clinical utility of the risk assessment model

Obtain the corresponding IPS data of THCA samples from The Cancer Immunome Atlas (TCIA) database. The main immunological measures determining immunogenicity were effector cells, immunosuppressive cells, major histocompatibility complex (MHC) molecules, and selected immunomodulators. Calculate the expression levels of these four to acquire the patient's IPS. Spearman correlation analysis was used to detect the correlation between the risk score and the expression of 3 types of IPS and 8 ICIs-related genes (CD274/PDL1, CD40, CTLA4, PDCD1/PD1, CD96, TIGIT, LAG3, and HAVCR2)[27]. Then use the R package "ggplot2" to visualize the analysis results. The ridge regression model was constructed using the cell line expression data in the Genomics of Drug Sensitivity in Cancer (GDSC) and the RNA sequence data in the TCGA database. We used the R package "pRophetic" to predict the IC50 of the five chemotherapy drugs (sorafenib, Axitinib, mesylate, pazopanib, and sunitinib) recommended by the National Comprehensive

Cancer Network (NCCN) guidelines in the high-risk and the low-risk group[27]. Then we analyzed the IC50 difference between the two groups. Eventually, Spearman correlation analysis was used again to evaluate the correlation between the risk scores of the five chemotherapeutic drugs and IC50.

Establishment and verification of Nomogram

Clinically, an operational decision aid tool was needed. Nomogram is one of the best choices and is widely used to predict cancer's 1, 3, and 5-year overall survival [26]. Integrate the risk score and other factors that may affect survival (age, stage, gender), and use the R package "rms" to build a multi-factor cox regression model to create a Nomogram. To determine the accuracy of the predicted value of the Nomogram compared to the actual value, we drew calibration curves. Finally, to understand the independent predictive power of the Nomogram on the prognosis of THCA, we drew ROC curves of 1-, 3-, and 5-year.

Immunohistochemistry assay

To further check the differential expression of the 14 genes used to build the model between thyroid cancer and normal thyroid tissue, we used the results of IHC staining from the Human Protein Atlas (HPA) database. Finally, we obtained immunohistochemical images of the protein expression of 11 genes in thyroid cancer and normal thyroid tissues.

Differential expression verification of PR-DE-FRGs based on external cohort

On September 5, 2021, we downloaded the external cohort data from the Gene Expression Synthesis (GEO) database. The RNA sequencing data of 82 samples from the GSE33630 cohort (78 THCA And 4 normal thyroid tissues) was used to verify the differential expression of the 14 PR-DE-FRGs used to construct the model. The R package "limma" was used to analyze the differential expression of 14 PR-DE-FRGs between THCA tissue and adjacent normal tissues for creating the model, according to $|\log_2FC| > 0$ and $FDR < 0.05$ [27]. And we showed the differential expression of these PR-DE-FRGs through heat map and volcano map.

Discovery and verify the ceRNA regulatory network related to GPX4

To explore gene functions and regulatory mechanisms at a deeper level, we selected GPX4 in the model to predict and verify the complete ceRNA regulatory axis. To confirm the differential expression of GPX4 between the cancerous and normal groups, we downloaded the RNA-seq data of 18 common cancers from the UCSC Xena database. Gene Expression Profiling Interactive Analysis (GEPIA) is a network tool server for cancer and normal gene expression profiling and interactive analysis [36], which was used to verify further the survival difference on the expression of GPX4 in THCA. Subsequently, the miRNA expression data of 514 THCA and 59 adjacent noncancerous tissues were obtained from the TCGA database. After obtaining the mature miRNA files from mirbase for annotating, we received the miRNA

expression matrix of these samples. Run multiple target gene prediction programs (including PITA, RNA22, miRmap, microT, miRanda, PicTar, and TargetScan in StarBase), and miRNAs that appear more than twice were included as the candidate upstream binds miRNAs of GPX4. Cytoscape (v3.7.2) was used to map the co-expression network of miRNA and GPX4. The R packages "ggExtra" and "reshape2" were performed to explore the expression correlation between GPX4 and the upstream binding miRNA. The differential expression of miRNAs (correlation coefficient $r < -0.2$, $p < 0.001$) in cancerous and normal tissues depends on the difference analysis. The Kaplan-Meier survival curve was used to understand the difference in survival probability between the high and low miRNA expression subgroups.

hsa-miR-1287-5p, as the only miRNA with statistical significance ($p < 0.05$), was used for subsequent analysis. StarBase was again performed to predict candidate lncRNAs that bind to hsa-miR-1287-5p. Using the above methods, we mapped the co-expression network of lncRNAs and hsa-miR-1287-5p and explored the correlation between lncRNAs and hsa-miR-1287-5p or GPX4 expression, as well as the difference and survival analysis of lncRNAs. Eventually, a ceRNA regulatory network composed of AL928654.4, hsa-miR-1287-5p, and GPX4 was drawn by running Cytoscape (v3.7.2).

Statistical analysis

According to the distribution characteristics, Student's t-test or Mann-Whitney test were used to compare the differences between continuous variables. The chi-square test or Fisher's exact test was used to compare the differences between categorical variables. Wilcoxon rank-sum test compares the differences of mDNasi among different subgroups. Univariable Cox regression analysis was used to screen out PR-DE-FRGs, and lasso regression was used to screen out the best genes for constructing a risk model. The 10-fold cross-validation was used to determine the optimal penalty parameter (λ). The Kaplan-Meier curve of the log-rank test was used to analyze the differences in OS between groups. This research used the multivariate Cox regression method to construct Nomogram. Univariate and multivariate Cox regression analyses were used to evaluate the independent predictive value of the risk score of the risk assessment model on survival. The ROC curve was used to assess the predictive ability of each factor on the prognosis. Spearman correlation analysis was used to analyze the correlation between scores, such as evaluating the risk scores and five chemotherapy drugs or IC50. The OCLR algorithm calculates the tumor cell stemness index. All analyses in our study were performed using the R programming language (version 4.0.3) and SPSS Statistics 22 software²². If not specified in the text, a two-sided P value less than 0.05 is statistically significant in all the statistical results.

Results

Data collection and DE-FRGs extraction

The flowchart of this research is shown in Fig. 1. This study included 510 THCA samples and 58 adjacent normal tissue samples in TCGA database. To use the R software to analyze the data, we use the gene transfer format to annotate the shared genome. According to the list of FRGs obtained from the FerrDb database, 246 FRGs were screened out. After analyzing gene expression and quantification, 176 DE-FRGs

were obtained. These DE-FRGs were presented in Supplementary Fig. S1A and S1B with a heat map and a volcano chart. As shown in the figure, 78 genes were identified as up-regulated, and 98 genes were down-regulated.

Establishment of regulatory network and risk assessment model

The data obtained from TCGA was screened, and we retained 501 samples with complete OS and RNA sequence data. The clinical features are shown in Table 1. Fifteen DE-FRGs related to OS were screened out by univariate Cox analysis (Fig. 2A and 2B). Among them, the prognostic risk factors were identified as ANGPTL7, DRD4, SRXN1, TXNRD1, CDKN2A, MIOX, PGD, TF, and TFRC (HR > 1). The factors identified as prognostic protective factors are CAPG, GPX4, ARNTL, ISCU, BID, and DPP4 (HR < 1). The lasso regression analysis was performed on the expression values of fifteen PR-DE-FRGs in the training group to screen out highly related genes to prognosis. The optimal penalty parameter (λ) determined by the smallest k-fold cross-validation with K = 10, which is based on a risk assessment model constructed by fourteen PR-DE-FRGs (GPX4, CAPG, TFRC, PGD, MIOX, DPP4, CDKN2A, BID, DRD4, ISCU, ARNTL, TXNRD1, ANGPTL7, and SRXN1) (Supplementary Fig. S1 C-D). The PR-DE-FRGs used to construct the model, and the corresponding coefficients are shown in Supplementary Table 1. Seven central genes (PGD, SRXN1, GPX4, ISCU, TFRC, TXNRD1, and CDKN2A) were found (Fig. 2C). By constructing 12 PR-DE-FRGs PPIs. Through the construction of thirteen PR-DE-FRGs related networks, the attributes of genes and their effects on prognosis are shown (Fig. 2D). By examining the correlation between fourteen PR-DE-FRGs and survival probability, we found that high expression of BID, CAPG, GPX4 genes and low expression of TFRC showed better OS (Fig. 3).

Table 1
Clinical characteristics of the THCA samples in training, test, whole set.

	Training set (n = 301)	Test set (n = 200)	Whole set (n = 501)
Gender (%)			
Male	86(28.6%)	49(24.5%)	135(26.9%)
Female	215(71.4%)	151(75.5%)	366(73.1%)
Age (median, range)	48(15–88)	47(18–89)	47(15–89)
Survival status			
OS-days (median, range)	1026(9-5423)	1055(6-5150)	1038(6-5423)
OS-state (alive (%)/ Dead (%))	291(96.7%)/10(3.3%)	194(97.0%)/6(3.0%)	485(96.8%)/16(3.2%)
Stage (%)			
I	167(55.5%)	114(57.0%)	281(56.1%)
II	37(12.3%)	15(7.5%)	52(10.4%)
III	65(21.6%)	46(23.0%)	111(22.1%)
IV	32(10.6%)	23(11.5%)	55(11.0%)
unknown	0(0.0%)	2(1.0%)	2(0.4%)
T (%)			
1	81(26.9%)	61(30.5%)	142(28.4%)
2	102(33.9%)	62(31.0%)	164(32.7%)
3	105(34.9%)	65(32.5%)	170(33.9%)
4	12(4.0%)	11(5.5%)	23(4.6%)
unknown	1(0.3%)	1(0.5%)	2(0.4%)
M (%)			
0	175(58.1%)	107(53.5%)	282(56.3%)
1	6(2.0%)	3(1.5%)	9(1.8%)
unknown	120(39.9%)	90(45.0%)	210(41.9%)
N (%)			
0	148(49.2%)	81(40.5%)	229(45.7%)

	Training set (n = 301)	Test set (n = 200)	Whole set (n = 501)
1	119(39.5%)	103(51.5%)	222(44.3%)
unknown	34(11.3%)	16(8.0%)	50(10.0%)

Evaluation of Risk assessment model

To evaluate the predictive ability of the risk assessment model, we simultaneously analyze the relationship between the distribution of risk scores in the training set, test set, and the whole set and the patient's OS and OS status. By drawing risk plots, we can find that the survival of the low-risk group is dramatically better than that of the high-risk group (Fig. 4A-C). Next, draw Kaplan-Meier survival curves, and the results show that the overall survival of the low-risk group is significantly better in the training set, test set, and whole set (Fig. 4D-F). The training set showed that the high-risk and low-risk samples are distributed in the two principal component directions of the PCA discrete trend graph (Fig. 4G). We found similar results in the test and the whole set (Fig. 4H-I). Univariate and multivariate Cox regression analyses, including risk score and OS, were performed to evaluate whether the model can be used as an independent OS predictor of THCA patients. Univariate Cox regression analysis results show that risk score is significantly related to OS. After adjusting for other prognostic-related factors, we performed a multivariate Cox regression analysis, and the results still showed a significant correlation between risk score and THCA (Table 2). Two studies suggest that this risk score can be used as an independent prognostic indicator of THCA.

Table 2
The results of univariate and multivariate Cox regression analysis, including clinical factors and risk scores in three sets.

	Univariate Cox		p ¹	Multivariate cox		p ²
	HR	95%CI		HR	95%CI	
Training set						
Age	1.11	1.04–1.19	0.002	1.09	1.00-1.20	0.06
Gender	3.79	0.75–19.08	0.105	3.92	0.49–31.26	0.197
Stage	3.53	1.44–8.62	0.006	2.06	0.52–8.09	0.301
Risk score	1.08	1.04–1.12	< 0.001	1.08	1.03–1.13	< 0.001
Test set						
Age	1.16	1.07–1.26	< 0.001	1.11	1.03–1.20	0.006
Gender	1.13	0.22–5.86	0.89	0.65	0.07–6.28	0.71
Stage	2.21	1.10–4.44	0.03	0.85	0.28–2.54	0.78
Risk score	2.20	1.54–3.14	< 0.001	1.64	1.03–2.62	0.038
Whole set						
Age	1.40	1.08–1.20	< 0.001	1.14	1.07–1.21	< 0.001
Gender	2.02	0.65–6.21	0.222	2.17	0.58–8.18	0.252
Stage	2.75	1.59–4.74	< 0.001	1.44	0.67–3.07	0.351
Risk score	1.07	1.04–1.10	< 0.001	1.07	1.04–1.11	< 0.001

To make the results more reliable and test the predictive performance of the risk model, we then draw the ROC curve. The risk score is more sensitive to OS prediction in the training, test, and whole set. The training set's 1-year, 2-year, and 3-year Area Under Curve (AUC) were 0.939, 0.968, and 0.922, respectively (Fig. 5A-C). In the test set, they were 0.921, 0.943, and 0.918 (Fig. 5D-F). In the whole set, they were 0.926, 0.959, and 0.912 (Fig. 5G-I). It is worth noting that the 1-year, 2-year, and 3-year AUC in the three groups were all greater than 0.9, and the risk score was the largest AUC among all factors. It is again confirmed that the risk assessment model can be used as a prognostic indicator of THCA, which has outstanding predictive value for the prognosis of patients.

Stratified analysis

It is worth mentioning that many clinically valuable results were found when analyzing the relationship between risk scores and various clinical characteristics. Age (Fig. 6A, $p < 0.01$), survival status (Fig. 6B, $p < 0.001$), gender (Fig. 6C, $p < 0.01$), stage (Fig. 6D, $p < 0.05$) and N stage (Fig. 6G, $p < 0.001$) was observed

to be significantly association with the risk score. Among them, the risk scores of patients in the age > 70-year-old group, the male group, and the N0 group were higher. Figure 6D shows that the patient's risk score also decreases as the stage increases. Unfortunately, the risk scores of different ages and fustat subgroups did not show significant differences (Fig. 6E-F). It is found that the model still has the better predictive ability for OS of each subgroup with different clinical characteristics in further survival analysis. It was observed that the OS of patients in the high-risk group in the age \leq 70 years old group, age > 70 years old group, Female, Male, Stage I-II, T3-4, M0, N0, and N1 subgroups was poor (Fig. 6H, $p < 0.05$). Among them, Male, Age > 70, Stage I-II, T3-4, N1 subgroups are most closely related to the prognosis of patients.

Gene set enrichment analysis

Gene set enrichment analysis to understand the differences in the gene expression between the THCA high-risk and the low-risk groups. As shown in Fig. 7A, the GO enrichment analysis of the two groups found that the main enrichment pathways for the high-risk group are hormone biosynthetic process, thyroid hormone metabolic process, vascular transport, cis-Golgi network, TOR complex, anion transmembrane transporter activity, organic anion transmembrane transporter activity. The low-risk group responds to interleukin-4, T-helper 1 cell differentiation, T-helper 2 cell differentiation, BLOC-1 complex, BLOC complex, cell-cell adhesion mediator activity, and lipase inhibitor activity. Interestingly, the hormone biosynthetic process and thyroid hormone metabolic process, closely related to thyroid function, were significantly enriched in the high-risk group (p -value < 0.001). T-helper 1 cell differentiation, T-helper 2 cell differentiation, and lipase inhibitor activity related to ferroptosis were found to be significantly enriched in the low-risk group ($p < 0.05$). We also performed a KEGG pathway analysis to verify further the difference of Biological Function and Pathway between the high-risk and low-risk groups. As shown in Fig. 7B, the results are as expected. The Adipocytokine signaling pathway, Fatty acid metabolism, Glycine, serine and threonine metabolism, Hedgehog signaling pathway, mTOR signaling pathway, regulation of autophagy enriched in the high-risk group are all closely related to ferroptosis. Through the above GSEA, we can find that the risk assessment model is closely associated with changes in thyroid function and ferroptosis.

The relationship between risk assessment model and tumor microenvironment

There is a significant correlation between risk score and immune cell score (Fig. 8A). Further difference analysis found that the immune cell score of the low-risk group was significantly higher than that of the high-risk group (Fig. 8B). Disappointingly, there is no significant relationship between risk score and stromal score (Fig. 8C-D). The correlation and difference between risk scores and the scores of 16 immune cells or 13 immune functions can help us understand the impact of ferroptosis on THCA immunity. Except for CD8 + T cells, Tfh, and B cells, immune cells and immune functions were significantly correlated with risk score (Fig. 8E, G). The results of the different analysis showed immune cells related to ferroptosis such as T-helper-cells, Th1 cells, Th2 cells, TIL, Treg, aDCs, DCs, iDCs, pDCs,

Macrophages, neutrophils, mast cells, and Natural killer (NK) cells in the low-risk group were significantly higher than those in the low-risk group (all adjusted $P < 0.05$) (Fig. 8F). Immune functions related to ferroptosis, such as T cell co-stimulation and T cell co-inhibition, and Type II Interferon (INF) Response, were also significantly higher in the low-risk group than in the high-risk group (all adjusted $P < 0.05$). In addition, Antigen-presenting cells (APC) co-inhibition, APC co-stimulation, CCR, checkpoint, inflammation promoting activity, HLA and MHC, Parainflammation, and Type I INF Response of inflammation-related pathways were also significantly enriched in the low-risk group (Fig. 8G-H). The significant differences in ferroptosis-related immune cells and immune pathways between the high and low-risk groups are similar to the results of GSEA, which suggests that in THCA, ferroptosis participates in the occurrence and development of tumors participating in the immune activities.

Stemness Features of THCA

We observed a negative correlation between mDNasi and risk score (Fig. 9A). In further difference analysis, we observed that the mDNasi of patients in the low-risk group was significantly higher, supporting the correlation analysis's conclusion (Fig. 9B). In addition, we also found that patients with stage N1 had significantly higher mDNasi (Fig. 9C).

Gene mutation and TMB analysis

We draw waterfall plots to explore differences in mutation characteristics between the high-risk and low-risk groups. The results of the waterfall plots show the top 20 most frequently mutated genes in the high-risk group and low-risk group (Fig. 10A-B). The boxplot results showed no significant difference in TMB between the two groups (Fig. 10C). To determine the relation between TMB and the prognosis of THCA patients, we further performed survival analysis. The results showed that the OS of the low TMB samples was significantly better ($p < 0.001$, Fig. 10D). We found that the NRAS mutation group was significantly correlated with a higher risk score by analyzing the relationship between the mutated gene and the risk score (Fig. 10E). Further survival analysis found that patients in the NRAS mutation group were associated with better OS (Fig. 10F). Although not significant, similar results were also observed in the HRAS mutation group (Fig. 10G-H).

The clinical utility of risk assessment models

This study thoroughly studied the correlation between the risk assessment model evaluated the effect of ICIs therapy on THCA patients through IPS. Patients with higher IPS scores will get better results in ICIs treatment [37]. We can observe that IPS-PD1, IPS-CTLA4, and IPS-PD1 + CTLA4 were negatively correlated with the risk score (Fig. 11A). Compared with the low-risk group, the three IPS scores of the high-risk group were lower. (Fig. 11B). CD274 (PD-L1), TIGIT, CTLA4, CD40, and HAVCR2 were significantly negatively correlated with risk scores (all p values < 0.05 ; Fig. 11C).

Further difference analysis also confirmed the relevant results. Except for CTLA4, the expression levels of the other four genes were lower in the high-risk group (all p values <0.05; Fig. 11D). This result indicates that our model can be used to predict the effect of immunotherapy. We expected the IC50 of each sample of the five chemotherapy drugs recommended by the NCCN guidelines (sorafenib, axitinib, mesylate, pazopanib, and sunitinib). When axitinib, pazopanib, and sorafenib were negatively correlated with the risk score, the IC50 of mesylate and sunitinib was positively correlated with the risk score (Fig. 11E). Further difference analysis results support the correlation conclusion (Fig. 11F). These conclusions mean that our risk score model can effectively predict the sensitivity of patients receiving these five chemotherapy drugs.

Establishment and verification of Nomogram

The Nomogram constructed from age, gender, stage, and risk score showed that risk score is the most important factor affecting patient survival, followed by age, stage, and gender (Fig. 12A). Through the internal calibration curve, it can be confirmed that Nomogram has a brilliant agreement between the predicted and actual 1-, 3-, and 5-year OS (Fig. 12B), showing its independent predictive ability for the prognosis of THCA. The ROC curves confirm that the Nomogram has excellent predictive power (AUC values are all > 0.9, Fig. 12C).

Immunohistochemistry assay

We obtained IHC staining images of the protein expression of 11 PR-DE-FRGs (PGD, DPP4, CDKN2A, BID, ISCU, TXNRD1, CAPG, TFRC, MIOX, ARNTL, and GPX4) in thyroid cancer and normal thyroid tissues. By comparing the immunohistochemical staining images of thyroid cancerous and normal tissues, we found that the protein expressions of 7 PR-DE-FRGs (PGD, DPP4, CDKN2A, BID, ISCU, TXNRD1, and CAPG) are different (Fig. 13A-G). The protein expression of DPP4, KN2A, BID, ISCU, and CAPG in THCA tissue is higher than that in normal tissue. The expression of PGD and TXNRD1 protein is higher in normal tissues. All of these support the results of the differential expression of these genes in our analysis. The images of the remaining 4 PR-DE-FRGs are shown in Supplementary Figure S2, and there is no difference in protein expression.

Differential expression verification of PR-DE-FRGs based on external cohort

We finally obtained RNA sequencing data of 12 PR-DE-FRGs from the GSE33630 cohort. After running the differential expression analysis, these 12 genes all showed significant differences. MIOX, SRXN1, PGD, TXNRD1, and DRD4 are down-regulated. At the same time, GPX4, CAPG, DPP4, CDKN2A, BID, ISCU, and ARNTL are up-regulated (Fig. 14A-14B), which is consistent with the differential expression results of PR-DE-FRGs from the TCGA database (Fig. 14C).

The performance of GPX4-related ceRNA regulatory network in THCA

Except for Breast invasive carcinoma (BRCA), Cholangiocarcinoma (CHOL), Glioblastoma multiforme (GBM), Head and Neck squamous cell carcinoma (HNSC), Lung squamous cell carcinoma (LUSC), Stomach adenocarcinoma (STAD), GPX4 was up-regulated in the other 12 cancers (Fig. 15A). Up-regulation of the displayed GPX4 survival curve is associated with a better prognosis (Fig. 15B), consistent with our previous results (Fig. 3). We predicted a total of 10 upstream miRNAs that may bind to GPX4. Figure 15C shows the miRNAs-GPX4 regulatory network. According to the mechanism by which miRNA regulates target gene expression, there is a negative correlation between miRNA and GPX4. Figure 15D shows a significant negative correlation between hsa-miR-1287-5p and GPX4 expression. We also noticed that hsa-miR-1287-5p was significantly down-regulated in THCA (Fig. 15E) and had a better prognosis (Fig. 15F). The above findings indicate that hsa-miR-1287-5p may be the most potential regulatory miRNA for GPX4 in THCA. Through the starBase database, 79 lncRNAs were obtained. Further screening got one lncRNA (AL928654.4) upstream of hsa-miR-1287-5p. The expression of AL928654.4 was negatively correlated with the expression of hsa-miR-1287-5p (Fig. 15G). Unfortunately, we did not find a significant correlation between AL928654.4 and GPX4 (Fig. 15H). Nevertheless, the expression of AL928654.4 in tumor tissues was significantly increased, and the high expression of AL928654.4 indicated a better prognosis (Fig. 15I-J). Figure 15K shows the lncRNAs-hsa-miR-1287-5p regulatory network. Figure 15L offers the CERNA regulatory network composed of AL928654.4/hsa-miR-1287-5p/GPX4.

Discussion

Numerous studies have focused on finding one or more genes to build tumor risk assessment models in recent years. At the same time, the impact of ferroptosis on tumor biology is also a hot topic today[13, 14]. Those bring us a new idea for FRGs combinations to construct a reasonable and practical risk assessment model.

This study obtained complete RNA sequence data of 246 FRGs in thyroid cancer through TCGA and the FerrDb databases. Next, we analyzed the relationship between FRGs and the patient's OS and the difference in FRGs expression between cancerous and normal tissues. According to the association between DE-FRGs and the patient's OS, we obtained 14 PR-DE-FRGs. Motivated by this, a risk assessment model with superior performance was proposed. Since this study data only comprised samples from a single cohort of TCGA, we randomly divided the samples into the training set, test set, and whole set to achieve the effect of mutual multi-set verification of the results. We split each training/test/whole set of samples into high-risk and low-risk groups based on the median value of the risk score. We found that low risk is significantly associated with better OS. Surprisingly, when we drew the ROC curve to verify this result, the 1-year, 2-year, and 3-year AUC of the training/test/whole sets were more outstanding than 0.9.

The AUC of the risk score is the largest among all prognostic-related factors. It proves that the risk assessment model has exceptional predictive value for the prognosis of patients.

Previous studies have demonstrated that ferroptosis was involved in cancer development and severely impacted the prognosis of cancer patients. For example, Liang et al. and Tang et al. found that the prognosis model constructed by ferroptosis-related genes in the high-risk group of Hepatocellular Carcinoma had significantly more inferior OS than the low-risk group [38, 39]. Mou et al. found that when genes' expression level that promotes ferroptosis is low, the OS of clear cell renal cell carcinoma is also inferior [40]. Similar results were found in lung adenocarcinoma by Yao et al. [41] and Tian et al. [42]. However, few studies have established a risk assessment model based on ferroptosis-related genes to predict the prognosis of THCA. This study used up to 14 genes to build models to minimize the risk of overfitting screening features, accurately predicting THCA patients' prognosis.

These 14 genes can be roughly divided into oxidative metabolism (PGD, DPP4, CDKN2A, MIOX, ARNTL, BID, CAPG, BID), lipid metabolism (ANGPTL7, TXNRD1, GPX4, SRXN1, DRD4), iron metabolism (TFRC, ISCU, DRD4) three categories [17][13] [43][44][45][22]. In the risk assessment model, eight genes (ANGPTL7, DRD4, SRXN1, TXNRD1, CDKN2A, MIOX, PGD, and TFRC) were identified as prognostic risk factors ($HR > 1$), while the remaining six genes were opposite (CAPG, GPX4, ARNTL, ISCU, BID and DPP4) ($HR < 1$). PGD reduced NADP to NADPH, preventing erastin-induced ferroptosis in Calu-1 cells [13]. Xie et al. found that high expression of DPP4 promotes plasma-membrane-associated DPP4-dependent lipid peroxidation, which ultimately leads to ferroptosis [46]. P53, through the sequestration of DPP4, achieves anti-ferroptotic function [47]. ARF was initially identified as an alternative transcript of the Ink4a tumor suppressor locus that encodes p16INK4a, an inhibitor of CDKN2A [48]. Combination of ARF induction and ROS treatment-induced ferroptotic cell death and Knockdown of endogenous ARF protected cells from ROS-induced cell death [49]. MIOX is a proximal tubular enzyme that promotes ROC production when overexpressed, thereby promoting ferroptosis [50]. ARNTL in the blockade of ferroptotic cancer cell death through control of the EGLN2-HIF1A pathway [44]. CAPG may inhibit ferroptosis by regulating GSH [51]. For lipid metabolism, SRXN1 and TXNRD1's expression were up-regulated during ferroptosis induced by erastin or RSL3 [22]. In contrast, ANGPTL7's expression was downregulated during ferroptosis caused by erastin or RSL3[22]. Those results suggested that the expression of these three genes may be related to cancer ferroptosis, but the specific mechanism of action is still unclear. According to reports, TXNRD1 can enhance the cytotoxicity of Lysine oxidase, which can activate ferroptosis [52]; this suggests that TXNRD1 may promote ferroptosis. The transcriptional activation of the SRXN1 gene is strongly responsive to oxidative stress [53], and SRXN1 plays a vital role in removing excess ROS [54]. The study of Kim et al. found that SRNX1 may be related to tumor progression [54], consistent with our research that SRNX1 is a prognostic risk factor for thyroid cancer. Song et al. also observed a higher expression level of ANGPTL7 in the esophageal squamous cell carcinoma study, and it was used as a prognostic risk gene [55]. Our research results are consistent with these studies. GPX4 is the phospholipid hydroperoxidase that inhibits ferroptosis because it can convert lipid hydroperoxides into non-toxic lipid alcohols[39]. BID is the mediator of mitochondrial dysfunction downstream of glutathione (GSH) depletion in the model of oxytosis [56]. For iron metabolism, dopamine reduced erastin-induced ferrous

iron accumulation, glutathione depletion, and malondialdehyde production [57]. This is why DRD4 may inhibit ferroptosis. Transferrin is an iron carrier protein in serum that can be transported into the cell via receptor-mediated endocytosis, RNAi of transferrin receptor (TFRC) inhibited ferroptosis [58]. Transferrin can only interact with TFRC and be transported into the cell when loaded with iron [58]. Overexpression of ISCU (Iron-sulfur cluster assembly enzyme, a mitochondrial protein) significantly attenuated Dihydroartemisinin induced ferroptosis by regulating iron metabolism, rescuing the mitochondrial function, and increasing the level of GSH[43]. Oncogenic mutations in RAS family members (included KRAS, HRAS, and NRAS) are common in 30% of human tumors[59]. Many studies have shown that NRSA gene mutations may be associated with the occurrence of ferroptosis [13, 25]. Dietrich et al. proved that NRAS is overexpressed in hepatocellular carcinoma and is a biomarker of poor patient prognosis and found that it contributes to Sorafenib resistance in hepatocellular carcinoma [60]. Many studies have proven that NRAS mutations in the thyroid are important for cancer, but the relationship between NRAS mutations and patient prognosis has not yet been determined [61, 62]. We confirmed a significant association between NRAS and risk score, but no significant association was found between NRAS and OS in patients with thyroid cancer.

GPX4 is an essential target for studying the mechanism of ferroptosis in cancer and a promising strategy for exploring potential treatments for cancer[63]. At the same time, more and more shreds of evidence showed that miR-1287-5p plays a substantial carcinogenic or anti-cancer effect in human tumors, such as Cervical Cancer[64], Breast cancer[65], THCA[66], and so on. Many previous studies have demonstrated that the miR-1287-5p/GPX4 axis plays a vital role in regulating the apoptosis and ferroptosis of cancer cells[67][68]. We verified the AL928654.4/ miR-1287-5p/ GPX4 regulatory axis, and the survival curve revealed the high expression of GPX4, the low expression of miR-1287-5p, and the increased expression of AL928654.4, and the better prognosis is relevant. AL928654.4 is significantly negatively correlated with miR-1287-5p and similar to the miR-1287-5p with GPX4. The above results may explain the regulation mode of GPX4 in the prognosis of THCA. The study by Wang et al. showed similar results to ours[68]. We have not found previous studies on AL928654.4, which led to our lack of sufficient evidence, but this also provides us with a new idea for studying the regulatory role of AL928654.4/ miR-1287-5p/ GPX4 in THCA.

Although ferroptosis has become a research focus, there are few studies on ferroptosis from thyroid cancer. Moreover, investigating the biological processes and pathways of thyroid cancer ferroptosis is even more limited. Therefore, based on 14 PR-DE-FRGs, GSEA was performed. Similar to expectations, Adipocytokine signaling pathway, Fatty acid metabolism, Hedgehog signaling pathway, mTOR signaling pathway, regulation of autophagy enriched, and Wnt signaling pathway in the high-risk group are all closely related to ferroptosis. Adipocytokine signaling pathway, Fatty acid metabolism can affect ferroptosis by regulating lipid metabolism [69]. Research by Yang et al. found that the Hedgehog signaling pathway, regulation of autophagy enriched, and Wnt signaling pathway play a vital role in the cancer stem cell process [70]. FZD7 is a classic Wnt receptor that can alter GSH metabolism and protect ovarian cancer from oxidative stress [71]. Studies have shown that promoting the mTOR signaling pathway can induce ferroptosis in trophoblast cells [72].

Since we found in GO analysis that immune-related biological processes and pathways were enriched, we further analyzed the impact of ferroptosis on immunomodulatory. Surprisingly, except for B cells, CD8 + T cells, and Tfh, the scores of the other 13 immune cells and 13 immune functions in the high-risk group were significantly lower than those in the low-risk group. The study has speculated that ferroptotic cells will release distinct signals, including lipid mediators, which will attract APC and other immune cells to the site of ferroptotically dying cells[47]. A study found that ferroptotic cancer cells are efficiently engulfed by macrophages in vitro[73]. In addition, oxidized lipids and lipid droplets regulate the anti-tumor immune response [47]. Lipid bodies containing oxidatively truncated lipids block antigen cross-presentation by dendritic cells in cancer, leading to anti-tumor immunodeficiency[74]. NK cells play a central role in cancer immune insurance by killing cancer cells. [75]. Many studies have shown that tumor-associated neutrophils[76] and lymphocytic[77] are associated with a better prognosis of tumors, but the specific mechanism is still controversial. In addition, 13 immune-related functions in the high-risk group were significantly impaired. Taken together, these results demonstrate that the poor prognosis of thyroid cancer patients in our high-risk group is probably related to low anti-tumor immunity. Stem cell characteristics reveal new drug targets for anti-cancer therapy[12][11]. Cancer stem cells (CSCs) usually show high iron levels in the cells [70]. The Ferroptosis pathway can selectively induce CSC death [78]. We found that the low-risk and N0 stage groups had higher mDNAsi scores. This conclusion indicated that cancerous cells in our low-risk group are less aggressive and can invade surrounding tissues, which is also reflected in the study of glioblastoma. [79].

Another advantage of our research is combining the risk assessment model with clinical applications. For example, the risk score is significantly correlated with the expression of three IPS, five ICIs, and the IC50 of five chemotherapeutics. These results all implied that our risk scoring model could effectively predict the chemosensitivity of patients receiving these five chemotherapy agents and predict the effect of immunotherapy. Clinically, an operational decision support tool was needed. Therefore, we used factors that may affect the prognosis (age, stage, gender, and risk score) to build a Nomogram for predicting thyroid cancer's 1, 3, and 5-year overall survival.

There are several limitations to this study. First, we only obtained the differential expression results of the 14 PR-DE-FRGs used to construct the model by analyzing the sharing data. However, due to the difficulty of collecting clinical specimens and the limitation of experimental conditions, we, unfortunately, failed to verify these gene expression differences through experiments. To make up for this limitation as much as possible, we used images of PR-DE-FRGs protein expression between thyroid cancerous and normal tissues obtained by immunohistochemical staining based on clinical specimens to verify our results. Second, the conclusion of the relationship between the risk score and the tumor microenvironment may need to be verified by experiments with bulk samples and a substantial workload. Third, we have not obtained transcriptome data with survival information in other databases to validate our model. To make up for this shortcoming, we randomly divided the TCGA sample into three sets. Then we verified the model similar to the external set and finally achieved brilliant verification results.

Conclusions

In conclusion, this study used 14 PR-DE-FRGs to build a risk assessment model for thyroid cancer. And through multiple analytic approaches, it has been verified that the model has a brilliant predictive ability for OS in patients with thyroid cancer. In discussing the mechanism, we found that the risk assessment model has a significant relationship with the TME, TMB, NRAS mutation, and stemness features.

AL928654.4/ miR-1287-5p/ GPX4 regulation axis explains its regulation mode in the progression and prognosis of THCA. At the same time, the model has proved to have clinically predictive value in drug sensitivity, IPS, ICIs. Nomogram has good predictive efficacy for 1-, 3-, and 5-year OS in patients with thyroid cancer.

Abbreviations

APC, antigen-presenting cells; AUC, area Under Curve; BRCA, Breast invasive carcinoma; CHOL, cholangiocarcinoma; CSCs, cancer stem cells; DE-FRGs, differentially expressed ferroptosis-related genes; FRGs, ferroptosis-related genes; GSEA, gene set enrichment analysis ; GEO, gene expression omnibus; GO, gene ontology; GDSC, genomics of drug sensitivity in cancer; GEPIA, gene expression profiling interactive analysis; GBM, glioblastoma multiforme; GSH, glutathione; HPA, human protein atlas; HNSCC, head and neck squamous cell carcinoma; IHC, Immunohistochemical; IPS, immunophenoscore; ICIs, immune checkpoint inhibitors; IC50, half-maximal inhibitory concentrations; INF, Interferon; KEGG, kyoto encyclopedia of genes and genomes; LUSC, lung squamous cell carcinoma; MHC, major histocompatibility complex; NES, normalized enrichment score; NCCN, national comprehensive cancer network; NK, natural killer; OS, overall survival; OCLR, one-class logistic regression; PR-DE-FRGs, prognostic-related differentially expressed ferroptosis-related genes; PCA, principal component analysis; ROS, reactive oxygen species; STAD, stomach adenocarcinoma; TME, tumor microenvironment ; THCA, thyroid cancer; TCIA, the cancer immunome atlas;

Declarations

Acknowledgements

The authors sincerely thank all the participants and the supporting from the Research Fund Project of Jiangxi Provincial Department of Education (180017).

Funding

Research Fund Project of Jiangxi Provincial Department of Education [180017].

Availability of data and materials

The datasets analyzed in this study came from databases shared publicly. Data can be obtained from FerrDb <http://www.zhounan.org/ferrdb>, GDSC <https://www.cancerrxgene.org/>, GEO <https://www.ncbi.nlm.nih.gov/geo/>, GEPIA <http://gepia.cancer-pku.cn/>, HPA <http://www.proteinatlas.org/>, mirbase <https://www.mirbase.org/>, StarBase <http://starbase.sysu.edu.cn/>, TCGA <https://cancergenome.nih.gov/>, TCIA <https://www.cancerimagingarchive.net/>, UCSC Xena database <https://xena.ucsc.edu/>.

Ethics approval and consent to participate

This bioinformatics study does not involve human subjects based on publicly obtained data from the TCGA and GEO databases. Therefore, it does not involve additional intervention on the human body and the collection of specimens, and it will not cause unnecessary damage and additional costs to potential research subjects. Therefore, this study does not involve the ethics approval process of the local ethics committee.

Competing interests

The author reports no conflicts of interest in this work.

Consent for publication

Not applicable

Authors' information

Affiliations

¹Department of Otolaryngology Head and Neck Surgery, The First Affiliated Hospital of Nanchang University, Nanchang, Jiangxi Province, China

²The First Clinical Medical College of Nanchang University, Nanchang, Jiangxi Province, China

³School of Stomatology, Nanchang University, Nanchang, Jiangxi Province, China

Authors' Contributions

X. F. and Z. Z. designed the research; X. F. prepared the figures and drafted the manuscript; X. F. analyzed the data; X. F., Y. F., and L. Z. contributed analytic tools and finalized the manuscript. X. F., Y.Z. and F. X. participated in the writing of the manuscript. All authors have read and approved the final manuscript.

Corresponding author

Correspondence to Zhi-yuan Zhang

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020 : GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. 2021;71:209–49.
2. Lortet-Tieulent J, Franceschi S, Dal Maso L, Vaccarella S. Thyroid cancer “epidemic” also occurs in low- and middle-income countries. *Int J Cancer*. 2019;144:2082–7.
3. Hwang T, Franklin JM. Increases in Thyroid Cancer Incidence and Mortality. 2017;318:389–91.
4. Apostolides AY. *Cancer Epidemiology and Prevention. J. Occup. Environ. Med.* 1983. p. 647.
5. Brito JP, Morris JC, Montori VM. Thyroid cancer : zealous imaging has increased detection and treatment of low risk tumours. 2013;4706:1–6.
6. Cabanillas ME, McFadden DG, Durante C. Thyroid cancer. *Lancet*. Elsevier Ltd; 2016;388:2783–95.
7. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid*. 2016;26:1–133.
8. Barney BM, Hitchcock YJ, Sharma P, Shrieve DC, Tward JD. Overall and cause-specific survival for patients undergoing lobectomy, near-total, or total thyroidectomy for differentiated thyroid cancer. *Head Neck*. United States; 2011;33:645–9.
9. Nixon IJ, Ganly I, Patel SG, Palmer FL, Witcher MM, Tuttle RM, et al. Thyroid lobectomy for treatment of well differentiated intrathyroid malignancy. *Surgery*. United States; 2012;151:571–9.
10. Haigh PI, Urbach DR, Rotstein LE. Extent of thyroidectomy is not a major determinant of survival in low- or high-risk papillary thyroid cancer. *Ann Surg Oncol*. United States; 2005;12:81–9.
11. Friedmann-Morvinski D, Verma IM. Dedifferentiation and reprogramming: origins of cancer stem cells. *EMBO Rep*. 2014;15:244–53.
12. Visvader JE, Lindeman GJ. Cancer stem cells: current status and evolving complexities. *Cell Stem Cell*. United States; 2012;10:717–28.
13. Dixon SJ, Lemberg KM, Lamprecht MR, Skouta R, Zaitsev EM, Gleason CE, et al. Ferroptosis: An iron-dependent form of nonapoptotic cell death. *Cell*. Elsevier Inc.; 2012;149:1060–72.
14. Yang WS, Stockwell BR. Ferroptosis: Death by Lipid Peroxidation. *Trends Cell Biol*. 2016;26:165–76.
15. Fearnhead HO, Vandenabeele P, Berghe T Vanden. How do we fit ferroptosis in the family of regulated cell death? *Cell Death Differ*. Nature Publishing Group; 2017;24:1991–8.
16. Mou Y, Wang J, Wu J, He D, Zhang C, Duan C, et al. Ferroptosis, a new form of cell death: Opportunities and challenges in cancer. *J Hematol Oncol. Journal of Hematology & Oncology*;

- 2019;12:1–16.
17. Hassannia B, Vandenabeele P, Vanden Berghe T. Targeting Ferroptosis to Iron Out Cancer. *Cancer Cell*. United States; 2019;35:830–49.
 18. Huang Y, Dai Z, Barbacioru C, Sadée W. Cystine-glutamate transporter SLC7A11 in cancer chemosensitivity and chemoresistance. *Cancer Res*. United States; 2005;65:7446–54.
 19. Dolma S, Lessnick SL, Hahn WC, Stockwell BR. Identification of genotype-selective antitumor agents using synthetic lethal chemical screening in engineered human tumor cells. *Cancer Cell*. United States; 2003;3:285–96.
 20. Sakitama K, Ozawa Y, Aoto N, Tomita H, Ishikawa M. Effects of a new centrally acting muscle relaxant, NK433 (lanperisone hydrochloride) on spinal reflexes. *Eur J Pharmacol*. Netherlands; 1997;337:175–87.
 21. Shaw AT, Winslow MM, Magendantz M, Ouyang C, Dowdle J, Subramanian A, et al. Selective killing of K-ras mutant cancer cells by small molecule inducers of oxidative stress. *Proc Natl Acad Sci U S A*. 2011;108:8773–8.
 22. Yang WS, SriRamaratnam R, Welsch ME, Shimada K, Skouta R, Viswanathan VS, et al. Regulation of ferroptotic cancer cell death by GPX4. *Cell*. 2014;156:317–31.
 23. Stockwell BR, Jiang X, Gu W. Emerging Mechanisms and Disease Relevance of Ferroptosis. *Trends Cell Biol*. 2020;30:478–90.
 24. Matsushita M, Freigang S, Schneider C, Conrad M, Bornkamm GW, Kopf M. T cell lipid peroxidation induces ferroptosis and prevents immunity to infection. *J Exp Med*. 2015;212:555–68.
 25. Lin R, Fogarty CE, Ma B, Li H, Ni G, Liu X, et al. Identification of ferroptosis genes in immune infiltration and prognosis in thyroid papillary carcinoma using network analysis. *BMC Genomics*. *BMC Genomics*; 2021;22:1–16.
 26. Freedman AN, Seminara D, Gail MH, Hartge P, Colditz GA, Ballard-Barbash R, et al. Cancer risk prediction models: A workshop on development, evaluation, and application. *J Natl Cancer Inst*. 2005;97:715–23.
 27. Fan X, Ou Y, Liu H, Zhan L, Zhu X, Cheng M, et al. A Ferroptosis-Related Prognostic Signature Based on Antitumor Immunity and Tumor Protein p53 Mutation Exploration for Guiding Treatment in Patients With Head and Neck Squamous Cell Carcinoma. *Front Genet*. 2021;12:732211.
 28. Liu C, Min L, Kuang J, Zhu C, Qiu X-Y, Zhu L. Bioinformatic Identification of miR-622 Key Target Genes and Experimental Validation of the miR-622-RNF8 Axis in Breast Cancer. *Front Oncol*. 2019;9:1114.
 29. Rydén M, Englund M, Ali N. ProteoMill: Efficient network-based functional analysis portal for proteomics data. *Bioinformatics*. Oxford University Press; 2021;37:3491–3.
 30. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102:15545–50.

31. Panja S, Hayati S, Epsi NJ, Parrott JS, Mitrofanova A. Integrative (epi) Genomic Analysis to Predict Response to Androgen-Deprivation Therapy in Prostate Cancer. *EBioMedicine*. 2018;31:110–21.
32. Huo J, Wu L, Zang Y. Development and validation of a CTNNB1-associated metabolic prognostic model for hepatocellular carcinoma. *J Cell Mol Med*. 2020/12/09. John Wiley and Sons Inc.; 2021;25:1151–65.
33. Mai H, Xie H, Luo M, Hou J, Chen J, Hou J, et al. Implications of Stemness Features in 1059 Hepatocellular Carcinoma Patients from Five Cohorts: Prognosis, Treatment Response, and Identification of Potential Compounds. *Cancers (Basel)*. 2022;14.
34. Hu X, Wu L, Liu B, Chen K. Immune Infiltration Subtypes Characterization and Identification of Prognosis-Related lncRNAs in Adenocarcinoma of the Esophagogastric Junction. *Front Immunol*. Frontiers Media S.A.; 2021;12:651056.
35. Mata DA, Benhamida JK, Lin AL, Vanderbilt CM, Yang S-R, Villafania LB, et al. Genetic and epigenetic landscape of IDH-wildtype glioblastomas with FGFR3-TACC3 fusions. *Acta Neuropathol Commun*. BioMed Central; 2020;8:186.
36. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017;45:W98–102.
37. Charoentong P, Angelova M, Charoentong P, Finotello F, Angelova M, Mayer C, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade Resource Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkp. *CellReports*. ElsevierCompany.; 2017;18:248–62.
38. Tang B, Zhu J, Li J, Fan K, Gao Y, Cheng S, et al. The ferroptosis and iron-metabolism signature robustly predicts clinical diagnosis, prognosis and immune microenvironment for hepatocellular carcinoma. *Cell Commun Signal*. 2020;18:174.
39. Liang J-Y, Wang D-S, Lin H-C, Chen X-X, Yang H, Zheng Y, et al. A Novel Ferroptosis-related Gene Signature for Overall Survival Prediction in Patients with Hepatocellular Carcinoma. *Int J Biol Sci*. 2020;16:2430–41.
40. Mou Y, Wu J, Zhang Y, Abdihamid O, Duan C, Li B. Low expression of ferritinophagy-related NCOA4 gene in relation to unfavorable outcome and defective immune cells infiltration in clear cell renal carcinoma. *BMC Cancer*. 2021;21:18.
41. Yao J, Chen X, Liu X, Li R, Zhou X, Qu Y. Characterization of a ferroptosis and iron-metabolism related lncRNA signature in lung adenocarcinoma. *Cancer Cell Int*. 2021;21:340.
42. Tian Q, Zhou Y, Zhu L, Gao H, Yang J. Development and Validation of a Ferroptosis-Related Gene Signature for Overall Survival Prediction in Lung Adenocarcinoma. *Front cell Dev Biol*. 2021;9:684259.
43. Du J, Wang T, Li Y, Zhou Y, Wang X, Yu X, et al. DHA inhibits proliferation and induces ferroptosis of leukemia cells through autophagy dependent degradation of ferritin. *Free Radic Biol Med*. United States; 2019;131:356–69.

44. Yang M, Chen P, Liu J, Zhu S, Kroemer G, Klionsky DJ, et al. Clockophagy is a novel selective autophagy process favoring ferroptosis. *Sci Adv.* 2019;5:eaaw2238.
45. Shau H, Butterfield LH, Chiu R, Kim A. Cloning and sequence analysis of candidate human natural killer-enhancing factor genes. *Immunogenetics.* United States; 1994;40:129–34.
46. Xie Y, Zhu S, Song X, Sun X, Fan Y, Liu J, et al. The Tumor Suppressor p53 Limits Ferroptosis by Blocking DPP4 Activity. *Cell Rep.* ElsevierCompany.; 2017;20:1692–704.
47. Friedmann Angeli JP, Krysko D V., Conrad M. Ferroptosis at the crossroads of cancer-acquired drug resistance and immune evasion. *Nat Rev Cancer.* Springer US; 2019;19:405–14.
48. Sherr CJ. Divorcing ARF and p53: an unsettled case. *Nat Rev Cancer.* England; 2006;6:663–73.
49. Chen D, Tavana O, Chu B, Erber L, Chen Y, Baer R, et al. NRF2 Is a Major Target of ARF in p53-Independent Tumor Suppression. *Mol Cell.* 2017;68:224–232.e4.
50. Deng F, Sharma I, Dai Y, Yang M, Kanwar YS. Myo-inositol oxygenase expression profile modulates pathogenic ferroptosis in the renal proximal tubule. *J Clin Invest.* 2019;129:5033–49.
51. Zhang X, Du L, Qiao Y, Zhang X, Zheng W, Wu Q, et al. Ferroptosis is governed by differential regulation of transcription in liver cancer. *Redox Biol.* 2019;24:101211.
52. Chepikova OE, Malin D, Strekalova E, Lukasheva E V, Zamyatnin AAJ, Cryns VL. Lysine oxidase exposes a dependency on the thioredoxin antioxidant pathway in triple-negative breast cancer cells. *Breast Cancer Res Treat.* 2020;183:549–64.
53. Kim H, Jung Y, Shin BS, Kim H, Song H, Bae SH, et al. Redox regulation of lipopolysaccharide-mediated sulfiredoxin induction, which depends on both AP-1 and Nrf2. *J Biol Chem.* 2010;285:34419–28.
54. Kim T, Li D, Terasaka T, Nicholas DA, Knight VS, Yang JJ, et al. SRXN1 Is Necessary for Resolution of GnRH-Induced Oxidative Stress and Induction of Gonadotropin Gene Expression. *Endocrinology.* 2019;160:2543–55.
55. Song J, Liu Y, Guan X, Zhang X, Yu W, Li Q. A Novel Ferroptosis-Related Biomarker Signature to Predict Overall Survival of Esophageal Squamous Cell Carcinoma. *Front Mol Biosci.* 2021;8:675193.
56. Neitemeier S, Jelinek A, Laino V, Hoffmann L, Eisenbach I, Eying R, et al. BID links ferroptosis to mitochondrial cell death pathways. *Redox Biol.* 2017;12:558–70.
57. Wang D, Peng Y, Xie Y, Zhou B, Sun X, Kang R, et al. Antiferroptotic activity of non-oxidative dopamine. *Biochem Biophys Res Commun.* United States; 2016;480:602–7.
58. Gao M, Monian P, Quadri N, Ramasamy R, Jiang X. Glutaminolysis and Transferrin Regulate Ferroptosis. *Mol Cell.* 2015;59:298–308.
59. Wilson CY, Tolias P. Recent advances in cancer drug discovery targeting RAS. *Drug Discov Today.* England; 2016;21:1915–9.
60. Dietrich P, Gaza A, Wormser L, Fritz V, Hellerbrand C, Bosserhoff AK. Neuroblastoma RAS Viral Oncogene Homolog (NRAS) Is a Novel Prognostic Marker and Contributes to Sorafenib Resistance in Hepatocellular Carcinoma. *Neoplasia.* 2019;21:257–68.

61. Huang M, Yan C, Xiao J, Wang T, Ling R. Relevance and clinicopathologic relationship of BRAF V600E, TERT and NRAS mutations for papillary thyroid carcinoma patients in Northwest China. *Diagn Pathol.* 2019;14:74.
62. Fakhruddin N, Jabbour M, Novy M, Tamim H, Bahmad H, Farhat F, et al. BRAF and NRAS Mutations in Papillary Thyroid Carcinoma and Concordance in BRAF Mutations Between Primary and Corresponding Lymph Node Metastases. *Sci Rep.* 2017;7:4666.
63. Viswanathan VS, Ryan MJ, Dhruv HD, Gill S, Eichhoff OM, Seashore-Ludlow B, et al. Dependency of a therapy-resistant state of cancer cells on a lipid peroxidase pathway. *Nature.* 2017;547:453–7.
64. Ji F, Du R, Chen T, Zhang M, Zhu Y, Luo X, et al. Circular RNA circSLC26A4 Accelerates Cervical Cancer Progression via miR-1287-5p/HOXA7 Axis. *Mol Ther Nucleic Acids.* 2020;19:413–20.
65. Schwarzenbacher D, Klec C, Pasculli B, Cerk S, Rinner B, Karbiener M, et al. MiR-1287-5p inhibits triple negative breast cancer growth by interaction with phosphoinositide 3-kinase CB, thereby sensitizing cells for PI3Kinase inhibitors. *Breast Cancer Res.* 2019;21:20.
66. Shao J, Xu Y, Li H, Chen L, Wang W, Shen D, et al. LMCD1 antisense RNA 1 (LMCD1-AS1) potentiates thyroid cancer cell growth and stemness via a positive feedback loop of LMCD1-AS1/miR-1287-5p/GLI2. *Ann Transl Med.* 2020;8:1508.
67. Zhang H-Y, Zhang B-W, Zhang Z-B, Deng Q-J. Circular RNA TTBK2 regulates cell proliferation, invasion and ferroptosis via miR-761/ITGB8 axis in glioma. *Eur Rev Med Pharmacol Sci. Italy;* 2020;24:2585–600.
68. Shanshan W, Hongying M, Jingjing F, Yiming Y, Yu R, Rui Y. CircDTL Functions as an Oncogene and Regulates Both Apoptosis and Ferroptosis in Non-small Cell Lung Cancer Cells. *Front Genet.* 2021;12:743505.
69. Von Roemeling CA, Copland JA. Targeting lipid metabolism for the treatment of anaplastic thyroid carcinoma. *Expert Opin Ther Targets.* 2016;20:159–66.
70. Yang Y, Li X, Wang T, Guo Q, Xi T, Zheng L. Emerging agents that target signaling pathways in cancer stem cells. *J Hematol Oncol.* 2020;13:60.
71. Wang Y, Zhao G, Condello S, Huang H, Cardenas H, Tanner EJ, et al. Frizzled-7 Identifies Platinum-Tolerant Ovarian Cancer Cells Susceptible to Ferroptosis. *Cancer Res.* 2021;81:384–99.
72. Han D, Jiang L, Gu X, Huang S, Pang J, Wu Y, et al. SIRT3 deficiency is resistant to autophagy-dependent ferroptosis by inhibiting the AMPK/mTOR pathway and promoting GPX4 levels. *J Cell Physiol. United States;* 2020;235:8839–51.
73. Klöditz K, Fadeel B. Three cell deaths and a funeral: macrophage clearance of cells undergoing distinct modes of cell death. *Cell death Discov.* 2019;5:65.
74. Veglia F, Tyurin VA, Mohammadyani D, Blasi M, Duperret EK, Donthireddy L, et al. Lipid bodies containing oxidatively truncated lipids block antigen cross-presentation by dendritic cells in cancer. *Nat Commun.* 2017;8:2122.
75. Guillerey C, Huntington ND, Smyth MJ. Targeting natural killer cells in cancer immunotherapy. *Nat Immunol. United States;* 2016;17:1025–36.

76. Ferrari SM, Fallahi P, Galdiero MR, Ruffilli I, Elia G, Ragusa F, et al. Immune and Inflammatory Cells in Thyroid Cancer Microenvironment. *Int J Mol Sci.* 2019;20.
77. Matsubayashi S, Kawai K, Matsumoto Y, Mukuta T, Morita T, Hirai K, et al. The correlation between papillary thyroid carcinoma and lymphocytic infiltration in the thyroid gland. *J Clin Endocrinol Metab.* United States; 1995;80:3421–4.
78. Elgendy SM, Alyammahi SK, Alhamad DW, Abdin SM, Omar HA. Ferroptosis: An emerging approach for targeting cancer stem cells and drug resistance. *Crit Rev Oncol Hematol.* Netherlands; 2020;155:103095.
79. Huang Z, Cheng L, Guryanova OA, Wu Q, Bao S. Cancer stem cells in glioblastoma–molecular signaling and therapeutic targeting. *Protein Cell.* 2010;1:638–55.

Figures

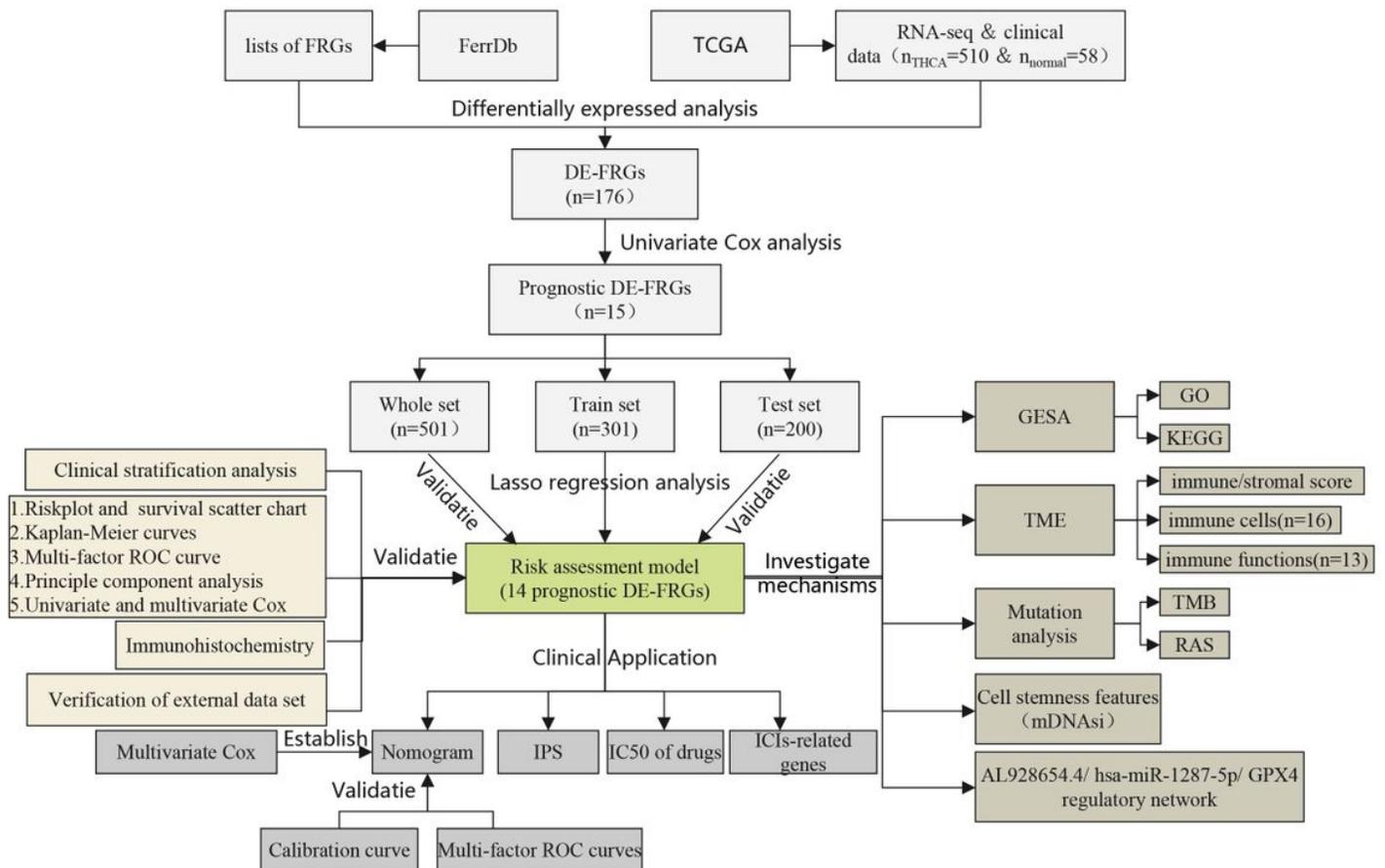


Figure 1

The workflow of the whole research.

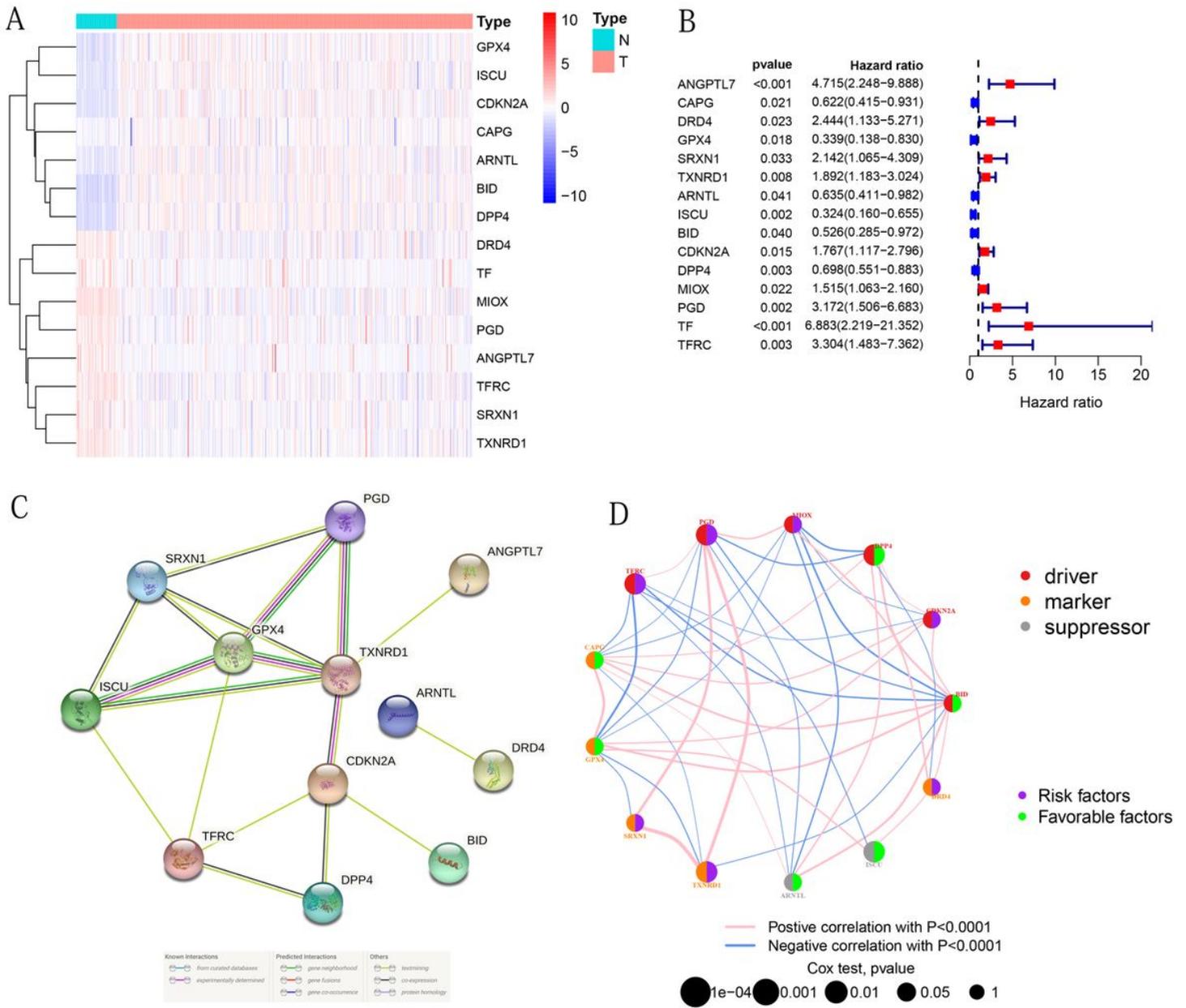


Figure 2

Identification of PR-DE-FRGs used to construct association networks. **(A)** The heatmap was used to show the expression profile of 15 PR-DE-FRGs. **(B)** The forest plot was used to indicate the univariate Cox regression analysis results based on 15 PR-DE-FRGs and survival. **(C)** PPI of 12 PR-DE-FRGs in model. The bottom of the figure shows that the different colored lines between nodes represent various sources of evidence. **(D)** Correlation network of 13 PR-DE-FRGs. The left and right halves of the circle represent the attributes of the gene and its influence on the prognosis, respectively.

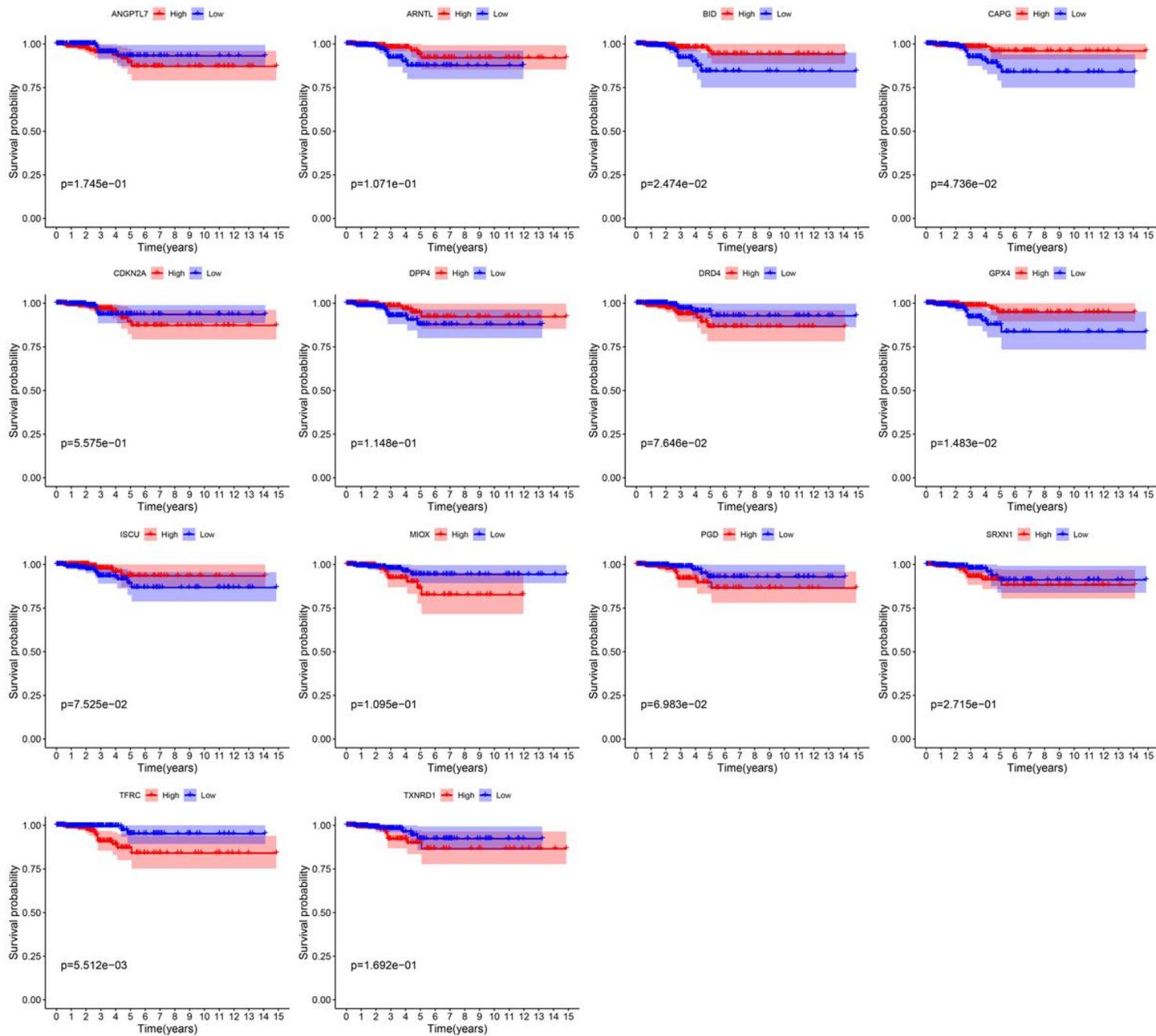


Figure 3

Survival curves of 14 PR-DE-FRGs were used to construct the model.

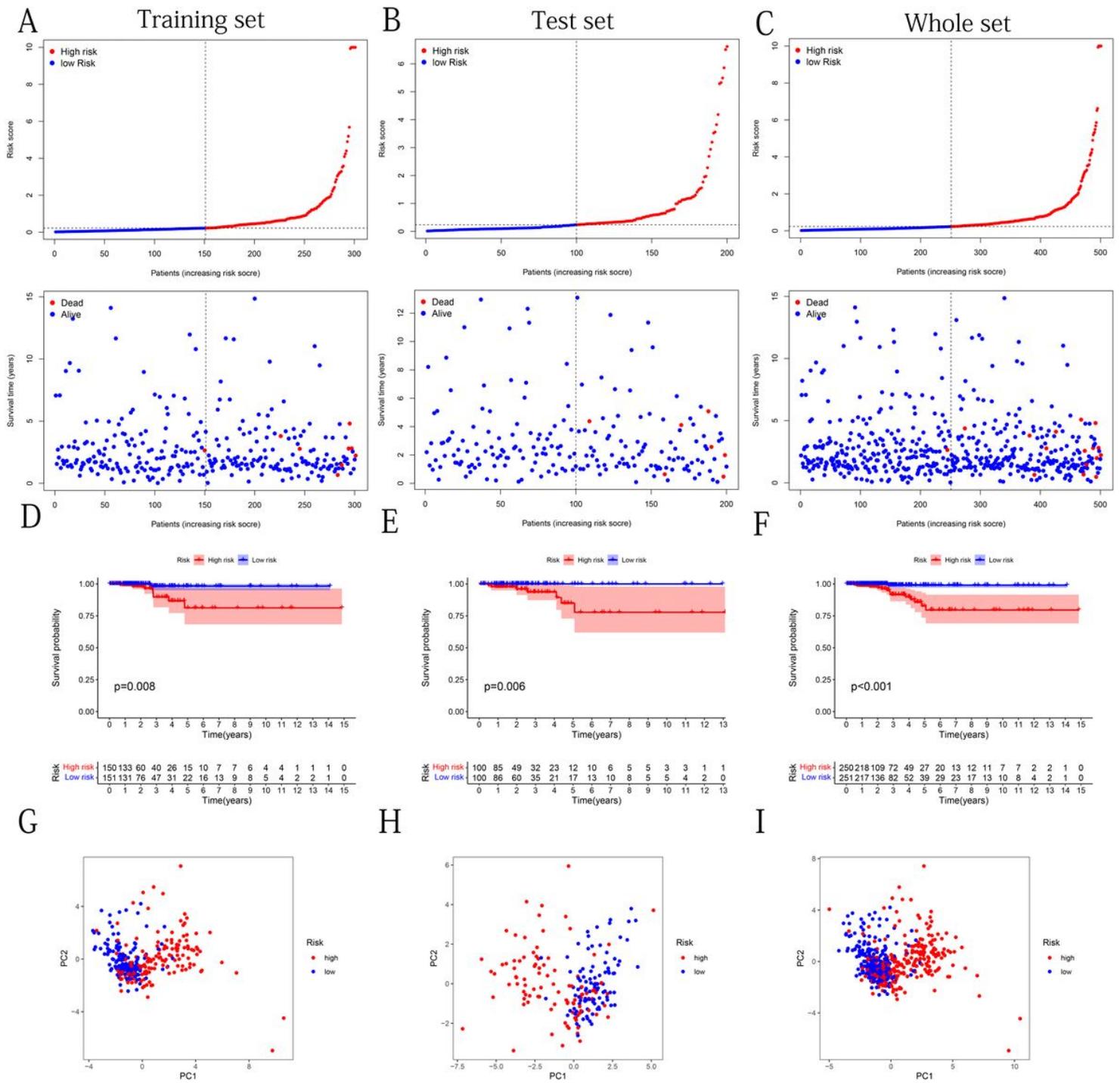


Figure 4

Riskplots, Kaplan-Meier survival curves, and PCA discrete trend charts based on the predictive model. **(A-C)** Riskplots of the training, test, and whole set, respectively. **(D-F)** Survival curves of the training/test/whole sets, respectively. **(G-I)** PCA discrete trend charts of the training/test/whole sets, respectively.

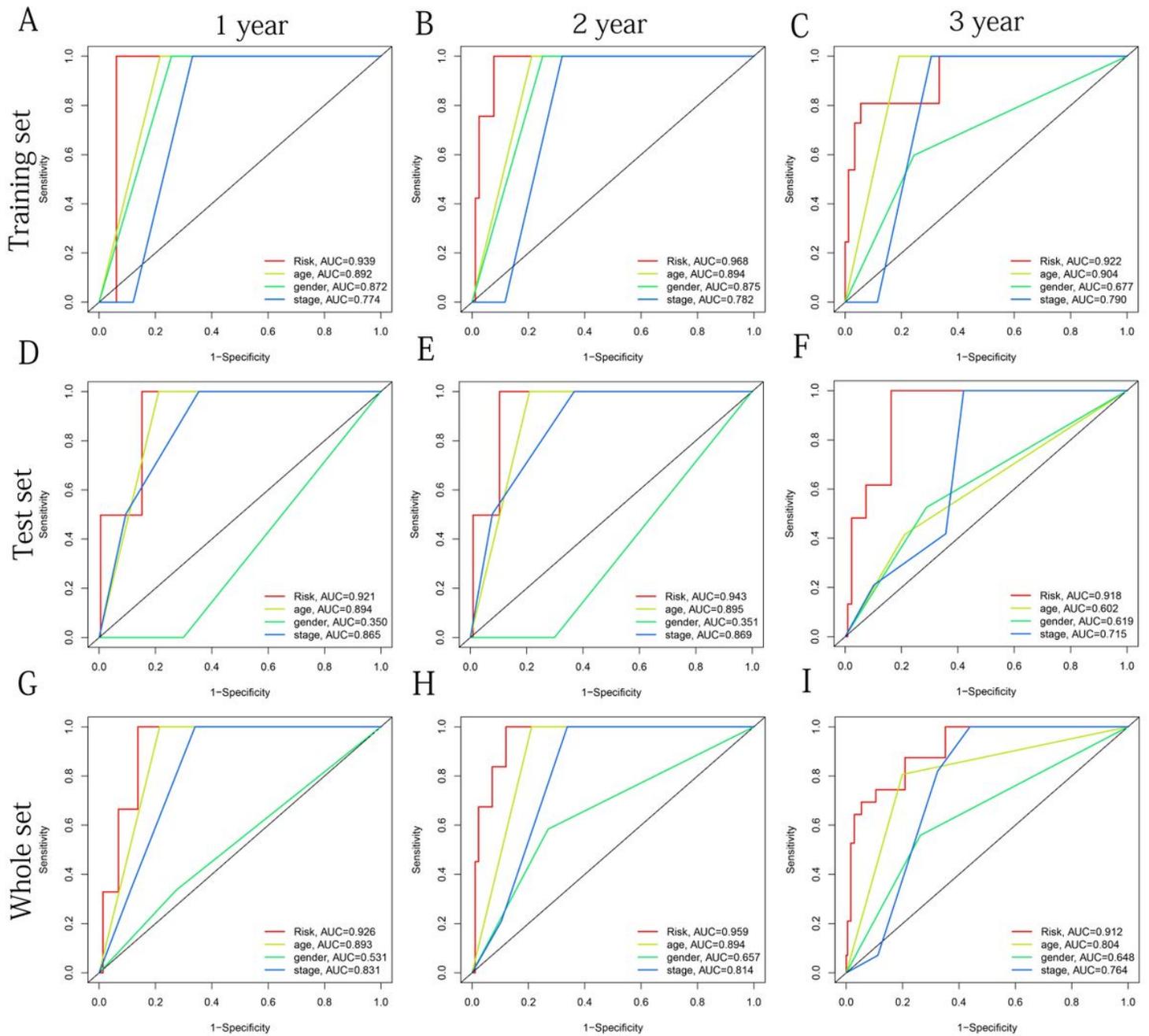


Figure 5

The multi-factor ROC curves confirm the best predictive performance of the model. **(A-C)** 1-, 2-, and 3-years of multi-factor ROC curves based on the training set, respectively. **(D-F)** 1-, 2-, and 3-years of multi-factor ROC curves based on the test set, respectively. **(G-H)** 1-, 2-, and 3-years of multi-factor ROC curves based on the whole set, respectively.

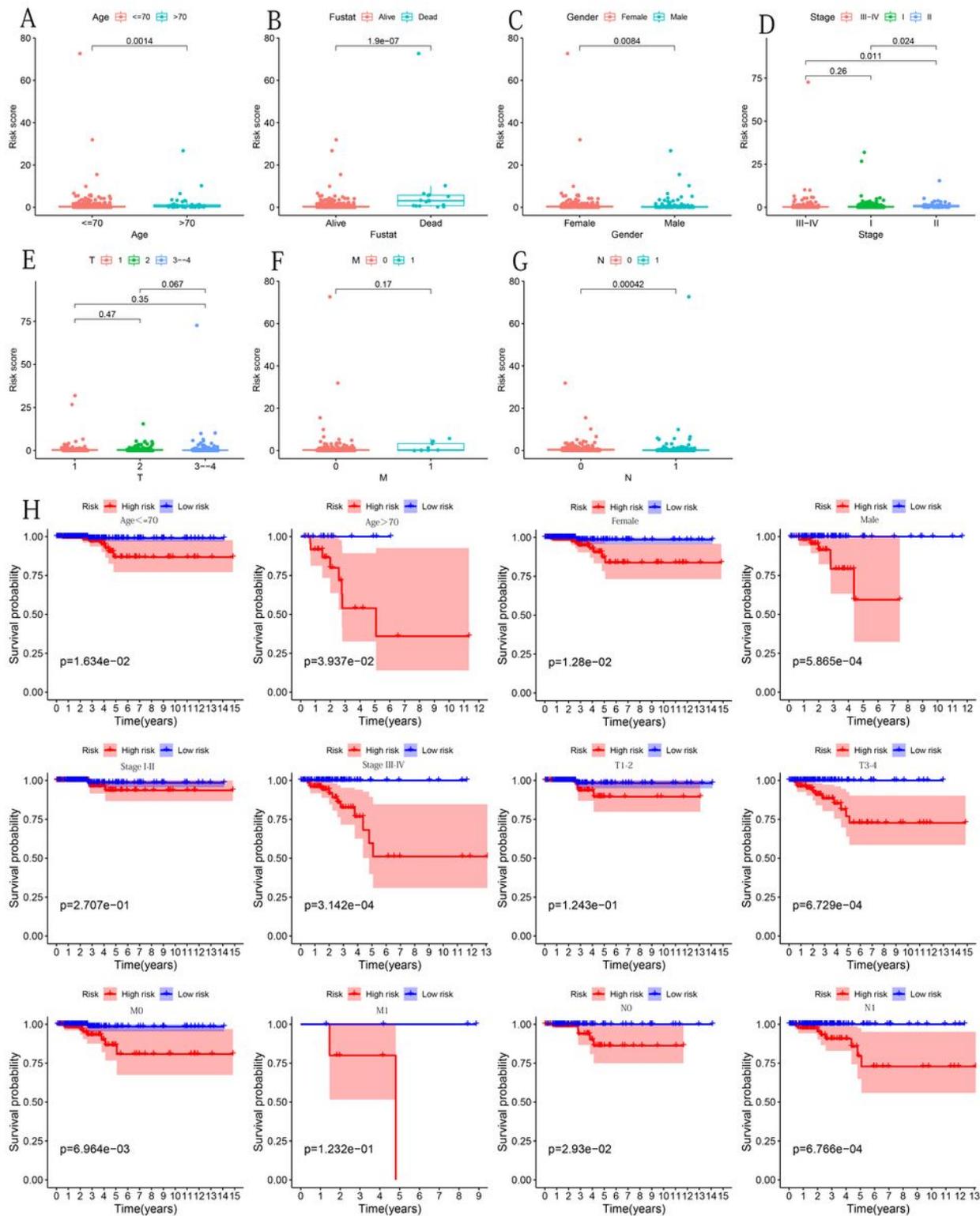


Figure 6

Stratification analysis of the risk assessment model. **(A)** Differences in risk score between patients with each subtype of each clinicopathological feature. **(B)** The risk assessment model still maintains the ability to predict the prognosis in patients with various subtypes of THCA.

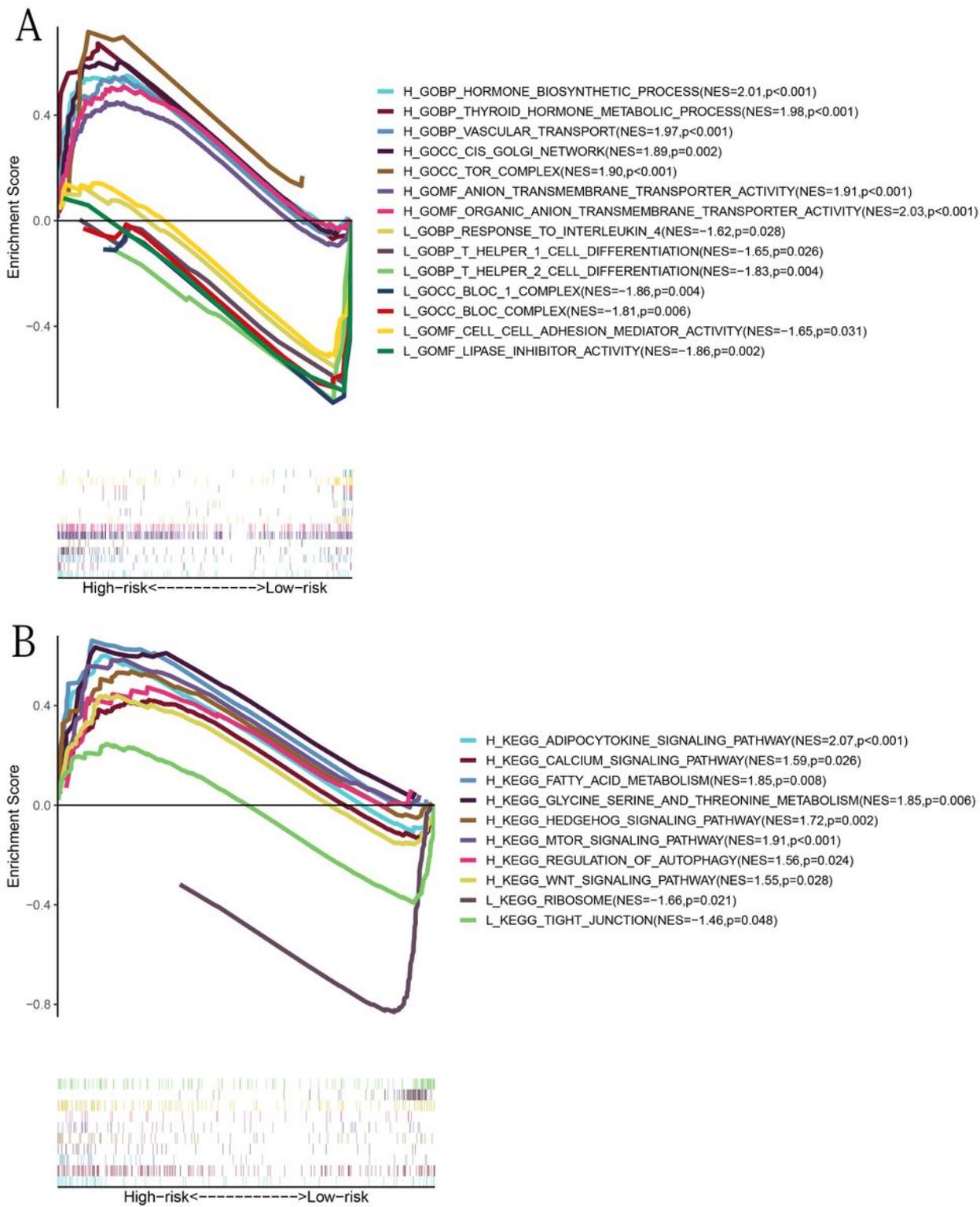


Figure 7

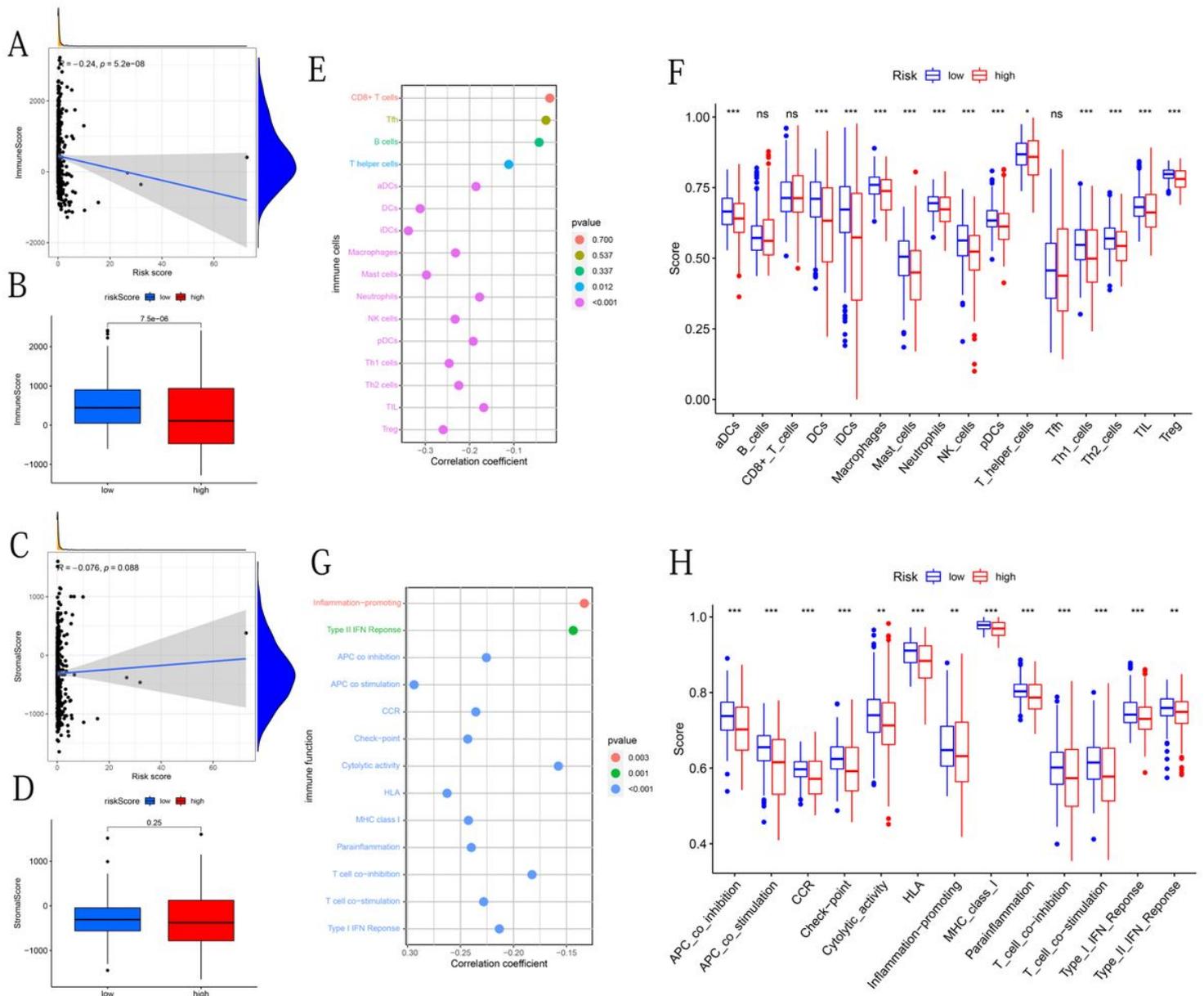


Figure 8

The correlation analysis between risk score and immune cells score, stroma cells score, 16 immune infiltrating cells, and 13 immune functions and their comparison between different risk groups. **(A, B)** Immune cells score. **(C, D)** Stroma cells score. **(E, F)** 16 immune infiltrating cells. **(G, H)** 13 immune functions. The symbol above the histogram shows the significance of the difference. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ns: no significance.

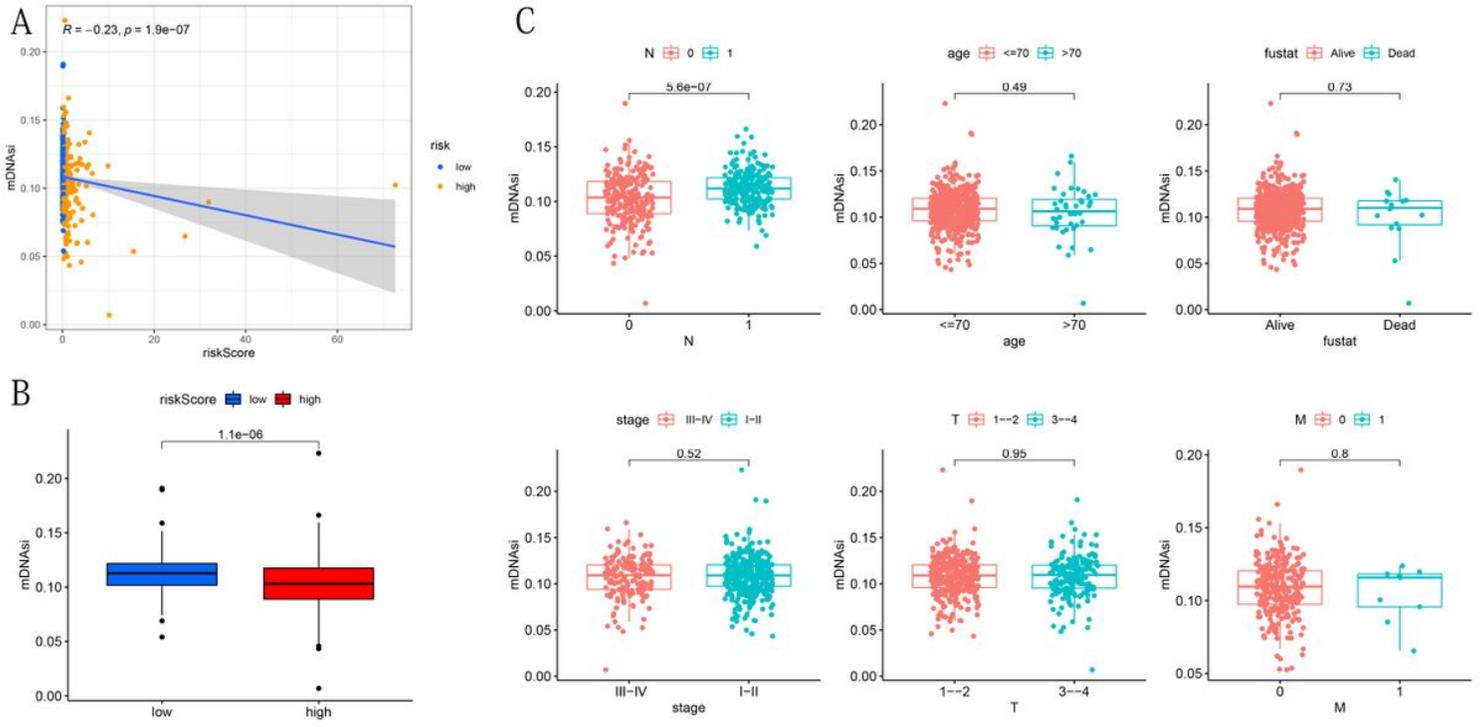


Figure 9

mDNAsi associated with the risk score and clinicopathological characteristics. **(A)** Correlation between mDNAsi and risk score. **(B)** Differences in mDNAsi between different risk groups. **(C)** Differences in mDNAsi between different subtypes of each clinicopathological feature.

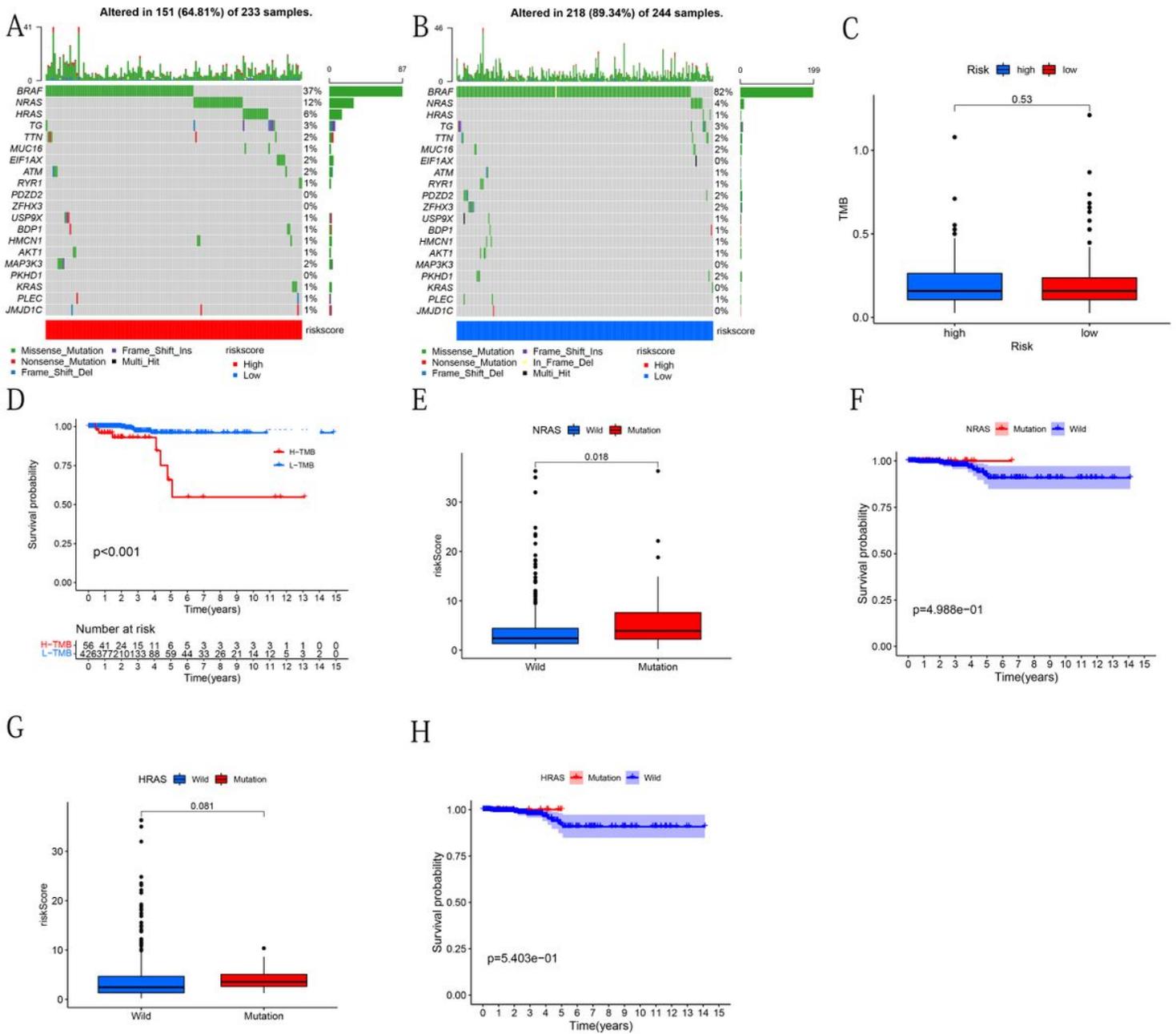


Figure 10

Mutation analysis. **(A, B)** The waterfall charts show the mutation distribution of patients in the high-risk group and low-risk group, respectively. The right panel of the waterfall plot shows the mutation frequency, and their mutation frequency sorts the genes. The upper and lower parts of the figure offer the TMB statistics and mutation types for each sample, respectively. **(C)** TMB difference between high-risk group and low-risk group. **(D)** Survival difference curve of patients between the high TMB and low TMB group. **(E, G)** The risk score difference between RAS (NRAS and HRAS) mutation group and RAS (NRAS and HRAS) wild group, respectively. **(F, H)** Survival difference curve of patients between RAS (NRAS and HRAS) mutation group and RAS (NRAS and HRAS) wild group respectively.

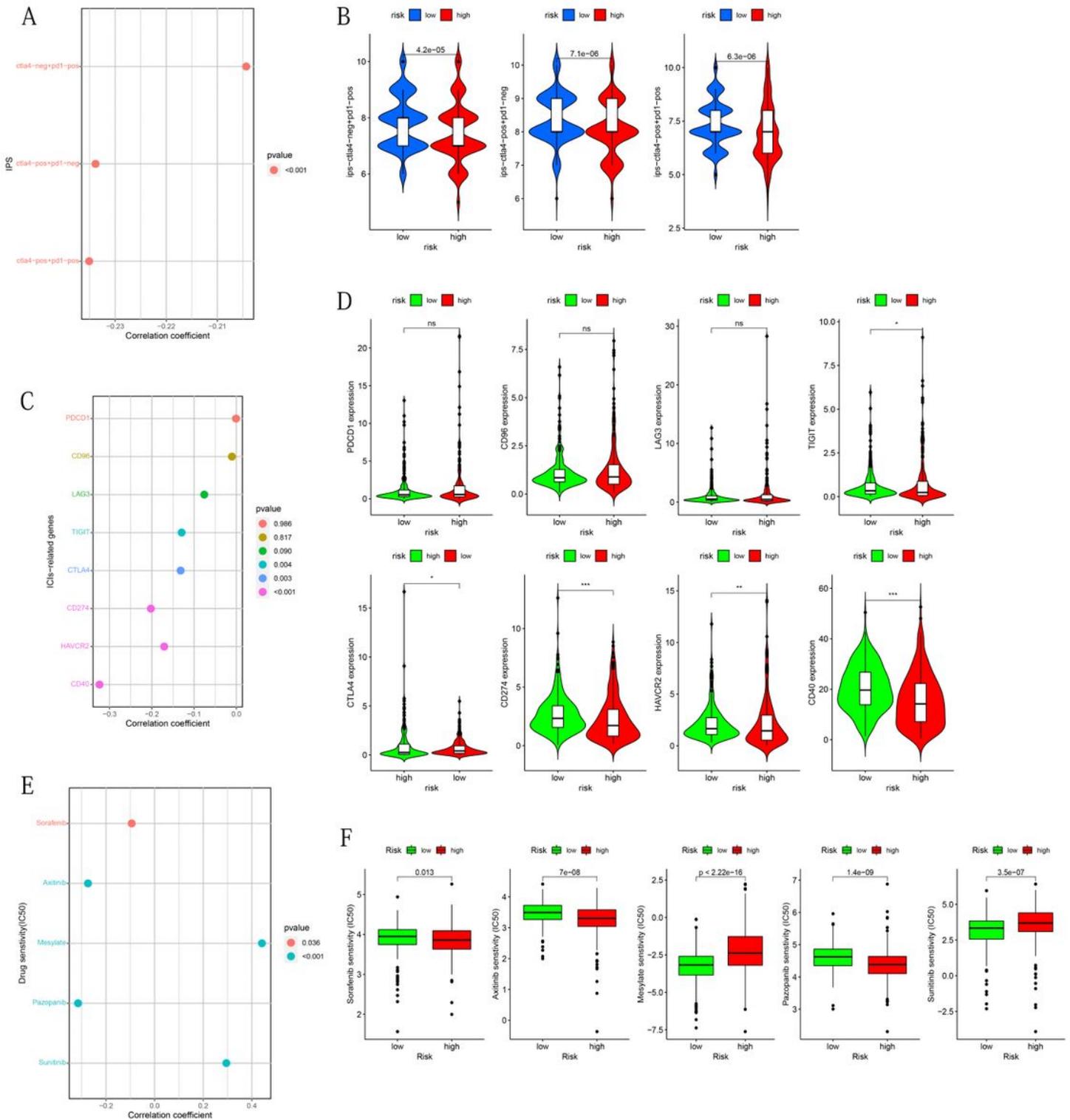


Figure 11

The correlation analysis between risk score and three kinds of IPSs, expression value of 8 ICIs-related genes and IC50 of the five chemotherapy drugs, and comparison between different risk groups. **(A, B)** Three kinds of IPSs. **(C, D)** The expression value of 8 ICIs-related genes. **(E, F)** The IC50 of the five chemotherapy drugs.

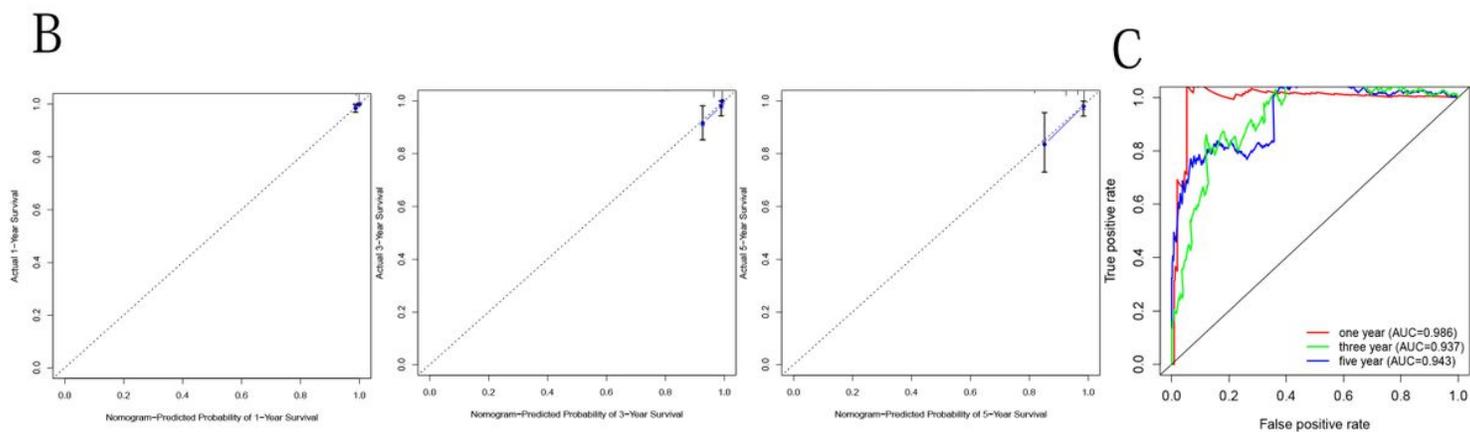
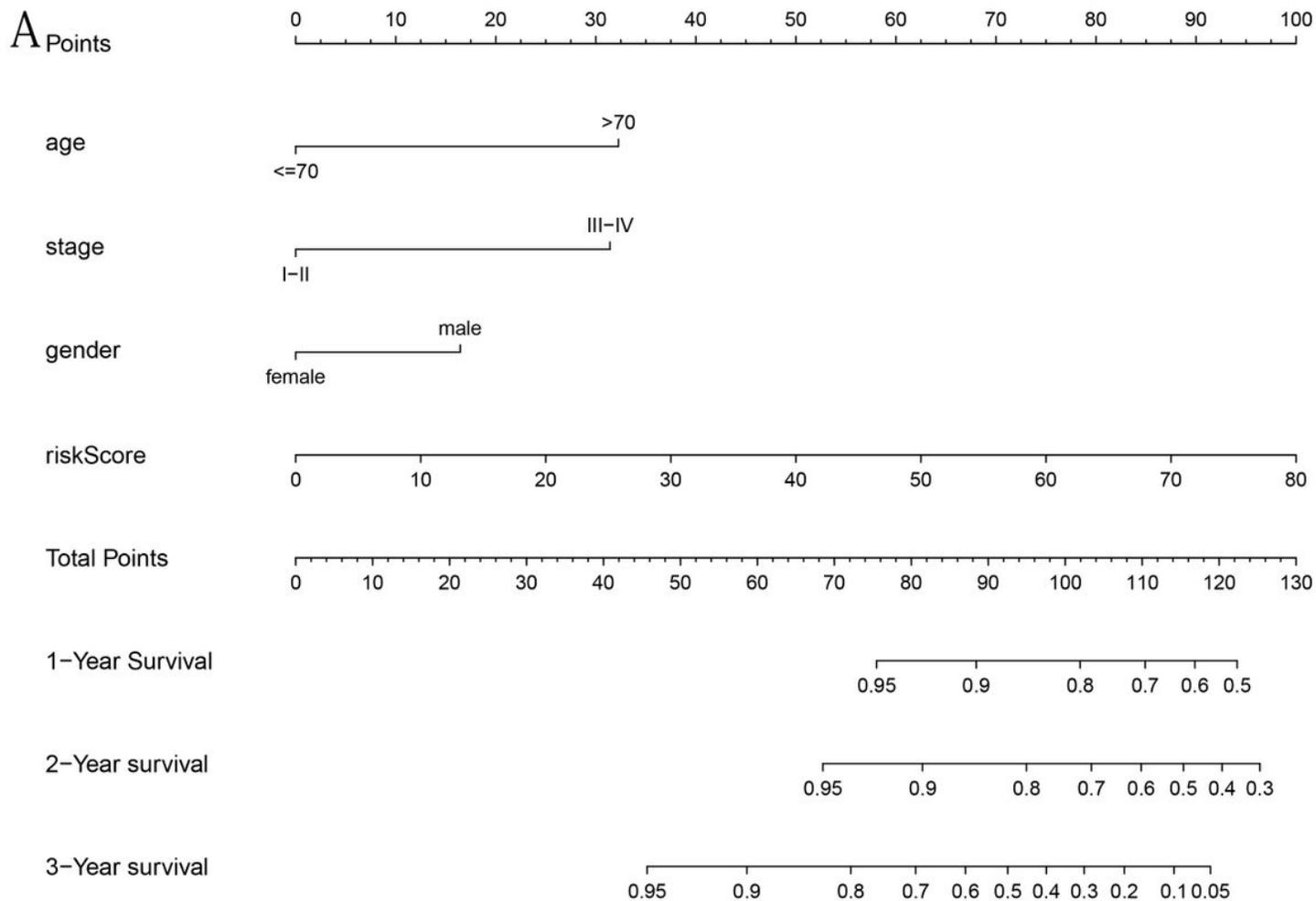


Figure 12

Nomogram based on risk score accurately predicts the OS of THCA patients. **(A)** The Nomogram comprised various clinical risk factors, and the risk score indicated the 1-, 3-, and 5-year survival probability of THCA patients. **(B)** 1-, 3-, and 5-years internal calibration curve. **(C)** 1-, 3-, and 5-years of ROC curves.

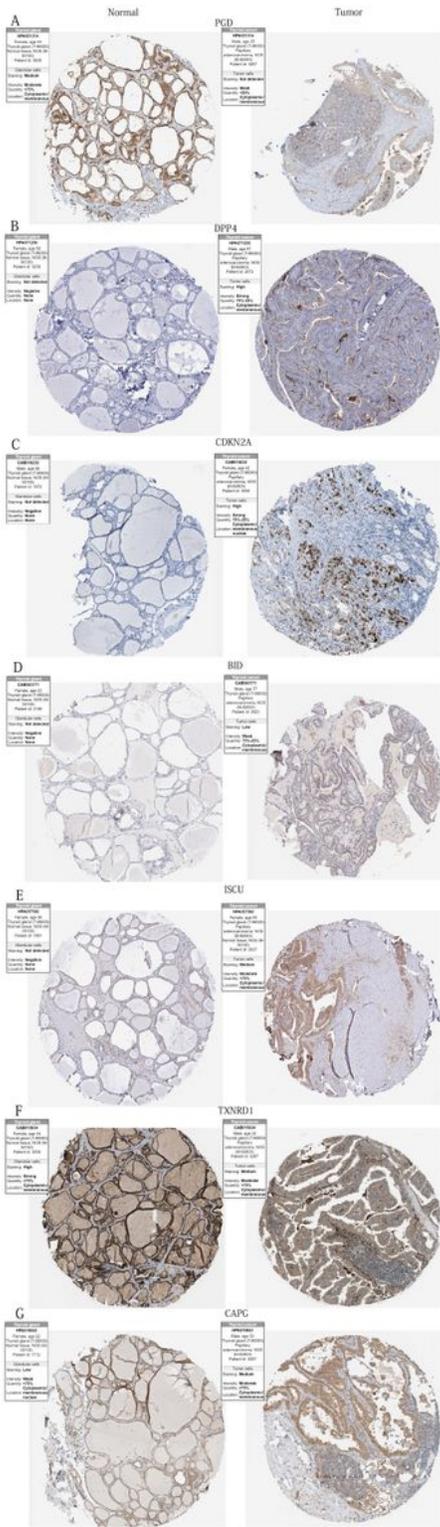


Figure 13

IHC images of 7 differentially expressed DE-FRGs in THCA and normal thyroid tissues. **(A)** PGD, **(B)** DPP4, **(C)** CDKN2A, **(D)** BID, **(E)** ISCU, **(F)** TXNRD1. **(G)** CAPG.

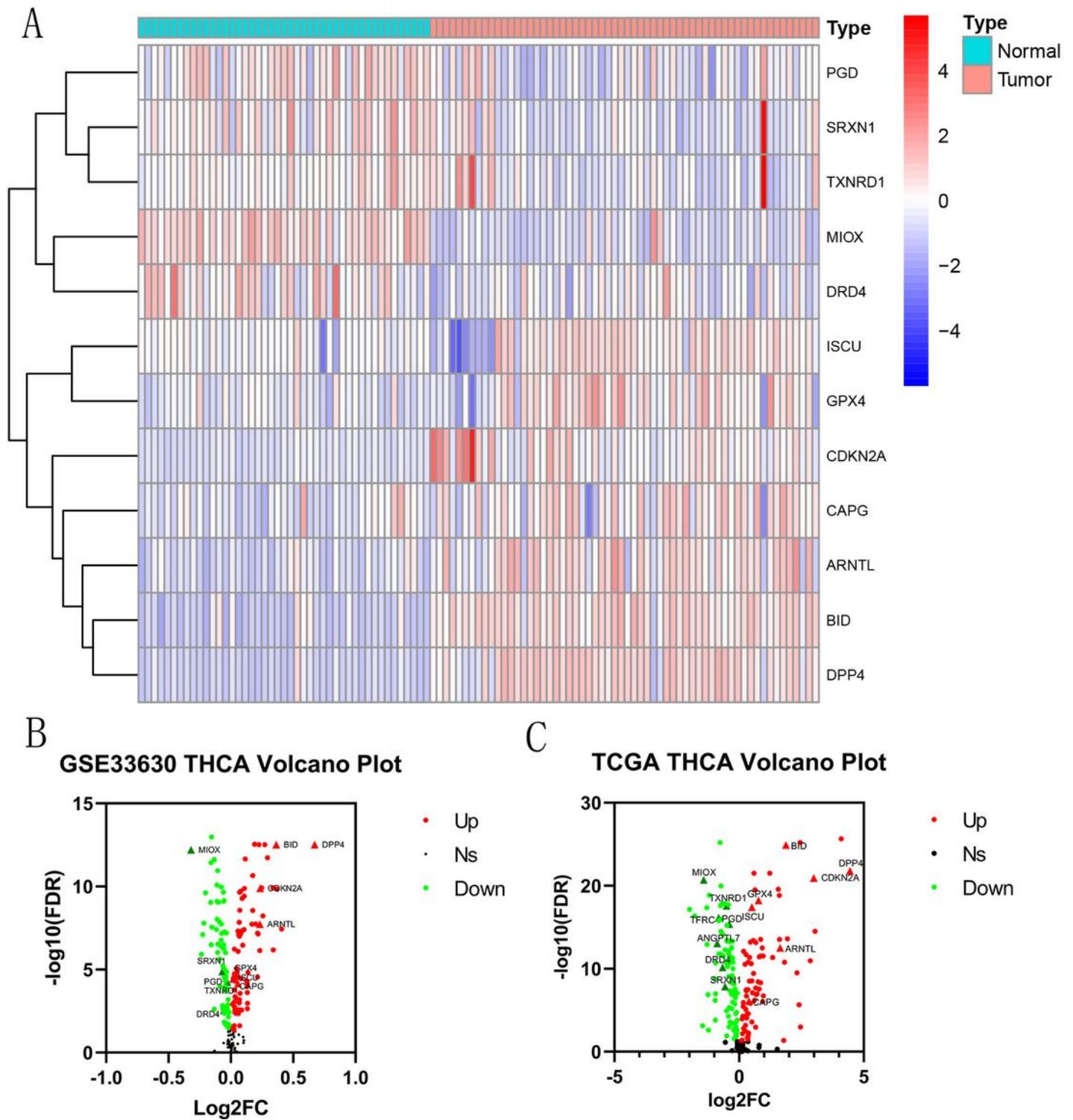


Figure 14

Differential expression verification of PR-DE-FRGs. **(A)** Expression heat map of 12 PR-DE-FRGs from GSE33630 cohort. **(B)** The volcano map shows the differential expression of 12 PR-DE-FRGs from the GSE33630 cohort. **(C)** The volcano map shows the differential expression of 14 PR-DE-FRGs from the TCGA cohort. The triangles represent differentially expressed genes.

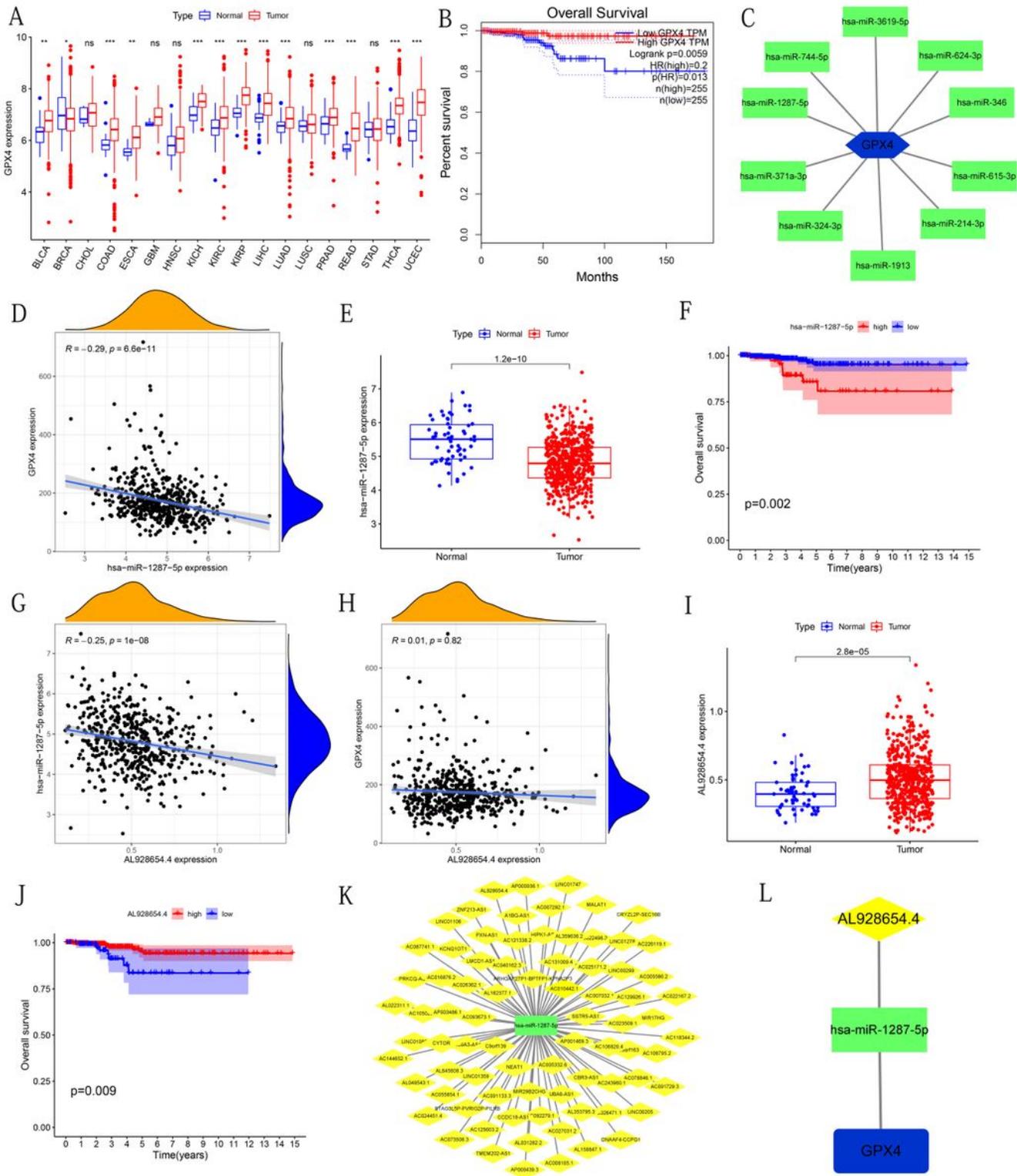


Figure 15

Discovery and verify the ceRNA regulatory network related to GPX4. **(A)** Differences in the expression of GPX4 between 18 cancers and normal tissues, respectively. **(B)** Kaplan-Meier curve based on GPX4 expression and OS obtained from GEPIA. **(C)** The miRNAs-GPX4 regulatory network was established by Cytoscape software. **(D)** Correlation results of GPX4 with hsa-miR-1287-5p. **(E)** Difference analysis results of hsa-miR-1287-5p between THCA and normal group. **(F)** KM survival analysis results of hsa-miR-1287-

5p. **(G)** Correlation results of AL928654.4 with hsa-miR-1287-5p. **(H)** Correlation results of AL928654.4 with GPX4. **(I)** Difference analysis results of AL928654.4 between THCA and normal group. **(J)** KM survival analysis results of AL928654.4. **(K)** The lncRNAs-hsa-miR-1287-5p regulatory network. **(L)** ceRNA regulatory network composed of AL928654.4/ hsa-miR-1287-5p/ GPX4. $P < 0.05$ was considered statistically significant.

Supplementary Files

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

- [SupplementaryFigureS1.jpg](#)
- [SupplementaryFigureS2.jpg](#)
- [SupplementaryTable1.docx](#)