

A signature of Ten Chromodomain Helicase DNA (CHD) Genes Related LncRNAs Predicts Overall Survival in Gastric Cancer Patients

xiaodong Wang

First Affiliated Hospital of Anhui Medical University

yaxian Li

First Affiliated Hospital of Anhui Medical University

yida Lu

First Affiliated Hospital of Anhui Medical University

Yongxiang Li (✉ liyongxiang@ahmu.edu.cn)

First Affiliated Hospital of Anhui Medical University <https://orcid.org/0000-0001-7183-9485>

Research Article

Keywords: gastric cancer, CHDs, LncRNA, bioinformatics, gastric cancer signature

Posted Date: April 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-428006/v1>

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

[Read Full License](#)

Abstract

Background: Gastric cancer (GC) is one of the most common tumors in the world and is often found at an advanced stage. Gene mutations and changes in lncRNA expression often led to the occurrence of GC. In this study, we explored the function of Chromodomain Helicase DNA (CHD) genes and a multi-gene prognostic model was established for its co-expressed lncRNA.

Methods: We used the "Limma" package of R software to analyze the expression differences of CHDs genes. Use Kaplan-Meier curve to plot the relationship between risk score and patient survival time. Receiver operating characteristic (ROC) curve was used to describe the reliability of the model. The data used is obtained from The Cancer Genome Atlas (TCGA) database and CellMiner database.

Results: We found that CHDs gene mutations had significant statistical differences with tumor mutation load and were involved in biological functions such as DNA winding and transcription. 10 CHDs gene-related lncRNAs were used to establish a prognosis model for GC patients. Univariate and multivariate Cox regression analysis proved that risk score could be an independent prognostic factor.

Conclusion: In our study, we developed a novel prognostic model using 10 CHDs-related lncRNAs to predict individualized survival in patients with GC.

1. Introduction

The latest statistics show that gastric cancer is the fifth most common cancer worldwide (5.6%) and the fourth most common cause of cancer death (7.7%) (1). GC patients usually have a poor prognosis, once it is found that patients are usually at an advanced stage and have a high rate of chemotherapy resistance, so it is necessary to determine new therapeutic targets. Although Long non-coding RNAs (lncRNAs) lack the ability to directly encode proteins, they can regulate cell physiology and function through complex epigenetic mechanisms (2). In recent years, with the in-depth studies of lncRNA, it has been proved that lncRNA is widely involved in the growth, metastasis, and invasion of GC (3, 4). lncRNA also plays a role in the regulation of gene expression and autophagy, and these dysfunctions can promote human tumorigenesis (5). Therefore, it is necessary to infer the prognosis of patients from lncRNA and other biological information.

Chromodomain Helicase DNA (CHD) binding protein family, consisting of CHD1-CHD9 genes, is characterized by the chromatin domain and SNF2-associated helicase /ATPase domain of its proteins. CHDs gene can modify chromatin structure to change gene expression thus altering access to the transcriptional apparatus to its Chromosomal DNA template. In recent years, the research on CHDs gene family in GC has been paid more and more attention, which has brought interest and provided some references for us to study CHDs gene family. CHDs is an important regulator of chromatin remodeling, frameshift mutations and deletion of CHDs gene expression are often seen in GC detection, and these changes usually promote tumor development (6). Some researchers examined the serum of patients with GC and found that CHD1 was highly methylated in the serum, the higher methylation level indicating a

worse prognosis (7). It has been reported that CHD5 promoter was methylated in 7 different GC cell lines and was identified as a direct target of miR-454. The low expression of CHD5 promoted tumor growth and had a worse survival time (8–10), compared with the normal tissue CHD8 lower the expression in GC, and can promote tumor cell proliferation (11).

At present, imbalance of lncRNA levels has been reported for each type of cancer, and the role of lncRNA in tumor development has been newly understood, which is expected to become the next type of diagnostic and therapeutic tool in oncology (12). The study of lncRNA transcription has become a new marker for cancer research. We have extracted CHDs-related lncRNA expression by means of co-expression and established a survival model of GC patients by using least absolute selection operator (LASSO) regression analysis, to provide a new method to predict GC patient's survival time, hoping to provide some new help for clinical treatment.

2. Materials And Methods

2.1 Online database

RNA-seq data, clinical information, tumor mutation data and lncRNA data of GC patients were all derived from The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>). Sensitivity data for antitumor agents were obtained from the CellMiner Database (<https://discover.nci.nih.gov/cellminer/home.do>).

2.2 Differentially expressed CHDs, Tumor mutation load and enrichment analysis

We first used the "Limma" package of R software(version 3.6.3) to analyze the differential expression of CHDs gene in normal samples and tumor samples, and to analyze the tumor mutation load (13). Then we conducted Gene Ontology (GO) enrichment analysis on the potential molecular mechanism of CHDs. The "Corrplot" package of R software was used to investigate the correlation between CHDs genes.

2.3 Selection and expression of OS-related lncRNAs

Select the lncRNAs related to CHDs, we still used the "Limma" package of R software to extract the expression data of the lncRNAs, and then used the "igraph" package to draw the node diagram of the lncRNAs and CHDs genes and used the heat map and box diagram to display the expression amount of these lncRNAs. "Survival" package and Univariate Cox regression analysis were used to make forest maps to show OS-related lncRNAs.

2.4 Establishment of the prognostic signature

To improve the accuracy of the prediction model established by CHDs-related lncRNAs, we divided the data of GC patients in TCGA into two groups according to 6.5 : 3.5 and used multiple data to establish the prognosis model. Lasso regression was used to find 10 lncRNAs that could be modeled, and the

correlation coefficient corresponding to each lncRNA was calculated. For each patient: Risk score = expression level of lncRNA1 \times β 1 + expression level of lncRNA2 \times β 2 +...+ expression level of lncRNA10 \times β 10; Where β is the coefficient induced by the multivariate Cox regression model (14).

Subsequently, patients were divided into high-risk group and low-risk group according to the median risk score, and the "SurvivalRoc" package of R software was used to construct time-dependent receiver operating characteristic (ROC) curve, to evaluate the predictive value of the model. Kaplan-Meier survival curves were used to show the difference in survival between the high-risk and low-risk groups, and univariate and multivariate Cox regression analyses were performed to demonstrate that risk score could be used as an independent prognostic factor to predict survival.

2.5 Verification of the prognostic signature

We divided the patients into two groups according to age, sex, stage, and other characteristics. Using the same predictive model, a risk score was calculated for each patient, and then Kaplan-Meier curves were plotted for the number of people in the high-risk and low-risk groups to show how patients survived in the different groups. Next, we used the data of GC patients who were not used to establish the model to draw the survival curve, ROC curve and risk curve respectively according to the same calculation method and used univariate and multivariate Cox regression analysis to further verify the reliability of the model.

2.6 Drug sensitivity analysis

We obtained the RNA-seq and drug activity data of CHDs gene from the CellMiner database, obtained the R package we were going to use from the Bioconductor website to extract the data, and analyzed the correlation of the data by Pearson method to explore the sensitivity of CHDs gene expression to various anti-tumor drugs.

3. Results

3.1 Tumor mutation load, enrichment analysis and differential expression of CHDs.

We first downloaded RNA-seq data from the TCGA database, including 375 tumor samples and 32 normal samples. Then we processed the data and analyzed the expression levels of CHD1-9 in the samples by using the "limma" package of R software. Except for CHD3 and CHD9, the rest of the CHD genes showed significant statistical differences (Fig. 1A). Mutation of tumor genes is the key to the occurrence of tumors. So we analyzed the tumor mutation load, when the CHDs gene was mutated, the tumor mutation load of samples in the mutant group increased to a large extent (Fig. 1B), which means that the CHDs gene may play a promoting role in the mutation of tumor genes.

In addition, to further explore the potential role of CHDs in GC patients, GO analysis was conducted in the R software. As shown in (Fig. 2A), the results of GO enrichment analysis showed that these CHDs genes play an important role in DNA winding, including DNA duplex unwinding, DNA geometric change, DNA

conformation change, DNA helicase activity, helicase activity, and catalytic activity acting on DNA. Co-expression analysis also showed a high co-expression relationship among CHDs genes, such as CHD2 and CHD6, CHD8 and CHD4, with correlation coefficients of 0.52 and 0.53, respectively (Fig. 2B).

3.2 Identification of OS-related lncRNA of CHDs and its expression.

We extracted lncRNAs co-expressed with CHDs genes by R software and obtained a total of 536 lncRNAs associated with CHDs genes (Fig. 3A). In addition, according to the results of univariate Cox regression analysis, a total of 13 differentially expressed lncRNAs were significantly correlated with OS in GC patients (Fig. 3B). Compared with normal tissues, the expressions of these lncRNAs were statistically different in tumors and the results were shown by box diagram (Fig. 4A) and heat map (Fig. 4B).

3.3 Establishment of the ten-lncRNAs risk signature.

Firstly, all the samples were randomly divided into 2 groups and the prognostic model of GC was established by using one group of data of 13 OS-related lncRNAs. Lasso Cox regression analysis was used to select 10 lncRNAs that could be used to establish a prognostic model (Fig. 5A,B), and the coefficients of each lncRNA were obtained (Fig. 5C). Then, according to the coefficient determined by Lasso Cox regression and multiplied by the corresponding lncRNA expression level, the weighted sum was the risk value of each patient. Patients were divided into high-risk and low-risk groups according to the median of the risk value.

Figure 6 shows the changing trend of the number of deaths in patients with GC with the increase of risk value. With the increase of risk value, the red points in the figure increase, indicating the increase of the number of deaths (Fig. 6A,B). Univariate Cox regression analysis showed a significant correlation between risk score and survival in patients with GC ($P < 0.001$, HR = 1.130, 95%CI 1.057–1.209) (Fig. 6E). Multivariate Cox regression analysis confirmed that risk score was an independent prognostic indicator ($P = 0.003$, HR = 1.113, 95%CI 1.038–1.194) (Fig. 6F). The Kaplan-Meier curve shows the difference in survival between patients with high and low risk scores, with a higher score indicating a shorter survival time (Fig. 6G). Then we use the “SurvivalRoc” package of R software to plot ROC curve, the areas under the curve (AUCs) for the risk score at 3-year in predicting OS were 0.805. In conclusion, risk score has a relatively good performance in predicting OS in TCGA dataset.

3.4 Validation of the ten-lncRNAs prognostic model.

Subsequently, to further verify the reliability of the model we established, GC patients in the TCGA database were divided into different groups according to different indicators. Such as by age into less than or equal to 65 and more than 65 in the two groups (Fig. 7A), according to the gender is divided into two groups of men and women (Fig. 7B), according to the stages of early phase and later period is also divided into two groups (Fig. 7C), next we use the same model formula to calculate the group in their

respective risk scores, found in different groups, get high risk score still has a shorter survival time of GC patients, which is consistent with our previous analysis results.

At the same time, we also used the data of another group of TCGA GC patients other than the established model to further verify the validity of the model by using the same computational model. As shown in Fig. 8, the number of deaths in GC patients increased along with the increase in risk worthiness (Figs. 8A,B,C,G). Univariate and multivariate Cox regression analysis also proved that risk score could be used as a predictor to predict the survival of GC patients alone (Figs. 8E,F), and the area under the ROC curve $AUC = 0.777$ also indicated that the model previously established by us had good reliability (Fig. 8D).

3.5 Drug sensitivity analysis of CHD genes.

We jointly analyzed the transcriptional expression of CHDs genes in NCI-60 cancer cell lines and the drug activity of 263 antitumor drugs in the CellMiner database and used Pearson correlation analysis to explore the potential influence of CHDs gene family on drug response, the results are shown in the Fig. 9.

Notably, there was a positive correlation between CHD1 and the sensitivity to Nelarabine (correlation coefficient:0.452, $P < 0.001$), and CHD7 expression was positively associated with Hypothemycin, Trametinib, Selumetinib, and Cobimetinib (correlation coefficient:0.383, $P = 0.003$; 0.359, $P = 0.005$; 0.356, $P = 0.005$, 0.354, $P = 0.006$). However, CHD3 and CHD2 are resistant to Entinostat and Depsipeptide (correlation coefficient:-0.391, $P = 0.002$;-0.382, $P = 0.003$) (Fig. 9).

4. Discussion

In recent years, the incidence and mortality of GC have decreased with the improvement of treatment methods (15). However, the diagnosis and treatment of GC still face great challenges, especially the survival of patients. We cannot accurately predict the survival time only with the current staging system (16). The development of tumors depends on mutations in multiple genes, which means that models composed of multiple genes may be better predictors of patient prognosis than single indicators.

Random errors often occur in the process of DNA replication, and the mismatched DNA will be repaired in time by the body's repair system. When this dynamic balance is broken, gene mutations will often occur. Such replication errors are also affected by external environment, such as radiation, smoking, diet, etc. When the mutated genes accumulate gradually or the key parts of the genes have missense mutations, the growth of cells will begin to be not regulated by the body normally, and the division and proliferation of cells will change, which means the beginning of cancer. CHDs gene family is a family of genes closely related to DNA duplex unwinding, DNA geometric change, DNA conformation change, DNA helicase activity and other biological functions. Studies have found that the occurrence of GC is often accompanied by changes in CHDs gene family. In other types of cancer have similar reports, such as high expression of CHD4 will promote colorectal cancer cell proliferation, invasion and metastasis (17), which were also determined for regulating the proliferation and migration of breast cancer oncogene (18,

19). There also have studies reported mutations of CHD4 through TGF- β signal pathway in promoting endometrial cancer (20). The loss of CHD1 increases the risk of postoperative metastasis of prostate cancer (21), and the expression of CHD5 and CHD9 may be independent biomarkers for prognosis of colorectal cancer (22, 23). Another study reported that patients with high expression of CHD9 had a worse prognosis compared with patients with low expression of CHD9 (24). This attracts us to use TCGA and other public databases for further analysis of CHDs gene, so as to provide new reference for future clinical treatment.

Tumor mutation load refers to the number of somatic mutations removed from the germ line of the tumor genome. Theoretically, higher TMB would also produce more neoantigens, and targeted immunotherapy might be more effective (25). Defective DNA mismatch repair will lead to increased mutation load, and the occurrence of GC and drug resistance are related to this (26, 27). Genomic instability is closely associated with defective repair of DNA damage, and CHD4, which is involved in chromatin relaxation, may affect DNA repair when mutated (28). CDH1 is also involved in chromatin repair, and the loss of CHD1 leads to chromatin dysregulation (29). In addition, CHD1 has been reported to promote DNA damage repair in prostate epithelial cells (30). We analyzed the relationship between CHDs genes and tumor mutation load and showed that mutations in CHDs directly contributed to the increase in TMB, which may be related to the involvement of CHDs genes in DNA duplex unwinding and other biological functions.

With the development of gene research, it has been found that the regulation of protein synthesis is not only the unique function of coding genes. In addition to the role of genetic information carrier, RNA also plays a variety of regulatory functions. LncRNA also plays a complex and precise regulatory role in body development, gene expression, transcriptional activation, transcriptional interference, nuclear transport and other regulatory processes, which have attracted extensive attention. With the further study of lncRNA, the changes of lncRNA transcriptome are expected to become a new indicator in tumor diagnosis and become a new diagnostic and therapeutic tool (12). In this study, CHDs gene-related lncRNAs were studied. Firstly, OS-related lncRNAs were extracted, and a prognostic model based on these lncRNAs was established by Lasso regression method to predict the survival of GC patients. Secondly, ROC curve was used to represent the reliability of the model. Risk curve and survival curve were used to show the survival status of GC patients with different risks scores. Univariate and multivariate Cox regressions analysis were used to test whether risk score could be used as an independent prognostic indicator to predict patient survival. Next, we divided the patients into different groups and used the remaining TCGA data to test the correctness of the model. The results showed good performance, with the increase of risk score, the mortality rate of the patients also increased.

However, our current study still has several limitations to consider. First, our patient data comes from the TCGA database, and the sample size is relatively small with regional distribution. Therefore, we need more data to verify the reliability of the model. Secondly, we need more and larger research centers to verify the value of CHDs-related lncRNA model. Thirdly, we have not yet studied the mechanism of these lncRNAs, and more experiments are needed to clarify the molecular mechanism.

5. Conclusion

In conclusion, a reliable prognostic model for GC patients was established in this study to predict the OS of GC patients.

Declarations

Acknowledgement

Thanks to Xin Xu, Youliang Wu for guiding the format modification and submission of the magazine.

Statement of Ethics

All analyses were based on Public database; thus, no ethical approval and patient consent are required.

Conflict of Interest Statement

The authors declare no conflict of interest.

Consent for publication

Not applicable

Funding Sources

This work was supported by a grant from the National Natural Science Foundation of China (81874063).

Authors' Contributions

Xiaodong Wang collects all the article data and is responsible for writing the full text. Yaxian Li participated in the writing of the article and the modification of the article format. Yida Lu was responsible for the editing of the pictures and participated in the writing of the full text. Yongxiang Li provided the ideas for the research and all the funding. All authors read and approved the final manuscript.

Availability of data and materials

The data included in the current study are available in the TCGA database (<https://cancergenome.nih.gov/>) and the CellMiner database (<https://discover.nci.nih.gov/cellminer/home.do>).

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.

- CA: a cancer journal for clinicians. 2021.
2. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *Cell*. 2009;136(4):629–41.
 3. Dong Y, Li X, Lin Z, Zou W, Liu Y, Qian H, et al. HOXC-AS1-MYC regulatory loop contributes to the growth and metastasis in gastric cancer. *J Exp Clin Cancer Res*. 2019;38(1):502.
 4. Xie S, Chang Y, Jin H, Yang F, Xu Y, Yan X, et al. Non-coding RNAs in gastric cancer. *Cancer Lett*. 2020;493:55–70.
 5. Li G, Deng L, Huang N, Sun F. The Biological Roles of lncRNAs and Future Prospects in Clinical Application. *Diseases*. 2021;9(1).
 6. Kim MS, Chung NG, Kang MR, Yoo NJ, Lee SH. Genetic and expressional alterations of CHD genes in gastric and colorectal cancers. *Histopathology*. 2011;58(5):660–8.
 7. Li Z, Guo Z. Comparison of CDH1 Gene Hypermethylation Status in Blood and Serum among Gastric Cancer Patients. *Pathol Oncol Res*. 2020;26(2):1057–62.
 8. Hashimoto T, Kurokawa Y, Wada N, Takahashi T, Miyazaki Y, Tanaka K, et al. Clinical significance of chromatin remodeling factor CHD5 expression in gastric cancer. *Oncol Lett*. 2020;19(1):1066–73.
 9. Wang X, Lau KK, So LK, Lam YW. CHD5 is down-regulated through promoter hypermethylation in gastric cancer. *J Biomed Sci*. 2009;16:95.
 10. Xu G, Zhu H, Zhang M, Xu J. Histone deacetylase 3 is associated with gastric cancer cell growth via the miR-454-mediated targeting of CHD5. *Int J Mol Med*. 2018;41(1):155–63.
 11. Sawada G, Ueo H, Matsumura T, Uchi R, Ishibashi M, Mima K, et al. CHD8 is an independent prognostic indicator that regulates Wnt/beta-catenin signaling and the cell cycle in gastric cancer. *Oncol Rep*. 2013;30(3):1137–42.
 12. Torsin LI, Petrescu GED, Sabo AA, Chen B, Brehar FM, Dragomir MP, et al. Editing and Chemical Modifications on Non-Coding RNAs in Cancer: A New Tale with Clinical Significance. *Int J Mol Sci*. 2021;22(2).
 13. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
 14. Wang Z, Gao L, Guo X, Feng C, Lian W, Deng K, et al. Development and validation of a nomogram with an autophagy-related gene signature for predicting survival in patients with glioblastoma. *Aging*. 2019;11(24):12246–69.
 15. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J Clin*. 2018;68(6):394–424.
 16. Zeng D, Zhou R, Yu Y, Luo Y, Zhang J, Sun H, et al. Gene expression profiles for a prognostic immunoscore in gastric cancer. *Br J Surg*. 2018;105(10):1338–48.
 17. Xia L, Huang W, Bellani M, Seidman MM, Wu K, Fan D, et al. CHD4 Has Oncogenic Functions in Initiating and Maintaining Epigenetic Suppression of Multiple Tumor Suppressor Genes. *Cancer Cell*.

- 2017;31(5):653–68. e7.
18. D'Alesio C, Bellese G, Gagliani MC, Lechiara A, Dameri M, Grasselli E, et al. The chromodomain helicase CHD4 regulates ERBB2 signaling pathway and autophagy in ERBB2(+) breast cancer cells. *Biol Open*. 2019;8(4).
 19. Wang Y, Chen Y, Bao L, Zhang B, Wang JE, Kumar A, et al. CHD4 Promotes Breast Cancer Progression as a Coactivator of Hypoxia-Inducible Factors. *Cancer Res*. 2020;80(18):3880–91.
 20. Li Y, Liu Q, McGrail DJ, Dai H, Li K, Lin SY. CHD4 mutations promote endometrial cancer stemness by activating TGF-beta signaling. *Am J Cancer Res*. 2018;8(5):903–14.
 21. Oh-Hohenhorst SJ, Tilki D, Ahlers AK, Suling A, Hahn O, Tennstedt P, et al. CHD1 loss negatively influences metastasis-free survival in R0-resected prostate cancer patients and promotes spontaneous metastasis in vivo. *Cancer Gene Ther*. 2021.
 22. Zhou Y, Qiu YE, Qin SL, Luo Y, Cui R, Qin J, et al. Expression of CHD5 may serve as an independent biomarker of prognosis in colorectal cancer via immunohistochemical staining. *Int J Clin Exp Pathol*. 2018;11(5):2792–8.
 23. Xu L, Peng H, Huang XX, Xia YB, Hu KF, Zhang ZM. Decreased expression of chromodomain helicase DNA-binding protein 9 is a novel independent prognostic biomarker for colorectal cancer. *Braz J Med Biol Res*. 2018;51(9):e7588.
 24. Guan B, Ran XG, Du YQ, Ren F, Tian Y, Wang Y, et al. High CHD9 expression is associated with poor prognosis in clear cell renal cell carcinoma. *Int J Clin Exp Pathol*. 2018;11(7):3697–702.
 25. Zhang C, Chen Z, Chong X, Chen Y, Wang Z, Yu R, et al. Clinical implications of plasma ctDNA features and dynamics in gastric cancer treated with HER2-targeted therapies. *Clin Transl Med*. 2020;10(8):e254.
 26. Buttura JR, Provisor Santos MN, Valieris R, Drummond RD, Defelicibus A, Lima JP, et al. Mutational Signatures Driven by Epigenetic Determinants Enable the Stratification of Patients with Gastric Cancer for Therapeutic Intervention. *Cancers (Basel)*. 2021;13(3).
 27. Chang W, Wang H, Kim W, Liu Y, Deng H, Liu H, et al. Hormonal Suppression of Stem Cells Inhibits Symmetric Cell Division and Gastric Tumorigenesis. *Cell Stem Cell*. 2020;26(5):739–54. e8.
 28. Hou T, Cao Z, Zhang J, Tang M, Tian Y, Li Y, et al. SIRT6 coordinates with CHD4 to promote chromatin relaxation and DNA repair. *Nucleic Acids Res*. 2020;48(6):2982–3000.
 29. Zhang Z, Zhou C, Li X, Barnes SD, Deng S, Hoover E, et al. Loss of CHD1 Promotes Heterogeneous Mechanisms of Resistance to AR-Targeted Therapy via Chromatin Dysregulation. *Cancer Cell*. 2020;37(4):584–98. e11.
 30. Zhu Y, Wen J, Huang G, Mittlesteadt J, Wen X, Lu X. CHD1 and SPOP synergistically protect prostate epithelial cells from DNA damage. *Prostate*. 2021;81(1):81–8.

Figures

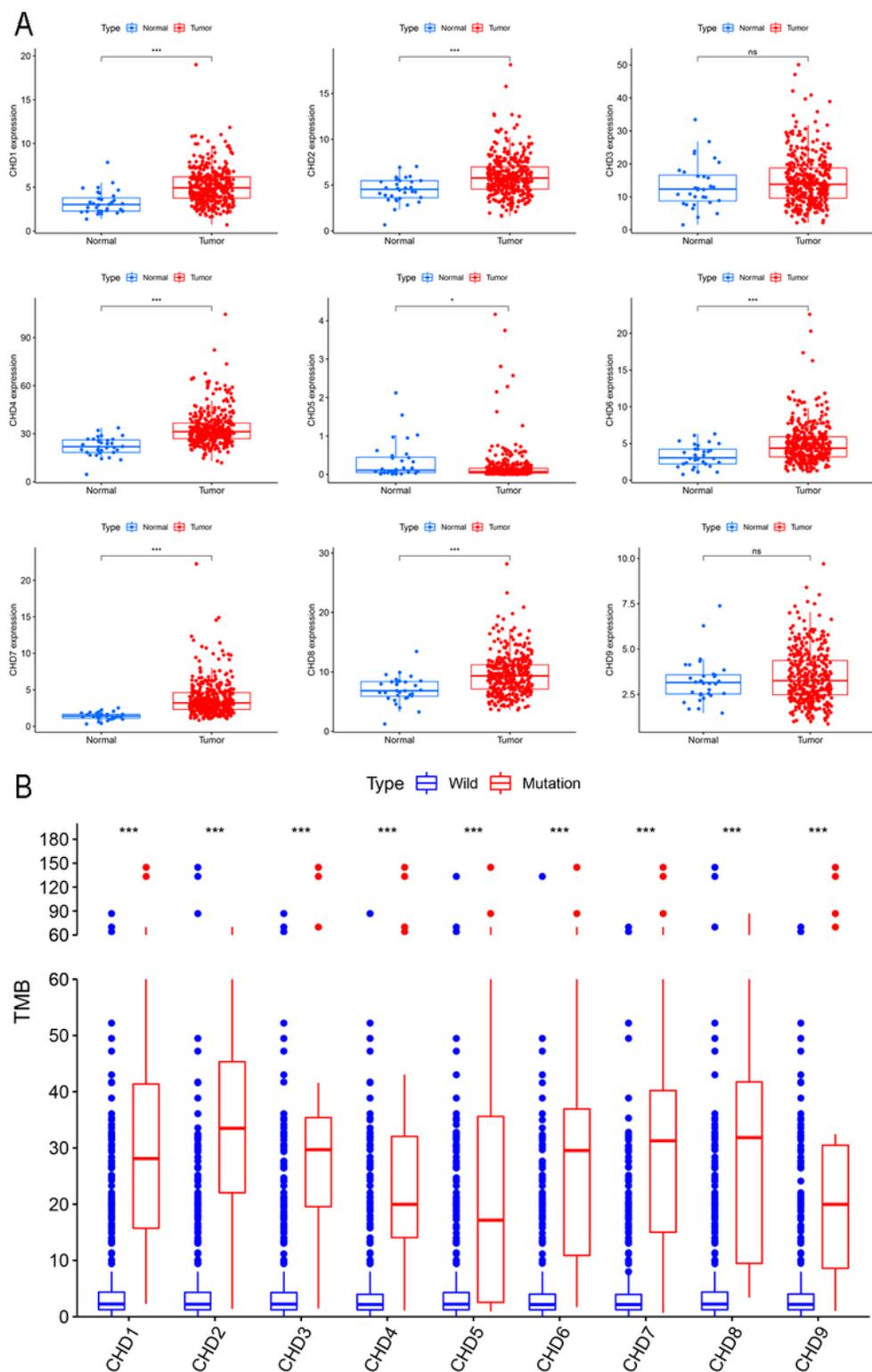


Figure 1

Differential expression of genes and their association with TMB. A: scatter plot shows the expression level of CHDs gene compared with that of adjacent cancer; B: box plot shows the relationship between CHDs and TMB. (***) $p < 0.001$; (**) $p < 0.01$; (*) $p < 0.05$)

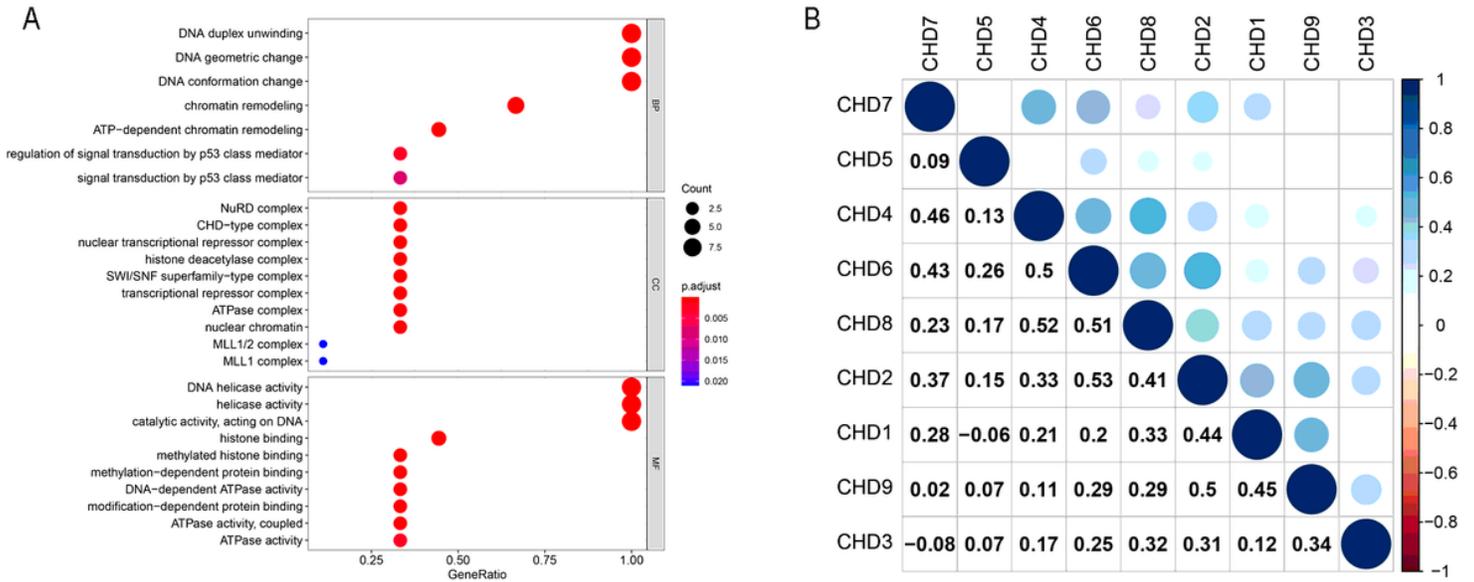


Figure 2

GO enrichment and protein interaction analysis. A: using bubble diagram to show the results of GO enrichment analysis; B: showing the interaction relationship between CHDs genes.

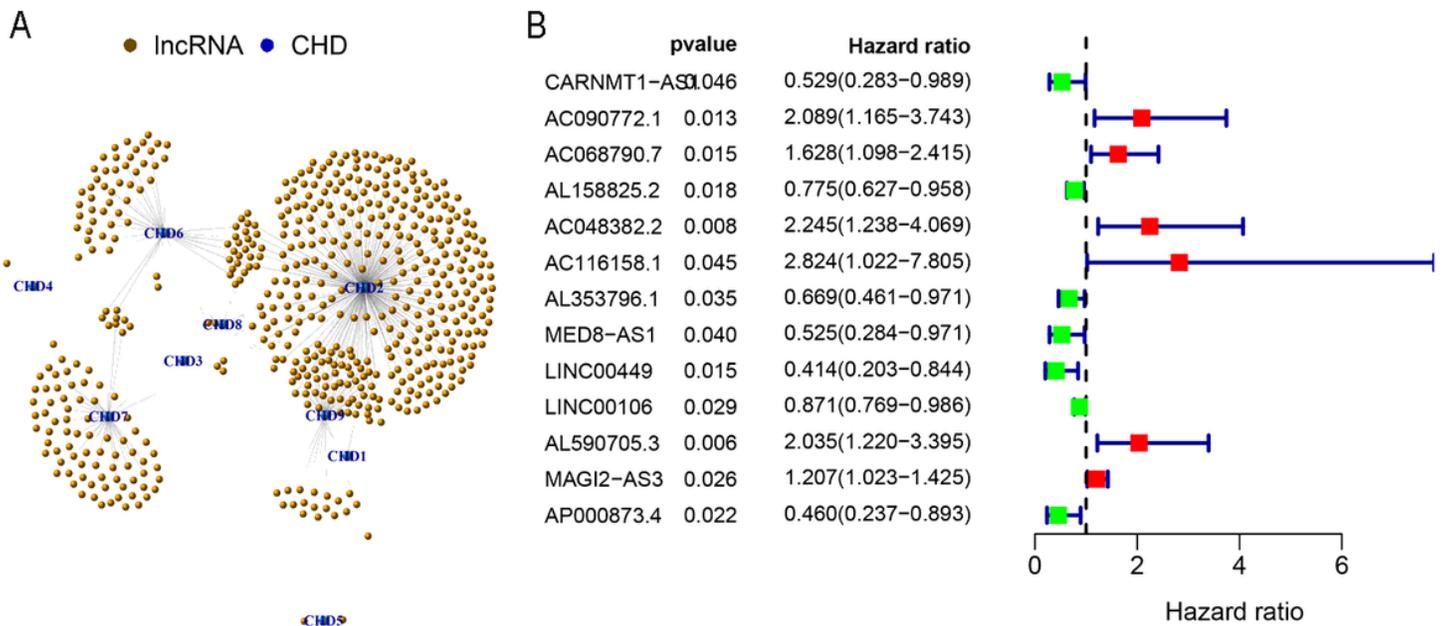


Figure 3

Node plot of lncRNAs and its univariate Cox regression analysis. A: node diagram showing the association between CHDs genes and related lncRNAs; B: Univariate Cox regression analysis and forest map were used to display OS-related lncRNAs.

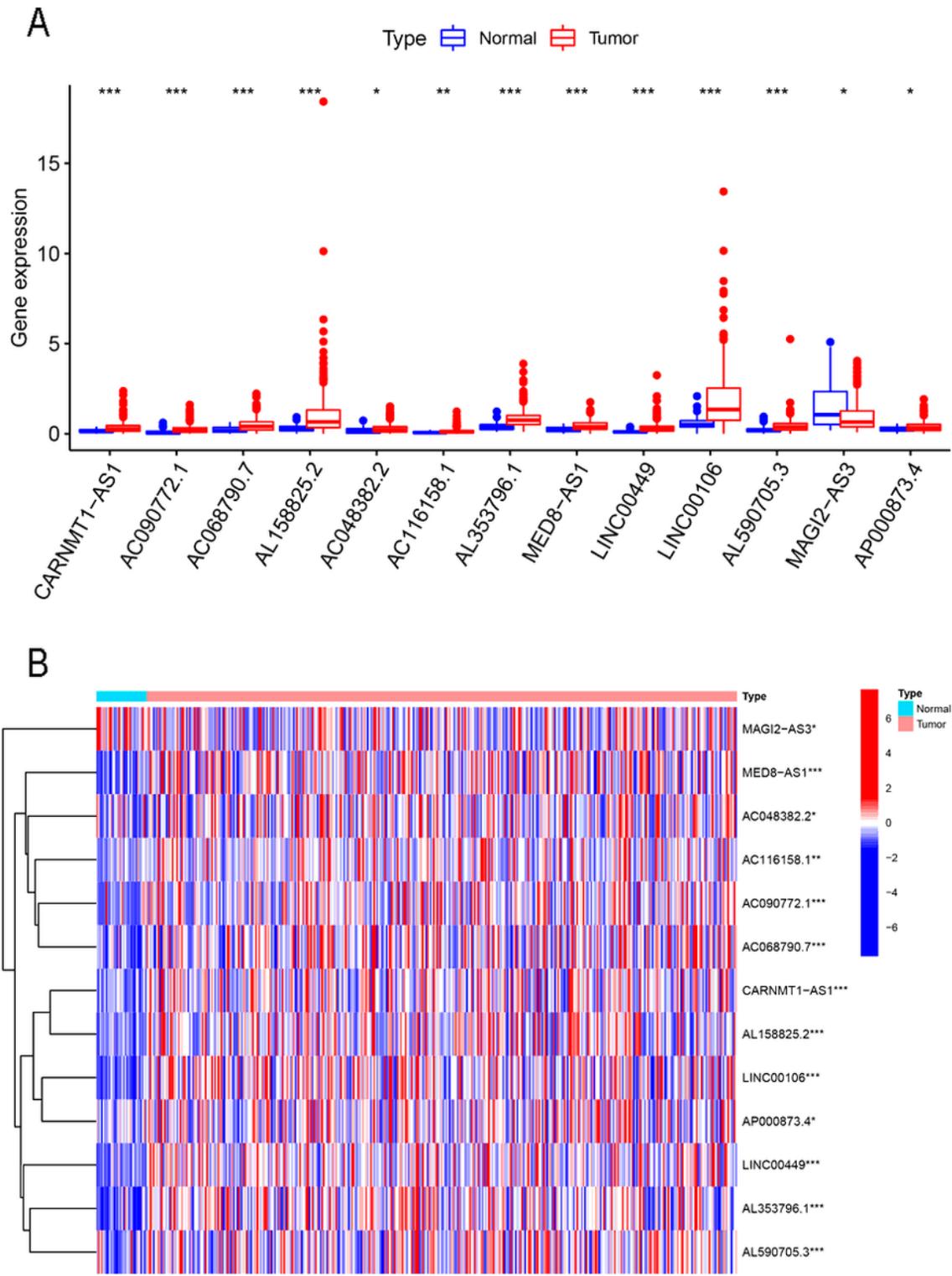


Figure 4

Differential expression analysis of lncRNAs. A: Box diagram shows the difference in lncRNA expression levels; B: Heat map shows the expression of lncRNA in normal samples and tumor samples. (** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$)

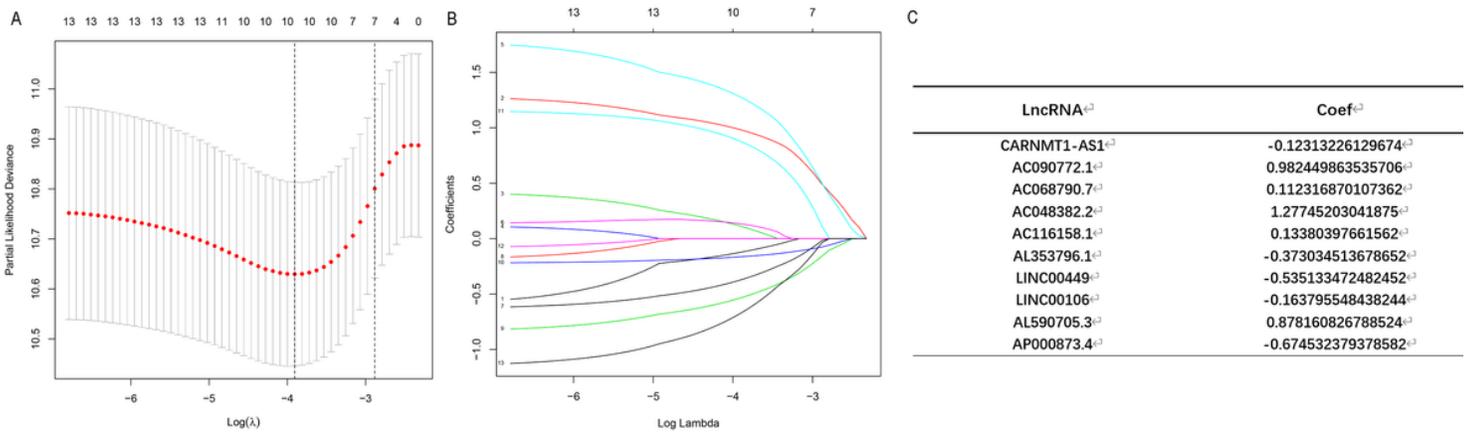


Figure 5

Lasso regression was used to establish a prognostic model. A-C: The Last absolute selection operator (Lasso) Cox regression for CHDs-related lncRNAs that meets the criteria of $P < 0.05$ in the univariate Cox regression analysis.

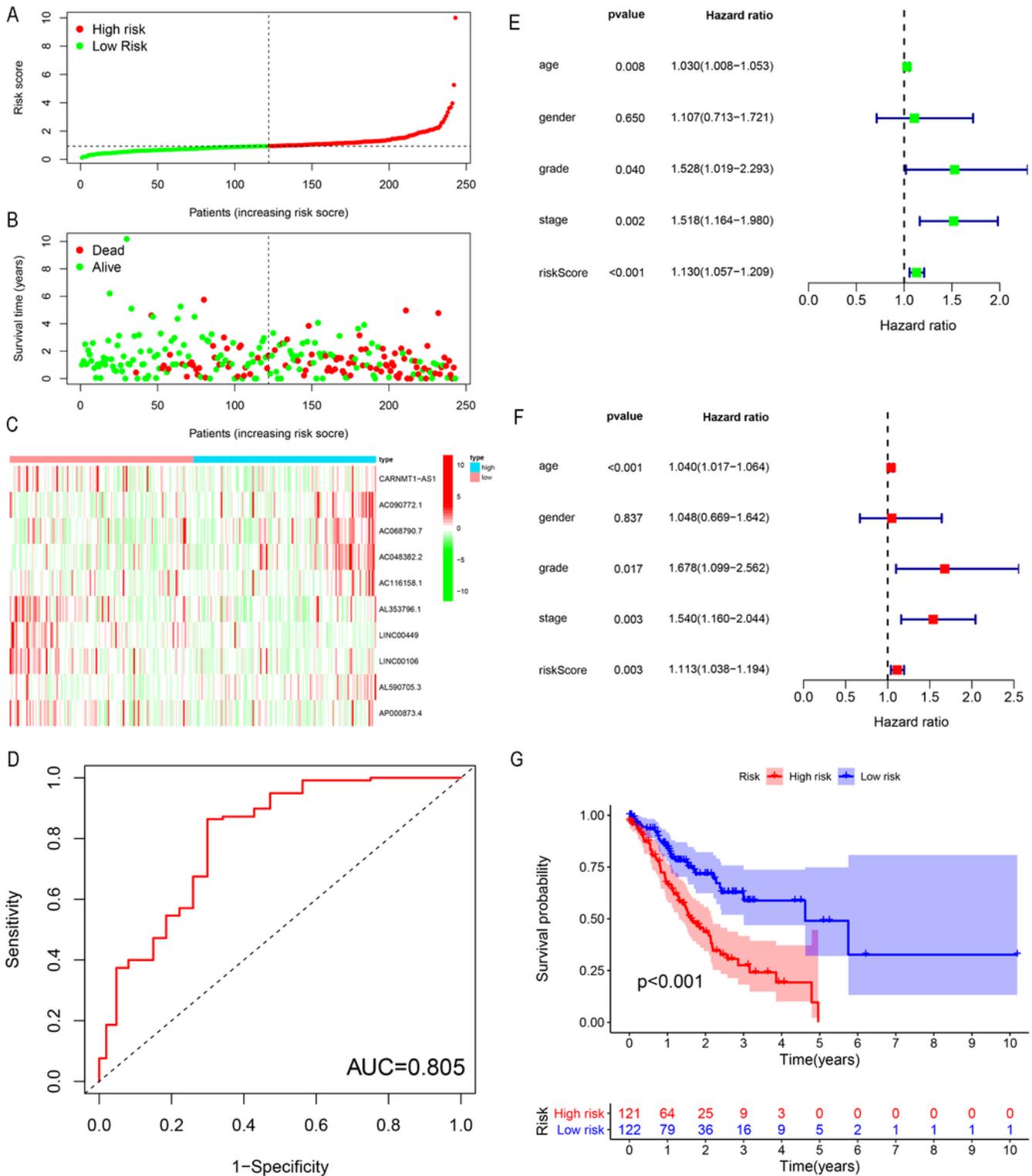


Figure 6

The prognostic model was established using TCGA database and verified its prognostic value. A-B: risk curve shows the survival status of high-risk and low-risk patients; C: heat map shows the expression of lncRNAs in the high and low risk groups; D: receiver operating characteristic (ROC) curve for patients with higher and lower risk score ($P < 0.001$, $AUC = 0.805$). E-F: univariate, and multivariate Cox regression

analysis tested the reliability of risk score; G: Kaplan - Meier curves for patients with higher and lower risk score.

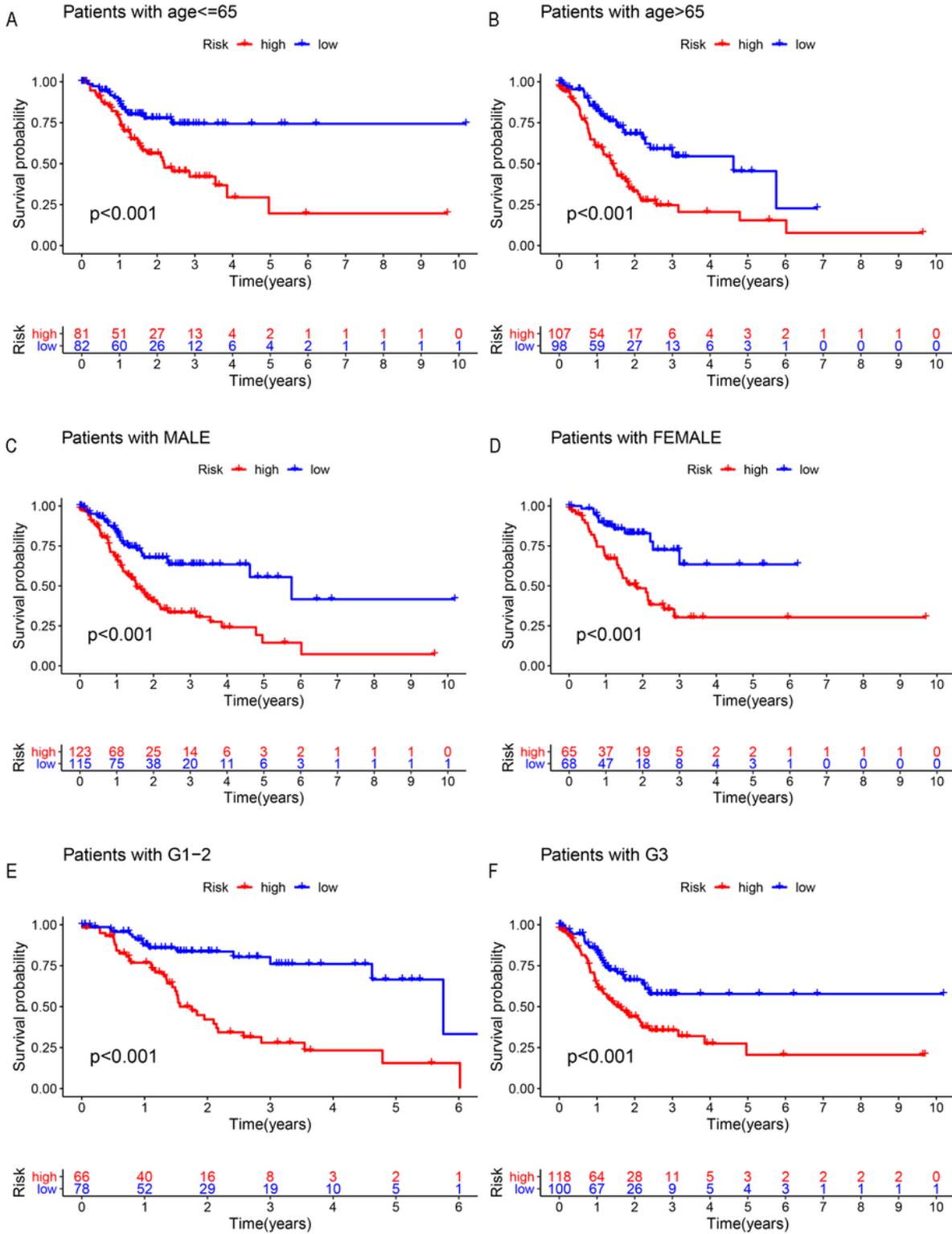


Figure 7

Validation of the model by patients in different groups. A-F: Kaplan -Meier curves show survival between low-risk and high-risk patients in different groups.

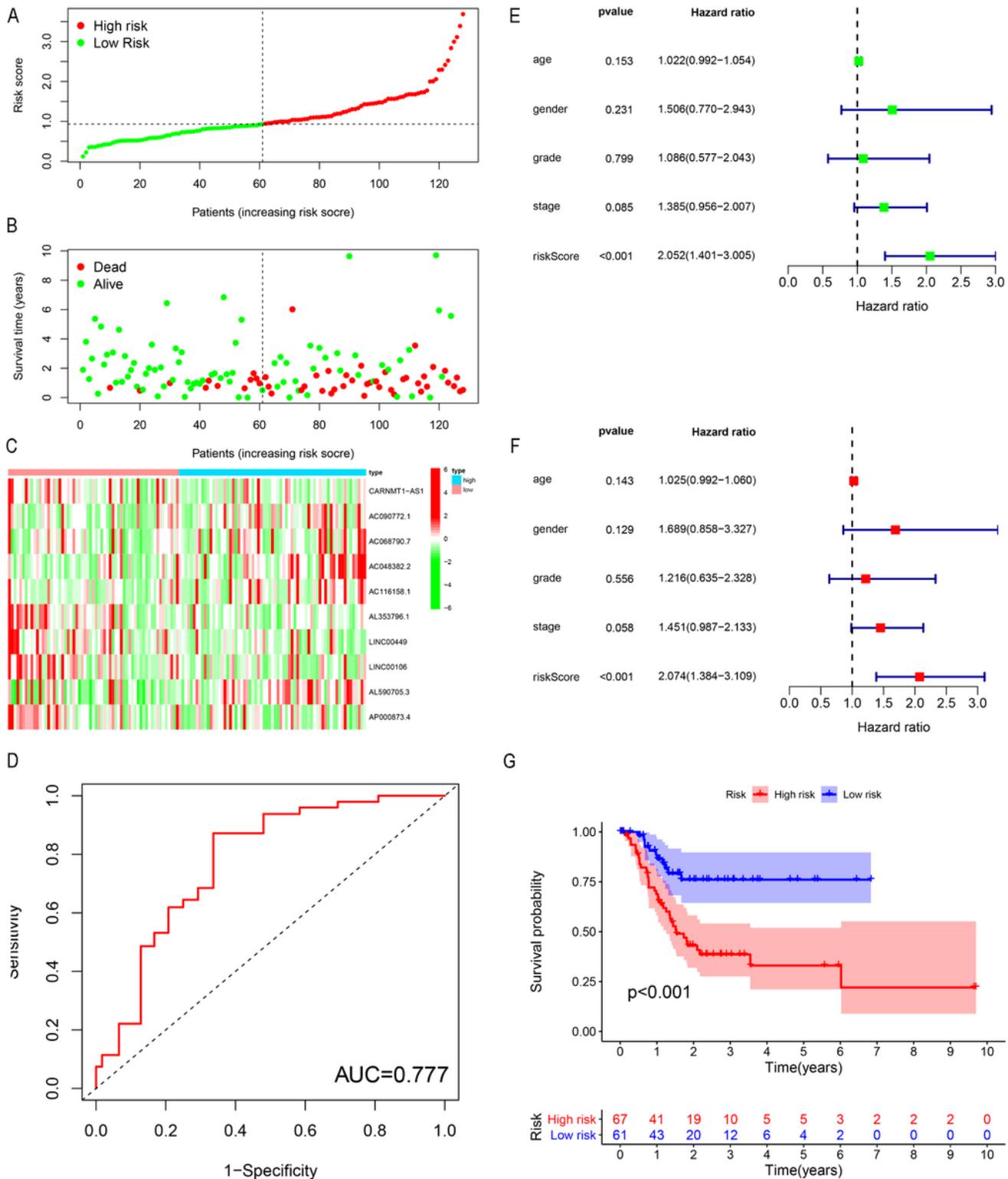


Figure 8

Other TCGA data were used to validate the model. A-B: Risk distribution and survival status of patients with gastric cancer. The dotted line is the best dividing line for GC patients into high-risk and low-risk groups; C: Heat map shows the expression of lncRNAs; D: receiver operating characteristic (ROC) curve verifies the accuracy of the model (AUC = 0.777); E-F: univariate, and multivariate Cox regression analysis, tested the reliability of the signature; G: Kaplan -Meier curves for patients with higher and lower risk score.

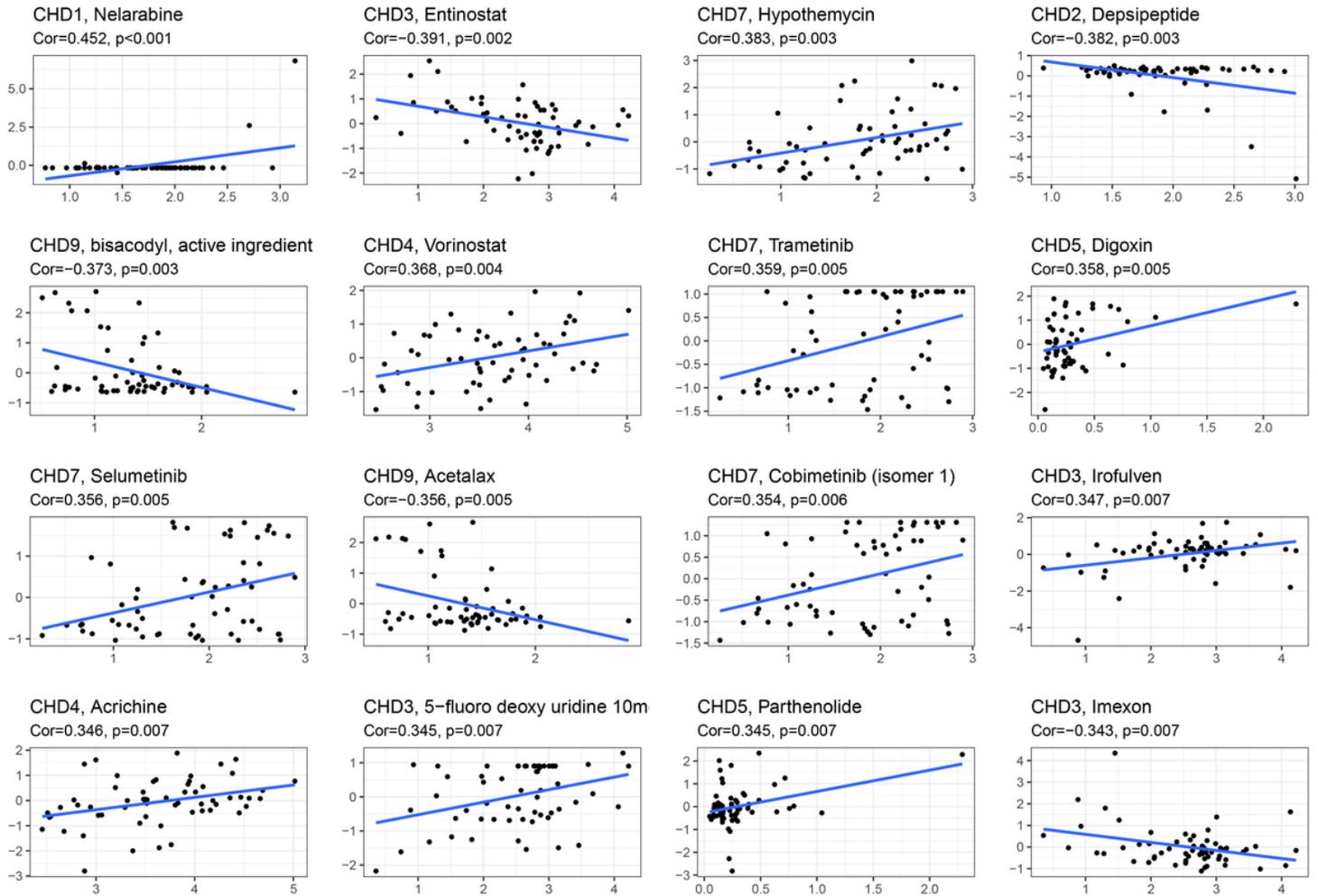


Figure 9

Drug response analysis. Correlation between TCGA tumor drug sensitivity and CHDs, Scatter plot sorted by P value.