Prognostic value of an immune long non-coding RNA signature in liver hepatocellular carcinoma

rui kong
Tongji University School of Medicine https://orcid.org/0000-0002-8202-1619
Nan Wang
Shanghai Tenth People's Hospital
Wei Han
Shanghai Tenth People's Hospital
Yuejuan Zheng
Shanghai University of Traditional Chinese Medicine
Jie Lu
kennisren@hotmail.com

https://orcid.org/0000-0002-1814-2125

Research article

Keywords: immune prognostic signature, long non-coding RNA, liver hepatocellular carcinoma, overall survival

Posted Date: December 19th, 2019

DOI: https://doi.org/10.21203/rs.2.19313/v1

Version of Record: A version of this preprint was published at Journal of Microbiology and Biotechnology on January 30th, 2024. See the published version at https://doi.org/10.4014/jmb.2308.08022.

Abstract

Background: In recent years, long non-coding RNAs (lncRNAs) are emerging as crucial regulators in the immunological process of liver hepatocellular carcinoma (LIHC). Increasing studies have found that some lncRNAs could be used as a diagnostic or therapeutic target for clinical management, but little research has investigated the role of immune-related lncRNA in tumor prognosis. In this study, we aimed to develop an immune lncRNA signature for the precise diagnosis and prognosis of liver hepatocellular carcinoma.

Methods: Gene expression profiles of LIHC samples obtained from TCGA were screened for immune-related genes using two reference gene sets. The optimal immune-related lncRNA signature was built via correlational analysis, univariate and multivariate cox analysis. Then the Kaplan-Meier plot, ROC curve, clinical analysis, gene set enrichment analysis, and principal component analysis were carried out to evaluate the capability of immune lncRNA signature as a prognostic indicator.

Results: Six long non-coding RNA MSC–AS1, AC009005.1, AL117336.3, AL031985.3, AL365203.2, AC099850.3 were identified via correlation analysis and cox regression analysis considering their interactions with immune genes. Next, tumor samples were separated into two risk groups by the signature with different clinical outcomes. Stratification analysis showed the prognostic ability of this signature acted as an independent factor. The AUC value of ROC curve was 0.779. The Kaplan-Meier method was used in survival analysis and results showed a statistical difference between the two risk groups. The predictive performance of this signature was validated by principal component analysis (PCA). Data from gene set enrichment analysis (GSEA) further unveiled several potential biological processes of these biomarkers may involve in.

Conclusion: In summary, the study demonstrated the potential role of the six-lncRNA signature served as an independent prognostic factor for LIHC patients.

Background

Liver hepatocellular carcinoma is one of the most common malignancies worldwide with increasing mortality and incidence rate [1, 2]. Surgical operation is the main method in LIHC administrations, but many patients are on the middle-advanced stage at first diagnosed and missed the chance of accepting surgery [3–5]. The liver is defined as a lymphoid organ in a broad sense [6]. It has been reported that cytokines, hepatic non-parenchymal cells, dendritic cells, and lymphocytes played a modulating role in immunological tolerance during tumor progression [7–10]. Meanwhile, immunology therapy comprises of the immune checkpoint, adoptive cellular immunotherapy (ACT) and vaccines have opened new horizons in LIHC treatment [11, 12]. Several studies clarified the clinical application of immune checkpoint blockade programmed cell death-1(PD-1) and cytotoxic T-lymphocyte antigen-4 (CTLA4) have enhanced the survival rate of some advanced patients [13–15]. Thus, immune biomarkers of high sensitivity and specificity in diagnosing and predicting the prognosis of LIHC are in urgent need of research.

Long non-coding RNA is a type of RNA poorly conserved with a length of 200 bp to 100kbp. It is capable of modulating gene expression at four main levels: epigenetic regulation, epigenetic transcriptional, post-transcriptional and translational regulations [16–21]. According to its location with protein-coding mRNA, lncRNA could be classified into four categories: antisense, pseudogene, long intergenic ncRNA and intronic lncRNA [22]. Recent publications have revealed the pivotal role of lncRNA implicated in innate immune response and T cell development, differentiation, activation as well. [23, 24] Furthermore, some studies have reviewed the relationship between aberrant expression of lncRNA and tumorigenesis, metastasis, diagnosis or prognosis [25–27]. For example, HULC, a lncRNA specifically located on cell plasma is found with high expression in hepatoma cells and promote cell proliferation [28, 29]. H19 has been reported at an elevated level in HBV-associated HCC and suppresses tumor metastasis [30].

In the current study, a robust immune lncRNA model was set up to assist diagnosis and prognosis in LIHC via bioinformatic analysis. Long non-coding RNA: MSC – AS1, AC009005.1, AL117336.3, AL031985.3, AL365203.2, AC099850.3 were critical components of the whole model. Data suggested a high expression level of these biomarkers was positively correlated with poor survival and malignancy phenotypes in TCGA data. Univariate cox regression and multivariate cox regression analysis further clarified this signature had an independent influence on overall survival. Results from the KM plot, ROC curve, PCA further proved the sensitivity and reliability of this prognostic model.

Methods

Data source and processing

RNA-sequencing data of LIHC cases were retrieved from the TCGA repository [31], which were operated on the Illumina HiSeq RNA-Seq platform. The corresponding clinical data including survival time, TNM classification information, risk factors, etc. were also downloaded. The data contained 424 samples, of which 50 were normal, and 374 were primary hepatocellular carcinoma. The use and acquisition of these data are according to TCGA data access policies and publication guidelines.

The human general transfer format obtained from the ensemble website [32] was used to match the ensemble name of mRNA and long non-coding RNA. Clinical samples without precise outcome or follow-up time less than 30 days were excluded from our study.

Selection of immune-related long non-coding RNAs

Two gene sets “immune response” and “immune system process” were acquired from the Molecular Signatures Database [33] and then employed to identify the immune-related genes of LIHC samples. Afterward, these genes were fitted into correlation analysis using “limma” packages and “cor function” in R to find the correlated lncRNAs, with the filter criteria set to absolute cor (corresponding coefficients) > 0.4, adjust *P* value < 0.001. Then the network was visualized by Cytoscape software [34]. In this study, lncRNAs failed to meet the following conditions were deleted: (1). exhibited expression (FPKM ≥ 1) in at least 50% of the LIHC samples; (2). exhibited no fluctuation in expression levels between samples [35].

Construction of prognostic lncRNA signature and statistical analysis

Univariate cox regression analysis was performed to select a subset of candidate lncRNAs, whose expression level and patient's overall survival (OS) are closely associated ($P < 0.001$). Afterward, a multivariate cox regression analysis was conducted to screen these candidates to build the prognostic lncRNA model. The formulae of the risk score were given below. Based on the cut-off of risk score, tumor specimens were divided into high-risk group and low-risk group. Kaplan-Meier survival method was used to calculate the survival rate. Receiver Operating Characteristic curve (ROC) was carried out to assess the effectiveness of this lncRNA model. Akaike Information Criterion (AIC) was adopted in the optimization process. All analyses were implemented using the following packages "survival", "survminer", "survival ROC", "pheatmap" on R platform.

lncRNA risk score = |cor|* expression value

risk score of samples = $\sum$ (lncRNA risk score)

Clinical feature analysis of lncRNA signature

Based on the foregoing analyses, we obtained the expression matrix of lncRNA biomarkers in LIHC patients and then combined the expression data with clinical information. Subsequently, clinical features such as T-stage (tumor), G-stage (histologic grade) and S-stage (pathological stage) were analyzed to investigate the correlation between these clinic factors and expression of lncRNA biomarkers ($P < 0.05$) using "ggpubr" package in R.

Gene set enrichment analysis

Gene set enrichment analysis (GSEA) was applied to figure out the underlying mechanism between risk scores, which were derived from the co-expression analysis. Two immune gene sets "immune system process" (M13664 Genes annotated by the GO term GO:0002376), "immune response" (M19817 Genes annotated by the GO term GO:0006955) mentioned above were downloaded from MsigDB database as reference gene sets. The enrichment results were considered statistically significant at a level of FDR < 0.25.

Principal component analysis

Principal component analysis was employed to show the cluster of samples with risk scores separated by different levels of six-lncRNA signature, immune lncRNAs, immune-related genes, and all genes, respectively. The expression matrixes were pre-processed before PCA analysis. De-duplication via taking an average value and disregarded the data exhibited no fluctuation in expression level. Graphs displayed first, second and third major components (PC1, PC2, PC3) in three dimensions. The "limma", "scatterplot3d" packages were conducted in the analysis.

Statistics

Mean ± standard deviation (SD) was used for data expression. Two-group comparisons were analyzed by student's t-test, while multi-group comparisons were analyzed via one-way ANOVA. Spearman's correlation analysis was performed to evaluate expression correlation. Kaplan Meier analysis was carried out to analyze overall survival. *P* value < 0.05 was considered statistically significant.

Results

Construction of six-lncRNA signature

In the current study, transcriptome data of LIHC tissues (n = 374) along with normal tissues (n = 50) were acquired from the TCGA database. As procedures presented in figure 1, immune gene sets consisted of "immune response" and "immune system process" from GSEA were applied to classify the immune genes from patient's gene expression patterns. A total of 331 immune-related genes were retained for further correlational analysis. lncRNAs with |cor| > 0.4 and *P* < 0.001 were identified as candidates using R. As shown in figure 2A, connections of some immune genes with linked lncRNAs were visualized via Cytoscape software. The coefficient and attributes of positive or negative regulation were listed in Table 1. Furthermore, candidate lncRNAs were subjected to univariate cox regression and multivariate cox regression analysis. 16 lncRNAs were examined highly associated with survival outcomes (*P* < 0.0001) and 6 of them were selected as high-risk prognostic ones (figure 2B). AIC value in the prognostic optimization model was 1156.35. Meanwhile, detailed information of the six-lncRNA signature was shown in Table 2. According to the risk score, LIHC samples were divided into a high-risk group (n = 171) and a low-risk group (n = 172). Following this, relative RNA expression of the six-lncRNA signature was investigated among different groups, the results showed that six lncRNAs were statistically upregulated in LIHC tissues (figure 2C-H).

Prognostic value of lncRNA signature for assessing clinical outcome

To evaluate the capability of the six-lncRNA signature in predicting the prognosis of LIHC patients, the finial survival state and six-lncRNA expression profiles of each sample were presented in figure 3A. The scatter graph illustrated that the survival estimate of samples was growing worse along with increasing risk score. Besides, survival analysis indicated that death rates of the low-risk group were significantly lower than the high-risk group (*P* = 6.75E−11), five years survival rate was 62.2% in the low-risk group and 34.1% in the high-risk group, respectively (figure 3B). Moreover, the ROC curve clarified the better predictive accuracy of combined six-lncRNA signature compared with other clinical parameters, containing age, gender, grade, and TNM staging. The AUC of the risk score system was 0.779 (figure 3C).

Correlation between lncRNA signature and clinical characteristics

To find out the relationship of clinical indicators and outcomes of patients grouped by risk score, a stratified analysis was carried out of 373 samples from the TCGA cohort (Table 3). Furthermore, univariate analysis revealed that TNM staging and six-lncRNA signature were significantly associated with OS, multivariate analysis indicated that T-staging and six-lncRNA signature were potential

independent prognostic factors (figure 4A-B). Besides, the association of six-lncRNA expression with tumor stage and tumor grade were explored. We found that expression of AC009005.1, AC099850.3, AL031985.3, MSC−AS1 kept a close link with clinic grade, AC009005.1, AC099850.3 were related with TNM staging, and AC009005.1, AC099850.3, AL031985.3 levels were correlated with T-staging (figure 4C-E).

Gene set enrichment analysis

GSEA was applied to explore the potential mechanism between biomarkers and biological processes based on the risk score. As shown in figure 5 A-B, data from GSEA analysis testified co-expressed protein-coding genes of the high-risk group were statistically enriched in the two reference gene sets “immune response” and “immune system process” (FDR q-value < 0.25). Furthermore, the top ten highly enriched biological processes were displayed in figure 5C-L, which suggested that co-expressed genes of the six-lncRNA signature may engage in cell cycle phase transition, RNA splicing, regulation of chromosome organization, etc.

Principal component analysis

Principal component analysis (PCA) was performed for aggregation of samples by similarity in gene expression patterns from two risk groups. Utilizing reducing the dimensions of multiple indices and extracting main parameters, tumor samples were analyzed at four levels: lncRNA signature, immune lncRNAs, immune genes, and all genes. Red points on the graph represented samples of the high-risk group, while green points represented samples of the low-risk group. We found that the six-lncRNA signature could make two groups of sample data in the greatest degree of separate compared with the other three levels (figure 6A-D). These findings might to some extent confirm the prognosis accuracy of the immune lncRNA model.

Discussion

In current years, multi-omics have developed rapidly with the application of high-throughput screening. Long non-coding RNA has drawn widespread attention as a prospective molecular target for cancer diagnosis and therapeutics [36–38]. Several studies have proposed that lncRNAs were involved in the immune process contributed to tumorigenesis and progression [39–41]. For example, Jiang et al. have reported a high level of lnc-EGFR in Tregs could enhance immunosuppression in hepatoma cells via AP-1/NF-AT1signal [42]. Moreover, great progress has been made in the field of immunotherapy, administration of anti-PD-1 showed definite efficacy on virus-associated hepatocellular carcinoma [43, 44]. In this study, the potential role of immune lncRNA as prognostic or therapeutic biomarkers has been examined and may give some insights for further investigation.

We identified a prognosis immune signature based on six lncRNAs including MSC − AS1, AC009005.1, AL117336.3, AL031985.3, AL365203.2, AC099850.3 and verified its prognostic value for HCC patients from TCGA dataset. In the present study, patients were assigned into low-risk and high-risk groups based

on the risk score of the above-mentioned lncRNA. Univariate cox analysis and multivariate cox analysis were performed to evaluate the availability of these biomarkers as independent risk factors. Hepatocellular carcinoma was classified based on the TNM classification of the American Joint Committee on Cancer (AJCC) [45]. The ROC results demonstrated a higher AUC value more than that of TNM classification and other indices, which further ascertain the validity of the proposed risk system to predict overall survival independently.

In clinical practice, Alpha-fetoprotein (AFP) is widely used to diagnose liver cancer and used to evaluate therapeutic efficacy [46–48]. In a healthy human body, AFP expresses at a low level, since it is synthesized by the fetal liver and down-regulated after birth [49]. While about half of liver malignancy is detected with radical changes of AFP serological level [50]. A concentration of more than 400 ng/ml may indicate the existence of liver cancer [51]. However, increased AFP concentrations have also been described in other diseases, such as cerebellar ataxia and testicular cancer [52, 53]. In present stratification analysis, samples with AFP level more than 20 ng/ml (pathological threshold) roughly concentrated in the high-risk group, these results inspired us that maybe the serum level of six-lncRNA signature could cooperate with AFP for early diagnosis, efficacy assessment, and prognostic evaluation, serve as a more sensitive circulating biomarker.

In recent years, great efforts have been made on the study of lncRNAs in tumor immunology. Research on uncovering the mechanism of lncRNA regulates immune response has developed a lot [54]. For example, lncRNA NKILA plays a crucial role in promoting immune evasion via associating STAT1 and NF-κB signaling in several tumors [55, 56]. LncRNA ID2-AS1 suppresses HCC metastasis by activating the HDAC8/ID2 pathway [57]. Herein, six prognosis biomarkers were highly expressed in tumor samples and data from GSEA unveiled that corresponding genes were enriched in a series of GO biological process pathways, for instance, cell cycle, RNA splicing, DNA repair, and chromosome regulation. However, the concrete roles of these lncRNAs in regulating biological pathways are needed for further investigation.

Conclusions

In conclusion, we developed an immune model including six lncRNAs for predicting the outcome of LIHC samples. Based on bioinformatics analysis, the function and effectiveness of this lncRNA signature were analyzed in our study. Data from findings indicated this signature may provide a new approach for accurate diagnosis and prognosis. However, clinical trials and functional tests are required to explore the underlying mechanism before clinical application.

Declarations

Acknowledgements

Not applicable.

Authors' contributions

Rui Kong and Nan Wang did data analyses. Rui Kong and Yuejuan Zheng wrote and revised the manuscript. Jie Lu, Rui Kong and Nan Wang designed the study. All authors read and approved the final manuscript.

Author details

1. Department of Gastroenterology, Shanghai Tenth People's Hospital, Tongji University School of Medicine, Shanghai 200072, China.
2. Department of Immunology and Microbiology, School of Basic Medical Sciences, Shanghai University of Traditional Chinese Medicine, Shanghai, China.

Funding

This study received support from National Natural Science Foundation of China (81471537) and the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Education.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no conflicts of interest.

Abbreviations

LIHC: liver hepatocellular carcinoma, TCGA: The Cancer genome atlas, lncRNA: long non-coding RNA, AUC: area under curve, ROC: receiver operating characteristic, PCA: principal component analysis, GSEA: gene set enrichment analysis, ACT: adoptive cellular immunotherapy, PD-1: programmed cell death-1, CTLA4: cytotoxic T-lymphocyte antigen-4, FPKM: fragments per kilobase of exon model per million mapped fragments, OS: overall survival, AIC: akaike information criterion, BMI: body mass index, AFP: alpha-fetoprotein, AJCC: American Joint Committee on Cancer, GO: gene ontology.

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F: Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *INT J CANCER* 2015, 136(5):E359-E386.

2. Llovet JM, Zucman-Rossi J, Pikarsky E, Sangro B, Schwartz M, Sherman M, Gores G: Hepatocellular carcinoma. *NAT REV DIS PRIMERS* 2016, 2:16018.
3. Grandhi MS, Kim AK, Ronnekleiv-Kelly SM, Kamel IR, Ghasebeh MA, Pawlik TM: Hepatocellular carcinoma: From diagnosis to treatment. *SURG ONCOL* 2016, 25(2):74-85.
4. Poon D, Anderson BO, Chen LT, Tanaka K, Lau WY, Van Cutsem E, Singh H, Chow WC, Ooi LL, Chow P *et al*: Management of hepatocellular carcinoma in Asia: consensus statement from the Asian Oncology Summit 2009. *LANCET ONCOL* 2009, 10(11):1111-1118.
5. Forner A, Reig M, Bruix J: Hepatocellular carcinoma. *LANCET* 2018, 391(10127):1301-1314.
6. Li S, Yang F, Ren X: Immunotherapy for hepatocellular carcinoma. *Drug Discov Ther* 2015, 9(5):363-371.
7. Wu J, Meng Z, Jiang M, Zhang E, Trippler M, Broering R, Bucchi A, Krux F, Dittmer U, Yang D *et al*: Toll-like receptor-induced innate immune responses in non-parenchymal liver cells are cell type-specific. *IMMUNOLOGY* 2010, 129(3):363-374.
8. Pinna F, Bissinger M, Beuke K, Huber N, Longerich T, Kummer U, Schirmacher P, Sahle S, Breuhahn K: A20/TNFAIP3 Discriminates Tumor Necrosis Factor (TNF)-Induced NF-kappaB from JNK Pathway Activation in Hepatocytes. *FRONT PHYSIOL* 2017, 8:610.
9. Lee JH, Lee JH, Lim YS, Yeon JE, Song TJ, Yu SJ, Gwak GY, Kim KM, Kim YJ, Lee JW *et al*: Adjuvant immunotherapy with autologous cytokine-induced killer cells for hepatocellular carcinoma. *GASTROENTEROLOGY* 2015, 148(7):1383-1391.
10. Greten TF, Wang XW, Korangy F: Current concepts of immune based treatments for patients with HCC: from basic science to novel treatment approaches. *GUT* 2015, 64(5):842-848.
11. Rosenberg SA, Restifo NP: Adoptive cell transfer as personalized immunotherapy for human cancer. *SCIENCE* 2015, 348(6230):62-68.
12. Takayama T, Sekine T, Makuuchi M, Yamasaki S, Kosuge T, Yamamoto J, Shimada K, Sakamoto M, Hirohashi S, Ohashi Y *et al*: Adoptive immunotherapy to lower postsurgical recurrence rates of hepatocellular carcinoma: a randomised trial. *LANCET* 2000, 356(9232):802-807.
13. Ringelhan M, Pfister D, O'Connor T, Pikarsky E, Heikenwalder M: The immunology of hepatocellular carcinoma. *NAT IMMUNOL* 2018, 19(3):222-232.
14. Kim HJ, Park S, Kim KJ, Seong J: Clinical significance of soluble programmed cell death ligand-1 (sPD-L1) in hepatocellular carcinoma patients treated with radiotherapy. *RADIOTHER ONCOL* 2018, 129(1):130-135.
15. Sangro B, Gomez-Martin C, de la Mata M, Inarrairaegui M, Garralda E, Barrera P, Riezu-Boj JI, Larrea E, Alfaro C, Sarobe P *et al*: A clinical trial of CTLA-4 blockade with tremelimumab in patients with hepatocellular carcinoma and chronic hepatitis C. *J HEPATOL* 2013, 59(1):81-88.
16. Li CH, Chen Y: Targeting long non-coding RNAs in cancers: progress and prospects. *Int J Biochem Cell Biol* 2013, 45(8):1895-1910.

17. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F *et al*: Landscape of transcription in human cells. *NATURE* 2012, 489(7414):101-108.
18. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG *et al*: The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *GENOME RES* 2012, 22(9):1775-1789.
19. Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q *et al*: Long noncoding RNAs with enhancer-like function in human cells. *CELL* 2010, 143(1):46-58.
20. Holoch D, Moazed D: RNA-mediated epigenetic regulation of gene expression. *NAT REV GENET* 2015, 16(2):71-84.
21. Clark MB, Mattick JS: Long noncoding RNAs in cell biology. *SEMIN CELL DEV BIOL* 2011, 22(4):366-376.
22. Kung JT, Colognori D, Lee JT: Long noncoding RNAs: past, present, and future. *GENETICS* 2013, 193(3):651-669.
23. Heward JA, Lindsay MA: Long non-coding RNAs in the regulation of the immune response. *TRENDS IMMUNOL* 2014, 35(9):408-419.
24. Peng X, Gralinski L, Armour CD, Ferris MT, Thomas MJ, Proll S, Bradel-Tretheway BG, Korth MJ, Castle JC, Biery MC *et al*: Unique signatures of long noncoding RNA expression in response to virus infection and altered innate immune signaling. *MBIO* 2010, 1(5).
25. Gibb EA, Brown CJ, Lam WL: The functional role of long non-coding RNA in human carcinomas. *MOL CANCER* 2011, 10:38.
26. Peng WX, Koirala P, Mo YY: LncRNA-mediated regulation of cell signaling in cancer. *ONCOGENE* 2017, 36(41):5661-5667.
27. Cheetham SW, Gruhl F, Mattick JS, Dinger ME: Long noncoding RNAs and the genetics of cancer. *Br J Cancer* 2013, 108(12):2419-2425.
28. Wang BG, Lv Z, Ding HX, Fang XX, Wen J, Xu Q, Yuan Y: The association of lncRNA-HULC polymorphisms with hepatocellular cancer risk and prognosis. *GENE* 2018, 670:148-154.
29. Wang Y, Chen F, Zhao M, Yang Z, Li J, Zhang S, Zhang W, Ye L, Zhang X: The long noncoding RNA HULC promotes liver cancer by increasing the expression of the HMGA2 oncogene via sequestration of the microRNA-186. *J BIOL CHEM* 2017, 292(37):15395-15407.
30. Zhang J, Han C, Ungerleider N, Chen W, Song K, Wang Y, Kwon H, Ma W, Wu T: A Transforming Growth Factor-beta and H19 Signaling Axis in Tumor-Initiating Hepatocytes That Regulates Hepatic Carcinogenesis. *HEPATOLOGY* 2019, 69(4):1549-1563.
31. The Cancer genome atlas. https://portal.gdc.cancer.gov/. Accessed 26 October 2019.
32. Ensemble website. http://asia.ensembl.org/index.html. Accessed 26 October 2019.
33. Gene Set Enrichment Analysis. http://software.broadinstitute.org/gsea/msigdb/index.jsp. Accessed 02 November 2019.

34. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: Cytoscape: a software environment for integrated models of biomolecular interaction networks. *GENOME RES* 2003, 13(11):2498-2504.
35. Wang C, Gu Y, Zhang K, Xie K, Zhu M, Dai N, Jiang Y, Guo X, Liu M, Dai J *et al*: Systematic identification of genes with a cancer-testis expression pattern in 19 cancer types. *NAT COMMUN* 2016, 7:10499.
36. Li S, Zhang M, Zhang H, Hu K, Cai C, Wang J, Shi L, Ma P, Xu Y, Zheng P: Exosomal long noncoding RNA lnc-GNAQ-6:1 may serve as a diagnostic marker for gastric cancer. *CLIN CHIM ACTA*
37. Elsayed ET, Salem PE, Darwish AM, Fayed HM: Plasma long non-coding RNA HOTAIR as a potential biomarker for gastric cancer. *Int J Biol Markers* 2018:349275596.
38. Vishnubalaji R, Hibah S, Elango R, Alajez NM: Noncoding RNAs as potential mediators of resistance to cancer immunotherapy. *SEMIN CANCER BIOL*
39. Yang X, Xie X, Xiao YF, Xie R, Hu CJ, Tang B, Li BS, Yang SM: The emergence of long non-coding RNAs in the tumorigenesis of hepatocellular carcinoma. *CANCER LETT* 2015, 360(2):119-124.
40. Ni W, Zhang Y, Zhan Z, Ye F, Liang Y, Huang J, Chen K, Chen L, Ding Y: A novel lncRNA uc.134 represses hepatocellular carcinoma progression by inhibiting CUL4A-mediated ubiquitination of LATS1. *J HEMATOL ONCOL* 2017, 10(1):91.
41. Zhang D, Sun G, Zhang H, Tian J, Li Y: Long non-coding RNA ANRIL indicates a poor prognosis of cervical cancer and promotes carcinogenesis via PI3K/Akt pathways. *BIOMED PHARMACOTHER* 2017, 85:511-516.
42. Jiang R, Tang J, Chen Y, Deng L, Ji J, Xie Y, Wang K, Jia W, Chu WM, Sun B: The long noncoding RNA lnc-EGFR stimulates T-regulatory cells differentiation thus promoting hepatocellular carcinoma immune evasion. *NAT COMMUN* 2017, 8:15129.
43. Liu X, Li M, Wang X, Dang Z, Jiang Y, Wang X, Kong Y, Yang Z: PD-1(+) TIGIT (+) CD8(+) T cells are associated with pathogenesis and progression of patients with hepatitis B virus-related hepatocellular carcinoma. *Cancer Immunol Immunother*
44. Pauken KE, Wherry EJ: Overcoming T cell exhaustion in infection and cancer. *TRENDS IMMUNOL* 2015, 36(4):265-276.
45. Matos LL, Dedivitis RA, Kulcsar M, de Mello ES, Alves V, Cernea CR: External validation of the AJCC Cancer Staging Manual, 8th edition, in an independent cohort of oral cancer patients. *ORAL ONCOL* 2017, 71:47-53.
46. Schieving JH, de Vries M, van Vugt JM, Weemaes C, van Deuren M, Nicolai J, Wevers RA, Willemsen MA: Alpha-fetoprotein, a fascinating protein and biomarker in neurology. *Eur J Paediatr Neurol* 2014, 18(3):243-248.
47. Song PP, Xia JF, Inagaki Y, Hasegawa K, Sakamoto Y, Kokudo N, Tang W: Controversies regarding and perspectives on clinical utility of biomarkers in hepatocellular carcinoma. *World J Gastroenterol* 2016, 22(1):262-274.

48. Tzartzeva K, Obi J, Rich NE, Parikh ND, Marrero JA, Yopp A, Waljee AK, Singal AG: Surveillance Imaging and Alpha Fetoprotein for Early Detection of Hepatocellular Carcinoma in Patients with Cirrhosis: A Meta-analysis. *GASTROENTEROLOGY* 2018, 154(6):1706-1718.
49. Galle PR, Foerster F, Kudo M, Chan SL, Llovet JM, Qin S, Schelman WR, Chintharlapalli S, Abada PB, Sherman M *et al*: Biology and significance of alpha-fetoprotein in hepatocellular carcinoma. *LIVER INT*
50. Van Hees S, Michielsen P, Vanwolleghem T: Circulating predictive and diagnostic biomarkers for hepatitis B virus-associated hepatocellular carcinoma. *World J Gastroenterol* 2016, 22(37):8271-8282.
51. Sterling RK, Wright EC, Morgan TR, Seeff LB, Hoefs JC, Di Bisceglie AM, Dienstag JL, Lok AS: Frequency of elevated hepatocellular carcinoma (HCC) biomarkers in patients with advanced hepatitis C. *AM J GASTROENTEROL* 2012, 107(1):64-74.
52. Nanetti L, Cavalieri S, Pensato V, Erbetta A, Pareyson D, Panzeri M, Zorzi G, Antozzi C, Moroni I, Gellera C *et al*: SETX mutations are a frequent genetic cause of juvenile and adult onset cerebellar ataxia with neuropathy and elevated serum alpha-fetoprotein. *ORPHANET J RARE DIS* 2013, 8:123.
53. Wymer KM, Daneshmand S, Pierorazio PM, Pearce SM, Harris KT, Eggener SE: Mildly elevated serum alpha-fetoprotein (AFP) among patients with testicular cancer may not be associated with residual cancer or need for treatment. *ANN ONCOL* 2017, 28(4):899-902.
54. Sun W, Yang Y, Xu C, Guo J: Regulatory mechanisms of long noncoding RNAs on gene expression in cancers. *Cancer Genet* 2017, 216-217:105-110.
55. Huang D, Chen J, Yang L, Ouyang Q, Li J, Lao L, Zhao J, Liu J, Lu Y, Xing Y *et al*: NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death. *NAT IMMUNOL* 2018, 19(10):1112-1125.
56. Wu W, Chen F, Cui X, Yang L, Chen J, Zhao J, Huang D, Liu J, Yang L, Zeng J *et al*: LncRNA NKILA suppresses TGF-beta-induced epithelial-mesenchymal transition by blocking NF-κB signaling in breast cancer. *INT J CANCER* 2018, 143(9):2213-2224.
57. Zhou Y, Huan L, Wu Y, Bao C, Chen B, Wang L, Huang S, Liang L, He X: LncRNA ID2-AS1 suppresses tumor metastasis by activating the HDAC8/ID2 pathway in hepatocellular carcinoma. *CANCER LETT* 2019.

Tables

Table 1. Correlation of immune genes and associated lncRNAs

Immune Gene	lncRNA	Cor	*P* value	Regulation
GLMN	AC007541.1	0.445	1.35E-19	positive
KMT2A	AP001318.2	0.460	5.45E-21	positive
NCOA6	AC012510.1	0.462	3.55E-21	positive
TRAF2	SREBF2-AS1	0.451	3.66E-20	positive
RPS19	AC132192.2	0.614	4.03E-40	positive
TRAF2	AL035446.1	0.463	3.12E-21	positive
HELLS	AL360181.2	0.455	1.55E-20	positive
PRKRA	AP002884.1	0.416	4.67E-17	positive
RPS19	AL109615.3	-0.401	7.64*10E-16	negative
APOA2	FLJ42351	-0.460	6.18*10E-21	negative

Table 2. Detailed information of six-lncRNA signature

Gene symbol	Gene_ID	Location	AIC
MSC-AS	ENSG00000235531.8	chr8: 71828167-72002405	1160.1
AC009005.1	ENSG00000267751.4	chr19: 567212-571745	1159.8
AL117336.3	ENSG00000271335.4	chr10: 35314552-35320998	1157.0
AL031985.3	ENSG00000260920.2	chr1: 40464319-40466767	1157.8
AL365203.2	ENSG00000273038.2	chr10:32887255-32889311	1156.9
AC099850.3	ENSG00000265415.1	chr17:59202677-59203829	1156.7

Table 3. Relationship between risk score of lncRNA signature for OS and clinical features

	low risk/high risk	Pearson $\chi 2$	P
Age		0.072	0.788
> = 55	119/116		
< 55	53/55		
Gender		3.565	0.059
female	47/63		
male	125/108		
TNM stage		8.404	0.004
I/II	130/108		
III/IV	30/53		
G		12.032	0.001
G1/G2	123/91		
G3/G4	47/77		
AFP (ng/mL)		5.917	0.015
> = 20	58/94		
< 20	64/57		
BMI		0.213	0.645
> =25	80/73		
< 25	81/82		
Race		0.268	0.605
white	83/86		
Asian	77/71		

Figures

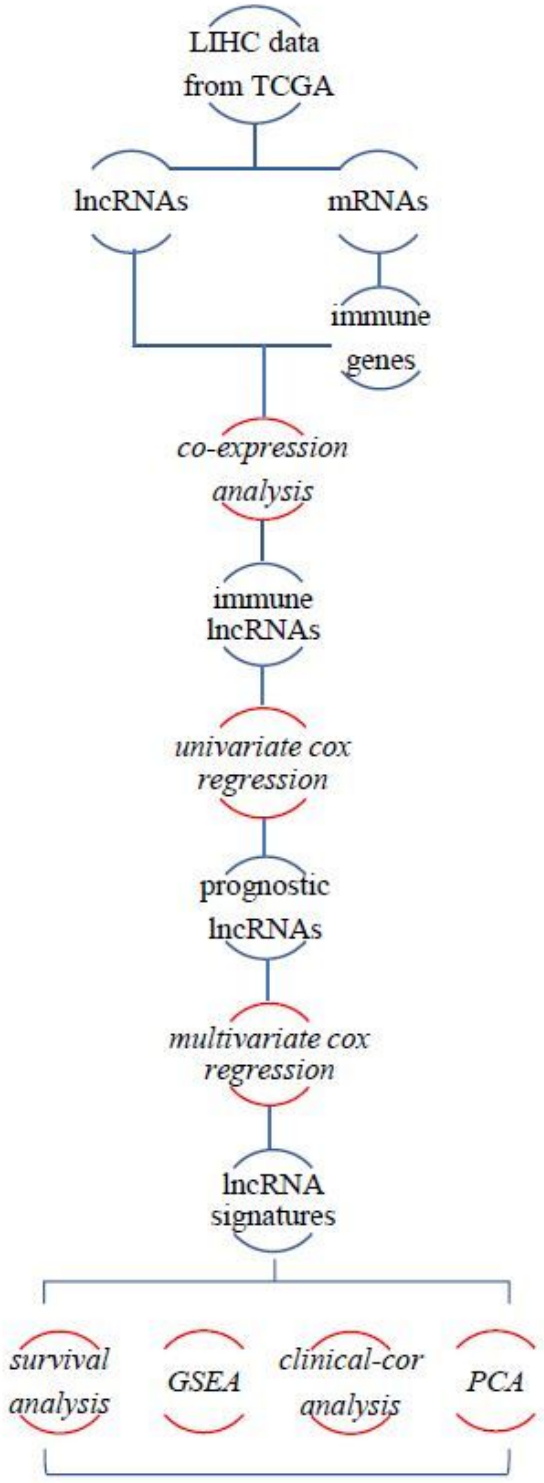

Figure 1

Flowchart showing the steps of constructing a prognostic lncRNA signature.

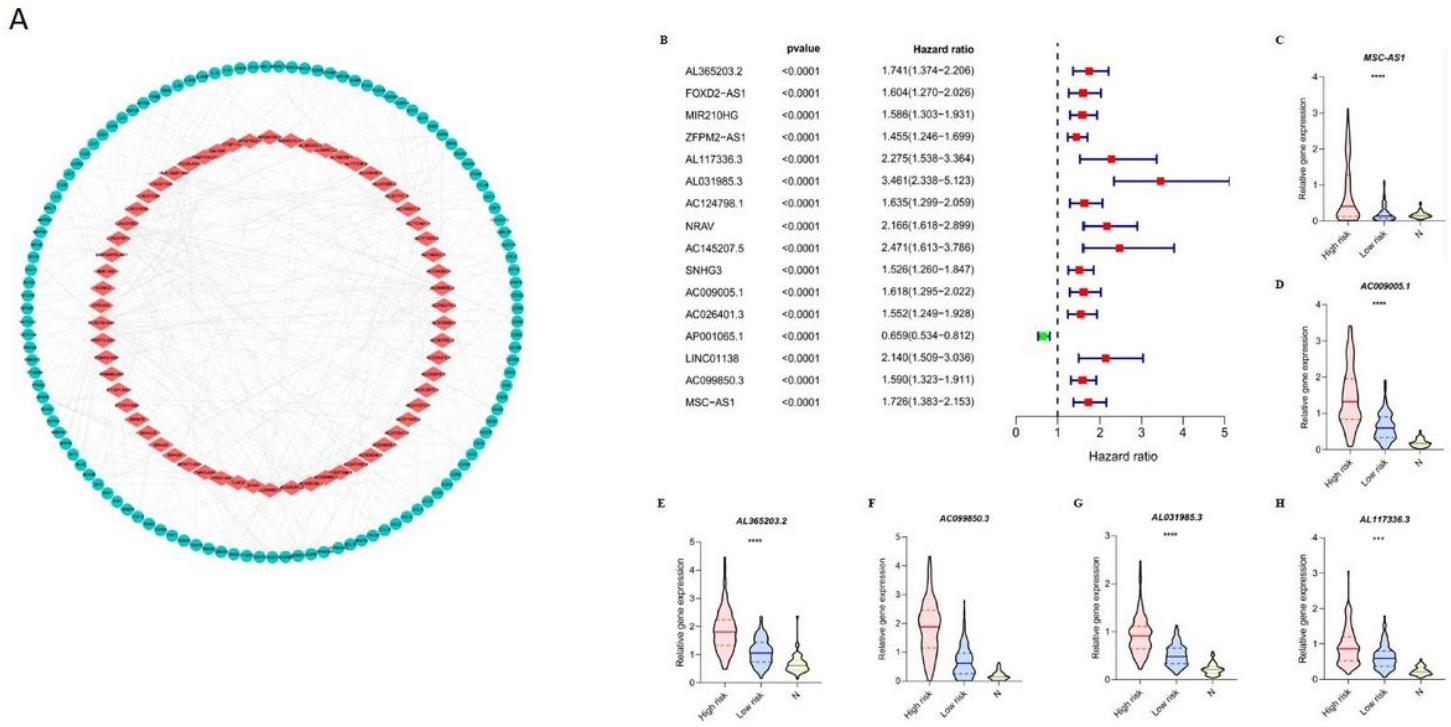

Figure 2

Construction of an immune lncRNA signature for liver hepatocellular carcinoma A. The network of partial immune genes and associated lncRNAs. B. Candidate lncRNA for the prognostic model with information about hazard ratio. C. Relative gene expression of MSC-AS1 among the low-risk group, high-risk group, and non-tumor samples. D. Relative gene expression of AC009005.1 among the low-risk group, high-risk group, and non-tumor samples. E. Relative gene expression of AL117336.3 among the low-risk group, high-risk group, and non-tumor samples. F. Relative gene expression of AL031985.3 among the low-risk group, high-risk group, and non-tumor samples. G. Relative gene expression of AL365203.2 among the low-risk group, high-risk group, and non-tumor samples. H. Relative gene expression of AC099850.3 among the low-risk group, high-risk group, and non-tumor samples.

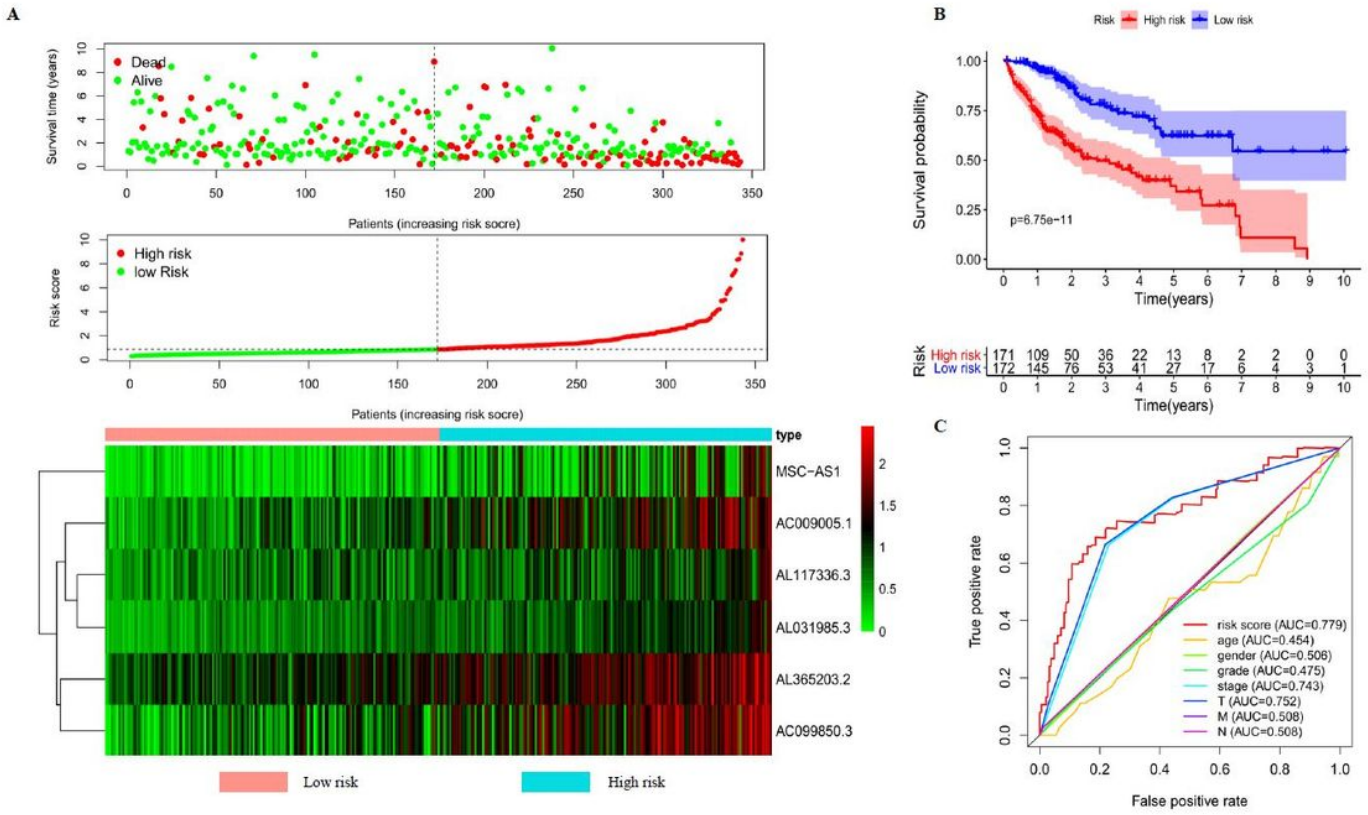

Figure 3

Validation of the prognostic lncRNA signature for hepatocellular carcinoma A. The upper graph shows the relationship between survival time and risk score, the medium graph shows the distribution of lncRNA risk score, the bottom heat map shows expression patterns of six-lncRNA signature for LIHC patients. B. The Kaplan-Meier curve of different tumor groups based on the median risk score. C. The ROC curve for the risk score, age, gender, grade, and TNM stage.

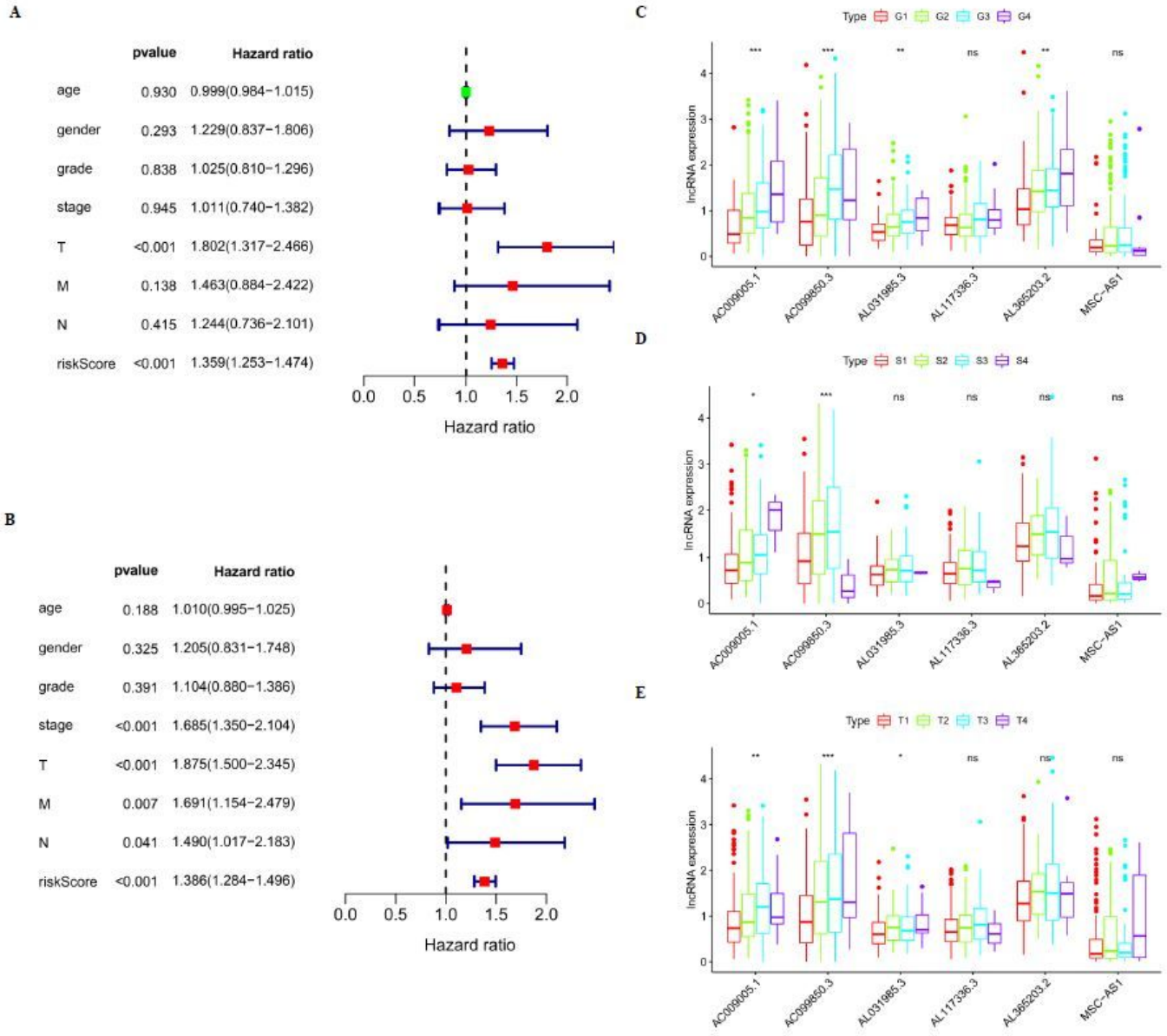

Figure 4

Evaluation of the six-lncRNA as an independent factor. A. The tree diagram shows the statistical significance and hazard ratios of several indices as prognostic factors. B. The tree diagram shows the statistical significance and hazard ratios of several indices as prognostic factors. C. Boxplot indicates the correlation of lncRNA biomarkers' expression and tumor grade. D. Boxplot indicates the correlation of lncRNA biomarkers expression and TNM stage. E. Boxplot indicates the correlation of lncRNA biomarkers expression and T-staging.

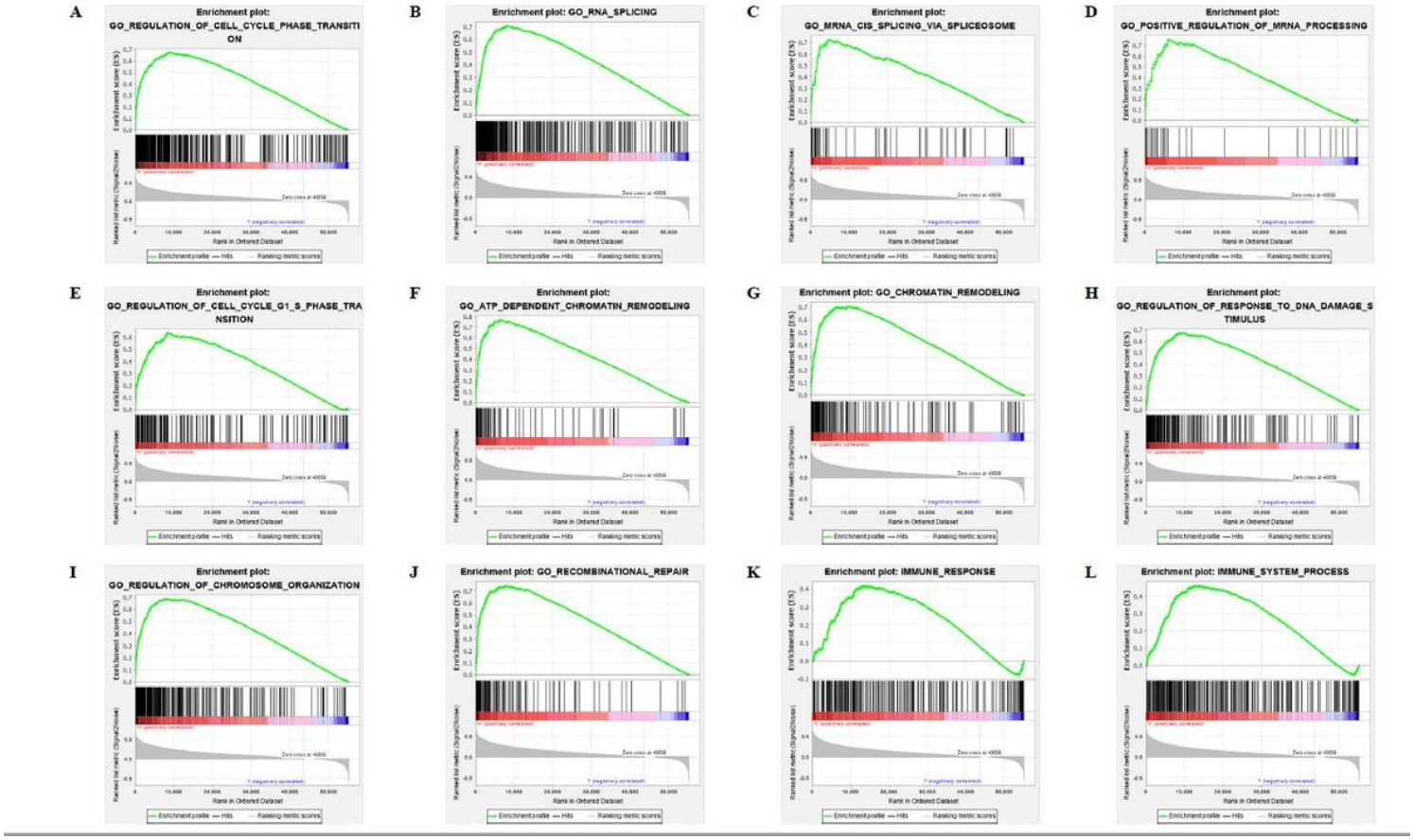

Figure 5

Highly enriched biological pathways for corresponding immune genes of six-lncRNA signature in TCGA. (A) Regulation of cell cycle phase transition (FDR < 0.007), (B) RNA splicing (FDR < 0.004), (C) mRNA CIS splicing via spliceosome (FDR < 0.002) (D) Positive regulation of mRNA processing (FDR < 0.002) (E) Regulation of cell cycle G1 S phase transition (FDR < 0.002) (F) ATP dependent chromatin remodeling (FDR < 0.002) (G) Chromatin remodeling (FDR < 0.002) (H) Regulation of response to DNA damage stimulus (FDR < 0.002) (I) Regulation of chromosome organization (FDR < 0.002) (J) Recombinational repair (FDR < 0.001) (K) Immune response (FDR < 0.203) (L) Immune system process (FDR < 0.087).

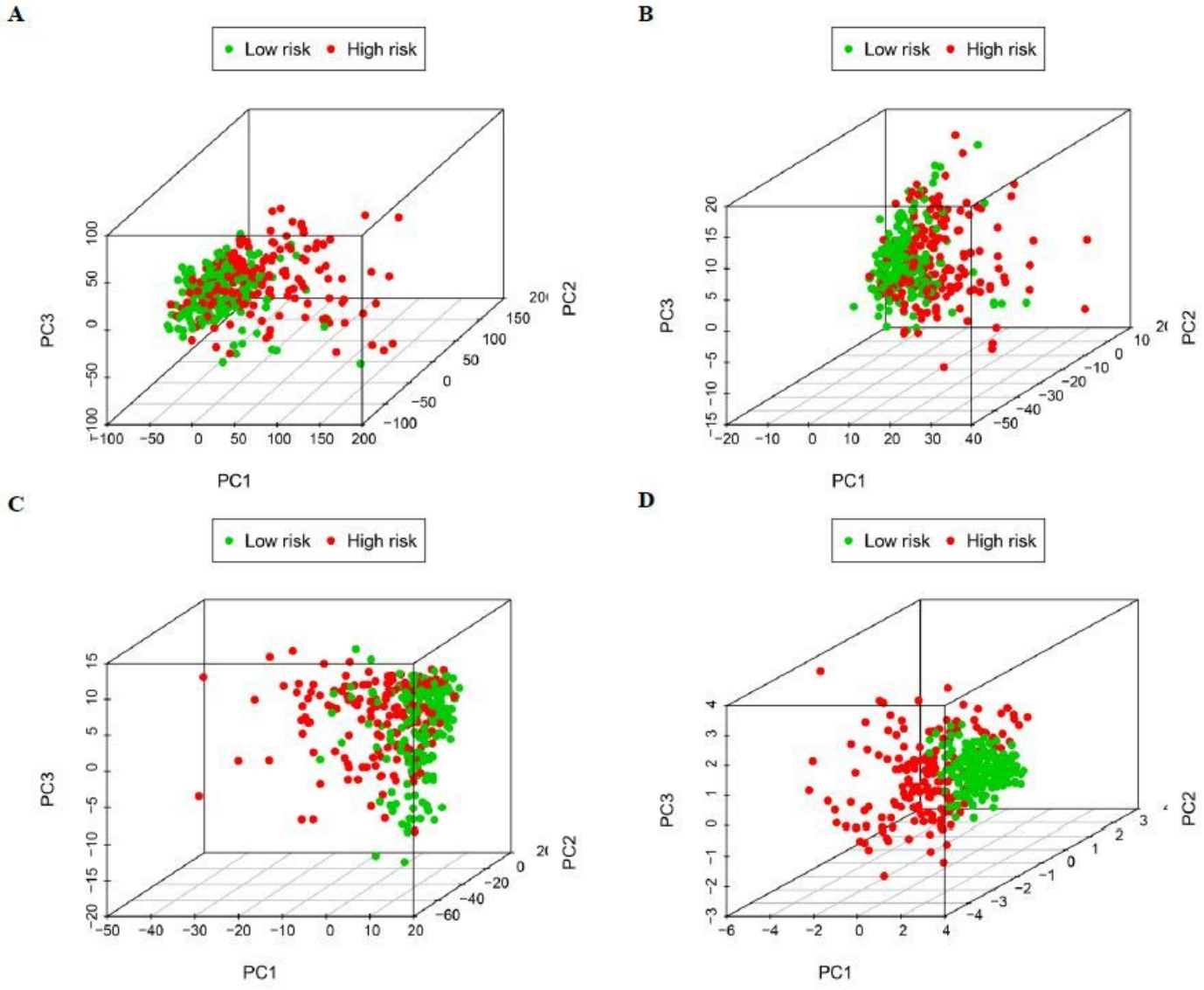

Figure 6

Principal component analysis of samples for TCGA. A. PCA shows samples divisibility from the low risk- and high-risk group based on all gene expression. B. PCA shows samples divisibility from the low risk- and high-risk group based on immune gene expression. C. PCA shows samples divisibility from the low risk- and high-risk group based on immune lncRNA expression. D. PCA shows samples divisibility from the low risk- and high-risk group based on lncRNA signature expression.