Identification of Lipid Metabolism-Related Genes as Prognostic Indicators in Papillary Thyroid Cancer

Shi-Shuai Wen
Fudan University Shanghai Cancer Center
Yi Luo
Fudan University Shanghai Cancer Center
Wei-Li Wu
Wenzhou Medical University
Ting Zhang
Fudan University Shanghai Cancer Center
Yi-Chen Yang
Fudan University Shanghai Cancer Center
Qing-Hai Ji
Fudan University Shanghai Cancer Center
Yi-Jun Wu
Zhejiang University
Rong-Liang Shi
Fudan University Shanghai Cancer Center
Ben Ma
Fudan University Shanghai Cancer Center
Midie Xu (xumd27202003@sina.com)
Fudan University Shanghai Cancer Center
Ning Qu
Fudan University Shanghai Cancer Center

Research Article

Keywords: papillary thyroid cancer, lipid metabolism, signature, risk score, prognosis

Posted Date: March 18th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-307303/v1

Version of Record: A version of this preprint was published at Acta Biochimica et Biophysica Sinica on October 23rd, 2021. See the published version at https://doi.org/10.1093/abbs/gmab145.

Abstract

Purpose

Lipid metabolism plays important roles not only in the structural basis and energy supply of healthy cells but also in the oncogenesis and progression of cancer. In this study, we investigate the prognostic value of lipid metabolism related genes in papillary thyroid cancer (PTC).

Methods

The in time to recurrence predictive gene signature was developed, internally and externally validated based on PTC datasets including The Cancer Genome Atlas (TCGA) and GSE33630 datasets. Univariate, LASSO and multivariate Cox regression analysis were applied to assess prognostic genes and build the prognostic gene signature. The expression profiles of prognostic genes were further determined by immunohistochemistry by using in-house cohorts which enrolled 97 patients. Kaplan-Meier curve, time-dependent receiver operating characteristic curve, nomogram and decision curve analysis were used to assess the performance of the gene signature.

Results

We identified four recurrence-related genes, PDZK1IP1, TMC3, LRP2 and KCNJ13, and established a 4-gene signature recurrence risk model. The expression profile of the 4 genes in the TCGA and in-house cohort indicated that stage T1/T2 PTC and locally advanced PTC exhibited notable associations not only with clinicopathological parameters but also with recurrence. Calibration analysis plots indicated the excellent predictive performance of the prognostic nomogram constructed based on the gene signature. GSEA showed that high-risk cases exhibited changes in several important tumorigenesis-related pathways, such as the intestinal immune network and the p53 and Hedgehog signalling pathways.

Conclusion

Our findings indicate that lipid metabolism-related gene profiling represents a potential marker for prognosis and treatment decisions for PTC patients.

Introduction

Metabolic reprogramming has been recognized as one of the ten new hallmarks of cancer cells. Increased glycolysis (Warburg effect) under normoxic conditions and glutamine metabolism have been determined to be among the main characteristics of human malignancies (1). Recently, accumulating evidence raised the possibility that dysregulation of lipid metabolism may also be one of the key metabolic markers of cancer cells(2). Cancer cells with high proliferation rates can obtain more lipids

through enhanced lipid uptake, lipolysis, and *de novo* fatty acid synthesis. Lipids play important roles not only in cell structure but also in energy supply and signal transfer(3). Papillary thyroid cancer (PTC) is a common malignant tumour that is increasing notably worldwide(4). Although most PTCs have good prognoses, approximately 10% of patients do not respond favourably to radioactive iodine therapy, and poorly differentiated or undifferentiated tumours are more likely to be associated with disease recurrence and death(5, 6). Thus, it is important to identify factors that may predict the prognosis such that we can distinguish the patients with higher recurrence risk and perform targeted treatment of these patients. Previously, we reported that lysine methyltransferase 5A (KMT5A), a lipid-modulating gene, could affect the oncogenesis of PTC *in vitro*(5, 6). Other studies also show the alteration of fatty acid metabolism influences energy storage, affects drug resistance, modulates cell proliferation and survival and stimulates the extracellular environment(7, 8).

Previously, we performed a transcriptome analysis of lipid metabolism-related genes (LMRGs) in PTC and constructed two subtypes with significant differences in tumour immunity and matrix microenvironment and patients' outcomes. Moreover, we identified 182 key genes that were functional clustered in some pivotal PTC related pathways(9). In this present study, we try to determine whether these 182 genes can be used to develop a gene signature that can predict the risk of recurrence in PTC patients. We devised a lipid metabolism-related 4-gene signature for evaluating the risk of recurrence. The risk model was demonstrated to precisely evaluate the recurrence risk in PTC patients and was subsequently verified to possess a satisfactory prediction ability.

Methods And Materials

Collection of datasets and patient information

A total of 497 PTCs with RNA-seq data and clinical information from the *TCGA* thyroid carcinoma (THCA) dataset (all PTCs) were collected. Gene expression profiles of PTC (GSE33630) in the GEO database were downloaded from NCBI (http://www.ncbi.nlm.nih.gov/geo/). This dataset contains 60 PTC samples and 45 normal samples. The *TCGA* cohort was divided randomly into a training cohort and a validation cohort with an equal number of cases. To avoid random allocation bias and ensure that the distribution of age and clinical stage of randomly selected samples was consistent with all samples, each sample was returned 100 times in advance. Patient information in the training set is shown in **SupplementaryTable S1**. The other 50% of samples in the TCGA PTC dataset and the GSE33630 dataset were used as internal and external validation cohorts, respectively. This study was approved by the ethics committee of Fudan University Shanghai Cancer Center.

Construction of the recurrence risk model

For each differentially coexpressed gene and associated survival data, a univariate Cox proportional hazard regression model was constructed using the R package 'survival coxph function', and log rank $P < 0.05$ was selected as the threshold. To narrow the gene range and maintain accuracy, Lasso Cox regression analysis was performed by using the R package 'glmnet' to identify the signature gene.

Multivariate Cox survival analysis was performed to construct the recurrence risk model. The R package ‘timeROC’ was used to perform an ROC analysis of the recurrence classification of the risk score. Next, the risk score was z-scored, samples with risk scores > 0 were classified into the high-risk group, and those with risk scores < 0 were classified into the low-risk group.

TMA construction and patient information

A total of 127 formalin-fixed, paraffin-embedded tissue samples were obtained from 97 patients who underwent primary surgery in the Department of Head and Neck Surgery of Fudan University Shanghai Cancer Center (FUSCC) between January 2010 and January 2015. A total of 127 samples were divided into three groups: 45 histologically proven PTCs in T1/T2 stage, 48 locally advanced PTCs and 34 normal thyroid tissues (29 of these tissues are parathyroid tissues of T1/T2 PTCs, and 5 are parathyroid tissues of LA-PTCs). Stage T1/T2 PTC was selected according to the eighth edition of the AJCC staging system. Locally advanced PTC (LA-PTC) is defined as PTC with extension of muscles, blood vessels, nerves, oesophagus and trachea. The follow-up was performed before TMA construction. There were 5 patients with recurrence in the T1/T2 PTC group and 14 patients with recurrence in the LA-PTC group. The sample information is presented in **SupplementaryTable S1**. Stained tumour sections were reviewed by one pathologist to define diagnostic areas. Two 1.5-mm-sized core samples were obtained from the representative areas of the samples and later embedded in microarray slices. None of the patients received any chemotherapy or radiotherapy before surgery. Tissue samples were collected from the Tissue Bank of FUSCC, which is allowed to perform translational research with the approval of the Ethics Committee and Institutional Review Board of FUSCC in compliance with ethics standards and patient confidentiality.

Immunohistochemical staining and evaluation

The TMA were deparaffinized and hydrated. Citric acid antigen retrieval buffer (G1202, Servicebio, China) was used to perform antigen retrieval. Hydrogen peroxide (0.3%, 10011218, Sinopharm Chemical Reagent Co., Ltd.) was used to block endogenous peroxidase activity. After blocking with 3% BSA (A8020, Solarbio) at room temperature for 30 minutes, the TMA was incubated with primary antibody at 4°C overnight (PDZK1IP1, dilution of 1:150, Ab156014, Abcam; LRP2, dilution of 1:500, Ab76969, Abcam; KCNJ13, dilution of 1:100, 12657-1-AP, Proteintech; TMC3, dilution of 1:4000, Ab243537, Abcam) followed by an HRP-conjugated goat anti-rabbit secondary antibody (GB23303, Servicebio). The TMA was subsequently detected by DAB and counterstained with Harris haematoxylin.

The mean optical density (MOD) and integral optical density (IOD) are widely used in semiquantitative analysis. Considering that the processing condition of immunohistochemical staining of each TMA is the same, while the positive staining area of each core in the TMA is different, MOD was used to indicate the expression level of the protein in this study. A Pannoramic scanner (Pannoramic MIDI, 3DHISTECH Ltd.) was used to scan TMAs. 3D Histech Quant Center2.1 analysis software (3D Histech, Hungary) was used to calculate the MOD. The PatternQuant module in the Quant Center2.1 analysis software was used to distinguish the brown area (positive area) and the blue area (negative area), and the area and greyscale

of the positive part were subsequently quantified. IOD was calculated from total greyscale. The final indicator MOD was obtained by dividing the IOD by the positive area.

Bioinformatic analysis

Single-sample gene set enrichment analysis (GSEA) was applied to identify the relationship between the risk scores of different samples and biological functions using the R package 'GSVA'.

Statistical analysis

Kaplan-Meier curves were applied to assess the difference in time to recurrence (TTR) between different groups. Univariate and multivariate Cox regression analyses were performed to assess the independent prognostic factors. A receiver operating characteristic (ROC) curve was used to analyse the sensitivity and specificity of the prediction of survival by the gene signature. The association between the risk score and clinicopathological parameters and the association between the expression of the four genes and clinicopathological parameters were analysed by the chi-square test and the t-test. Comparison of MOD displayed by TMA immunohistochemical staining among stage T1/T2 PTC, locally advanced PTC and normal thyroid tissues was performed using the Wilcoxon rank-sum test. Comparison of MOD between patients with recurrence and patients without recurrence was performed using the Wilcoxon rank-sum test and the t-test. Nomogram, a statistical method that can present the influence of each variable on the outcome using the length of a straight line, and the effect of different values of each variable on the outcome (10), was established using the R package RMS to predict the survival rate of patients. Data preparation, statistical analyses and heatmaps were performed using SPSS for Windows (version 22.0; IBM Corp., Armonk, NY), R software (version 3.5.1; R Foundation for Statistical Computing, Vienna, Austria) and GraphPad Prism (version 6.01; GraphPad Software Inc., La Jolla, Calif). In all cases, a P value < 0.05 was considered to be significant.

Results

Construction of a recurrence risk model based on differentially coexpressed genes

We applied a Cox proportional hazard regression model to analyse the 182 differentially coexpressed genes(9) and TTR data in the training cohorts in TCGA, and then we submitted these genes that were significantly associated with the TTR of PTC (log rank $P < 0.05$) to Lasso Cox regression analysis and 10-fold cross-validation for model construction and identified 5 optimal genes when lambda = 0.020 (**Supplementary Figure 1A-B**). By applying Kaplan-Meier analysis to these five genes, we confirmed that LRP2, KCNJ13 and TMC3 were significantly associated with TTR ($P < 0.05$), PDZK1IP1 had a marginally significant positive correlation with TTR ($P = 0.062$), and PDE10A could not predict TTR in the training cohort ($P = 0.108$, **Supplementary Figure 1C-G, Supplementary Table S2**). Thus, we developed a new 4-gene signature calculated using multivariate Cox survival analysis, with the formula being as follows: $RiskScore = 0.07 \times expression^{PDZK1IP1} - 0.129 \times expression^{LRP2} - 0.468 \times expression^{KCNJ13} - 0.595 \times expression^{TMC3}$. The risk score distribution of the sample showed that patients in the high-risk score

group had worse recurrence than those in the low-risk score group; PDZK1IP1 expression was positively associated with recurrence risk, whereas the expression levels of LRP2, KCNJ13, and TMC3 were negatively associated with recurrence risk (**Figure 1A**). These data indicated that PDZK1IP1 was a risk factor and that LRP2, KCNJ13, and TMC3 were protective factors. Furthermore, we analysed the prediction efficiency of the 4-gene signature on one-, two-, and three-year TTR of PTC patients and found that the area under the ROC curve (AUROC) of the 4-gene signature was greater than 0.70 (**Figure 1B**). With this risk-score formula, patients in the training set were divided into high-risk or low-risk score groups, with the z-score (0) serving as the cut-off. Kaplan-Meier curves showed that patients in the high-risk group (n = 130) had a shorter TTR than those in the low-risk group (n = 118, P < 0.0001; **Figure 1C**).

TCGA and GEO datasets test and validation

We calculated the risk score of each sample in the other 50% of *TCGA* PTC samples (*TCGA* validation dataset) according to the expression level of the sample, which also showed that patients in the high-risk-score group had worse recurrence than those in the low-risk-score group (**Figure 2A**). Furthermore, we analysed the prediction efficiency of the 4-gene signature on one-, two-, and three-year TTR in the *TCGA* validation dataset and found that the AUROC of the 4-gene signature on one-, two-, three-year TTR was 0.84, 0.64 and 0.51, respectively (**Figure 2B**). Patients in the *TCGA* validation dataset were classified into high-risk and low-risk groups with the same cut-off as that used in the training set. As expected, patients in this cohort with high-risk scores (n = 143) had a shorter TTR than those with low-risk scores (n = 106, P = 0.044; **Figure 2C**).

We further assessed the prognostic value of these 4-gene signatures in the whole *TCGA* PTC dataset with the same algorithm as that used in the training set. The spectrum clustering analysis also successfully divided all patients into two groups (**Figure 3A**). In the ROC analysis, the AUROC for the one-, two-, three-year TTR was 0.73, 0.71 and 0.60, respectively (**Figure 3B**). With the same risk-score formula, patients in the whole *TCGA* PTC cohort with high-risk scores (n = 273) had a shorter TTR than those with low-risk scores (n = 224, P < 0.001, **Figure 3C**). We further investigated the prognostic significance of the 4-gene signature in patients stratified by clinicopathological parameters. Kaplan-Meier analysis showed that patients with high-risk scores had shorter overall survival than those with low-risk scores in all age groups, among females, in stages I-II, and in all T stages (**Supplementary Figure 2**). These data demonstrated that the 4-gene signature could accurately predict the recurrence risk in most patients.

The robustness of the 4-gene signature was also assessed in the GSE33630 dataset. However, TMC3 was absent from this database, and we could not further analyse the predictive capacity of the 4-gene signature. Nevertheless, we also observed high expression levels of PDZK1IP1 and low expression levels of LRP2 and KCNJ13 in PTC samples compared to normal tissue samples ($P < 0.001$, **Figure 3D**). These data indicate that high expression of PDZK1IP1 and low expression of LRP2 and KCNJ13 are associated with the development of PTC.

linicopathological associations of the risk score in the TCGA cohort (**Supplementary Table S3,Figure 4**). There was no significant difference in age or sex between the low-risk subgroup and the high-risk

subgroup. The high-risk-score subgroup was positively correlated with progressive linicopathological features, including extrathyroidal extension (ETE), T3/T4 stage and lymphatic metastasis (LNM). The low-risk-score subgroup was negatively correlated with these linicopathological features. Next, we analysed the associations between the four genes and clinicopathological variables (**Supplementary Table S4)**. PDZK1IP1 was determined to be positively correlated with advanced T stage, lymphatic metastasis, advanced TNM stage and ETE. LRP2 was observed to be negatively correlated with advanced T stage, lymphatic metastasis, advanced TNM stage and ETE. KCNJ13 was observed to be negatively correlated with advanced T stage, lymphatic metastasis, advanced TNM stage and ETE. TMC3 was determined to be negatively correlated with advanced T stage, advanced TNM stage and ETE.

IHC analysis of protein expression of 4-gene signature

Four-gene signature expression at the protein level is presented in **Figure 5**. PDZK1IP1 was significantly higher in PTC than in normal thyroid tissue and was higher in locally advanced PTCs than in T1/T2 PTCs. For LRP2, although there was no significant difference between PTC and NT, T1/T2 PTC and LA-PTC, we still observed that the expression in PTC was slightly lower than that in NT, and that in LA-PTC was slightly lower than that in T1/T21 PTC. The protein expression of KCNJ13 was notably downregulated in LA-PTC in comparison with T1/T2 PTC and NT, while there was no significant distinction between T1/T2 PTC and NT. The IHC staining of KCNJ13 and TMC3 was almost completely negative in NT. In PTC, the expression of TMC3 in T1/T2 PTC notably exceeds that in LA-PTC. The MODs of the 4 genes in the recurrent group and nonrecurrent group are compared in **SupplementaryTable S5**. Recurrent patients were compared with an equal number of patients without recurrence, who were randomly selected as the control group. In both T1/T2 PTC and LA-PTC, PDZK1IP1 was significantly upregulated in the recurrent group, and LRP2 and KCNJ13 were significantly downregulated in the recurrent group.

Univariate and multivariate Cox regression analyses of the 4-gene signature

To identify the independence of the 4-gene signature in clinical application, we analysed its prognostic value in the whole *TCGA* PTC cohort by using univariate and multivariate Cox regression analysis. Clinicopathological parameters, including age, sex, pathological T stage, N stage, M stage and TNM stage, were included in the analysis. In the univariate analysis, age, T stage (T3/T4 vs. T1/T2), tumour stage (stage III/IV vs. stage I/II), and the 4-gene signature were significantly related to the TTR in PTC (**Figure 6A**). However, the multivariate analysis indicated that only the 4-gene signature was significantly related to the recurrence risk in PTC (HR = 3.606, 95% CI = 1.659-7.839, P = 0.001, **Figure 6B**). To demonstrate the influence degree of different variables and different values of each variable on the outcome by the length of the line, a nomogram was constructed based on the clinicopathological parameters and 4-gene signature, and the 4-gene signature showed the greatest prediction value on the TTR (**Figure 6C**). The calibration plots for predicting 2-year, 3-year and 5-year RFS were modified to visualize the nomogram (**Figure 6D**).

Potential relative pathways

Single-sample GSEA (ssGSEA) was applied to identify the relationship between the risk scores and biological functions in PTC. Seventy pathways with a coefficient greater than 0.5 are presented in **Figure 7A**, and most of them were positively correlated with the risk score. Among these pathways, the ssGSEA scores of 13 metabolic pathways, including the intestinal immune network and p53 signalling pathway, increased with increasing risk score, while the ssGSEA scores of the other four pathways, such as the Hedgehog signalling pathway, decreased with increasing risk score (**Figure 7B**). LRP2, KCNJ13, TMC3 and PDZK1IP1 may be involved in these signalling pathways, and the dysfunction of these pathways is closely related to the progression and recurrence of PTC.

Discussion

Although most PTCs have a good prognosis, and satisfying outcomes are obtained after surgery, some well-differentiated thyroid cancer cases have significantly more aggressive behaviours. Given the diversity of PTC patients with different risks, it is necessary to identify factors for risk stratifications, which is very important for formulating treatment and follow-up plans. Lipid metabolism is an important and complex biochemical reaction in the human body. There are two ways to take up fatty acids: from extracellular sources through membrane-associated transport proteins and from production through *de novo* fatty acid synthesis. Normal cells preferentially use circulating exogenous lipids, whereas tumour cells exhibit a high rate of *de novo* lipid synthesis(11). The importance of lipid metabolic reprogramming in carcinogenesis has been demonstrated in recent studies(12). It was reported that lipid metabolism is closely related to the malignancy behaviours of several cancers, including lung, breast and prostate cancer(13–15). However, the association between lipid metabolism and tumour progression has been less heavily researched in thyroid cancer. Previously, by investigated the expression profile of LMRGs in PTC and constructed two subtypes with distinct different characteristics(9). In the present study, we investigated the gene expression of our previously identified 182 key genes involving in PTC progression(9), and established a risk model.

In our results, we found that PDZK1IP1, TMC3, LRP2 and KCNJ13 were significantly associated with TTR in PTC. PDZK1-interacting protein 1 (PDZK1IP1) was first identified as an epithelium-specific molecule (16, 17) and is frequently overexpressed in a number of human carcinomas (18, 19). It was reported that the overexpression of LRP2 partially abolished the inhibitory effect of metformin on the viability of TPC-1 cells(20). The potassium channel Kir7.1 (KCNJ13) gene encodes a protein low-conductance potassium channel related to photoreceptor sensitivity and tracheal tubulogenesis (21, 22). TMC3 belongs to the transmembrane channel-like gene family. However, little known regarding the role of LRP2, TMC3 and KCNJ13 in human malignancy, and these gene requires further exploration. A single nucleotide polymorphism in TMC3 is associated with LDL cholesterol/HDL cholesterol, which influences blood lipid levels (23). This property provides a novel perspective to understand how TMC3 may affect fat metabolism in tumour cells.

Based on these four genes, a 4-gene signature recurrence prediction model was established. Using the TCGA and GEO cohort as validation datasets, a 4-gene signature model was proven to have strong

robustness. Moreover, by analysing the influence of clinical factors on the prediction efficiency of the 4-gene signature, we confirmed that this signature was a significantly independent prognostic factor in PTC. Further validation with the expression profile of all four genes in the TCGA cohort and the PTC/normal thyroid tissue microarray showed a significant association between these genes and clinicopathological parameters at the mRNA and protein levels. In addition, a comparison of the MOD of the four genes in patients with recurrence and those without recurrence demonstrated that PDZK1IP1, LRP2, and KCNJ13 are notably associated with PTC recurrence at the protein level. A limitation of the current study was that we did not consider the differences in population characteristics, such as race and lifestyle, and their effects on patient survival in this study. Further research is warranted to further confirm the validity of the 4-gene signature.

In conclusion, to the best of our knowledge, our report describes the first use of lipid metabolism-related genes to assess their ability to predict the risk of recurrences PTC. We profiled lipid metabolism-related genes in PTC and identified four genes associated with the recurrence risk and clinicopathological features of PTC, which was further verified by IHC of 127 samples containing normal thyroid tissue, T1/T2 PTC and locally advanced PTC. Moreover, we employed a lipid metabolic 4-gene signature TTR prediction model that showed satisfactory prediction ability for PTC.

Abbreviations

Papillary thyroid cancer (PTC), Lipid-metabolism-related genes (LMRG), The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), Non-negative matrix factorization (NMF), Differentially expressed genes (DEGs), Thyroid carcinoma (THCA), Weighted gene correlation network analysis (WGCNA), Gene set enrichment analysis (GSEA), Relapse free survival (RFS)

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

The datasets generated and analyzed during the current study are available in the TCGA repository (https://portal.gdc.cancer.gov/) with the accession number TCGA- THCA, and the GEO repository (https://www.ncbi.nlm.nih.gov/geo/) the accession number was GSE33630. Coherently expressed signatures of human lipid metabolism-related pathways were all download from the Reactome pathway database (https://reactome.org/) and derived by aggregating MSigDB version 7.0 gene sets.

Competing interests

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

Funding

This study was supported by grants from the National Natural Science Foundation of China (no. 81702649 to NQ, nos. 81472498 and 81772851 to YLW) and Zhejiang Basic Public Interest Research Project (LGF20H160012).

Authors' contributions

Ning Qu, Mi-die Xu and Rong-liang Shi put forward concept and conducted the work. Shi-shuai Wen and Ben Ma drafted and revised the review. Qing-hai Ji and Yi-jun Wu offered technical and administrative support. Ting-ting Zhang helped data analysis and interpretation. Shi-shuai Wen, Yi Luo and Wei-Li Wu did basic experiments. All authors gave final approval of the version to be published.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Acknowledgements

The authors would like to thank all researchers contributed to the TCGA and GEO data sets included.

References

1. N. Hay, Reprogramming glucose metabolism in cancer: can it be exploited for cancer therapy? Nature reviews Cancer **16**(10), 635–649 (2016)
2. Q. Liu, Q. Luo, A. Halim, G. Song, Targeting lipid metabolism of cancer cells: A promising therapeutic strategy for cancer. Cancer Lett. **401**, 39–45 (2017)
3. J.J. Wiltshire, T.M. Drake, L. Uttley, S.P. Balasubramanian, Systematic Review of Trends in the Incidence Rates of Thyroid Cancer. Thyroid. **26**(11), 1541–1552 (2016)
4. M.N. Nikiforova, Y.E. Nikiforov, Molecular genetics of thyroid cancer: implications for diagnosis, treatment and prognosis. Expert Rev Mol Diagn **8**(1), 83–95 (2008)
5. F. Pacini, F. Cetani, P. Miccoli, F. Mancusi, C. Ceccarelli, F. Lippi, E. Martino, A. Pinchera, Outcome of 309 patients with metastatic differentiated thyroid carcinoma treated with radioiodine. World J Surg **18**(4), 600–604 (1994)

6. T. Liao, Y.J. Wang, J.Q. Hu, Y. Wang, L.T. Han, B. Ma, R.L. Shi, N. Qu, W.J. Wei, Q. Guan, J. Xiang, J.Y. Chen, G.H. Sun, D.S. Li, X.M. Mu, Q.H. Ji, Histone methyltransferase KMT5A gene modulates oncogenesis and lipid metabolism of papillary thyroid cancer in vitro. Oncol. Rep. **39**(5), 2185–2192 (2018)
7. F. Baenke, B. Peck, H. Miess, A. Schulze, Hooked on fat: the role of lipid synthesis in cancer metabolism and tumour development. Dis Model Mech **6**(6), 1353–1363 (2013)
8. J.P. Brunet, P. Tamayo, T.R. Golub, J.P. Mesirov, Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci U S A **101**(12), 4164–4169 (2004)
9. M. Xu, T. Sun, S. Wen, T. Zhang, X. Wang, Y. Cao, Y. Wang, X. Sun, Q. Ji, R. Shi, N. Qu, Characteristics of lipid metabolism-related gene expression-based molecular subtype in papillary thyroid cancer. Acta Biochim. Biophys. Sin. (Shanghai) **52**(10), 1166–1170 (2020)
10. J. Lubsen, J. Pool, E. van der Does, A practical device for the application of a diagnostic or prognostic function. Methods Inf. Med. **17**(2), 127–129 (1978)
11. E. Werbrouck, D. Van Gansbeke, A. Vanreusel, M. De Troch, Temperature Affects the Use of Storage Fatty Acids as Energy Source in a Benthic Copepod (Platychelipus littoralis, Harpacticoida). PLoS One **11**(3), e0151779 (2016)
12. E. Currie, A. Schulze, R. Zechner, T.C. Walther, R.V. Farese, Jr. Cellular fatty acid metabolism and cancer. Cell Metabol. **18**(2), 153–161 (2013)
13. G. Deep, I.R. Schlaepfer, Aberrant Lipid Metabolism Promotes Prostate Cancer: Role in Cell Survival under Hypoxia and Extracellular Vesicles Biogenesis. *Int J Mol Sci.* 2016;17(7)
14. Y.J. Cha, H.M. Kim, J.S. Koo, Expression of Lipid Metabolism-Related Proteins Differs between Invasive Lobular Carcinoma and Invasive Ductal Carcinoma. *Int J Mol Sci*. 2017;18(1)
15. S. Kim, Y. Lee, J.S. Koo, Differential expression of lipid metabolism-related proteins in different breast cancer subtypes. PLoS One **10**(3), e0119473 (2015)
16. C. Jaeger, B.M. Schaefer, R. Wallich, M.D. Kramer, The membrane-associated protein pKe#192/MAP17 in human keratinocytes. J Invest Dermatol **115**(3), 375–380 (2000)
17. O. Kocher, P. Cheresh, S.W. Lee, Identification and partial characterization of a novel membrane-associated protein (MAP17) up-regulated in human carcinomas and modulating cell replication and tumor growth. Am. J. Pathol. **149**(2), 493–500 (1996)
18. M.V. Guijarro, J.F. Leal, J. Fominaya, C. Blanco-Aparicio, S. Alonso, M. Lleonart, J. Castellvi, L. Ruiz, Y.C.S. Ramon, A. Carnero, MAP17 overexpression is a common characteristic of carcinomas. Carcinogenesis. **28**(8), 1646–1652 (2007)
19. M.V. Guijarro, M.E. Castro, L. Romero, V. Moneo, A. Carnero, Large scale genetic screen identifies MAP17 as protein bypassing TNF-induced growth arrest. Journal of cellular biochemistry **101**(1), 112–121 (2007)
20. Y. He, L. Cao, L. Wang, L. Liu, Y. Huang, X. Gong, Metformin Inhibits Proliferation of Human Thyroid Cancer TPC-1 Cells by Decreasing LRP2 to Suppress the JNK Pathway. Onco Targets Ther **13**, 45–50 (2020)

21. H. Zhong, Y. Chen, Y. Li, R. Chen, G. Mardon, CRISPR-engineered mosaicism rapidly reveals that loss of Kcnj13 function in mice mimics human disease phenotypes. Sci Rep **5**, 8366 (2015)
22. W. Yin, H.T. Kim, S. Wang, F. Gunawan, L. Wang, K. Kishimoto, H. Zhong, D. Roman, J. Preussner, S. Guenther, V. Graef, C. Buettner, B. Grohmann, M. Looso, M. Morimoto, G. Mardon, S. Offermanns, D.Y.R. Stainier, The potassium channel KCNJ13 is essential for smooth muscle cytoskeletal organization during mouse tracheal tubulogenesis. Nat Commun **9**(1), 2815 (2018)
23. C. Dong, A. Beecham, L. Wang, S. Slifer, C.B. Wright, S.H. Blanton, T. Rundek, R.L. Sacco, Genetic loci for blood lipid levels identified by linkage and association analyses in Caribbean Hispanics. J Lipid Res **52**(7), 1411–1419 (2011)

Figures

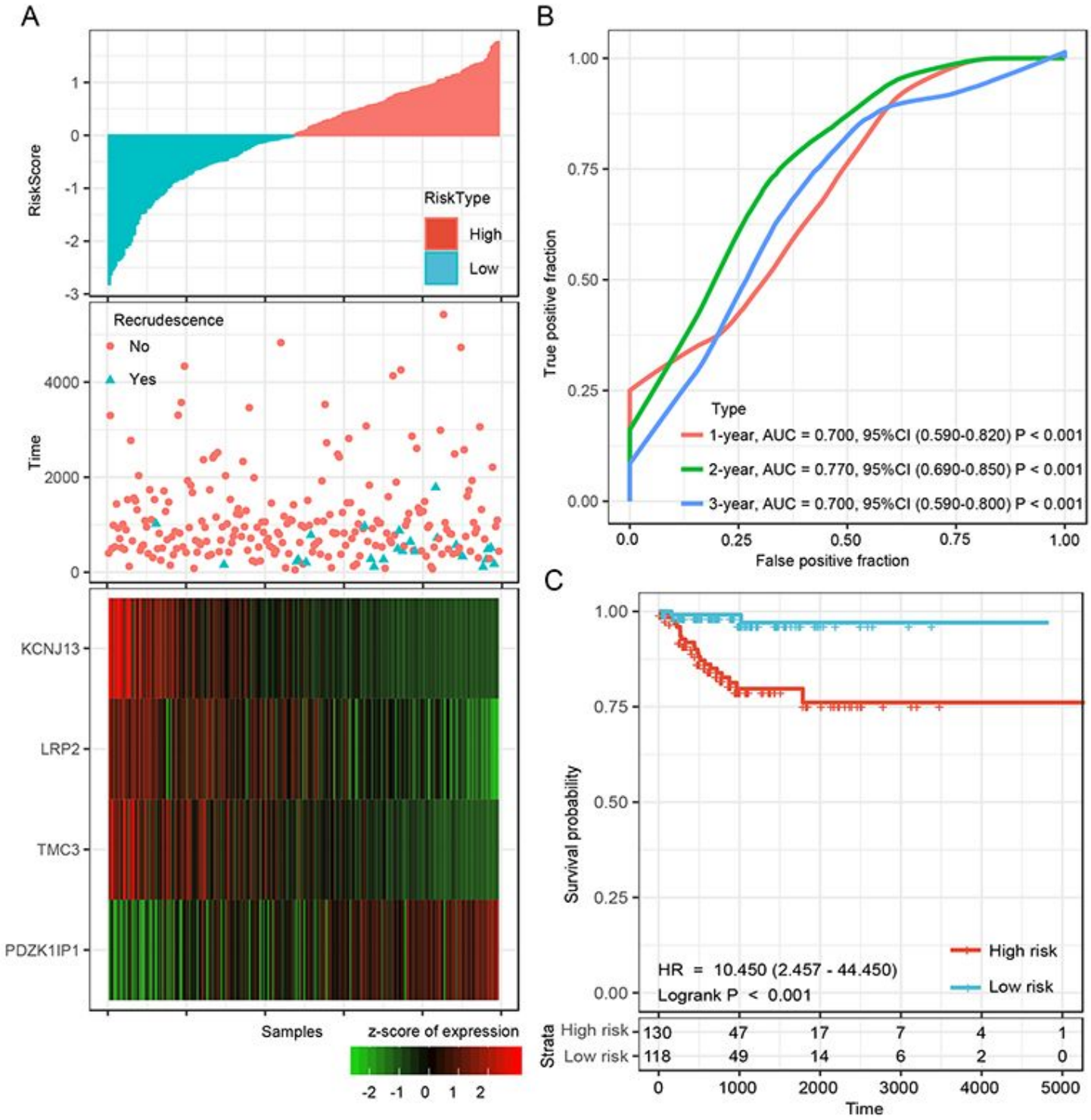

Figure 1

Performance of the 4-gene signature in the training set. A: Distribution of patients' risk score, TTR time and recrudescence state and 4-gene expression in the training set; B: Receiver operating characteristic (ROC) curves for TTR at different time points in the training set. P-values show the area under the ROC curve (AUROC) of the 4-gene signature; C: Kaplan-Meier method and log-rank test of the prognosis difference between patients in the indicated group.

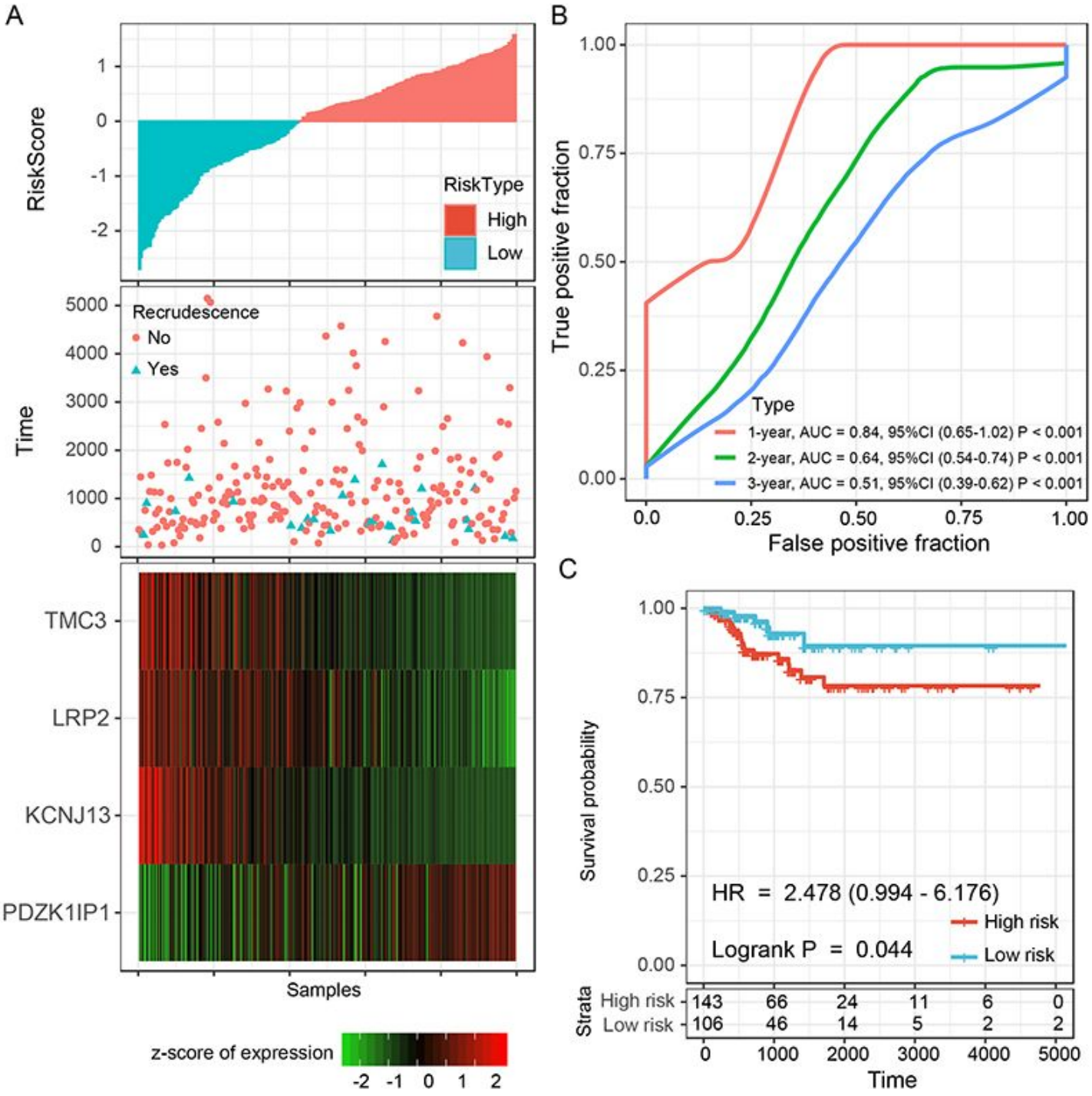

Figure 2

Performance of the 4-gene signature in the TCGA validation set A: Distribution of patients' risk score, TTR time and recrudescence state and 4-gene expression in the TCGA validation set; B: ROC curves for TTR at different time points in the TCGA validation set. P-values show AUROC of the 4-gene signature; C: Kaplan-Meier method and log-rank test of the prognosis difference between patients in the indicated group.

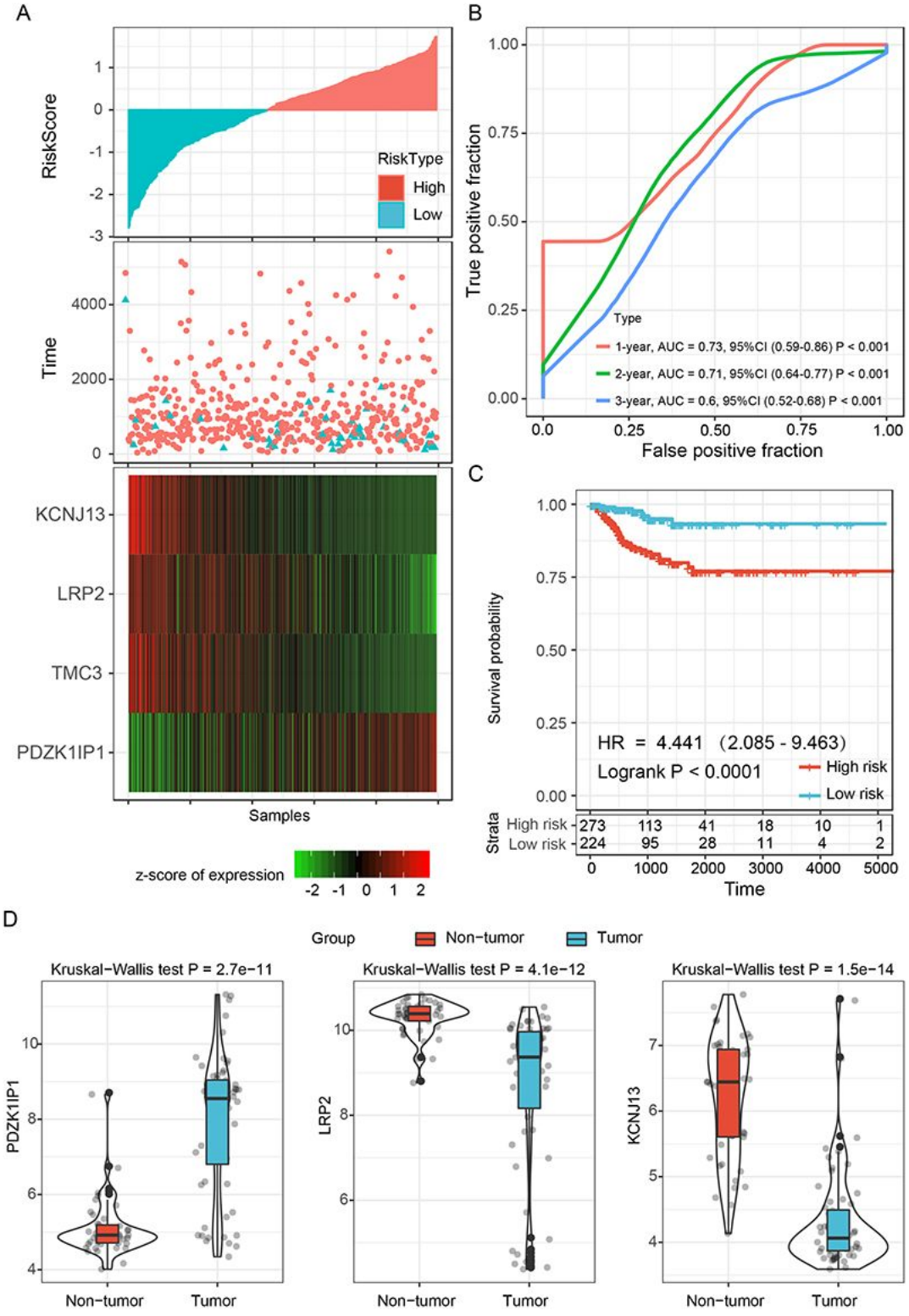

Figure 3

Performance of the 4-gene signature in the entire TCGA dataset and the GSE33630 datasets A: Distribution of patients' risk score, TTR time and recrudescence state and 4-gene expression in the entire TCGA set; B: ROC curves for TTR at different time points in the entire TCGA set. P-values show the AUROC of the 4-gene signature; C: Kaplan-Meier method and log-rank test of the prognosis difference between

patients in the indicated group; D: The expression of PDZK1IP1, LRP2 and KCNJ13 in the GSE33630 dataset.

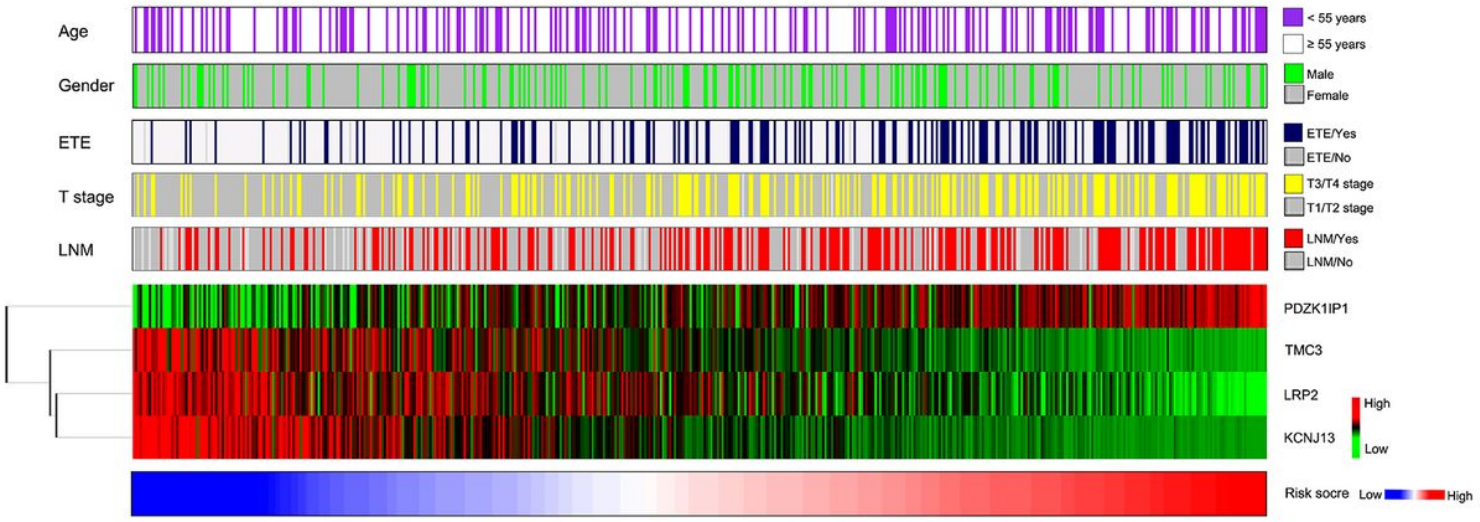

Figure 4

Clinicopathological associations of the signature risk score in the TCGA cohort Patients were ranked by age, sex, ETE, T stage, signature risk score and heat map of the expression of the four genes.

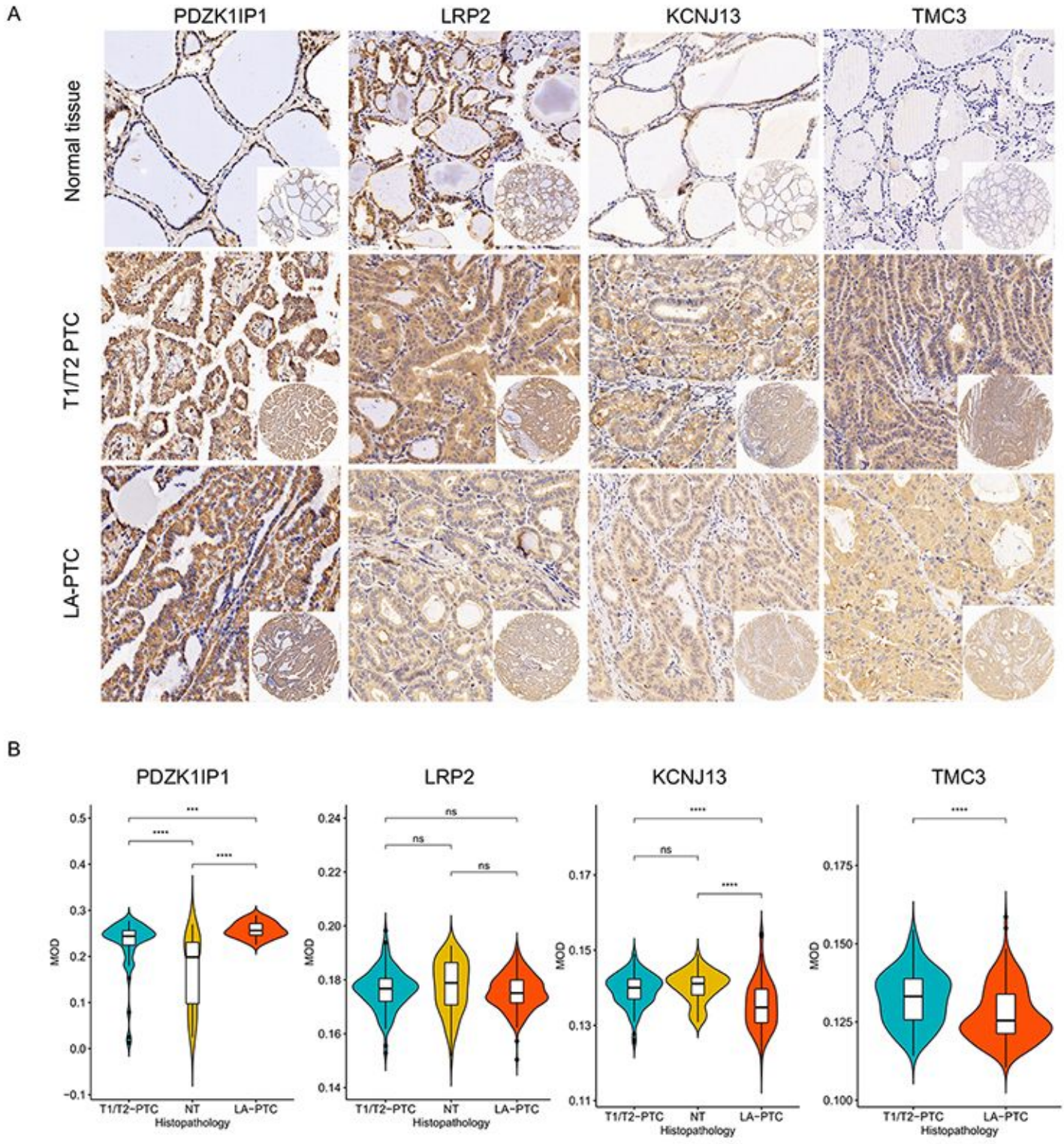

Figure 5

IHC analysis of the protein expression of the 4-gene signature in T1/T2 PTC, LA-PTC and NT A: Representative IHC staining of 4 genes in stage T1/T2 PTC, locally advanced PTC and normal tissues in the TMA cohort. B: Comparison of the protein expression of 4 genes among stage T1/T2 PTC, locally advanced PTC and normal tissues in the TMA cohort (ns: not significant, * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001).

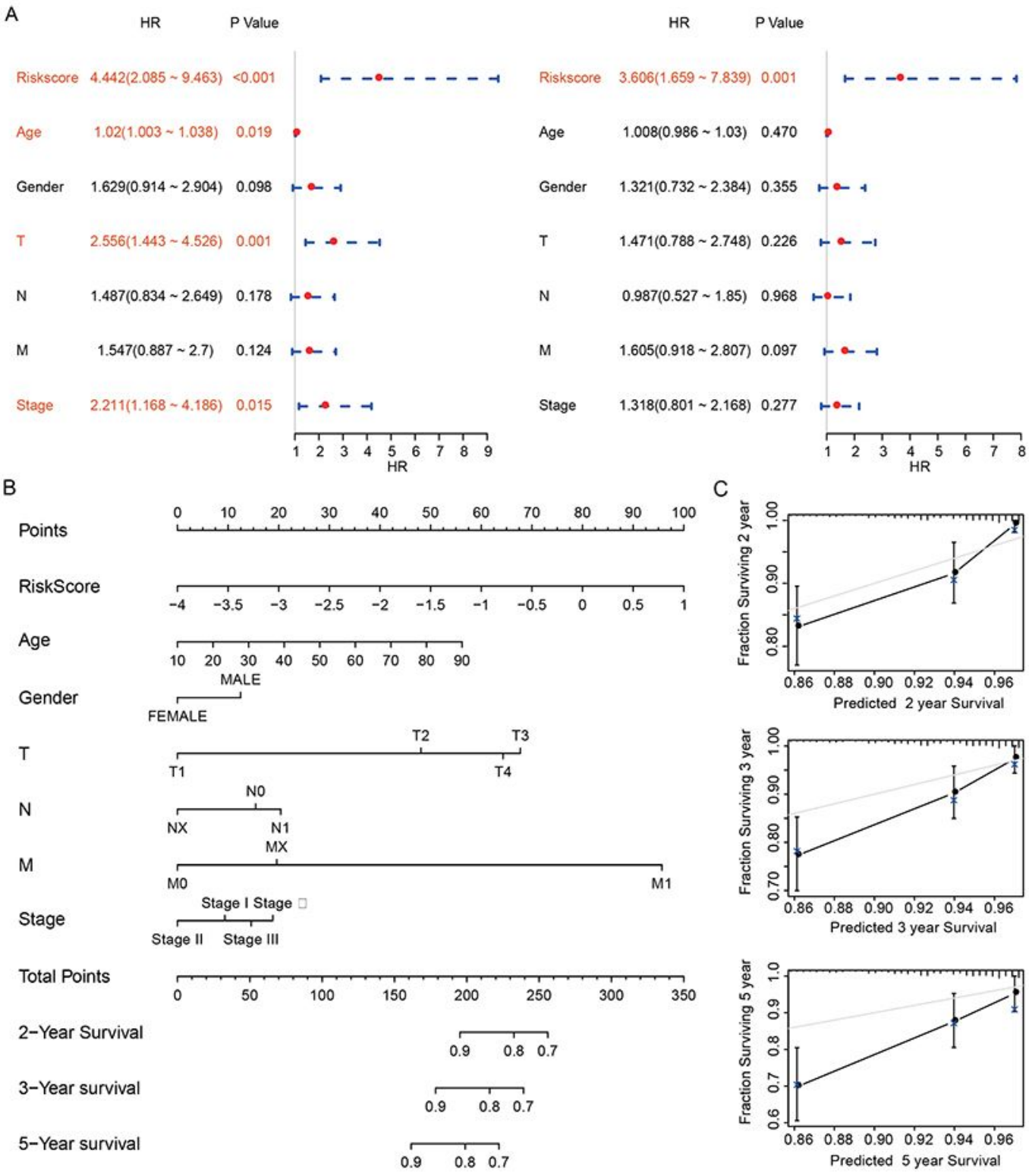

Figure 6

A. Forest plot of the univariate and multivariate Cox regression analyses. ‘Orange’ indicates a significant relationship with relapse-free survival. B: A nomogram based on the risk score and clinical factors. C: Calibration plots for predicting 2-year, 3-year and 5-year relapse-free survival.

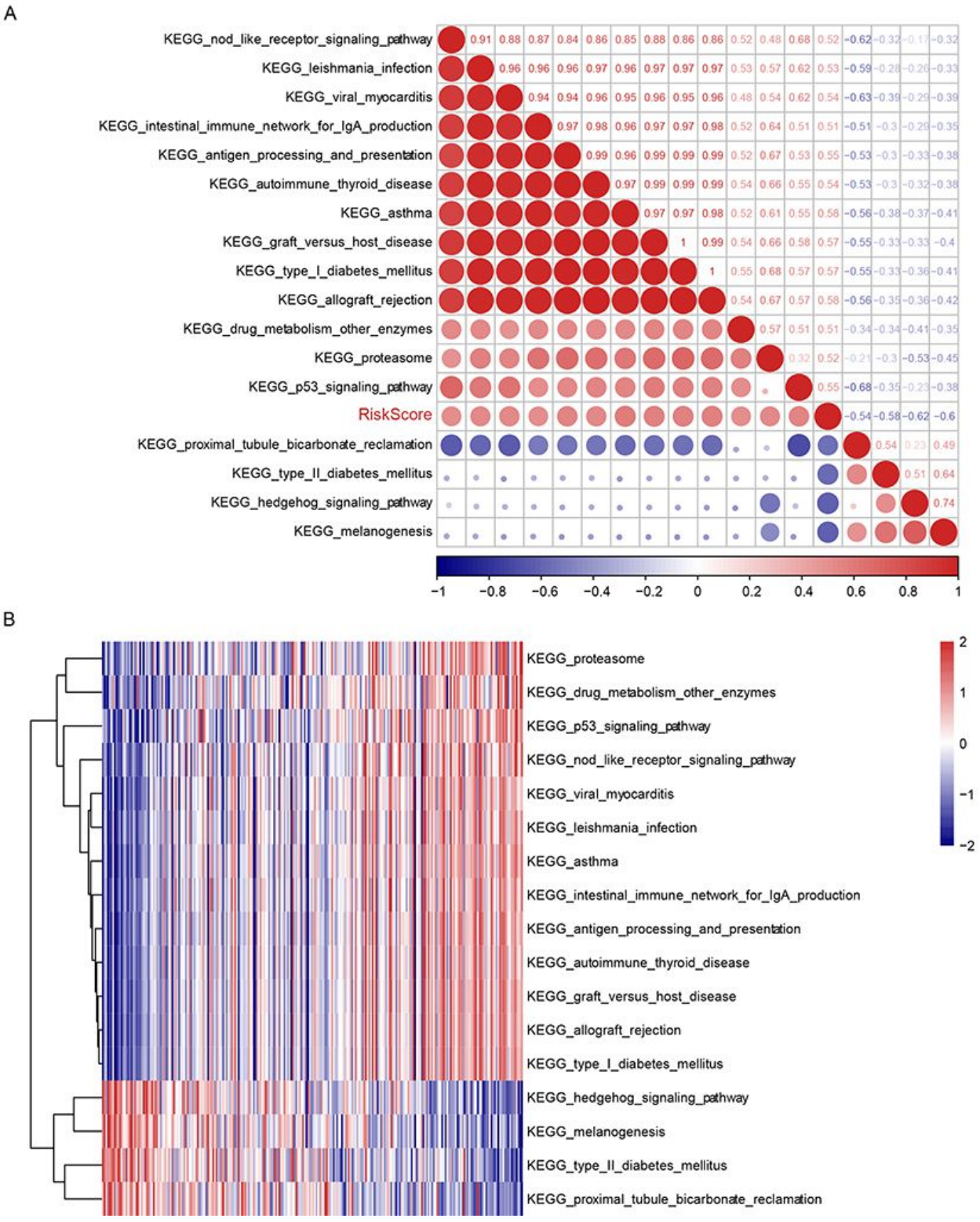

Figure 7

Potential pathways relative to the 4 genes A: Cluster analysis of the risk score and the 17 KEGG pathways with coefficient > 0.5. B: The relationship between the risk score and the ssGSEA score of the 17 KEGG pathways with a coefficient > 0.5.

Supplementary Files

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

- Supplementarymaterials.docx