

m5C-Related lncRNAs impact on Prognosis and Tumor Microenvironment in Papillary Thyroid Carcinoma

WenLong Wang

Xiangya Hospital Central South University <https://orcid.org/0000-0003-4441-6027>

Xinying Li (✉ lixinyingcn@126.com)

Xiangya Hospital Central South University

Research article

Keywords: papillary thyroid carcinoma (PTC), tumor immune microenvironment, prognosis, 5-methylcyosine (m5C), nomogram, immunotherapy

Posted Date: October 25th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-966667/v1>

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

[Read Full License](#)

Abstract

Background

Emerging evidence has indicated that 5-methylcyosine (m5C) modification plays a vital role in cancer development. However, the function of m5C RNA methylation-related long noncoding RNAs (m5C-lncRNAs) in papillary thyroid carcinoma (PTC) has never been reported. We aimed to investigate the role of m5C-lncRNAs in the prognosis and tumor immune microenvironment of PTC.

Methods

Pearson correlation analysis, univariate and multivariate Cox regression analysis to filter the prognostic significance of m5C-lncRNAs. Corresponding risk score and immune risk score models were established. ESTIMATE, CIBERSOFT, ssGSEA, and MCP counter algorithms to analysis the tumor microenvironment. Finally, the expression level of m5C-lncRNAs were validated by PTC tissues and cell lines.

Results

A total of 6 m5C-related lncRNAs and five immune cell types were selected to construct risk score and immune risk score (IRS) model, respectively. Patients with a high-risk score had a worse prognosis and the ROC indicated a reliable prediction performance (AUC=0.796). As expected, the ESTIMATE and immune scores were higher ($P < 0.001$) and the tumor purity ($P < 0.05$) was significantly lower in the low-risk subgroup. CIBERSORT analysis showed Tregs, M0 macrophages, dendritic cells resting, and eosinophils were positively correlated to the risk score. Moreover, the expression levels of PD-1, PD-L1, CTLA-4, TIM-3, LAG-3, and KLRB1 were lower in the high-risk subgroup. Importantly, patients in high-risk subgroup tended to have a better response to immunotherapy than those in low-risk subgroup ($P = 0.022$). Similar to the above risk score, IRS model also showed favorable prognosis predictive performance (AUC=0.764). The abundance levels and proportions of B cell naive, plasma cell, M1 macrophages, mast cells activated were higher in high-IRS subgroup. Similarly, MCP-counter algorithms analysis showed that the level of B cell, monocytes, myeloid dendritic cells were higher in high-IRS subgroup. Finally, an integrated nomogram combining risk score, IRS, and age exhibited good prognostic predictive performance. Additionally, 20 clinical samples and 7 PTC cell lines were validated.

Conclusions

Our research confirm that m5C-lncRNAs not only contribute to evaluate the prognosis of patients with PTC but also help predict immune cell infiltration and immunotherapy response.

Introduction

Papillary thyroid cancer (PTC) is a common malignant endocrine tumor, and its incidence have increased over the past decades (W. Wang, et al. 2020). In 2015, there were approximately 201,000 new cases in China (151,000 females and 49,000 males), accounting for 10.44/100,000 in the general

population, and become the fastest-growing malignant tumor among women under 30 years old (S. Zhang, et al. 2021, (W. Wang, et al. 2019). Most patients with PTC (PTCs) have a favorable prognosis by surgical resection, followed by either radioactive iodine or observation. However, a small number of PTCs still suffer from radioactive iodine refractory or progress to distant metastasis, resulting in high mortality rates and poor life quality (W. Wang, et al. 2021). Hence, elucidating the underlying molecular mechanisms of PTC progression, and developing novel therapeutic targets for PTCs treatment is crucial.

5-methylcytosine (m5C) modification, which was discovered in 1925, is a common RNA modification of mRNA or non-coding RNAs (ncRNAs), and plays essential roles in structural stability, tRNA recognition, ribosome assembly, RNA export, and translation (P. Nombela, et al. 2021, (D. Dominissini, et al. 2017, (L. Trixl, et al. 2019). At the post-transcriptional level, m5C RNA modifications can be reversibly and dynamically controlled by methyltransferases (m5C writers), demethylases (m5C erasers), and binding proteins (m5C readers) (J. Pan, et al. 2021, (H. Chen, et al. 2021). The methyltransferase complex includes *NOL1/NOP2*/sun family (*NSUN1-7*), DNA methyltransferase (*DNMT*) member (*DNMT1, DNMT2, DNMT3A, and DNMT3B*). *TET2* is an important “erasers” protein which acts as a demethylase. The main “readers” is ALYREF, which can read the m5C motif. Existing evidence has indicated that dysregulation of m5C modification plays an important role in the tumor initiation, progression, and drug responses (Q. Li, et al. 2017, (H. Chen, et al. 2020). For instance, aberrant *NSUN2* promotes the pathogenesis by targeting the m5C methylation site in the untranslated region of *HDGF3* in human bladder urothelial carcinoma (X. Chen, et al. 2019). Moreover, *NSUN2* was also demonstrated to promote tumor progression and increase drug resistance by methylation of *NMR* in esophageal squamous cell carcinoma (Y. Li, et al. 2018). Additionally, studies showed that downregulation of *TET3* promotes glioblastoma tumorigenesis through the genome-wide m5C levels (A. Carella, et al. 2020). These studies reveal that m5C regulators are highly involved in tumor progression and metastasis, thus serving as useful therapeutic targets with promising prognostic values. However, the role of m5C in PTC still unclear.

In recent years, increasing evidence has revealed that tumor microenvironment (TME) plays a critical role in PTC aggressiveness and progression, especially in radioiodine-refractory and advanced PTC (S. M. Ferrari, et al. 2019, (Z. Hurst, et al. 2019). Novel therapies are necessary for these PTCs. Antitumor efficacy of immune checkpoint blockade (ICB) therapy, targeting cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) and/or programmed cell death protein 1 (PD-1)/PD-ligand 1 (PD-L1), has been observed promising clinical results in non-small-cell lung carcinoma, metastatic melanoma, and (R. J. Motzer, et al. 2020, (A. Prat, et al. 2017). In the light of these successes, there has been strong interest in radioiodine-refractory and advanced PTCs. To optimize this therapeutic approach, there is a critical need to identify high-risk PTCs who likely to achieve benefit in the early stage, but it currently performs poorly in PTCs.

In this study, we performed in-depth sequence analyses to identify the prognostic significance of m5C-related lncRNAs (m5C- lncRNAs) in PTCs based on the Cancer Genome Atlas (TCGA) dataset. Corresponding risk models and immune risk score (IRS) of m5C-lncRNAs prognostic signature was established to improve prognostic risk stratification and the possible response to ICB therapy of patients.

This study sought to provide a new insight into the regulatory mechanisms associated with TME and the strategies for PTC immunotherapy.

Materials And Methods

Gene Datasets and Clinical Data Collection

We downloaded the RNA-seq dataset from the UCSC Xena browser (<http://xena.ucsc.edu/>), which containing 58 normal and 492 thyroid cancer samples with complete clinical information. The expression data were normalized by Fragment Per Kilobase Million (FPKM). The corresponding patients' clinical information including gender, age, extrathyroidal extension, bilaterality, multifocality, histological type, TNM stage, lymph node metastasis (LNM), and survival time. According to published literature, there were 14 m5C regulators, including *SUN1-7*, *DNMT1*, *DNMT1*, *DNMT2*, *DNMT3A*, *DNMT3B*, *ALYREF*, *TRDMT1*, *TET2*, and *YBX1*. Among them, *DNMT1* and *DNMT2* were not found in the TCGA THCA data. Next, differential expression analysis ($|\text{Log}_2\text{FC}| > 1$, adjusted p-value < 0.05) between PTC and normal samples by the limma software package in R version (4.0.2). Subsequently, Spearman correlation coefficient analysis were used to show the relationship of m5C regulators. The work flow was shown in Figure 1.

Protein-Protein Interaction (PPI) Network

To further analysis the potential PPI relationship, we use the STRING database (<https://string-db.org/>) to analyze the m5C regulators and construct PPI network, The threshold of genes at the center of the PPI network was the minimum gene interaction score > 0.4 .

Construction of Risk Assessment Model Construction

Pearson's correlation method was performed to filter the m6A-lncRNA based on the threshold criteria of Pearson's coefficient $R^2 > 0.5$ and $P < 0.001$. Then, univariate and multivariate Cox analysis were used to identify the prognostic significance of m5C-lncRNAs. A total of 9 prognostic m6A-lncRNAs were extracted and for further analyzed. The least absolute shrinkage and selection operator (LASSO) Cox regression was employed to select candidate prognostic m5A-lncRNAs by the "glmnet" package. The optimal penalty parameter λ was set as the minimum 10-fold cross-validation. The risk score for each patient was calculated according to the following algorithm: Risk score = $\sum_{i=1}^n \text{coef}(i) * a(i)$. The a_i represented the enrichment scores m5C-lncRNAs and the enrichment scores of immune cells in the risk score model and immune risk score model, respectively. The patients were divided into high - and low -risk groups based on the median risk score. Kaplan-Meier curves with the log-rank test were used to draw the survival curves. The area under the ROC curve (AUC) was used to assess the overall survival predictive accuracy of the risk score model.

Evaluation of Immune Infiltration and Enrichment Functional Analysis

single-sample Gene Set Enrichment Analysis (ssGSEA) was used to calculate the enrichment scores of types of infiltrating immune cells with the "GSVA" and "GSEABase" software package. ESTIMATE

algorithm (<https://bioinformatics.mdanderson.org/estimate/>) was employed to calculate the immune and stroma scores for each patient, and tumor purity was calculated by genomic methods. Furthermore, CIBERSOFT and MCP counter algorithms to assess the immune response of immune cells and components. The response of PTCs to ICB therapy was predicted by ImmuCellAI (<http://bioinfo.life.hust.edu.cn/ImmuCellAI/>) (Y. R. Miao, et al. 2020). Differentially expressed genes (DEGs) between high- and low- immune risk score (IRS) subgroups were screened by the “limma” package based on P values <0.05 and $|\log_2FC| \geq 1$. GO enrichment and KEGG pathway analysis were realized through Metascape (<http://metascape.org>).

Quantitative real-time polymerase chain reaction validation

In this study, eight human thyroid cancer cell lines (B-CPAP, K1, TPC-1, IHH4, KTC-1, BHT101, WRO, CAL-62) and human normal thyroid epithelial cell line (nthy-ori3-1) were used and cultured in a complete medium composed of 10% fetal bovine serum (Gibco) and 90% RPMI1640 (Gibco, Carlsbad, USA) or DMEM (Gibco, Carlsbad, USA), supplemented with 100 U/mL penicillin (HyClone) and 100 mg/mL streptomycin. These cells were cultured in 37 °C with 95% O₂ and 5% CO₂. Besides, we collected 20 PTC tissue samples and paired adjacent tissues from patients who underwent thyroidectomy in the Thyroid Surgery Department of Xiangya hospital from March 2020 to June 2020. The fresh tissues were stored at -80°C. Informed consent was obtained from all the participants and approved by the Ethics Committee of Xiangya Hospital of Central South University.

Briefly, total RNA was extracted using the AG RNAex Pro RNA extraction kit (AG21101). cDNA was synthesized using a reverse transcription kit (Takara, Japan). Subsequently, quantitative real-time polymerase chain reaction (qRT-PCR) assay was carried out by the TBGreen Premix Ex Taq™II (Cat # RR047A-5, TaKaRa, Japan). Relative expression levels were calculated using the $2^{-\Delta\Delta C_q}$ method. Primer sequences for m5C-lncRNAs are listed in Supplementary Table 1.

Statistical Analysis

Statistical tests were performed using SPSS 22.0 (IBM, NY, USA) and R software (version 4.0.2). Chi-square and Student's t -tests were performed to compare the differences between the two subgroups. The Kruskal–Wallis test was used to compare immune scores, stroma scores, tumor purity, and ESTIMATE scores among different risk score subgroups. Survival curves were depicted by using the Kaplan–Meier method. Principal component analysis (PCA) was conducted to reveal the distribution of samples. In addition, we performed univariate and multivariate analyses to identify independent prognostic factors and establish an integrated nomogram combining predictable clinicopathological factors and risk scores. The predictive performance of the nomogram was validated by calibrated plots and the concordance index (C index). The receiver operating characteristic (ROC) curve was used to verify the prognostic ability of the nomogram for 3-/5-year OS. The calibration plot was used to comparison between actual and nomogram-predicted outcome and decision curve analysis (DCA) was employed to assess the clinical values. The statistical significance was shown as followed: $P < 0.05$ (*), $P < 0.01$ (**), $P < 0.001$ (***) .

Table 1 Univariate and multivariate Cox regression analysis the prognosis factors of PTC

Variable	Univariate analysis		Multivariate analysis	
	HR (95%CI)	<i>P</i>	OR (95%CI)	<i>P</i>
Age	1.16(1.10-1.22)	<0.001	1.15 (1.08-1.23)	<0.001
Gender				
Female	ref		-	
Male	1.92(0.69-5.29)	0.21	-	-
TNM stage				
I	ref		ref	
II	5.67(0.80-40.27)	0.083	-	-
III	10.25(2.13-49.45)	0.004	0.29 (0-51.65)	0.638
IV	16.47(3.18-85.33)	<0.001	0.08 (0-29.85)	0.408
T Stage				
T1	ref		ref	
T2	1.10(0.18-6.59)	0.920	-	-
T3	1.69(0.33-8.73)	0.531	-	-
T4	11.52(2.31-57.57)	0.003	66.34(0.21-298.54)	0.154
N Stage				
N0	ref		-	-
N1	1.14(0.43-3.05)	0.078	-	-
M Stage				
M0	ref		ref	
M1	5.39(1.22-23.83)	0.026	1.44(0.10-21.10)	0.789
Multifocality				
Multifocal	ref		-	-
Unifocal	3.91(0.88-17.34)	0.073	-	-
Bilaterality				
Bilateral	ref		-	-
Unilateral	0.95(0.21-4.26)	0.942	-	-
Isthmus	1.05(0.09-11.79)	0.967	-	-
Risk score	2.22(1.57-3.14)	<0.001	2.72(1.54-4.79)	<0.001

Immune risk score	1.78(1.34-2.38)	<0.001	1.69(1.13-2.54)	0.011
-------------------	-----------------	--------	-----------------	-------

Abbreviations: CI: confidence intervals,

Results

The Profile of m5C-lncRNA Regulators

To clarify the interaction and related functions of m5C regulators in PTC, we analyzed 492 PTC tissues and 58 normal tissues from TCGA. The results showed that the expression of *NSUN1*, *NSUN3*, *NSUN4*, *NSUN6*, *NSUN7*, *DNMT3B*, *ALYREF*, *TRDMT1*, *TET2*, and *YBX1* were significantly down-regulated in PTC samples compared with normal tissues (all $p < 0.05$). While *NSUN5* and *DNMT1* were significantly upregulated in PTCs (both $p < 0.001$). However, the expression of *NSUN2* and *DNMT3A* showed no significant differences (Fig 2A). Next, to identify the relationship between 14 m5C regulators, we used the STRING database to build a PPI network (Fig 2B), and revealed the *TRDMT1* as a hub gene among m5C regulators that interacted with the other 12 genes (Fig 2C). However, *TRDMT1* did not show a strong correlation with other m5C regulators in the subsequent correlation analysis. Interestingly, *NSUN3* and *TET2* had the strongest correlation, and its expression was significantly associated with other eight genes (Fig 2D). These results indicated *NSUN3* may be a key gene that involved in the developed of PTC.

Then, Pearson correlation analysis was employed to search m5C-lncRNAs based on the criterion of Pearson $R > 0.5$ and $p < 0.001$, and 556 m5C-lncRNAs were uncovered. Integrated with the prognostic information, univariate and multivariate Cox regression were to screen m5C-lncRNAs prognostic signatures. Finally, 9 m5C-lncRNAs were significantly related with the overall survival. Fig 2E showed that most of m5C regulators positively correlated with the expression level of lncRNAs. Among 9 m5C-lncRNAs prognostic signatures, the expression of *ZNF451-AS1*, *AC018647.2*, *POLR2J4*, *PVT1*, *AL359532.1*, *A2M.AS1*, *PPP1R12A-AS1*, and *AC090515.3* showed significantly lower in PTC samples than normal tissues ($p < 0.05$). Whereas only *AC024257.3* was upregulated in PTCs (Fig 2F).

Risk Signature Was Associated with PTC Prognosis and Tumor Immune Microenvironment

To evaluate the prognostic value of m5C-lncRNAs in PTC, we performed LASSO Cox analysis that 6 m5C-lncRNAs exhibited strong prognostic value (Fig S1). The risk score for each patient was calculated based on the expression profile of 6 m5C-lncRNAs weighted by their regression coefficients. Risk score = $(4.0 \times ZNF451-AS1) + (0.68 \times AC018647.2) + (2.02 \times POLR2J4) + (0.53 \times PVT1) - (4.64 \times AC024257.3) + (0.58 \times PPP1R12A-AS1)$. The patients were divided into low- and high-risk subgroups based on the median risk scores, and further survival analysis showed that patients in the low-risk subgroups had better prognosis than high-risk subgroups ($P=0.016$) (Fig 3B), indicating that the risk model had a prognostic value. Also, the number of death cases in the low-risk subgroup was less than that in the high-risk subgroup (Fig 3A). To assess the prognostic accuracy of the risk score, the AUC of ROC analysis was 0.796 (Fig 3C). At the same time, principal component analysis could distinct the distribution of the two

risk groups (Fig S2). As for the tumor microenvironment of PTC, compared with high-risk subgroup, the ESTIMATE and immune scores were higher ($P < 0.001$) and the tumor purity ($P < 0.05$) was significantly lower in the low-risk subgroup (Fig 3D). CIBERSORT analysis showed Tregs, M0 macrophages, dendritic cells resting, and eosinophils were positively correlated to the risk score (Fig 3E). Further, elevated levels of T cell, cytotoxicity score, myeloid dendritic cells and decreased levels of neutrophil and endothelial cell were found in the low-risk subgroup when compared with the high-risk subgroup (Fig 3F). These results indicated that risk score model has a promising prognostic predictive value and closely associated with tumor microenvironment in PTCs.

In addition, to investigate the relationship between immune cell scores and prognosis. We performed univariate Cox regression analyses for 28 immune cell types, and 26 cell types showed significance differences ($P < 0.05$) (Fig 4A). With the aid of the LASSO Cox regression analysis, we generated five cell types to calculate the risk score for each patient. Immune risk score (IRS) = $(44.37 \times \text{central memory CD4 T cell}) + (19.83 \times \text{eosinophil}) + (60.06 \times \text{monocyte}) - (49.01 \times \text{T follicular helper cell}) - (29.86 \times \text{type17 T helper cell})$. Similar to the above risk score, patients were also classified into low- and high-IRS subgroup based on the median IRS. The heatmap results indicated that central memory CD4 T cell, eosinophil and monocyte were highly enrichment in the high-IRS group, whereas the low-IRS group had higher infiltration levels of T follicular helper cell and type17 T helper cell (Fig 4B). Kaplan–Meier analysis showed that the overall survival of high-IRS was poorer than low-IRS subgroup ($p = 0.004$) (Fig 4C). The AUC value for the IRS model was 0.764 (Fig 4D), which showed favorable prognostic predictive performance. Combining clinical features, we found that age, TNM stage, T stage, N stage, M stage, bilaterality, and multifocality were not significance difference between two IRS groups, only gender was associated with IRS ($P < 0.05$) (Fig S3). Besides, the ESTIMATE and stroma scores were notably higher ($P < 0.001$), whereas the tumor purity was significantly lower in the high-IRS than those in the low-IRS subgroup ($P < 0.05$) (Fig 4E). Utilizing the CIBERSORT algorithm, we found the level of B cell naive, plasma cell, M1 macrophages, mast cells activated were positively correlated to the IRS while Tregs, T cells gamma delta, M1 macrophages, mast cells resting, monocytes, and neutrophils exhibited negative correlations (Fig 4F). Further, MCP-counter analysis showed B cell, monocytes, myeloid dendritic cells, and cancer associated fibroblast were enrichment in high-IRS subgroup (Fig 4G).

Molecular basis of the IRS Signature

To better comprehend the potential biological mechanisms between the high- and low-IRS subgroups, we performed gene enrichment analysis by Metascape to explore their mechanism pathways. The top 6 GO terms used were TNF α signaling via NF- κ B, epithelial mesenchymal transition, response to growth factor, blood vessel development, leukocyte differentiation, KRAS signaling up (Fig 5).

Association between Risk Score and Response to Immunotherapy

To explore the role of the risk score in the immunotherapy response of PTCs. We compared the expression levels of PD-1, PD-L1, CTLA-4, TIM-3, LAG-3, and KLRB1 between high- and low-risk subgroup. The expression of the above immune checkpoints in the low-risk subgroup were higher than the high-risk

subgroup (Fig 6A). Meanwhile, we found that patients in high-risk subgroup have better to immunotherapy compared with those in low-risk subgroup ($P=0.022$) (Fig 6B). Besides, we separated patients into response and no response groups according to the predicted response to ICB. Interestingly, the ssGSEA scores including central memory CD4 T cell, eosinophil, monocyte, T follicular helper cell, and type17 T helper cell in the no response groups were significantly higher than the response groups ($P < 0.001$) (Fig 6C).

Construction of the m5C-lncRNAs Prognosis Nomogram

To further improve the accuracy of our survival prognostic model, we developed an integrated nomogram by combining immune risk model and several predictable clinical features. Univariate and multivariate Cox regression analysis showed that age (OR = 1.15; 95% CI, 1.08–1.22; and $P < 0.001$), risk score (OR = 2.72; 95% CI, 1.54–4.79; and $P < 0.001$), and IRS (OR = 1.69; 95% CI, 1.13–2.54; and $P = 0.011$) were the independent prognostic factors for PTCs (Table 1). Subsequently, an integrated nomogram was established (Fig 7A). The C-index and AUC of the nomogram were 0.968 and 0.969 (Fig 7B), respectively, indicating good discrimination. Moreover, the calibration plots showed excellence in predicting the 3- and 5-year overall survival (Fig 7C). Besides, decision curve analysis demonstrated the integrated nomogram can achieve a greater net benefit than any others (Fig 7D). These data suggesting advantageous usability of our prognostic prediction model in clinical practice.

Validation the expression level of m5C-lncRNAs

To verify the m5C-lncRNAs expression level of PTC tissues and cells for subsequent functional and molecular experiments, we performed RT-qPCR analysis on m5C-lncRNAs involved in the construction of nomogram. Our results showed that the *ZNF451-AS1* and *PPP1R12A-AS1* were downregulated in PTC tissues compared with

normal samples (Fig 8A). Furthermore, we further utilized 9 cell lines verify m5C-lncRNAs expression at mRNA level in vitro (Fig 8B).

Discussion

Despite with relatively favorable prognosis, PTCs still have a risk of disease recurrence and persistence (H. Kim, et al. 2018). Identifying the high-risk patients in the early stage is vital for clinicians to adopt more aggressive treatment. Currently, the risk stratification and prognostic prediction for PTCs were mainly dependent on the AJCC staging system (S. A. Ghaznavi, et al. 2018). Considering the heterogeneity of PTCs, it is limited to recognize patients with progression in the early phase. Hence, developing novel molecular indicators that effectively predict the prognosis of PTCs are necessary and urgent. In this study, we found that m5C regulators were significantly lower expressed in PTC, and then used six m5C-lncRNAs signatures constructed a prognostic risk model and IRS model. Meantime, we analyzed the role of the risk score in tumor immune infiltration and immunotherapy response of PTCs. Finally, we constructed an integrated nomogram by combining immune risk model and several

predictable clinical features and it yielded a superior performance. To our knowledge, this is the first study to report the role of m5C-lncRNAs in PTC.

In recent years, with the development of high-throughput sequencing and mass spectrometry (A. R. Quinlan, et al. 2010), more and more RNA modifications have been detected. m5C is one of the common post-transcriptional modifications, which closely related to cancer tumorigenesis and disease pathogenesis (P. Nombela, et al. 2021, (L. Trixl, et al. 2019). Previous studies have demonstrated that m5C modification is implicated in the initiation and progression of liver carcinoma (Z. Sun, et al. 2020, (C. Xue, et al. 2021). Also, there is existing evidence showed that m5C related-gene regulators can be used as diagnostic and prognostic biomarkers of carcinoma (S. Xiang, et al. 2020, (L. Sun, et al. 2020, (W. Zhao, et al. 2019). As mentioned above, m5C is found in a wide range of RNAs in all life domains but has also been detected in lncRNA and small ncRNAs (T. Amort, et al. 2014). m5C deposition in ncRNA and affects its processing into regulatory small RNA (S. Hussain, et al. 2013, (A. A. Sajini, et al. 2019). In one report, conventional bisulfite sequencing revealed marked disappearance of known target m⁵C methylation in the non-coding *RPPH1* when NSUN2 knockdown (J. E. Squires, et al. 2012). More recently, another study reported that aberrant NSUN2 mediated methylation of *H19* lncRNA by recruiting the G3BP1 oncoprotein and thus affects tumor proliferation and invasion (Z. Sun, et al. 2020). Besides, Chen et al (J. Pan, et al. 2021). implemented bioinformatic analysis to mining m5C related prognostic lncRNAs in TCGA datasets and successfully establish a risk score model to forecast the overall survival of lung adenocarcinoma patients. Taken together, these researches confirmed m5C methylation modification can target or modulate lncRNAs to influence carcinoma development. In this study, we provide a new insight into the interplay between lncRNAs and m5C modification in PTCs.

In addition, the tumor microenvironment has received extensive attention. The tumor microenvironment consists of stromal cells, immune cells, extracellular matrix, lymphatic and vascular system, and bioactive molecules, which play a key role in cancer development, immune escape, and therapeutic efficacy (M. Binnewies, et al. 2018, (X. Xu, et al. 2019). Several studies have shown a complexity and strong interrelationship between tumor microenvironment infiltrating immune cells and m5C modification. Schoeler et al (K. Schoeler, et al. 2019). demonstrated that *TET* enzymes promote a permissive chromatin state in antibody-mediated immunity. *TET3* and *TET2* guide the transition of germinal center B cells to antibody-secreting plasma cells. In breast cancer, the aberrant *NSUN2* was associated with the upregulation of macrophages/monocytes and Tregs in the TME (W. Mou, et al. 2015). To investigate the role of immune microenvironment on prognosis. We constructed a prognostic IRS model, and the AUC value was 0.764, indicating favorable discrimination performance. Besides, the ESTIMATE and stroma scores were notably higher, whereas the tumor purity was significantly lower in the high-IRS than those in the low-IRS subgroup. The abundance levels and proportions of B cell naive, plasma cell, M1 macrophages, mast cells activated were higher in high-IRS subgroup. Similarly, MCP-counter algorithms analysis showed that the level of B cell, monocytes, myeloid dendritic cells were higher in high-IRS subgroup. The above results demonstrated high-IRS patients have a lower proportion of antitumor immune cells and higher proportion of tumor-promoting immune cells.

Currently, immune checkpoint inhibitors (ICB), such as anti-CTLA-4 and anti-PD-L1/PD-1 antibodies, elicit durable and effective responses in some solid tumors (Q. Lei, et al. 2020, (S. Kumagai, et al. 2020). However, the efficacy and side effects of each patient in the treatment show individual differences (T. Akin Telli, et al. 2020). Therefore, it is necessary to identify patients who might benefit from ICB therapy. In our study, the expression levels of PD-1, PD-L1, CTLA-4, TIM-3, LAG-3, and KLRB1 were significantly lower in high-risk subgroup, and patients with high-risk score tended to have better responses to ICB. This finding indicates that ICB therapy might be more effective for this specific population. Therefore, m5C-lncRNAs risk score may be useful immunotherapy predictor of PTC outcome in clinical practice. The underlying mechanism of the relationship between m5C modification and ICB require further exploration.

Furthermore, to improve the accuracy of our survival prognostic model, we developed an integrated nomogram by combining immune risk model and several predictable clinical features. Although the TNM 8th edition has made great improvements in predicting prognosis on the basis of the TNM 7th edition (B. Cavalheiro, et al. 2021), there is still no concept of molecular profile for prognostic risk stratification and insufficient to distinguish the risk of mortality in PTCs. Several studies attempted to introduce BRAF V600E and TERT mutations to make up for the shortcomings of the existing tumor staging (J. D. Prescott, et al. 2012, (J. Park, et al. 2021). In this study, the C-index and AUC of the nomogram indicated good discrimination of the model. Moreover, the calibration plots showed excellence in predicting the 3- and 5-year overall survival and decision curve analysis demonstrated the integrated nomogram can achieve a greater net benefit than any others. This study constructed the nomogram based on the expression of m5C-lncRNAs and revealed that risk score was significantly related with prognosis, immune infiltration, and the efficacy of immunotherapy.

However, there are also several limitations in the present study. First, our main findings were based on accessed databases, resulting in an inevitable selection bias in clinical and genetic data. Second, we only analysis the correlation between m5C regulators and lncRNAs, lacking of experiments such as RNA-BisSeq, miCLIP, and RNA-seq to further confirm m5C modification sites on lncRNAs. Third, due to the limited research budgets, we only used real-time polymerase chain reaction to validate the level expression of 4 m5C-lncRNAs, lacking of function verification and regulation mechanism research. Fourth, the cutoff value of the risk score model was defined as the median risk score. Actually, this definition lacks rigor and requires larger clinical patient cohorts to verify. Last but not least, the research on the mechanism of immune difference is based on the inference of a variety of algorithms and has not been verified by experiments. Hence, future clinical and experimental studies are necessary to validate the application of our survival prediction model in clinical practice.

Conclusion

In summary, with the aid of bioinformatics analysis methods, our study comprehensively interpreted the influence of m5C modification on PTC from the perspective of immune infiltration and lncRNA. We found that m5C-lncRNAs could predict the clinical prognostic risk and regulate the TME in PTCs. Besides, risk score model based on the expression level of m5C-lncRNAs could be used to screen suitable patients who

can benefit from for ICB therapy. Additionally, an integrated nomogram combining infiltrating immune cells and m5C-lncRNAs exhibited good predictive performance. These findings indicated that m5C-lncRNAs have a promising value in tumor prognosis and individual immunotherapy.

Declarations

Acknowledgments

We thank the TCGA database for gene expression and clinical data.

Author Contributions

W.W conducted the formal analysis and wrote the original draft; All authors contributed to statistical analysis or interpretation of the data, drafting of the manuscript. All authors read and approved the final submitted manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (grant No. 81672885), the Natural Science Foundation of Hunan Province (grant No. 2019JJ40500) and Innovative Foundation for graduate students of Hunan Province (grant No. 2020zzts259).

Availability of data and materials

All data not included in the manuscript is available from the supplementary materials.

Ethics approval and consent to participate

Informed consent was obtained from all the participants and this study was approved by the Ethics Committee of Xiangya Hospital of Central South University.

Consent for publication

All authors read and agreed to the content of the final manuscript, and consented to publish this material.

Competing interests

The authors declare that they have no competing interests.

References

1. Wang W., Yang Z. and Ouyang Q. A nomogram to predict skip metastasis in papillary thyroid cancer. *World J Surg Oncol.*2020.18:167.
2. Zhang S., et al. Cancer incidence and mortality in China, 2015. *Journal of the National Cancer Center.*2021.1:2-11.
3. Wang W., Meng C., Ouyang Q., Xie J. and Li X. Magnesium: an independent risk factor of hypocalcemia after thyroidectomy. *Cancer Manag Res.*2019.11:8135-8144.
4. Wang W., et al. Identification and validation of potential novel biomarkers to predict distant metastasis in differentiated thyroid cancer. *Annals of Translational Medicine.*2021.9:1053-1053.
5. Nombela P., Miguel-Lopez B. and Blanco S. The role of m(6)A, m(5)C and Psi RNA modifications in cancer: Novel therapeutic opportunities. *Mol Cancer.*2021.20:18.
6. Dominissini D. and Rechavi G. 5-methylcytosine mediates nuclear export of mRNA. *Cell Res.*2017.27:717-719.
7. Trixl L. and Lusser A. The dynamic RNA modification 5-methylcytosine and its emerging role as an epitranscriptomic mark. *Wiley Interdiscip Rev RNA.*2019.10:e1510.
8. Pan J., Huang Z. and Xu Y. m5C RNA Methylation Regulators Predict Prognosis and Regulate the Immune Microenvironment in Lung Squamous Cell Carcinoma. *Front Oncol.*2021.11:657466.
9. Chen H., et al. M(5)C regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in lung adenocarcinoma. *Transl Lung Cancer Res.*2021.10:2172-2192.
10. Li Q., et al. NSUN2-Mediated m5C Methylation and METTL3/METTL14-Mediated m6A Methylation Cooperatively Enhance p21 Translation. *J Cell Biochem.*2017.118:2587-2598.
11. Chen H., et al. m(5)C modification of mRNA serves a DNA damage code to promote homologous recombination. *Nat Commun.*2020.11:2834.
12. Chen X., et al. 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. *Nat Cell Biol.*2019.21:978-990.
13. Li Y., et al. Novel long noncoding RNA NMR promotes tumor progression via NSUN2 and BPTF in esophageal squamous cell carcinoma. *Cancer Lett.*2018.430:57-66.
14. Carella A., et al. Epigenetic downregulation of TET3 reduces genome-wide 5hmC levels and promotes glioblastoma tumorigenesis. *Int J Cancer.*2020.146:373-387.
15. Ferrari S.M., et al. Immune and Inflammatory Cells in Thyroid Cancer Microenvironment. *International journal of molecular sciences.*2019.20.

16. Hurst Z., et al. Risk Haplotypes Uniquely Associated with Radioiodine-Refractory Thyroid Cancer Patients of High African Ancestry. *Thyroid*.2019.29:530-539.
17. Motzer R.J., et al. Molecular Subsets in Renal Cancer Determine Outcome to Checkpoint and Angiogenesis Blockade. *Cancer Cell*.2020.38:803-817 e804.
18. Prat A., et al. Immune-Related Gene Expression Profiling After PD-1 Blockade in Non-Small Cell Lung Carcinoma, Head and Neck Squamous Cell Carcinoma, and Melanoma. *Cancer Res*.2017.77:3540-3550.
19. Miao Y.R., et al. ImmuCellAI: A Unique Method for Comprehensive T-Cell Subsets Abundance Prediction and its Application in Cancer Immunotherapy. *Adv Sci (Weinh)*.2020.7:1902880.
20. Kim H., et al. Prognosis of Differentiated Thyroid Carcinoma with Initial Distant Metastasis: A Multicenter Study in Korea. *Endocrinol Metab (Seoul)*.2018.33:287-295.
21. Ghaznavi S.A., et al. Using the American Thyroid Association Risk-Stratification System to Refine and Individualize the American Joint Committee on Cancer Eighth Edition Disease-Specific Survival Estimates in Differentiated Thyroid Cancer. *Thyroid*.2018.28:1293-1300.
22. Quinlan A.R.and Hall I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*.2010.26:841-842.
23. Sun Z., et al. Aberrant NSUN2-mediated m(5)C modification of H19 lncRNA is associated with poor differentiation of hepatocellular carcinoma. *Oncogene*.2020.39:6906-6919.
24. Xue C., Zhao Y., Li G.and Li L. Multi-Omic Analyses of the m(5)C Regulator ALYREF Reveal Its Essential Roles in Hepatocellular Carcinoma. *Front Oncol*.2021.11:633415.
25. Xiang S., et al. m(5)C RNA Methylation Primarily Affects the ErbB and PI3K-Akt Signaling Pathways in Gastrointestinal Cancer. *Front Mol Biosci*.2020.7:599340.
26. Sun L., et al. Large-scale transcriptome analysis identified RNA methylation regulators as novel prognostic signatures for lung adenocarcinoma. *Ann Transl Med*.2020.8:751.
27. Zhao W., et al. Predictive Factors of Lateral Lymph Node Metastasis in Papillary Thyroid Microcarcinoma. *Pathol Oncol Res*.2019.25:1245-1251.
28. Amort T., et al. Long non-coding RNAs as targets for cytosine methylation. *RNA Biology*.2014.10:1002-1008.
29. Hussain S., et al. NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. *Cell Rep*.2013.4:255-261.
30. Sajini A.A., et al. Loss of 5-methylcytosine alters the biogenesis of vault-derived small RNAs to coordinate epidermal differentiation. *Nat Commun*.2019.10:2550.

31. Squires J.E., et al. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res.*2012.40:5023-5033.
32. Pan J., Huang Z.and Xu Y. m5C-Related lncRNAs Predict Overall Survival of Patients and Regulate the Tumor Immune Microenvironment in Lung Adenocarcinoma. *Front Cell Dev Biol.*2021.9:671821.
33. Binnewies M., et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med.*2018.24:541-550.
34. Xu X., et al. Association of Germline Variants in Natural Killer Cells With Tumor Immune Microenvironment Subtypes, Tumor-Infiltrating Lymphocytes, Immunotherapy Response, Clinical Outcomes, and Cancer Risk. *JAMA Netw Open.*2019.2:e199292.
35. Schoeler K., et al. TET enzymes control antibody production and shape the mutational landscape in germinal centre B cells. *FEBS J.*2019.286:3566-3581.
36. Mou W., et al. Expression of Sox2 in breast cancer cells promotes the recruitment of M2 macrophages to tumor microenvironment. *Cancer Lett.*2015.358:115-123.
37. Lei Q., Wang D., Sun K., Wang L.and Zhang Y. Resistance Mechanisms of Anti-PD1/PDL1 Therapy in Solid Tumors. *Front Cell Dev Biol.*2020.8:672.
38. Kumagai S., et al. The PD-1 expression balance between effector and regulatory T cells predicts the clinical efficacy of PD-1 blockade therapies. *Nat Immunol.*2020.21:1346-1358.
39. Akin Telli T., et al. PD-1 and PD-L1 inhibitors in oesophago-gastric cancers. *Cancer Lett.*2020.469:142-150.
40. Cavalheiro B., et al. Survival in differentiated thyroid carcinoma: A comparison between the 7th and 8th editions of the AJCC/UICC TNM staging system and the ATA initial risk stratification system. *Head & neck.*2021.
41. Prescott J.D., et al. BRAF V600E status adds incremental value to current risk classification systems in predicting papillary thyroid carcinoma recurrence. *Surgery.*2012.152:984-990.
42. Park J., et al. TERT Promoter Mutations and the 8th Edition TNM Classification in Predicting the Survival of Thyroid Cancer Patients. *Cancers (Basel).*2021.13.

Figures

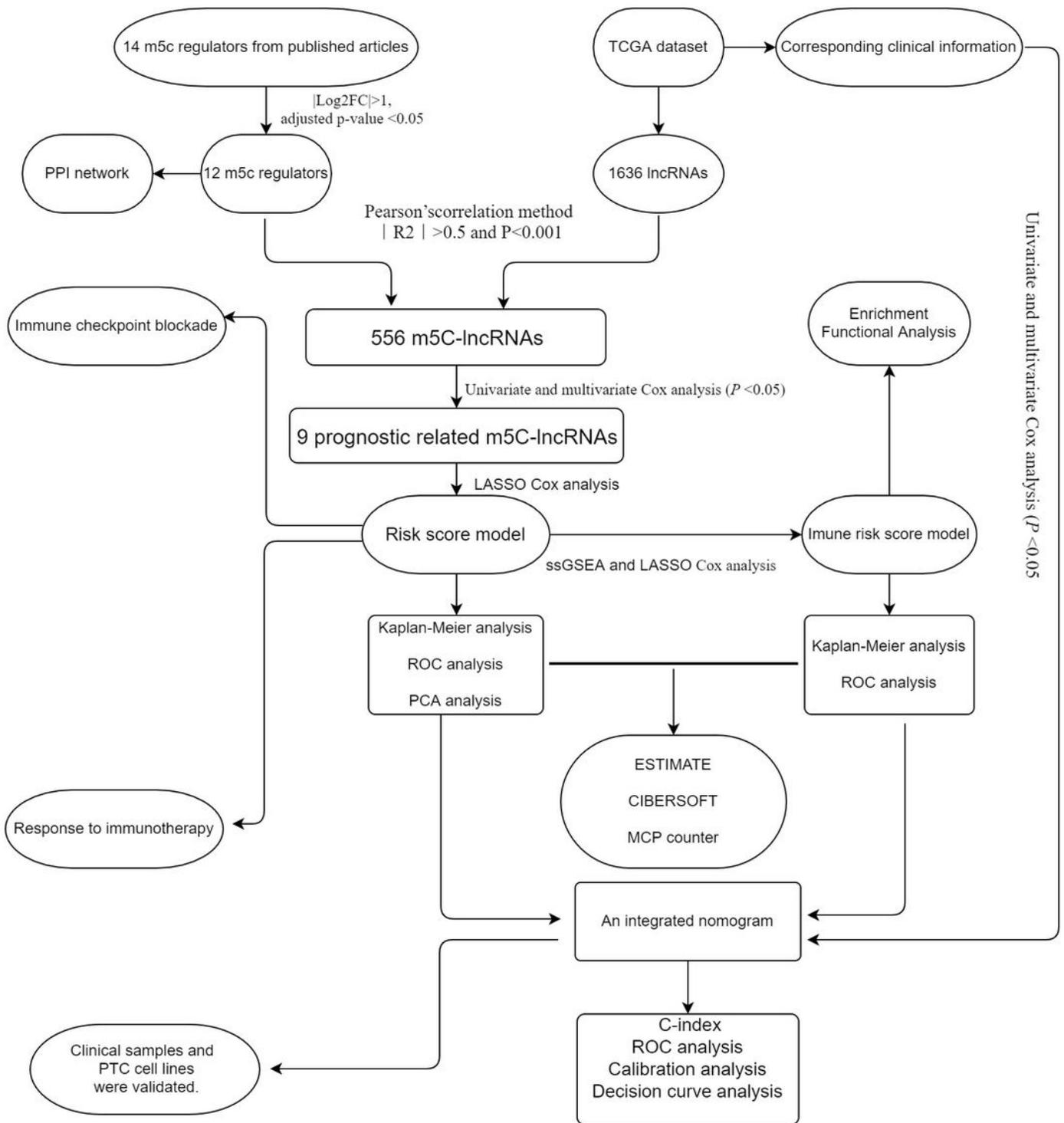


Figure 1

Study flow chart.

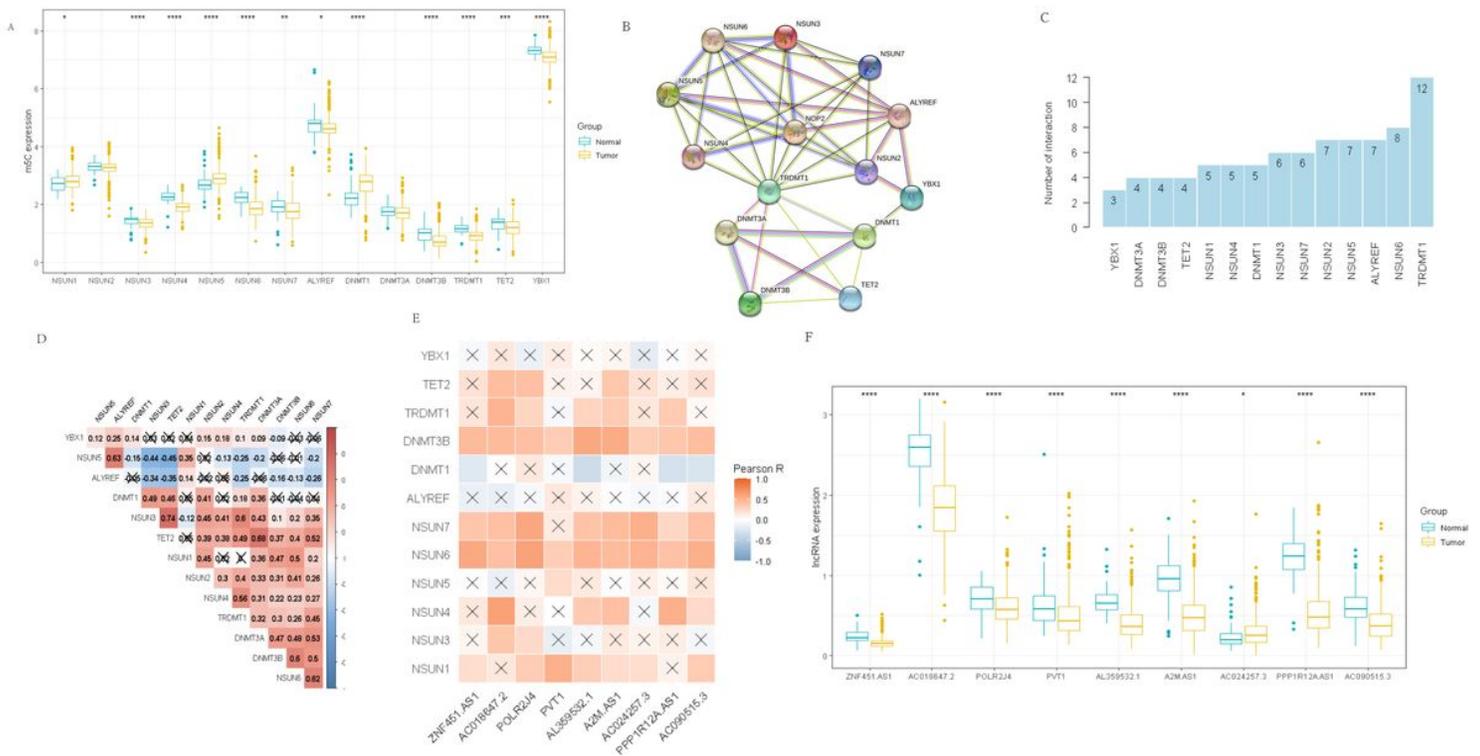


Figure 2

The profile of m5C-IncRNAs regulators. A: Differential expression of 14 m5C regulators; B-C: PPI network shows the interaction between the differentially expressed genes of m5C regulators. D: Pearson correlation analysis of the m5C regulators. E: The correlations between m5C-IncRNAs and m5C regulators. F: Differential expression of 9 m5C-IncRNAs regulators. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

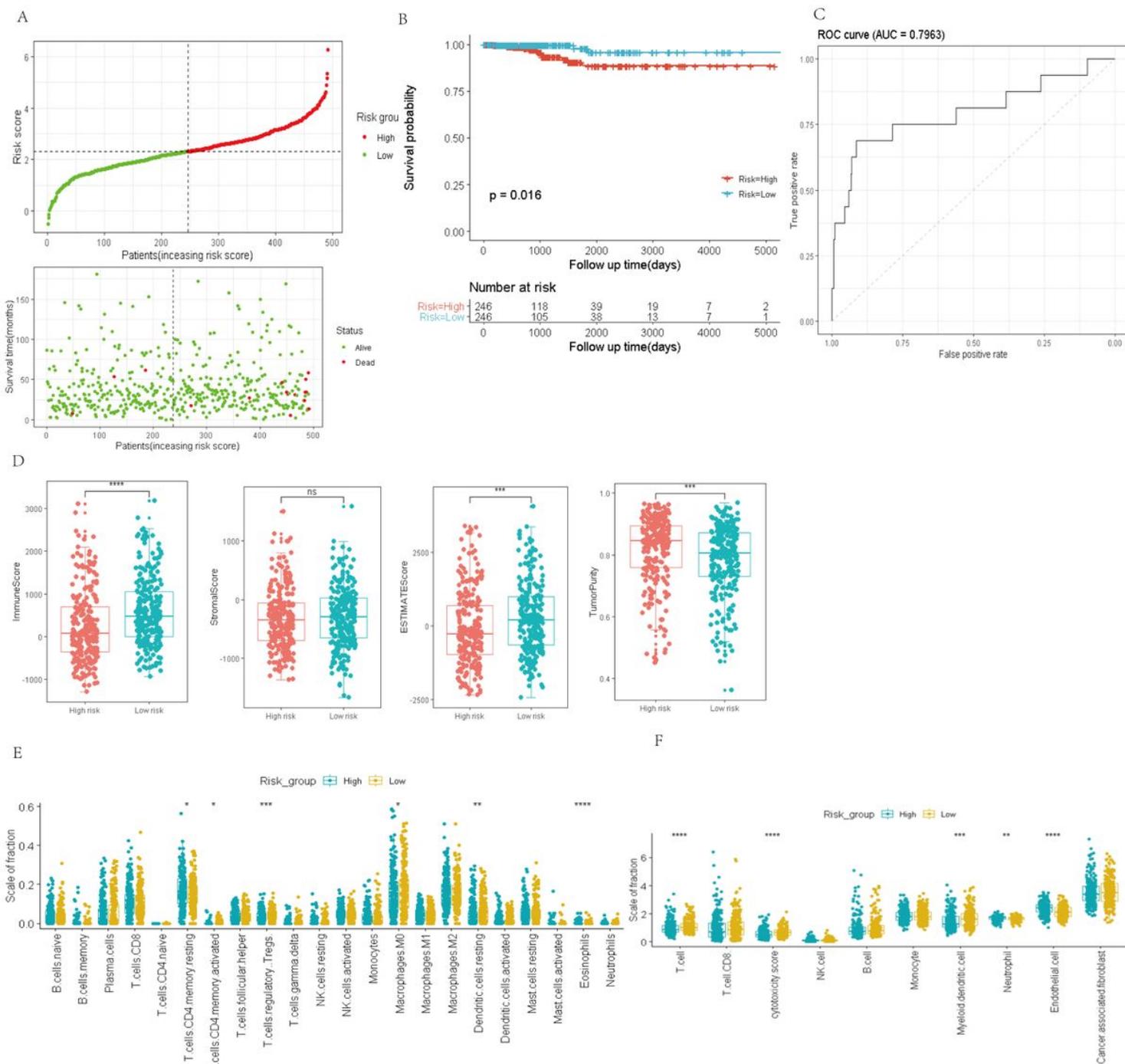


Figure 3

The relationship between risk score and PTC prognosis and tumor immune microenvironment. A: Distributions of risk scores and survival status of PTCs. B: Kaplan-Meier curves showed that patients in the low-risk subgroups had better prognosis than high-risk subgroups ($P = 0.016$). C: The area under the ROC curve (AUC) of the risk score. D-F: The immune cell infiltration landscape in the risk score subgroups was analyzed by ESTIMATE, CIBERSOFT, MCP counter, respectively. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and ns: no significance.

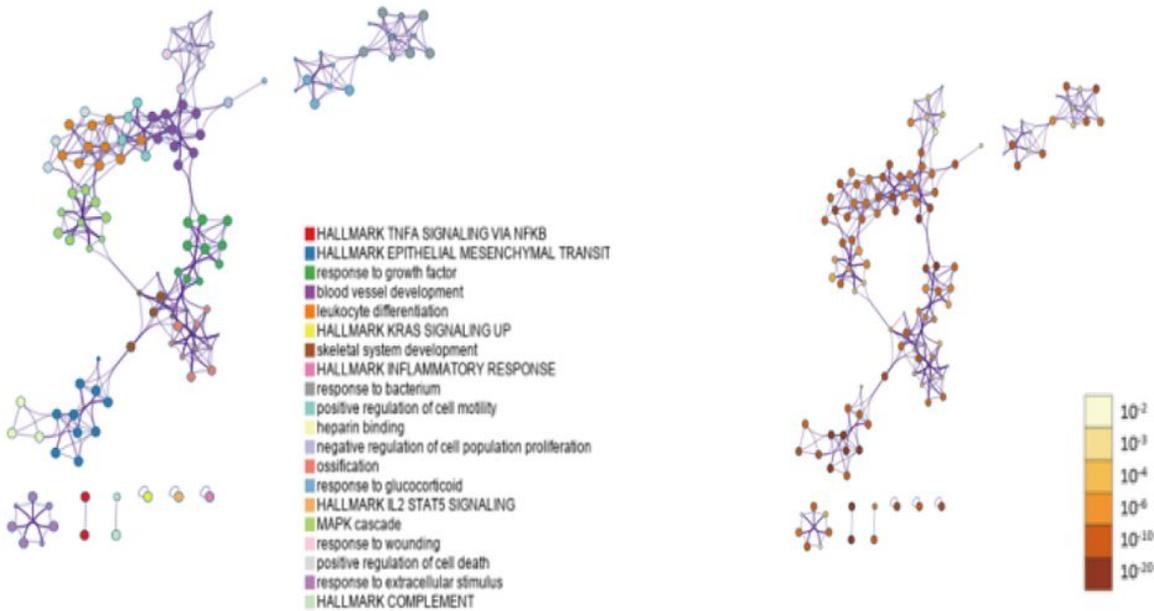
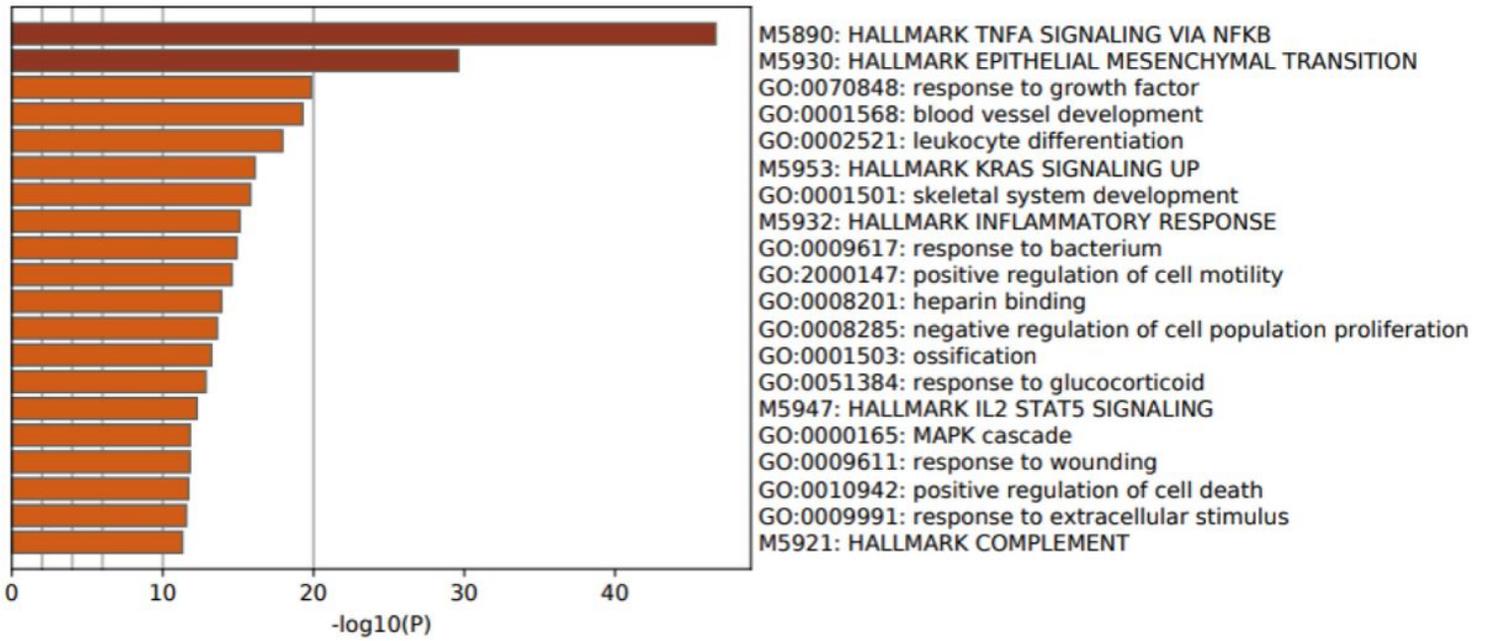


Figure 5

The functional and pathway enrichment analysis of m5C-lncRNAs by Metascape.

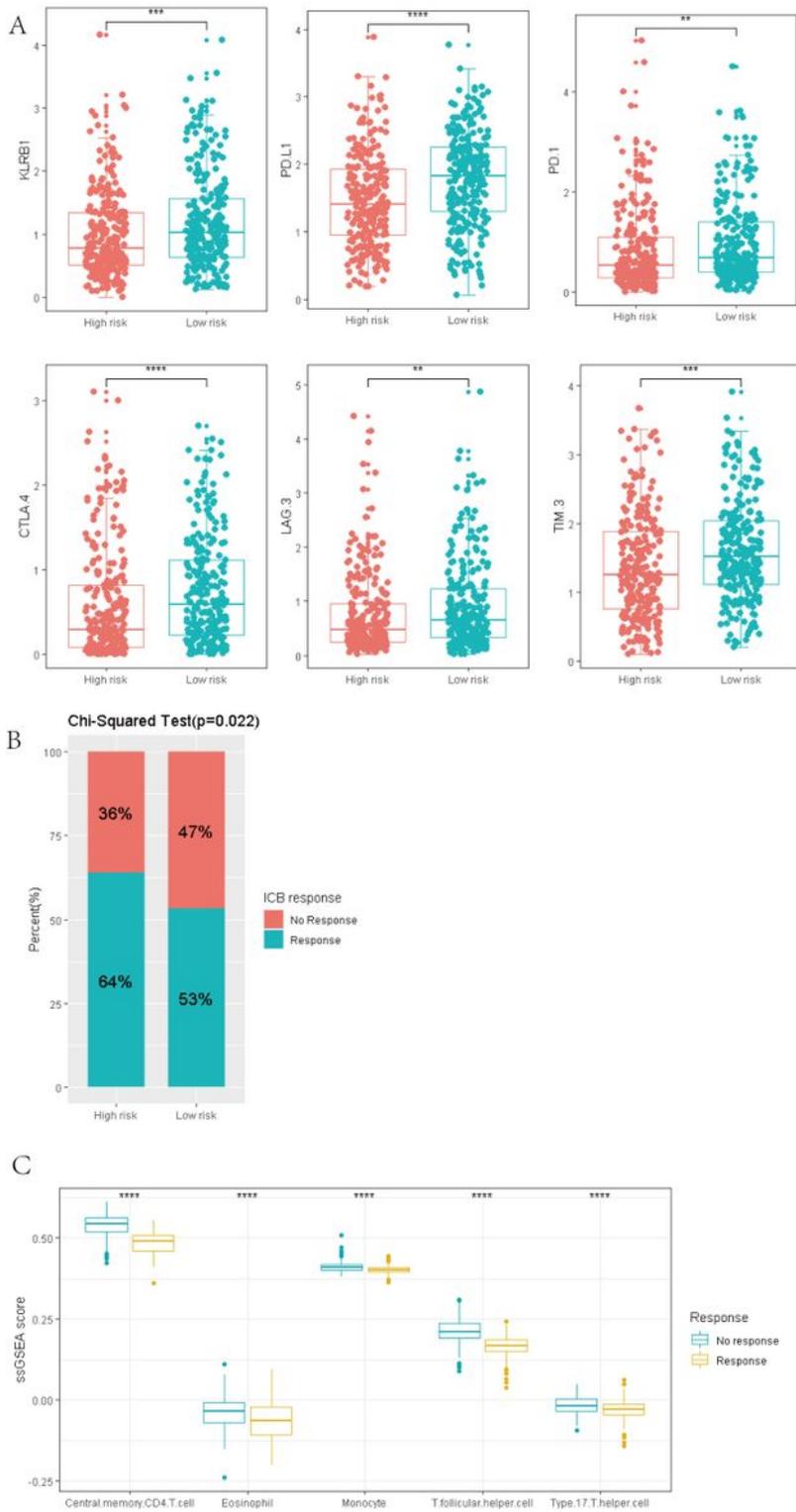


Figure 6

Prediction of the response to immunotherapy. A: The expression levels of six immune checkpoints between high- and low-risk subgroups. B: The rate of response to immune checkpoints blockade therapy. C: The ssGSEA scores of five immune cells were lower in the response group. **p < 0.01, ***p < 0.001.

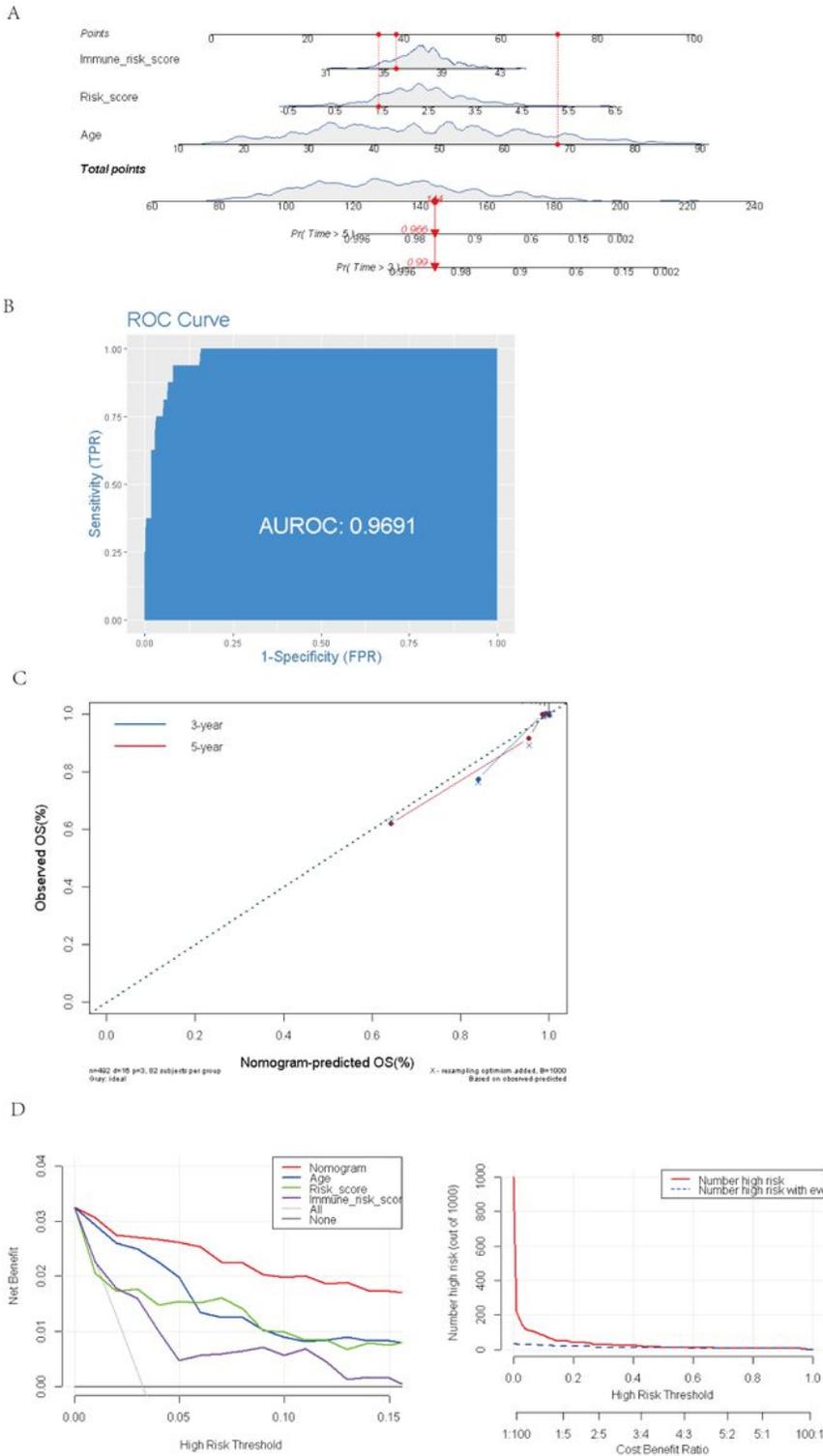


Figure 7

Construction of the m5C-IncRNAs prognosis nomogram. A: Nomogram was developed to predict the survival probability. B: The AUC value of the nomogram. C: The calibration plots of nomogram model. D: Decision curve analysis and clinical impact curve of the nomogram.

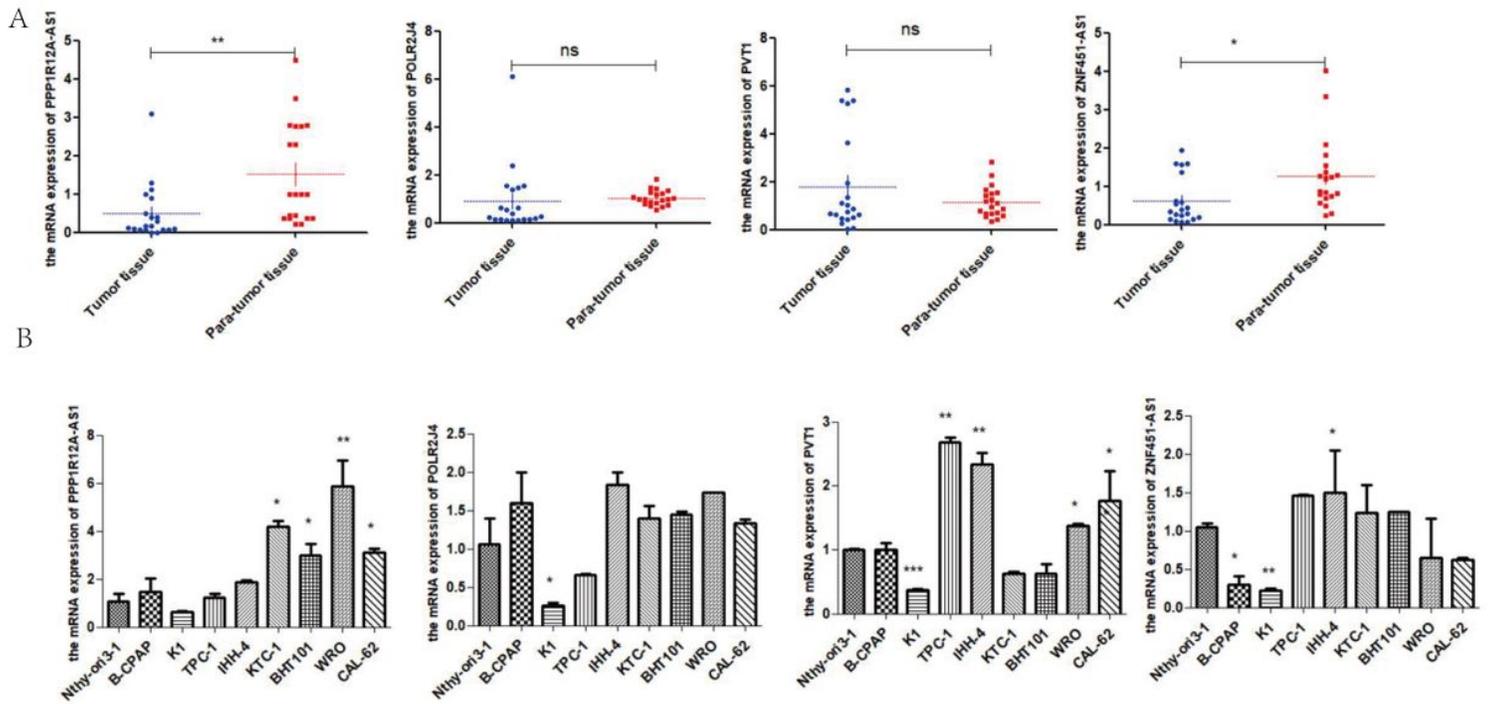


Figure 8

Validation the expression level of m5C-lncRNAs. A-B: the expression level of m5C-lncRNAs in 20 pairs PTC tissues and nine cells line. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and ns: no significance.

Supplementary Files

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

- [FigureS1.tif](#)
- [Supplementarymaterials.docx](#)
- [FigureS2.tif](#)
- [FigureS3.tif](#)