UHRF1, NSUN2, and NEIL1 were Detected as Clinical Biomarker Candidates in Prostate Adenocarcinoma with Potential Roles in Disease Pathogenesis

Buse Nur Kahyaoğulları[1], Turan Demircan[2,3]

1) Institute of Health Sciences, Muğla Sıtkı Koçman University, Muğla, Turkey

2) Medical Biology Department, School of Medicine, Muğla Sıtkı Koçman University, Muğla, Turkey

3) Department of Bioinformatics, Muğla Sıtkı Koçman University, Muğla, Turkey

Correspondence: Assoc.Prof.Turan Demircan

E-mail: turandemircan@gmail.com

Tel.: +90 252 211 4846

Abstract

Prostate cancer (PCa) is the most commonly diagnosed cancer in men. Expanding evidence suggests a significant association between cancer progression and RNA modifications. However, our knowledge of the link between m5C and hm5C pathways with PCa is limited. Therefore, we aimed to explore the diagnostic and prognostic values of m5C and hm5C regulators in PCa. In this study, genetic alterations in m5C and hm5C regulators were identified using publicly available databases. Differentially expressed genes in these pathways between tumor and nontumor samples, correlation among m5C and hm5C pathway members, and prognostic value of the regulators were evaluated. Furthermore, enrichment of gene ontology (GO) terms and KEGG pathways was carried out. Obtained results unveiled the mRNA level differences as the key genetic alterations for m5C and hm5C regulators between tumor and nontumor samples. *UHRF1, TET3,* and *NEIL1* were significantly upregulated in tumor samples compared to nontumor ones, whereas *EGR1* was significantly downregulated. *UHRF1, DNMT1, NSUN2, NSUN4, C1orf77, C3orf37, WDR77, NEIL1*, and *TDG* genes were identified as candidate prognostic markers of overall survival. The upregulated genes in patient samples with genetic alterations in m5C and hm5C pathways enriched cell cycle-related processes. In summary, our findings suggest that the m5C and hm5C regulators might play a role in PRAD development by activation of proliferation, and the *UHRF1, NSUN2*, and *NEIL1* genes have the potential to be utilized as clinical biomarkers.

Keywords: Prostate cancer, Epitranscriptomics, Biomarker, TCGA, 5-methylcytosine (m5C), 5-hydroxymethylcytosine (hm5C)

Introduction

Prostate cancer (PCa) is the second most diagnosed cancer (13.5%) in men after lung cancer (14.5%), with a 6.7% death rate[1]. Incidence and mortality of prostate cancer increase with age [2, 3]. Biopsy, digital rectal examination, transrectal ultrasound, and prostate-specific antigen test are used to diagnose PCa [4]. Histological imaging is mainly used in the staging of PCa, and the Gleason scoring system based on the morphological structure of prostate cancer is commonly adopted [5]. In this scoring method, a Gleason score ≤6 represents a good prognosis (grade 1), seven is used for grades 2 and 3, eight refers to grade 4, and finally, nine and ten portrey the worst prognosis with grade 5 [6]. Current treatment options for PCa are androgen hormone deprivation therapy (ADT), surgical removal of the tumor, and radiotherapy [7]. Androgen hormone has a vital role in PCa progression, and therefore several treatments to target androgen hormone have been developed [8]. ADT has a prevalent use to treat PCa [9], and although patients initially respond positively to ADT, almost all patients eventually develop

castrate-resistant prostate cancer [10]. Hence, identification of new potential candidate genes to cure the PCa is required.

Epitranscriptomics, employing next-generation sequencing to decode the RNA modifications, has recently become one of the emerging fields for researchers [11]. By now, more than 170 types of RNA modifications have been described [12], and among them, N6-methyladenosine (m6A), N1-methyladenosine (m1A), 5-methylcytosine (m5C), 5-hydroxymethylcytosine (hm5C), pseudouridine (Ψ), inosine (I), and uridine (U) are commonly detected mRNA modifications [13]. Due to RNA modifications, mRNA stability, export, localization, and translation rate are affected [13]. Proteins that regulate the dynamic and reversible RNA modifications are classified into three groups. Enzymes that add methyl groups are called 'writers', binding proteins that recognize and bind the methylated residues are grouped as 'readers', and the proteins which catalyze the demethylation reaction are known as 'erasers' [14].

m5C was first identified in DNA as one of the epigenetic regulation mechanisms [15]. Recently, m5C modification was also determined in various RNA species, including rRNA, tRNA, non-coding RNA, and mRNA, with potential effects on RNA export, ribosome assembly, translation, and RNA stability [16]. m5C modification is abundantly found in coding and non-coding regions along a transcript, suggesting that m5C modification may play a significant role in translation rate and the post-transcriptional regulation of the mRNA ,[16]. m5C modification is provided by the writer enzymes member of DNA methyltransferases (DNMTs), and NOP2/SUN RNA methyltransferase family (NSUN), particularly by the DNMT2 and NSUN2 [17, 18]. Localization of DNMT2 as a shuttling protein in the nucleus and cytosol during the cell cycle indicates that this methyltransferase actively modifies nucleic acids in both cellular compartments, which has been validated experimentally for a variety of RNA types besides the DNA [17, 19]. Moreover, previous reports demonstrated that the NSUN family could target mRNA, rRNA, mitochondrial DNA, and non-coding RNAs, apart from tRNAs, to add a methyl group to the cytosine base [20]. Aly/REF Export Factor (ALYREF), a reader protein in the m5C pathway, takes a role in exporting mRNAs modified primarily by NSUN2 [21]. Based on the accumulated evidence, the ten-eleven translocation (TET) protein family is considered the eraser of the m5C pathway by catalyzing the oxidation reaction on m5C in DNA and RNA to form a hm5C group [22, 23]. On the contrary, TET enzymes are classified as 'writers' in the hm5C pathway. The formation of hm5C in mammalian RNAs is evident in the previous studies [23]. Although the mechanism is not fully understood yet, translation efficiency is increased for the mRNAs modified with hm5C [24].

RNA modifications affect many cellular processes and play a role in various human cancers [25]. The poor prognosis of bladder cancer due to the 5mC methylation of the *HDGF* gene is associated with the activity of NSUN2 writer protein [26]. In hepatocellular carcinoma (HCC), the detected copy number variation (CNV) of m5C regulators and their expression levels have been correlated with the progression of the disease [27]. In addition, 5mC modifications in HCC, catalyzed by NSUN4, and recognized by the reader ALYREF protein, resulted in dysregulation of

cell cycle and mitosis and therefore accounted for HCC development [27]. In glioblastoma multiforme (GBM), 5mC modification on miRNAs decreases the binding efficiency of miRNAs to target mRNAs, and hence methylation of miRNAs increases the gene expression [28]. m5C modification in miRNA-181a-5p exhibited a tumor-growth effect and has been related to a poor prognosis of GBM [28]. In ovarian cancer, differential expression of m5C regulators was detected in subtypes of cancer, and they are used to characterize the ovarian cancer subtypes [29]. Increased hm5C modification by TET enzymes has been demonstrated to prevent melanoma development [30]. In another study, the hm5C level decreased significantly in thyroid cancer patients with TERT promoter mutation [31].

The current study examined diagnostic and prognostic values of m5C and hm5C RNA regulators in PCa using TCGA pan-cancer datasets. An in-depth bioinformatics analysis was employed to unveil the link between genetic alteration in RNA regulators and PCa progression. Expression level differences between tumor and non-tumor samples were investigated, and a list of differentially expressed genes (DEGs) was identified. Several genes whose altered expression levels were associated with poor PCa progression, were detected as potential prognostic biomarkers. We believe that the findings of this study would be useful to understand the roles of RNA modification regulators in PCa development, and some of these could serve as valuable biomarkers of PCa.

Materials and Methods

TCGA Pan-cancer PRAD Dataset

The expression and clinical data for 550 samples (498 tumor and 52 non-tumor samples) from 498 PRAD patients were extracted from the TCGA database [32] by using the RTCGA package [33] in the R environment.

Genetic alterations for m5C and hm5C regulators

The cBioportal database [34] was used to identify genetic alterations of the m5C and hm5C regulators in PRAD patients. cBioportal is an open-access website containing multidimensional cancer genomic data and aims to provide researchers with access to various molecular profiles and clinical data. Frequency and type of genetic alterations for m5C and hm5C regulators were profiled on an oncoprint graph using the PRAD Pancancer dataset, and then we selected seventeen studies (Figure 1b, d) for analyzing alteration frequency.

Generated Heatmaps for m5C and hm5C Genes

To evaluate the expression levels of the m5C and hm5C regulators in prostate cancer, RNAseq data was extracted from the TCGA-PRAD cohort, and heatmaps were generated using 'ComplexHeatmap' package [35] in R language to display the expressions of m5C and hm5C regulators according to patients.

Principle Component Analysis and Expression Level Differences of m5C and hm5C Regulators

Principal component analysis (PCA) was performed using the 'ggfortify' package of R language to assess the discrimination capacity of m5C and hm5C regulators' expression levels between tumor and non-tumor samples. The same package was employed to visualize the results.

To test whether the expression levels of analyzed genes are significantly different in m5C and hm5C pathways between compared groups, the Wilcoxon test in 'rstatix' package was adopted. 'ggplot' package was implemented to display the results with significance levels on boxplots.

'Limma' package [36] in R language was exploited to identify the differentially expressed m5C and hm5C RNA methylation regulators between tumor and non-tumor samples. The cut-off was set as |FC| > 2 and p-value <0.05. The Differentially expressed genes (DEGs) in m5C and hm5C pathways were listed on separate tables.

Correlation Analysis

Correlation maps and heatmaps were generated to explore the expression trend of the genes among tumor samples for m5C and hm5C pathways. Furthermore, a similar analysis was conducted using the unified list of the genes gathered from these two pathways. 'Hmisc' package was utilized for correlation analysis, and 'corrplot' package was used along with the base functions in the R language for visualization of the results. Correlation coefficient values were obtained between -1 to 1. The values close to -1 or 1 for the genes were considered as a strong correlation, whereas those close to 0 indicated a weak correlation. A positive correlation was shown with a value between 0 and 1, and -1 to 0 values signified the negative correlation.

Kaplan-Meier overall survival analysis

For generated overall survival graphs, the clinical data (PRAD.clinical) was extracted from the 'RTCGA.clinical' package in the R environment. Kaplan-Meier estimation method was applied to test the prognostic value of the m5C and hm5C genes regarding the overall survival (OS) of prostate cancer patients with low or high m5C-hm5C gene expression. To determine the optimal cut-off and calculate the P-value for each gene 'survival' package in the R environment was used, and the list of obtained p-values was shown on a table. Significant genes (p-value <0.05) for m5C

or hm5C pathways were selected to draw the Kaplan-Meier survival curves by utilization of 'survminer' and 'ggplot2' R packages.

ROC Curves

Receiver operating characteristic (ROC) curve analysis was conducted utilizing the 'pROC' package [37] in R language to evaluate whether the expression level of selected m5C and hm5C regulators (*UHRF1, NSUN2*, and *NEIL1*) was distinctive for tumor and nontumor samples. Area under curve (AUC) value was calculated for each gene to estimate the sensitivity and specificity of the performed analysis.

Expression Levels of *UHRF1*, *NSUN2*, and *NEIL1* Genes in PRAD Patients with Different Gleason Score

Gleason classification is used for morphological grading for prostate adenocarcinoma. The association of clinical features with expression levels of the genes was assessed through the UALCAN online database [38]. Expression levels in tumor samples with different Gleason scores (from 6 to 10) were compared with expression levels in non-tumor samples. The results were illustrated on the graphs generated by the UALCAN database.

Enrichment Analyses for Altered and Unaltered Groups

We classified the tumor samples into two groups for both m5C and hm5C pathways. Tumor samples with genetic alterations for at least one gene formed the altered group and the unaltered group consisted of tumor samples without mutations for the genes in m5C or hm5C pathways. A similar grouping was used for the genes (*UHRF1*, *NSUN2*, and *NEIL1*) considered as putative prognostic markers.

Gene Ontology (GO) analysis was carried out to get all ontological terms enriched by DEGs for altered and unaltered groups. Molecular function (MF), biological process (BP), and cellular component (CC) terms were identified for each group with 'clusterProfiler' package [39] in R. The same package was implemented to unveil the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways in altered and unaltered groups. The applied parameters for GO and KEGG were as following: p-value <0.01, q-value <0.2, padjustmethod; 'BH'

Statistically significant terms and pathways were visualized on dotplot and barplot graphs using R2s 'ggplot2' and 'enrichplot' packages as performed elsewhere [40].

The communality of Enriched GO Terms for Altered and Unaltered Groups

Unique and common GO terms among altered and unaltered groups were found via Jvenn tool [41], as did elsewhere [42]. For both m5c and hm5C pathways, separate venn diagrams for altered and unaltered groups were generated.

The Human Protein Atlas Data

The human protein atlas database [43] was utilized to map the expression of the selected proteins in normal and tumor tissues by immunohistochemical staining (IHC). Representative images of m5C (*UHRF1* and *NSUN2*) and hm5C (*NEIL1*) proteins in tumor and normal tissues were retrieved and visualized. While the proteins were labeled with protein-specific antibodies, a blue-colored hematoxylin dye was used to create contrast.

Results

m5C and hm5C Regulators were Found as Frequently Mutated in Prostate Cancer

Genetic alterations of m5C and hm5C pathway genes in prostate cancer were visualized on an oncoprint map using the cBioportal database (Figure 1a, d). For the m5C pathway, the top five genes with the highest alteration frequency were *NSUN2* (52%), *NSUN5* (38%), *DNMT3A* (30%), *NOP2* (26%), and *TET3* (23%) (Figure 1a). Besides *TET3*, which is an intersected member of both RNA modification pathways, *PRMT5* (31%), *WDR77* (28%), *THYN1* (25%), and *SMUG1* (21%) were found as commonly altered genes among hm5C regulators (Figure 1d).

Alteration profiles at genetic levels for m5C and hm5C regulators were demonstrated on barplots generated by the cBioportal database (Figure 1b, e). For both of the pathways, amplification was the most frequently observed alteration type, followed by mutation and deep deletion alterations (Figure 1b, d). Furthermore, it was found that the majority of the m5C and hm5C genes were upregulated in most of the tumor samples compared to non-tumor ones in PRAD patients (Figure 1c, f). Lower expression of MECP2 and WDR76 was noteworthy (Figure 1c, f).

UHRF1, TET3, NEIL1, and *EGR1* were Identified as DEGs Between Tumor and non-Tumor Samples

PCA analysis was implemented to examine whether the m5C or hm5C regulators have a remarkable power to discriminate the tumor and nontumor samples (Figure 2a, d). As can be seen on PCA plots for the m5C (Figure 2a) and hm5C (Figure 2d) regulators, although the expression level of these genes has a considerable discrimination rate for both pathways (for m5C:PC1 = 28.8%, PC2 = 18.6%; for hm5C: PC1 = 24.2%, PC2 = 15.1%), tumor and non-tumor samples were still intermingled.

Then we applied the Wilcoxon test to calculate the p-value of each gene in m5C and hm5C pathways for tumor and nontumor samples without a fold change cut-off. Obtained results for each gene with significance level were displayed on boxplots (Figure 2b, e). Out of 5mC

regulators, *DNMT3A, DNMT3B, NOP2, NSUN2, NSUN5, TET3,* and *UHRF1* expression levels were significantly higher in tumor samples compared to the nontumor group (Figure 2b). On the other hand, the expression level of the *MECP2* gene was detected as significantly lower in tumor samples (Figure 2b). The hm5C regulators *ERH, NEIL1, PRMT5, SMUG1, TDG, TET3, WDR76,* and *WDR77* were upregulated significantly in tumor samples (Figure 2e). Differential expression analysis between tumor and non-tumor samples using LIMMA resulted in one gene (*UHRF1*) in the m5C pathway and three genes (*TET3, NEIL1,* and *EGR*) in the hm5C pathway, all having an adjusted p-value < 0.01 and FC>1.5 (Supplementary Table 1). *UHRF1, TET3,* and *NEIL1* were significantly upregulated in tumor samples, and *EGR1* was identified as a significantly downregulated gene of RNA modification pathway genes in tumor samples compared to nontumor samples.

Expression Levels of m5C and hm5C Regulatory Genes were Highly Correlated in PRAD Tumor Samples

To explore whether the genes in m5C or hm5C pathways were correlated 'Spearman' correlation test was applied to the analyzed genes, and the obtained results were presented in Figure 2. We observed a positive correlation trend for most of the genes in the m5C pathway (Figure 2c and Supplementary Table 2). Particularly, *TET2* and *TET3* (correlation coefficient: 0.71, p-value<0.001), *DNMT1* and *UHRF1* (correlation coefficient: 0.64, p-value<0.001), and *DNMT3A* and *TET3* (correlation coefficient: 0.62, p-value<0.001) were the most positively correlated gene pairs among the m5C regulators. In contrast, the top negatively correlated gene pairs were found as follows: *NSUN5* and *TET2* (correlation coefficient: -0.52, p-value<0.001), *NSUN5* and *MECP2* (correlation coefficient: -0.44, p-value<0.001), and *NSUN3* and *NSUN5* (correlation coefficient: -0.43, p-value<0.001).

On the other hand, for the hm5C pathway, the number of negatively and positively correlated gene pairs was close (Figure 2f and Supplementary Table 3). Aside from *TET2* and *TET3*, a positive correlation was manifested mainly in *TET3* and *TDG* (correlation coefficient: 0.64, p-value<0.001), *TET1* and *TET3* (correlation coefficient: 0.57, p-value<0.001), and *TET1* and *TET2* (correlation coefficient: 0.57, p-value<0.001) gene pairs. Contrary, the most negatively correlated gene pairs were detected as *TET2* and *SMUG1* (correlation coefficient: -0.55, p-value<0.001), *TET2* and *WDR77* (correlation coefficient: -0.46, p-value<0.001), and *TET3* and *SMUG1* (correlation coefficient: -0.45, p-value<0.001).

Correlation analysis was also conducted for the merged list of the genes in m5C and hm5C pathways (Supplementary Figure 1 and Supplementary Table 4.) to get insights into the potential cross-talk between these two pathways. The observed three clusters on the correlation heatmap indicated a strong positive correlation among the regulators of epitranscriptome pathways (Supplementary Figure 1). Relatively stronger positive correlation between the gene pairs *UHRF1* and *WDR76* (correlation coefficient: 0.78, p-value<0.001) and *DNMT1* and *WDR76* (correlation coefficient: 0.72, p-value<0.001) was an interesting finding. Notably, *MECP2* and *WDR77* (correlation coefficient: -0.51, p-value<0.001), *MECP2* and *SMUG1* (correlation

coefficient: -0.47, p-value<0.001), and *NSUN3* and *SMUG1* (correlation coefficient: -0.46, p-value<0.001) gene pairs were revealed as negatively correlated genes in PRAD.

UHRF1, NSUN2, and *NEIL1* were Identified as Potential Prognostic Biomarkers

The prognostic value of the m5C and hm5C regulators in PRAD was evaluated using the Kaplan-Meier estimation method (Supplementary Table 5). *UHRF1, DNMT1, NSUN2, NSUN4, C1orf77, C3orf37, WDR77, NEIL1,* and *TDG* genes were recognized as statistically significant (p-value<0.05) to predict the PRAD prognosis considering the overall survival data of patients (Figure 4a and Supplementary Figure 2). Low expression levels of *UHRF1, NSUN2, C1orf77, C3orf37, WDR77,* or *TDG* genes were associated with a better disease prognosis (Figure 3a and Supplementary Figure 2). However, low expression levels of *DNMT1, NSUN4,* or *NEIL1* genes were decreased the overall survival (Figure 3a and Supplementary Figure 2).

The genes commonly identified in overall survival and LIMMA analyses (*UHRF1, NSUN2*, and *NEIL1*) were considered as candidate genes with potential prognostic value and roles in disease pathogenesis. Prediction scores of these genes were validated by plotting their ROC curves (Figure 3b). The area under the ROC curve (auROC) was calculated to estimate these genes' sensitivity and specificity to distinguish the tumor and nontumor samples (Figure 3b). The auROC of the *UHRF1, NSUN2*, and *NEIL1* was calculated as %74.1, %86.8, and %75.5, respectively.

Gleason classification was performed in order to assess the correlation between mRNA expression levels of selected genes and clinical features of prostate cancer patients (Figure 3c). The expression level of the *UHRF1* gene was significantly higher in the samples with Gleason scores 8 and 9 compared to normal samples. Increased *NSUN2 and NEIL1* expression levels were also in accordance with Gleason grading scores. The expression levels of these genes were elevated in the samples with a higher Gleason score.

The Human Protein Atlas database was used to access IHC staining images of the *UHRF1, NSUN2*, and *NEIL1* genes to further examine how the expression level of m5C and hm5C regulators change in normal and tumor tissues. The selected three genes were more expressed in PRAD tissues (Figure 3d) than the nontumor samples. Increased IHC staining in the *UHRF1, NSUN2*, and *NEIL1* genes confirmed that the expression level of these three genes was increased in PRAD.

Proliferation Related Biological Processes are Enriched in Altered m5C and hm5C Groups

To unveil the effect of m5C and hm5C regulators on disease pathogenesis, we focused on DE genes between altered and unaltered groups. The transcriptome data of patients with mutated genes in the 5mC pathway (altered group) was compared with the transcriptome data of patients without mutations in the 5mC pathway (unaltered group) using the cBioPortal database. The downloaded expression level list was filtered by setting p-value and fold-change cut-offs. Resultant genes were used to enrich the GO terms for altered and unaltered groups. The same

approach was followed for hm5C pathway genes as well. GO enrichment analysis was separately carried out on the lists of altered and unaltered m5C and hm5C pathways groups.

A total of 145 biological processes (BP), 24 cellular components (CC), and 4 molecular functions (MF) were enriched by the 281 DEGs (Supplementary Table 6) in the altered group of the 5mC pathway. In contrast, the other 1810 DEGs in the unaltered group enriched a total of 607 BPs, 80 CCs, and 79 MFs (Supplementary Table 6). The 168 DEGs in the altered group of the hm5C pathway enriched a total of 93 BPs, 18 CCs, and 2 MFs (Supplementary Table 7), whereas the other 2204 DEGs in the unaltered group enriched a total of 893 BPs, 82 CCs, and 90 MFs (Supplementary Table 7).

Processes related to proliferation such as 'chromosome segregation', 'nuclear division', 'sister chromatid segregation', and 'mitotic nuclear division' were the top BPs enriched by the DEGs in the altered group of both m5C and hm5C pathways (Figure 4a, c), while 'muscle contraction', 'extracellular matrix organization', 'extracellular structure organization', and 'muscle system process' were the top BPs enriched by the DEGs in unaltered groups of both m5C and hm5C pathways (Figure 4b, d). Molecular functions such as 'microtubule binding' and 'tubulin binding' were the common MFs enriched by the DEGs in altered group of m5C and hm5C pathways (Figure 4a, c), however 'passive transmembrane transporter activity', 'channel activity', 'ion-channel activity', and 'gated channel activity' were among the common top MFs enriched by the DEGs in unaltered group of m5C and hm5C pathways (Figure 4a, c). On the other hand, 'Chromosomal region', 'spindle', 'chromosome, centromeric region', and 'kinetochore' were identified as the enriched top CCs in the altered group of m5C and hm5C pathways (Figure 4a, c), whereas 'collagen-containing extracellular matrix', 'synaptic membrane', and 'transmembrane transporter complex' were detected as the top CCs enriched by the DEGs in the unaltered group (Figure 4b, d).

Enrichment of the 'cell cycle' and 'oocyte meiosis' KEGG pathways for both m5C and hm5C altered groups was remarkable (Supplementary Figure 3a, c). On the contrary, for the unaltered groups, 'calcium signaling pathway', 'dilated cardiomyopathy', and 'hypertrophic cardiomyopathy' pathways were the top enriched KEGG pathways (Supplementary Figure 3b, d).

To further dissect the putative roles of selected genes in PRAD pathology, we compared the gene expression profile of the PRAD patients classified according to the mutation status of the selected genes. As presented in Figure 5a, enriched terms for *UHRF1* or *NSUN2* altered groups were highly alike. 'Organelle fission', 'nuclear division', and 'chromosome segregation' were the top significant BPs, 'mitotic spindle' and 'chromosomal region' were the most significant enriched CCs, and 'microtubule binding', 'motor activity', and 'ATPase activity' were detected as the common MFs for *UHRF1* or *NSUN2* altered groups.

It is remarkable to observe that *NEIL1* altered group was diversified from the other two genes considering the enriched GO terms. For the *NEIL1* altered group, the most significant detected BPs were as follows: 'viral transcription', 'viral gene expression', and 'protein targeting to membrane' as BPs, 'ribosome', 'ribosomal subunit', and 'polysome' as CCs, and 'structural constituent of ribosome' as MF (Figure 5a).

Similar top GO terms were enriched in *UHRF1* or *NSUN2* unaltered groups, and the top-enriched GO terms were different in the *NEIL1* unaltered group (Figure 5b). BPs such as 'muscle system

process', 'muscle contraction', and 'muscle tissue development', were the top BPs enriched by the upregulated genes in the unaltered group of *UHRF1* or *NSUN2*. Those upregulated genes in *NEIL1* unaltered group enriched 'extracellular matrix organization', 'extracellular structure organization', and 'external encapsulating structure organization' terms defined top enriched BPs.

Shared molecular functions such as 'glycosaminoglycan binding', 'sulfur compound binding', and 'heparin binding' were among the top MFs enriched by the genes in *UHRF1, NSUN2*, or *NEIL1* unaltered groups (Figure 5b). 'Collagen-containing extracellular matrix', 'contractile fiber', and 'myofibril' were among the top CCs enriched by the genes in *UHRF1* or *NSUN2* unaltered groups (Figure 5b), while 'Collagen-containing extracellular matrix', 'external side of plasma membrane', and 'synaptic membrane' were among the top CCs enriched by the genes in *NEIL1* unaltered group (Figure 5b).

KEGG results showed the 'cell cycle' was the common enriched pathway for *UHRF1* and *NSUN2* altered groups. 'DNA replication', 'oocyte meiosis', and 'p53 signaling pathway' were the examples of observed pathways only in the *UHRF1* altered group. In *NSUN2* altered group, the other pathway was 'maturity onset diabetes of the young' (Supplementary Figure 4a). In unaltered groups, *UHRF1* and *NSUN2* enriched pathways were alike. 'calcium signaling pathway', 'dilated cardiomyopathy', and 'hypertrophic cardiomyopathy' pathways were enriched in both unaltered groups (Supplementary Figure 4b). On the other hand, the profile of KEGG pathways enriched by the genes in altered and unaltered groups of *NEIL1* was remarkably different from *UHRF1* and *NSUN2* profiles (Supplementary Figure 4b). The only two KEGG pathways identified in the altered group were 'ribosome' and 'coronavirus disease', and the top three unaltered pathways were 'cell adhesion molecules', 'hematopoietic cell linage', and 'malaria'.

High Number of Shared Enriched BPs in m5C and hm5C Groups is Noteworthy

Shared and unique BP terms among the altered and unaltered groups were displayed on venn diagrams (Figure 6). m5C, *UHRF1*, and *NSUN2* altered groups had 63 shared BPs, whereas 167 BPs were found as unique to one of the altered groups (Figure 6a). For unaltered m5C, *UHRF1*, and *NSUN2* groups, 344 BPs were detected as common, and 367 BPs were identified as unique BPs (Figure 6b). Furthermore, the same analysis was applied for hm5C and *NEIL1* comparisons. Intriguingly, we found no shared BP between altered hm5C and *NEIL1* groups. On the other hand, 563 and 564 BPs were classified as shared or unique BPs between unaltered hm5C and *NEIL1* groups. The high number of shared BPs among unaltered groups was notable.

Discussion

PCa is one of the widely detected cancer types in man [44]. Although the death rate of PCa is not as high as some highly aggressive cancers due to its common diagnosis, studies aim to explore the pathology and prognosis of PCa has great importance in the cancer field. Recent studies have demonstrated that genetic alterations in epitranscriptomic regulators result in many diseases, including cancer [45]. Besides m6A, mutations in m5C and hm5C regulators have been also

linked to cancer pathology. In this study, we examined diagnostic and prognostic values of hm5C and m5C regulators in PRAD.

mRNA level changes were the most frequently observed alteration type at the genetic level, whereas amplification was detected as the most common change at the DNA level. The uncovered high mutation rate and increased expression level of *NSUN2* in PRAD were consistent with previous cancer studies [46, 47]. The increased amplification and mutation levels of *NSUN2* in ovarian carcinoma, oesophageal carcinoma, bladder carcinoma (BLCA), cervical squamous cell carcinoma (CESC), lung squamous cell carcinoma (LUSC), and lung adenocarcinoma (LUAD) have been indicated in a pan-cancer study [47]. Moreover, in another study, a high *NSUN2* mRNA level in LUAD and LUSC compared to nontumor samples was found [46]. Overexpression of *NSUN2* in colorectal carcinoma (CRC) [48] and head and neck squamous cell carcinoma (HNSCC) [49] due to copy-number amplification was documented earlier.

Identification of *NSUN5* as another frequently altered regulator in PRAD is also well aligned with the literature. A pan-cancer study illustrated that *NSUN5* was overexpressed in various cancers, like glioblastoma multiforme, BLCA, PRAD, and breast invasive carcinoma (BIC) [50]. Furthermore, it has been documented that *NSUN5* promotes CRC development by enhancing cell proliferation [51], which is in line with another study underlining the confining effect of *NSUN5* deficiency on the proliferation and growth rate of mammalian cells [14]. We detected a high *NOP2*, also known as *NSUN1*, mRNA level in PRAD samples which ties well with previous findings. The amplification of *NOP2* and its overexpression is frequently found in renal cell carcinoma (RCC) [52]. Increased *NOP2* expression in diverse cancer types such as LUAD, LUSC, adrenocortical carcinoma, and ovarian serous cystadenocarcinoma was described in a reported pan-cancer study [53]. *NSUN2, NSUN5*, and *NOP2* are members of NOP2/SUN domain (NSUN) protein family, which act as DNA methyltransferase [17]. These enzymes participate in essential cellular functions such as cellular proliferation, cellular senescence, cellular migration, cellular differentiation, mRNA nuclear export, mRNA stabilization, tRNA stabilization, miRNA processing, and enhanced mRNA translation [46]. The significant roles of *NSUN2* and *NSUN5* in vital biological processes may explain the effect of dysregulation in these genes on PRAD development.

DNMT3A, a member of the DNA methyltransferase (DNMT) family, was identified as another m5C regulator with high mRNA expression level in PRAD samples. The increased mutation rate of *DNMT3A* in several cancers, including endometrial carcinoma (EC), acute myeloid leukemia (AML), LUAD, CESC, rectum adenocarcinoma, mesothelioma, was displayed in a previous pan-cancer study [54]. As a regulator of de novo methylation, *DNMT3A* functions in stem cell activity, proliferation, and differentiation [55]. Based on our analysis, the increased level of *DNMT3A* in PRAD samples provides evidence for the requirement of *DNMT3A* over-activation for cancer progression to alter the epigenetic methylation mark of cancer cells.

It was noteworthy to obtain high *TET3* and low *TET2* mRNA levels in PRAD samples. Ten-eleven translocation (TET) family members possess the catalytic activity to convert the m5C to hm5C. Therefore, TET proteins are not only considered as m5C erasers but are also accepted as writers of the hm5C pathway [14]. In many cancers, mutations in TET enzymes with different outcomes have been highlighted. Acting as a tumor-suppressor, *TET2* inhibits the BIC progression [56], and it is frequently mutated in hematopoietic malignancies [57]. On the other

hand, higher *TET3* levels are associated with a poor prognosis of ovarian cancer [58] and BIC [59], whereas its downregulation promotes GBM progression [60]. Even though our results are generally consistent with older studies, further investigation of the *TET*s' role in PRAD pathology is required. Moreover, high levels of *TET3* related proteins, *PRMT5* (a member of protein arginine methyltransferases family) and *WDR77* (androgen receptor cofactor p44), in PRAD patients are broadly in line with what has been reported in previous researches. The increased mutation rate and expression level of *PRMT5* was defined for many cancers, including uterine, ovarian, bladder, lung, glioma, and melanoma [61]. Another study revealed that inhibition of *WDR77* activity in CRC suppresses cancer cell growth [62]. The roles of *PRMT5* and *WDR77* in contribution to prostate growth, differentiation, and modulation by affecting the expression of androgen receptor-related genes [63], support the putative link between high expression of these genes with PRAD pathology. The observed high mRNA level of *SMUG1* in PRAD samples is in accordance with the previously reported findings in melanoma [64]. *SMUG1*, single-strand-selective monofunctional uracil-DNA glycosylase 1, is an eraser of the hm5C pathway [65]. As a base excision repair enzyme to catalyze the removal of the uracil and oxidized bases, *SMUG1* plays multiple roles to alter the RNA modification status and maintain genome stability [65]. Hence, its general cancer-suppressive roles contradict the high expression profile in PRAD samples which could be related to additional uncovered molecular functions of *SMUG1* that should be interrogated in future studies. Taken together, high genetic alteration rates of m5C and hm5C regulators in PRAD detected in this study are broadly consistent with previous findings.

The list of identified DEGs for m5C and hm5C pathways was noteworthy. Some of the readers, writers, or erasers exhibited different expression trends despite having similar activities. The observed opposite trend of the regulators with similar activity might be due to the expression under specific conditions or acting on different targets. Also, the expression trend of m5C and hm5C regulators may differ among various cancer types. For instance, *DNMT3A* is found as overexpressed in gastric adenocarcinoma [66], pancreatic adenocarcinoma (PDAC) [67], AML [68], gastroenteropancreatic cancer [69], HCC [70]. On the contrary, the loss of *DNMT3A* is responsible for a preleukemic phenotype on murine hematopoietic stem cells [71]. Downregulation of *DNMT3A* leads to tumor growth in BIC and CRC cell lines [72], and the lack of *DNMT3A* was related to tumor progression in LUAD [73]. Therefore, complementary to pathway scale examination, each gene in the RNA modification pathways should be evaluated individually.

The performed analysis of m5C and hm5C regulators based on gene expression levels in PRAD patients pinpointed a strong positive correlation among TET enzymes, which was also reported in earlier studies conducted in gastric cancer [74] and HNSCC [75]. Findings of positive correlation among DNMT enzymes, and between *TET3* and *DNMT3A* genes support the previous results in HNSCC [75] and non-small cell lung carcinoma [76]. Likewise, the detected negative correlation between *NSUN5* and *TET2*, and *NSUN5* and *TRDMT1* in this study were well aligned with the obtained data in pan-cancer [54] and RCC [77] studies. These results indicated that interactions among m5C and hm5C regulators might play crucial roles in PRAD. The significance of these correlations among the m5C and hm5C regulators is an open question.

Interestingly, enriched top BPs in the altered m5C and hm5C pathways groups were highly similar and related to cell division, such as chromosome segregation, nuclear division, and mitotic nuclear division. Affected proliferation-related pathways in patients with genetic alteration in m5C and hm5C imply the potential roles of these pathways in PRAD progression. Previous in-silico studies suggested that the cancer patients with a mutation in mRNA modification regulators upregulate the cell cycle-related pathways in gastrointestinal cancer [78] and BIC [42]. Experimental evidence demonstrated that an increased level of the m5C regulator *NSUN2* enhances the proliferation of gastric cancer cells [79]. Moreover, *PRMT5* plays an essential role in epithelial-mesenchymal transition in pancreatic cancer by regulating the EGFR and B-catenin pathways [80]. Knock-down of *PRMT5* in the pancreatic cancer cell line inhibits the growth of the cancer cells [25]. Since the hm5C pathway has not been extensively studied, we were not able to compare GO results for PRAD results with other cancer types due to the lack of reports on the enrichment of gene ontology terms considering the mutation status in hm5C regulators in different cancers.

Relation of the expression level of m5C and hm5C regulators and OS of PRAD patients was assessed to explore the prognostic value of these genes. Among the examined genes, *UHRF1, DNMT1, NSUN2, NSUN4, C1orf77, C3orf37, WDR77, NEIL1*, and *TDG* genes were detected as putative prognostic biomarkers. The genes common in DEG and OS lists (*UHRF1, NSUN2*, and *NEIL1*) were considered as strong candidates, which is confirmed by ROC curve analysis. The association of expression levels of these genes with overall survival as prognostic markers has also been demonstrated previously in multiple cancers. For instance, high levels of *UHRF1* resulted in lower overall survival in melanoma [81], HCC [82], and LUAD [82]. *NSUN2* high expression was also related to poor survival rates in hHNSCC [83], gastric cancer [79]. High *NEIL1* expression is correlated with low overall survival in gastric cancer [84]. Consistency of our findings with earlier reports prompted us to check whether *UHRF1, NSUN2*, and *NEIL1* expression levels of these genes are compatible with Gleason classifications. According to the obtained scores, we can conclude that the elevated expression levels of these genes are associated with clinically poor prognosis except the grade 10. The low sample number of Grade 10 is a limiting factor for a statistical evaluation of the relationship between expression level and Gleason score. High expression in PRAD was further validated by the comparison of protein expression profiles of *UHRF1, NSUN2*, and *NEIL1* in tumor and nontumor tissues using the Human Protein database. Taken together, we believe that *UHRF1, NSUN2*, and *NEIL1* can be considered as promising diagnostic and prognostic biomarkers of PRAD.

Therefore, we performed GO and KEGG analyses for altered and unaltered groups of *UHRF1, NSUN2*, and *NEIL1* genes to provide new clues on the impact of RNA modification pathways in PRAD development. BPs enriched by the upregulated genes in *UHRF1* and *NSUN2* mutated patients were greatly alike and considerably resembled the BPs obtained by the altered group of m5C. Most of the common altered BPs among m5C, *UHRF1*, and *NSUN2*, such as chromosome segregation, nuclear division, meiotic cell cycle processes, were related to cell cycle and proliferation. *UHRF1* has roles in *DNMT1* degradation by functioning as an E3-ubiquitin-ligase [85]. Therefore, for many cancer types, global DNA hypomethylation occurs through the upregulation of *UHRF1* [86]. It has been reported that *UHRF1* participates in the development and progression of many cancers, such as gastric cancer [87], non-small lung carcinoma [88], and EC [89]. Moreover, its high expression is associated with a poor prognosis in bladder cancer [90],

pancreatic cancer [89], and HCC [87]. Likewise, the pivotal roles of *UHRF1* in PRAD development have been experimentally corroborated [88, 91]. A similar association with cancer progression has been also described for *NSUN2* overexpression. High expression of *NSUN2* was linked to the development of several cancers such as esophageal squamous cell carcinoma (ESCC) [92], gastric cancer [79], and hypopharyngeal squamous cell carcinoma and [93]. Since our results indicated the high mRNA expression level as the most commonly detected alteration of *UHRF1* and *NSUN2* genes in PRAD patients, the obtained cell cycle-related BPs for the altered group of *UHRF1* and *NSUN2* strongly favors the findings of previously conducted studies. On the other hand, surprisingly, no common BPs were found between the *NEIL1* and hm5C altered groups. Diverse BPs enriched in altered *NEIL1* gene, and hm5C regulators groups imply the limited contribution of *NEIL1* to enrich the cell cycle-related pathways. Considering the *NEIL1* roles in the removal of carboxyl- and formylcytosine residues to alter the epigenetic profile [94], further functional studies should be addressed to describe its association with PRAD pathology.

In conclusion, our study comprehensively analyzes the m5C and hm5C regulators in PRAD development and progression. Our findings revealed that regulators in m5C and hm5C pathways are frequently mutated, and a high mRNA expression level was detected as a general mark in tumor samples. *UHRF1, NSUN2*, and *NEIL1* genes appeared as candidate biomarkers to use in PRAD. However, to further support the in-silico results of the present study, future in-vitro and in-vivo researches are required to elucidate the relationship between PRAD development and RNA modification mechanisms.

References

[1]   Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.10.3322/caac.21492

[2]   Merriel SWD, Funston G, Hamilton W. Prostate Cancer in Primary Care. Adv Ther 2018;35:1285-94.10.1007/s12325-018-0766-1

[3]   Rawla P. Epidemiology of Prostate Cancer. World J Oncol 2019;10:63-89.10.14740/wjon1191

[4]   Selley S, Donovan J, Faulkner A, Coast J, Gillatt D. Diagnosis, management and screening of early localised prostate cancer. Health Technol Assess 1997;1:i, 1-96.

[5]   Gleason DF. Classification of prostatic carcinomas. Cancer Chemother Rep 1966;50:125-8.

[6]   Sehn JK. Prostate Cancer Pathology: Recent Updates and Controversies. Mo Med 2018;115:151-5.

[7]   Janiczek M, Szylberg L, Kasperska A, Kowalewski A, Parol M, Antosik P, Radecka B, Marszalek A. Immunotherapy as a Promising Treatment for Prostate Cancer: A Systematic Review. J Immunol Res 2017;2017.Artn 4861570.10.1155/2017/4861570

[8]   Huggins C, Hodges CV. Studies on prostatic cancer. I. The effect of castration, of estrogen and androgen injection on serum phosphatases in metastatic carcinoma of the prostate. CA Cancer J Clin 1972;22:232-40.10.3322/canjclin.22.4.232

[9]   Spetsieris N, Boukovala M, Patsakis G, Alafis I, Efstathiou E. Neuroendocrine and Aggressive-Variant Prostate Cancer. Cancers (Basel) 2020;12.10.3390/cancers12123792

[10]   Chandrasekar T, Yang JC, Gao AC, Evans CP. Mechanisms of resistance in castration-resistant prostate cancer (CRPC). Transl Androl Urol 2015;4:365-80.10.3978/j.issn.2223-4683.2015.05.02

[11]   Helm M, Motorin Y. Detecting RNA modifications in the epitranscriptome: predict and validate. Nat Rev Genet 2017;18:275-91.10.1038/nrg.2016.169

[12]   Boccaletto P, Machnicka MA, Purta E, Piatkowski P, Baginski B, Wirecki TK, de Crecy-Lagard V, Ross R, Limbach PA, Kotter A, Helm M, Bujnicki JM. MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 2018;46:D303-D7.10.1093/nar/gkx1030

[13]   Shi HH, Chai PW, Jia RB, Fan XQ. Novel insight into the regulatory roles of diverse RNA modifications: Re-defining the bridge between transcription and translation. Mol Cancer 2020;19.ARTN 78.10.1186/s12943-020-01194-6

[14]   Chen YS, Yang WL, Zhao YL, Yang YG. Dynamic transcriptomic m(5) C and its regulatory role in RNA processing. Wiley Interdiscip Rev RNA 2021;12:e1639.10.1002/wrna.1639

[15]   Wyatt GR. Occurrence of 5-methylcytosine in nucleic acids. Nature 1950;166:237-8.10.1038/166237b0

[16]   Squires JE, Patel HR, Nousch M, Sibbritt T, Humphreys DT, Parker BJ, Suter CM, Preiss T. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 2012;40:5023-33.10.1093/nar/gks144

[17]   Bohnsack KE, Hobartner C, Bohnsack MT. Eukaryotic 5-methylcytosine (m(5)C) RNA Methyltransferases: Mechanisms, Cellular Functions, and Links to Disease. Genes (Basel) 2019;10.10.3390/genes10020102

[18]   Song P, Tayier S, Cai Z, Jia G. RNA methylation in mammalian development and cancer. Cell Biol Toxicol 2021.10.1007/s10565-021-09627-8

[19]   Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang X, Golic KG, Jacobsen SE, Bestor TH. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. Science 2006;311:395-8.10.1126/science.1120976

[20]   Zhang Q, Liu F, Chen W, Miao H, Liang H, Liao Z, Zhang Z, Zhang B. The role of RNA m(5)C modification in cancer metastasis. Int J Biol Sci 2021;17:3369-80.10.7150/ijbs.61439

[21]   Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, Li A, Wang X, Bhattarai DP, Xiao W, Sun HY, Zhu Q, Ma HL, Adhikari S, Sun M, Hao YJ, Zhang B, Huang CM, Huang N, Jiang GB, Zhao YL, Wang HL, Sun YP, Yang YG. 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. Cell Res 2017;27:606-25.10.1038/cr.2017.55

[22]   Liutkeviciute Z, Lukinavicius G, Masevicius V, Daujotyte D, Klimasauskas S. Cytosine-5-methyltransferases add aldehydes to DNA. Nat Chem Biol 2009;5:400-2.10.1038/nchembio.172

[23]   Fu L, Guerrero CR, Zhong N, Amato NJ, Liu Y, Liu S, Cai Q, Ji D, Jin SG, Niedernhofer LJ, Pfeifer GP, Xu GL, Wang Y. Tet-mediated formation of 5-hydroxymethylcytosine in RNA. J Am Chem Soc 2014;136:11582-5.10.1021/ja505305z

[24]   Lewis CJ, Pan T, Kalsotra A. RNA modifications and structures cooperate to guide RNA-protein interactions. Nat Rev Mol Cell Biol 2017;18:202-10.10.1038/nrm.2016.163

[25]   Barbieri I, Kouzarides T. Role of RNA modifications in cancer. Nat Rev Cancer 2020;20:303-22.10.1038/s41568-020-0253-2

[26]   Chen X, Li A, Sun BF, Yang Y, Han YN, Yuan X, Chen RX, Wei WS, Liu Y, Gao CC, Chen YS, Zhang M, Ma XD, Liu ZW, Luo JH, Lyu C, Wang HL, Ma J, Zhao YL, Zhou FJ, Huang Y, Xie D, Yang YG. 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. Nat Cell Biol 2019;21:978-90.10.1038/s41556-019-0361-y

[27]   He Y, Yu X, Li J, Zhang Q, Zheng Q, Guo W. Role of m(5)C-related regulatory genes in the diagnosis and prognosis of hepatocellular carcinoma. Am J Transl Res 2020;12:912-22.

[28]   Cheray M, Etcheverry A, Jacques C, Pacaud R, Bougras-Cartron G, Aubry M, Denoual F, Peterlongo P, Nadaradjane A, Briand J, Akcha F, Heymann D, Vallette FM, Mosser J, Ory B, Cartron PF. Cytosine

methylation of mature microRNAs inhibits their functions and is associated with poor prognosis in glioblastoma multiforme. Mol Cancer 2020;19:36.10.1186/s12943-020-01155-z

[29]   Wang L, Gao S. Identification of 5-methylcytosine-related signature for predicting prognosis in ovarian cancer. Biol Res 2021;54:18.10.1186/s40659-021-00340-8

[30]   Lian CG, Xu Y, Ceol C, Wu F, Larson A, Dresser K, Xu W, Tan L, Hu Y, Zhan Q, Lee CW, Hu D, Lian BQ, Kleffel S, Yang Y, Neiswender J, Khorasani AJ, Fang R, Lezcano C, Duncan LM, Scolyer RA, Thompson JF, Kakavand H, Houvras Y, Zon LI, Mihm MC, Jr., Kaiser UB, Schatton T, Woda BA, Murphy GF, Shi YG. Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. Cell 2012;150:1135-46.10.1016/j.cell.2012.07.033

[31]   Oishi N, Vuong HG, Mochizuki K, Kondo T. Loss of 5-Hydroxymethylcytosine is an Epigenetic Hallmark of Thyroid Carcinomas with TERT Promoter Mutations. Endocr Pathol 2020;31:359-66.10.1007/s12022-020-09652-z

[32]   Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013;45:1113-20.10.1038/ng.2764

[33]   Kosinski M BP. The Cancer Genome Atlas Data Integration. 2016.

[34]   Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2:401-4.10.1158/2159-8290.CD-12-0095

[35]   Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016;32:2847-9.10.1093/bioinformatics/btw313

[36]   Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.10.1093/nar/gkv007

[37]   Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77.10.1186/1471-2105-12-77

[38]   Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, Varambally S. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. Neoplasia 2017;19:649-58.10.1016/j.neo.2017.05.002

[39]   Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS: A Journal of Integrative Biology 2012;16:284-7.10.1089/omi.2011.0118

[40]   Sibai M, Parlayan C, Tuglu P, Ozturk G, Demircan T. Integrative Analysis of Axolotl Gene Expression Data from Regenerative and Wound Healing Limb Tissues. Sci Rep 2019;9:20280.10.1038/s41598-019-56829-6

[41]   Bardou P. jvenn: an interactive Venn diagram viewer. BMC Bioinformatics 2014;293.10.1186/1471-2105-15-293

[42]   Demircan T, Yavuz M, Akgül S. m 6 A Pathway Regulators Are Frequently Mutated in Breast Invasive Carcinoma and May Play an Important Role in Disease Pathogenesis. OMICS: A Journal of Integrative Biology 2021;5(10):660-678.2021.10.1101/2021.06.07.447359

[43]   Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjostedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Ponten F. Proteomics. Tissue-based map of the human proteome. Science 2015;347:1260419.10.1126/science.1260419

[44]   Kucera R, Pecen L, Topolcan O, Dahal AR, Costigliola V, Giordano FA, Golubnitschaja O. Prostate cancer management: long-term beliefs, epidemic developments in the early twenty-first century and 3PM dimensional solutions. EPMA J 2020;11:399-418.10.1007/s13167-020-00214-1

[45]   Kanwal R, Gupta K, Gupta S. Cancer epigenetics: an introduction. Methods Mol Biol 2015;1238:3-25.10.1007/978-1-4939-1804-1_1

[46]   Chellamuthu A, Gray SG. The RNA Methyltransferase NSUN2 and Its Potential Roles in Cancer. Cells 2020;9.10.3390/cells9081758

[47]   Manning M, Jiang Y, Wang R, Liu L, Rode S, Bonahoom M, Kim S, Yang ZQ. Pan-cancer analysis of RNA methyltransferases identifies FTSJ3 as a potential regulator of breast cancer progression. RNA Biol 2020;17:474-86.10.1080/15476286.2019.1708549

[48]   Okamoto M, Hirata S, Sato S, Koga S, Fujii M, Qi G, Ogawa I, Takata T, Shimamoto F, Tatsuka M. Frequent increased gene copy number and high protein expression of tRNA (cytosine-5-)-methyltransferase (NSUN2) in human cancers. DNA Cell Biol 2012;31:660-71.10.1089/dna.2011.1446

[49]   Xue M. Gene signatures of m5C regulators may predict prognoses of patients with head and neck squamous cell carcinoma. Am J Transl Res 2020;12:6841–52.

[50]   Zhang X-w, Huang Y, Wu L-y, Qi Q, Zhong R, Li S-s, Zhao J-m, Liu H-r, Wu H-g. NSUN5 is Upregulated and Positively Correlated with Translation in Human Cancers: A Bioinformatics-based Study. 2020.10.21203/rs.3.rs-92585/v1

[51]   Jiang Z, Li S, Han MJ, Hu GM, Cheng P. High expression of NSUN5 promotes cell proliferation via cell cycle regulation in colorectal cancer. Am J Transl Res 2020;12:3858-70.

[52]   Wang G, Qu F, Liu S, Zhou J, Wang Y. Nucleolar protein NOP2 could serve as a potential prognostic predictor for clear cell renal cell carcinoma. Bioengineered 2021;12:4841-55.10.1080/21655979.2021.1960130

[53]   Du E, Li J, Sheng F, Li S, Zhu J, Xu Y, Zhang Z. A pan-cancer analysis reveals genetic alterations, molecular mechanisms, and clinical relevance of m(5) C regulators. Clin Transl Med 2020;10:e180.10.1002/ctm2.180

[54]   He Y, Yu X, Zhang M, Guo W. Pan-cancer analysis of m(5)C regulator genes reveals consistent epigenetic landscape changes in multiple cancers. World J Surg Oncol 2021;19:224.10.1186/s12957-021-02342-y

[55]   Zhang W, Xu J. DNA methyltransferases and their roles in tumorigenesis. Biomark Res 2017;5:1.10.1186/s40364-017-0081-z

[56]   Zhu X, Li S. TET2 inhibits tumorigenesis of breast cancer cells by regulating caspase-4. Sci Rep 2018;8:16167.10.1038/s41598-018-34462-z

[57]   Xu YP, Lv L, Liu Y, Smith MD, Li WC, Tan XM, Cheng M, Li Z, Bovino M, Aube J, Xiong Y. Tumor suppressor TET2 promotes cancer immunity and immunotherapy efficacy. J Clin Invest 2019;129:4316-31.10.1172/JCI129317

[58]   Cao T, Pan W, Sun X, Shen H. Increased expression of TET3 predicts unfavorable prognosis in patients with ovarian cancer-a bioinformatics integrative analysis. J Ovarian Res 2019;12:101.10.1186/s13048-019-0575-4

[59]   Wu MZ, Chen SF, Nieh S, Benner C, Ger LP, Jan CI, Ma L, Chen CH, Hishida T, Chang HT, Lin YS, Montserrat N, Gascon P, Sancho-Martinez I, Izpisua Belmonte JC. Hypoxia Drives Breast Tumor Malignancy through a TET-TNFalpha-p38-MAPK Signaling Axis. Cancer Res 2015;75:3912-24.10.1158/0008-5472.CAN-14-3208

[60]   Carella A, Tejedor JR, Garcia MG, Urdinguio RG, Bayon GF, Sierra M, Lopez V, Garcia-Torano E, Santamarina-Ojeda P, Perez RF, Bigot T, Mangas C, Corte-Torres MD, Saenz-de-Santa-Maria I, Mollejo M, Melendez B, Astudillo A, Chiara MD, Fernandez AF, Fraga MF. Epigenetic downregulation of TET3 reduces

genome-wide 5hmC levels and promotes glioblastoma tumorigenesis. Int J Cancer 2020;146:373-87.10.1002/ijc.32520

[61]   Stopa N, Krebs JE, Shechter D. The PRMT5 arginine methyltransferase: many roles in development, cancer and beyond. Cell Mol Life Sci 2015;72:2041-59.10.1007/s00018-015-1847-9

[62]   Qi H, Shi X, Yu M, Liu B, Liu M, Song S, Chen S, Zou J, Zhu WG, Luo J. Sirtuin 7-mediated deacetylation of WD repeat domain 77 (WDR77) suppresses cancer cell growth by reducing WDR77/PRMT5 transmethylase complex activity. J Biol Chem 2018;293:17769-79.10.1074/jbc.RA118.003629

[63]   Gao S, Wu H, Wang F, Wang Z. Altered differentiation and proliferation of prostate epithelium in mice lacking the androgen receptor cofactor p44/WDR77. Endocrinology 2010;151:3941-53.10.1210/en.2009-1080

[64]   Malvi P, Wang B, Shah S, Gupta R. Dissecting the role of RNA modification regulatory proteins in melanoma. Oncotarget 2019;10:3745-59.10.18632/oncotarget.26959

[65]   Raja S, Van Houten B. The Multiple Cellular Roles of SMUG1 in Genome Maintenance and Cancer. Int J Mol Sci 2021;22.10.3390/ijms22041981

[66]   Kataoka I, Funata S, Nagahama K, Isogaya K, Takeuchi H, Abe N, Shibahara J. DNMT3A overexpression is associated with aggressive behavior and enteroblastic differentiation of gastric adenocarcinoma. Ann Diagn Pathol 2020;44:151456.10.1016/j.anndiagpath.2019.151456

[67]   Jing W, Song N, Liu YP, Qu XJ, Qi YF, Li C, Hou KZ, Che XF, Yang XH. DNMT3a promotes proliferation by activating the STAT3 signaling pathway and depressing apoptosis in pancreatic cancer. Cancer Manag Res 2019;11:6379-96.10.2147/CMAR.S201610

[68]   Yang Y, Dai Y, Yang X, Wu S, Wang Y. DNMT3A Mutation-Induced CDK1 Overexpression Promotes Leukemogenesis by Modulating the Interaction between EZH2 and DNMT3A. Biomolecules 2021;11.10.3390/biom11060781

[69]   Rahman MM, Qian ZR, Wang EL, Yoshimoto K, Nakasono M, Sultana R, Yoshida T, Hayashi T, Haba R, Ishida M, Okabe H, Sano T. DNA methyltransferases 1, 3a, and 3b overexpression and clinical significance in gastroenteropancreatic neuroendocrine tumors. Hum Pathol 2010;41:1069-78.10.1016/j.humpath.2010.01.011

[70]   Oh BK. DNA methyltransferase expression and DNA methylation in human hepatocellular carcinoma and their clinicopathological correlation. Int J Mol Med 2007;20(1):65-73.10.3892/ijmm.20.1.65

[71]   Mayle A, Yang L, Rodriguez B, Zhou T, Chang E, Curry CV, Challen GA, Li W, Wheeler D, Rebel VI, Goodell MA. Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. Blood 2015;125:629-38.10.1182/blood-2014-08-594648

[72]   Jia Y, Li P, Fang L, Zhu H, Xu L, Cheng H, Zhang J, Li F, Feng Y, Li Y, Li J, Wang R, Du JX, Li J, Chen T, Ji H, Han J, Yu W, Wu Q, Wong J. Negative regulation of DNMT3A de novo DNA methylation by frequently overexpressed UHRF family proteins as a mechanism for widespread DNA hypomethylation in cancer. Cell Discov 2016;2:16007.10.1038/celldisc.2016.7

[73]   Husni RE, Shiba-Ishii A, Iiyama S, Shiozawa T, Kim Y, Nakagawa T, Sato T, Kano J, Minami Y, Noguchi M. DNMT3a expression pattern and its prognostic value in lung adenocarcinoma. Lung Cancer 2016;97:59-65.10.1016/j.lungcan.2016.04.018

[74]   Zhang J, Piao HY, Wang Y, Meng XY, Yang D, Zhao Y, Zheng ZC. To Develop and Validate the Combination of RNA Methylation Regulators for the Prognosis of Patients with Gastric Cancer. Onco Targets Ther 2020;13:10785-95.10.2147/OTT.S276239

[75]   Han Z, Yang B, Wang Y, Zeng X, Tian Z. Identification of Expression Patterns and Potential Prognostic Significance of m(5)C-Related Regulators in Head and Neck Squamous Cell Carcinoma. Front Oncol 2021;11:592107.10.3389/fonc.2021.592107

[76]   Daskalos A, Oleksiewicz U, Filia A, Nikolaidis G, Xinarianos G, Gosney JR, Malliri A, Field JK, Liloglou T. UHRF1-mediated tumor suppressor gene inactivation in nonsmall cell lung cancer. Cancer 2011;117:1027-37.10.1002/cncr.25531

[77]   Wu J, Hou C, Wang Y, Wang Z, Li P, Wang Z. Comprehensive Analysis of m(5)C RNA Methylation Regulator Genes in Clear Cell Renal Cell Carcinoma. Int J Genomics 2021;2021:3803724.10.1155/2021/3803724

[78]   Xiang S, Ma Y, Shen J, Zhao Y, Wu X, Li M, Yang X, Kaboli PJ, Du F, Ji H, Zheng Y, Li X, Li J, Wen Q, Xiao Z. m(5)C RNA Methylation Primarily Affects the ErbB and PI3K-Akt Signaling Pathways in Gastrointestinal Cancer. Front Mol Biosci 2020;7:599340.10.3389/fmolb.2020.599340

[79]   Hu Y, Chen C, Tong X, Chen S, Hu X, Pan B, Sun X, Chen Z, Shi X, Hu Y, Shen X, Xue X, Lu M. NSUN2 modified by SUMO-2/3 promotes gastric cancer progression and regulates mRNA m5C methylation. Cell Death Dis 2021;12:842.10.1038/s41419-021-04127-3

[80]   Ge L, Wang H, Xu X, Zhou Z, He J, Peng W, Du F, Zhang Y, Gong A, Xu M. PRMT5 promotes epithelial-mesenchymal transition via EGFR-beta-catenin axis in pancreatic cancer cells. J Cell Mol Med 2020;24:1969-79.10.1111/jcmm.14894

[81]   Wei C, Lu N, Wang L, Zhang Y, Feng Z, Yang Y, Qi F, Gu J. Upregulation of UHRF1 promotes the progression of melanoma by inducing cell proliferation. Oncol Rep 2018;39:2553-62.10.3892/or.2018.6356

[82]   Liang D, Xue H, Yu Y, Lv F, You W, Zhang B. Elevated expression of UHRF1 predicts unfavorable prognosis for patients with hepatocellular carcinoma. Int J Clin Exp Pathol 2015;8:9416-21.

[83]   Lu L, Zhu G, Zeng H, Xu Q, Holzmann K. High tRNA Transferase NSUN2 Gene Expression is Associated with Poor Prognosis in Head and Neck Squamous Carcinoma. Cancer Invest 2018;36:246-53.10.1080/07357907.2018.1466896

[84]   Ni F, Tang H, Wang C, Wang Z, Yu F, Chen B, Sun L. Berzosertib (VE-822) inhibits gastric cancer cell proliferation via base excision repair system. Cancer Manag Res 2019;11:8391-405.10.2147/CMAR.S217375

[85]   Hong Q, Shao ZM. Ubiquitination/deubiquitination and acetylation/deacetylation: making DNMT1 stability more coordinated. Acta Pharmacol Sin 2011;32:139-40.10.1038/aps.2011.3

[86]   Bronner C, Krifa M, Mousli M. Increasing role of UHRF1 in the reading and inheritance of the epigenetic code as well as in tumorogenesis. Biochem Pharmacol 2013;86:1643-9.10.1016/j.bcp.2013.10.002

[87]   Zhuo H, Tang J, Lin Z, Jiang R, Zhang X, Ji J, Wang P, Sun B. The aberrant expression of MEG3 regulated by UHRF1 predicts the prognosis of hepatocellular carcinoma. Mol Carcinog 2016;55:209-19.10.1002/mc.22270

[88]   Patnaik D, Esteve PO, Pradhan S. Targeting the SET and RING-associated (SRA) domain of ubiquitin-like, PHD and ring finger-containing 1 (UHRF1) for anti-cancer drug development. Oncotarget 2018;9:26243-58.10.18632/oncotarget.25425

[89]   Cui L, Chen J, Zhang Q, Wang X, Qu J, Zhang J, Dang S. Up-regulation of UHRF1 by oncogenic Ras promoted the growth, migration, and metastasis of pancreatic cancer cells. Mol Cell Biochem 2015;400:223-32.10.1007/s11010-014-2279-9

[90]   Unoki M, Kelly JD, Neal DE, Ponder BA, Nakamura Y, Hamamoto R. UHRF1 is a novel molecular marker for diagnosis and the prognosis of bladder cancer. Br J Cancer 2009;101:98-105.10.1038/sj.bjc.6605123

[91]   Babbio F, Pistore C, Curti L, Castiglioni I, Kunderfranco P, Brino L, Oudet P, Seiler R, Thalman GN, Roggero E, Sarti M, Pinton S, Mello-Grand M, Chiorino G, Catapano CV, Carbone GM, Bonapace IM. The SRA protein UHRF1 promotes epigenetic crosstalks and is involved in prostate cancer progression. Oncogene 2012;31:4878-87.10.1038/onc.2011.641

[92]   Su J, Wu G, Ye Y, Zhang J, Zeng L, Huang X, Zheng Y, Bai R, Zhuang L, Li M, Pan L, Deng J, Li R, Deng S, Zhang S, Zuo Z, Liu Z, Lin J, Lin D, Zheng J. NSUN2-mediated RNA 5-methylcytosine promotes esophageal squamous cell carcinoma progression via LIN28B-dependent GRB2 mRNA stabilization. Oncogene 2021;40:5814-28.10.1038/s41388-021-01978-0

[93]   Chen L, Ding J, Wang B, Chen X, Ying X, Yu Z, Dong P. RNA methyltransferase NSUN2 promotes hypopharyngeal squamous cell carcinoma proliferation and migration by enhancing TEAD1 expression in an m(5)C-dependent manner. Exp Cell Res 2021;404:112664.10.1016/j.yexcr.2021.112664

[94]   Slyvka A, Mierzejewska K, Bochtler M. Nei-like 1 (NEIL1) excises 5-carboxylcytosine directly and stimulates TDG-mediated 5-formyl and 5-carboxylcytosine excision. Sci Rep 2017;7:9001.10.1038/s41598-017-07458-4

Acknowledgments

Author Contributions

Buse Nur Kahyaoğulları: Methodology, Formal Analysis, Investigation, Writing-Original Draft, Visualization**. Turan Demircan:** Conceptualization, Methodology, Formal Analysis, Investigation, Writing-Original Draft, Writing- Reviewing&Editing.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Funding
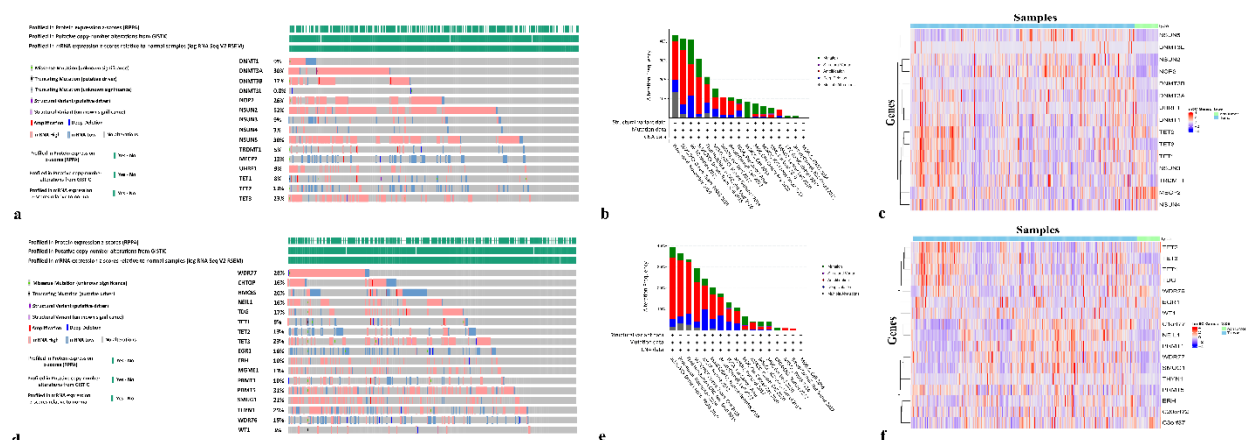
Figure 1.

Figure 1. Genomic alterations of m5C and hm5C regulators in TCGA-PRAD cohort (494 samples). (a, d) Oncoprints of m5C and hm5C regulators for 15 and 17 genes in PRAD, respectively. (b, e) Cancer type summary for m5C and hm5C regulators in seventeen studies. (c, e) Expression levels of regulators in tumor and non-tumor samples of PRAD patients.
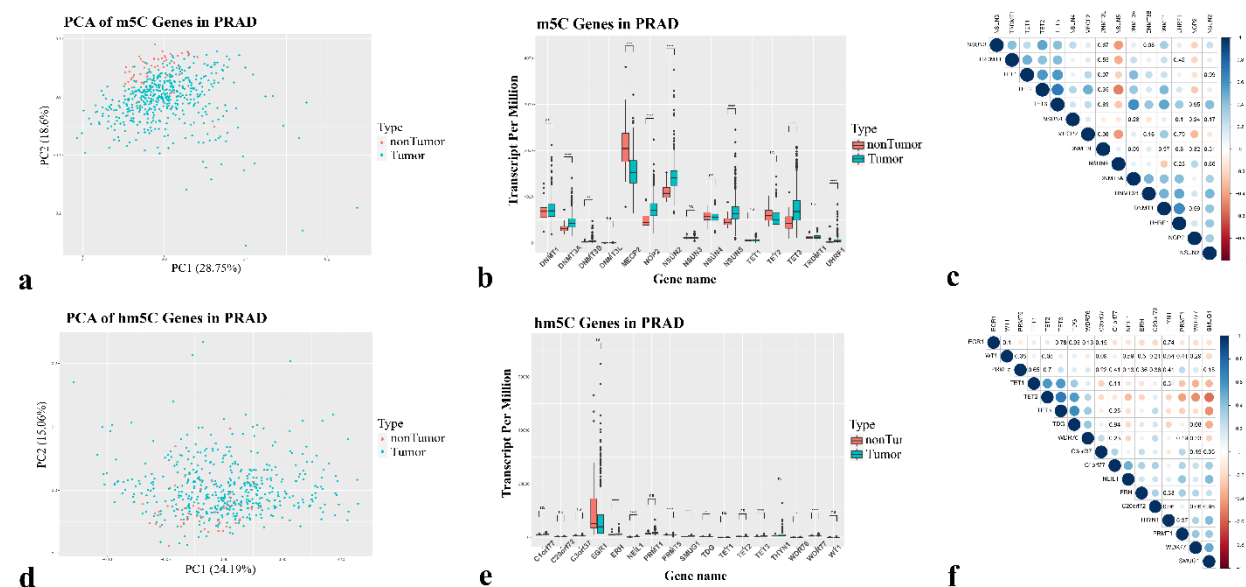
Figure 2.

Figure 2. PCA, boxplots, and correlation maps for m5C-hm5C regulators. (a, d) PCA plots for m5C and hm5C genes, respectively. (b, e) Boxplots exhibit the expression levels of m5C and hm5C regulators between tumor and non-tumor samples. (c, f) Correlation of m5C and hm5C genes. Red and blue colors symbolize high and low correlation, respectively.

Figure 3.
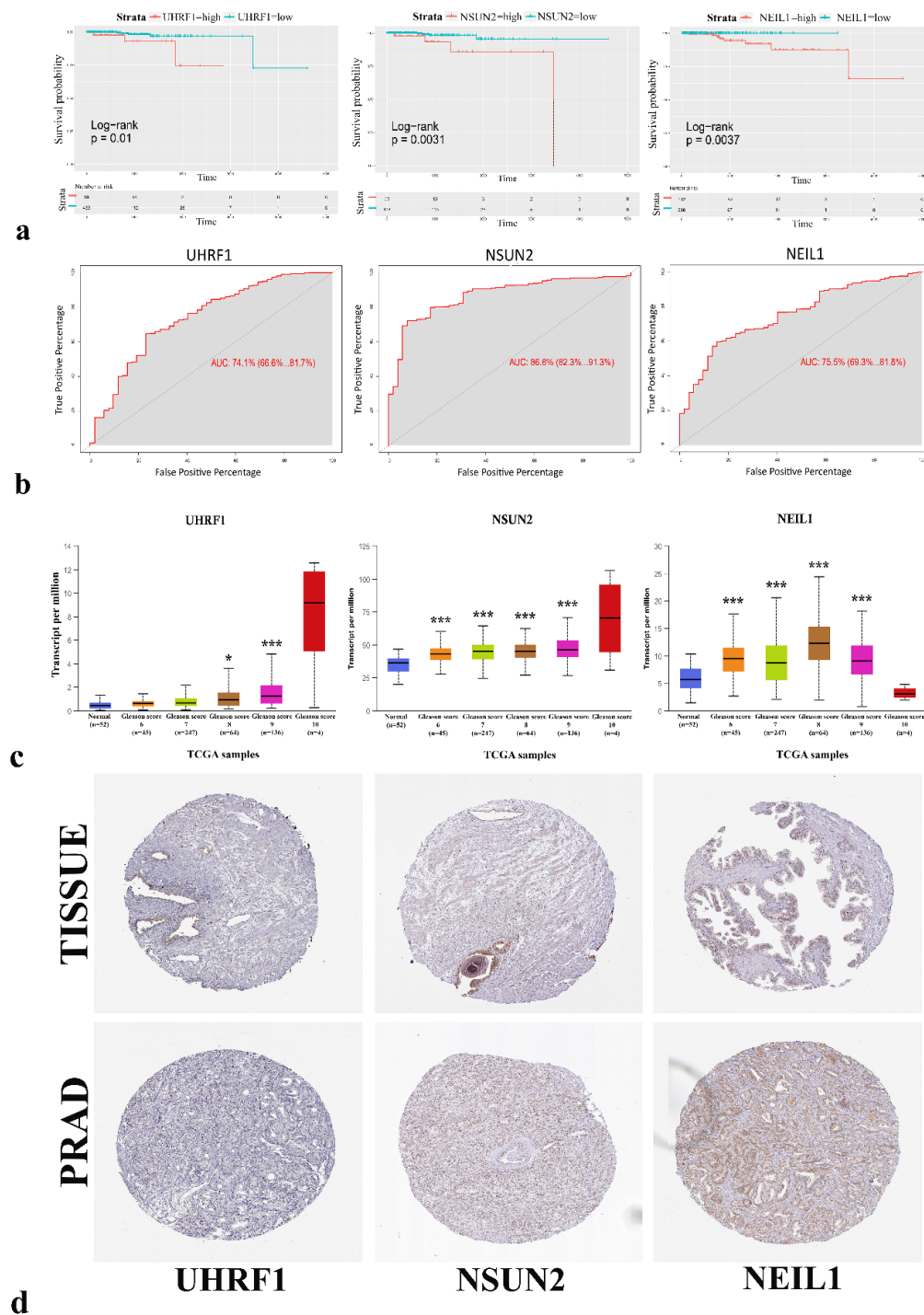
Figure 3. Represented pathways for altered and unaltered groups of m5C and hm5C genes. (a, c) Gene ontology (GO) analyses of m5C and hm5C altered groups, respectively. (b, d) Related pathways for m5C and hm5C unaltered groups on GO plots, respectively. BP; biological process, CC; cellular component, MF; molecular function.
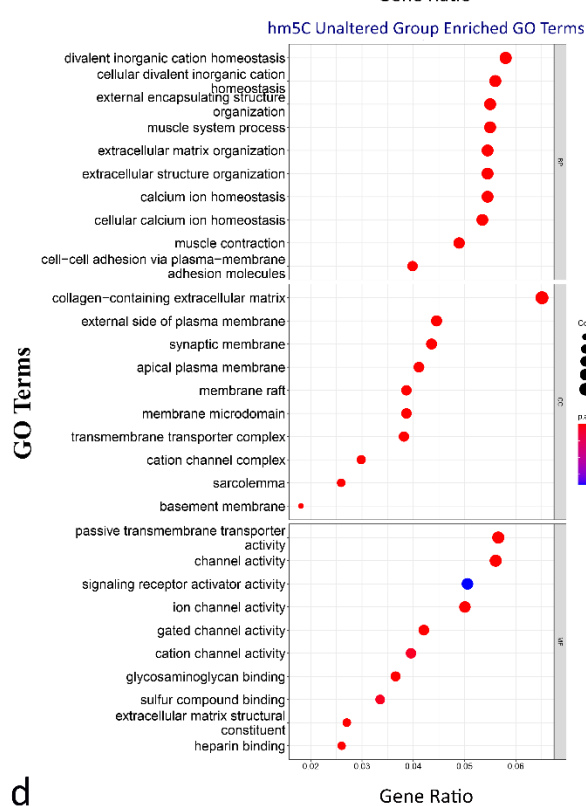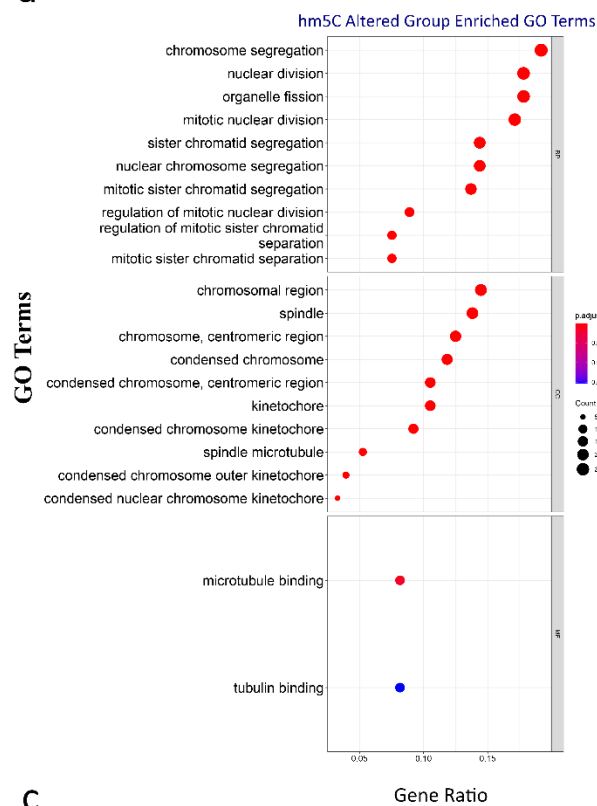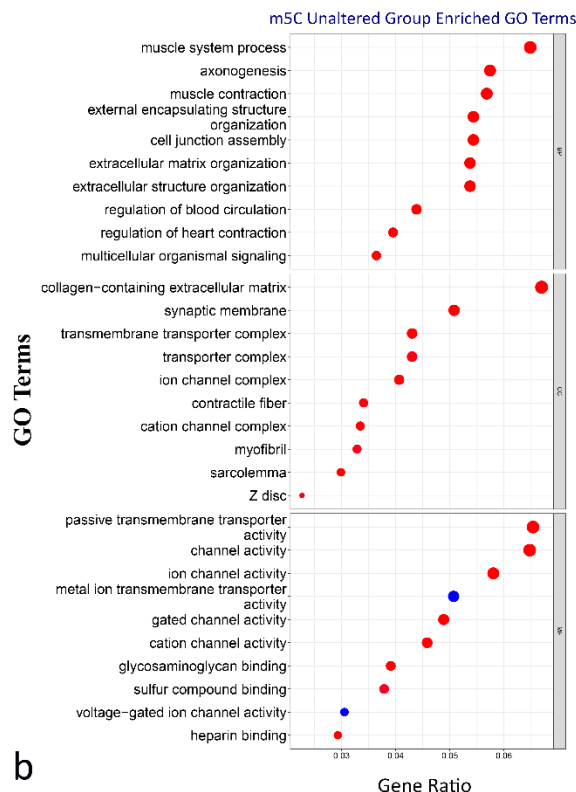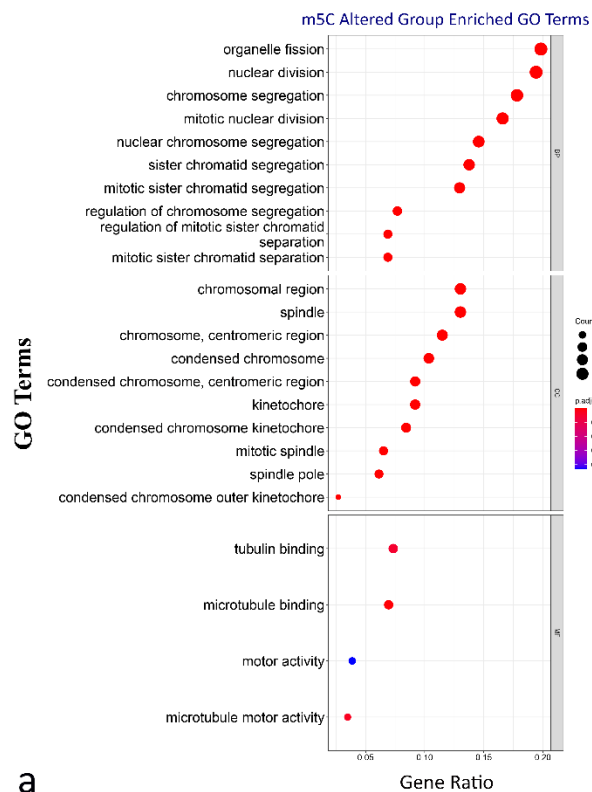
Figure 4.

Figure 4. Potential biomarkers of PRAD (a, b) Overall survival and ROC curve analysis of selected genes. (c) Expression levels of *UHRF1, NSUN2,* and *NEIL1* in different Gleason scores. (d) Immunohistochemical staining images of *UHRF1, NSUN2,* and *NEIL1*
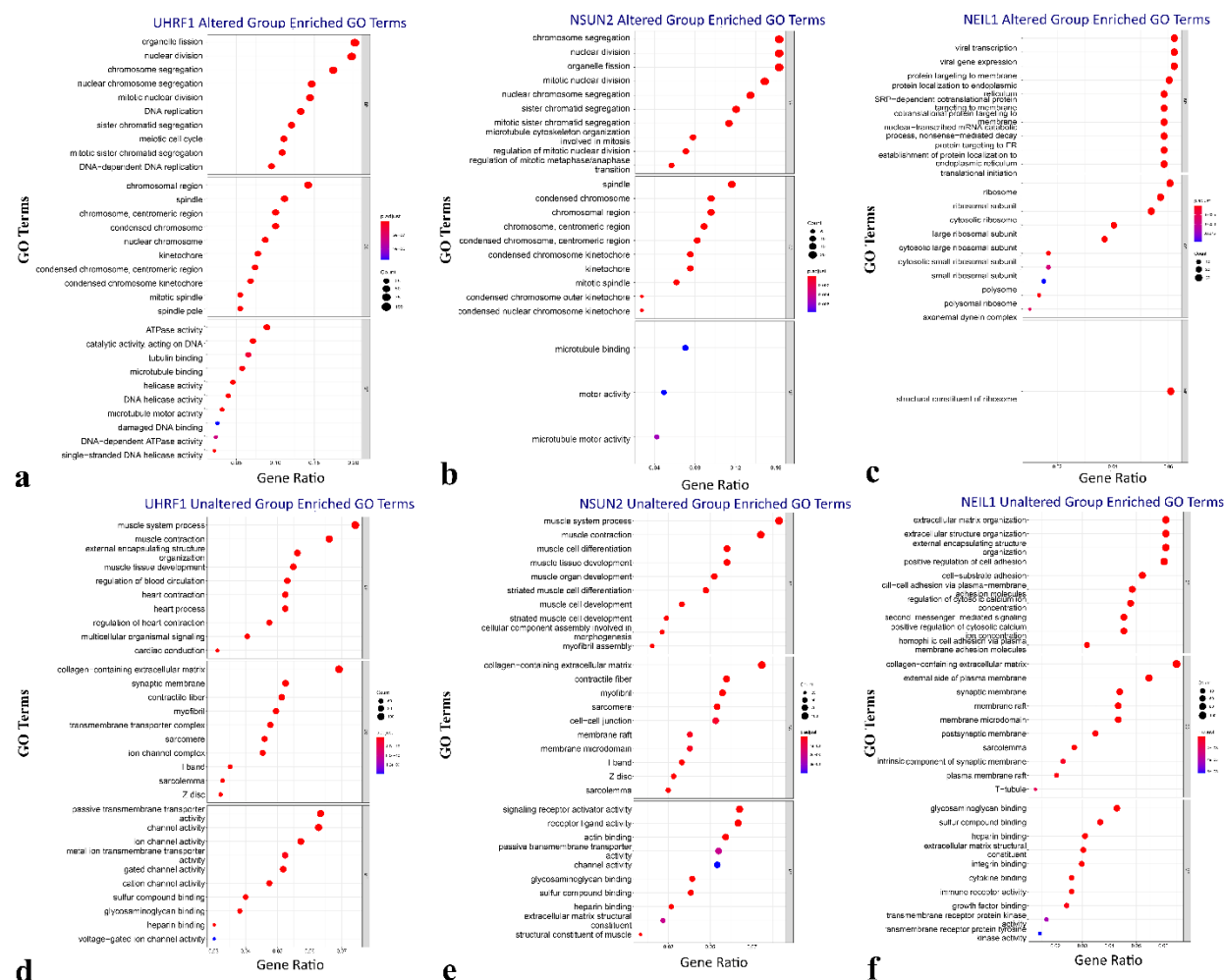
Figure 5.

Figure 5. Enriched GO Terms for *UHRF1* altered (a) and unaltered (b), *NSUN2* altered (c) and unaltered (d), and *NEIL1* altered (e) and unaltered (f), groups. BP; biological process, CC; cellular component, MF; molecular function.
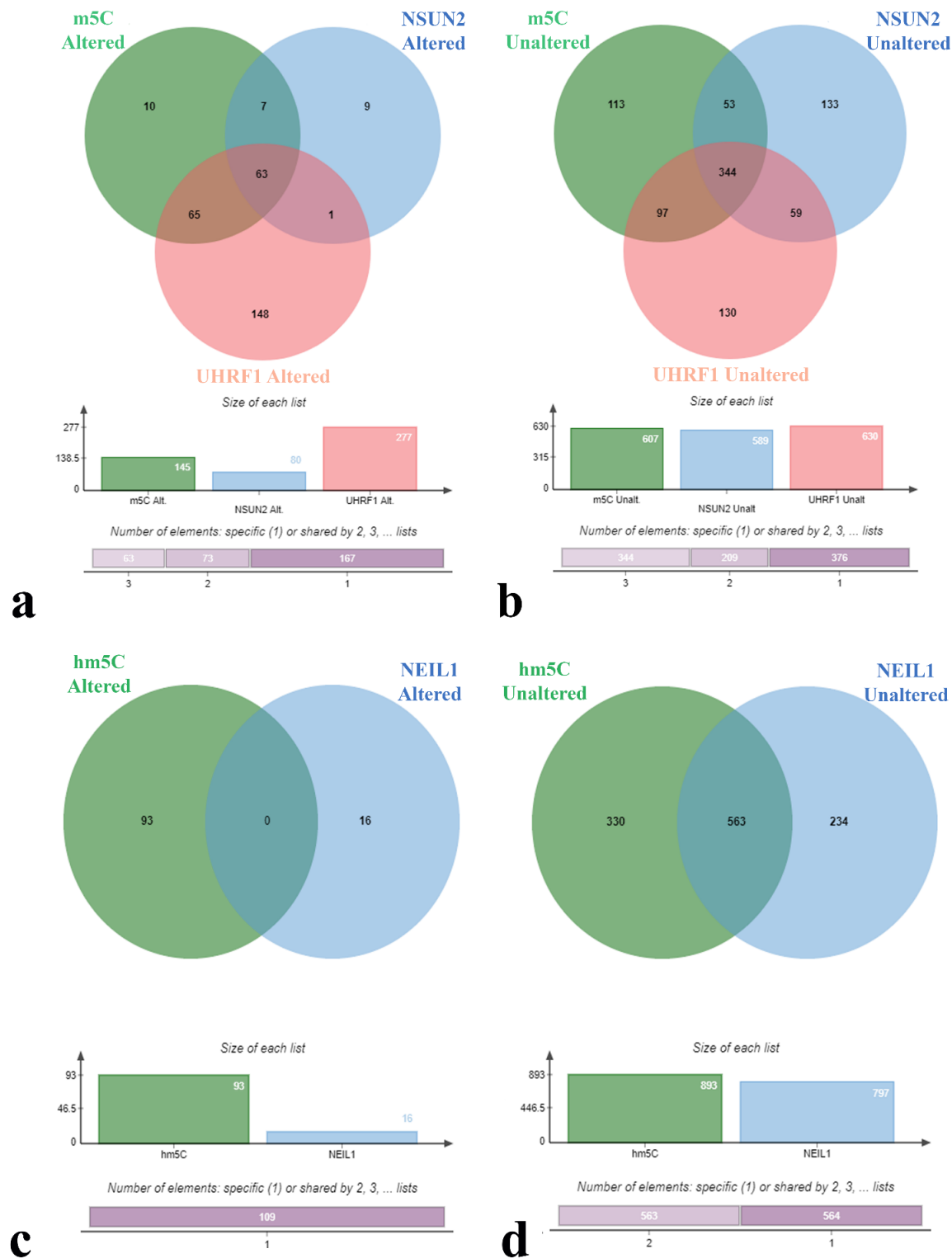
Figure 6.

Fig 6. Commonly enriched and unique BPs among m5C, *UHRF1, NSUN2* altered (a) and unaltered (b), and between hm5C and *NEIL1* altered (c) and unaltered (d) groups.