Research article

Implications of m^{5}C modifications in ribosomal proteins on oxidative stress, metabolic reprogramming, and immune responses in patients with mid-to-late-stage head and neck squamous cell carcinoma: Insights from nanopore sequencing

Gongbiao Lin [b], Haoxi Cai [c], Yihong Hong [d], Min Yao [a], Weiwei Ye [a], Wenzhi Li [a], Wentao Liang [a], Shiqiang Feng [a], Yunxia Lv [e,**], Hui Ye [a,***], Chengfu Cai [a,f,****], Gengming Cai [a,f,*]

[a] Department of Otolaryngology-Head and Neck Surgery, Haicang Hospital of Xiamen, Affiliated Haicang Hospital of Xiamen Medical College, The Sixth Hospital of Xiamen City, China
[b] Department of Otolaryngology-Head and Neck Surgery, The First Affiliated Hospital of Fujian Medical University, China
[c] School of Stomatology, Ningxia Medical University, China
[d] Community Health Service Center of Xidu Street, Fengxian District, Shanghai, China
[e] Department of Thyroid and Head and Neck Surgery, The Second Affiliated Hospital of Nanchang University, China
[f] Department of Clinical Medical, Fujian Medical University, China

ARTICLE INFO

ABSTRACT

Background: Head and Neck Squamous Cell Carcinoma (HNSCC) is a malignancy characterized by a high incidence and recurrence rate. 5-methylcytosine (m^{5}C) RNA modification is a common alteration affecting cancer progression; however, how m^{5}C operates within the tumor microenvironment of HNSCC remains to be elucidated.
Methods: We conducted Nanopore sequencing on 3 pairs of cancer and paracancerous tissues from mid- and late-stage HNSCC, obtaining 132 upregulated genes (transcriptomically upregulated, m^{5}C elevated) and 129 downregulated genes (transcriptomically downregulated, m^{5}C reduced). Subsequent Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed; a differential gene interaction network (PPI) was constructed, revealing the interactions of each gene with others in the network. Co-expression analysis was performed on the genes within the PPI, unveiling their expression and regulatory relationships. Through GSVA analysis, variations in related pathways under different states were identified. Furthermore, results of m^{5}C in lncRNA were screened, followed by target gene prediction.
Results: Sequencing results from the 3 pairs of mid- and late-stage HNSCC cancer and paracancerous tissues demonstrated that *RPS27A*, *RPL8*, and the lncRNAs including differentiation

---

* Corresponding authors. Department of Otolaryngology-Head and Neck Surgery, Haicang Hospital of Xiamen, Affiliated Haicang Hospital of Xiamen Medical College, The Sixth Hospital of Xiamen City, China.
** Corresponding author.
*** Corresponding authors.
**** Corresponding author. Department of Otolaryngology-Head and Neck Surgery, Haicang Hospital of Xiamen, Affiliated Haicang Hospital of Xiamen Medical College, The Sixth Hospital of Xiamen City, China.
E-mail addresses: 83394045@qq.com (Y. Lv), entyh@126.com (H. Ye), caicf@xmu.edu.cn (C. Cai), cgmkgx@hotmail.com (G. Cai).

antagonizing nonprotein coding RNA (DANCR), DCST1 antisense RNA 1 (CCDC144NL-AS1), Growth Arrest-Specific Transcript 5 (GAS5), Nuclear Paraspeckle Assembly Transcript 1 (NEAT1), and Small Nucleolar RNA Host Gene 3 (SNHG3), etc., under m^{5}Cregulation, have close connections with surrounding genes. The differentially m^{5}Cmodified genes are primarily involved in ribosomal protein synthesis, oxidative stress response, metabolic reprogramming, immunity, and other life processes; pathways like mitochondrial protein import and photodynamic therapy induced unfolded protein response are upregulated in the tumor, while pathways, including the classic P53, are suppressed. Analysis on m^{5}C-regulated long non-coding RNAs (lncRNAs) revealed tight associations with *RPS27*A and *RPL8* as well.

Conclusion: Our study identifies the key factors and signaling pathways involving m^{5}C in HNSCC. The findings suggest that ribosome-related genes might regulate ribosomal protein synthesis, oxidative stress response, metabolic reprogramming, and immune response through m^{5}C RNA modification by means like hypoxia and ferroptosis, thereby playing a pivotal role in the onset and progression of HNSCC. Hence, attention should be paid to the role of ribosomes in HNSCC. These findings may facilitate the precision and individualized treatment of patients with mid- and late-stage HNSCC in clinical settings.

1. Introduction

HNSCC is a highly heterogeneous malignant tumor, originating from various regions within the upper respiratory and digestive tracts. In 2018, on a global scale, head and neck cancer ranked as the seventh most common cancer, with 890,000 new cases and 450,000 deaths reported. In the United States, head and neck cancer comprises 3% of all cancers, with 51,540 new cases, accounting for over 1.5% of all cancer deaths, which totals to 10,030 deaths [1]. In China, approximately 90,000 new patients are diagnosed annually, with about 50,000 succumbing to the disease [2]. Over 60% of HNSCC patients are diagnosed at stage III or IV of the disease and possess a heightened risk of local recurrence in advanced stages of head and neck cancer (15% – 40%) as well as distant metastasis, presenting a poor prognosis (with an overall 5-year survival rate of less than 50%) [1].

Currently, surgical resection remains the most effective treatment approach for HNSCC. Patients who are not surgical candidates may undergo adjuvant radiotherapy or chemoradiotherapy (CRT) based on the staging of the disease. CRT has consistently been the primary modality for the treatment of pharyngeal carcinomas. Aside from early-stage oral carcinomas (treated solely by surgery) or carcinomas (which can be treated by either surgery or radiation), the majority of HNSCC patients require multimodal therapeutic strategies, thus necessitating multidisciplinary care. The monoclonal antibody cetuximab, which targets the Epidermal Growth Factor Receptor (*EGFR*; also known as *HER1*), has been approved by the U.S. Food and Drug Administration as a radiosensitizer, and can be used alone or in combination with chemotherapy for the treatment of recurrent or metastatic diseases. Immune checkpoint inhibitors have been sanctioned for the treatment of platinum-refractory recurrent or metastatic HNSCC. Nevertheless, the prognosis for metastatic, recurrent, and advanced HNSCC remains unsatisfactory, and there's a dearth of knowledge concerning the mechanisms behind HNSCC progression and metastasis. Hence, there's an urgent need to investigate the mechanisms underlying HNSCC progression to formulate novel therapeutic strategies [3].

RNA modifications, such as N6-methyladenosine (m^{6}A), play pivotal roles in epigenetic gene regulation and cellular functions and are intimately associated with a multitude of human diseases, including cancers, neurological disorders, and immune dysregulations. Recently, an RNA modification that has garnered increasing attention is m^{5}C modification. Analogous to m^{6}A, m^{5}C involves methylation at the fifth nitrogen of cytosine nucleotide, predominantly located within CG-enriched regions, and a myriad of m^{5}C methylation sites are present in both coding RNAs and non-coding RNAs. Currently identified genes related to m^{5}C are majorly categorized into three types: 1) writers: methyltransferases that facilitate RNA methylation modification, including *NSUN1*, *NSUN2*, *NSUN3*, *NSUN4*, *NSUN5*, *NSUN6*, *NSUN7*, *DNMT1*, *DNMT2*, *DNMT3A*, and *DNMT3B*. Erasers: demethylases, which negate methylation, include *TET2*. 3) readers: m5C-binding proteins that recognize and bind to m^{5}C sites on mRNA, including ALYREF. The "writer" forms a methyltransferase complex and amplifies m^{5}C levels, while the "eraser" operates as an m^{5}C demethylase, counteracting the "writer" function. Furthermore, the "reader" acts as a decoding effector, translating m^{5}C methylation information into functional signals. It is known that these regulators collaborate with m^{5}C modification to participate in the progression of various tumors. Abundant evidence suggests that *NSUN2*, being a crucial component of the m^{5}C methyltransferase complex, exerts non-negligible roles in numerous biological functions, such as cell differentiation, proliferation, migration, and is involved in tumorigenesis in an m^{5}C-dependent manner. Immunohistochemical analyses have verified the overexpression of *NSUN2* in various tumors, including those of the esophagus, stomach, liver, pancreas, cervix, prostate, kidney, bladder, thyroid, and breast. Using a specialized universal primer presents an opportunity to enhance understanding of expression patterns within the regulator family genes and potentially serve as predictive markers [4]. However, the regulatory mechanisms and distribution of m^{5}C in human HNSCC remain largely uncharted, and the expression and function of these regulators in HNSCC are yet to be elucidated [5].

Ribosomes represent a complex assembly, composed of ribosomal RNA and ribosomal proteins, functioning as translational machinery that converts messenger RNA into proteins. Disruptions in these checkpoints and pathways can potentially lead to the hyperactivation of ribosomal biogenesis. The onset of ribosomopathies is associated with mutations in ribosomal proteins, rRNA processing, and ribosomal assembly factors, elevating the risk of malignant tumor development. Recent findings suggest that cancer cells possess a distinct set of ribosomes that promote the translational program of oncogenes, regulate cellular functions, and drive

metabolic reprogramming. Contemporary research has linked ribosomal protein mutations and aberrant ribosomes to poor prognosis, underscoring the promise of ribosome-targeted therapies for cancer patients [6]. Ribosomal protein S27a (RPS27a) encodes a fusion protein comprising an N-terminal ubiquitin and a C-terminal ribosomal protein S27a. RPS27a is a component of the ribosomal 40S subunit and belongs to the ribosomal protein S27AE family. It contains a C4-type zinc finger domain and is localized in the cytoplasm (https://www.genecards.org/cgi-bin/carddisp.pl?gene=RPS27A&keywords=RPS27a). RPS27a, as a ribosomal protein (RP), not only assembles into the ribosome but also functions independently of it. As an RNA-binding protein, RPS27a undertakes extra-ribosomal functions, encompassing ribosomal biogenesis and post-translational modification processes. RPS27a has been substantiated as a direct transcriptional target of p53. It plays roles in promoting proliferation, regulating cell cycle progression, and inhibiting apoptosis through overexpression in DNA damage and in various cancers, such as renal, breast, and colon cancer. Furthermore, RPS27a participates in the progression of numerous diseases or cancers, such as liver cancer, chronic myeloid leukemia, and is associated with poor prognosis. It potentially serves as a latent target for EBV-induced LMP1-positive cancer cells, with its upregulation fostering the growth of colorectal cancer cells and inhibiting apoptosis, implicating in the pathogenesis mechanism of diabetes pancreatic ductal adenocarcinoma (PDAC), and it is also one of the pathway elements promoting the proliferation of HPV-immortalized cervical epithelial cells (H8), facilitating the occurrence of cervical cancer. The murine double minute 2 (MDM2)/P53 axis is the principal pathway through *RPS27A* regulates cancer development [7,8]. Liao et al. uncovered that *RPS27A/USP9X* induces the upregulation of *PPP1R14B*, elevates STMN1 activity, and results in breast cancer progression and resistance to paclitaxel [9]. Research by Li et al. suggests that RPS27A interacts with RPL11; the knockout of *RPS27A* enhances the binding of RPL11 and MDM2, thereby inhibiting MDM2-mediated p53 ubiquitination and degradation. Moreover, *RPS27A* plays a vital role in the progression and prognosis of lung adenocarcinoma (LUAD), positioning itself as a potential therapeutic target for LUAD patients [10]. Chen et al. elucidated the oncogenic role of *GYS1* in the proliferation and glycogen metabolism of clear cell renal cell carcinoma (ccRCC) cells. Overexpression of *GYS1* facilitated tumor growth, while its silencing inhibited tumor growth through the activation of the canonical NF-κB pathway. An indirect interaction between GYS1 and NF-κB, mediated by RPS27a, promoted the phosphorylation and nuclear import of p65 [11]. Ribosomal Protein S8 (RPS8), encoding a ribosomal protein which is a constituent of the 40S subunit, belongs to the ribosomal protein S8E family (https://www.genecards.org/cgi-bin/carddisp.pl?gene=RPS27A&keywords=RPS8). The human RPS8 gene spans 3,161 bp and encompasses six exons. It is chromosomally located at 1p32-p34.1 [12]. RPS8 is associated with pancreatic ductal adeno-carcinoma (PDAC) and may serve as a potential therapeutic target to enhance survival rates in pancreatic cancer [13]; it is upregulated in astrocytoma [14], significantly correlates with recurrence-free survival and overall survival time in lung adenocarcinoma (LUAD) patients [15], and can act as a novel biomarker for patients with alcohol-related liver cancer [16]. Ribosomal Protein L8 (RPL8), which encodes a ribosomal protein and is a component of the 60S subunit, belongs to the ribosomal protein L2P family and is located in the cytoplasm (https://www.genecards.org/cgi-bin/carddisp.pl?gene=RPL8). RPL8, a ferroptosis-related gene, exhibits expression significantly altered with the grading of pancreatic tumors, correlates with the prognosis of pancreatic cancer, and is associated with the progression of glioma [17]. Platelets, which participate in tumor angiogenesis and cancer progression, stand as potential sources of cancer biomarkers. Studies have revealed an upregulation of RPL8 in the platelet proteome of cancer patients [18].

In this study, we aim to elucidate the regulatory mechanisms and distribution of m^{5}C in HNSCC using nanopore sequencing. By analyzing differential gene expression and constructing interaction networks, we seek to uncover underlying biological processes and signaling pathways. Additionally, we aim to predict target genes for selected lncRNAs, offering potential avenues for personalized therapeutic strategies in HNSCC patients.

2. Materials and methods

2.1. Data processing

2.1.1. RNA extraction, library preparation, and sequencing

Prepared RNA of each sample was used for a DRS library preparation using the Oxford Nanopore DRS protocol (SQK-RNA002, Oxford Nanopore Technologies). For reversed connector connection, 9 μL prepared RNA, 3 μL NEBNext Quick Ligation Reaction Buffer (NEB), 1 μL RT Adapter (RTA) (SQK-RNA002) and 2 μL T4 DNA Ligase (NEB) were mixed together and incubated under 25 °C for 10 min. Afterwards, 8 μL 5 × first-strand buffer (NEB), 2 μL 10 mM dNTPs (NEB), 9 μL Nuclease-free water, 4 μL 0.1 M DTT (Thermo Fisher) and 2 μL SuperScript III Reverse Transcriptase (Thermo Fisher Scientific, 18080044) were added into the above 15 μL reaction system and incubated under 50 °C for 50 min then 70 °C for 10 min. Reverse-transcribed mRNA was purified with 1.8 × Agencourt RNAClean XP beads and washed with 23 μL Nuclease-free water, and subsequently sequencing adapters were added using 8 μL NEBNext Quick Ligation Reaction Buffer, 6 μL RNA Adapter (RMX) and 3 μL T4 DNA Ligase. The mix was purified and washed again as above, and 50 μL RRB (SQK-RNA002) were added with 35 μL Nuclease-free water. This final reaction system was loaded into Nanopore R9.4 sequencing micro-array and sequenced for 48 – 72 h using PromethION sequencer (Oxford Nanopore Technologies). This was performed by Seqhealth Technology Co., LTD (Wuhan, China).

2.1.2. Analysis of RNA sequence

Initially, the raw sequencing data were filtered through SOAPnuke (version 1.6.0), excising low-quality and contaminated reads. Clean Reads were first clustered based on Unique Molecular Identifier (UMI) sequences, grouping identical UMI sequences into the same cluster. Upon generating all sub-clusters, multiple sequence alignments were executed, deriving a consensus sequence for each sub-cluster. Preprocessing, Alignment, and Analysis of New Genes and Transcripts. The Guppy software (version 5.0.16) was employed for base calling of raw reads and evaluating read quality, with Guppy also executing trimming of adapter sequences. Nanofilt (version

2.7.1) filtered low-quality (Q-value < 7) and short-length reads (< 50 bp) [19]. Remaining clean reads were corrected using Fclmr2 (version 0.1.2) and short-length RNA-Seq data [20]. Subsequently, clean reads from each library were aligned with the Homo sapiens genome assembly T2T-CHM13v2.0 genome using Minimap2 (version 2.17-r941) [21]. Samtools (version 1.10) calculated the alignment rate of clean reads with reference genes [22]. Fair was utilized to obtain consensus sequences based on alignment results (version: 1.5.0; parameters: t 20). To minimize result redundancy, alignments differing only in 5′ exons were merged, and non-redundant transcript sequences were attained using StringTie software (version: 2.1.4; parameters: L –R conservative) [23]. Gffcompare software (version: 0.12.1; parameters: R –C -K -M) was utilized to compare transcripts with known genome transcripts and identify new transcripts and genes [24]. In order to acquire comprehensive functional information for new transcripts, annotations were performed on transcripts across seven databases, including NR, Pfam, UniProt, KEGG, GO, KOG/COG, and PATHWAY, based on sequence similarity and base sequence similarity. The differential expression genes were screened by R package (DEseq2) based on a cutoff criterion of fold change ≥ 2 and FDR < 0.05 [25].

2.1.3. Analysis of m^5C RNA methylation

m^5C site identification was calculated based on the alternative model of methylation model constructed by Tombo software. Methylkit software was used to analyze differential methylation sites. Logistic regression test is used detect differential methylation; When there was only one sample in each group.

2.1.4. Structural analysis of gene transcripts

Compare the transcript with the known transcript of the genome. If there is a region outside the boundary of the original transcript, extend the untranslated region of the transcript upstream and downstream, so as to correct the boundary of the transcript. Suppa2 software (https://github.com/comprna/SUPPA; parameters: f IOE - e se SS MX RI FL) were used to obtain the variable splice type of each sample, and supa2 (DiffSplice) was used to identify the difference variable splice between groups. Tofu software (version: 13.0.0; parameter: default) was used to compare and find the fusion transcript. CNCI (version: 2.0; parameter: default) and CPC2 (version: standalone_python 3 v1.0.1) software and Pfam were used to predict the coding potential of the newly identified transcripts. Tapas software(https://github.com/arefeen/TAPAS Parameter: L 50) were usded to identifies APA in combination with the reference sequence of RefSeq database.

2.1.5. Functional annotation (GO) and pathway enrichment (KEGG) analysis

The differentially expressed genes (DEGs) identified in the aforementioned control groups underwent functional annotation, utilizing the annotation and visualization of GO terms and meta-views. Subsequently, the DEGs were introduced into FunRich, a tool for functional enrichment analysis (http://www.funrich.org/), utilized for the KEGG pathway analysis. GENEMANIA was employed to establish a gene-gene interaction network of the DEGs to evaluate their functionalities.

2.1.6. Construction of the PPI network

The overlapping genes were analyzed for intergene interactions using STRING v11.0 (https://string-db.org/) with default parameters. This analysis provided insights into their relationships, with a selection criterion set at a combined score of ≥ 0.4. Subsequently, network visualization was generated using Cytoscape software (v3.5.1; http://cytoscape.org/) with default parameters. Co-expressed hub genes: Co-expression analysis of genes within the PPI facilitated the elucidation of their expression relationships, subsequently uncovering their regulatory associations within this experiment.

2.1.7. lncRNA analysis

The newly identified transcripts underwent coding potential prediction using CNCI (version: 2.0; default parameters), CPC2 (version: standalone_python3 v1.0.1) software, and Pfam. CPC2, a coding potential calculation software, aligns the transcripts with known protein databases using BLAST alignment. It evaluates the coding potential of each transcript based on the biological sequence features of its coding frame using a support vector machine classifier. CNCI effectively distinguishes coding and non-coding sequences based on the frequency spectrum of adjacent tri-nucleotides and can efficiently predict the coding potential of incomplete transcripts and antisense transcript pairs. The Pfam database contains high-quality domain structures of most known proteins. The number of transcripts predicted as noncoding by each software was statistically analyzed for intersection. Target genes corresponding to lncRNA were forecasted using the ENCORI database (https://starbase.sysu.edu.cn/index.php) with default parameters.

2.1.8. Gene Set Variation Analysis (GSVA)

The GSVA package in R software (version 3.6) was utilized for the analysis of m^5C. Variations in related pathways under distinct conditions were identified from the TPM (Transcripts Per Million) values obtained from RNA-seq data of two sets of six samples. Pathways with divergent enrichment were filtered using the *t*-test method.

3. Results

3.1. Clinical information of head and neck squamous cell carcinoma specimens collection

In our study of HNSCC, we first collect three clinical specimens to elucidate the molecular signatures associated with this disease (Table 1). Transcriptomic and epitranscriptomic sequencing were performed using nanopore sequencing technology for these

specimens.

3.2. Diffrentially methylated m⁵C genes and differential expressed genes participate in ribosome related pathways

We identified the global mRNA expression patterns in normal and cancer groups. The results showed that a total of 1,416 mRNAs were significantly dysregulated in cancer group compared with normal group, including 833 down-regulated mRNAs and 583 up-regulated mRNAs (Supplementary Table 1). Additionally, we identified 114,200 m^5C sites (Supplementary Table 2). Then, we filtered for both differentially expressed genes and differentially expressed genes associated with m^5C.; Through the comparative filtering of two sets of differential genes, genes which were upregulated in both the transcriptome and m^5C, as well as genes which were downregulated in both, were identified. We acquired 132 upregulated genes (upregulated in both transcriptome and m^5C) and 129 downregulated genes (downregulated in both transcriptome and m^5C) (Supplementary Table 3).

Subsequently, we explored the potential function of differentially expressed genes and the overlap of these genes with differential methylated m^5C. The GO biological process analysis results showed that the upregulated genes (upregulated in the transcriptome and m^5C) were significantly associated with cytosolic ribosome; ribosomal subunit; ribosome; polysome and so on (Fig. 1A). The ribosome functions as a molecular machine, orchestrating the translation of mRNA into proteins, which indicating that these genes would be associated with translation process. Additionally, the downregulated genes (downregulated in the transcriptome and m^5C) were enriched in blood microparticle; proton-transporting ATP synthase complex, coupling factor F(0); cytosolic small ribosomal subunit and so on (Fig. 1B). The findings suggest that ribosomes could play crucial roles in the cancer process, while m^5C modifications may influence transcript processing.

3.3. KEGG analysis

Pathway analysis unveiled those upregulated genes, concurrent in both the transcriptome and m^5C datasets, exhibited significant enrichment in pathways related to ribosome function, coronavirus disease (COVID-19), viral myocarditis, and more (Fig. 2A). Conversely, downregulated genes, showing consistent downregulation in both datasets, were notably enriched in pathways associated with staphylococcus aureus infection, complement activation, coagulation cascades, among others (Fig. 2B). Notably, these genes play pivotal roles in diverse biological processes, including ribosomal protein synthesis, response to oxidative stress, metabolic reprogramming, and modulation of immune responses.

3.4. Key gene identification and the prediction of regulation

In order to understand the interactions between proteins encoded by differentially methylated and differentially expressed genes, these mRNAs were uploaded to the STRING database for PPI network construction. By constructing the interaction network of differentially methylated and differentially expressed genes, we ascertain the relationship each gene within the network shared with others (Fig. 3A). Following the extraction and sorting of the degree table for gene PPI (Table 2), we meticulously scrutinized the data to identify hub genes within the network. Employing co-expression analysis, we delved into the intricate relationships among these hubs and their neighboring genes within the PPI network (Table 3). Remarkably, our analysis illuminated the pronounced interconnectedness of RPS27A and RPL8 with surrounding genes, underscoring their pivotal roles in the regulatory landscape, particularly under the influence of m^5C modification. Furthermore, through conducting co-expression analysis on the genes within the PPI, we were able to discern their expressive interrelations, thereby revealing their regulatory relationships (Fig. 3B).

3.5. GSVA analysis

To analyze the functional dynamics and pathway alterations across a spectrum of samples, GSVA was employed. We utilized TPM values obtained from RNA-seq data of two sets of six samples to investigate variations in related pathways under distinct conditions. Through the analysis of gene signal values and their associated pathways, we predicted pathways associated with varied disease states. Notably, our findings indicate the upregulation of several signaling pathways in tumors, such as mitochondrial protein import and photodynamic therapy induced unfolded protein response. Conversely, numerous pathways, including the classical P53 pathway, were observed to act as inhibitory pathways within the tumors (Fig. 4).

3.6. Analysis of lncRNA associated with m⁵C regulation

We identified a total of 154 m^5C gene-related lncRNAs within HNSCC samples compared with normal samples (Supplementary

Table 1
Clinical data of the samples.

Patients	Gender	Age	Tobacco	Alcohol	Stage
I	male	58	+	+-	T3N2$_C$M0
II	male	60	–	–	T4N3$_b$M0
III	male	47	+	+	T4aN1M0
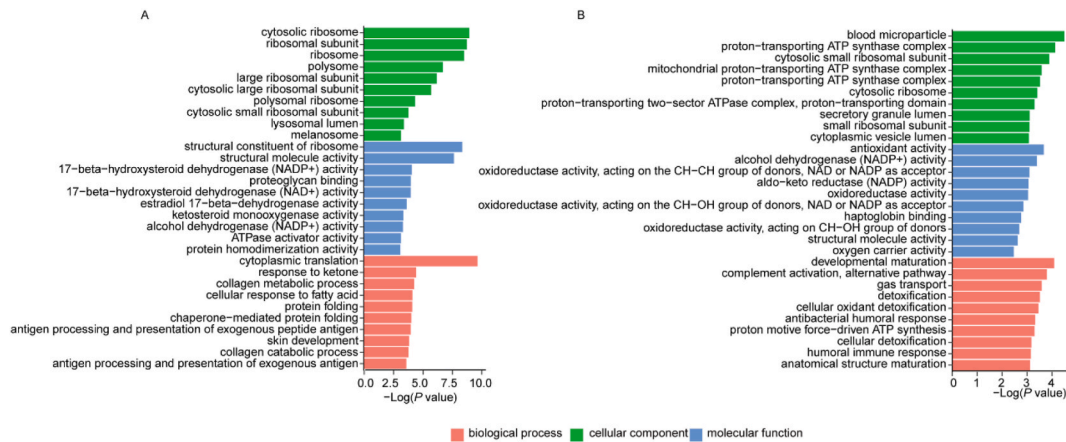
Fig. 1. The GO analysis of differential expressed genes. A. the upregulated genes (upregulated in the transcriptome and m⁵C). B. the downregulated genes (downregulated in the transcriptome and m⁵C).
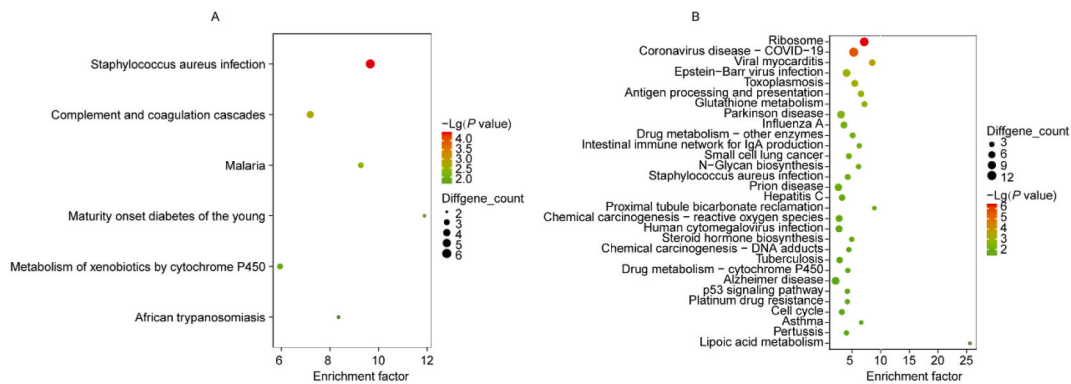
Fig. 2. The KEGG analysis for differential expressed genes. A. The upregulated genes (with elevated m⁵C transcription). B. The downregulated genes (with reduced m⁵C translation).

Table 4). The m⁵C alterations pertaining to lncRNA are sifted from the differential m⁵C table. Subsequent to this, target genes of lncRNA are prognosticated. LncRNA with upregulated m⁵C, mRNA with transcriptional upregulation, lncRNA with downregulated m⁵C, and mRNA with transcriptional downregulation are sieved. Among these genes NEAT1, SNHG3 and GAS5 operate as crucially down-regulated lncRNAs, while EPB41L4A-AS1, PSMA3-AS1 and SLX1B-SULT1A4 act as pivotal up-regulated lncRNAs (Fig. 5).

4. Discussion

Despite m⁵C being a prevalent modification in RNA, its regulation and biological functions in pathological conditions, such as cancer, remain elusive [26]. m⁵C, a common mRNA modification, is located in the untranslated regions (UTRs) of mRNA transcripts. m⁵C has been identified in rRNA and tRNA and was recently discovered in mRNA. The methylation process of m⁵C is analogous to m⁶A, and m⁵C RNA methyltransferases (RNMT) belong to the DNA methyltransferase family. Although the roles of these enzymes in tumor development are still unclear, an increasing number of individuals believe that m⁵C levels are associated with cancer [27].

In order to investigate the regulatory mechanisms and distribution of m⁵C in HNSCC, we employed nanopore sequencing technology to conduct Direct RNA sequencing on a total of six specimens, including both cancerous and adjacent non-cancerous tissues from patients with advanced-stage HNSCC. We filtered the transcriptomic differentially expressed genes and those with m⁵C modifications, identifying 132 upregulated genes (with concurrent transcriptomic and m⁵C upregulation) and 129 downregulated genes (with concurrent transcriptomic and m⁵C downregulation). Genes at the intersection were subjected to intergene interaction prediction. Subsequent co-expression analysis of genes within the PPI network facilitated the discernment of their expression relationships, thereby elucidating their regulatory correlations. Notably, through the co-expression analysis of central genes in the PPI, RPS27A and RPL8 were found to be intimately linked with surrounding genes under m⁵C regulation.

In our study, we undertook an analysis of lncRNAs and identified those that are subject to RNA m⁵C modification. This group includes the lncRNA differentiation antagonizing nonprotein coding RNA (DANCR), the lncRNA DCST1 antisense RNA 1 (CCDC144NL-AS1), Growth Arrest Specific 5 (GAS5), Nuclear Paraspeckle Assembly Transcript 1 (NEAT1), and Small Nucleolar RNA Host Gene 3
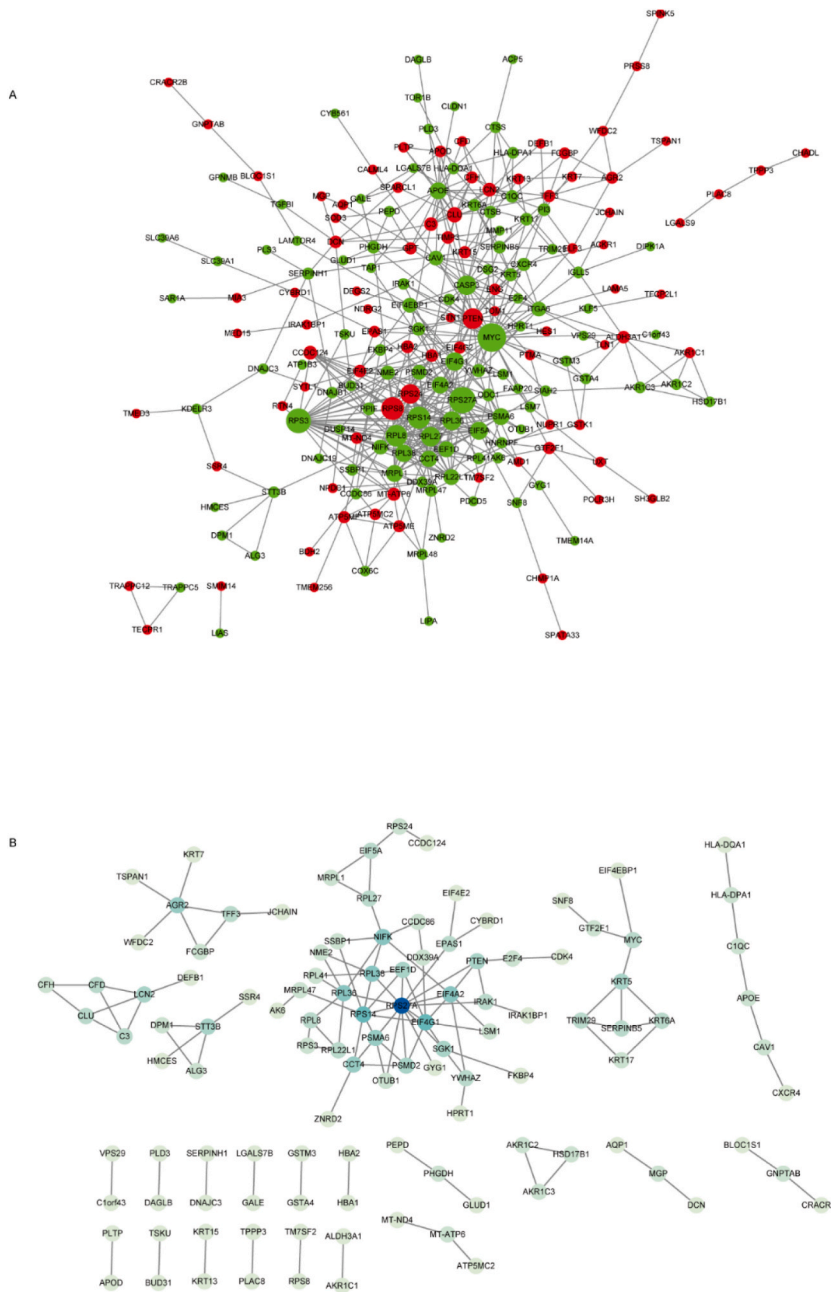
Fig. 3. The PPI of intersecting genes and co-expression relationships among genes within the PPI. The colors indicate the magnitude of upregulation or downregulation, with red signifying upregulated genes and green indicating downregulated genes. Deep blue denotes genes of higher degree, while lighter shades indicate those of a lesser degree. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(SNHG3). DANCR, a cancer-associated lncRNA, is upregulated in various cancer types such as lung cancer, gastric cancer, breast cancer, and hepatocellular carcinoma (HCC) [28]. Its dysregulation plays a pivotal role in in vitro cancer cell proliferation, apoptosis, migration, invasion, chemotherapy resistance, as well as in vivo tumor growth and metastasis [29]. CCDC144NL-AS1, characterized by its 144N-terminal coiled-coil domain-like structure, is aberrantly expressed in HCC, ovarian cancer (OC), gastric cancer (GC), non-small cell lung cancer (NSCLC), and osteosarcoma. It has potential as a prognostic biomarker and therapeutic target for cancers [30]. GAS5 regulates various cellular functions in diverse human cancers including leukemia, cervical cancer, breast cancer, OC, prostate cancer, bladder cancer, lung cancer, GC, colorectal cancer, liver cancer, osteosarcoma, and brain cancer. It acts as a tumor suppressor in various tumor types, with its upregulation increasing tumor sensitivity to chemotherapy or radiotherapy, playing a previously underestimated but crucial role in therapy-induced resistance [31]. GAS5 can modulate intracellular levels of autophagy,

Table 2

Degree of gene in PPI. Degree signifies the quantity of other genes related to the gene in question; for instance, a degree of 10 indicates that there are 10 genes interacting with it. A larger degree denotes a higher number of genes interacting with it (Only the top 10 rankings are displayed).

Name	Style	Degree
MYC	up	37
RPS27A	up	33
RPS3	up	30
RPS8	down	26
RPS14	up	25
RPL8	up	24
PTEN	down	22
RPL27	up	20
CASP3	up	20
RPS24	down	20

Table 3

Co-expression analysis of hub genes in the PPI.

Name	Degree
RPS27A	16
RPL8	14
RPS3	12
RPS14	11
RPL36	10
MYC	10
RPL22L1	8
NIFK	8
RPL38	8
PSMD2	8
EIF4G1	7
RPL27	7
NME2	6
CCT4	6

Fig. 4. Variations in relevant pathways under diverse sample conditions. The color scheme transitions from blue to red in the figure, representing an elevation in enrichment score values, indicating pathway activation. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

oxidative stress, and immune cell function, playing a role in cell proliferation, invasion, and apoptosis [32]. NEAT1, a long non-coding RNA (lncRNA), promotes cell growth, migration, and invasion while inhibiting apoptosis in cancer cells when elevated [33]. NEAT1 modulates the invasive phenotype of cancer cells, promotes tumor glycolysis and cancer metastasis, and interacts with tumor
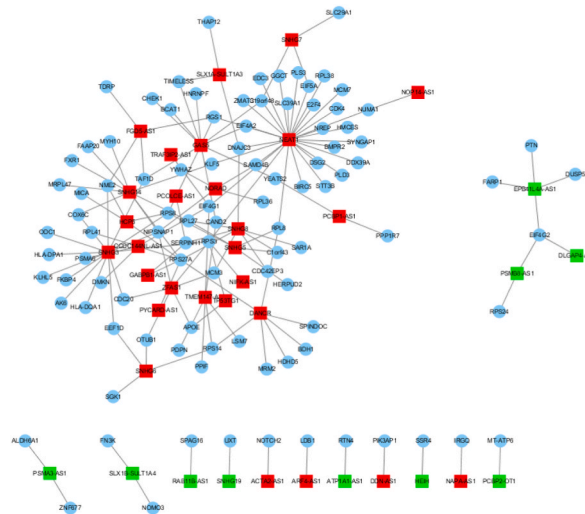
Fig. 5. Network of m^{5}C-associated lncRNA, mRNA, and Corresponding Target Genes. Green signifies lncRNA is downregulated, red indicates lncRNA is upregulated, and blue represents mRNA. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

suppressor transcriptional inhibitors. It has also been shown to interact with cyclin-dependent kinases and enzymes associated with glucose metabolism [34,35]. SNHG3 is overexpressed in the majority of tumors, including breast cancer, osteosarcoma, glioma, laryngeal cancer, GC, colorectal cancer, renal cell carcinoma, HCC, lung cancer, acute myeloid leukemia, and epithelial ovarian cancer. It plays a crucial role in the biology of tumor cells, impacting processes like proliferation, migration, invasion, and apoptosis [36]. Through the analysis of lncRNAs associated with m^{5}C regulation, we also identified a strong association with ribosomal genes such as RPS27a and RPL8.

Consequently, we posit that *RPS27A* and *RPL8* exert a significant influence on mid-to-late-stage squamous cell carcinoma of the head and neck through m^{5}C modifications. An analysis of the sequencing data via Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis reveals that genes exhibiting m^{5}C modifications predominantly participate in vital biological processes such as ribosomal protein synthesis, oxidative stress response, metabolic reprogramming, and immunity. Research by Han et al. identified that in high-risk subgroups, the leading terms were notably correlated with "ribosome," "oxidative phosphorylation," "non-alcoholic fatty liver disease," and "synaptic vesicle cycle," among others [37], mirroring our results closely. This signifies that ribosomal proteins play a crucial regulatory role in life processes such as oxidative stress response, metabolism, and immune response in mid-to-late-stage head and neck squamous cell carcinoma.

Further investigation through GSVA of relevant pathways reveals that signal pathways, such as mitochondrial protein import and photodynamic therapy induced unfolded protein response, are upregulated in tumors, while pathways, including the classical P53, serve as inhibitory pathways in the tumor milieu. Through GO, KEGG analysis, PPI, and co-expression analyses, it was discerned that hub genes, which maintain a close association and co-expression with adjacent genes, predominantly encode ribosomal proteins. Notably, classic oncogenes and tumor suppressor genes such as *MYC* and *PTEN* respectively, also undergo significant alterations. Our research comprehensively underscores the pivotal role of ribosomal proteins in cancer. They may influence life processes such as ribosomal protein synthesis, oxidative stress response, metabolic reprogramming, and immune response through mechanisms of ferroptosis and classical pathways such as P53 and NF-kappaB (NF-κB), thereby playing a substantial role in cancer dynamics.

In summary, we utilized Nanopore sequencing technology to conduct direct RNA sequencing on both cancerous and adjacent noncancerous tissues derived from clinical cases of advanced-stage HNSCC. Co-expression analysis of pivotal genes in PPI revealed that *RPS27A* and *RPL8*, along with LncRNAs such as DANCR, CCDC144NL-AS1, GAS5, NEAT1, and SNHG3, establish robust connections with neighboring genes under the regulation of m^{5}C. Genes undergoing m^{5}C variations principally participate in life processes, such as ribosomal protein synthesis, oxidative stress response, metabolic reprogramming, and immune response; signaling pathways such as mitochondrial protein import and photodynamic therapy induced unfolded protein response are upregulated in tumors, whereas several pathways, including the classical P53 pathway, are suppressed in tumor contexts. This implies that ribosomal-related genes regulate ribosomal protein synthesis, oxidative stress, metabolic reprogramming, and immune responses through m^{5}C RNA modifications, adopting mechanisms such as hypoxia and ferroptosis, thereby playing a significant role in the occurrence and development of HNSCC. Furthermore, *RPS27A* and *RPL8* may exert central regulatory effects therein. Therefore, we posit that attention should be accorded to the role of ribosomes in HNSCC. However, the sample size of this study was relatively small, with only six specimens analyzed. This might limit the generalizability of our findings to a broader population of HNSCC patients. Additionally, while nanopore sequencing technology offers several advantages, including long-read capabilities and direct RNA sequencing, it also has inherent limitations such as higher error rates compared to other sequencing platforms. Therefore, caution should be exercised in interpreting the results, and validation studies using alternative methodologies are warranted to confirm our findings. Furthermore, our study

focused on the analysis of m⁵C modifications and their association with gene expression changes in HNSCC. While we identified significant correlations between m⁵C-modified genes and ribosomal-related genes such as *RPS27A* and *RPL8*, further mechanistic studies are needed to elucidate the functional implications of these modifications in cancer progression.

Data availability

The raw sequence data have been deposited in the Genome Sequence Archive in the National Genomics Data Center (NGDC), Beijing Institute of Genomics, Chinese Academy of Sciences, under accession number (HRA007694).

Ethics statement

Our research was conducted in accordance with the principles of the Declaration of Helsinki. The study was approved by ethics review form for branch for medical research and clinical technology application, ethics committee of first affiliated hospital of fujian medical university under the approval number |2015|084-2. Written informed consent was obtained from all participants prior to collecting clinical specimens for research purposes. All specimens were stored securely and handled using appropriate procedures to maintain their integrity. De-identified data were used in all analyses to protect patient privacy.

CRediT authorship contribution statement

Gongbiao Lin: Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Haoxi Cai:** Writing – review & editing, Writing – original draft, Validation, Software, Investigation, Formal analysis, Data curation. **Yihong Hong:** Writing – review & editing, Writing – original draft, Software, Resources, Investigation, Data curation, Formal analysis. **Min Yao:** Writing – review & editing, Writing – original draft, Software, Resources, Data curation. **Weiwei Ye:** Writing – review & editing, Writing – original draft, Software, Formal analysis, Data curation. **Wenzhi Li:** Writing – review & editing, Writing – original draft, Formal analysis, Data curation. **Wentao Liang:** Writing – review & editing, Writing – original draft, Formal analysis, Data curation. **Shiqiang Feng:** Writing – review & editing, Writing – original draft, Formal analysis, Data curation. **Yunxia Lv:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Formal analysis, Data curation. **Hui Ye:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology. **Chengfu Cai:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology. **Gengming Cai:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Cai gengming reports financial support was provided by The Natural Science Foundation of Fujian Provincial Department of Science and Technology (No. 2020J011281). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e34529.

References

[1] L.Q.M. Chow, Head and neck cancer, N. Engl. J. Med. 382 (2020) 60–72.
[2] C. Xu, et al., Association between cancer incidence and Mortality in Web-based data in China: Infodemiology study, J. Med. Internet Res. 21 (2019) e10677.
[3] D.E. Johnson, et al., Head and neck squamous cell carcinoma, Nat. Rev. Dis. Prim. 6 (2020) 92.
[4] Xia, H. et al. MultiPrime: a reliable and efficient tool for targeted next-generation sequencing. iMeta, e143.
[5] Y. Hu, et al., NSUN2 modified by SUMO-2/3 promotes gastric cancer progression and regulates mRNA m5C methylation, Cell Death Dis. 12 (2021) 842.
[6] A.R. Elhamamsy, B.J. Metge, H.A. Alsheikh, L.A. Shevde, R.S. Samant, Ribosome biogenesis: a central player in cancer metastasis and therapeutic resistance, Cancer Res. 82 (2022) 2344–2353.
[7] Q. Wang, Y. Cai, X. Fu, L. Chen, High RPS27A expression predicts poor prognosis in patients with HPV type 16 cervical cancer, Front. Oncol. 11 (2021) 752974.
[8] J. Luo, H. Zhao, L. Chen, M. Liu, Multifaceted functions of RPS27a: an unconventional ribosomal protein, J. Cell. Physiol. 238 (2023) 485–497.
[9] L. Liao, et al., Protein Phosphatase 1 subunit PPP1R14B Stabilizes STMN1 to promote progression and paclitaxel resistance in Triple-Negative breast cancer, Cancer Res. 83 (2023) 471–484.
[10] H. Li, et al., Loss of RPS27a expression regulates the cell cycle, apoptosis, and proliferation via the RPL11-MDM2-p53 pathway in lung adenocarcinoma cells, J. Exp. Clin. Cancer Res. 41 (2022) 33.
[11] S.L. Chen, et al., GYS1 induces glycogen accumulation and promotes tumor progression via the NF-κB pathway in Clear Cell Renal Carcinoma, Theranostics 10 (2020) 9186–9199.

[12] B. Davies, M. Fried, The structure of the human intron-containing S8 ribosomal protein gene and determination of its chromosomal location at 1p32-p34.1, Genomics 15 (1993) 68–75.
[13] R. Chen, et al., Proteins associated with pancreatic cancer survival in patients with resectable pancreatic ductal adenocarcinoma, Lab. Invest. 95 (2015) 43–55.
[14] T.J. MacDonald, I.F. Pollack, H. Okada, S. Bhattacharya, J. Lyons-Weiler, Progression-associated genes in astrocytoma identified by novel microarray gene expression data reanalysis, Methods Mol. Biol. 377 (2007) 203–222.
[15] C. Liu, et al., A novel recurrence-associated metabolic prognostic model for risk stratification and therapeutic response prediction in patients with stage I lung adenocarcinoma, Cancer Biol Med 18 (2021) 734–749.
[16] N. Bi, et al., Identification of 40S ribosomal protein S8 as a novel biomarker for alcohol-associated hepatocellular carcinoma using weighted gene co-expression network analysis, Oncol. Rep. 44 (2020) 611–627.
[17] J.M. Park, et al., A case-control study in Taiwanese cohort and meta-analysis of serum ferritin in pancreatic cancer, Sci. Rep. 11 (2021) 21242.
[18] D. Yun, et al., A novel prognostic signature based on glioma Essential ferroptosis-related genes predicts clinical Outcomes and indicates treatment in glioma, Front. Oncol. 12 (2022) 897702.
[19] W. De Coster, S. D'Hert, D.T. Schultz, M. Cruts, Van Broeckhoven, C. NanoPack: visualizing and processing long-read sequencing data, Bioinformatics 34 (2018) 2666–2669.
[20] J.R. Wang, J. Holt, L. McMillan, C.D. Jones, FMLRC: Hybrid long read error correction using an FM-index, BMC Bioinf. 19 (2018) 50.
[21] H. Li, Minimap2: pairwise alignment for nucleotide sequences, Bioinformatics 34 (2018) 3094–3100.
[22] H. Li, et al., The sequence alignment/Map format and SAMtools, Bioinformatics 25 (2009) 2078–2079.
[23] M. Pertea, D. Kim, G.M. Pertea, J.T. Leek, S.L. Salzberg, Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown, Nat. Protoc. 11 (2016) 1650–1667.
[24] G. Pertea, M. Pertea, GFF Utilities: GffRead and GffCompare, F1000Res 9 (2020).
[25] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (2014) 550.
[26] X. Chen, et al., 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs, Nat. Cell Biol. 21 (2019) 978–990.
[27] P. Haruehanroengra, Y.Y. Zheng, Y. Zhou, Y. Huang, J. Sheng, RNA modifications and cancer, RNA Biol. 17 (2020) 1560–1575.
[28] X. Zhou, M. Zhao, Y. Fan, Y. Xu, Identification of a necroptosis-related gene signature for making clinical predictions of the survival of patients with lung adenocarcinoma, PeerJ 12 (2024) e16616.
[29] M. Wang, et al., Long non-coding RNA DANCR in cancer: roles, mechanisms, and implications, Front. Cell Dev. Biol. 9 (2021) 753706.
[30] S. Ghafouri-Fard, et al., A review on the role of DANCR in the carcinogenesis, Cancer Cell Int. 22 (2022) 194.
[31] H. Fan, et al., Long non-coding RNA CCDC144NL-AS1 sponges miR-143-3p and regulates MAP3K7 by acting as a competing endogenous RNA in gastric cancer, Cell Death Dis. 11 (2020) 521.
[32] G.I. Lambrou, K. Hatziagapiou, A. Zaravinos, The non-coding RNA GAS5 and its role in tumor therapy-induced resistance, Int. J. Mol. Sci. 21 (2020).
[33] G. Lin, T. Wu, X. Gao, Z. He, W. Nong, Research progress of long non-coding RNA GAS5 in malignant tumors, Front. Oncol. 12 (2022) 846497.
[34] K. Li, T. Yao, Y. Zhang, W. Li, Z. Wang, NEAT1 as a competing endogenous RNA in tumorigenesis of various cancers: role, mechanism and therapeutic potential, Int. J. Biol. Sci. 17 (2021) 3428–3440.
[35] J. Gu, et al., Molecular interactions of the long noncoding RNA NEAT1 in cancer, Cancers 14 (2022).
[36] B. Xu, et al., LncRNA SNHG3, a potential oncogene in human cancers, Cancer Cell Int. 20 (2020) 536.
[37] Z. Han, B. Yang, Y. Wang, X. Zeng, Z. Tian, Identification of expression patterns and potential prognostic significance of m(5)C-related regulators in head and neck squamous cell carcinoma, Front. Oncol. 11 (2021) 592107.