Am J Cancer Res 2023;13(11):5513-5530
www.ajcr.us /ISSN:2156-6976/ajcr0151603

Brief Communication

Transcriptomics data mining to uncover signature genes in head and neck squamous cell carcinoma: a bioinformatics analysis and RNA-sequencing based validation

Wenjie Yu[1,2], Xiaoling He[1,2], Chunming Zhang[1,2], Hui Huangfu[1,2]

[1]First School of Clinical Medcine, Shanxi Medical University, Taiyuan 030001, Shanxi, China; [2]Shanxi Key Laboratory of Otorhinolaryngology Head and Neck Cancer, First Hospital of Shanxi Medical University, Taiyuan 030001, Shanxi, China

Received June 6, 2023; Accepted September 8, 2023; Epub November 15, 2023; Published November 30, 2023

Abstract: Due to its heterogeneous nature, head and neck squamous cell carcinoma (HNSC) had the worst prognosis. Hence, there is an urgent need to develop novel diagnostic and prognostic models for effective disease management. A multi-layer dry-lab and wet-lab methodologies were adopted in the present study to identify novel diagnostic and prognostic biomarkers of HNSC. Initially, the GSE6631 gene microarray HNSC dataset was retrieved from the Gene Expression Omnibus (GEO) database. The R language-based "limma" package was employed to identify differentially expressed genes (DEGs) between HNSC and control samples. The Cytohubba plug-in software was used to identify the top four hub genes based on the degree score method. The Cancer Genome Atlas (TCGA) datasets, Gene Expression Omnibus (GEO) datasets, clinical HNSC tissue samples, HNSC cell line (FaDu), and normal cell line (HOK) were used to validate the expressions of hub genes. Moreover, additional bioinformatics analyses were performed to further evaluate the mechanisms of hub genes in the development of HNSC. In total, 1372 reliable DEGs were screened from the GSE6631 dataset. Out of these DEGs, only based on the four up-regulated hub genes, including UBE2C (Ubiquitin-conjugating enzyme E2C), BUB1B (BUB1 Mitotic Checkpoint Serine/Threonine Kinase B), MCM4 (Minichromosome Maintenance Complex Component 4), and KIF23 (Kinesin family member 23), we developed and validated a diagnostic and prognostic model for HNSC patients. Moreover, some interesting correlations observed between hub gene expression and infiltration level of immune cells may also improve our understanding of HNSC immunotherapy. In conclusion, we developed a novel diagnostic and prognostic model consisting of the UBE2C, BUB1B, MCM4, and KIF23 genes for HNSC patients. However, the efficiency of this model needs to be verified through more experimental studies.

Keywords: HNSC, immunotherapy, biomarker, hub gene

Introduction

Head and neck squamous cell carcinoma (HNSC) is the 6th most frequent cause of cancer-associated mortalities around the globe [1, 2]. HNSC denotes tumors that originate from the neck region, mouth, larynx, and nasopharynx [3]. HNSC is one of the most prevalent cancers and accounts for nearly 600,000 new cases and approximately 350,000 mortalities worldwide each year [4]. Due to significant advances over the past decade in the treatment methods of HNSC, which include surgery, radiation, and chemotherapy, treatment of this disease achieved good results [5, 6]. However, the survival rate of HNSC patients with advanced cancer stages is less than 34% [7, 8]. Therefore, the urgent discovery of some potential HNSC diagnostic and prognostic biomarkers is critical conducting effective diagnosis, treatment, improving prognosis, and reducing HNSC-associated mortality rate.

As it is well acknowledged by previous studies, the immune system is a crucial contributing factor to the development and progression of cancer [9-13]. The immunotherapy treatment method, which is primarily recognized by the programmed cell death 1 (PD-1) pathway, has highlighted a major breakthrough in the treat-

ment of many types of cancers [14-17]. Therefore, it is also critical to evaluate the associations of novel HNSC biomarkers with the infiltration levels of the immune cells to develop additional treatment options for HNSC patients.

Recent reports utilizing whole-genome sequencing (WES) technology have indicated an increasing volume of evidence supporting the ability of gene markers to detect the incidence and prognosis of HNSC patients. For instance, Shen et al. in their study have revealed a 7-gene-based diagnostic and prognostic system of potential biomarkers for the early detection and predicting the prognosis of HNSC patients [18]. Liu et al. identified 5 potential lncRNAs to construct a prognosis-predictive model for HNSC patients [19]. Similarly, She et al. introduced a new immune-associated genes-based prognosis predictive model for HNSC patients [20]. However, the limited sample sizes of HNSC cohorts in those studies and the utilization of different sequencing platforms may have led to doubtful results. The use of a multi-omics integrated strategy can greatly assist in avoiding these issues. This approach allows the use of multiple expression datasets consisting of large sample sizes for the exploration and validation of cancer signatures. While there have been numerous studies focused on developing single gene-based signatures for HNSC, there has been limited research on the development of robust multi-gene-based signatures using combined in silico, in vivo, and in vitro approaches. We believe that the creation of multi-gene-based signatures could be particularly effective in accurately predicting HNSC diagnosis and monitoring patient prognosis.

In this work, we aimed to identify differentially expressed genes (DEGs) and determine a few important hub genes between HNSC and normal tissue samples via a multi-layer dry-lab and wet-lab experimental approach.

Methodology

Data source and DEGs determination

The GEO database was thoroughly searched using the keyword “Head and Neck Cancer”. The selection criteria employed to identify a suitable HNSC dataset for hub gene selection were as follows: (i) studies involving any type of pharmacological manipulation were excluded; (ii) studies employing interfering molecules such as miRNAs, siRNAs, or any form of gene therapies were excluded; (iii) datasets encompassing knockdown cultures or artificially induced mutations were omitted; (iv) studies with a minimum of fifteen control and fifteen experimental samples were chosen; (v) studies exclusively conducted in Homo sapiens were preferred; and (vi) studies employing xenograft techniques were eliminated. In the end, all studies up until the conclusion of 2022 underwent individual examination and meticulous curation. A total of four microarray datasets, including both single- and dual-channel experiments, met the inclusion criteria. Based on enough sample size, the GSE6631 [21] dataset, containing 22 normal samples and 22 HNSC patient samples, was chosen as the experimental dataset. The standardized matrix data was obtained straight away from the GEO, and probes in this data were matched to the gene symbols based on the manufacturer’s annotation file that was provided with the data. If a single gene symbol was matched to multiple probes, the median value of the sequence was considered.

DEGs determination among HNSC and control tissue samples in the GSE6631 gene microarray dataset was carried out using the R language-based “limma” package [22, 23]. All the genes having “|LogFC| > 1.0, false discovery rate (FDR) < 0.05 and P < 0.05” were determined as the DEGs.

Construction and analysis of the protein-protein interaction (PPI) network and the selection of hub genes

Based on the DEGs obtained from the GSE6631, the analysis of interactions and construction of the PPI between DEGs was done by STRING [24], and the network type was a full STRING network, and the minimum required interaction score: medium confidence (0.400), and the max number of interactions to show: 1st shell was none/query proteins only, and 2nd shell was none. STRING database revealed protein interaction information and nodes notes, and then PPI networks were visualized using the Cytoscape 3.7.1 software for module and hub gene identification [25]. Firstly, the significant module was identified in the PPI network using the MCODE application of the Cytoscape,

then, based on the degree scores, the top four genes from the significant module were selected by Cytohubba application as potential hub genes for further validation and analysis [26].

UALCAN

Boxplots depicting the expression of hub genes across the HNSC TCGA dataset were plotted using expression values from the UALCAN database [27]. Furthermore, the UALCAN tool was employed to evaluate the expression of hub genes in the form of boxplots, across a range of clinical parameters associated with HNSC patients affected.

Verification of hub gene expression and survival analysis

To carry out hub genes expression validation analysis on additional HNSC GEO and TCGA cohorts, we assessed the expression values of the hub genes in HNSC and control samples via the GEO datasets (GSE65858 and GSE58911), GEPIA [28], OncoDb [29], and GENT2 [30] databases. Moreover, the GEPIA was also utilized to assess the survival outcomes of the hub genes in HNSC patients.

Development of hub genes-based prognostic model

To develop the prediction model, we utilized the least absolute shrinkage and selection operator (Lasso) and multivariate Cox proportional hazard regression analysis with the "survival" package in the R language [31]. The EMTAB-8588 dataset was used as the training dataset, while the GSE75538 dataset was used as the validation dataset. The prognostic model formula for HNSC patients' prognosis was calculated using the sum of the multivariate Cox regression coefficient variation of each mRNA.

Sub-cellular localization, methylation, mutational, and co-express gene analyses

To explore the sub-cellular localization of hub genes in HNSC cells, the Human Protein Atlas (HPA) database was utilized, while for methylation analysis, the MEXPRESS [32] and OncoDB [29] databases were used. Moreover, for conducting mutational and co-express gene analysis, the cBioPortal database [33] was considered in the present study.

Enrichment and immune cell infiltration analyses

To explore HNSC-related pathobiological mechanisms in more detail, the GSEA analysis [34] was performed to reveal hub gene-enriched GO and KEGG terms. Moreover, in this study, the abundance of three immune cells, including CD8+ T cells, CD4+ T cells, and macrophages, with respect to hub gene expression across the HNSC sample was calculated via the CIBERSORT database [35]. Additionally, GSE65858 and GSE58911 datasets consisted of 285 HNSC and 15 control samples were used to validate the expression levels of genes (CD4, CD8A, and CSF1R) encoding different immune cells, including CD8+ T cells, CD4+ T cells, and macrophages.

miRNA network and drug prediction analyses

In the current study, the miRNA network of the hub genes was constructed with the help of the ENCORI database [36]. We used the DrugBank database [37] to uncover a variety of drugs associated with the identified hub genes because we believe that the identified hub genes could be promising therapeutic targets.

Collection of clinical specimens and total RNA isolation

We acquired paired fresh cancer tissue specimens along with control samples from 10 patients who underwent surgical resection of HNSC at the First Hospital of Shanxi Medical University, Taiyuan, China between 2022 and 2023. None of the patients had received any chemotherapy or radiation therapy prior to the surgery. The collected tissue samples were promptly frozen in liquid nitrogen and stored at -80°C until DNA and RNA isolation. The study received ethical approval in accordance with the Helsinki Declaration and informed written consent was obtained from all participants.

Total RNA from the collected specimens was extracted using the TRIzol method [38]. NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) was employed to assess the concentration and purity of the extracted RNA, ensuring that the A260/A280 ratio fell within the range of 1.8 to 2.0.

Real time quantitative PCR (RT-qPCR)

The RNA that was isolated underwent conversion into cDNA utilizing the Prime-Script RT reagent kit (TaKaRa, Dalian, China). Subsequently, RT-qPCR analysis was carried out on an ABI 7500 Real-Time PCR System (Applied Biosystems, USA) employing the SYBR Premix Ex Taq TM II kit (TaKaRa). The expression levels were standardized against β-actin (ACTB). Each experiment was independently performed in triplicate. To evaluate the relative expression of each hub gene [39], the 2^(-ΔΔCt) method was utilized. The following primer pairs were employed for the amplification of both hub and control genes.

ACTB-F 5'-CACCATTGGCAATGAGCGGTTC-3', ACTB-R 5'-AGGTCTTTGCGGATGTCCACGT-3'; UBE2C-F TGATGTAGCGGAGCAGTGGAAG-3', UBE2C-R 5'-GGAGGAAGTCTTGGCAGAACAG-3'; BUB1B-F 5'-GTGGAAGAGACTGCACAACAGC-3', BUB1B-R 5'-TCAGACGCTTGCTGATGGCTCT-3'; MCM4-F 5'-TGTTTTCCAGCCCTCCCCAAATG-3', MCM4-R 5'-GAGTGCCGTATGTCAGTGGTGAAC-3'; KIF23-F 5'-CAGATTTCCAACGGCCAGCA-3', KIF23-R 5'-TCATGGCTTTTTGCGCTTGG-3'.

RNA sequencing-based hub gene expression analysis

A total of one HNSC cell line, including FaDu, and one normal human oral keratinocyte (HOK) cell line were purchased from the ATCC (American Type Culture Collection). The purchased cell lines were cultured in DMEM (HyClone), supplemented with 10% fetal bovine serum (FBS; TBD), 1% glutamine, and 1% penicillin-streptomycin in 5% CO_2 at 37°C. Total RNA extraction from these two cells lines was done using the TRIzol method [38], and the RNA samples were sent to the Beijing Genomics Institute (BGI) company for RNA-sequencing analysis.

After RNA-seq analysis, the gene expression values of the hub genes were normalized using fragments per kilo base million reads (FPKM). The obtained FPKM values against hub genes in HNSC (FaDu) and normal oral keratinocyte (HOK) cell line were compared to identify differences in the expression levels.

Immunohistochemistry (IHC)

Immunohistochemistry (IHC) was conducted with slight modifications as previously outlined [40]. In brief, sections of paraffin-embedded tissues were deparaffinized using xylene and subsequently rehydrated via a graded ethanol series. Following antigen retrieval, inactivation of endogenous peroxidase, and blocking with normal goat serum, the sections were subjected to an overnight incubation at 4°C with anti-UBE2C (CAT#: TA329811), anti-BUB1B (CAT#: CF500533), anti-MCM4 (CAT#: TA321984), and anti-KIF23 (CAT#: TA334708) antibodies. Subsequently, a dextran carrying anti-rabbit IgG conjugated to horseradish peroxidase (HRP) was added, and positive staining was visualized using the Dako REAL EnVision detection system. Images of the stained sections were acquired using an Olympus CX31 digital microscope (Olympus, Japan) for the quantification of stained cells. IHC staining was evaluated by taking into account the intensity of staining.

Statistics details for in silico analyses

DEGs were identified using a t-test [41]. For GO and KEGG enrichment analysis, we used Fisher's Exact test for computing statistical difference [42]. Correlational analyses were carried out using the Pearson method. For comparisons, a student t-test was adopted in the current study. All the analyses were carried out in R version 3.6.3 software.

Results

DEGs screening

After a careful evaluation of HNSC and normal samples' expression profiles in the GSE6631 HNSC gene microarray dataset (**Figure 1A**) and a comprehensive analysis using defined parameters (details are given in the method section), 1372 DEGs were screened out between HNSC and normal samples (**Figure 1B**).

Determination of the hub genes

The top 250 DEGs with the lowest *p*-values were used in the present study for hub gene selection. To do so, initially, the PPI network of the selected DEGs was constructed using the STRING database. After removing disconnected nodes, the constructed PPI had 249 nodes and 776 edges (**Figure 1C**). Then, the constructed PPI was analyzed via the MCODE application for identifying the significant module. The identified module was enriched in 24 DEGs (**Figure 1D** and **1E**). Finally, the identified module was undertaken the Cytohubba analysis for

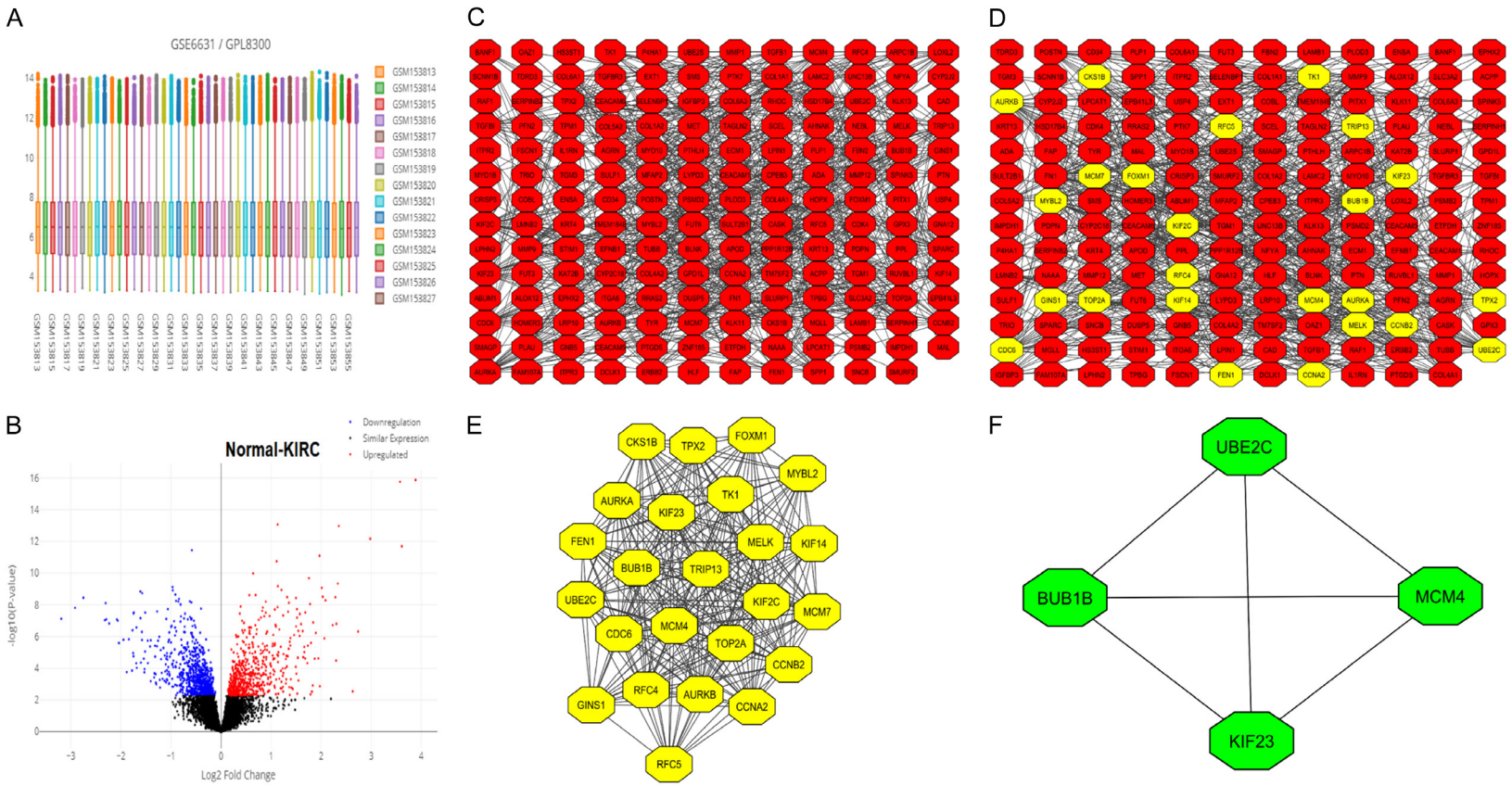


Figure 1. A comparison between expression profiles of samples, and volcano graph of DEGs, a PPI network of the top 250 DEGs, a significant module in the constructed PPI network, and a PPI network of the identified hub genes in GSE6631 microarray dataset. (A) A comparison between expression profiles of samples in GSE6631 microarray dataset, (B) A volcano graph of the DEGs observed in GSE6631 microarray dataset, (C) A PPI network of the top 250 DEGs in GSE6631 microarray dataset, (D, E) A PPI network of the most significant module, and (F) A PPI network of identified four hub genes.

hub genes selection. The top 4 genes with the highest degree scores were selected as hub genes (**Figure 1F**). The hub genes included up-regulated UBE2C (Ubiquitin-conjugating enzyme E2C), BUB1B (BUB1 Mitotic Checkpoint Serine/Threonine Kinase B), MCM4 (Minichromosome Maintenance Complex Component 4), and KIF23 (Kinesin family member 23) (**Figure 1F**).

Expression analysis of hub genes in The Cancer Genome Atlas

To perform expression analysis of hub genes in TCGA, the mRNA and protein expression data of the hub genes across HNSC and control samples were obtained from the ULACAN database (**Figure 2**). Based on the obtained data, it was observed that expressions of the UBE2C, BUB1B, MCM4, and KIF23 genes at both mRNA and protein levels were significantly up-regulated in terms of *p*-value in HNSC samples compared to the controls (**Figure 2A-C**). By looking at the significant overexpression of the hub genes, we further analyzed UBE2C, BUB1B, MCM4, and KIF23 in HNSC patients with diverse clinical parameters to confirm the expression status of those genes. All selected hub genes were also overexpressed in HNSC patients with different clinical variables relative to controls (**Figure 2D-G**).

Validation and survival analysis results

Expression validation analysis of the hub genes on additional GEO and TCGA HNSC datasets was carried out using the GEO, GEPIA, OncoDB, and GENT2 databases for more reliable analytic outcomes. Collective results from the analyzed GEO and TCGA HNSC datasets showed that the expressions of UBE2C, BUB1B, MCM4, and KIF23 were notably higher across HNSC tissues compared to normal tissues (**Figure 3A-D**). Moreover, via the GEPIA database, we also revealed that the higher expressions of the UBE2C, BUB1B, MCM4, and KIF23 were associated with the negative prognosis in HNSC patients, with UBE2C and KIF23 hub genes having the greater prognostic values (**Figure 3E**).

Development of hub genes-based prognostic model

To analyze the prognostic model that is based on UBE2C, BUB1B, MCM4, and KIF23 genes, we used the EMTAB-8588 dataset as our training dataset and the GSE75538 dataset as our validation dataset. Our prognostic model was constructed using a stepwise Cox regression model that incorporated the parameters of the hazard ratio, c-index, and risk score. Through evaluating our predictive prognostic model using the c-index, we determined that it effectively and accurately assessed the prognosis of HNSC patients, which is demonstrated in Supplementary Figure 1.

Sub-cellular localization, methylation, mutational, and co-express gene analysis results

The HPA database was used in the present study to evaluate the sub-cellular localization of UBE2C, BUB1B, MCM4, and KIF23 in HNSC cells. The protein encoded by the UBE2C gene was localized in the plasma membrane and cytosol (Supplementary Figure 2A). The localization of the protein encoded by the BUB1B gene was seen in the cytosol (Supplementary Figure 2B). For a protein encoded by MCM4, the localization was found in the nucleoplasm (Supplementary Figure 2C) and for a protein coded by the KIF23 gene, the localization was in the nucleoplasm, midbody ring, and the mitosis spindle (Supplementary Figure 2D).

Hub genes' promoter methylation level was explored next using the MEXPRESS and OncoDB databases. As shown in Supplementary Figure 3A and 3B, relative to the normal tissue samples, the promoter methylation levels of UBE2C, BUB1B, MCM4, and KIF23 were notably lowered in the HNSC samples. Therefore, we speculate that the expressions of UBE2C, BUB1B, MCM4, and KIF23 hub genes were negatively correlated with their promoter methylation levels in HNSC tissue samples.

To find the potential roles of genetic alterations in the UBE2C, BUB1B, MCM4, and KIF23 dysregulations, we further focused on their genetic alterations via the cBioPortal database. We selected "TCGA, Firehose Legacy (530 cases)" datasets for exploring UBE2C, BUB1B, MCM4, and KIF23 for analysis. As a result, we observed slight frequencies of genetic alterations in UBE2C (0.6%), BUB1B (1.2%), MCM4 (7%), and KIF23 (1.2%) in the HNSC samples of the utilized TCGA dataset (**Figure 4**). As shown in **Figure 4A**, deep amplification was the most frequent form of genetic alteration in the UBE2C and MCM4 genes, while missense mutation

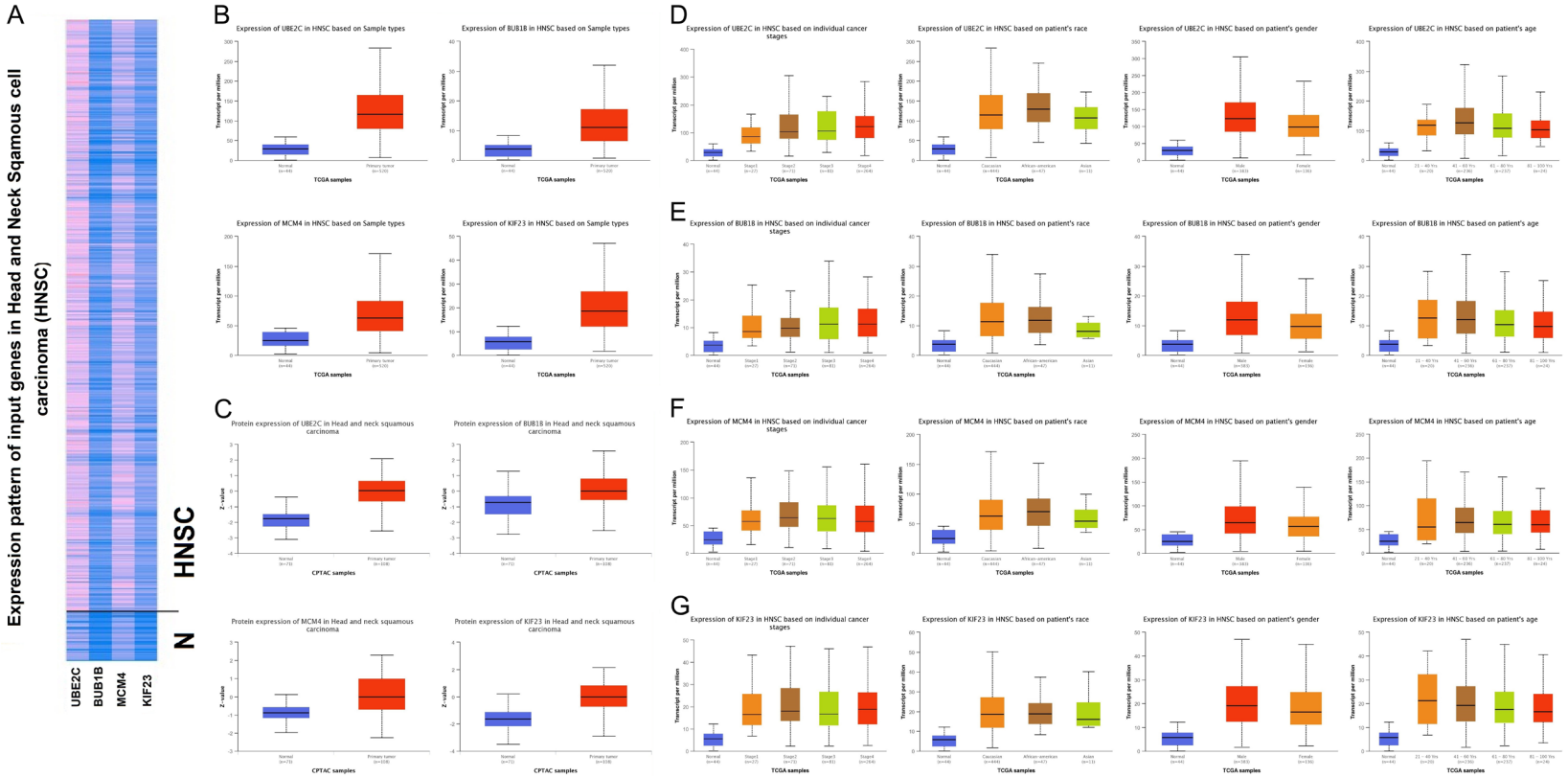


Figure 2. mRNA and protein expression profiling of UBE2C, BUB1B, MCM4, and KIF23 in HNSC and normal samples across TCGA dataset, and expression profiling of UBE2C, BUB1B, MCM4, and KIF23 in HNSC samples of different clinical variables relative to controls via UALCAN. (A) A heatmap of UBE2C, BUB1B, MCM4, and KIF23 hub gene mRNA expression in HNSC sample group and normal control group, (B) Box plot presentation of UBE2C, BUB1B, MCM4, and KIF23 hub gene mRNA expression in HNSC sample group and normal control group, (C) Box plot presentation of UBE2C, BUB1B, MCM4, and KIF23 hub gene protein expression in HNSC sample group and normal control group, (D) mRNA expression profiling of UBE2C in HNSC samples of different clinical variables and normal controls, (E) mRNA xpression profiling of BUB1B in HNSC samples of different clinical variables and normal controls, (F) mRNA expression profiling of MCM4 in HNSC samples of different clinical variables and normal controls, and (G) mRNA expression profiling of KIF23 in HNSC samples of different clinical variables and normal controls.

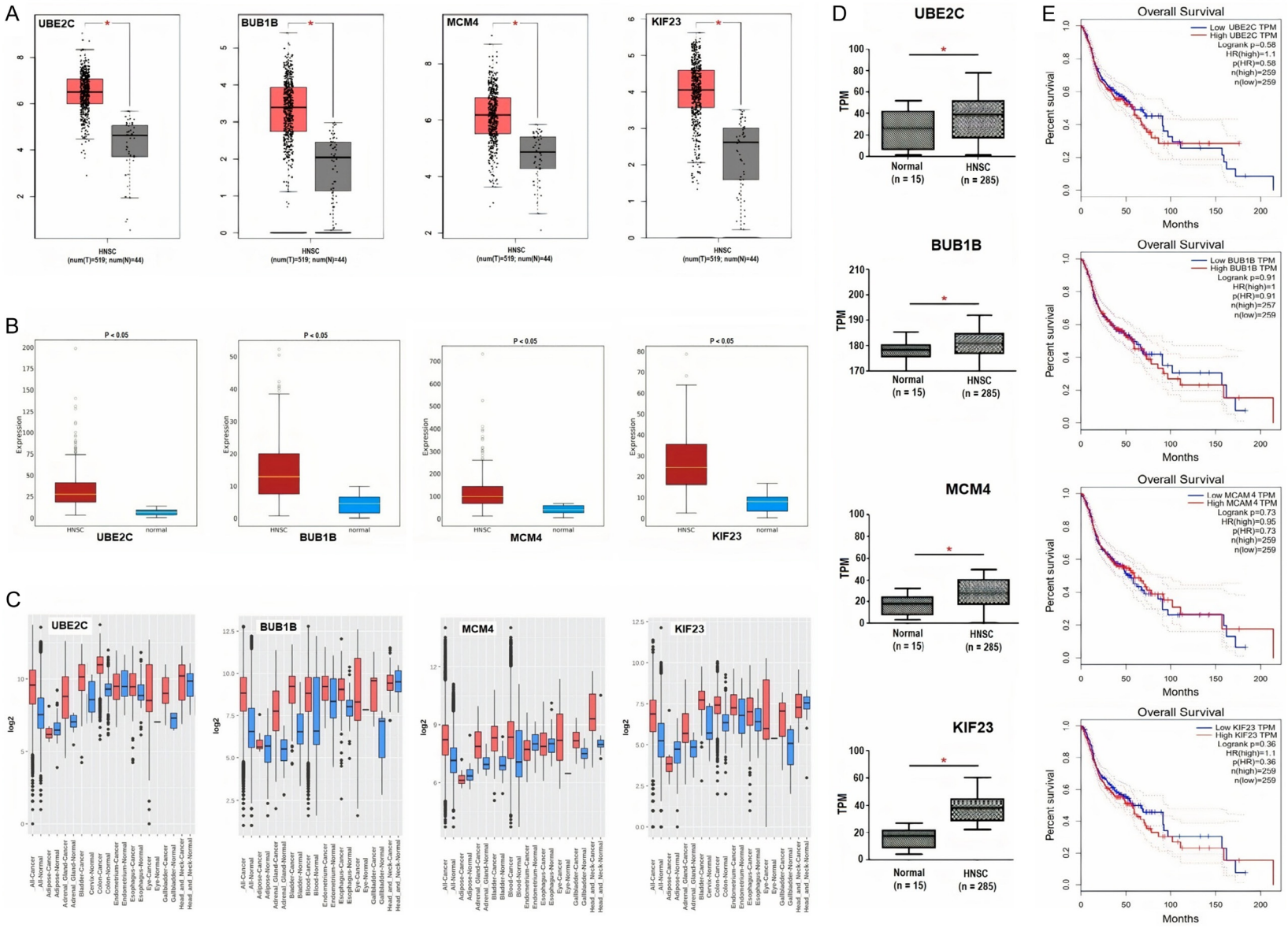


Figure 3. Expression validation and survival analysis of UBE2C, BUB1B, MCM4, and KIF23. (A) Expression validation of UBE2C, BUB1B, MCM4, and KIF23 in HNSC and normal samples via GEPIA database, (B) Expression validation of UBE2C, BUB1B, MCM4, and KIF23 in HNSC and normal samples via OncoDB database, (C) Expression validation of UBE2C, BUB1B, MCM4, and KIF23 via GENT2 database, (D) Expression validation of UBE2C, BUB1B, MCM4, and KIF23 via GEO database using GSE65858 and GSE58911 datasets, and (E) Survival analysis of UBE2C, BUB1B, MCM4, and KIF23 in HNSC and normal samples via GEPIA database.

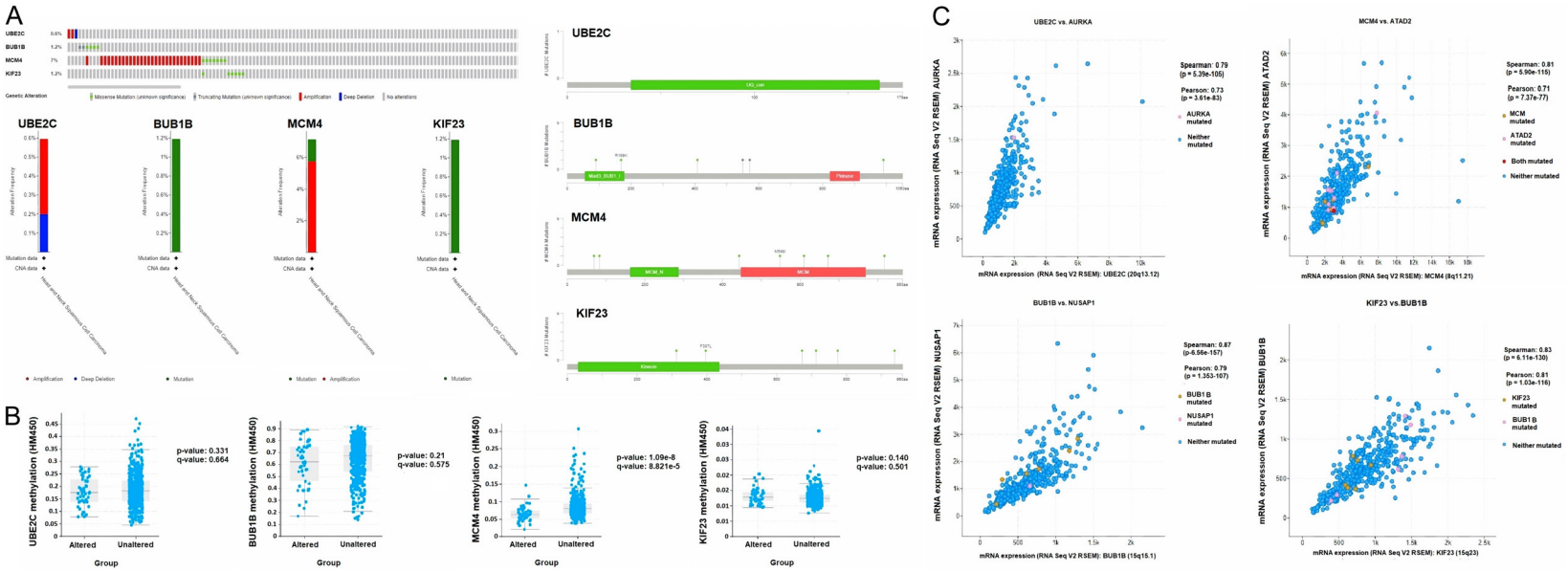


Figure 4. Exploration of genetic alteration frequencies, co-expressed genes, and methylatiom analysis of UBE2C, BUB1B, MCM4, and KIF23 in HNSC samples via cBioPortal. (A) Types, frequencies, and location of the genetic alterations in UBE2C, BUB1B, MCM4, and KIF23, (B) Identification of co-expressed genes with UBE2C, BUB1B, MCM4, and KIF23 in HNSC samples, and (C) Promoter methylation analysis of UBE2C, BUB1B, MCM4, and KIF23 in HNSC samples.

was the most frequent form of genetic alteration in the BUB1B and KIF23 genes. We also noticed that missense mutations were majorly concentrated in the Mad3-BuB1 domain of the BUB1B, the MCM domain of the MCM4, and the Kinesin domain of the KIF23 gene (**Figure 4A**). Further analysis showed that in the HNSC patient group with genetically altered hub genes, the promoter methylation levels of the UBE2C, BUB1B, MCM4, and KIF23 genes were lowered as compared to the unaltered group (**Figure 4B**). In addition to this, co-expression analysis of the UBE2C, BUB1B, MCM4, and KIF23 genes revealed that UBE2C vs. AURKA, BUB1B vs. NUSAP1, MCM4 vs. ATAD2, and KIF23 vs. BUB1B are the highly co-expressed genes along with hub genes in HNSC samples (**Figure 4C**).

Enrichment and immune cell infiltration analysis

GO analysis of UBE2C, BUB1B, MCM4, and KIF23 hub genes indicated that those genes were involved in the "germinal vesicle, centralspindlin, female germ cell nucleus, chromosome passenger complex" CC terms (Supplementary Figure 4A), "histone serine kinase activity, DNA replication origin binding, ubiquitin conjugating enzyme activity, ubiquitin-like protein conjugating enzyme activity" MF terms (Supplementary Figure 4B), "anaphase-promoting complex dependent catabolic process, Reg. of mitotic nuclear division, mitotic sister chromatid segregation, sister chromatid segregation" BP terms (Supplementary Figure 4C). KEGG analysis of UBE2C, BUB1B, MCM4, and KIF23 showed that these genes were dominantly involved in the "DNA replication, cell cycle, progesterone-mediated oocyte maturation, oocyte meiosis, ubiqutin mediated proteolysis, and MicroRNAs in cancer" pathways (Supplementary Figure 4D).

To further highlight the oncogenic roles of the UBE2C, BUB1B, MCM4, and KIF23 hub genes in HNSC, the CIBERSORT algorithm was utilized in our work to investigate immunocytotic infiltration. The CIBERSORT algorithm highlighted some interesting correlations among UBE2C, BUB1B, MCM4, and KIF23 expressions and immunocytotic infiltration. The expressions of UBE2C, BUB1B, MCM4, and KIF23 hub genes were significantly correlated with the abundant infiltration of CD8+ T cells, CD4+ T cells, and macrophages across HNSC (Supplementary Figure 5A-D). To further validate this scenario, we conducted an analysis of gene expression levels in two GEO datasets, GSE65858 and GSE58911; comprising a total of 285 HNSC samples and 15 control samples. Specifically, we examined the expression of genes (CD4, CD8A, and CSF1R), responsible for encoding CD8+ T cells, CD4+ T cells, and macrophages, respectively. The results from this analysis clearly demonstrated a significant up-regulation of CD4, CD8A, and CSF1R genes within the HNSC sample group when compared to the control group (Supplementary Figure 5E). These findings provide robust confirmation that the heightened expression of UBE2C, BUB1B, MCM4, and KIF23 hub genes in HNSC patients is closely associated with the increased expression of CD4, CD8A, and CSF1R genes, which in turn are responsible for higher production of CD8+ T cells, CD4+ T cells, and macrophages immune cells.

miRNA network of the hub genes

Via ENCORI and Cytoscape, we constructed the lncRNA-miRNA-mRNA co-regulatory networks of the UBE2C, BUB1B, MCM4, and KIF23. In the constructed networks, the total counts of lncRNAs, miRNAs, and mRNAs were 78, 151, and 4, respectively (Supplementary Figure 6). Based on the constructed networks, we have identified one miRNA (hsa-mir-16-5p) that targets all hub genes simultaneously (Supplementary Figure 6). Therefore, we speculate that the identified lncRNAs, has-mir-16-5p, and hub genes (UBE2C, BUB1B, MCM4, and KIF23) (Supplementary Figure 6) as an axis, might also be the potential inducers of the HNSC.

Drug prediction analysis of the hub genes

For patients suffering from HNSC, medical treatment is the first option for treatment. Therefore, a selection of appropriate candidate drugs is required. In the current study, via the DrugBank database, we explored some potential drugs that can reverse the gene expressions of identified hub genes for the treatment of HNSC. We noted that cyclosporine along with many other drugs are the negative expression regulators of UBE2C, BUB1B, MCM4, and KIF23 mRNA expressions (**Table 1**).

Table 1. DrugBank-based hub genes-associated drugs

Sr. No	Hub gene	Drug name	Effect	Reference	Group
1	UBE2C	Amsacrine	Decrease expression of UBE2C mRNA	A20666	Approved
		Cyclosporine		A21092	
		Dronabinol		A22083	
2	BUB1B	Dasatinib	Decrease expression of BUB1B mRNA	A21899	Approved
		Cyclosporine		A21092	Approved
3	MCM4	Cyclosporine	Decrease expression of MCM4 mRNA	A21092	Approved
		Calcitriol		A22301	
4	KIF23	Cyclosporine	Decrease expression of KIF23 mRNA	A21092	Approved
		Estradiol		A21424	

Hub gene expression validation via RT-qPCR, RNA sequencing, and Immunohistochemistry analyses using clinical samples and cell lines of HNSC patients

In the present study, we conducted RT-qPCR analysis of four hub genes (UBE2C, BUB1B, MCM4, and KIF23) utilizing paired samples consisting of 10 cases of HNSC and their corresponding control tissues. Our findings revealed a significant ($P < 0.05$) up-regulation of these genes in the HNSC tissue samples compared to their respective controls (**Figure 5A**).

Moreover, using expression data obtained from RNA-seq analysis of the HNSC cell line (FaDu) and one normal human oral keratinocyte (HOK) cell line, the expression levels of identified four hub genes (BUB1B, MCM4, and KIF23) were further analyzed. As depicted in the figure, we observed a conspicuous increase in FPKM values for UBE2C, BUB1B, MCM4, and KIF23 within HNSC cell lines (FaDu) when compared to their levels in normal control cell line (HOK). This observation strongly indicates an up-regulation of these hub genes in the FaDu cell line in contrast to the normal control cell line (HOK) (**Figure 5B**).

Next, we conducted an IHC analysis to examine the expression of proteins encoded by the hub genes UBE2C, BUB1B, MCM4, and KIF23 in one selective HNSC tissue sample paired with its respective control. Our assessment of the IHC images revealed a significant disparity in the staining intensity between the two analyzed samples. Specifically, we observed a pronounced increase in the levels of UBE2C, BUB1B, MCM4, and KIF23 proteins within the HNSC tissue samples compared to the control tissues (**Figure 5C**). These findings strongly indicate an up-regulation of these proteins in the context of HNSC, highlighting their potential involvement in the development or progression of this cancer.

Discussion

Due to its heterogeneous nature, HNSC has the worst prognosis, and the prevalence of this disease is steadily increasing around the globe with each passing year [43]. Moreover, the five-year survival rate of HNSC patients is less than 40% [7]. With advancements in genome-sequencing technology, expression profiling of cancer-causing genes has grabbed huge attention across the world and has thus been used to investigate the potential diagnostic and prognostic biomarkers of cancer patients [44-48]. Taking into account the high prevalence of HNSC cases, the investigation of molecular biomarkers that may be fully associated with tumor development and assist clinicians in treating HNSC by providing some novel therapeutic targets is urgently needed.

In this work, we initially screened out DEGs between HNSC and adjacent non-tumor tissue samples present in the GSE6631 [21] gene microarray dataset. Then, we processed DEGs for hub gene determination analysis, and the outcomes of this analysis revealed four hub genes, including UBE2C, BUB1B, MCM4, and KIF23, between HNSC and normal samples. Later on, via in silico and in vitro analyses, we confirmed the significant up-regulation of UBE2C, BUB1B, MCM4, and KIF23 in TCGA and cell line (HOK and FaDu) samples of HNSC patients compared to controls. Higher expressions of UBE2C, BUB1B, MCM4, and KIF23 were significantly associated with the poor sur-

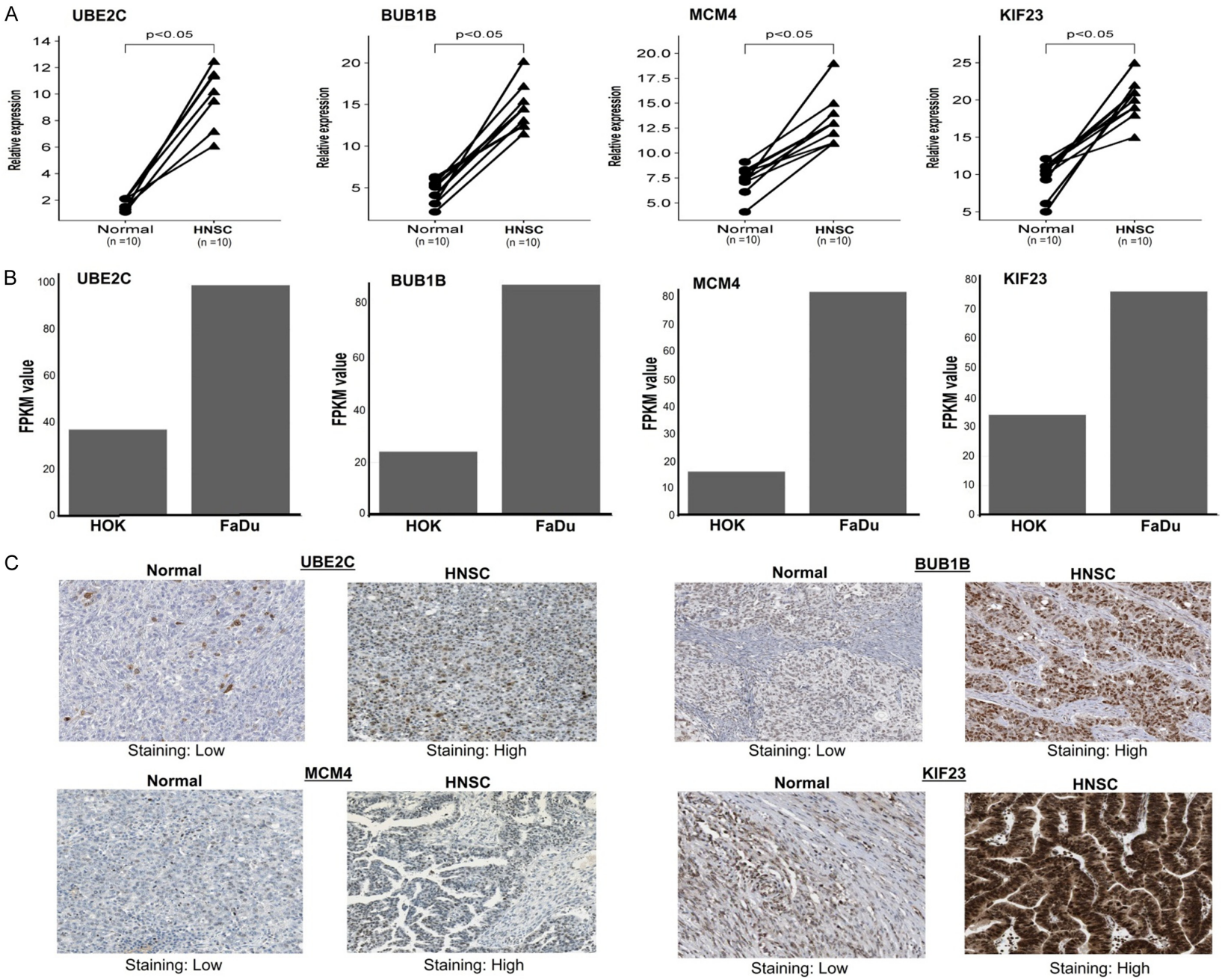

A
UBE2C
BUB1B
MCM4
KIF23
p<0.05
Relative expression
Normal
(n =10)
HNSC
(n =10)
B
UBE2C
BUB1B
MCM4
KIF23
FPKM value
HOK
FaDu
C
Normal
HNSC
UBE2C
BUB1B
MCM4
KIF23
Staining: Low
Staining: High

Figure 5. Validating UBE2C, BUB1B, MCM4, and KIF23 mRNA and protein expression on clinical samples and cell lines via RT-qPCR, RNA sequencing, and Immunohistochemistry. (A) RT-qPCR analysis results of UBE2C, BUB1B, MCM4, and KIF23 across clinical HNSC samples paired with controls, (B) RNA-seq analysis results of UBE2C, BUB1B, MCM4, and KIF23 across HOK and FaDu cell lines, and (C) IHC analysis results of UBE2C, BUB1B, MCM4, and KIF23 across HNSC sample paired with control.

vival of HNSC patients and hypomethylation of their promoter regions.

UBE2C is a member of the E2 family of proteins and performs a variety of roles in mitotic cyclin destruction and the progression of the cell cycle [49]. The dysregulation of UBE2C was earlier implicated in the development and progression of a variety of cancers [50, 51]. For example, the higher expression of UBE2C genes was previously reported as an oncogenic factor in colon cancer, ovarian cancer, breast cancer, liver cancer, and esophageal cancer [52-57]. Moreover, the overexpression of UBE2C was declared as a potential biomarker in brain tumor patients and found to be associated with poor prognosis [58, 59]. In addition to this, elevated UBE2C expression was also noticed in high-grade astrocytomas [59].

BUB1B protein belongs to the SAC family of proteins [60, 61]. BUB1B protein is the critical part of the mitotic checkpoint complex that consists of Cdc20 protein and different other SAC proteins, such as Mad2 and Bub3 [62, 63]. It was earlier revealed that low expression of the BUB1B gene participates in the development and progression of different human cancers, including colon and lung cancers [64, 65]. However, the latest body of data highlighted that the higher expression of the BUB1B gene also participates in the development of a variety of other human cancers, including gastric cancer [66], breast cancer [67], bladder cancer [68], hepatocellular carcinoma [69], esophageal squamous cell carcinoma [70], and some other types of cancer. The relationship between dysregulation of the BUB1B gene and the development and progression of HNSC has not yet been established.

MCM4 protein belongs to the MCM family of proteins [71, 72]. The overexpression of the MCM4 gene is largely associated with the development and progression of esophageal cancer [73]. Therefore, this gene may serve as a potential biomarker for patients with esophageal cancer [73]. In addition to this, higher MCM4 expression is also reported to be an oncogenic factor in patients with melanoma [74]. Overexpression of MCM4 in lung cancer patients was associated with metastasis and higher clinical grades, therefore, it may be utilized as a novel molecular biomarker and therapeutic target in patients with lung cancer [75]. MCM4 gene-based knockout experiments revealed that MCM4 down-regulation inhibited cell growth and invasion in cell lines of lung cancer patients [76].

KIF23 protein belongs to the family of kinesin proteins [77, 78]. A recent body of data suggested that KIF23 dysregulation is an oncogenic factor of various cancers. For example, overexpression of KIF23 could stimulate cell proliferation in gastric cancer patients by activating a few diverse pathways, including the Wnt/β-catenin signaling pathway [79]. In patients with lung cancer, KIF23 overexpression was associated with poor prognosis [80, 81]. In patients with ovarian cancer, KIF23 overexpression was revealed as an independent diagnostic signature [82]. The higher expression of KIF23 was also correlated with the worst survival in patients with breast cancer [83]. The role of overexpressed KIF23 in the development and progression of HNSC is not clear yet.

Previous studies reported that the immune system plays a vital role in controlling tumor growth and metastasis [84]. As a promising anti-cancer treatment option, immunotherapy has highlighted a huge therapeutic potential across different cancers [85]. In HNSC patients, a higher infiltration level of CD8+ T immune cells was correlated with a better prognosis [85]. In our work, we noticed that expressions of UBE2C, BUB1B, MCM4, and KIF23 hub genes were significantly correlated with the abundant infiltration of CD8+ T cells, CD4+ T cells, and macrophages across the HNSC samples. This valuable information may provide novel ideas for designing immunotherapy treatment for HNSC patients.

KEGG analysis of UBE2C, BUB1B, MCM4, and KIF23 showed that these genes were dominantly involved in the “DNA replication, cell

cycle, progesterone-mediated oocyte maturation, oocyte meiosis, ubiquitin mediated proteolysis, and MicroRNAs in cancer" pathways. The dysregulation of these pathways is already implicated in cancer development by previously published studies [86-88]. We further noticed that UBE2C, BUB1B, MCM4, and KIF23 hub genes' expression were regulated simultaneously by hsa-mir-16-5p miRNA in HNSC patients. Previously, the dysregulation of hsa-mir-16-5p in multiple human cancers has been reported in published studies, for example in breast cancer, bladder cancer, glioblastoma, and lung cancer [89, 90]. However, any tumor suppressor or tumor-causing role of hsa-mir-16-5p in HNSC is not reported anywhere.

Conclusion

We established a novel model of the HNSC biomarkers consisting of UBE2C, BUB1B, MCM4, and KIF23. Significant overexpression of these genes was closely associated with the development and progression of HNSC in the GEO datasets, TCGA datasets, clinical tissue samples, and cell lines. However, additional functional studies are needed to further confirm the biomarker roles of these genes in HNSC patients.

Disclosure of conflict of interest

None.

Address correspondence to: Hui Huangfu, Shanxi Key Laboratory of Otorhinolaryngology Head and Neck Cancer, First Hospital of Shanxi Medical University, Taiyuan 030001, Shanxi, China. E-mail: Huangfuhuisx@126.com

References

[1] Geng X, Zhang Y, Zeng Z, Zhu Z, Wang H, Yu W and Li Q. Molecular characteristics, prognostic value, and immune characteristics of m6A regulators identified in head and neck squamous cell carcinoma. Front Oncol 2021; 11: 629718.

[2] Weber BZC, Agca S, Domaniku A, Bilgic SN, Arabaci DH and Kir S. Inhibition of epidermal growth factor receptor suppresses parathyroid hormone-related protein expression in tumours and ameliorates cancer-associated cachexia. J Cachexia Sarcopenia Muscle 2022; 13: 1582-1594.

[3] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018; 68: 7-30.

[4] Zheng H, Liu H, Lu Y and Li H. Identification of a novel signature predicting overall survival in head and neck squamous cell carcinoma. Front Surg 2021; 8: 717084.

[5] Kim MM and Califano JA. Molecular pathology of head-and-neck cancer. Int J Cancer 2004; 112: 545-553.

[6] Kowalski LP, Carvalho AL, Martins Priante AV and Magrin J. Predictive factors for distant metastasis from oral and oropharyngeal squamous cell carcinoma. Oral Oncol 2005; 41: 534-541.

[7] Leemans CR, Braakhuis BJ and Brakenhoff RH. The molecular biology of head and neck cancer. Nat Rev Cancer 2011; 11: 9-22.

[8] Gildener-Leapman N, Ferris RL and Bauman JE. Promising systemic immunotherapies in head and neck squamous cell carcinoma. Oral Oncol 2013; 49: 1089-1096.

[9] Karin M and Greten FR. NF-κB: linking inflammation and immunity to cancer development and progression. Nat Rev Immunol 2005; 5: 749-759.

[10] Thomas S, Izard J, Walsh E, Batich K, Chongsathidkiet P, Clarke G, Sela DA, Muller AJ, Mullin JM, Albert K, Gilligan JP, DiGuilio K, Dilbarova R, Alexander W and Prendergast GC. The host microbiome regulates and maintains human health: a primer and perspective for non-microbiologists. Cancer Res 2017; 77: 1783-1812.

[11] Usman M, Okla MK, Asif HM, AbdElgayed G, Muccee F, Ghazanfar S, Ahmad M, Iqbal MJ, Sahar AM, Khaliq G, Shoaib R, Zaheer H and Hameed Y. A pan-cancer analysis of GINS complex subunit 4 to identify its potential role as a biomarker in multiple human cancers. Am J Cancer Res 2022; 12: 986-1008.

[12] Usman M, Hameed Y, Ahmad M, Rehman JU, Ahmed H, Hussain MS, Asif R, Murtaza MG, Jawad MT and Iqbal MJ. Breast cancer risk and human papillomavirus infection: a Bradford Hill criteria based evaluation. Infect Disord Drug Targets 2022; 22: e200122200389.

[13] Hameed Y, Usman M, Liang S and Ejaz S. Novel diagnostic and prognostic biomarkers of colorectal cancer: capable to overcome the heterogeneity-specific barrier and valid for global applications. PLoS One 2021; 16: e0256020.

[14] Ohaegbulam KC, Assal A, Lazar-Molnar E, Yao Y and Zang X. Human cancer immunotherapy with antibodies to the PD-1 and PD-L1 pathway. Trends Mol Med 2015; 21: 24-33.

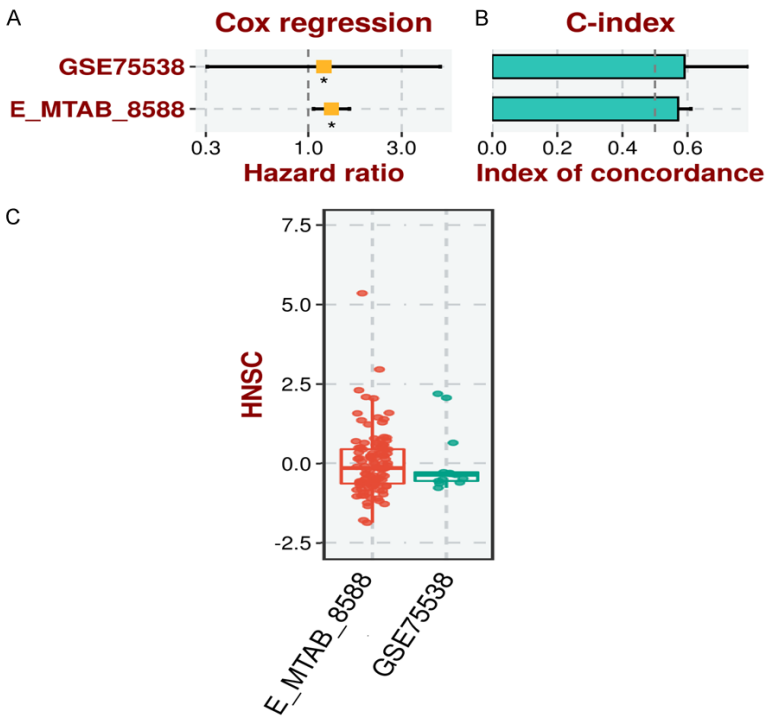
[15] Lin W, Chen M, Hong L, Zhao H and Chen Q. Crosstalk between PD-1/PD-L1 blockade and its combinatorial therapies in tumor immune microenvironment: a focus on HNSCC. Front Oncol 2018; 8: 532.

[16] Ahmad M, Hameed Y, Khan M, Usman M, Rehman A, Abid U, Asif R, Ahmed H, Hussain MS, Rehman JU, Asif HM, Arshad R, Atif M, Hadi A, Sarfraz U and Khurshid U. Up-regulation of GINS1 highlighted a good diagnostic and prognostic potential of survival in three different subtypes of human cancer. Braz J Biol 2021; 84: e250575.
[17] Sial N, Saeed S, Ahmad M, Hameed Y, Rehman A, Abbas M, Asif R, Ahmed H, Hussain MS, Rehman JU, Atif M and Khan MR. Multi-omics analysis identified TMED2 as a shared potential biomarker in six subtypes of human cancer. Int J Gen Med 2021; 14: 7025-7042.
[18] Shen S, Bai J, Wei Y, Wang G, Li Q, Zhang R, Duan W, Yang S, Du M, Zhao Y, Christiani DC and Chen F. A seven-gene prognostic signature for rapid determination of head and neck squamous cell carcinoma survival. Oncol Rep 2017; 38: 3403-3411.
[19] Liu G, Zheng J, Zhuang L, Lv Y, Zhu G, Pi L, Wang J, Chen C, Li Z, Liu J, Chen L, Cai G and Zhang X. A prognostic 5-lncRNA expression signature for head and neck squamous cell carcinoma. Sci Rep 2018; 8: 15250.
[20] She Y, Kong X, Ge Y, Yin P, Liu Z, Chen J, Gao F and Fang S. Immune-related gene signature for predicting the prognosis of head and neck squamous cell carcinoma. Cancer Cell Int 2020; 20: 22.
[21] Kuriakose MA, Chen WT, He ZM, Sikora AG, Zhang P, Zhang ZY, Qiu WL, Hsu DF, McMunn-Coffran C, Brown SM, Elango EM, Delacure MD and Chen FA. Selection and validation of differentially expressed genes in head and neck cancer. Cell Mol Life Sci 2004; 61: 1372-1383.
[22] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43: e47.
[23] Hameed Y, Usman M and Ahmad M. Does mouse mammary tumor-like virus cause human breast cancer? Applying Bradford Hill criteria postulates. Bull Natl Res Cent 2020; 44: 1-13.
[24] Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, Jensen LJ and von Mering C. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res 2021; 49: D605-D612.
[25] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.
[26] Chin CH, Chen SH, Wu HH, Ho CW, Ko MT and Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 2014; 8 Suppl 4: S11.
[27] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK and Varambally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017; 19: 649-658.
[28] Tang Z, Li C, Kang B, Gao G, Li C and Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017; 45: W98-W102.
[29] Tang G, Cho M and Wang X. OncoDB: an interactive online database for analysis of gene expression and viral infection in cancer. Nucleic Acids Res 2022; 50: D1334-D1339.
[30] Park SJ, Yoon BH, Kim SK and Kim SY. GENT2: an updated gene expression database for normal and tumor tissues. BMC Med Genomics 2019; 12 Suppl 5: 101.
[31] Xu Y, Wang X, Huang Y, Ye D and Chi P. A LASSO-based survival prediction model for patients with synchronous colorectal carcinomas based on SEER. Transl Cancer Res 2022; 11: 2795-2809.
[32] Koch A, De Meyer T, Jeschke J and Van Criekinge W. MEXPRESS: visualizing expression, DNA methylation and clinical TCGA data. BMC Genomics 2015; 16: 636.
[33] Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C and Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013; 6: pl1.
[34] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102: 15545-15550.
[35] Chen B, Khodadoust MS, Liu CL, Newman AM and Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. Methods Mol Biol 2018; 1711: 243-259.
[36] Huang DP, Zeng YH, Yuan WQ, Huang XF, Chen SQ, Wang MY, Qiu YJ and Tong GD. Bioinformatics analyses of potential miRNA-mRNA regulatory axis in HBV-related hepatocellular carcinoma. Int J Med Sci 2021; 18: 335-346.
[37] Freshour SL, Kiwala S, Cotto KC, Coffman AC, McMichael JF, Song JJ, Griffith M, Griffith OL and Wagner AH. Integration of the drug-gene interaction database (DGIdb 4.0) with open crowdsource efforts. Nucleic Acids Res 2021; 49: D1144-D1151.

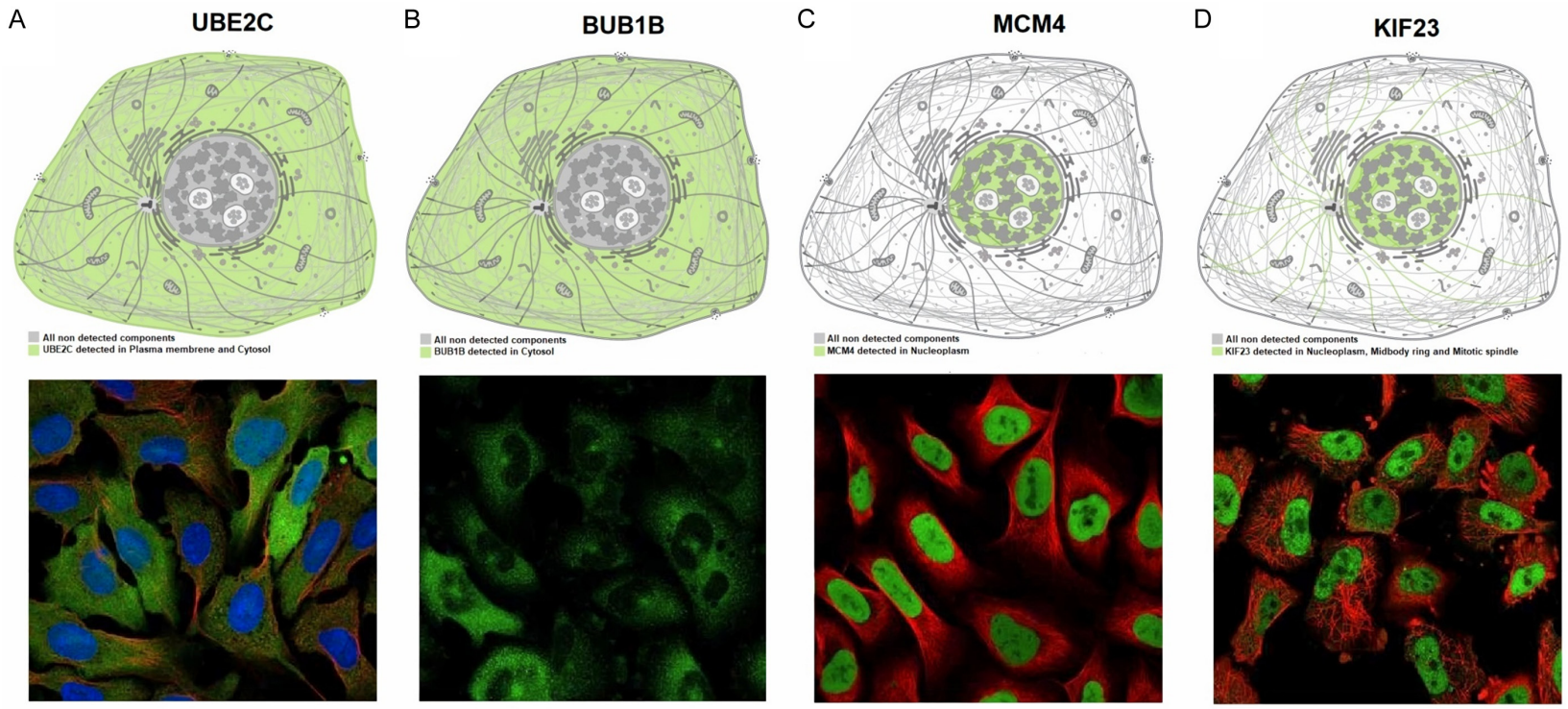
[38] Rio DC, Ares M Jr, Hannon GJ and Nilsen TW. Purification of RNA using TRIzol (TRI reagent). Cold Spring Harb Protoc 2010; 2010: pdb.prot5439.
[39] Livak KJ and Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001; 25: 402-408.
[40] Dong XY, Peng JR, Ye YJ, Chen HS, Zhang LJ, Pang XW, Li Y, Zhang Y, Wang S, Fant ME, Yin YH and Chen WF. Plac1 is a tumor-specific antigen capable of eliciting spontaneous antibody responses in human cancer patients. Int J Cancer 2008; 122: 2038-2043.
[41] Kim TK. T test as a parametric statistic. Korean J Anesthesiol 2015; 68: 540-546.
[42] Kim HY. Statistical notes for clinical researchers: chi-squared test and Fisher's exact test. Restor Dent Endod 2017; 42: 152-155.
[43] Jung AC, Job S, Ledrappier S, Macabre C, Abecassis J, de Reyniès A and Wasylyk B. A poor prognosis subtype of HNSCC is consistently observed across methylome, transcriptome, and miRNome analysis. Clin Cancer Res 2013; 19: 4174-4184.
[44] Karimi MR, Karimi AH, Abolmaali S, Sadeghi M and Schmitz U. Prospects and challenges of cancer systems medicine: from genes to disease networks. Brief Bioinform 2022; 23: bbab343.
[45] Mao J, Huang X, Okla MK, Abdel-Maksoud MA, Mubarak A, Hameed Z, Noreen R, Chaudhary A, Ghazanfar S, Liao Y, Hameed Y and Li C. Risk factors for TERT promoter mutations with papillary thyroid carcinoma patients: a meta-analysis and systematic review. Comput Math Methods Med 2022; 2022: 1721526.
[46] Usman M, Hameed Y, Ahmad M, Iqbal MJ, Maryam A, Mazhar A, Naz S, Tanveer R, Saeed H, Bint-E-Fatima, Ashraf A, Hadi A, Hameed Z, Tariq E and Aslam AS. SHMT2 is associated with tumor purity, CD8+ T immune cells infiltration, and a novel therapeutic target in four different human cancers. Curr Mol Med 2023; 23: 161-176.
[47] Sial N, Rehman JU, Saeed S, Ahmad M, Hameed Y, Atif M, Rehman A, Asif R, Ahmed H, Hussain MS, Khan MR, Ambreen A and Ambreen A. Integrative analysis reveals methylenetetrahydrofolate dehydrogenase 1-like as an independent shared diagnostic and prognostic biomarker in five different human cancers. Biosci Rep 2022; 42: BSR20211783.
[48] Khan M and Hameed Y. Discovery of novel six genes-based cervical cancer-associated biomarkers that are capable to break the heterogeneity barrier and applicable at the global level. J Cancer Ther 2023; 9000: 231-239.
[49] Shen Z, Jiang X, Zeng C, Zheng S, Luo B, Zeng Y, Ding R, Jiang H, He Q, Guo J and Jie W. High expression of ubiquitin-conjugating enzyme 2C (UBE2C) correlates with nasopharyngeal carcinoma progression. BMC Cancer 2013; 13: 192.
[50] Rape M and Kirschner MW. Autonomous regulation of the anaphase-promoting complex couples mitosis to S-phase entry. Nature 2004; 432: 588-595.
[51] Rape M, Reddy SK and Kirschner MW. The processivity of multiubiquitination by the APC determines the order of substrate degradation. Cell 2006; 124: 89-103.
[52] Parris TZ, Danielsson A, Nemes S, Kovács A, Delle U, Fallenius G, Möllerström E, Karlsson P and Helou K. Clinical implications of gene dosage and gene expression patterns in diploid breast carcinoma. Clin Cancer Res 2010; 16: 3860-3874.
[53] Lin J, Raoof DA, Wang Z, Lin MY, Thomas DG, Greenson JK, Giordano TJ, Orringer MB, Chang AC, Beer DG and Lin L. Expression and effect of inhibition of the ubiquitin-conjugating enzyme E2C on esophageal adenocarcinoma. Neoplasia 2006; 8: 1062-1071.
[54] Bavi P, Uddin S, Ahmed M, Jehan Z, Bu R, Abubaker J, Sultana M, Al-Sanea N, Abduljabbar A, Ashari LH, Alhomoud S, Al-Dayel F, Prabhakaran S, Hussain AR and Al-Kuraya KS. Bortezomib stabilizes mitotic cyclins and prevents cell cycle progression via inhibition of UBE2C in colorectal carcinoma. Am J Pathol 2011; 178: 2109-2120.
[55] Wagner KW, Sapinoso LM, El-Rifai W, Frierson HF, Butz N, Mestan J, Hofmann F, Deveraux QL and Hampton GM. Overexpression, genomic amplification and therapeutic potential of inhibiting the UbcH10 ubiquitin conjugase in human carcinomas of diverse anatomic origin. Oncogene 2004; 23: 6621-6629.
[56] Ieta K, Ojima E, Tanaka F, Nakamura Y, Haraguchi N, Mimori K, Inoue H, Kuwano H and Mori M. Identification of overexpressed genes in hepatocellular carcinoma, with special reference to ubiquitin-conjugating enzyme E2C gene expression. Int J Cancer 2007; 121: 33-38.
[57] Berlingieri MT, Pallante P, Guida M, Nappi C, Masciullo V, Scambia G, Ferraro A, Leone V, Sboner A, Barbareschi M, Ferro A, Troncone G and Fusco A. UbcH10 expression may be a useful tool in the prognosis of ovarian carcinomas. Oncogene 2007; 26: 2136-2140.
[58] Donato G, Iofrida G, Lavano A, Volpentesta G, Signorelli F, Pallante PL, Berlingieri MT, Pierantoni MG, Palmieri D, Conforti F, Maltese L, Tucci L, Amorosi A and Fusco A. Analysis of UbcH10 expression represents a useful tool for the diagnosis and therapy of astrocytic tumors. Clin Neuropathol 2008; 27: 219-223.
[59] Jiang L, Huang CG, Lu YC, Luo C, Hu GH, Liu HM, Chen JX and Han HX. Expression of ubiqui-

tin-conjugating enzyme E2C/UbcH10 in astrocytic tumors. Brain Res 2008; 1201: 161-166.
[60] Silva PMA and Bousbaa H. BUB3, beyond the simple role of partner. Pharmaceutics 2022; 14: 1084.
[61] Mullany LE, Herrick JS, Sakoda LC, Samowitz W, Stevens JR, Wolff RK and Slattery ML. miRNA involvement in cell cycle regulation in colorectal cancer cases. Genes Cancer 2018; 9: 53-65.
[62] Chen RH. BubR1 is essential for kinetochore localization of other spindle checkpoint proteins and its phosphorylation requires Mad1. J Cell Biol 2002; 158: 487-496.
[63] Musacchio A and Salmon ED. The spindle-assembly checkpoint in space and time. Nat Rev Mol Cell Biol 2007; 8: 379-393.
[64] Shin HJ, Baek KH, Jeon AH, Park MT, Lee SJ, Kang CM, Lee HS, Yoo SH, Chung DH, Sung YC, McKeon F and Lee CW. Dual roles of human BubR1, a mitotic checkpoint kinase, in the monitoring of chromosomal instability. Cancer Cell 2003; 4: 483-497.
[65] Park HY, Jeon YK, Shin HJ, Kim IJ, Kang HC, Jeong SJ, Chung DH and Lee CW. Differential promoter methylation may be a key molecular mechanism in regulating BubR1 expression in cancer cells. Exp Mol Med 2007; 39: 195-204.
[66] Ando K, Kakeji Y, Kitao H, Iimori M, Zhao Y, Yoshida R, Oki E, Yoshinaga K, Matumoto T, Morita M, Sakaguchi Y and Maehara Y. High expression of BUBR1 is one of the factors for inducing DNA aneuploidy and progression in gastric cancer. Cancer Sci 2010; 101: 639-645.
[67] Yuan B, Xu Y, Woo JH, Wang Y, Bae YK, Yoon DS, Wersto RP, Tully E, Wilsbach K and Gabrielson E. Increased expression of mitotic checkpoint genes in breast cancer cells with chromosomal instability. Clin Cancer Res 2006; 12: 405-410.
[68] Yamamoto Y, Matsuyama H, Chochi Y, Okuda M, Kawauchi S, Inoue R, Furuya T, Oga A, Naito K and Sasaki K. Overexpression of BUBR1 is associated with chromosomal instability in bladder cancer. Cancer Genet Cytogenet 2007; 174: 42-47.
[69] Liu AW, Cai J, Zhao XL, Xu AM, Fu HQ, Nian H and Zhang SH. The clinicopathological significance of BUBR1 overexpression in hepatocellular carcinoma. J Clin Pathol 2009; 62: 1003-1008.
[70] Tanaka K, Mohri Y, Ohi M, Yokoe T, Koike Y, Morimoto Y, Miki C, Tonouchi H and Kusunoki M. Mitotic checkpoint genes, hsMAD2 and BubR1, in oesophageal squamous cancer cells and their association with 5-fluorouracil and cisplatin-based radiochemotherapy. Clin Oncol (R Coll Radiol) 2008; 20: 639-646.
[71] Adachi Y, Usukura J and Yanagida M. A globular complex formation by Nda1 and the other five members of the MCM protein family in fission yeast. Genes Cells 1997; 2: 467-479.
[72] Gozuacik D, Chami M, Lagorce D, Faivre J, Murakami Y, Poch O, Biermann E, Knippers R, Bréchot C and Paterlini-Bréchot P. Identification and functional characterization of a new member of the human Mcm protein family: hMcm8. Nucleic Acids Res 2003; 31: 570-579.
[73] Huang XP, Rong TH, Wu QL, Fu JH, Yang H, Zhao JM and Fang Y. MCM4 expression in esophageal cancer from southern China and its clinical significance. J Cancer Res Clin Oncol 2005; 131: 677-682.
[74] Winnepenninckx V, Lazar V, Michiels S, Dessen P, Stas M, Alonso SR, Avril MF, Ortiz Romero PL, Robert T, Balacescu O, Eggermont AM, Lenoir G, Sarasin A, Tursz T, van den Oord JJ and Spatz A; Melanoma Group of the European Organization for Research and Treatment of Cancer. Gene expression profiling of primary cutaneous melanoma and clinical outcome. J Natl Cancer Inst 2006; 98: 472-482.
[75] Kikuchi J, Kinoshita I, Shimizu Y, Kikuchi E, Takeda K, Aburatani H, Oizumi S, Konishi J, Kaga K, Matsuno Y, Birrer MJ, Nishimura M and Dosaka-Akita H. Minichromosome maintenance (MCM) protein 4 as a marker for proliferation and its clinical and clinicopathological significance in non-small cell lung cancer. Lung Cancer 2011; 72: 229-237.
[76] Sanada H, Seki N, Mizuno K, Misono S, Uchida A, Yamada Y, Moriya S, Kikkawa N, Machida K, Kumamoto T, Suetsugu T and Inoue H. Involvement of dual strands of miR-143 (miR-143-5p and miR-143-3p) and their target oncogenes in the molecular pathogenesis of lung adenocarcinoma. Int J Mol Sci 2019; 20: 4482.
[77] Iolascon A, Heimpel H, Wahlin A and Tamary H. Congenital dyserythropoietic anemias: molecular insights and diagnostic approach. Blood 2013; 122: 2162-2166.
[78] Yu Y and Feng YM. The role of kinesin family proteins in tumorigenesis and progression: potential biomarkers and molecular targets for cancer therapy. Cancer 2010; 116: 5150-5160.
[79] Liu Y, Chen H, Dong P, Xie G, Zhou Y, Ma Y, Yuan X, Yang J, Han L, Chen L and Shen L. KIF23 activated Wnt/β-catenin signaling pathway through direct interaction with Amer1 in gastric cancer. Aging (Albany NY) 2020; 12: 8372-8396.
[80] Kato T, Wada H, Patel P, Hu HP, Lee D, Ujiie H, Hirohashi K, Nakajima T, Sato M, Kaji M, Kaga K, Matsui Y, Tsao MS and Yasufuku K. Overexpression of KIF23 predicts clinical outcome in

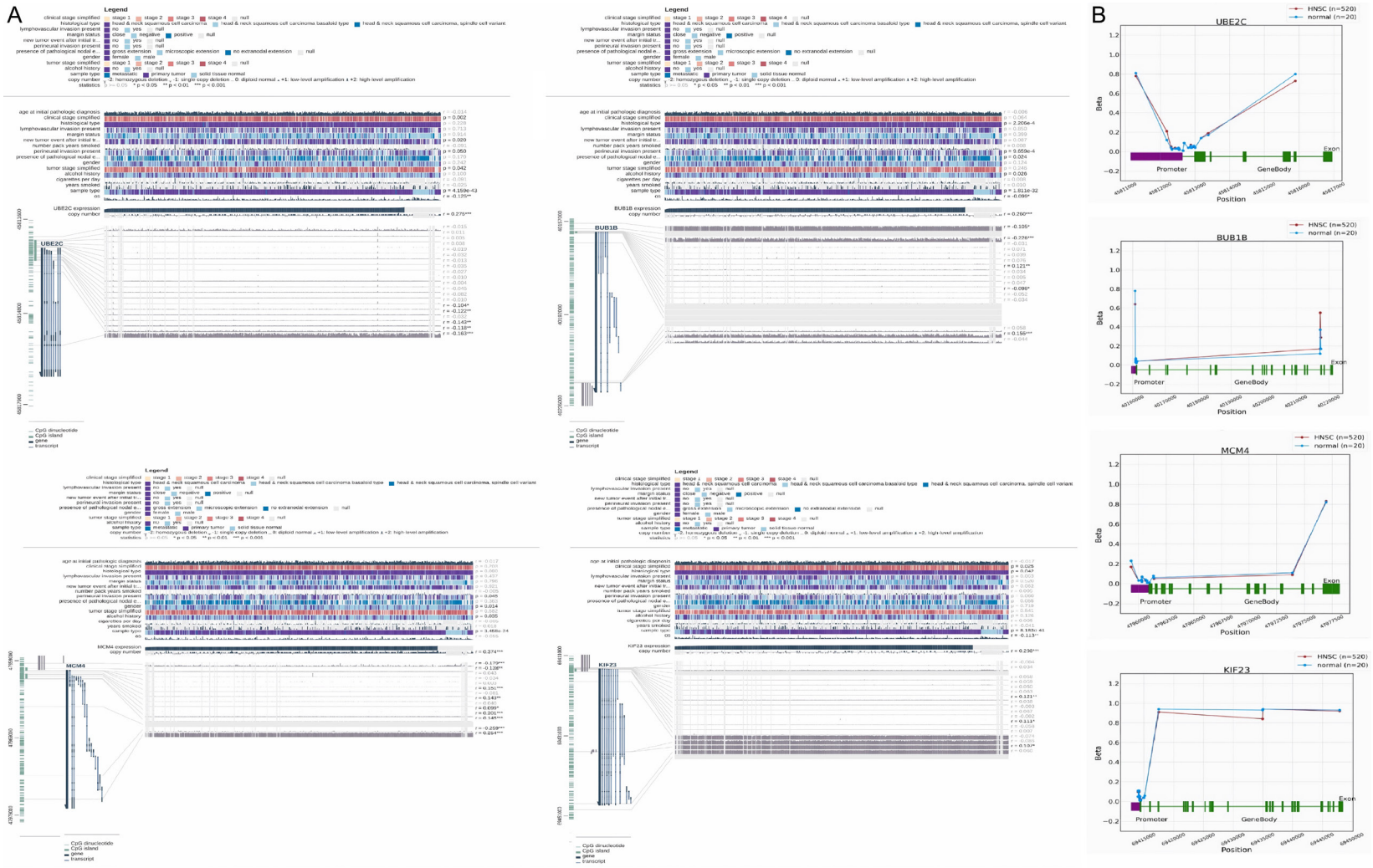
primary lung cancer patients. Lung Cancer 2016; 92: 53-61.
[81] Vikberg AL, Vooder T, Lokk K, Annilo T and Golovleva I. Mutation analysis and copy number alterations of KIF23 in non-small-cell lung cancer exhibiting KIF23 over-expression. Onco Targets Ther 2017; 10: 4969-4979.
[82] Hu Y, Zheng M, Wang C, Wang S, Gou R, Liu O, Li X, Liu J and Lin B. Identification of KIF23 as a prognostic signature for ovarian cancer based on large-scale sampling and clinical validation. Am J Transl Res 2020; 12: 4955-4976.
[83] Li TF, Zeng HJ, Shan Z, Ye RY, Cheang TY, Zhang YJ, Lu SH, Zhang Q, Shao N and Lin Y. Overexpression of kinesin superfamily members as prognostic biomarkers of breast cancer. Cancer Cell Int 2020; 20: 123.
[84] Gonzalez H, Hagerling C and Werb Z. Roles of the immune system in cancer: from tumor initiation to metastatic progression. Genes Dev 2018; 32: 1267-1284.
[85] de Ruiter EJ, Ooft ML, Devriese LA and Willems SM. The prognostic role of tumor infiltrating T-lymphocytes in squamous cell carcinoma of the head and neck: a systematic review and meta-analysis. Oncoimmunology 2017; 6: e1356148.
[86] Chen QF, Xia JG, Li W, Shen LJ, Huang T and Wu P. Examining the key genes and pathways in hepatocellular carcinoma development from hepatitis B virus-positive cirrhosis. Mol Med Rep 2018; 18: 4940-4950.
[87] Yamaguchi H, Hsu JL and Hung MC. Regulation of ubiquitination-mediated protein degradation by survival kinases in cancer. Front Oncol 2012; 2: 15.
[88] Peng Y and Croce CM. The role of MicroRNAs in human cancer. Signal Transduct Target Ther 2016; 1: 15004.
[89] Wang L, Liu Y, Du L, Li J, Jiang X, Zheng G, Qu A, Wang H, Wang L, Zhang X, Liu H, Pan H, Yang Y and Wang C. Identification and validation of reference genes for the detection of serum microRNAs by reverse transcription-quantitative polymerase chain reaction in patients with bladder cancer. Mol Med Rep 2015; 12: 615-622.
[90] Hu H, Chen C, Chen F and Sun N. LINC00152 knockdown suppresses tumorigenesis in non-small cell lung cancer via sponging miR-16-5p. J Thorac Dis 2022; 14: 614-624.



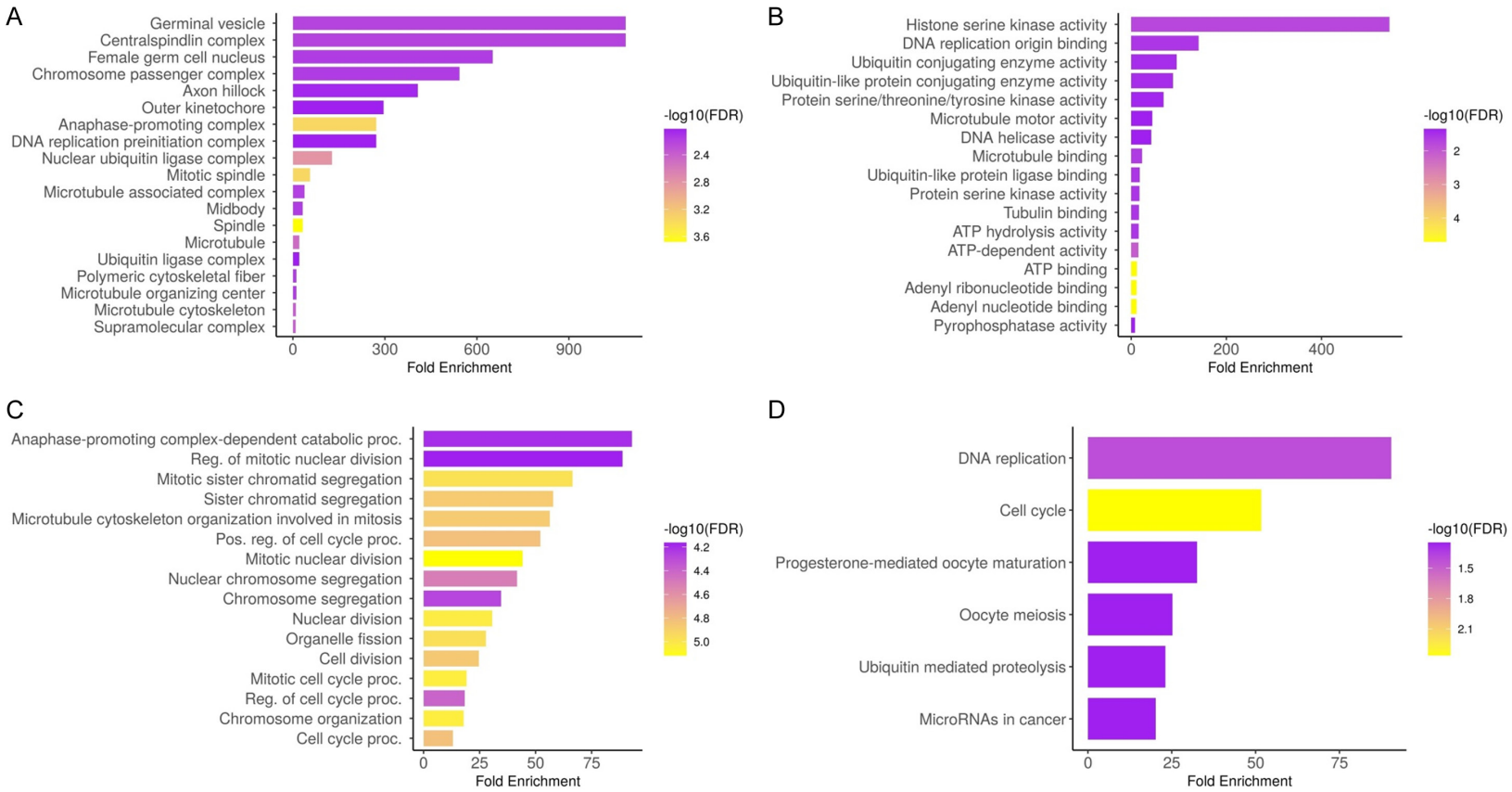
Supplementary Figure 1. Construction of the hub genes (UBE2C, BUB1B, MCM4, and KIF23) based prognostic model. (A) Univariate Cox regression analysis, (B) C-index scores, and (C) Risk scores.



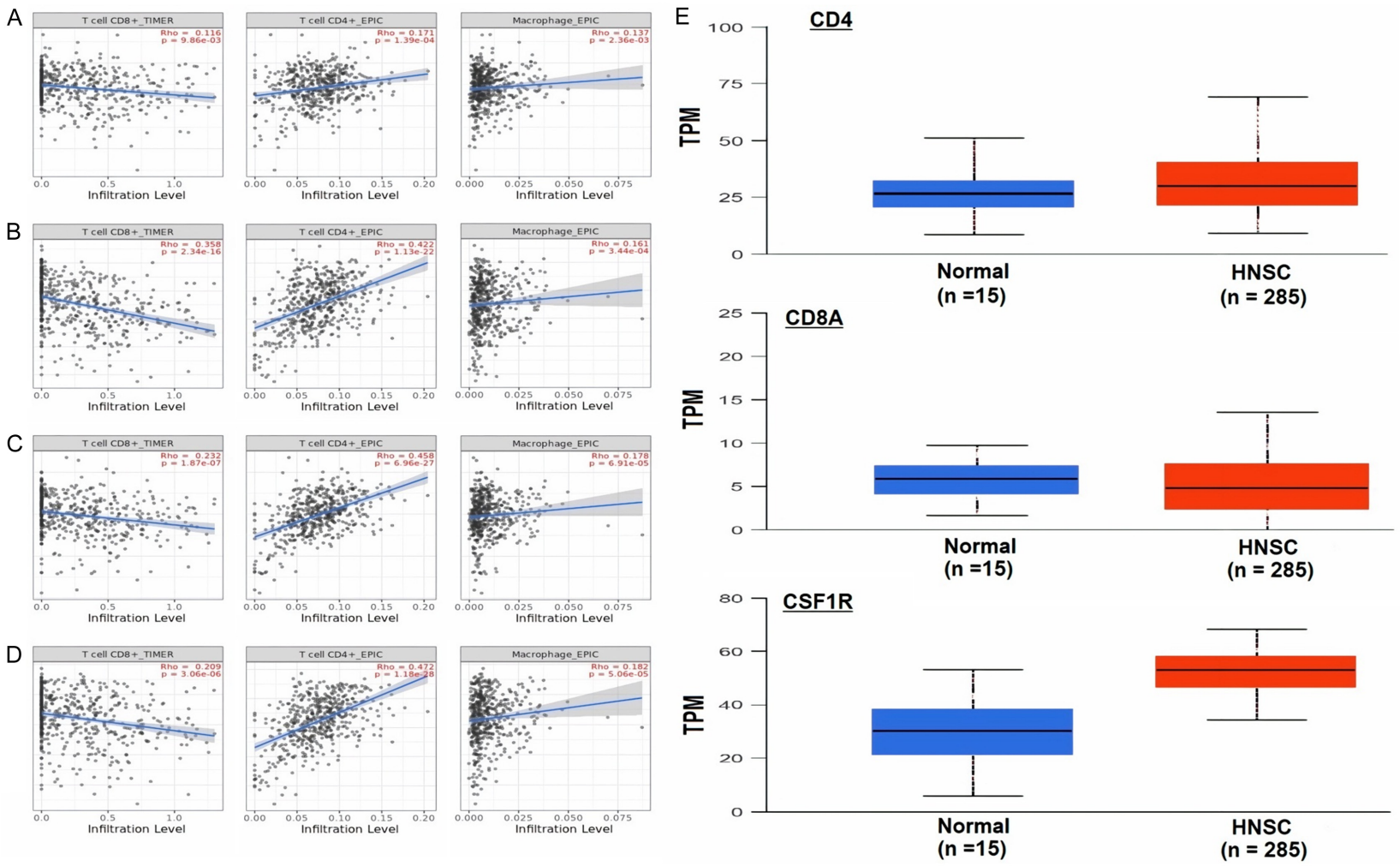
Supplementary Figure 2. Subcellular localization of the UBE2C, BUB1B, MCM4, and KIF23 via HPA database. (A) UBE2C, (B) BUB1B, (C) MCM4, and (D) KIF23.



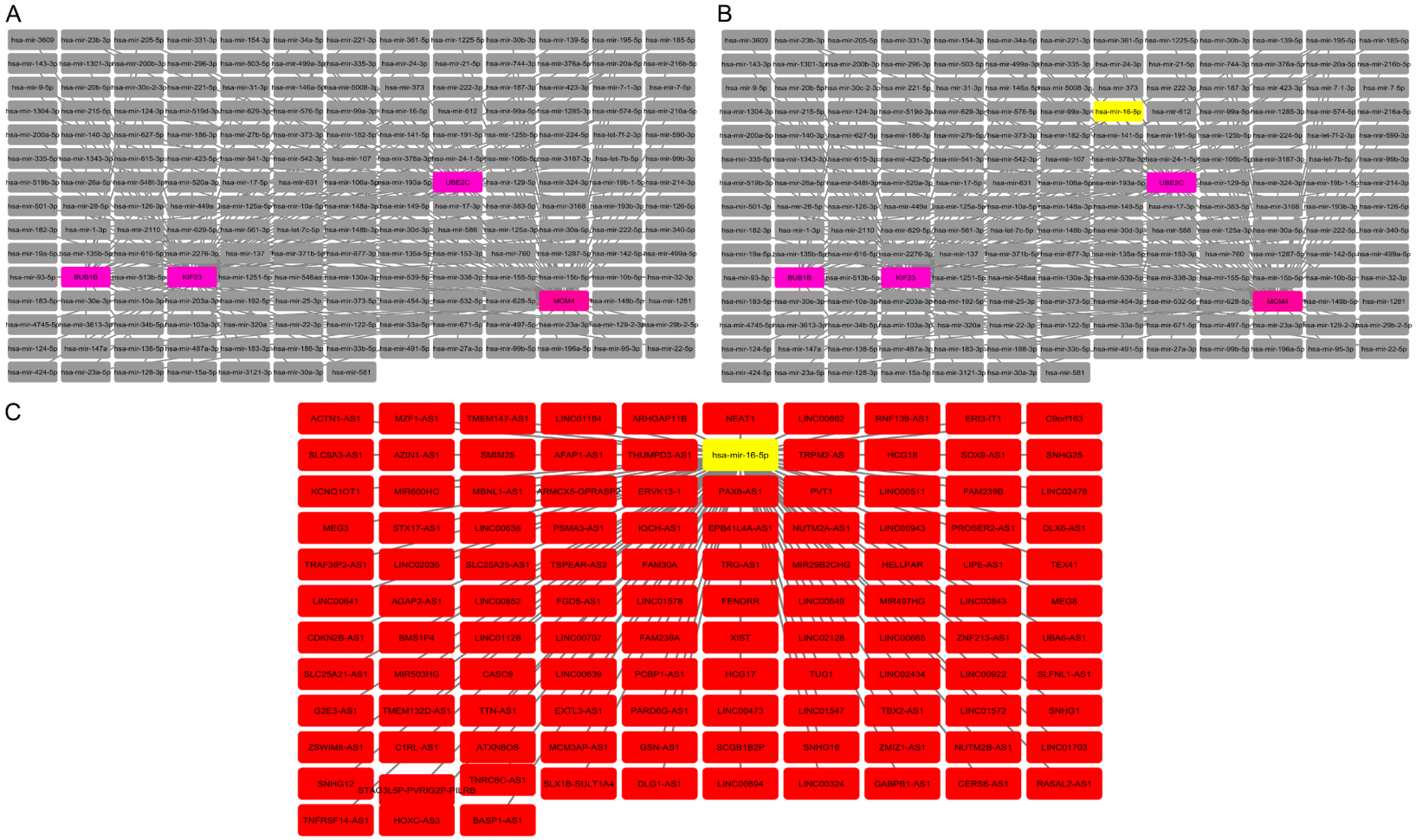
Supplementary Figure 3. Methylation status exploration of UBE2C, BUB1B, MCM4, and KIF23 via MEXPRESS and OncoDB in HNSC and normal samples. (A) Methylation status exploration of UBE2C, BUB1B, MCM4, and KIF23 via MEXPRESS, and (B) Methylation status exploration of UBE2C, BUB1B, MCM4, and KIF23 via OncoDB.



Supplementary Figure 4. Gene enrichment analysis of UBE2C, BUB1B, MCM4, and KIF23. (A) UBE2C, BUB1B, MCM4, and KIF23 associated CC terms, (B) UBE2C, BUB1B, MCM4, and KIF23 associated MF terms, (C) UBE2C, BUB1B, MCM4, and KIF23 associated BP terms, and (D) UBE2C, BUB1B, MCM4, and KIF23 associated KEGG terms.



Supplementary Figure 5. Correlation analysis of UBE2C, BUB1B, MCM4, and KIF23 hub genes expression with different immune cells (CD8+ T, CD4+ T, and Macrophages) infiltration level and GEO datasets-based expression analysis of CD4, CD8A, and CSF1R genes across HNSC and normal cells. (A) Correlation analysis of UBE2C with immune cells in HNSC, (B) Correlation analysis of BUB1B with immune cells in HNSC, (C) Correlation analysis of MCM4 with immune cells in HNSC, (D) Correlation analysis of KIF23 with immune cells in HNSC, and (E) GEO datasets-based expression analysis of CD4, CD8A, and CSF1R genes across HNSC and normal controls.



Supplementary Figure 6. lncRNA-miRNA-mRNA co-regulatory network of UBE2C, BUB1B, MCM4, and KIF23 hub genes. (A) A PPI of miRNAs targeting hub genes, (B) A PPI highlighting most important miRNA (hsa-mir-16-5p) targeting all hub genes, and (C) A PPI of lncRNAs targeting hsa-mir-16-5p. Grey color nodes: miRNAs, pink color nodes: mRNAs, yellow color nodes: significant miRNA, and red color nodes: lncRNAs.