Am J Cancer Res 2023;13(4):1259-1277
www.ajcr.us /ISSN:2156-6976/ajcr0149436

Original Article

Identification of hub genes associated with head and neck squamous cell carcinoma by integrated bioinformatics approach and RNA-seq validation analysis

Amir Shabeer[1], Sara Mustafa[2], Rimsha Sadia Bukhari[3], Mostafa A Abdel-Maksoud[4], Saeedah Musaed Almutairi[4], Wahidah H Al-Qahtani[5], Khushboo Chandio[6], Nazima Yousaf Khan[7], Jaweria Gul[8], Mohammed Aufy[9]

[1]Isra University Hyderabad, Sindh, Pakistan; [2]Tehsil Headquarter (THQ) Rojhan, Rajanpur, Pakistan; [3]Fatima Jinnah Medical University, Lahore, Pakistan; [4]Department of Botany and Microbiology, College of Science, King Saud University Riyadh, P.O. 2455, Riyadh 11451, Saudi Arabia; [5]Department of Food Sciences & Nutrition, College of Food and Agricultural Sciences, King Saud University, P.O. Box 270677, Riyadh 11352, Saudi Arabia; [6]Peoples Nursing School LUMHS, Jamshoro, Pakistan; [7]Institute of Biochemistry, University of Balochistan, Quetta, Pakistan; [8]Shaheed Benazir Bhutto University, Sheringal, Dir Upper, Pakistan; [9]Department of Pharmaceutical Sciences, Division of Pharmacology and Toxicology, University of Vienna, Vienna, Austria

Received January 31, 2023; Accepted March 19, 2023; Epub April 15, 2023; Published April 30, 2023

Abstract: Head and neck squamous cell carcinoma (HNSC) is one of the most lethal malignancies around the globe. Due to its complex nature, the diagnostic and prognostic signatures of HNSC remain poorly understood. This study was launched to identify signature genes and their signaling pathways related to the development of HNSC. In the current study, we retrieved the GSE53819 dataset from the Gene Expression Omnibus (GEO) database to determine the differentially expressed genes (DEGs) using the "Limma" R package. Adjusted *P* values $P < 0.05$ and $|logFC| \geq 1$ were selected as the filtering conditions. To identify hub genes, the protein-protein interaction (PPI) network construction of the DEGs was performed using STRING. We further used UALCAN, GEPIA, OncoDB, GENT2, MEXPRESS, and HPA databases for the expression, validation, survival, and methylation analyses of the hub genes. The cBioPortal tool was used to investigate the genetic alterations in hub genes. CancerSEA, TIMER, DAVID, ENCORI, and DrugBank were also used to explore a few more hub gene-associated parameters. Lastly, HOK, FaDu, and SCC25 cell lines were used to validate hub gene expression via RNA sequencing (RNA-seq) technique. A total of top 250 DEGs were selected for detailed analysis in this study. From these DEGs, prognostic and diagnostic associated four hub genes, which could serve as potential molecular biomarkers and therapeutic targets in HNSC patients were identified. Four hub genes, including down-regulated DNAH1 and DNALI1, while up-regulated DNAH9 and CCDC151 were strongly implicated in HNSC. We also validated the same expression pattern of the hub genes using RNA-seq analysis in HNSC and normal cell lines. Moreover, this study also revealed some novel links between DNAH1, DNALI1, DNAH9, and CCDC151 expression and genetic alterations, promoter methylation status, immune cell infiltration, miRNAs, gene enrichment terms, and various chemotherapeutic drugs. In conclusion, we indicated four hub genes (DNAH1, DNALI1, DNAH9, and CCDC151) and their associated signaling pathways, which may improve our understanding of HNSC and could be used as new therapeutic targets.

Keywords: HNSC, biomarker, gene expression, therapeutic target

Introduction

Head and neck squamous cell carcinoma (HNSC) is a group of different malignancies that originate in the oral epithelium, oropharynx, larynx, and hypopharynx [1]. HNSC is the 6th most prevalent cancer type around the globe [2]. In the United States (US), a total of 50,000 new HNSC cases, and nearly 10,000 mortalities due to HNSC were recorded in 2021 [3]. The rate of HNSC occurring in whites is increasing by about 1% a year, and the frequency of HNSC cases is reported to be more than twice as high in males compared to females [4]. Early detection of can-

cer is the key feature leading to less spreading of the disease, fruitful cancer treatment, and better survival outcomes for patients [5].

Cancer development is an extremely complex process involving a variety of genetic alterations in the genome, including up-regulation of oncogenes and inactivation or down-regulation of tumor suppressor genes (TSGs) [6]. Despite the availability of advanced treatment options, including surgery, radiotherapy, and chemotherapy, more than 50% of HNSC patients died around the world [7]. Late diagnosis, metastasis of cancer cells from primary to secondary sites, and tumor recurrences are the major causes of poor prognosis in HNSC patients, leading to an increased mortality rate [7]. Therefore, the discovery of authentic and more effective molecular markers in HNSC patients and an understanding of the oncogenic roles of these biomarkers is urgently needed.

In this research, we downloaded the GSE53819 [8] gene expression dataset from the Gene Expression Omnibus (GEO) website. This website is developed to publically house gene expression datasets of numerous malignancies [9]. After downloading, the expression profiles of HNSC patients and normal samples in this dataset were compared to determine the differentially expressed genes (DEGs). After DEGs determination, we constructed a protein-protein interaction (PPI) network of the DEGs and carried out the Molecular Complex Detection (MCODE) and Cytohubba analyses for identifying significant gene module and novel genes (hub genes) based on the degree-score method. Lastly, we employed a series of comprehensive bioinformatics analyses to uncover the oncogenic roles of the identified hub genes. Our study provides a novel piece of information on Dynein, axonemal, heavy chain 1 (DNAH1), Dynein Axonemal Light Intermediate Chain 1 (DNALI1), Dynein, axonemal, heavy chain 1 (DNAH9), and Coiled-coil domain containing 151 (CCDC151) genes as tumor-promoting factors, reliable potential molecular biomarkers, and therapeutic targets for treatment in HNSC patients.

Methodology

Microarray data

First of all, the GEO database [9] was searched to find a suitable HNSC dataset. Following the completion of the search process, GSE53819 datasets [8] having gene expression profiles of HNSC and normal samples were chosen for further investigation. The GSE53819 dataset is based on platform GPL6102 (Illumina human-6 v2.0 expression beadchip). This dataset is freely available on the GEO database.

Determination of the DEGs

The "Limma" package [10] is a widely used package for processing GEO datasets in order to identify DEGs between user-defined groups. After analyzing samples information in the GSE53819 dataset, only HNSC and their corresponding control samples were chosen from this dataset for analysis. In the identification process, an adjusted $P < 0.05$ and $|logFC| \geq 1$ were selected as the DEGs identification parameters.

Construction of PPI, module identification, and the selection of hub genes

The Search Tool for the Retrieval of Interacting Genes (STRING) database [11] with default settings has helped us in the construction of the PPI network of the top 250 DEGs, which were found in the HSNC samples. The Cytoscape software [12] was useful in visualizing and analyzing the constructed PPI networks of the DEGs. A Cytoscape application, Molecular Complex Detection (MCODE) [13], was helpful in recognizing the most significant gene module in the constructed PPI network. Another Cytoscape application, CytoHubba [14], has helped in identify hub genes across the recognized gene module based on the degree-ranked method [15].

UALCAN database

UALCAN [16], a detailed analysis resource, contains clinical and expression data for more than 30 types of cancer from the TCGA projects. We utilized this resource for the expression analysis of the hub genes in normal samples, and HNSC samples of different clinicopathological parameters.

GEPIA, OncoDB, and GENT2 databases

In order to further validate hub gene expressions in HNSC tissues and cell lines and to check the effect of these gene expression on the survival of HNSC patients, we used the GEPIA [17], OncoDb [18] and GENT2 [19] data-

bases for expression analysis. These databases use TCGA projects expression data and expression analysis results in the form of box plots.

MEXPRESS

OncoDB [18] and MEXPRESS [20] (https://mexpress.be/) resources were used in our work to assess the DNA promoter methylation levels of identified hub genes in HNSC.

cBioPortal

The cBioPortal [21] is a freely available resource for investigating cancer multi-omics data. In the present study, a TCGA HNSC dataset, labeled as, "TCGA, Firehose Legacy (530 cases)", was used for exploring genetic mutations, mutational hotspots, co-expressed genes, promoter methylation levels, and the effect of mutations on the survival durations of the HNSC patients.

The human protein atlas (HPA)

The HPA (https://www.proteinatlas.org/) online database [22] was used in the present study to find the subcellular localization of proteins encoded by the hub genes in HNSC cells. Moreover, this database has also helped to perform hub gene expression analysis at the protein level.

Functional enrichment analysis

GSEA [23], a freely available tool, was utilized in the current study to explore DEGs-associated GO and KEGG terms across HNSC. The enrichment outcomes of this analysis were presented as classical bar charts.

TIMER database

The TIMER database, which has a web-based interface, is used to assess immune cell infiltration in tumours [24]. In this work, the expression of hub genes was compared to the immune cell infiltration in HNSC.

The miRNA network of the hub genes

The ENCORI database [25] was used to predict the miRNA network of the identified hub genes in the current study.

Hub genes' drug prediction analysis

By considering hub genes as therapeutic targets in the treatment of HNSC, we conducted the DrugBank analysis to identify hub genes' associated drugs. This database provides a detailed account on the mechanistic details of different drugs targeting hub gene expression [26].

RNA-seq based in vitro validation of the hub genes expression

Three human cell lines, i.e., two HNSC (FaDu, and SCC25) and one normal human oral keratinocyte (HOK) cell lines were purchased from the ATCC (American Type Culture Collection). All cell lines were cultured in DMEM (HyClone), supplemented with 10% fetal bovine serum (FBS; TBD, Tianjin, China), 1% glutamine, and 1% penicillin-streptomycin in 5% CO_2 at 37°C. Total RNA was extracted from cells using TRIzol® reagent (Life Technologies [Thermo Fisher Scientific], Carlsbad, CA, USA) according to the manufacturer's protocol, and sent to Beijing Genomics Institute (BGI) company for further RNA-seq detection and analysis using the MGISEQ-2000 sequencer. The Dr. Tom network platform (http://report.bgi.com) developed by the BGI was used to obtain fragments per kilo base million reads FPKM expression values. The obtained FPKM values against hub genes in HNSC and normal oral keratinocyte (HOK) cell lines were compared using a t-test to identify differences in the expression levels.

Statistics details for in silico analyses

DEGs were identified using a t-test [27]. While for GO and KEGG enrichment analysis, we used Fisher's Exact test for computing statistical difference [28]. 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 identification

As highlighted in **Figure 1A-C**, based on the selection criteria (mentioned in the method section), a total of 11331 DEGs were found between the HNSC and normal samples group

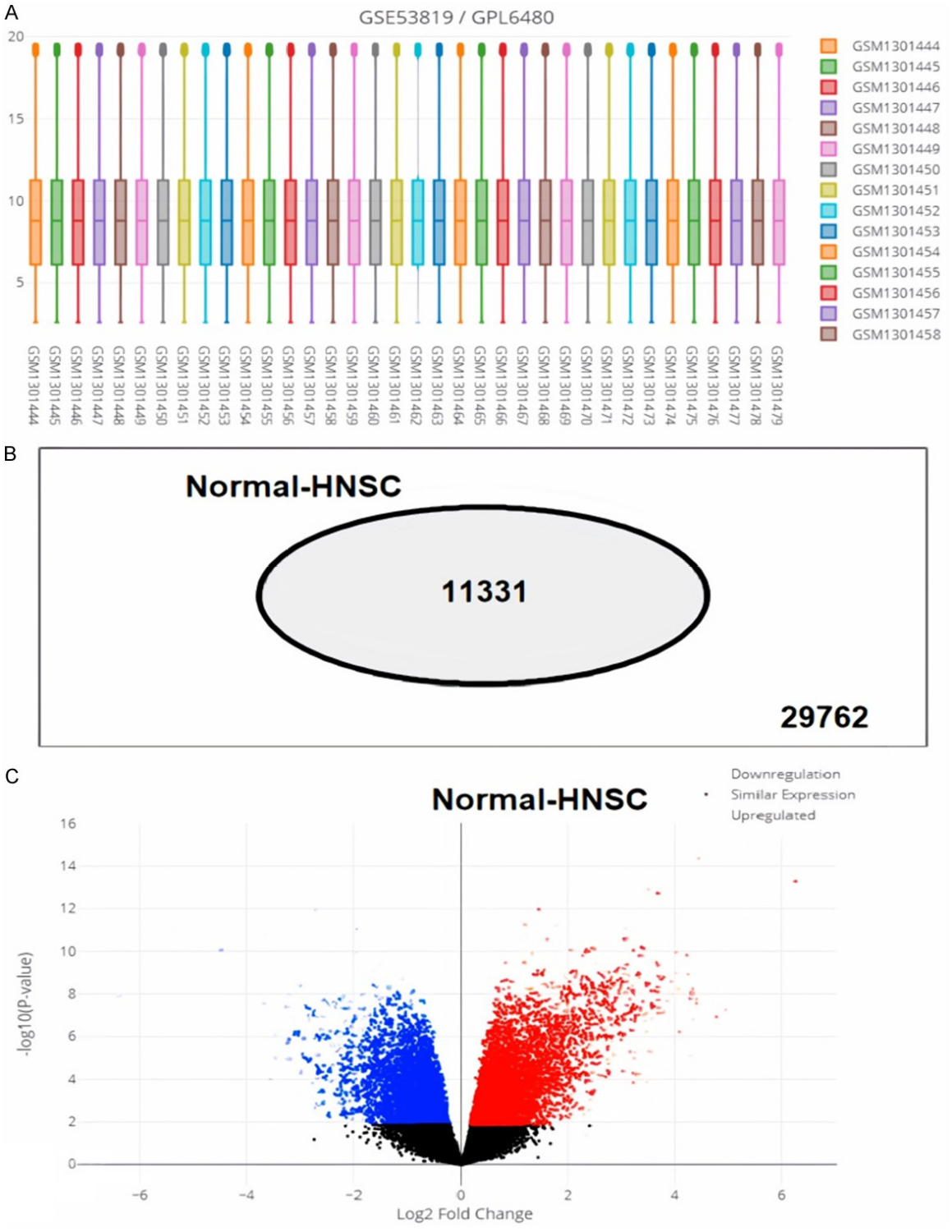


Figure 1. A comparison between expression profiles of samples, volcano graphs of DEGs, and a total count of DEGs in GSE53819 microarray dataset. (A) A comparison between expression profiles of samples in GSE53819 microarray dataset, (B) A volcano graph of the DEGs observed in GSE53819 microarray dataset, and (C) A total count of DEGs and non-DEGs in GSE53819 microarray dataset.

across GSE53819 dataset. Finally, out of the total identified 11331 DEGs, the top 250 DEGs with high reliability were chosen for further analysis in this study.

PPI network and hub genes selection

With the help of STRING, the PPI network of the top 250 DEGs was constructed. In this step, the isolated, non-interacting nodes were removed from the network. After construction, the ultimately visualized PPI network was consisted of 87 nodes and 212 edges (Supplementary Figure 1A). Cytoscap-based plug-in applications (MCODE and Cytohubba) were used to analyze significant module and hub genes in the network, respectively. According to the MCODE analysis, the significant module was consisted of 16 nodes in the network (Supplementary Figure 1B, 1C). According to the Cytohubba analysis, Dynein, axonemal, heavy chain 1 (DNAH1), Dynein Axonemal Light Intermediate Chain 1 (DNALI1), Dynein, axonemal, heavy chain 1 (DNAH9), and Coiled-coil domain containing 151 (CCDC151) genes with the most stable degree scores were the hub genes across the selected gene module (Supplementary Figure 1D).

mRNA expression analysis

Based on the DEGs analysis of the GSE53819 dataset, out of the four identified hub genes (DNAH1, DNALI1, DNAH9, and CCDC151), DNAH1 and DNALI1 were found to be down-regulated while DNAH9 and CCDC151 were up-regulated in HNSC samples. We further analyzed the mRNA expression of the hub genes via UALCAN to check the expression behavior of these genes in HNSC samples with different clinical parameters and normal controls across TCGA projects. As shown in **Figures 2** and **3**, the mRNA expression of the DNAH1 and DNALI was significantly down-regulated, while the mRNA expression of DNAH9 and CCDC151 was notably higher in HNSC samples with different clinical parameters (cancer stage, race, and gender) than in normal tissue (**Figures 2** and **3**). The findings of the UALCAN analysis were consistent with the findings obtained from the microarray dataset.

Verification of the hub genes expression and survival analysis

Three additional databases (GEPIA, OncoDB, and GENT2) were further employed in the current study for the mRNA expression validation of DNAH1, DNALI1, DNAH9, and CCDC151 hub genes in HNSC patients cohorts and cell lines. As shown in **Figure 4A** and **4B**, According to GEPIA, OncoDB databases, the mRNA expression of DNAH1 and DNALI1 was significantly down-regulated while mRNA expression of DNAH9 and CCDC151 was significantly down-regulated in HNSC patient samples compared to the controls. In addition to this, the mRNA expression profiling of the DNAH1, DNALI1, DNAH9, and CCDC151 genes across HNSC cell line samples via GENT2 analysis also confirmed the down-regulation of DNAH1 and DNALI1 while up-regulation of DNAH9 and CCDC151 (**Figure 4C**). Then, we explored the prognostic values of DNAH1, DNALI1, DNAH9, and CCDC151 gene in HNSC patients using "Survival" option of the GEPIA database. Results of the analysis highlighted that the lower expression of DNAH1 and DNALI1 while higher expression of DNAH9 and CCDC151 are linked with the worst OS of the HNSC patients (**Figure 4D**). Therefore, it is speculated that hub genes could be employed as an accurate prognostic model to predict HNSC patient survival rates.

Subcellular localization, protein expression validation, and survival analysis of the DNAH1, DNALI1, DNAH9, and CCDC151

Through HPA database, the subcellular location of DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC cells was noted. For DNAH1 and DNALI1, these proteins were mainly enriched in centrosome (**Figure 5A**), the DNAH9 localization was found nucleoplasm, plasma membrane, and cytosol (**Figure 5A**), and CCDC151 localization was enriched in Golgi apparatus and vesicles (**Figure 5A**). An immunohistochemistry (IHC)-based protein expression of the DNAH1, DNALI1, DNAH9, and CCDC151 was analyzed in HNSC samples relative to controls via HPA. As a result, the expressions of DNAH1, DNALI1 genes were not detected in the normal tissue (staining = not detected) relative to HNSC, whereas the medium expression of these genes were detected (staining = medium) in HNSC samples (**Figure 5B**). Moreover, the higher expressions of DNAH9 and CCDC-151 (staining = medium) were also found in HNSC samples relative control samples (staining = not detected) (**Figure 5B**).

Methylation status of DNAH1, DNALI1, DNAH9, and CCDC151

We examined the methylation status of DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC sam-

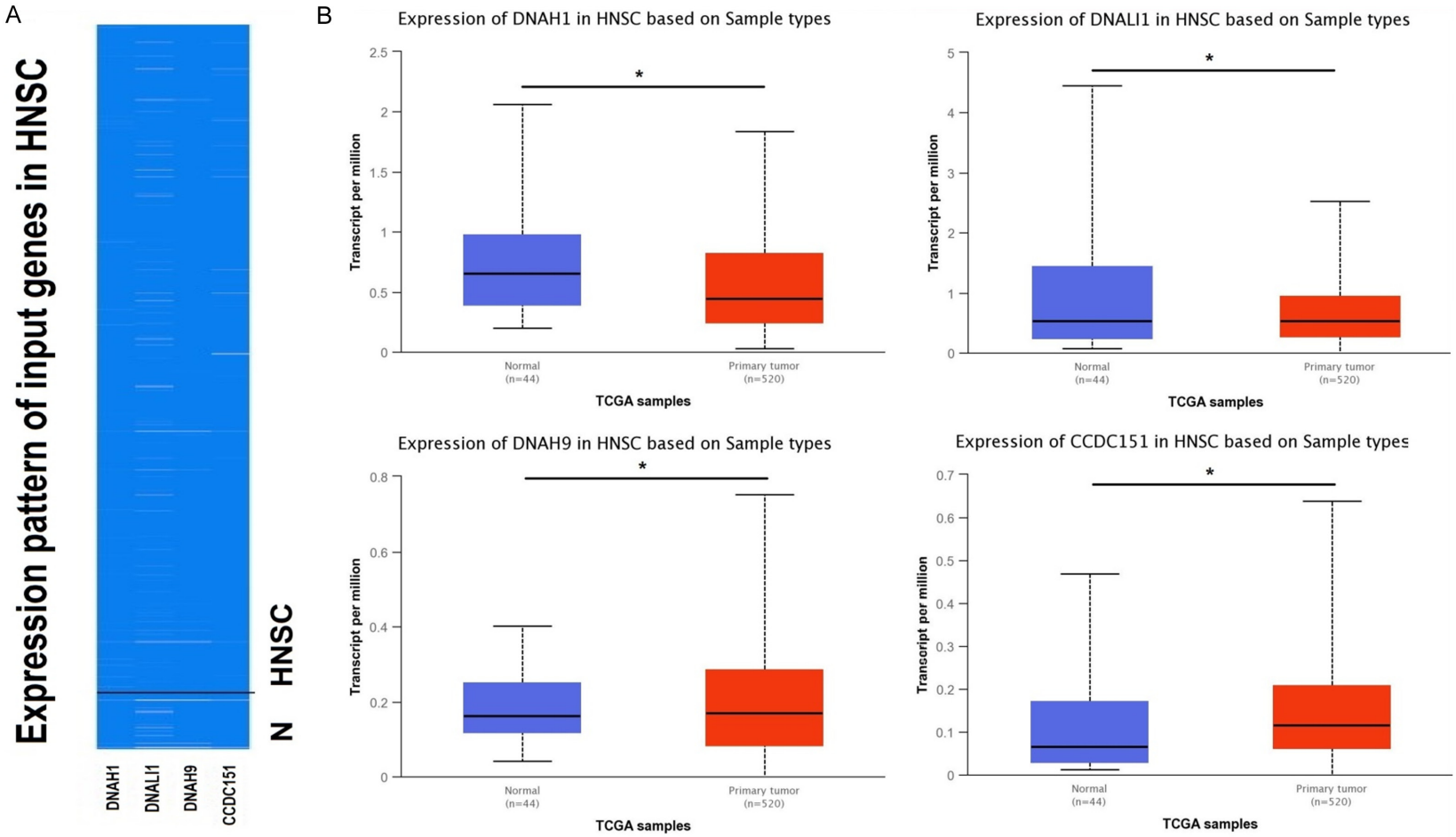
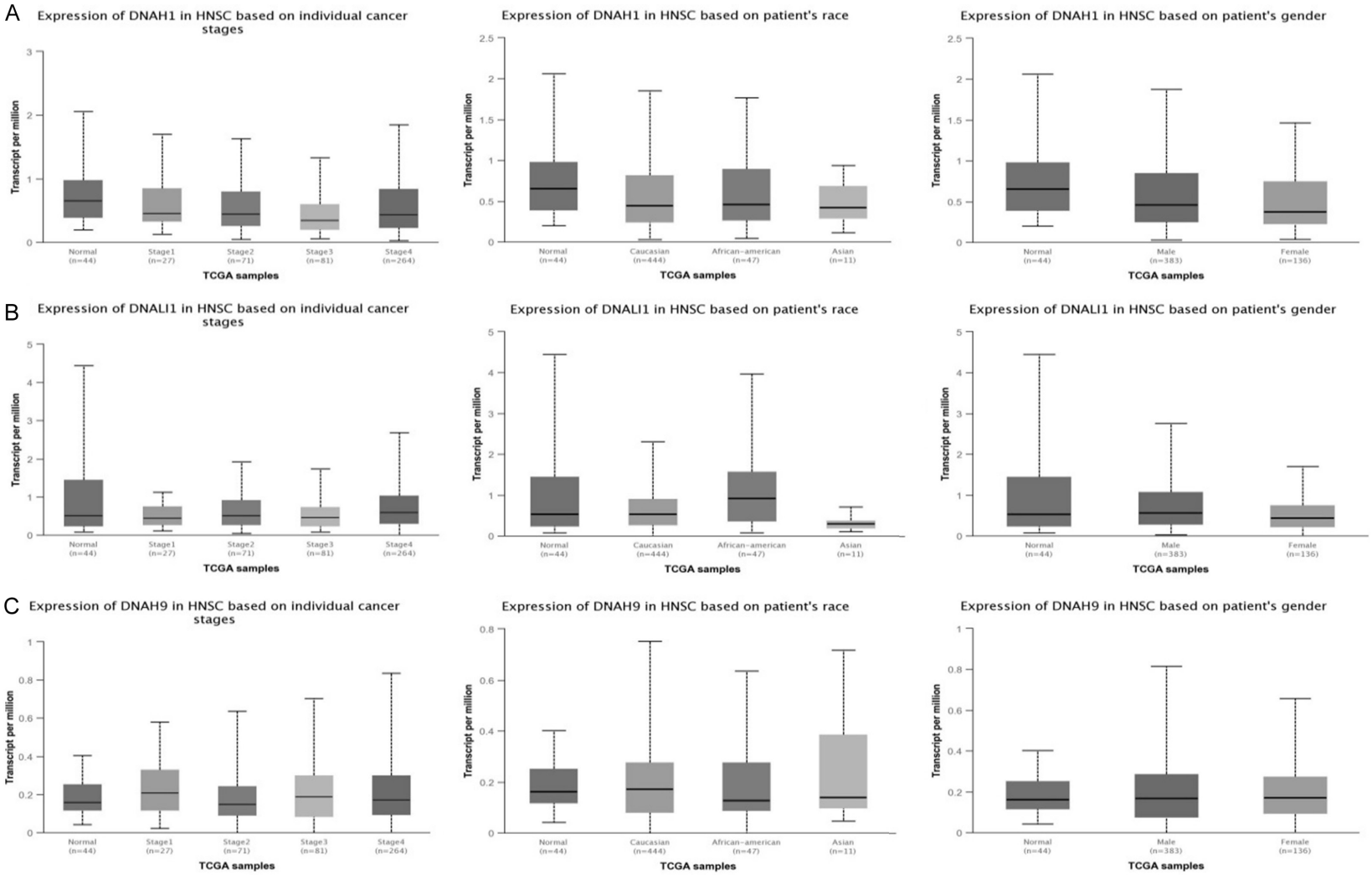


Figure 2. mRNA expression profiling of DNAH1, DNALI1, DNAH9, and CCDC151 via UALCAN. (A) A heatmap of DNAH1, DNALI1, DNAH9, and CCDC151 hub genes in HNSC sample group and normal control group, and (B) Box plot presentation of DNAH1, DNALI1, DNAH9, and CCDC151 hub genes mRNA expression in HNSC sample group and normal control group.


A
Expression of DNAH1 in HNSC based on individual cancer stages
Expression of DNAH1 in HNSC based on patient's race
Expression of DNAH1 in HNSC based on patient's gender
B
Expression of DNALI1 in HNSC based on individual cancer stages
Expression of DNALI1 in HNSC based on patient's race
Expression of DNALI1 in HNSC based on patient's gender
C
Expression of DNAH9 in HNSC based on individual cancer stages
Expression of DNAH9 in HNSC based on patient's race
Expression of DNAH9 in HNSC based on patient's gender
Transcript per million
TCGA samples
Normal (n=44)
Stage1 (n=27)
Stage2 (n=71)
Stage3 (n=81)
Stage4 (n=264)
Caucasian (n=444)
African-american (n=47)
Asian (n=11)
Male (n=383)
Female (n=136)

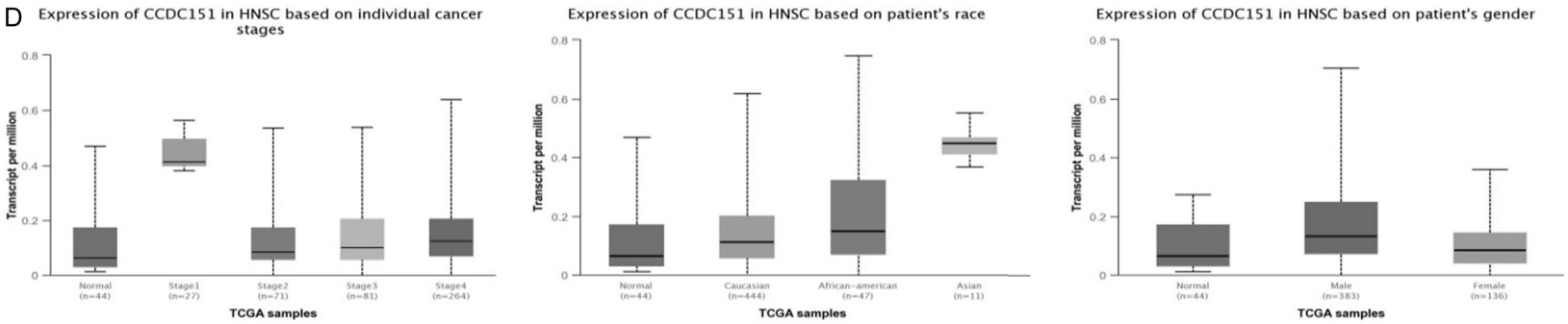


Figure 3. Expression profiling of DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC samples of different clinical variables relative to controls via UALCAN. (A) Expression profiling of DNAH1 in HNSC samples of different clinical variables, (B) Expression profiling of DNALI1 in HNSC samples of different clinical variables, (C) Expression profiling of DNAH9 in HNSC samples of different clinical variables, and (D) Expression profiling of CCDC151 in HNSC samples of different clinical variables.

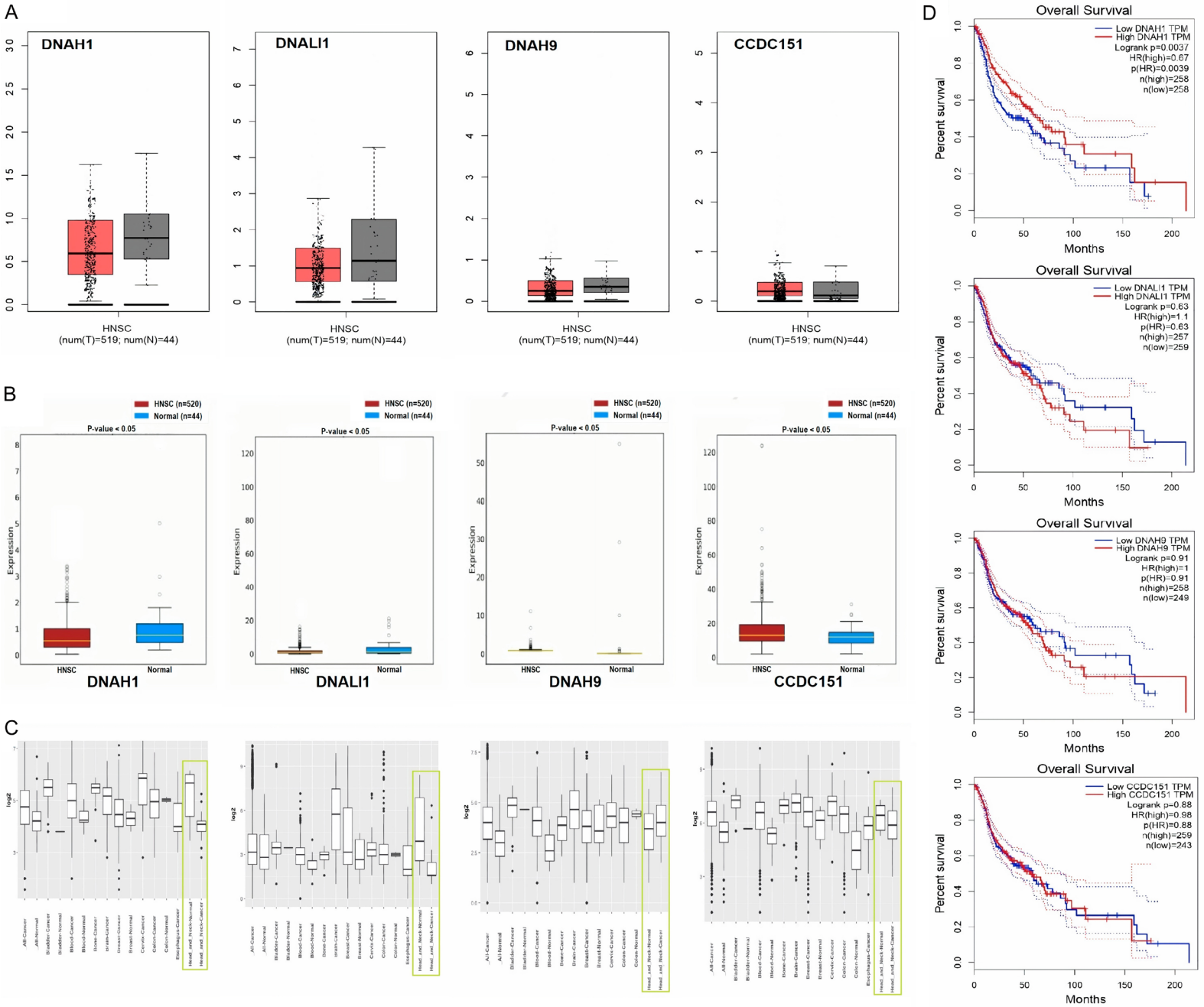

A
DNAH1
DNALI1
DNAH9
CCDC151
HNSC
(num(T)=519; num(N)=44)
B
HNSC (n=520)
Normal (n=44)
P-value < 0.05
Expression
HNSC
Normal
DNAH1
DNALI1
DNAH9
CCDC151
C
log2
D
Overall Survival
Low DNAH1 TPM
High DNAH1 TPM
Logrank p=0.0037
HR(high)=0.67
p(HR)=0.0039
n(high)=258
n(low)=258
Low DNALI1 TPM
High DNALI1 TPM
Logrank p=0.63
HR(high)=1.1
p(HR)=0.63
n(high)=257
n(low)=259
Low DNAH9 TPM
High DNAH9 TPM
Logrank p=0.91
HR(high)=1
p(HR)=0.91
n(high)=258
n(low)=249
Low CCDC151 TPM
High CCDC151 TPM
Logrank p=0.88
HR(high)=0.98
p(HR)=0.88
n(high)=259
n(low)=243
Percent survival
Months

Figure 4. Expression validation and survival analysis of DNAH1, DNALI1, DNAH9, and CCDC151. (A) Expression validation of DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC and normal samples via GEPIA database, (B) Expression validation of DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC and normal samples via OncoDB database, (C) Expression validation of DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC cell line via GENT2 database, and (D) Survival analysis of DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC and normal samples via GEPIA database.

ples via the MEXPRESS and OncoDB databases to understand the mechanisms behind the dysregulation of these genes. Results showed that DNAH1, DNALI1, DNAH9, and CCDC151 hub genes were aberrantly methylated, i.e., DNAH1 and DNALI1 were hypomethylated while DNAH9 and CCDC151 were hypermethylated. These aberrant methylation patterns of hub genes were extremely correlated with their expression levels in HNSC samples (**Figure 6**).

Hub genes-associated alterations, co-expressed genes analysis, and methylation level confirmation

Next, we further analyzed hub genes-associated alterations in HNSC samples via cBioPortal. All hub genes, including DNAH1, DNALI1, DNAH9, and CCDC151 were detected in this database. Altogether, 530 HNSC samples were analyzed with the help of cBioPortal, and the results of this analysis showed that DNAH9 was the most frequently altered (10% altered frequency) gene out of the analyzed four hub genes in HNSC samples (**Figure 7A**). In the analyzed HNSC samples, the altered frequencies of other hub genes including DNAH1, DNALI1, and CCDC151 were 3%, 1.2%, and 0.5%, respectively (**Figure 7A**). Furthermore, analysis from cBioPortal also revealed that HNSC patients with alterations in hub genes had worse OS and DFS than those HNSC patients who did not have any kind of alterations in the hub genes (**Figure 7B**). In terms of highly co-expressed genes with DNAH1, DNALI1, DNAH9, and CCDC151, it was noted that DNAH1-ZMAT1, DNALI1-SERPINF1, DNAH9-ZNF18, and CCDC151-CCDC74A were the top positively correlated genes in HNSC samples (**Figure 7C**). Ultimately, methylation level confirmation analysis from the cBioPortal database confirmed that there were substantial negative correlations exist between hub genes expressions and their promoter methylation levels across HNSC samples (**Figure 7D**).

Gene enrichment analysis

GO enrichment analysis showed that DNAH1, DNALI1, DNAH9, and CCDC151 hub genes along with their co-expressed genes were enriched in CC, including "Inner dynein arm, Axonemal dynein complex, Outer dynein arm, Dynein axonemal particle, Dynein complex, and Axoneme" (Supplementary Figure 2A). MF analysis showed that DNAH1, DNALI1, DNAH9, and CCDC151 hub genes along with their co-expressed genes were enriched in "Minus-end-directed microtubule activity, Dynein heavy chain binding, Dynein intermediate chain binding, Microtubule activity, and Cytoskeletal activity" (Supplementary Figure 2B). BP analysis showed that these genes were involved in "Inner dynein arm assembly, Axonemal dynein arm assembly, Outer dynein arm assembly, and Sperm axoneme assembly" (Supplementary Figure 2C). Pathway analysis revealed that these genes were the part of "Hungtington disease, Amyotrophic lateral sclerosis, and pathways of neurodegeneration" pathways (Supplementary Figure 2D).

Immune infiltration analysis

Infiltrating immune cells are the core components of the tumor microenvironment and, therefore, may contribute to the development of cancer [29]. Through TIMER database, we revealed the associations among DNAH1, DNALI1, DNAH9, and CCDC151 hub genes expression and CD8+ T, CD4+ T, and Macrophages immune cell populations in HNSC patients. Results highlighted that CD8+ T immune cell population has significant negative correlation with DNAH1, DNALI1, DNAH9, and CCDC151 hub genes expression, while CD4+ T, and Macrophages immune cell populations have significant positive correlations with DNAH1, DNALI1, DNAH9, and CCDC151 hub genes expressions in HNSC samples (Supplementary Figure 3).

lncRNA-miRNA-mRNA interaction network

Via ENCORI and Cytoscape, we constructed the lncRNA-miRNA-mRNA co-regulatory networks of DNAH1, DNALI1, DNAH9, and CCDC151. In the constructed networks, the total count of lncRNAs, miRNAs, and mRNAs were 56, 54, and 4, respectively (Supplementary Figure 4).

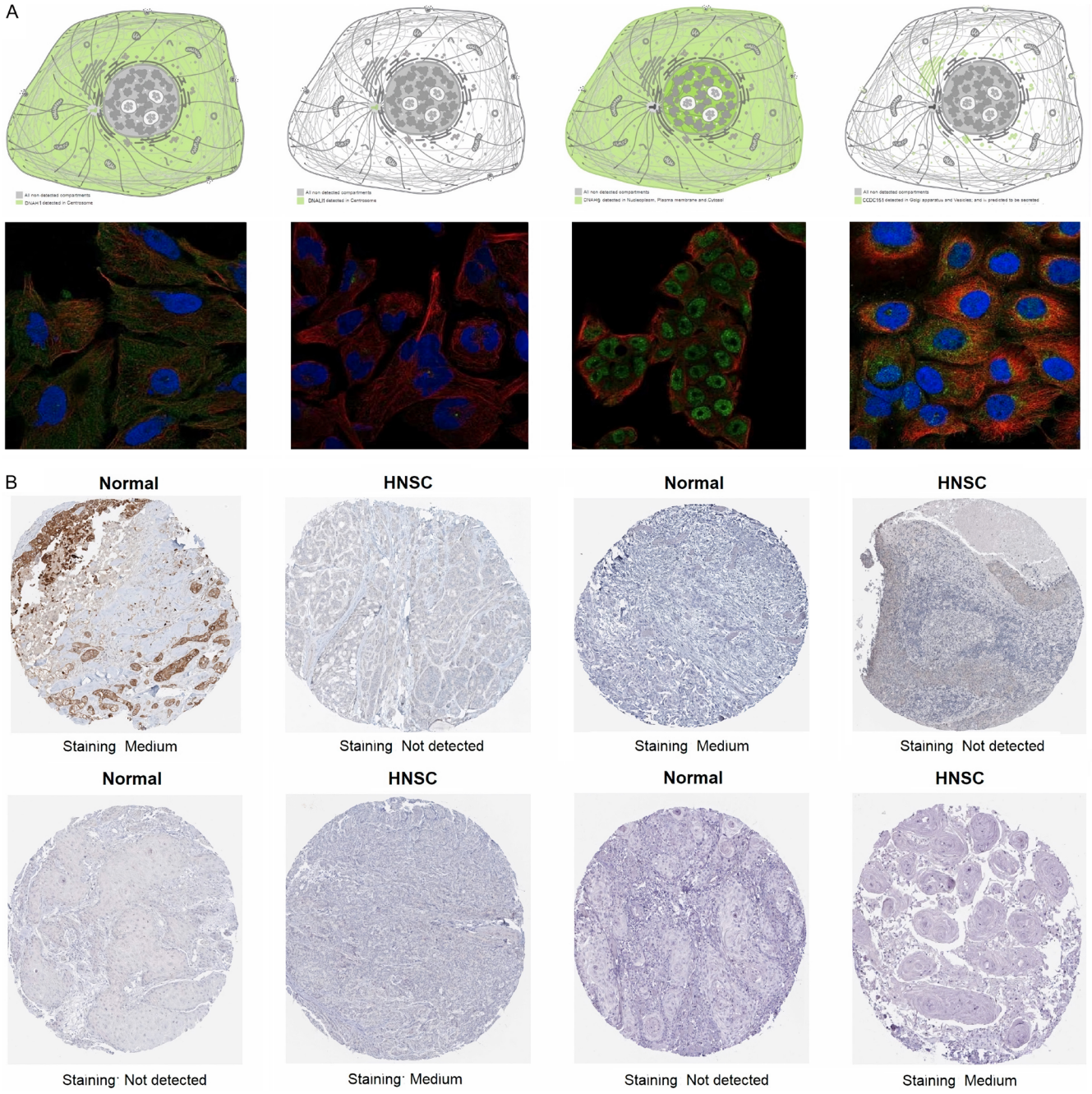

A
B
Normal
HNSC
Normal
HNSC
Staining Medium
Staining Not detected
Staining Medium
Staining Not detected
Normal
HNSC
Normal
HNSC
Staining· Not detected
Staining· Medium
Staining Not detected
Staining Medium

Figure 5. Subcellular localization and protein expression validation of DNAH1, DNALI1, DNAH9, and CCDC151 via HPA database. (A) Subcellular localization prediction of DNAH1, DNALI1, DNAH9, and CCDC151, and (B) Protein expression validation of DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC and normal samples.

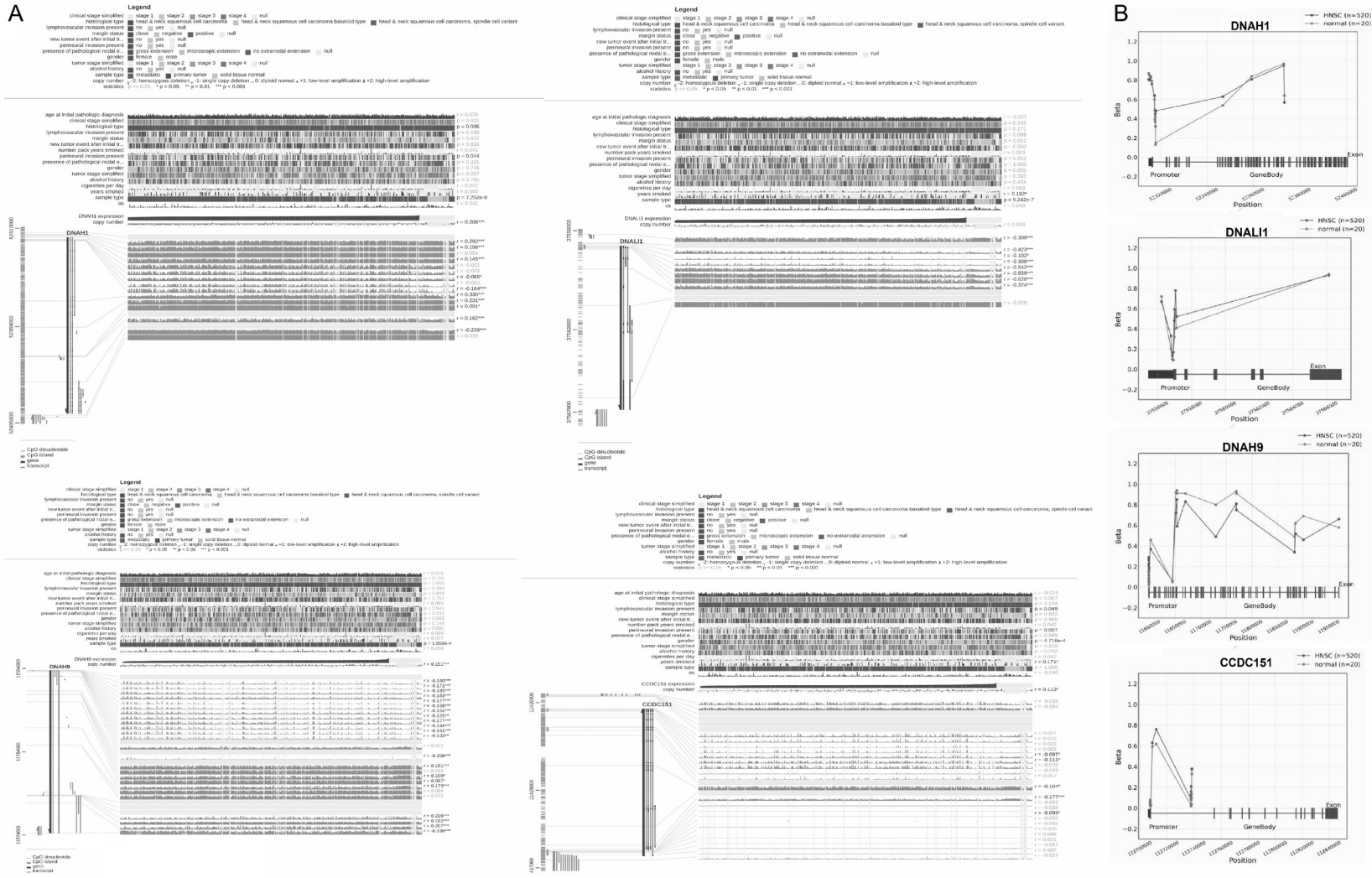


Figure 6. Methylation status exploration of DNAH1, DNALI1, DNAH9, and CCDC151 via MEXPRESS and OncoDB in HNSC and normal samples. (A) Methylation status exploration of DNAH1, DNALI1, DNAH9, and CCDC151 via MEXPRESS, and (B) Methylation status exploration of DNAH1, DNALI1, DNAH9, and CCDC151 via OncoDB.

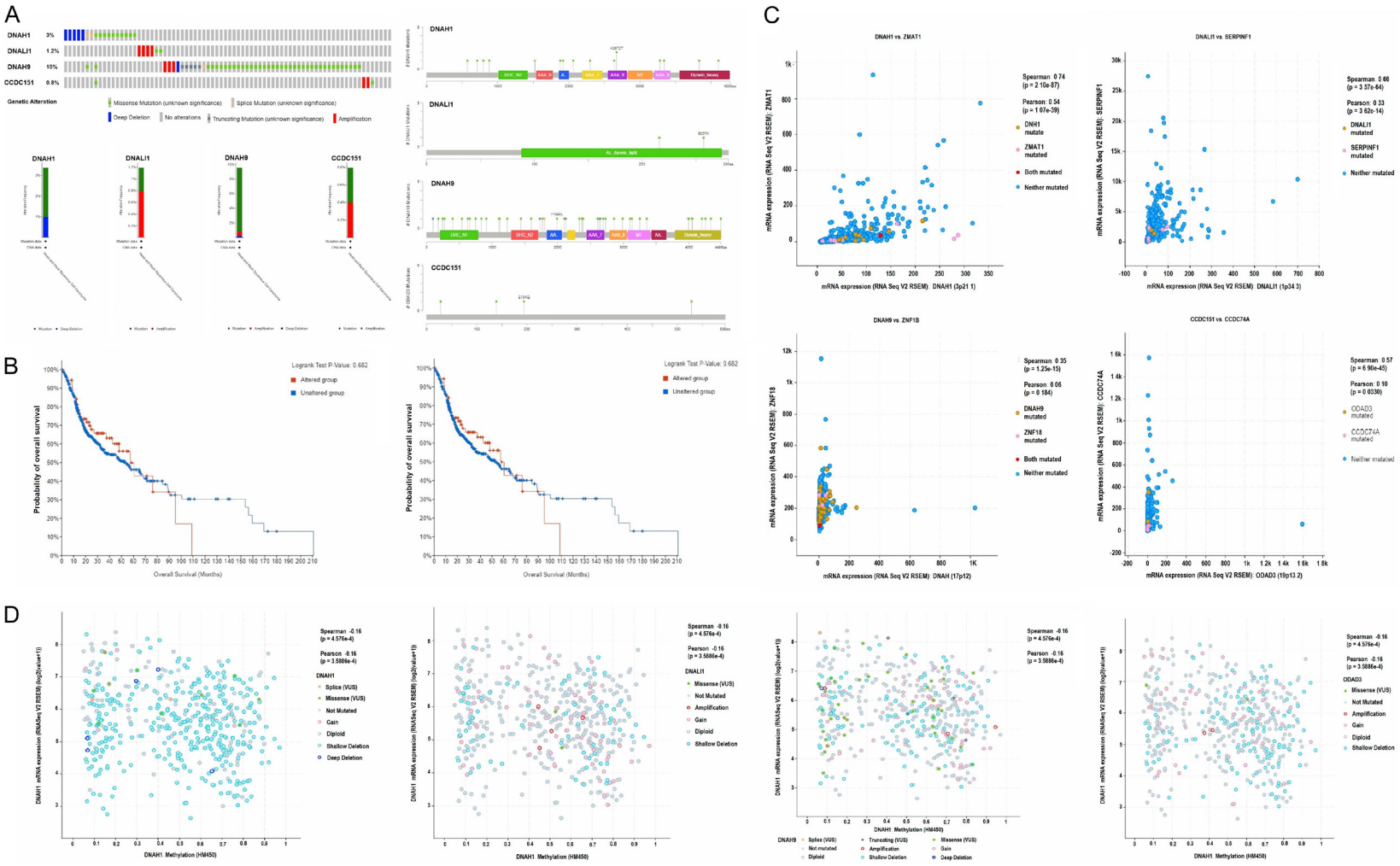


Figure 7. Exploration of genetic alteration frequencies, mutational hotspots, OS, DFS analyses, co-expressed genes, and methylatiom analysis of DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC samples via cBioPortal. (A) Types, frequencies, and location of the genetic alterations in DNAH1, DNALI1, DNAH9, and CCDC151, (B) OS and DFS analysis of DNAH1, DNALI1, DNAH9, and CCDC151 in genetically altered and unaltered HNSC groups, (C) Identification of co-expressed genes with DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC samples, and (D) Promoter methylation analysis of DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC samples.

Table 1. DrugBank-based hub genes-associated drugs

Sr. No	Hub gene	Drug name	Effect	Reference	Group
1	DNAH1	Acetaminophen	Increase expression of DNAH1 mRNA	A20420	Approved
		Acetylcysteine		A20451	
2	DNALI1	Valproic acid	Increase expression of DNALI1 mRNA	A24688	Approved
3	DNAH9	Valproic acid	Decrease expression of DNAH9 mRNA	A24688	Approved
		Genistein		A21119	
4	CCDC151	Valproic acid	Decrease expression of CCDC151 mRNA	A24688	Approved
		Estradiol		A21424	

Based on the constructed networks, we have identified one miRNA (has-let-7b-5p), that targets all hub genes simultaneously. Therefore, we speculate that the identified lncRNAs, has-let-7b-5p, and hub genes (DNAH1, DNALI1, DNAH9, and CCDC151) (Supplementary Figure 4) as an axis, might also be the potential inducers of the HNSC.

Drug prediction analysis

A selection of suitable candidate drugs for reversing DNAH1, DNALI1, DNAH9, and CCDC151 hub genes expression was made by querying the DrugBank database. After the comprehensive search, we explored some suitable therapeutic drugs for the treatment of HNSC. For example, Acetaminophen and Valproic acid drugs were identified as the positive expression regulators of DNAH1 mRNA expression (**Table 1**) while Valproic acid was identified as the negative expression regulator of DNAH1 mRNA expression (**Table 1**).

Experimental in vitro validation of the hub genes expression

In the current study, through RNA-seq analysis of two HNSC (FaDu, and SCC25) and one normal human oral keratinocyte (HOK) cell line, the expression levels of identified four hub genes were validated. The expression levels of these genes were validated using FPKM, which is a quantitative value with widespread use in the RNA-seq analysis. As shown in **Figure 8**, it was noticed that DNAH1, DNALI1, DNAH9, and CCDC151 hub genes were expressed in all three analyzed cell lines, and RPKM values of DNAH1 and DNALI1 were significantly lower while FPKM values of DNAH9 and CCDC151 were significantly higher in HNSC cell lines (FaDu and SCC25) as compared to normal cell line (HOK) (**Figure 8**).

Discussion

The aim of the current research was to explore HNSC involving hub genes and a variety of their associated parameters by analyzing and comparing the expression profiles of tumor and normal tissues in the GSE53819 dataset. Finally, based on the degree-score method, a total of four genes, namely DNAH1, DNALI1, DNAH9, and CCDC151 were identified as the hub genes between HNSC and normal tissue samples. Expression validation analysis at both mRNA and protein levels confirmed that DNAH1 and DNALI1 were significantly down-regulated while DNAH9 and CCDC151 were significantly up-regulated in HNSC samples relative to normal controls. Moreover, expression of these genes was also found to be associated with a worse prognosis of the HNSC patients.

The DNAH1 gene, that belongs to axonemal dynein gene family encodes for the DNAH1 protein [30]. This protein plays an important role in the human reproductive system [31]. Genetic alterations in DNAH1 gene were initially reported in different diseases, including ciliary dyskinesia disease [32, 33], sperm immobility [34] and a few more diseases caused by cilia dysfunction. Recently, genetic alterations (mutations and down-regulation) in the DNAH1 gene have been frequently observed among various types of malignant cancers, including breast cancer [35], colorectal cancer [35], and cervical cancer [36]. Moreover, mutations in the DNAH1 gene were also correlated with chemotherapy resistance in gastric cancer patients [37].

The DNALI1 gene is another important member of the axonemal dynein gene family [38]. The precise function of DNALI1 gene is still unclear [39]. However, this gene is thought to be responsible for the cilium movement [38]. Cilia

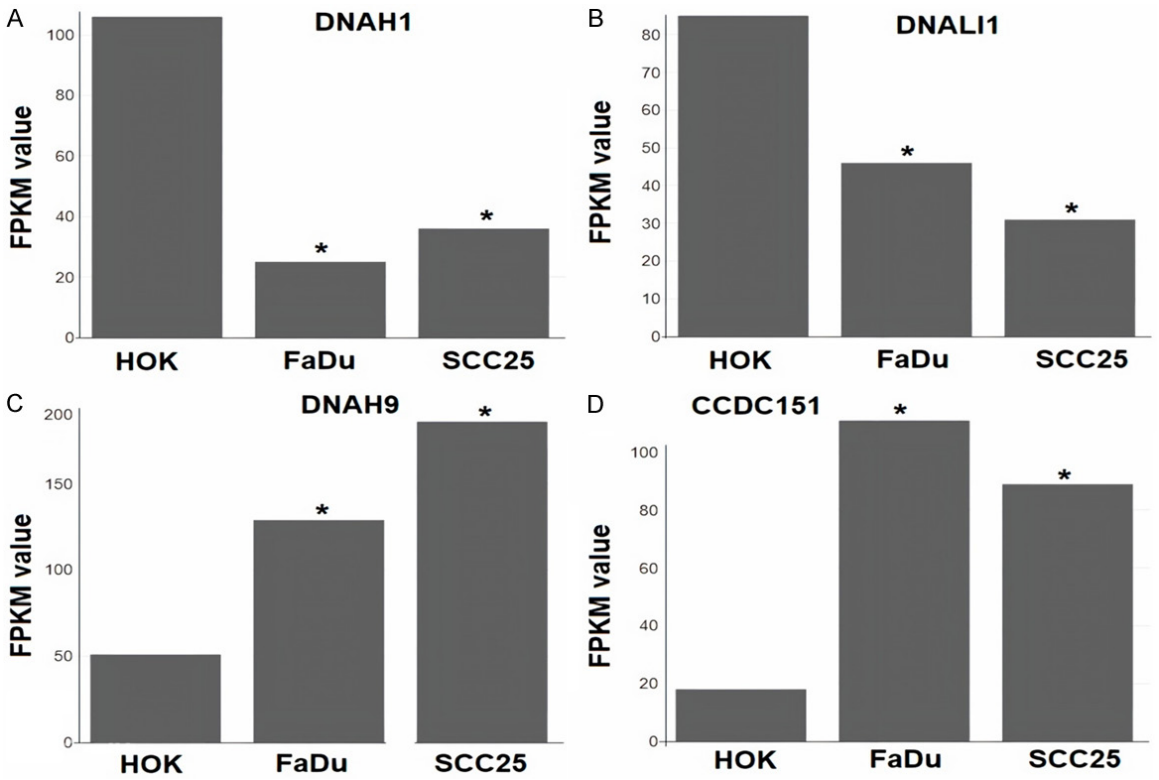


Figure 8. Validating DNAH1, DNALI1, DNAH9, and CCDC151 expressions using HOK, FaDu, and SCC25 cell lines via RNA-seq analysis. (A) DNAH1, (B) DNALI1, (C) DNAH9, and (D) CCDC151.

are the important little hair-like structures, that found on the surface of the cell and perform a variety of critical tasks, including the removal of unwanted inhaled particles and germs in the nose, ears, and lungs etc. [40]. Earlier, the genetic mutations and down-regulation of DNALI1 have been reported as the causative factors of the primary ciliary dyskinesia (PCD) disease, which cause defects in cilia development and infection of airways [41]. However, recent studies also reported the tumor suppressor function of the DNALI1 across different types of cancers [42]. The significant down-regulation of DNALI1 and its association with the poor survival of breast cancer patients was reported by *Parris et al.* [42]. The role of DNLI1 is still unknown in the HNSC development.

The DNAH9 protein in humans is encoded by the DNAH9 gene [43]. This protein is involved to medicate the movement of cilia on the cell surface [44]. There are several reports available in the medical literature, which highlighted the tumor-causing role of DNAH9. For example, a study by *Gruel et al.,* revealed that DNAH9 harbors genetic alterations such as missense mutations and overexpressed in breast cancer patients [45]. Similarly, another report by *Martini et al.,* has also highlighted the tumor-promoting role of DNAH9 overexpression in colorectal and cervical cancer patients [46].

The CCDC151 gene encodes for a protein that is an integral part of the ODA-complex assembly and is also essential to medicate the movement of cilia on the cell surface [47]. According to earlier studies, the genetic mutations and expression variations in the CCDC151 gene, especially nonsense mutations, can cause PCD disease, which is a very complex disease to treat [48-50]. Moreover, recent studies revealed the overexpression and functional loss of the CCDC151 gene in a variety of human cancers, including breast cancer [51], prostate cancer [52], and gastric cancer [51, 53]. In this study, we observed that the mRNA and protein expressions of DNAH1, DNALI1, DNAH9, and CCDC151 were significantly dysregulated and associated with the different clinicopathological parameters of the HNSC patients. Earlier, the relationship between the identified hub genes and HNSC had not been reported in the

medical literature, however, the findings of the current study indicate that the dysregulation of DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC samples may be novel biomarkers for the early detection of HNSC.

We also performed the promoter methylation and alteration analysis of the DNAH1, DNALI1, DNAH9, and CCDC151 genes' in HNSC patients. Results showed that the promoter methylation level of these genes had negative correlations with their expression levels. Moreover, the results of the alteration revealed that DNAH9 was the most frequently altered (10%) gene out of the four hub genes in HNSC samples. To the best of our knowledge, we are the first to report the relevancy of the DNAH1, DNALI1, DNAH9, and CCDC151 genes' expression, promoter methylation, and genetic alterations with the development and progression of HNSC.

Pathways analysis showed that DNAH1, DNALI1, DNAH9, and CCDC151 hub genes were part of the "Hungtington disease, Amyotrophic lateral sclerosis, and pathways of neurodegeneration" pathways in HNSC patients. The oncogenic roles of these pathways have earlier been well-acknowledged in cancer development [54-56]. We further confirmed that DNAH1, DNALI1, DNAH9, and CCDC151 hub genes' expression were regulated simultaneously by hsa-let-7b-5p miRNA in HNSC patients, and expressions of these genes were significantly related to the immune cell infiltration of CD8+ T, CD4+ T, and macrophages. Previously, the dysregulation of hsa-let-7b-5p in multiple human cancers has been reported in published studies, for example in breast cancer [57, 58], bladder cancer [59], glioblastoma [57, 60], and esophageal cancer [61]. However, any tumor suppressor or tumor-causing role of hsa-let-7b-5p with respect to DNAH1, DNALI1, DNAH9, and CCDC151 in HNSC is not reported anywhere. Therefore, to the best of our knowledge, this study is the first to report the probable cancer-driving role of the hsa-let-7b-5p miRNA with respect to DNAH1, DNALI1, DNAH9, and CCDC151 hub genes in HNSC.

Conclusion

In conclusion, this study identified and analyzed four core genes (DNAH1, DNALI1, DNAH9, and CCDC151) associated with the development and progression of HNSC, which could help us better understand the carcinogenesis process and provide indicators for prognosis and early detection of the disease.

Acknowledgements

The authors extend their appreciation to the Researchers Supporting Project number (RSPD2023R725) King Saud University, Riyadh, Saudi Arabia.

Disclosure of conflict of interest

None.

Address correspondence to: Mostafa A Abdel-Maksoud, Department of Botany and Microbiology, College of Science, King Saud University Riyadh, P.O. 2455, Riyadh 11451, Saudi Arabia. E-mail: mabdelmaksoud@ksu.edu.sa; Jaweria Gul, Department of Biotechnology, Shaheed Benazir Bhutto University, Sheringal, Dir Upper, Pakistan. E-mail: jaweria_hassan16@yahoo.com

References

[1] Solomon B, Young RJ and Rischin D. Head and neck squamous cell carcinoma: genomics and emerging biomarkers for immunomodulatory cancer treatments. Semin Cancer Biol 2018; 52: 228-240.

[2] Vokes EE, Weichselbaum RR, Lippman SM and Hong WK. Head and neck cancer. N Engl J Med 1993; 328: 184-194.

[3] Plath M, Gass J, Hlevnjak M, Li Q, Feng B, Hostench XP, Bieg M, Schroeder L, Holzinger D, Zapatka M, Freier K, Weichert W, Hess J and Zaoui K. Unraveling most abundant mutational signatures in head and neck cancer. Int J Cancer 2021; 148: 115-127.

[4] Lopes-Ramos CM, Quackenbush J and DeMeo DL. Genome-wide sex and gender differences in cancer. Front Oncol 2020; 10: 597788.

[5] Chai RC, Lambie D, Verma M and Punyadeera C. Current trends in the etiology and diagnosis of HPV-related head and neck cancers. Cancer Med 2015; 4: 596-607.

[6] Vogelstein B and Kinzler KW. Cancer genes and the pathways they control. Nat Med 2004; 10: 789-799.

[7] Pisani P, Airoldi M, Allais A, Aluffi Valletti P, Battista M, Benazzo M, Briatore R, Cacciola S, Cocuzza S, Colombo A, Conti B, Costanzo A, Della Vecchia L, Denaro N, Fantozzi C, Galizia D, Garzaro M, Genta I, Iasi GA, Krengli M, Landolfo V, Lanza GV, Magnano M, Mancuso M, Maroldi R,

Masini L, Merlano MC, Piemonte M, Pisani S, Prina-Mello A, Prioglio L, Rugiu MG, Scasso F, Serra A, Valente G, Zannetti M and Zigliani A. Metastatic disease in head & neck oncology. Acta Otorhinolaryngol Ital 2020; 40 Suppl 1: S1-S86.

[8] Bao YN, Cao X, Luo DH, Sun R, Peng LX, Wang L, Yan YP, Zheng LS, Xie P, Cao Y, Liang YY, Zheng FJ, Huang BJ, Xiang YQ, Lv X, Chen QY, Chen MY, Huang PY, Guo L, Mai HQ, Guo X, Zeng YX and Qian CN. Urokinase-type plasminogen activator receptor signaling is critical in nasopharyngeal carcinoma cell growth and metastasis. Cell Cycle 2014; 13: 1958-1969.

[9] Clough E and Barrett T. The gene expression omnibus database. Methods Mol Biol 2016; 1418: 93-110.

[10] 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.

[11] von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P and Snel B. STRING: a database of predicted functional associations between proteins. Nucleic Acids Res 2003; 31: 258-261.

[12] Demchak B, Hull T, Reich M, Liefeld T, Smoot M, Ideker T and Mesirov JP. Cytoscape: the network visualization tool for GenomeSpace workflows. F1000Res 2014; 3: 151.

[13] Bandettini WP, Kellman P, Mancini C, Booker OJ, Vasu S, Leung SW, Wilson JR, Shanbhag SM, Chen MY and Arai AE. MultiContrast Delayed Enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study. J Cardiovasc Magn Reson 2012; 14: 83.

[14] 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.

[15] Pan X, Chen S, Chen X, Ren Q, Yue L, Niu S, Li Z, Zhu R, Chen X, Jia Z, Zhen R and Ban J. UTP14A, DKC1, DDX10, PinX1, and ESF1 modulate cardiac angiogenesis leading to obesity-induced cardiac injury. J Diabetes Res 2022; 2022: 2923291.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] Thul PJ and Lindskog C. The human protein atlas: a spatial map of the human proteome. Protein Sci 2018; 27: 233-244.

[23] 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.

[24] Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B and Liu XS. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Res 2017; 77: e108-e110.

[25] 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.

[26] 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.

[27] Kim TK. T test as a parametric statistic. Korean J Anesthesiol 2015; 68: 540-546.

[28] Kim HY. Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. Restor Dent Endod 2017; 42: 152-155.

[29] Whiteside TL. The tumor microenvironment and its role in promoting tumor growth. Oncogene 2008; 27: 5904-5912.

[30] Ben Khelifa M, Coutton C, Zouari R, Karaouzène T, Rendu J, Bidart M, Yassine S, Pierre V, Delaroche J, Hennebicq S, Grunwald D, Escalier D, Pernet-Gallay K, Jouk PS, Thierry-Mieg N, Touré A, Arnoult C and Ray PF. Mutations in DNAH1, which encodes an inner arm heavy chain dynein, lead to male infertility from multiple morphological abnormalities of the sperm flagella. Am J Hum Genet 2014; 94: 95-104.

[31] Khan R, Zaman Q, Chen J, Khan M, Ma A, Zhou J, Zhang B, Ali A, Naeem M, Zubair M, Zhao D,

Shah W, Khan M, Zhang Y, Xu B, Zhang H and Shi Q. Novel loss-of-function mutations in DNAH1 displayed different phenotypic spectrum in humans and mice. Front Endocrinol (Lausanne) 2021; 12: 765639.

[32] Knowles MR, Leigh MW, Carson JL, Davis SD, Dell SD, Ferkol TW, Olivier KN, Sagel SD, Rosenfeld M, Burns KA, Minnix SL, Armstrong MC, Lori A, Hazucha MJ, Loges NT, Olbrich H, Becker-Heck A, Schmidts M, Werner C, Omran H and Zariwala MA; Genetic Disorders of Mucociliary Clearance Consortium. Mutations of DNAH11 in patients with primary ciliary dyskinesia with normal ciliary ultrastructure. Thorax 2012; 67: 433-441.

[33] Li Y, Yagi H, Onuoha EO, Damerla RR, Francis R, Furutani Y, Tariq M, King SM, Hendricks G, Cui C, Saydmohammed M, Lee DM, Zahid M, Sami I, Leatherbury L, Pazour GJ, Ware SM, Nakanishi T, Goldmuntz E, Tsang M and Lo CW. DNAH6 and its interactions with PCD genes in heterotaxy and primary ciliary dyskinesia. PLoS Genet 2016; 12: e1005821.

[34] Ben Khelifa M, Coutton C, Zouari R, Karaouzène T, Rendu J, Bidart M, Yassine S, Pierre V, Delaroche J, Hennebicq S, Grunwald D, Escalier D, Pernet-Gallay K, Jouk PS, Thierry-Mieg N, Touré A, Arnoult C and Ray PF. Mutations in DNAH1, which encodes an inner arm heavy chain dynein, lead to male infertility from multiple morphological abnormalities of the sperm flagella. Am J Hum Genet 2014; 94: 95-104.

[35] Matsuda S, Ueda M, Hashimoto O, Itano O, Hasegawa H and Kitagawa Y. Detection of DNAH1 fragment closely related to human colorectal cancer by RLGS and functional analysis of DNAH1 in cancer cells. Cancer Res 2011; 71: 4814.

[36] Wen Y, Zhang S, Yang J and Guo D. Identification of driver genes regulating immune cell infiltration in cervical cancer by multiple omics integration. Biomed Pharmacother 2019; 120: 109546.

[37] Zhu C, Yang Q, Xu J, Zhao W, Zhang Z, Xu D, Zhang Y, Zhao E and Zhao G. Somatic mutation of DNAH genes implicated higher chemotherapy response rate in gastric adenocarcinoma patients. J Transl Med 2019; 17: 109.

[38] Viswanadha R, Sale WS and Porter ME. Ciliary motility: regulation of axonemal dynein motors. Cold Spring Harb Perspect Biol 2017; 9: a018325.

[39] Yamamoto R, Hwang J, Ishikawa T, Kon T and Sale WS. Composition and function of ciliary inner-dynein-arm subunits studied in chlamydomonas reinhardtii. Cytoskeleton (Hoboken) 2021; 78: 77-96.

[40] Shapiro A, Davis S, Manion M and Briones K. Primary Ciliary Dyskinesia (PCD). Am J Respir Crit Care Med 2018; 198: P3-P4.

[41] Lucas JS, Barbato A, Collins SA, Goutaki M, Behan L, Caudri D, Dell S, Eber E, Escudier E, Hirst RA, Hogg C, Jorissen M, Latzin P, Legendre M, Leigh MW, Midulla F, Nielsen KG, Omran H, Papon JF, Pohunek P, Redfern B, Rigau D, Rindlisbacher B, Santamaria F, Shoemark A, Snijders D, Tonia T, Titieni A, Walker WT, Werner C, Bush A and Kuehni CE. European respiratory society guidelines for the diagnosis of primary ciliary dyskinesia. Eur Respir J 2017; 49: 1601090.

[42] 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.

[43] Loges NT, Antony D, Maver A, Deardorff MA, Güleç EY, Gezdirici A, Nöthe-Menchen T, Höben IM, Jelten L, Frank D, Werner C, Tebbe J, Wu K, Goldmuntz E, Čuturilo G, Krock B, Ritter A, Hjeij R, Bakey Z, Pennekamp P, Dworniczak B, Brunner H, Peterlin B, Tanidir C, Olbrich H, Omran H and Schmidts M. Recessive DNAH9 loss-of-function mutations cause laterality defects and subtle respiratory ciliary-beating defects. Am J Hum Genet 2018; 103: 995-1008.

[44] Bartoloni L, Blouin JL, Maiti AK, Sainsbury A, Rossier C, Gehrig C, She JX, Marron MP, Lander ES, Meeks M, Chung E, Armengot M, Jorissen M, Scott HS, Delozier-Blanchet CD, Gardiner RM and Antonarakis SE. Axonemal beta heavy chain dynein DNAH9: cDNA sequence, genomic structure, and investigation of its role in primary ciliary dyskinesia. Genomics 2001; 72: 21-33.

[45] Gruel N, Benhamo V, Bhalshankar J, Popova T, Fréneaux P, Arnould L, Mariani O, Stern MH, Raynal V, Sastre-Garau X, Rouzier R, Delattre O and Vincent-Salomon A. Polarity gene alterations in pure invasive micropapillary carcinomas of the breast. Breast Cancer Res 2014; 16: R46.

[46] Hassan Ibrahim I, Balah A, Gomaa Abd Elfattah Hassan A and Gamal Abd El-Aziz H. Role of motor proteins in human cancers. Saudi J Biol Sci 2022; 29: 103436.

[47] Jerber J, Baas D, Soulavie F, Chhin B, Cortier E, Vesque C, Thomas J and Durand B. The coiled-coil domain containing protein CCDC151 is required for the function of IFT-dependent motile cilia in animals. Hum Mol Genet 2014; 23: 563-577.

[48] Hjeij R, Onoufriadis A, Watson CM, Slagle CE, Klena NT, Dougherty GW, Kurkowiak M, Loges NT, Diggle CP, Morante NF, Gabriel GC, Lemke KL, Li Y, Pennekamp P, Menchen T, Konert F, Marthin JK, Mans DA, Letteboer SJ, Werner C, Burgoyne T, Westermann C, Rutman A, Carr IM, O'Callaghan C, Moya E, Chung EM; UK10K

Consortium, Sheridan E, Nielsen KG, Roepman R, Bartscherer K, Burdine RD, Lo CW, Omran H and Mitchison HM. CCDC151 mutations cause primary ciliary dyskinesia by disruption of the outer dynein arm docking complex formation. Am J Hum Genet 2014; 95: 257-274.

[49] Alsaadi MM, Erzurumluoglu AM, Rodriguez S, Guthrie PA, Gaunt TR, Omar HZ, Mubarak M, Alharbi KK, Al-Rikabi AC and Day IN. Nonsense mutation in coiled-coil domain containing 151 gene (CCDC151) causes primary ciliary dyskinesia. Hum Mutat 2014; 35: 1446-1448.

[50] Zhang W, Li D, Wei S, Guo T, Wang J, Luo H, Yang Y and Tan Z. Whole-exome sequencing identifies a novel CCDC151 mutation, c.325G>T (p.E109X), in a patient with primary ciliary dyskinesia and situs inversus. J Hum Genet 2019; 64: 249-252.

[51] Hu Y, He C, Liu JP, Li NS, Peng C, Yang-Ou YB, Yang XY, Lu NH and Zhu Y. Analysis of key genes and signaling pathways involved in Helicobacter pylori-associated gastric cancer based on The Cancer Genome Atlas database and RNA sequencing data. Helicobacter 2018; 23: e12530.

[52] Chen R, Li Y, Buttyan R and Dong X. Implications of PI3K/AKT inhibition on REST protein stability and neuroendocrine phenotype acquisition in prostate cancer cells. Oncotarget 2017; 8: 84863-84876.

[53] Wang G, Jia Y, Ye Y, Kang E, Chen H, Wang J and He X. Identification of key methylation differentially expressed genes in posterior fossa ependymoma based on epigenomic and transcriptome analysis. J Transl Med 2021; 19: 174.

[54] Sørensen SA, Fenger K and Olsen JH. Significantly lower incidence of cancer among patients with Huntington disease: an apoptotic effect of an expanded polyglutamine tract? Cancer 1999; 86: 1342-1346.

[55] Freedman DM, Curtis RE, Daugherty SE, Goedert JJ, Kuncl RW and Tucker MA. The association between cancer and amyotrophic lateral sclerosis. Cancer Causes Control 2013; 24: 55-60.

[56] Hawer H, Hammermeister A, Ravichandran KE, Glatt S, Schaffrath R and Klassen R. Roles of elongator dependent tRNA modification pathways in neurodegeneration and cancer. Genes (Basel) 2018; 10: 19.

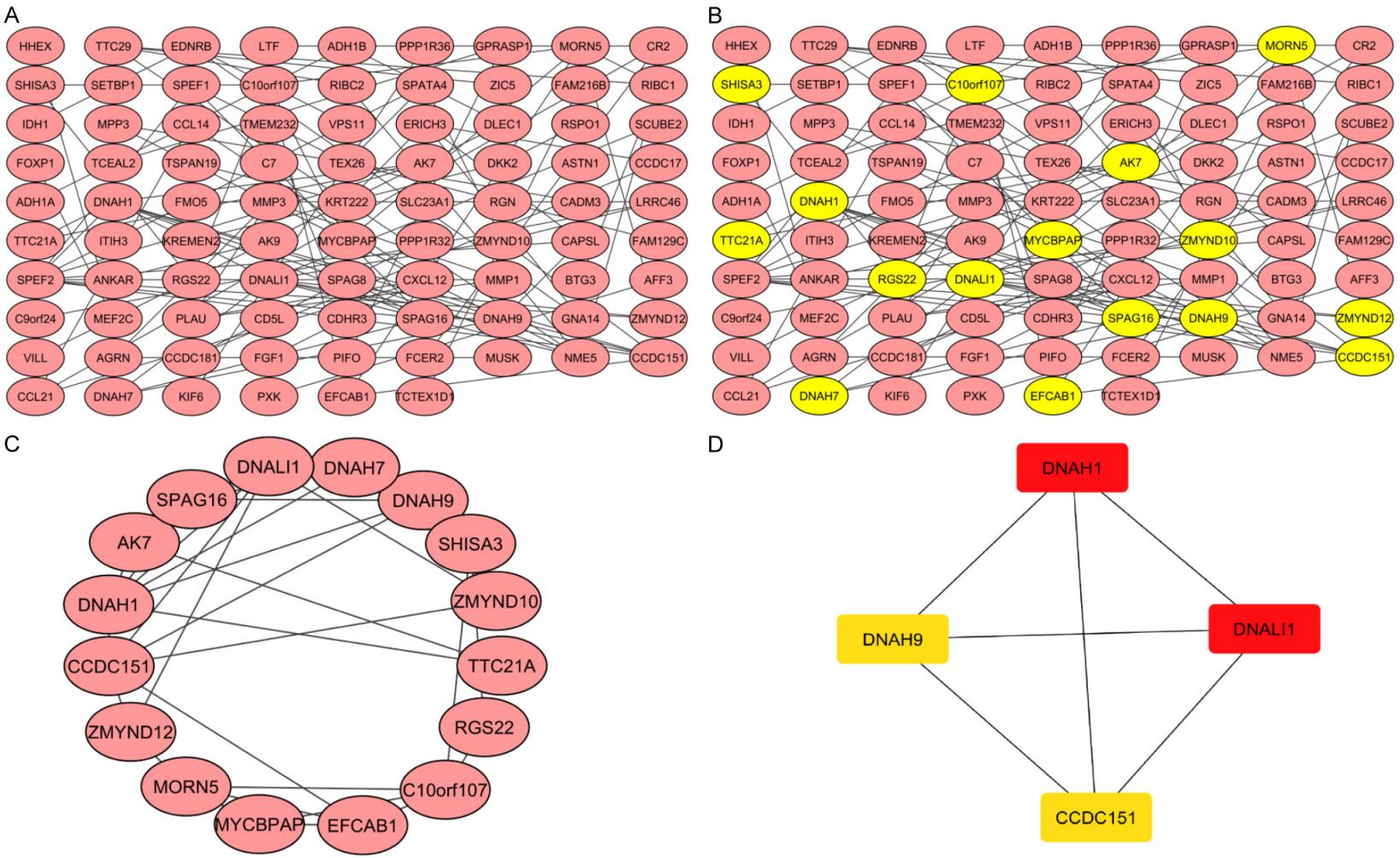
[57] Xi X, Chu Y, Liu N, Wang Q, Yin Z, Lu Y and Chen Y. Joint bioinformatics analysis of underlying potential functions of hsa-let-7b-5p and core genes in human glioma. J Transl Med 2019; 17: 129.

[58] Bozgeyik E. Bioinformatic analysis and in vitro validation of Let-7b and Let-7c in breast cancer. Comput Biol Chem 2020; 84: 107191.

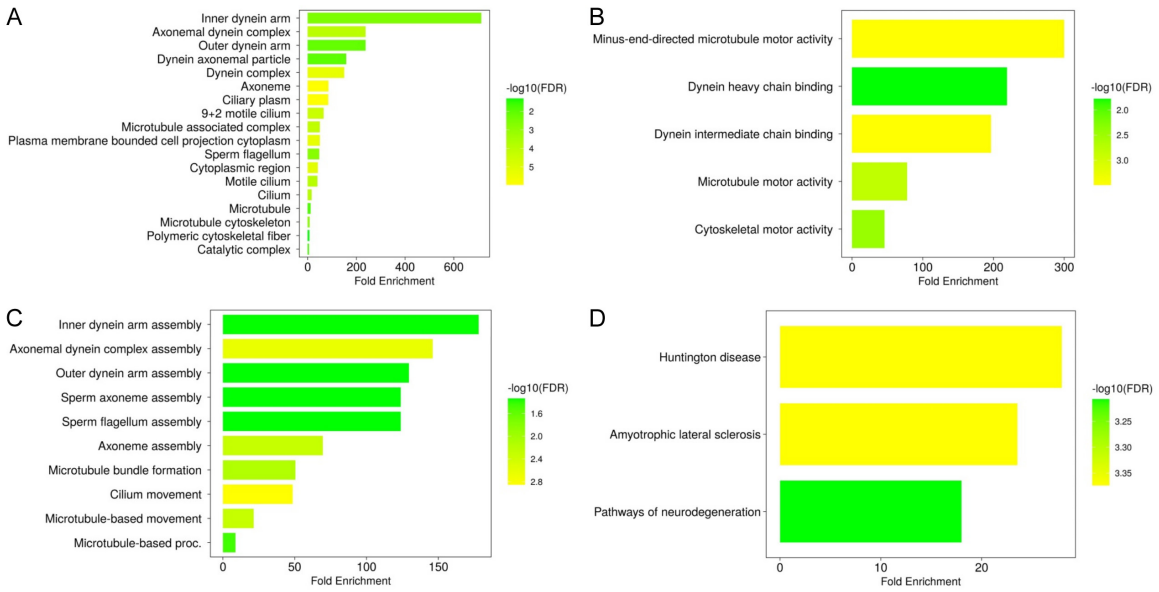
[59] Gu CI, Zhang Z, Fan WS, Li LA, Ye MX, Zhang Q, Zhang NN, Li Z and Meng YG. Identification of MicroRNAs as potential biomarkers in ovarian endometriosis. Reprod Sci 2020; 27: 1715-1723.

[60] Xue C, Liu C, Yun X, Zou X, Li X, Wang P, Li F, Ge Y, Zhang Q, Xie X, Li X and Luo B. Knockdown of hsa_circ_0008922 inhibits the progression of glioma. PeerJ 2022; 10: e14552.

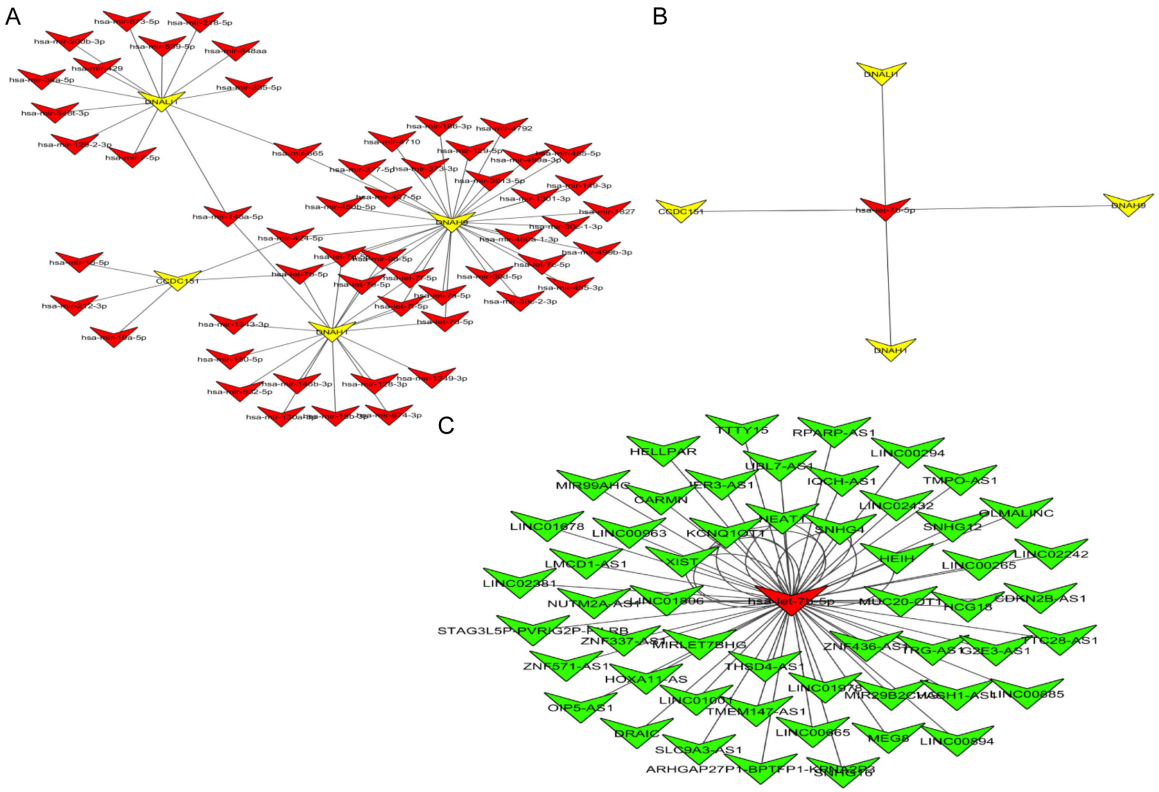
[61] Xue J, Jia E, Ren N and Xin H. Identification of prognostic miRNA biomarkers for esophageal cancer based on The Cancer Genome Atlas and Gene Expression Omnibus. Medicine (Baltimore) 2021; 100: e24832.



Supplementary Figure 1. 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 GSE53819 microarray dataset. (A) A PPI network of the top 250 DEGs in GSE53819 microarray dataset, (B, C) A PPI network of the most significant module, and (D) A PPI network of identified four hub genes.

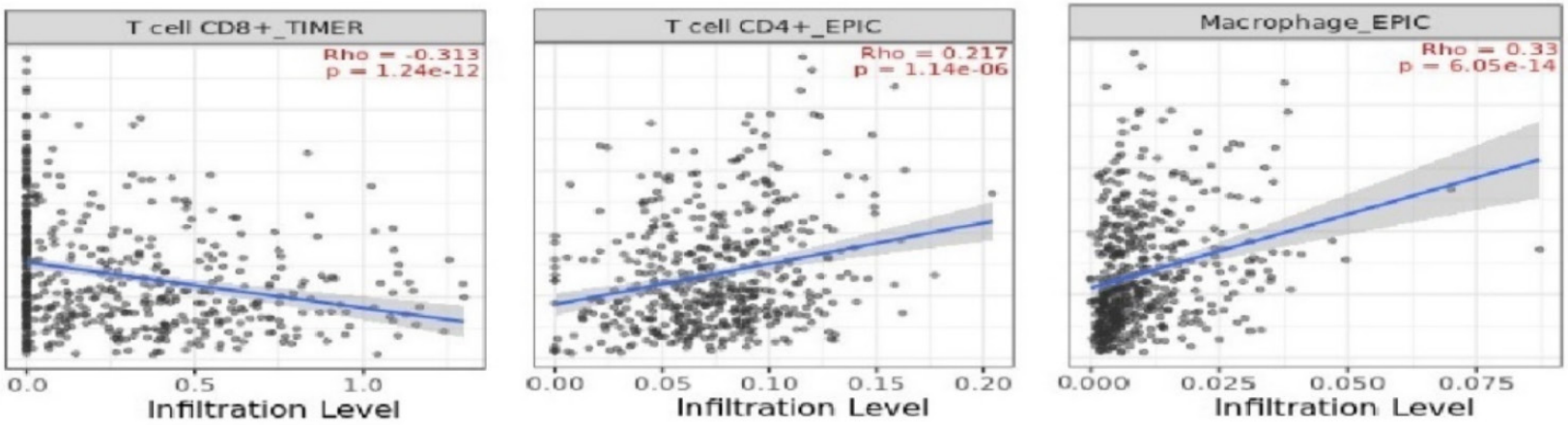


Supplementary Figure 2. Gene enrichment analysis of DNAH1, DNALI1, DNAH9, and CCDC151. (A) DNAH1, DNALI1, DNAH9, and CCDC151 associated CC terms, (B) DNAH1, DNALI1, DNAH9, and CCDC151 associated BP terms, (C) DNAH1, DNALI1, DNAH9, and CCDC151 associated MF terms, and (D) DNAH1, DNALI1, DNAH9, and CCDC151 associated KEGG terms.

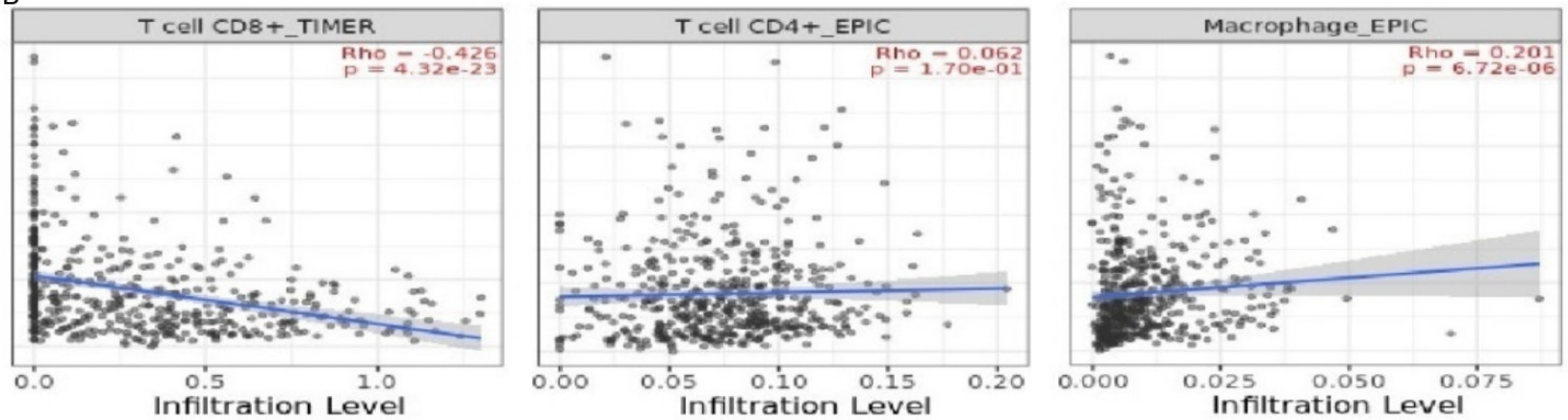


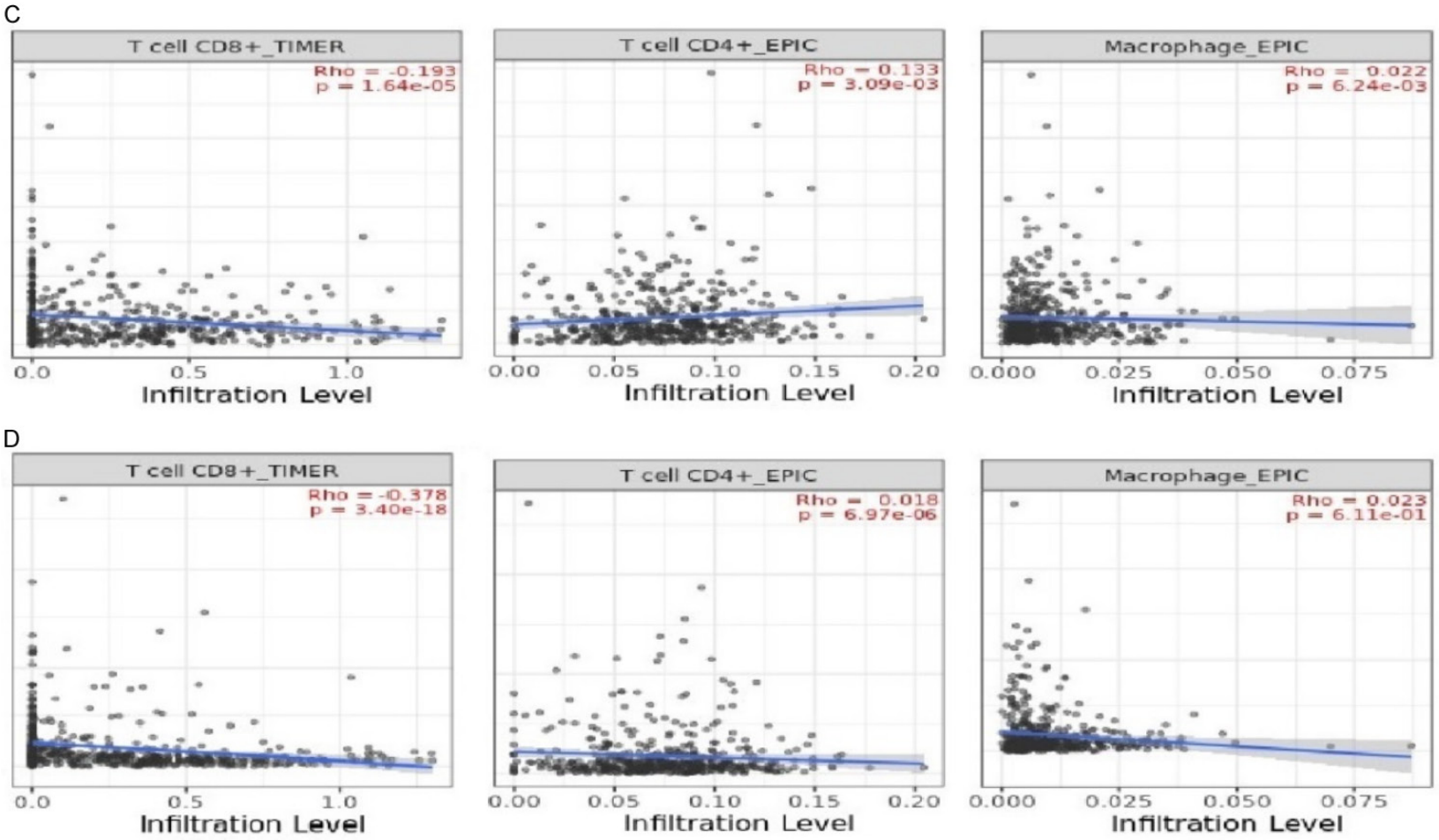
Supplementary Figure 3. lncRNA-miRNA-mRNA co-regulatory network of DNAH1, DNALI1, DNAH9, and CCDC151 hub genes. (A) A PPI of miRNAs targeting hub genes, (B) A PPI highlighting most important miRNA (hsa-let-7b-5p) targeting all hub genes, and (C) A PPI of lncRNAs targeting hsa-let-7b-5p. Red color nodes: miRNAs, yellow color nodes: mRNAs, and Green color nodes: lncRNAs.

A


T cell CD8+_TIMER
Rho = -0.313
p = 1.24e-12
0.0
0.5
1.0
Infiltration Level
T cell CD4+_EPIC
Rho = 0.217
p = 1.14e-06
0.00
0.05
0.10
0.15
0.20
Infiltration Level
Macrophage_EPIC
Rho = 0.33
p = 6.05e-14
0.000
0.025
0.050
0.075
Infiltration Level

B


T cell CD8+_TIMER
Rho = -0.426
p = 4.32e-23
0.0
0.5
1.0
Infiltration Level
T cell CD4+_EPIC
Rho = 0.062
p = 1.70e-01
0.00
0.05
0.10
0.15
0.20
Infiltration Level
Macrophage_EPIC
Rho = 0.201
p = 6.72e-06
0.000
0.025
0.050
0.075
Infiltration Level



Supplementary Figure 4. Correlation analysis of DNAH1, DNALI1, DNAH9, and CCDC151 hub genes expression with different immune cells (CD8+ T, CD4+ T, and Macrophages) infiltration level. (A) DNAH1, (B) DNALI1, (C) DNAH9, and (D) CCDC151.