frontiers | Research Topics

The role of DNA viruses in human cancers

Edited by
Ming Hu, Chengjun Wu and Bin Wang

Published in
Frontiers in Cellular and Infection Microbiology

ISSN 1664-8714
ISBN 978-2-83251-116-9
DOI 10.3389/978-2-83251-116-9

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

The role of DNA viruses in human cancers

Topic editors
Ming Hu — Qingdao University, China
Chengjun Wu — Dalian University of Technology, China
Bin Wang — Qingdao University, China

Citation
Hu, M., Wu, C., Wang, B., eds. (2023). *The role of DNA viruses in human cancers*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-116-9

Table of contents

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Editorial
PUBLISHED 08 December 2022
DOI 10.3389/fcimb.2022.1103505

OPEN ACCESS

EDITED AND REVIEWED BY
John Hiscott,
Istituto Pasteur Italia Cenci Bolognetti Foundation, Italy

*CORRESPONDENCE
Chengjun Wu
wcj5532@dlut.edu.cn
Bin Wang
Wangbin532@126.com

SPECIALTY SECTION
This article was submitted to Virus and Host, a section of the journal Frontiers in Cellular and Infection Microbiology

RECEIVED 20 November 2022
ACCEPTED 28 November 2022
PUBLISHED 08 December 2022

CITATION
Hu M, Wang B and Wu C (2022) Editorial: The role of DNA viruses in human cancers.
Front. Cell. Infect. Microbiol. 12:1103505.
doi: 10.3389/fcimb.2022.1103505

Editorial: The role of DNA viruses in human cancers

Ming Hu[1], Bin Wang[1]* and Chengjun Wu[2]*

[1]Department of Special Medicine, Basic Medicine College, Qingdao University, Qingdao, China, [2]School of Biomedical Engineering, Dalian University of Technology, Dalian, China

KEYWORDS

DNA virus, cancer, prognosis, oncolytic virus, therapy

Editorial on the Research Topic
The role of DNA viruses in human cancers

There is a wide variety of human tumor viruses, including small DNA viruses such as Merkel cell polyomavirus (MCPyV), hepatitis B virus (HBV) and human papillomaviruses (HPV), or large DNA viruses as Kaposi's sarcoma-associated herpesvirus (KSHV) and Epstein–Barr virus (EBV) (Chen et al., 2021). Around 85% of virus-induced cancers occur in developing countries, which has important implications for the translation of knowledge into public health treatment. As well, certain viruses are likely to contribute to sex-specific differences in tumorigenesis. Almost 90% of HPV-induced cancers are almost exclusively diagnosed in women, whereas HBV, HCV and EBV cancers are generally diagnosed in men (Plummer et al., 2016).

Oncoviruses also differ widely in their carcinogenic mechanisms. Cancers caused by viruses usually arise from chronic infections after many years, which indicates that the infection is just one step among many. Viral gene products regulate anti-apoptotic, proliferative and/or immune escape activities by interacting with cellular genes. Examples of continued expression of specific viral oncoproteins include LMP1 of EBV, E6 and E7 of HPVs and Tax of HTLV-1.

The HPV normally infects stratified epithelium and can cause a variety of anogenital carcinomas, including cervical, vaginal and anal and other mucosal carcinomas including oropharyngeal carcinomas (de Martel et al., 2017). Within two years, the host immune system will be able to clear the majority of HPV infections. In some rare cases, the high-risk HPV infection may develop into persistent infection and cause pre-cancerous lesions to the cervix, which may progress to malignant cervical cancer. HPV gains access to basal layer cells by interacting with heparan sulfate proteoglycans on the cell surface. Although endocytosis is necessary for virus entry into cytoplasm, the cellular entry mechanism is not thoroughly studied. Once infection occurs, HPV genome starts to replicate. In the early stage, the viral genome replicates at a low-level. Baedyananda et al. review summarized that the viral early protein E1 and E2 cooperate to enhance the affinity of viral genome binds to the host cellular DNA replication machinery. During this stage, viral oncoprotein E6 and E7 are generated, the high-risk E6 oncoprotein degrades the tumor suppressor protein p53 resulting in the inhibition of cell apoptosis (McBride, 2017). In an overview, Zheng et al. discussed how

several splicing factors control the splicing of high-risk HPV E6/E7 mRNA. E7 binds to tumor suppressor retinoblastoma-associate protein (pRB) and blocks the binding between E7 and its partner protein E2F activating the expression of DNA replication factor. Since the high-risk oncoprotein E6 and E7 are derived from same pre-mRNAs, therefore, the balance of E6 and E7 are vital to the cancer progression. The post-transcriptional modification especially the alternative splicing decides the ratio of E6 and E7. As known, four splice sites locate in HPV16 E6 and E7 coding region including one splicing donor (SD226) and splicing acceptors (SA409, SA526, and SA742) (Ajiro et al., 2012). The unspliced mRNA maintain the entire E6 coding region, it is therefore used for E6 production, while the pre-mRNAs spliced from SD226 to SA409 are used for E7 production. If the splicing is repressed or inhibited, the unspliced E6 level will increase at the expense of E7 level, insufficient E7 cannot completely reduce the pRB, so the infected cells will eventually go to apoptosis. On the contrary, if the splicing is becoming too efficient, the expression level of E7 increases, E6 is too less to completely destroy the apoptosis pathway mediated by p53 thereby shutting down the process of cell carcinogenesis (Olmedo-Nieva et al., 2018 and Cui et al., 2022). It is worth mentioning, besides hnRNPs earlier reported cellular factors including epidermal growth factor (EGF), 5' cap-binding factors, SRSF1/SRSF2, CTCF, and SF3B1 affect E6/E7 mRNA level further interrupt E6 and E7 protein ratio. In this topic, Lu et al. summarized the clinical features of 19 patients with cervical small cell carcinoma trying to explore more effective therapy for cervical SCNEC. Cui et al. identified five differentially expressed genes in HPV infection related cervical cancer using bioinformatics, which may be used as potential biomarker for cervical screening.

Over 50% of HCC cases are caused by HBV infections worldwide, making them the most significant carcinogenic factor of HCC. There are several factors that can increase the risk of HCC in HBV patients, including gender, age, alcohol consumption, exposure to carcinogens, and duration of infection. While HBV-encoded proteins may directly induce hepato-carcinogenesis, they may also cause cancer indirectly through chronic inflammation and tissue damage caused by persistent infection. Our Research Topic brings together articles that explore the mechanism of HBV infection in HCC, as well as treating HBV-related HCC. In a comprehensive analysis, Li et al. explored a microtubule nucleation factor named TPX2 as a new prognostic biomarker in HBV-related HCC. Specifically, this article identified 541 differential expressed lncRNAs from HBV-related TCGA-HCC cohorts in $TPX2^{low}$ and $TPX2^{high}$ groups and discussed how this could affect the prognosis for HCC. The findings claimed a ceRNA regulation network to elucidate how TPX2 affects the prognosis of HBV-related HCC. Unless HCC is detected early and completely resected or abated, the prognosis is grave. Yin et al. focused their study on treating HBV-related HCC with traditional Chinese medicine (TCM). In this study, using target-driven reverse network pharmacology, authors investigated TCM's therapeutic potential in treating HBV-related HCC. In order to understand the biological processes and pathways regulated targets, a network of 47 targets was established and a functional analysis was also conducted. Finally, they obtained a small library of chemical components and herbs against HBV-related HCC.

Using viruses to selectively destroy cancer cells is another established concept. With the approval of several oncolytic virus agents for clinical use, oncolytic viruses have shown promising anti-tumor efficacies, especially in combination with other traditional treatments. Li et al. focused on the characteristics of oncolytic viruses EV-A71 and their mechanisms in tumor treatment. They revealed the molecular mechanism by which EV-A71 infection are able to target glioma cells from the perspective of ABCD3, and also discovered the relationship with immune factor interactions in this process. These innovative virotherapy results strengthened virus-based targeted therapy, while at the same time broadened the idea of gene-modified oncolytic viruses.

In conclusion, this Research Topic addressed the mechanisms by which DNA tumor viruses operate on cells during transformation. Based on the findings in these papers, there are new therapeutic targets, new diagnostic tools, and strategies that need to be applied to clinical practice in order to cure tumors caused by oncogenic viruses.

Author contributions

MH, BW and CW conceived of the study and its design. MH and CW wrote the manuscript. All authors read and approved the final manuscript.

Funding

MH is supported by the Natural Science Foundation of Shandong Province (ZR2021MH177).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Ajiro, M., Jia, R., Zhang, L., Liu, X., and Zheng, Z. M. (2012). Intron definition and a branch site adenosine at nt 385 control RNA splicing of HPV16 E6*I and E7 expression. *PloS One* 7 (10), e46412. doi: 10.1371/journal.pone.0046412

Chen, C.-J., You, S.-L., Hsu, W.-L., Yang, H.-I., Lee, M.-H., Chen, H.-C., et al. (2021). "Epidemiology of virus infection and human cancer," in *Viruses and human cancer: From basic science to clinical prevention*, vol. 13-45 . Eds. T. C. Wu, M.-H. Chang and K.-T. Jeang (Cham: Springer International Publishing).

Cui, X., Hao, C., Gong, L., Kajitani, N., and Schwartz, S. (2022). HnRNP d activates production of HPV16 E1 and E6 mRNAs by promoting intron retention. *Nucleic Acids Res.* 50 (5), 2782–2806. doi: 10.1093/nar/gkac132

de Martel, C., Plummer, M., Vignat, J., and Franceschi, S. (2017). Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int. J. Cancer* 141 (4), 664–670. doi: 10.1002/ijc.30716

McBride, A. A. (2017). Mechanisms and strategies of papillomavirus replication. *Biol. Chem.* 398 (8), 919–927. doi: 10.1515/hsz-2017-0113

Olmedo-Nieva, L., Munoz-Bello, J. O., Contreras-Paredes, A., and Lizano, M. (2018). The role of E6 spliced isoforms (E6*) in human papillomavirus-induced carcinogenesis. *Viruses* 10 (1). doi: 10.3390/v10010045

Plummer, M., de Martel, C., Vignat, J., Ferlay, J., Bray, F., and Franceschi, S. (2016). Global burden of cancers attributable to infections in 2012: a synthetic analysis. *Lancet Glob Health* 4 (9), e609–e616. doi: 10.1016/S2214-109X(16)30143-7

frontiers | Frontiers in Cellular and Infection Microbiology

ORIGINAL RESEARCH
published: 17 June 2022
doi: 10.3389/fcimb.2022.926348

Bioinformatics Analysis Highlights Five Differentially Expressed Genes as Prognostic Biomarkers of Cervical Cancer and Novel Option for Anticancer Treatment

*Hongtu Cui[1], Ruilin Ma[1], Tao Hu[1], Gary Guishan Xiao[2] and Chengjun Wu[1]**

[1] School of Biomedical Engineering, Dalian University of Technology, Dalian, China, [2] School of Pharmaceutical Science and Technology, Dalian University of Technology, Dalian, China

Cervical cancer is one of the most common gynecological malignancies and is related to human papillomavirus (HPV) infection, especially high-risk type HPV16 and HPV18. Aberrantly expressed genes are involved in the development of cervical cancer, which set a genetic basis for patient prognosis. In this study, we identified a set of aberrantly expressed key genes from The Cancer Genome Atlas (TCGA) database, which could be used to accurately predict the survival rate of patients with cervical squamous cell carcinoma (CESC). A total of 3,570 genes that are differentially expressed between normal and cancerous samples were analyzed by the algorithm of weighted gene co-expression network analysis (WGCNA): 1,606 differentially expressed genes (DEGs) were upregulated, while 1,964 DEGs were downregulated. Analysis of these DEGs divided them into 7 modules including 76 hub genes. Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) enrichment analysis revealed a significant increase of genes related to cell cycle, DNA replication, p53 signaling pathway, cGMP-PKG signaling pathway, and Fanconi anemia (FA) pathway in CESC. These biological activities are previously reported to associate with cervical cancer or/and HPV infection. Finally, we highlighted 5 key genes (*EMEMP2, GIMAP4, DYNC2I2, FGF13-AS1*, and *GIMAP1*) as robust prognostic markers to predict patient's survival rate (p = 3.706e-05) through univariate and multivariate regression analyses. Thus, our study provides a novel option to set up several biomarkers for cervical cancer prognosis and anticancer drug targets.

Keywords: bioinformatics analysis, cervical cancer, DEGs, prognosis, anticancer drug

OPEN ACCESS

Edited by:
Qiliang Cai,
Fudan University, China

Reviewed by:
Jianyu Rao,
University of California, Los Angeles, United States
Kai Fu,
University at Buffalo, United States

****Correspondence:***
Chengjun Wu
wcj5532@dlut.edu.cn

Specialty section:
This article was submitted to Virus and Host, a section of the journal Frontiers in Cellular and Infection Microbiology

Received: *22 April 2022*
Accepted: *11 May 2022*
Published: *17 June 2022*

Citation:
Cui H, Ma R, Hu T, Xiao GG and Wu C (2022) Bioinformatics Analysis Highlights Five Differentially Expressed Genes as Prognostic Biomarkers of Cervical Cancer and Novel Option for Anticancer Treatment. Front. Cell. Infect. Microbiol. 12:926348. doi: 10.3389/fcimb.2022.926348

INTRODUCTION

Cervical cancer is one of the most common gynecological malignancies (Small et al., 2017), of which cervical squamous cell carcinoma (CESC) accounts for more than 80% of total cases (Cheng et al., 2020). Cervical cancer has become one of the most susceptible and fatal cancers in women, mainly through sexual contact (Ojesina et al., 2014; Szymonowicz and Chen, 2020). According to the latest

data from the International Cancer Center, about 600,000 new cases of cervical cancer have been reported worldwide in 2020; the reported death cases were more than 340,000 (Sung et al., 2021). Earlier studies have shown that the occurrence of cervical cancer is directly related to the persistent infection of high-risk human papillomavirus (HPV) (Schiffman et al., 2016); more than 99.7% of cervical cancer patients were infected by high-risk HPVs. High-risk subtype HPV16 and HPV18 are the most prevalent types (Walboomers et al., 1999). High-risk HPVs generate E6 and E7 oncoproteins. The E6 protein binds to the cellular factor p53, causing p53 degradation, thereby disrupting cellular apoptosis. The E7 protein interacts with pRB, causing pRB inactivation and altering the cell cycle regulatory pathways. Together, the E6 and E7 oncoproteins cause the transformation of infected cells and finally accumulate mutations to further develop into cancer cells (Burd, 2003).

As the most efficient screening methods Pap test and HPV test are widely used for cervical cancer diagnosis in the clinic (Kessler, 2017), most of the patients can be diagnosed and treated at an early stage of cervical cancer progression. Unfortunately, there is no efficient treatment for advanced or recurrent cervical cancer (Wang et al., 2014; Yang et al., 2020). To date, the common treatment of cervical cancer includes surgical resection, radiotherapy, and chemotherapy (Ellenson and Wu, 2004; Hazell et al., 2018), but the prognosis of patients does not meet expectations. Especially, the 5-year survival rate of patients is even less than 10% (Eskander and Tewari, 2014). Therefore, it is important to identify novel anticancer drug targets to improve therapy efficiency. Moreover, novel prognostic biomarkers are needed to promote the survival rate of cervical cancer patients.

Recent studies focus on screening differentially expressed genes (DEGs) but ignore the complex networks among genes and the gene-related clinical phenotypes (Liu et al., 2019). However, more and more evidence suggests that the arising of cervical cancer involves multiple abnormally expressed genes (Zhou and Wang, 2015; Bahrami et al., 2018). The high-throughput data mining algorithm weighted gene co-expression network analysis (WGCNA) identifies biological key modules by using high-throughput gene expression data (Langfelder and Horvath, 2008). In recent years, WGCNA has been increasingly used in the research of tumor markers. For example, Xing et al. (2020) used WGCNA to find PHY906 and CPT11 as key genes for colon cancer.

In this study, we obtained the DEG expression profiles from the public database The Cancer Genome Atlas (TCGA) to construct a co-expression network to identify cervical cancer progression-related hub genes. These highlighted hub genes can be applied to predict the 3- and/or 5-year survival rates of cervical cancer patients. The potential role of these genes as biomarkers needs to be further studied, and it may also provide a theoretical basis for the prognosis assessment of cervical cancer patients.

MATERIALS AND METHODS

The Cancer Genome Atlas Data Preprocessing

The plan of the study is shown in **Figure S1**. The original cervical cancer expression data (Htseq-counts), standardized data (Htseq-FPKM), and clinical data were obtained from TCGA database (https://cancergenome.nih.gov/), including three normal cervix tissue samples and 304 cervix cancer tissue samples. The clinical data contain information of phenotypes, including the age (20–88 years), gender (men, women), clinical stage (stage IV, stage III, stage II, stage I), neoplasm histologic grade (G4, G3, G2, G1, GX), survival time (0–6,408 days), body mass index (13–70), HPV status (positive, negative), smoking history grade (1–3), and vital status (alive, dead). Low-expression samples in raw data were excluded using the filterByExpr in the edgeR package (Law et al., 2016).

Analysis of Differentially Expressed Genes

The edgeR package, DESeq2 package, and limma package are used to analyze gene expression differences between normal cervical tissue samples and cancer samples (Robinson et al., 2010; Anders and Huber, 2010; Law et al., 2014). The analyzed genes with threshold p value <0.05 and |log2fold change (FC)| >1 were selected as the final DEGs of cervical cancer.

Gene Function Enrichment Analysis

Gene function enrichment analysis is adopted to compare genes or genomes with functional databases for overexpression analysis and functional annotation. GO database is mainly for the study of cellular component (CC), molecular function (MF), and biological process (BP) (The Gene Ontology Consortium, 2019). KEGG database is for understanding the functions and applications of biological systems according to genomics or the information at the molecular level (Kanehisa et al., 2017). In this study, the enrichGO and enrichKEGG functions in software R package cluster profile (Yu et al., 2012) were used to conduct enrichment analysis and pathway analysis of the DEGs to identify the BP of significant GO ($p < 0.05$) and the significant KEGG pathways ($p < 0.05$).

Weighted Gene Co-Expression Network Analysis

The WGCNA package in R software was selected to conduct the process of WGCNA (Langfelder and Horvath, 2008). The construction steps of the network mainly include the gene co-expression similarity matrix calculation, the adjacency function of the gene network calculation, the soft threshold selection, the topological overlap matrix (TOM), the dissimilarity matrix calculation, the gene module by dynamic branch cut method calculation, and the correlation analysis between the gene module and external information.

Core Gene Screening

We use Cytoscape software to visualize the gene network in the module (Doncheva et al., 2019); the highest connectivity genes are identified as hub genes according to the connectivity between genes in the module.

Survival Analysis

The survival package of R software is used for survival analysis of hub genes. According to the median value of the gene expression level, cervical cancer samples were divided into two groups. The survival curves of each group were generated.

Construction of Prognostic Markers of Cervical Cancer Based on Key Genes

Univariate Cox proportional hazards regression analysis was adopted to assess the relationship between hub genes and survival rate. The estimation of the prognostic risk score of each cervical cancer patient was analyzed by multivariate Cox regression analysis. The risk score model was built using the coxph function of the R software survival package. According to the risk score, patients can be divided into low-risk and high-risk groups. The software R survival package generated the survival curves of the two groups.

RESULTS

Differential Analysis Selected 1,606 Upregulated Genes and 1,964 Downregulated Differentially Expressed Genes

The cervical cancer gene expression profile was collected from TCGA database, including 304 cervical cancer tumor tissues and three normal tissues. EdgeR, DESeq2, and Limma analysis were applied to obtain the DEGs between the normal and cervical cancer groups. The batch information was successively added into the constructed model (George et al., 2014; Nygaard et al., 2016). EdgeR identified 2,146 upregulated genes and 2,478 downregulated genes. DESeq2 identified 3,013 upregulated genes and 2,125 downregulated genes. Limma identified 1,779 upregulated genes and 2,758 downregulated genes (**Figure 1A**). The heatmap of DEGs was generated according to the results from EdgeR, DESeq, and Limma analysis (**Figure S2**). We integrated the results of the three methods to reduce the error, and the results were visualized by Venn diagram (Stupnikov et al., 2021). As shown, 1,606 upregulated and 1,964 downregulated genes were identified (**Figures 1B, C**). To narrow down the sample size, we selected 50 highly expressed genes and 50 reduced genes according to the value of |logFC| to generate the heatmap as shown in **Figure 1D**.

To further expand the understanding of the role of all DEGs in the occurrence and progression of cervical cancer, KEGG and GO enrichment analyses were applied. KEGG analysis showed that the selected DEGs were significantly gathered in DNA replication, p53 signaling pathway, and cell cycle (**Figure 1E**). GO enrichment analysis was divided into three groups: BP, CC, and MF. The results of the GO enrichment analysis showed that all of the DEGs in the BP group were enriched in DNA replication, chromosome segregation, and organelle fission. In group CC, these DEGs were highly gathered in cell–cell junction, chromosomal region, actin-binding, single-stranded DNA helicase activity, and catalytic activity were significantly enriched in group MF (**Figure 1F**).

Weighted Gene Co-Expression Network Analysis Highlighted Seven Key Hub Modules Associated With Clinical Phenotypes of Cervical Cancer

The weighted gene co-expression network was established by the DEGs above methods screened to correlate with clinical phenotypes of cervical cancer. The outlier sample TCGA-LP-A4AV-01A was excluded in subsequent analysis (**Figure S3**). The selected samples were grouped in different clusters to form the distribution map of the clinical feature data, including the age of patients with cervical cancer, clinical stage (stage I–stage IV), histological tumor grade (G1–G3), HPV infection status (negative, positive), body mass index (BMI), patient smoking history, survival status (alive, dead), and survival time (**Figure 2A**). Next, to identify the optimal value of the threshold power from 1 to 30, we conducted network topology analysis to determine the relatively balanced scale independence and mean connectivity of the WGCNA. When the threshold power of $\beta = 4$ (scale-free R2 = 0.88) and cutoff modules size, more than 30 were set as the soft threshold to ensure a scale-free network (**Figure 2B**). The tree was grouped into 17 modules by a dynamic tree cut algorithm. All the selected genes were clustered using a TOM-based dissimilarity measurement (**Figure 2C**). Each module was represented by different colors, and the number of genes was concluded and shown in **Table 1**. The genes that did not belong to any of the modules were marked as a gray module. Therefore, this gray module was not included in the subsequent analysis. We analyzed the 16 modules to investigate the interaction between modules, and the heatmap of the network was created (**Figure 2D**). The results indicated that each module was highly independent, but the gene expression in each module was less independent. WGCNA further established the correlation of each module according to different phenotypic traits of cervical cancer by calculating the module significance of correlation of each module-trait (**Figure 2E**). Next, we investigated if each module was positively or negatively correlated, indicating the negatively correlated modules as the following: green-yellow vs. HPV infection status ($r = -0.47$, $p = 8e\text{-}18$), midnight-blue vs. survival status ($r = -0.16$, $p = 0.005$), blue vs. clinical stage ($r = -0.12$, $p = 0.04$), and survival time vs. smoking history ($r = 0.13$, $p = 0.02$); and the positively correlated modules as the following: tan vs. patient's age ($r = 0.19$, $p = 7e\text{-}04$), brown vs. histological grade of cervical cancer ($r = 0.15$, $p = 0.01$), pink blue vs. BMI ($r = 0.14$, $p = 0.01$), and light cyan vs. survival time ($r = 0.14$, $p = 0.02$). Taken together, we selected these seven modules as the most clinical phenotype-associated key hub modules.

Functional Enrichment Analysis Identified 76 Hub Genes From the Hub Modules

We imported the hub modules into Cytoscape software to screen the hub genes. According to the connectivity between genes in the module, we selected the highest connectivity genes in the modules as hub genes. A total of 76 hub genes were screened from 7 hub modules (**Figure 3A**). The KEGG enrichment analysis indicated that hub genes were gathered in "vascular smooth muscle contraction, gap junction, and prostate cancer" (**Figure 3B**). The GO enrichment analysis showed that "extracellular structure organization, extracellular matrix organization, and camera-type eye development" were significantly gathered in the BP group; "collagen-containing extracellular matrix, nucleosome and DNA packaging complex" were significantly enriched in the CC group; and "collagen binding, GTP binding, and extracellular matrix structural constituent" were significantly enriched in the MF group (**Figure 3C**).

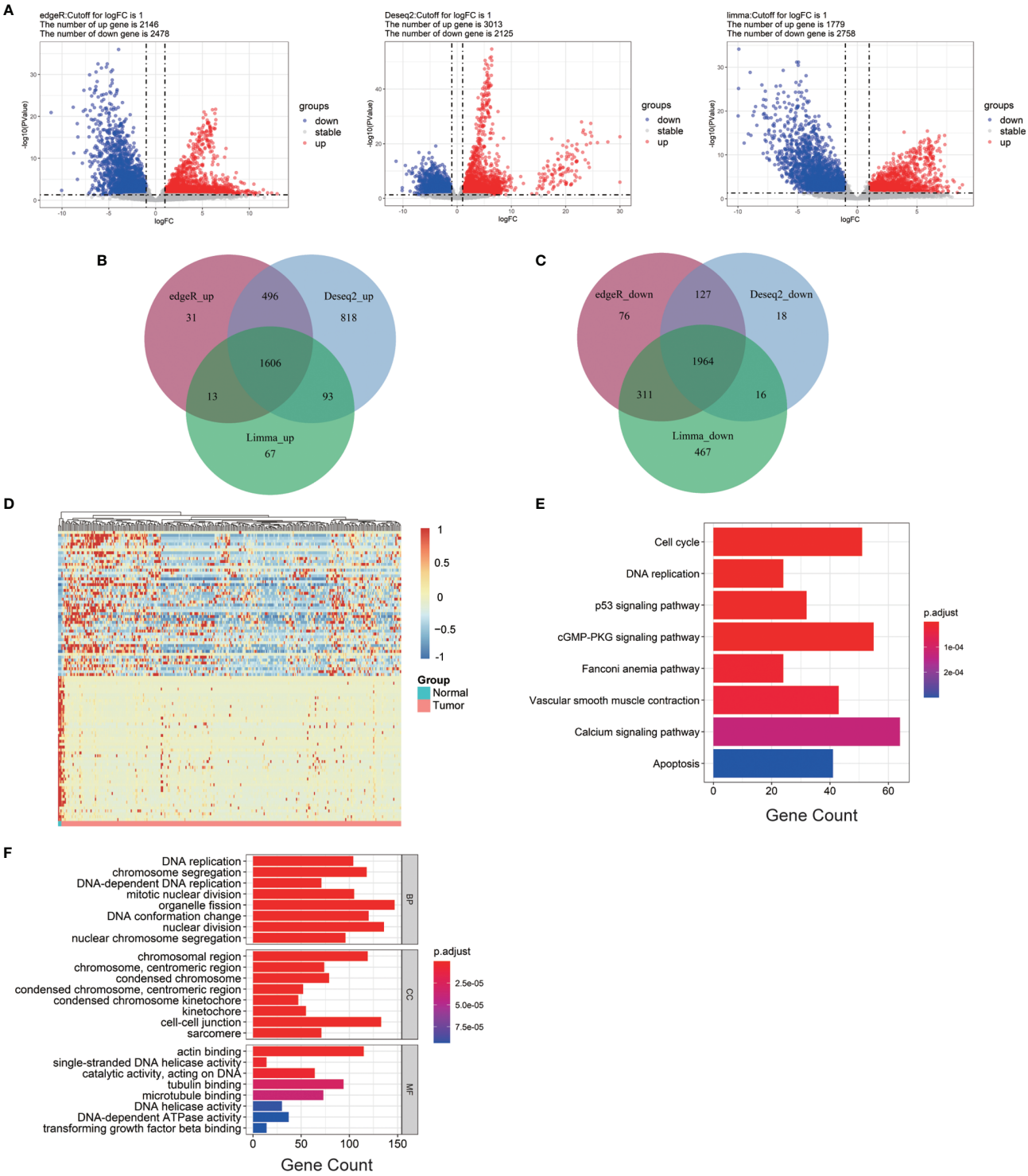


FIGURE 1 | Bioinformatics analysis of all DEGs from cervical cancer tissue samples. **(A)** Volcano plots of DEGs screened by three methods. "Down" refers to the genes that were downregulated. "Up" refers to the genes that were upregulated. "Stable" refers to the genes that have no difference in expression between tumor groups and normal groups. **(B)** Upregulated genes. **(C)** Downregulated genes. **(D)** The heatmap of the top 200 DEGs according to the value of |logFC|. **(E, F)** KEGG pathway enrichment and GO enrichment of DEGs. Each column bar on the y-axis represents an enrichment pathway, and the x-axis is the number of genes that were enriched in this pathway. DEGs, differentially expressed genes; logFC, log fold change; BP, biological process; CC, cellular component; MF, molecular function.

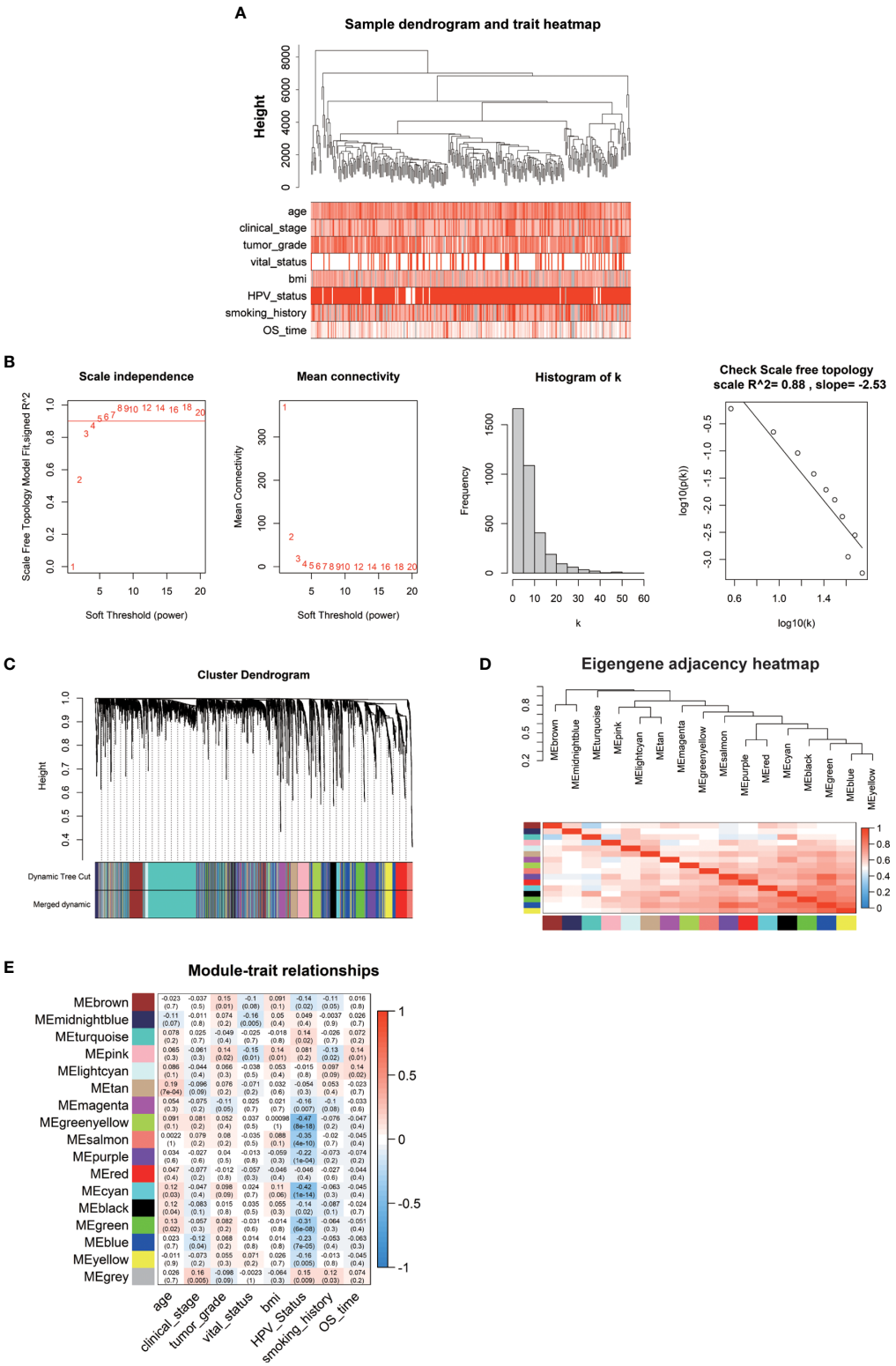


FIGURE 2 | **(A)** The clinical trait heatmap and tree dendrogram. **(B)** Determination of soft-threshold power in the WGCNA. **(C)** Clustering dendrogram of DEGs with assigned module colors. **(D)** Visualization of gene network using heatmap plot. **(E)** Heatmap of the correlation between module eigengenes and phenotype of CESC.

TABLE 1 | The number of genes in each module.

Module	Genes
Brown	279
midnightblue	41
turquoise	844
Pink	143
lightcyan	40
Tan	77
magenta	140
greenyellow	112
salmon	72
Purple	129
Red	167
Cyan	63
Black	158
Green	179
Blue	716
Yellow	218
grey	192

Five Prognostic Markers for Cervical Cancer Were Highlighted by Cox Regression Analysis

We performed a prognostic analysis on the above 76 hub genes to predict the probable prognostic markers. Univariate Cox proportional hazards regression analysis demonstrated that the most significant prognostic factors were EFEMP2, GIMAP4, DYNC2I2, ITM2A, GIMAP7, FGF13-AS1, H1-3, LTA, GIMAP1, and GIMAP5. Multivariate Cox proportional hazards regression analysis was performed to analyze these 10 prognostic factors. GIMAP4, DYNC2I2, EFEMP2, FGF13-AS1, and GIMAP1 showed significant prognostic values, as shown in **Table 2**. Risk values for each patient were obtained from the survival packages. Based on the median risk score, patients with cervical cancer were divided into high-risk and low-risk group. The Kaplan-Meier (KM) survival curve compares the survival time of the high- and low-risk group. As shown, the survival time of the high-risk group was significantly lower than that of the low-risk group ($p < 0.0001$, **Figure 4A**). To further verify the accuracy of the predicting model of the cervical cancer patient's survival time, we calculated the area under the curve (AUC) values of Receiver Operating Characteristic Curve (ROC) curves, which is normally adopted to reflect the reliability of the model (AUC >0.7). The AUC value of 1-, 3-, and 5-year survival time prediction model in this study was 0.712, 0.723, and 0.761, respectively, indicating that this model possessed optimal performance in predicting the survival time of cervical cancer patients at 1, 3, and 5 year (**Figure 4B**). We also analyzed the distribution of risk scores and survival status among cervical cancer patients to further confirm the accuracy of our model in predicting cervical cancer patients' survival time (**Figures 4C, D**).

DISCUSSION

Cervical cancer is one of the most common tumors among women globally. More than 600,000 new cases are diagnosed by the end of 2020; the death rate of HPV-related cervical cancer in Asia is more than 50%. Low- and middle-income countries have higher mortality rates due to poor medical conditions (Daniyal et al., 2015; Denny, 2015; Vu et al., 2018). The occurrence and progression of cervical cancer are highly associated with the infection of high-risk HPV (Walboomers et al., 1999; Schiffman et al., 2016). In addition, increasing evidence suggested that many DEGs are expressed by cancer cells (van Wieringen and van der Vaart, 2015). Aberrant gene expression levels in the cancer cell may lead to the dysregulation of cell signaling pathways by inhibiting or stimulating (Vogelstein and Kinzler, 2004). Therefore, we detected abnormally expressed genes of CESC from TCGA database. EdgeR, DESeq2, and Limma were applied to reduce errors. A total of 3,570 DEGs were obtained (1,606 were upregulated and 1,964 were downregulated). To further investigate the role of these DEGs in cervical cancer, KEGG analysis and GO enrichment analysis was performed. We observed that these DEGs were significantly enriched in DNA replication, cell cycle, cGMP-PKG signaling pathway, p53 signaling pathway, and Fanconi anemia (FA) pathway. The high-risk E6 oncoprotein has been shown to degrade p53, resulting in the inhibition of apoptosis, which is a key step to progress to cervical cancer (Howie et al., 2009; Kajitani et al., 2012; Zheng et al., 2020). Previous studies demonstrated that the E7 oncoprotein affects cell cycle progression by preventing the binding of the tumor suppressor protein pRB and E2F or degrading pRB and the pocket proteins, which together contribute to cervical cancer progression (Roman and Munger, 2013). Enriched DEGs in the p53 signaling pathway and cell cycle that were highlighted in this study are meaningful to future studies to reveal the mechanism of oncoproteins E6 and E7 in the occurrence of cervical cancer. Plenty of reports have shown that the cGMP/PKG pathway is involved in the proliferation, differentiation, and apoptosis of cancer cells (Fajardo et al., 2014). Furthermore, HPV infection highly depends on cell proliferation and differentiation; a previous study indicated that the cGMP/PKG pathway plays a key role in the malignant phenotype of cervical cancer cells (Gong et al., 2019), and our data are also in agreement with the role of the cGMP/PKG pathway in the development of cervical cancer. The disruption of the FA pathway has been shown to increase HPV16 E7 protein levels and viral genome amplification (Hoskins et al., 2012). Unfortunately, due to the insufficient number of normal samples in TCGA database, there will be a certain amount of errors in screening the DEGs. Even though the DEGs highlighted by our bioinformatics analysis need to be further investigated, the role of these DEGs may provide chances to develop novel drug targets or diagnostic markers.

WGCNA is a co-expression network algorithm, which is widely used in the research of cancer markers. Although a previous study identified many prognostic markers in cervical cancer using WGCNA, they mainly focus on DEGs between groups (tumor vs. normal) (Liu et al., 2019). Still, the clinical phenotype of patients has not been taken into consideration. In addition to HPV infection, previous studies have shown that age, smoking, and obesity are also associated with cervical cancer (Waggoner et al., 2006; Gu et al., 2013; Su et al., 2018; Zhang

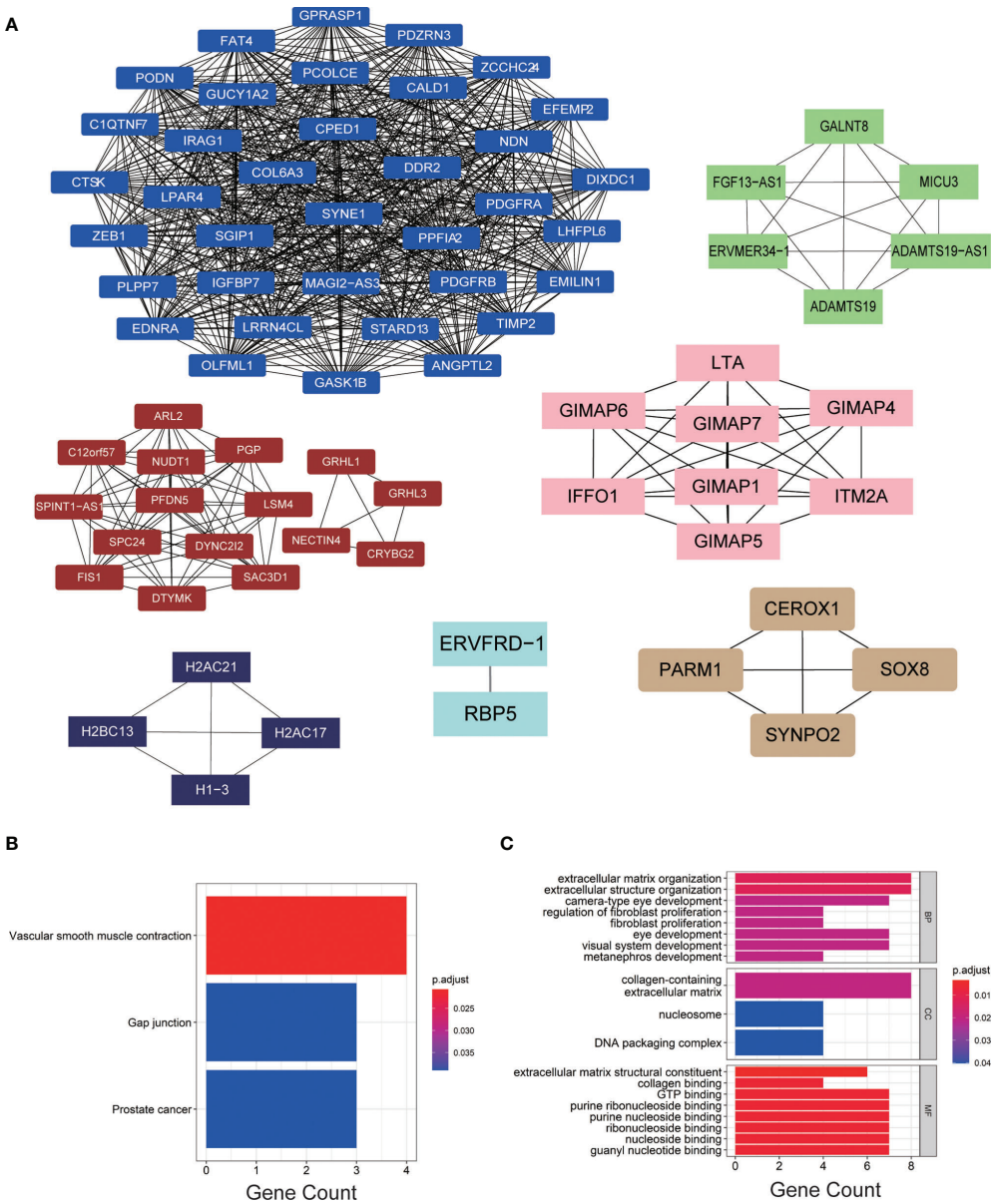


FIGURE 3 | **(A)** The hub genes of each module. **(B, C)** KEGG and GO enrichment analyses of hub genes.

TABLE 2 | The significantly prognostic genes revealed by univariate and multivariate Cox proportional hazards regression analysis.

characteristic	Univariate analysis			Multivariate analysis		
	HR	95%CI	P value	HR	95%CI	P value
EFEMP2	1.055	(1.015,1.097)	0.007	1.045	(1.005,1.088)	0.028
GIMAP4	0.941	(0.899,0.984)	0.008	0.829	(0.727,0.945)	0.005
DYNC2I2	0.989	(0.980,0.997)	0.008	0.986	(0.977,0.994	0.001
ITM2A	0.886	(0.808,0.970)	0.009	–	–	–
GIMAP7	0.931	(0.881,0.984)	0.012	–	–	–
FGF13-AS1	8.432	(1.506,47.224)	0.015	13.504	(2.161,84.378)	0.005
H1-3	0.623	(0.423,0.917)	0.016	–	–	–
LTA	0.548	(0.328,0.914)	0.021	–	–	–
GIMAP1	0.652	(0.453,0.939)	0.021	2.545	(0.980,6.614)	0.055
GIMAP5	0.111	(0.017,0.735)	0.023	–	–	–

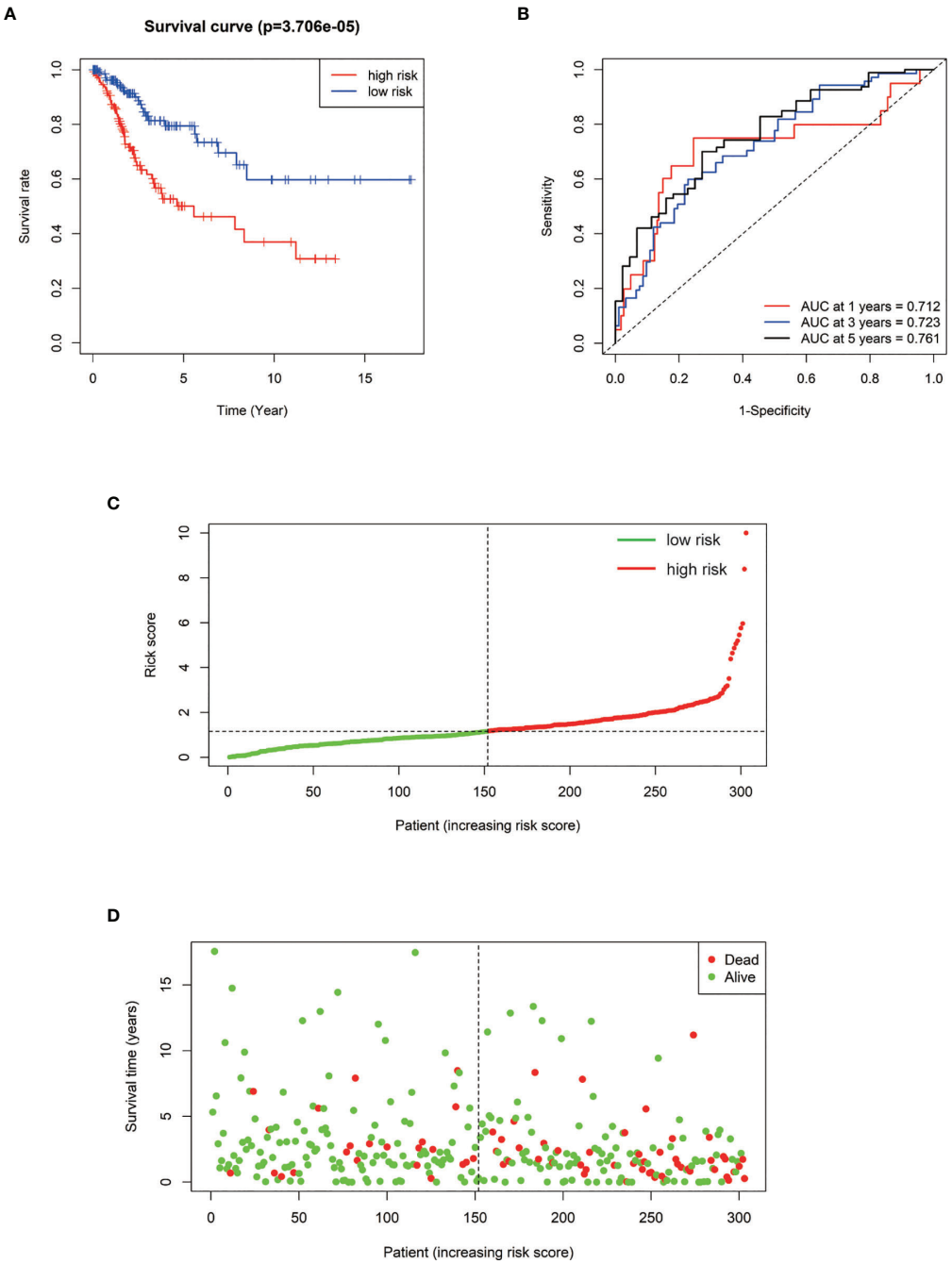


FIGURE 4 | **(A)** Kaplan–Meier survival curve of overall survival between the high-risk group and low-risk group. **(B)** The 1-, 3-, 5-year survival time-dependent ROC curve. **(C, D)** The distributions of the risk score and survival status for each patient.

et al., 2019). Moreover, we also considered that clinical indicators may associate with cervical cancer patients. Therefore, in this study, we firstly screened DEGs between normal tissues and tumor tissues, then carried out WGCNA by considering eight clinical phenotypes of the cervical cancer patient, including the age, clinical stage, tumor histology grade, vital status, BMI, HPV infection status, smoking history, and survival time of cervical cancer patients. Finally, a total of 7 modules were associated with these clinical traits. From the module–trait relationship in WGCNA, HPV infection is the most relevant phenotype. Normally, the HPV infection will be cleared by the host immune system within a year or two. In some rare cases, the HPV may develop into persistent infection, after several years, sometimes decades, may progress to cervical cancer (Hu and

Ma, 2018). Interestingly, in our study, WGCNA indicated that older age is positively related to cervical cancer. Considering that about 99.7% of cervical cancer cases are related to high-risk HPV infection, we suspected that older age that positively related to the occurrence of cervical cancer may be due to high-risk HPV persistent infection.

Currently, the prognosis of patients with cervical cancer is not ideal, and the survival rate of patients is too low (Eskander and Tewari, 2014). Therefore, prognostic markers are of great significance to enhance the overall survival rate of patients. To establish a reliable prognostic model, we performed a prognostic analysis of all key genes from the above seven modules. Finally, we identified that the model constructed by GIMAP4, GIMAP1, FGF13-AS1, EFEMP2, and DYNC2I2 could be used as the prediction model for the prognosis of cervical cancer. GIMAP1 and GIMAP4 are proteins of the GIMAP members (GTPase immune-associated proteins) family. GIMAP4 plays a key role in cellular apoptosis and T-cell development, which can be used as a prognostic marker of cervical cancer (Schnell et al., 2006; Heinonen et al., 2015; Xu et al., 2021). GIMAP1 is crucial for the survival of B cells used as the major marker of endometrial cancer (Krucken et al., 1999; Webb et al., 2016; Guo et al., 2021). The human EFEMP2 gene is located near the centromere of chromosome 11q13 and plays a role in the invasion and metastasis of tumors (Obaya et al., 2012; Papke and Yanagisawa, 2014). EFEMP2 has been found related to survival rate; downregulation of EFEMP2 leads to a higher death rate in bladder cancer (Zhou et al., 2019). FGF13-AS1 is a long-chain non-coding RNA (lncRNA); studies have shown that FGF13-AS1 inhibits the proliferation and migration of breast cancer by impairing glycolysis and dry properties. Reduction of FGF13-AS1 is associated with poor prognosis (Ma et al., 2019). DYNC2I2 is involved in cell cycle progression, apoptosis, and gene regulation [provided by RefSeq, March 2014], but this gene was not reported in the literature. Considering the features of DEGs in the expression, clinical phenotypes, and survival time of cervical cancer patients, we concluded that these five hub genes are likely to play a role in cervical cancer, which can be considered as potential biomarkers. However, the function of these five genes needs further experimental verification.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in The Cancer Genome Atlas (TCGA) at [https://cancergenome.nih.gov/]. These data were derived from the following resources available in the public domain: [https://xenabrowser.net/datapages/?dataset=TCGA-HNSC.htseq_counts.tsv&host=https%3A%2F%2Fgdc.xenahubs.net&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443]. [https://xenabrowser.net/datapages/?dataset=TCGA-HNSC.htseq_fpkm.tsv&host=https%3A%2F%2Fgdc.xenahubs.net&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443]

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: Study conception and design: CW; Data collection: HC and TH; Analysis and interpretation of results: HC and RM; Draft article preparation: HC, GX, and CW.All authors reviewed the results and approved the final version of the article.

FUNDING

This work was supported by the National Natural Science Foundation of China (grant 82072287).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2022.926348/full#supplementary-material

REFERENCES

Anders, S., and Huber, W. (2010). Differential Expression Analysis for Sequence Count Data. *Genome Biol.* 11 (10), R106. doi: 10.1186/gb-2010-11-10-r106

Bahrami, A., Hasanzadeh, M., Shahidsales, S., Farazestanian, M., Hassanian, S. M., Moetamani Ahmadi, M., et al. (2018). Genetic Susceptibility in Cervical Cancer: From Bench to Bedside. *J. Cell Physiol.* 233 (3), 1929–1939. doi: 10.1002/jcp.26019

Burd, E. M. (2003). Human Papillomavirus and Cervical Cancer. *Clin. Microbiol. Rev.* 16 (1), 1–17. doi: 10.1128/CMR.16.1.1-17.2003

Cheng, Y., Yang, S., Shen, Y., Ding, B., Wu, W., Zhang, Y., et al. (2020). The Role of High-Risk Human Papillomavirus-Related Long Non-Coding RNAs in the Prognosis of Cervical Squamous Cell Carcinoma. *DNA Cell Biol.* 39 (4), 645–653. doi: 10.1089/dna.2019.5167

Daniyal, M., Akhtar, N., Ahmad, S., Fatima, U., Akram, M., and Asif, H. M. (2015). Update Knowledge on Cervical Cancer Incidence and Prevalence in Asia. *Asian Pac J. Cancer Prev.* 16 (9), 3617–3620. doi: 10.7314/APJCP.2015.16.9.3617

Denny, L. (2015). Control of Cancer of the Cervix in Low- and Middle-Income Countries. *Ann. Surg. Oncol.* 22 (3), 728–733. doi: 10.1245/s10434-014-4344-8

Doncheva, N. T., Morris, J. H., Gorodkin, J., and Jensen, L. J. (2019). Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J. Proteome Res.* 18 (2), 623–632. doi: 10.1021/acs.jproteome.8b00702

Ellenson, L. H., and Wu, T. C. (2004). Focus on Endometrial and Cervical Cancer. *Cancer Cell* 5 (6), 533–538. doi: 10.1016/j.ccr.2004.05.029

Eskander, R. N., and Tewari, K. S. (2014). Targeting Angiogenesis in Advanced Cervical Cancer. *Ther. Adv. Med. Oncol.* 6 (6), 280–292. doi: 10.1177/1758834014543794

Fajardo, A. M., Piazza, G. A., and Tinsley, H. N. (2014). The Role of Cyclic Nucleotide Signaling Pathways in Cancer: Targets for Prevention and Treatment. *Cancers (Basel)* 6 (1), 436–458. doi: 10.3390/cancers6010436

George, B., Seals, S., and Aban, I. (2014). Survival Analysis and Regression Models. *J. Nucl. Cardiol.* 21 (4), 686–694. doi: 10.1007/s12350-014-9908-2

Gong, L., Lei, Y., Tan, X., Dong, Y., Luo, Z., Zhang, D., et al. (2019). Propranolol Selectively Inhibits Cervical Cancer Cell Growth by Suppressing the cGMP/PKG Pathway. *BioMed. Pharmacother.* 111, 1243–1248. doi: 10.1016/j.biopha.2019.01.027

Gu, W., Chen, C., and Zhao, K. N. (2013). Obesity-Associated Endometrial and Cervical Cancers. *Front. Biosci. (Elite Ed.)* 5, 109–118. doi: 10.2741/E600

Guo, C., Tang, Y., Zhang, Y., and Li, G. (2021). Mining TCGA Data for Key Biomarkers Related to Immune Microenvironment in Endometrial Cancer by

Immune Score and Weighted Correlation Network Analysis. *Front. Mol. Biosci.* 8, 645388. doi: 10.3389/fmolb.2021.645388

Hazell, S. Z., Stone, R. L., Lin, J. Y., and Viswanathan, A. N. (2018). Adjuvant Therapy After Radical Trachelectomy for Stage I Cervical Cancer. *Gynecol. Oncol. Rep.* 25, 15–18. doi: 10.1016/j.gore.2018.05.001

Heinonen, M. T., Kanduri, K., Lähdesmäki, H. J., Lahesmaa, R., and Henttinen, T. A. (2015). Tubulin- and Actin-Associating GIMAP4 is Required for IFN-Gamma Secretion During Th Cell Differentiation. *Immunol. Cell Biol.* 93 (2), 158–166. doi: 10.1038/icb.2014.86

Hoskins, E. E., Morreale, R. J., Werner, S. P., Higginbotham, J. M., Laimins, L. A., Lambert, P. F., et al. (2012). The Fanconi Anemia Pathway Limits Human Papillomavirus Replication. *J. Virol.* 86 (15), 8131–8138. doi: 10.1128/JVI.00408-12

Howie, H. L., Katzenellenbogen, R. A., and Galloway, D. A. (2009). Papillomavirus E6 Proteins. *Virology* 384 (2), 324–334. doi: 10.1016/j.virol.2008.11.017

Hu, Z., and Ma, D. (2018). The Precision Prevention and Therapy of HPV-Related Cervical Cancer: New Concepts and Clinical Implications. *Cancer Med.* 7 (10), 5217–5236. doi: 10.1002/cam4.1501

Kajitani, N., Satsuka, A., Kawate, A., and Sakai, H. (2012). Productive Lifecycle of Human Papillomaviruses That Depends Upon Squamous Epithelial Differentiation. *Front. Microbiol.* 3, 152. doi: 10.3389/fmicb.2012.00152

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: New Perspectives on Genomes, Pathways, Diseases and Drugs. *Nucleic Acids Res.* 45 (D1), D353–D361. doi: 10.1093/nar/gkw1092

Kessler, T. A. (2017). Cervical Cancer: Prevention and Early Detection. *Semin. Oncol. Nurs.* 33 (2), 172–183. doi: 10.1016/j.soncn.2017.02.005

Krücken, J., Stamm, O., Schmitt-Wrede, H. P., Mincheva, A., Lichter, P., and Wunderlich, F. (1999). Spleen-Specific Expression of the Malaria-Inducible Intronless Mouse Gene Imap38. *J. Biol. Chem.* 274 (34), 24383–24391. doi: 10.1074/jbc.274.34.24383

Langfelder, P., and Horvath, S. (2008). WGCNA: An R Package for Weighted Correlation Network Analysis. *BMC Bioinf.* 9, 559. doi: 10.1186/1471-2105-9-559

Law, C. W., Chen, Y., Shi, W., and Smyth, G. K. (2014). Voom: Precision Weights Unlock Linear Model Analysis Tools for RNA-Seq Read Counts. *Genome Biol.* 15 (2), R29. doi: 10.1186/gb-2014-15-2-r29

Law, C. W., Alhamdoosh, M., Su, S., Dong, X., Tian, L., Smyth, G. K., et al. (2016). RNA-Seq Analysis is Easy as 1-2-3 With Limma, Glimma and edgeR. *F1000Res* 5, ISCB Comm J–1408. doi: 10.12688/f1000research.9005.3

Liu, J., Nie, S., Gao, M., Jiang, Y., Wan, Y., Ma, X., et al. (2019). Identification of EPHX2 and RMI2 as Two Novel Key Genes in Cervical Squamous Cell Carcinoma by an Integrated Bioinformatic Analysis. *J. Cell Physiol.* 234 (11), 21260–21273. doi: 10.1002/jcp.28731

Ma, F., Liu, X., Zhou, S., Li, W., Liu, C., Chadwick, M., et al. (2019). Long non-Coding RNA FGF13-AS1 Inhibits Glycolysis and Stemness Properties of Breast Cancer Cells Through FGF13-AS1/IGF2BPs/Myc Feedback Loop. *Cancer Lett.* 450, 63–75. doi: 10.1016/j.canlet.2019.02.008

Nygaard, V., Rodland, E. A., and Hovig, E. (2016). Methods That Remove Batch Effects While Retaining Group Differences may Lead to Exaggerated Confidence in Downstream Analyses. *Biostatistics* 17 (1), 29–39. doi: 10.1093/biostatistics/kxv027

Obaya, A. J., Rua, S., Moncada-Pazos, A., and Cal, S. (2012). The Dual Role of Fibulins in Tumorigenesis. *Cancer Lett.* 325 (2), 132–138. doi: 10.1016/j.canlet.2012.06.019

Ojesina, A. I., Lichtenstein, L., Freeman, S. S., Pedamallu, C. S., Imaz-Rosshandler, I., Pugh, T. J., et al. (2014). Landscape of Genomic Alterations in Cervical Carcinomas. *Nature* 506 (7488), 371–375. doi: 10.1038/nature12881

Papke, C. L., and Yanagisawa, H. (2014). Fibulin-4 and Fibulin-5 in Elastogenesis and Beyond: Insights From Mouse and Human Studies. *Matrix Biol.* 37, 142–149. doi: 10.1016/j.matbio.2014.02.004

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). Edger: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics* 26 (1), 139–140. doi: 10.1093/bioinformatics/btp616

Roman, A., and Munger, K. (2013). The Papillomavirus E7 Proteins. *Virology* 445 (1-2), 138–168. doi: 10.1016/j.virol.2013.04.013

Schiffman, M., Doorbar, J., Wentzensen, N., de Sanjosé, S., Fakhry, C., Monk, B. J., et al. (2016). Carcinogenic Human Papillomavirus Infection. *Nat. Rev. Dis. Primers* 2, 16086. doi: 10.1038/nrdp.2016.86

Schnell, S., Démollière, C., van denBerk, P., and Jacobs, H. (2006). Gimap4 Accelerates T-Cell Death. *Blood* 108 (2), 591–599. doi: 10.1182/blood-2005-11-4616

Small, W. Jr., Bacon, M. A., Bajaj, A., Chuang, L. T., Fisher, B. J., Harkenrider, M. M., et al. (2017). Cervical Cancer: A Global Health Crisis. *Cancer* 123 (13), 2404–2412. doi: 10.1002/cncr.30667

Stupnikov, A., McInerney, C. E., Savage, K. I., McIntosh, S. A., Emmert-Streib, F., Kennedy, R., et al. (2021). Robustness of Differential Gene Expression Analysis of RNA-Seq. *Comput. Struct. Biotechnol. J.* 19, 3470–3481. doi: 10.1016/j.csbj.2021.05.040

Su, B., Qin, W., Xue, F., Wei, X., Guan, Q., Jiang, W., et al. (2018). The Relation of Passive Smoking With Cervical Cancer: A Systematic Review and Meta-Analysis. *Med. (Baltimore)* 97 (46), e13061. doi: 10.1097/MD.0000000000013061

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* 71 (3), 209–249. doi: 10.3322/caac.21660

Szymonowicz, K. A., and Chen, J. (2020). Biological and Clinical Aspects of HPV-Related Cancers. *Cancer Biol. Med.* 17 (4), 864–878. doi: 10.20892/j.issn.2095-3941.2020.0370

The Gene Ontology Consortium. (2019). The Gene Ontology Resource: 20 Years and Still GOing Strong. *Nucleic Acids Res.* 47 (D1), D330–D338. doi: 10.1093/nar/gky1055

van Wieringen, W. N., and van der Vaart, A. W. (2015). Transcriptomic Heterogeneity in Cancer as a Consequence of Dysregulation of the Gene-Gene Interaction Network. *Bull. Math Biol.* 77 (9), 1768–1786. doi: 10.1007/s11538-015-0103-7

Vogelstein, B., and Kinzler, K. W. (2004). Cancer Genes and the Pathways They Control. *Nat. Med.* 10 (8), 789–799. doi: 10.1038/nm1087

Vu, M., Yu, J., Awolude, O. A., and Chuang, L. (2018). Cervical Cancer Worldwide. *Curr. Probl. Cancer* 42 (5), 457–465. doi: 10.1016/j.currproblcancer.2018.06.003

Waggoner, S. E., Darcy, K. M., Fuhrman, B., Parham, G., Lucci, J. 3rd, Monk, B. J., et al. (2006). Association Between Cigarette Smoking and Prognosis in Locally Advanced Cervical Carcinoma Treated With Chemoradiation: A Gynecologic Oncology Group Study. *Gynecol. Oncol.* 103 (3), 853–858. doi: 10.1016/j.ygyno.2006.05.017

Walboomers, J. M., Jacobs, M. V., Manos, M. M., Bosch, F. X., Kummer, J. A., Shah, K. V., et al. (1999). Human Papillomavirus is a Necessary Cause of Invasive Cervical Cancer Worldwide. *J. Pathol.* 189 (1), 12–19. doi: 10.1002/(SICI)1096-9896(199909)189:1<12::AID-PATH431>3.0.CO;2-F

Wang, I. T., Chou, S. C., and Lin, Y. C. (2014). Zoledronic Acid Induces Apoptosis and Autophagy in Cervical Cancer Cells. *Tumour Biol.* 35 (12), 11913–11920. doi: 10.1007/s13277-014-2460-5

Webb, L. M., Datta, P., Bell, S. E., Kitamura, D., Turner, M., and Butcher, G. W. (2016). GIMAP1 Is Essential for the Survival of Naive and Activated B Cells In Vivo. *J. Immunol.* 196 (1), 207–216. doi: 10.4049/jimmunol.1501582

Xing, S., Wang, Y., Hu, K., Wang, F., Sun, T., and Li, Q. (2020). WGCNA Reveals Key Gene Modules Regulated by the Combined Treatment of Colon Cancer With PHY906 and CPT11. *Biosci. Rep.* 40 (9). doi: 10.1042/BSR20200935

Xu, F., Shen, J., and Xu, S. (2021). Integrated Bioinformatical Analysis Identifies GIMAP4 as an Immune-Related Prognostic Biomarker Associated With Remodeling in Cervical Cancer Tumor Microenvironment. *Front. Cell Dev. Biol.* 9, 637400. doi: 10.3389/fcell.2021.637400

Yang, H. J., Xue, J. M., Li, J., Wan, L. H., and Zhu, Y. X. (2020). Identification of Key Genes and Pathways of Diagnosis and Prognosis in Cervical Cancer by Bioinformatics Analysis. *Mol. Genet. Genomic Med.* 8 (6), e1200. doi: 10.1002/mgg3.1200

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). Clusterprofiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS* 16 (5), 284–287. doi: 10.1089/omi.2011.0118

Zhang, L., Jiang, Y., Lu, X., Zhao, H., Chen, C., Wang, Y., et al. (2019). Genomic Characterization of Cervical Cancer Based on Human Papillomavirus Status. *Gynecol Oncol.* 152 (3), 629–637. doi: 10.1016/j.ygyno.2018.12.017

Zheng, Y., Jönsson, J., Hao, C., Shoja Chaghervand, S., Cui, X., Kajitani, N., et al. (2020). Heterogeneous Nuclear Ribonucleoprotein A1 (hnRNP A1) and hnRNP A2 Inhibit Splicing to Human Papillomavirus 16 Splice Site SA409

Through a UAG-Containing Sequence in the E7 Coding Region. *J. Virol.* 94 (20). doi: 10.1128/JVI.01509-20

Zhou, Q., Chen, S., Lu, M., Luo, Y., Wang, G., Xiao, Y., et al. (2019). EFEMP2 Suppresses Epithelial-Mesenchymal Transition *via* Wnt/beta-Catenin Signaling Pathway in Human Bladder Cancer. *Int. J. Biol. Sci.* 15 (10), 2139–2155. doi: 10.7150/ijbs.35541

Zhou, X. L., and Wang, M. (2015). Expression Levels of Survivin, Bcl-2, and KAI1 Proteins in Cervical Cancer and Their Correlation With Metastasis. *Genet. Mol. Res.* 14 (4), 17059–17067. doi: 10.4238/2015.December.16.6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Cellular and Infection Microbiology

REVIEW
published: 27 June 2022
doi: 10.3389/fcimb.2022.929666

High-Risk Human Papillomavirus Oncogenic E6/E7 mRNAs Splicing Regulation

Yunji Zheng[1†], *Xue Li*[1†], *Yisheng Jiao*[2] *and Chengjun Wu*[2*]

[1] School of Pharmacy, Binzhou Medical University, Yantai, China, [2] School of Biomedical Engineering, Dalian University of Technology, Dalian, China

High-risk human papillomavirus infection may develop into a persistent infection that is highly related to the progression of various cancers, including cervical cancer and head and neck squamous cell carcinomas. The most common high-risk subtypes are HPV16 and HPV18. The oncogenic viral proteins expressed by high-risk HPVs E6/E7 are tightly involved in cell proliferation, differentiation, and cancerous transformation since E6/E7 mRNAs are derived from the same pre-mRNA. Hence, the alternative splicing in the E6/E7-coding region affects the balance of the E6/E7 expression level. Interrupting the balance of E6 and E7 levels results in cell apoptosis. Therefore, it is crucial to understand the regulation of E6/E7 splice site selection and the interaction of splicing enhancers and silencers with cellular splicing factors. In this review, we concluded the relationship of different E6/E7 transcripts with cancer progression, the known splicing sites, and the identified cis-regulatory elements within high-risk HPV E6/E7-coding region. Finally, we also reviewed the role of various splicing factors in the regulation of high-risk HPV oncogenic E6/E7 mRNA splicing.

Keywords: high-risk HPVs, cervical cancer, HNSCC, E6/E7, splicing, splicing factors

OPEN ACCESS

Edited by:
Qiliang Cai,
Fudan University, China

Reviewed by:
Tanel Punga,
Uppsala University, Sweden
Gloria Arriagada,
Andres Bello University, Chile
Zhenyong Wu,
University of Wisconsin—Madison, United States

***Correspondence:**
Chengjun Wu
wcj5532@dlut.edu.cn

[†]*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Virus and Host, a section of the journal Frontiers in Cellular and Infection Microbiology

Received: *27 April 2022*
Accepted: *19 May 2022*
Published: *27 June 2022*

Citation:
Zheng Y, Li X, Jiao Y and Wu C (2022) High-Risk Human Papillomavirus Oncogenic E6/E7 mRNAs Splicing Regulation. Front. Cell. Infect. Microbiol. 12:929666. doi: 10.3389/fcimb.2022.929666

HUMAN PAPILLOMAVIRUS INFECTION AND RELATED CANCERS

Human papillomaviruses (HPV) are non-enveloped, small, double-stranded DNA viruses. The genome size of HPVs consists of 8,000 nucleotides (zur Hausen, 2002; zur Hausen, 2009). As one of the most common sexually transmitted diseases among almost all sexually active populations, a HPV-caused viral infection usually starts at the cutaneous or mucosal epithelium at the basal layer through sexual activity or a small wound. In most cases, a HPV infection is usually asymptomatic and will disappear within 2 years (Richardson et al., 2003). However, it will develop into a persistent infection in some rare cases, further developing into malignant lesions and progressing to cancer if left untreated (Parkin and Bray, 2006; Zheng and Baker, 2006). Depending on the association of HPV subtypes with cancer progression, they are divided into high-risk HPVs and low-risk HPVs (Zheng and Baker, 2006; Stanley, 2008).

About 15% of human cancers are caused by a viral infection; a high-risk HPV infection nearly accounts for half of all cases. According to the global cancer report in 2021, about 600,000 cancer

cases worldwide are reported to be associated with high-risk HPV infections, causing 340,000 deaths. Cervical cancer is one of the most prevalent cancer types tightly related to HPVs; more than 99% of cases are diagnosed with HPV infections. Moreover, recent studies have shown that anal cancer and head and neck cancer increase rapidly; about 95% of anal cancer and 80% of head and neck cancer are associated with a high-risk HPV infection (Arbyn et al., 2012). Up to now, more than 20 high-risk HPVs have been identified, among which HPV16 and HPV18 are the most popular high-risk strains (Lorincz et al., 1992; de Villiers et al., 2004; Bzhalava et al., 2013; de Villiers, 2013). Nearly half of cervical cancer cases are HPV16-positive, while HPV18 accounts for 20% of the cases (de Sanjose et al., 2010; Li et al., 2011).

High-risk HPV viral oncoproteins E6 and E7 are essential for viral replication and the termination of cell differentiation (Doorbar, 2005; Johansson and Schwartz, 2013). Furthermore, E6 and E7 play a vital role in cancer progression since they interact with cellular tumor suppressor proteins p53 and pRb. Transcription factor p53 contains a DNA binding region to activate the expression of the genes involved in DNA damage repair and cell apoptosis (Howie et al., 2009). HPV16 E6 oncoprotein interacts with p53 with the help of E6-associated protein, causing p53 degradation through the ubiquitination pathway. E6 interrupts p53 DNA binding activity, resulting in transcription inhibition by directly binding to p53. P300/CBP is a histone acetyltransferase essential for p53 production. E6 indirectly downregulates the p53 level by interacting with p300/CBP and inhibiting the p300/CBP expression level (Li and Coffino, 1996; Patel et al., 1999; Zimmermann et al., 2000). The loss of p53 may result in immortalization, followed by cancer occurrence (Scheffner et al., 1993; Martinez-Zapien et al., 2016). The tumor suppressor factor pRb plays a significant role in the regulation of cell cycle. Within the G1 phase, unphosphorylated pRb interacts with transcription factor E2F to suppress the transcription. Once pRb starts to phosphorylate gradually, the cell cycle steps into the S phase (Kato et al., 1993). HPV16 E7 interacts with pRB, causing the inactivation or degradation of pRB, thus leading to the release of E2F and thereby forcing the cell cycle to progress from G1 to the S phase (Munger et al., 2001).

The mRNAs that encode E6 and E7 are bicistronic and generated from the same pre-mRNA (Mesplede et al., 2012). Various splice sites are identified in high-risk HPV E6- and E7-coding regions. Alternative splicing decides the selection of the splice site to produce different mRNA expression variants. However, no splicing in E6 and E7 in low-risk HPVs has been reported since all the spliced E6/E7 mRNA isoforms excluded the most part of the E6-coding sequence. Hence, E6 is only translated from unspliced mRNA, and E7 is generated from several spliced mRNAs but not the unspliced mRNAs. One can speculate that the production of E7 produced from the unspliced mRNA is too low to reach the threshold that can be detected (Zheng and Baker, 2006; Ajiro and Zheng, 2014). The balance of E6/E7 level is directly controlled by alternative splicing, thereby impacting the fate of infected cells developing to apoptosis or immortalization. This review concerns the regulation of E6/E7 splicing isoforms, summarizing the identified cis-regulatory elements and the splicing factors involved in E6/E7 splicing.

RNA PROCESSING IN GENERAL

The gaps in between the intronic region and the exonic coding regions are 5′ splice site (5′ ss) and 3′ splice site (3′ ss). The 5′ ss contains a highly conserved invariable "GU" dinucleotide, while the 3′ ss contains a conserved "AG" dinucleotide (Black, 2003; Hallegger et al., 2010). Besides the splice sites, a branch point sequence (BPS) contains a conserved adenine located at 1–40 nucleotides upstream of 3′ ss, and a polypyrimidine tract is also needed for splicing (Taggart et al., 2012). The general process of mRNA processing includes 5′ capping, polyadenylation, and mRNA splicing. 5′ capping adds a 7-methylguanosine cap to the 5′ end of the growing transcript by a 5′- to -5′ phosphate linkage (Topisirovic et al., 2011). If 5′ capping provides the ribosome with a "start-checkpoint" for translation and mRNA stabilization, the 3′ polyadenylation is an "end-check point" by adding a poly-A tail to the mRNA (Bienroth et al., 1993; Beaudoing et al., 2000). Pre-mRNA splicing is a process that brings exons together in different combinations by removing the introns to generate mature mRNA-encoding functional proteins. The occurrence of splicing requires cis-acting elements that function as splicing enhancers or silencers to interact with trans-acting factors and further constitute spliceosomes (Black, 2003; Hallegger et al., 2010). Besides RNA splicing, post-transcriptional modification is also essential to mature mRNA production, such as N6-methyl adenosine (m6A), which is the most prevalent conserved internal modification in prokaryotic and eukaryotic RNAs (Liu and Pan, 2016). However, we will only focus on the regulation of HPV oncogene E6/E7 pre-mRNA splicing in this review.

HPV GENE EXPRESSION IS REPLICATION CYCLE DEPENDENT

Various splice sites have been identified in the HPV genome; the selection of different splice sites generates a variety of HPV mRNA transcripts to ensure that HPV expresses enough viral protein to drive their replication cycle. Therefore, the regulation of HPV gene expression is highly linked to the HPV replication cycle (Doorbar, 2005; Johansson and Schwartz, 2013). The HPV replication cycle can be divided into early and late stages. The study of the HPV replication cycle is best characterized for HPV16. In HPV16, the early promoter p97 is active during the early phase, while the late promoter p670 is active at the late phase. The HPV16 infection usually starts by gaining entry to the host cell's nucleus in an epithelial cell located at the basal layer of the mucosal epithelium. The cellular transcription machinery helps the HPV16 genome to start replicating from early promoter p97 (Egawa, 2003; Borgogna et al., 2012; Howley and Pfister, 2015; Schafer et al., 2015). HPV16

expresses only early genes during the early stage, including E6, E7, E1, E2, E4, and E5. The HPV16 E2 protein accumulates during cell differentiation and shuts down the HPV early promoter p97. Consequently, HPV16 E6 and E7 expressions are shut off, and the HPV16 differentiation-dependent late promoter p670 is activated (Shin et al., 2007). The late promoter p670 mediates several early viral gene expressions, including E1, E2, and E4. However, when the HPV replication cycle is switched to the late stage, p670 mediates only late gene expression, including L1, L1i, and L2. All the early mRNAs produced by promoter p97 and p670 are polyadenylated at pAE, but high levels of HPV E2 protein inhibit HPV pAE, leading to the activation of HPV late L1 and L2 gene expression. The late mRNAs produced by HPV late promoter p670 are polyadenylated at pAL (Johansson and Schwartz, 2013) (**Figure 1**).

Although most of the HPV infections will be cleared by the host immune system within 2 years, some rare cases may become persistent infections and eventually progress to cancer if left untreated. It is worth mentioning that the occurrence of cancer needs to repress or mute the E2 expression since E2 plays a vital role in early gene expression (Hanahan and Weinberg, 2000). In the majority of cervical cancer cases, high-risk HPVs' genome integrates into the host genome, and the early viral gene expression regulator E2 protein is disrupted (Pyeon et al., 2009). Within this stage, progeny virion is not produced anymore, and the overexpression of E6 and E7 is induced (Groves and Coleman, 2015). A high level and constant E6 and E7 oncoprotein expression are required for tumor formation and maintenance. However, some cervical cancer cases demonstrated that high-risk-HPV DNA keeps in an episomal form. In these cases, the differential methylation of E2 binding sites 1, 3, and 4 in long control region is related to the overexpression of the E6 and E7 proteins, further suggesting that HPV genome integration into the host genome may not be essential for malignant cellular transformation (Matsukura et al., 1989; Goodwin and DiMaio, 2000; Vinokurova et al., 2008).

HPVS E6/E7 SPLICING ISOFORM AND FUNCTIONS

HPV expresses multiple viral proteins to maintain the viral replication cycle. The most common high-risk subtypes HPV16 and HPV18 have been studied thoroughly. In HPV16, various splicing donors and acceptors have been characterized (**Figure 2**). The selection of different splicing donors and splicing acceptors decides the generation of viral proteins. The splicing donor SD226 alternatively spliced to SA409, SA526, and SA742 generates the E6*I, E6*II, and E6^E7 variants (**Figure 2**) (Guccione et al., 2004; Ajiro et al., 2012; Manzo-Merino et al., 2014; Olmedo-Nieva et al., 2018; Zheng et al., 2020). In HPV18, splicing donor SD233 and splicing acceptors SA416, SA3434, SA2779, and SA791 are involved in producing E6/E7 transcript variants (Vazquez-Vega et al., 2013; Cerasuolo et al., 2017; Olmedo-Nieva et al., 2018). The coding region of high-risk HPV E6/E7 contains at least one splicing donor, and an acceptor is used for alternative splicing to generate one or several isoforms termed E6*, but low-risk HPVs and beta-papillomavirus do not have such splicing sites in the E6/E7-

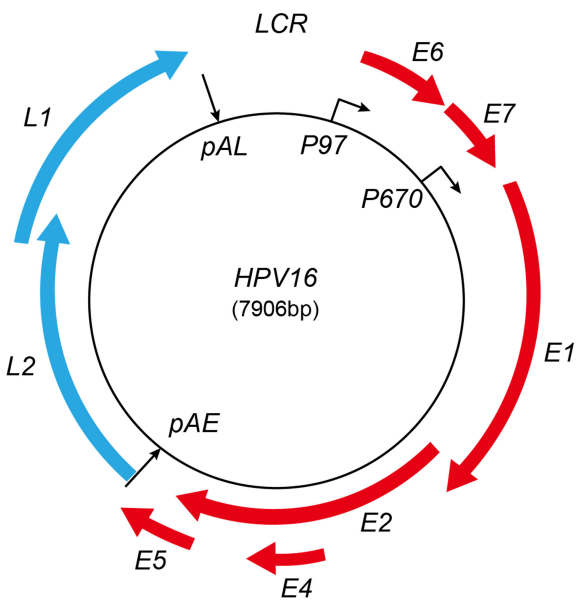


FIGURE 1 | Schematic map of HPV16 genome. Red represents early proteins; blue represents late proteins. LCR, long control region; pAE, early polyadenylation signal; pAL, late polyadenylation signal. The early and late promoters are indicated. Red boxes represent viral early-protein-encoding regions; blue boxes represent late-protein-encoding regions.

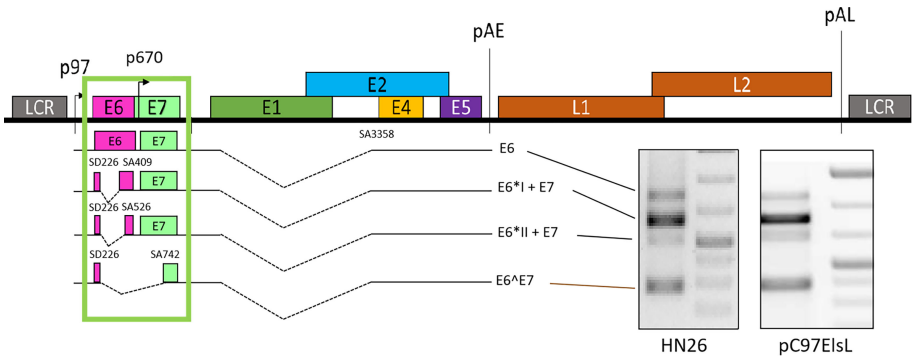


FIGURE 2 | Schematic presentation of E6/E7 transcripts. The unspliced mRNA in E6 coding region generates E6 (Zheng et al., 2020), SD226^SA409 generates E6*I and E7, SD226^SA526 generates E6*II and E7, and SD226^742 generates E6^E7. All the transcripts are detected in primary head and neck cancer cell line HN26 and subgenomic plasmid pC97ElsL transfected HeLa cells.

coding region to conduct alternative splicing. We concluded the known splice sites and E6/E7 mRNA variants of high-risk HPV as those shown in **Table 1**. Since E6 and E7 are oncoproteins and play a vital role in high-risk HPV-induced cancers, the splicing mechanism of E6 and E7 is of great interest for study (Graham and Faizo, 2017; Cerasuolo et al., 2020). HPV16 oncoprotein E6 and E7 are generated from the same pre-mRNA. It is disputable that the unspliced mRNA produces an early viral E6 protein since the unspliced mRNA contains the full-length E6-encoding gene. The mRNA spliced from SD226 to SA409 produces the E6*I transcript, used for E7 production (Bernard et al., 1987). In high-risk HPV, the splicing within the E6-coding region to produce E6*I mRNA is a landmark to distinguish high- and low-risk HPV (Guccione et al., 2004; Mesplede et al., 2012; Filippova et al., 2014) replication cycle. A previous study demonstrates that the most abundant transcript isoform in HPV16-positive cervical malignant and oropharyngeal lesion is E6*I. In agreement with this observation, consistent results were identified in various cancer cell lines (Guccione et al., 2004) (**Figure 2**).

The splicing between HPV16 splice site SD226 and SA409 results in the shortening of the E6-coding region. The shortened coding region produces a small protein E6*I; the E6*I open reading frame (Bernard et al., 1987) contains a weak ATG located upstream of the E7 ORF. Therefore, the ribosome may reinitiate translation from E7 ATG by leaky scanning. Zheng ZM et al. demonstrated that most E7 proteins are translated from E6*I mRNA in HPV16-positive cervical cancer cells (CaSki) and HPV18-positive HeLa cells (Schafer et al., 2015). Although increasing evidence supports this conclusion, several individual studies also presented different opinions that unspliced E6E7 RNA expresses more E7 than spliced variants. It is necessary to point out that these findings were drawn from a less stringent experimental system (Filippova et al., 2007; Manzo-Merino et al., 2014). E6*I is a multi-functional protein that has been detected in a HPV16-positive CaSki cell line and functions as an antagonist of the full-length E6 protein, further suggesting its antitumorigenic role. In HPV18, E6*I protein indirectly promotes the expression of the potential cervical cancer diagnostic marker p14ARF through p53 degradation. The overexpression of E6*I results in a slight increase of p14ARF. This result suggests that HPV18 E6*I protein may interact with p53 to prevent p53 from regulating p14ARF (Vazquez-Vega et al., 2013). Several earlier studies demonstrate that E6*I interacts with p53 to impact p53 degradation. Moreover, it has been shown that the E6*I protein does not drive keratinocyte immortalization and proliferation (Sedman et al., 1991; Pim et al., 1997). However, HPV16 produces other E6/E7 splicing isoforms, the mRNA SD226^SA526 encodes E6*II protein, while the mRNA SD226^SA742 produces E6^E7 protein. The ratio of unspliced E6, E6*I, E6*II, and E6^E7 mRNAs plays a vital role in HPV-related cancer progression. The variant E6^E7 is generally present at the lower level, while E6*I is the most abundant mRNA expressed in HPV16-infected cells or HPV16-related cancers (Ajiro et al., 2012; Ajiro et al., 2016; Olmedo-Nieva et al., 2018; Zheng et al., 2020). In addition, the splicing variant E6*II mRNA expression level is regularly higher than that of the unspliced E6 mRNA but lower than the major spliced product E6*I. However, the E6*I and E6*II mRNA levels vary in different cancers; cervical cancer samples express higher levels of E6*I and E6*II mRNAs than those in oropharyngeal cancer (Cerasuolo et al., 2017). E6*II is a suboptimal E7 producer (Schmitt et al., 2010; Schmitt and Pawlita, 2011) that has been associated with the grade of cervical lesions, but this association could not be confirmed. Very little is known about the functions of E6^E7 protein, but it has been reported to assist in the stabilization of E6, E7, and E6*I (Ajiro and Zheng, 2015). Furthermore, several suboptimal splice sites have been identified in HPV16 to generate E6/E7 mRNA variants within the E6 ORF, including SD191^SA409 (E6*VI), SD221^SA409 (E6*V), SD174^SA718 (E6^E7*I), and SD221^SA850 (E6^E7*II) (Ajiro et al., 2012; Islam et al., 2017). A previous study indicates that HPV16-positive breast cancer samples detect various E6/E7 transcript isoforms, including E6*I, E6*II, E6^E7, E6^E7*I, and E6^E7*II. This finding does not elucidate the role of these isoforms in breast cancer progression but suggests that HPV16 is

TABLE 1 | Summary of identified E6/E7 splicing variants in 20 high-risk human papillomaviruses (HPVs) (Olmedo-Nieva et al., 2018).

HPV type	E6/E7 transcripts	5′ ss–3′ ss (nucleotide positions)
HPV16	E6*I	226–409 (Doorbar et al., 1990)
	E6*II	226–526 (McFarlane et al., 2015)
	E6*III	226–3,358 (Doorbar et al., 1990)
	E6*IV	226–2,709 (Schmitt et al., 2010)
	E6*V	221–409 (Ajiro et al., 2012)
	E6*VI	191–409 (Ajiro et al., 2012)
	E6^E7 (E6*X)	226–742 (Ajiro and Zheng, 2015)
	E6^E7*I	174–718 (Islam et al., 2017)
	E6^E7*II	221–850 (Islam et al., 2017)
HPV18	E6*I	233–416 (Toots et al., 2014; Ajiro et al., 2016)
	E6*II	233–3,434 (Toots et al., 2014; Ajiro et al., 2016)
	E6*III	233–2,779 (Ajiro et al., 2016)
	E6^E7	233–791 (Ajiro and Zheng, 2015; Ajiro et al., 2016)
HPV26	E6*I	173–406 (Mesplede et al., 2012)
HPV31	E6*I	210–413 (Hummel et al., 1992; Ozbun, 2002)
	E6*III	210–3,295 (Ozbun, 2002)
HPV33	E6*I	231–509 (Snijders et al., 1992)
	E6*II	231–785 (Snijders et al., 1992)
	E6*III	231–3,351 (Snijders et al., 1992)
HPV35	E6*I	228–419 (Mesplede et al., 2012)
HPV39	E6*I	231–420 (Mesplede et al., 2012)
HPV45	E6*I	230–413 (Halec et al., 2013)
HPV51	E6*I	173–406 (Mesplede et al., 2012)
HPV52	E6*I	224–502 (Halec et al., 2013)
HPV53	E6*I	236–419 (Halec et al., 2013)
HPV56	E6*I	157–420 (Mesplede et al., 2012)
HPV58	E6*I	232–510 (Li et al., 2013)
	E6*II	232–3,355 (Li et al., 2013)
HPV59	E6*I	183–582 (Halec et al., 2013)
HPV66	E6*I	157–420 (Mesplede et al., 2012)
HPV67	E6*I	224–502 (Halec et al., 2013)
HPV70	E6*I	231–422 (Halec et al., 2013)
HPV73	E6*I	227–410 (Halec et al., 2013)
HPV82	E6*I	178–411 (Mesplede et al., 2012)

transcriptionally active (Islam et al., 2017). However, the functional feature of these mRNA variants in HPV16-related cancer progression is largely unknown.

HPV E6/E7 SPLICING REGULATION

The splicing regulation is tightly related to the splicing site selection regulated by the interaction of cis-acting elements and trans-acting factors. cis-acting elements provide binding motifs to interact with cellular splicing factors resulting in splicing enhancement or splicing inhibition. The cis-acting elements located in exons enhancing splicing are termed exonic splicing enhancers (ESE), while those of inhibiting splicing are termed exonic splicing silencers (ESS). Similarly, the cis-acting elements located in introns are also termed intronic splicing enhancers (Garland et al., 2007) and intronic splicing silencers (ISS). The trans-acting factors interact with cis-regulatory elements to facilitate the composition of spliceosomes, further conducting splicing from splice sites. The expression level of trans-acting factors varies in different tissues, and they may compete for binding to the cis-acting elements to promote splicing or inhibit splicing (Manley and Tacke, 1996; Wang et al., 2006; Zhou and Fu, 2013). The families of serine/arginine-rich proteins (SR proteins) and heterogeneous nuclear ribonucleoproteins (hnRNPs) are known splicing factors. The SR proteins usually function as splicing activators binding to ESE or ISE, while hnRNPs function as splicing silencers binding to ESS or ISS (Graveley et al., 1999; Hertel and Maniatis, 1999; Han et al., 2010). However, this classification is not strict; increasing evidence indicates that the role of SR and hnRNP proteins in splicing regulation may alter (Lim et al., 2011). Their opposite function in splicing depends on the positions of the splicing regulatory elements relative to the splice site. Besides SR and hnRNP proteins, various splicing factors participate in the pre-mRNA splicing process (Elliott et al., 1998).

Cis-Regulatory Elements in E6/E7-Coding Region

The study on the mechanism of HPV splicing was well performed on the most common subtypes, including HPV16 and HPV18. Several cis-regulatory elements have been mapped, including BPS and splicing silencers. Moreover, various splicing factors have been identified to impact the production of expression of major splicing variants E6*I and E7.

Researchers have identified the sequence of HPV16 BPS as AACAAAC located within the E6-coding region upstream of 3′ ssSA409. The adenosine at nucleotide 385 (underlined) is the branch site to dominate the splicing to 3′ ss SA409, affecting the E7 expression. Once the point mutation was introduced to the branch site in BPS, the BPS binding activity to U2 protein was interrupted, resulting in 3′ ss SA409 splicing inhibition to 3′ ss SA409. Therefore, the pre-mRNAs contain the mutated BPS, causing inefficient splicing activity to produce very little E7 protein. Moreover, since the BPS locates in the E6-coding region, the point mutation at BPS generates the mutated E6 protein, which has little effect on p53 degradation (Ajiro et al., 2012). An earlier study identified a suboptimal BPS (AGUGAGU) on HPV16, from nucleotide 323 to 329 in the E6-coding region. The underlined guanosine at nucleotide 328 is a cryptic branch site. This cryptic BPS may play its role in keeping the E6/E7 mRNA level, but no information about E6/E7 pre-mRNA splicing pattern change has been reported (De la Rosa-Rios et al., 2006) (**Figure 3A**). In HPV18, researchers have identified two alternative branch sites for E6*I generation and E7 production. The mapped BPS is AACUAAC (from nucleotide 383 to 389); both branch sites are adenosine: one adenosine locates at nucleotide 384, and the other locates four nucleotides downstream at 388 (underlined). The selection of two alternative branch sites decides the efficiency of the E6*I splice, further affecting the production of HPV18 E6 and E7. The study has demonstrated that the E6*I splice favors the branch site at nucleotide 388 rather than the adenosine at nucleotide 384. If the preferred branch site is inactivated or mutated, cryptic splicing to 3′ ss SA636 would be activated (Brant et al., 2019) (**Figure 3B**). In general, the identified HPV18 BPS in the E6-coding region is similar to the BPS in HPV16. However, the fourth nucleotide in the BPS is differed—uridine for HPV18 and adenosine for HPV16.

In addition to BPS identification, a splicing silencer in the HPV18 E7-coding region at nucleotide 612 to 639 has been

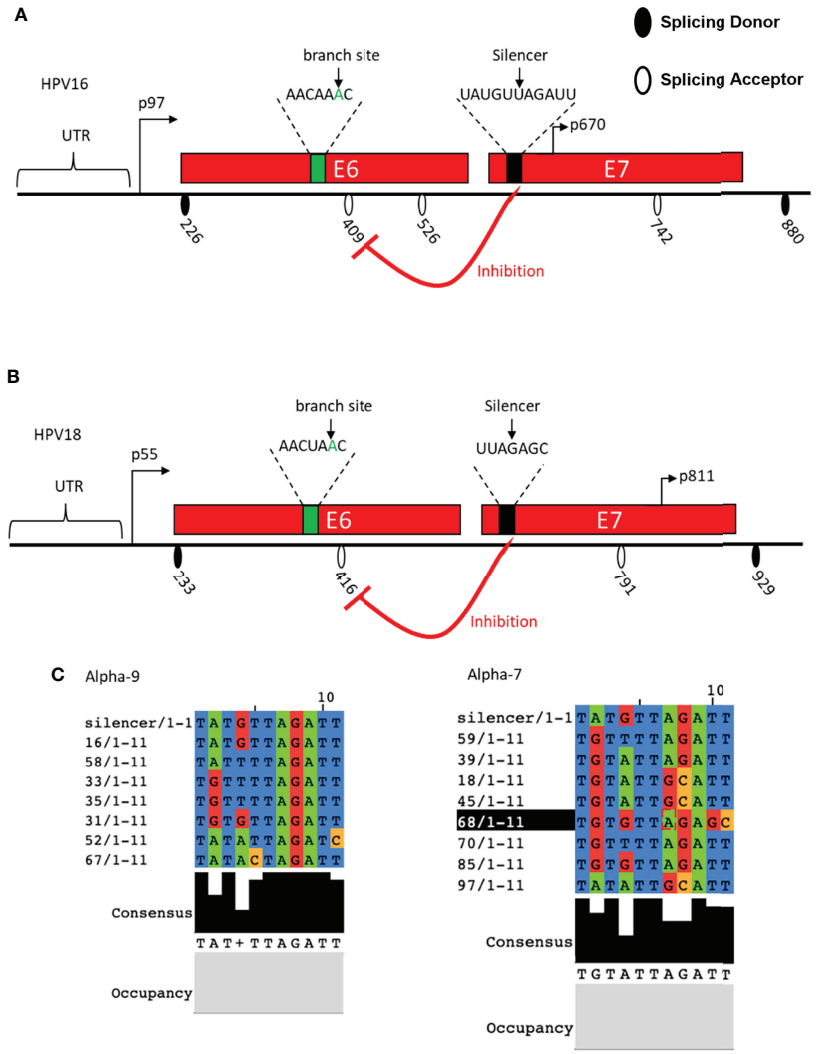


FIGURE 3 | Schematic presentation of cis-regulatory elements identified in HPV16 and HPV18. **(A)** The HPV16 E6- and E7-encoding regions and the splice sites are indicated. Identified branch site and SA409 silencer are indicated. **(B)** HPV18 E6- and E7-encoding regions and the splice sites are indicated. Identified branch site and SA416 silencer are indicated. **(C)** HPV16 silencer sequence alignment with alpha-9 and alpha-7 groups.

mapped to inhibit the splicing from 5′ ss SD233 to 3′ ss SA416 by binding to cellular splicing factor hnRNP A1. The interaction between the mapped silencer and hnRNP A1 inhibits HPV18 E6*I production and E7 expression. These 27 nucleotides in the length silencer contains two hnRNP A1 binding motifs: the sequence of motif one is AAGACA, and that of the other is UUAGAGC. The mutation of the motif AAGACA did not affect the SD233^SA416 splicing efficiency, but the mutation of UUAGAGC resulted in an increase in E6*I expression (Ajiro et al., 2016). This outcome suggested that the sequence UUAGAGC contributes to the interaction with hnRNP A1. Similarly, the HPV16 E7-coding region contains a splicing silencer interacting with hnRNP A1/A2 to inhibit splicing to 3′ ss SA409, thereby reducing the expression of E6*I and E7 production. This silencer locates at nucleotide 594 to 604, consisting of UAUGUUAGAUU. Since HPV16 E7 genetic conservation is essential to carcinogenesis, sequence alignment was therefore performed, and it indicated that the HPV16 11 nucleotide in the length silencer is not well conserved in all high-risk HPVs (Zheng et al., 2020). However, the UAGAU is completely conserved in HPV16 belonging to the alpha-9 subgroup. In alpha-7 group, UAGAU is conserved with HPV18, HPV39, HPV59, HPV70, and HPV85 but less conserved with HPV45, HPV68, and HPV97. Interestingly, all high-risk HPVs have a "UA"-rich region close to the E7 start codon, and one can speculate that the "UA"-rich region is necessary to the silencer's inhibitory activity, which interacted with hnRNP A1/A2 to manipulate the balancing of the E6/E7 protein level and is required for malignant transformation (Zheng et al., 2020) (**Figure 3C**).

Splicing Factors Involved in E6/E7 mRNA Expression

According to the effect of splicing regulation, splicing factors are classified as splicing activator and splicing repressor. However, this division is not strict; the effect of splicing factors can switch depending on the position of the splicing regulatory elements and their activity. The regulation of both HPVs E6/E7 gene expression is tightly regulated by several splicing factors, thereby affecting the splicing pattern change (**Figure 4**).

hnRNP A1/A2/G/D

The splicing factor hnRNP A1 was identified as a splicing repressor interacting with the mapped splicing silencer (UUAGAGC) located at the HPV18 E7-coding region to inhibit splicing to 3′ ss SA416, resulting in a reduction of E6*I and E7 protein (Ajiro et al., 2016). Nevertheless, this study did not reveal how hnRNP A1 interacts with the silencer. In a recent study, the researchers reported that hnRNP A1/A2 inhibits the splicing to 3′ ss SA409 on HPV16 but that this resulted in different consequences. The overexpression of hnRNP A1 leads to an increase of unspliced E6 mRNAs at the expense of E7 mRNAs, while the overexpression of hnRNP A2 leads to alternative splicing to downstream 3′ ss SA742. The splice site SA742 is used to produce E6^E7, E1, and E4 mRNAs (Cerasuolo et al., 2017). Moreover, researchers revealed that the hnRNP A1 inhibitory effect is contributed by the interaction of hnRNP A1 C-terminus with the splicing silencer (UAGAU) in the E7-coding region (Zheng et al., 2020). Taken together, as one of the most abundant splicing factors, hnRNP A1 is reported to participate in the regulation of HPV16 early and late gene expressions, causing an E6/E7 splicing pattern change and demonstrating its vital role in the HPV replication cycle as well as malignant transformation.

A recent study demonstrates that hnRNP G inhibits splicing from HPV16 3′ ss SA409 to prevent HPV16 E7 production. The reduction of E7 levels restores the pRB levels, while the absence of hnRNP G causes a reduction of pRB. The inhibitory effect of hnRNP G may be due to the direct interaction of hnRNP G with the previously reported splicing silencer in the E7-coding region. Meanwhile, the overexpression of hnRNP G promotes the splicing from HPV16 3′ ss SA2709 to increase the production of E2 protein. The different effects of hnRNP G in the regulation of E6, E7, and E2 splicing are contributed by the different domains of hnRNP G. The hnRNP G N-terminal RGG domain and the NTD domain play an essential role in the inhibition of E6/E7 mRNA splicing, while the C-terminal region is required for E2 production. Since E2 protein shuts down the HPV16 early promoter, it thereby repressed the expression of E6 and E7 oncoproteins to pave the way for differentiation and late gene expression. The high levels of pro-apoptotic E2 protein caused by the overexpression of hnRNP G would counteract cell transformation and HPV16-related cancer progression, suggesting that hnRNP G may be used for anticancer or antiviral treatment (Hao et al., 2022).

hnRNP D has four variants, including hnRNP D37, D40, D42, and D45. All four variants have been shown to promote the intron retention of HPV16 E6 mRNAs, resulting in increasing E6 mRNA levels at the expense of E7 levels. The RRM1 and RRM2 domains of hnRNP D40 are necessary for the interaction of hnRNP D40 with HPV16 mRNAs, while the inhibitory effect of hnRNP D40 is contributed by its N-terminal region and C-terminal RGG domain. Besides splicing inhibition, hnRNP D40 promotes the intron-retained E6 mRNA expression level in the cytoplasm (Cui et al., 2022).

Epidermal Growth Factor

The epidermal growth factor (EGF) regulates HPV16 E6/E7 splicing via inducing the activation of the ERK1–ERK2 pathway, which inhibits splicing to 3′ ss SA409. The inhibitory effect of EGF results in the reduction of E6*I but gives rise to unspliced E6 mRNA generation (Rosenberger et al., 2010). However, the mechanism of EGF in the regulation of E6/E7 splicing remains a subject for further study.

5′ Cap-Binding Factors

The previous study has reported that the splicing on 5′ capped E6/E7 mRNAs prefers the 5′ ss SD226 on HPV16. The cap-

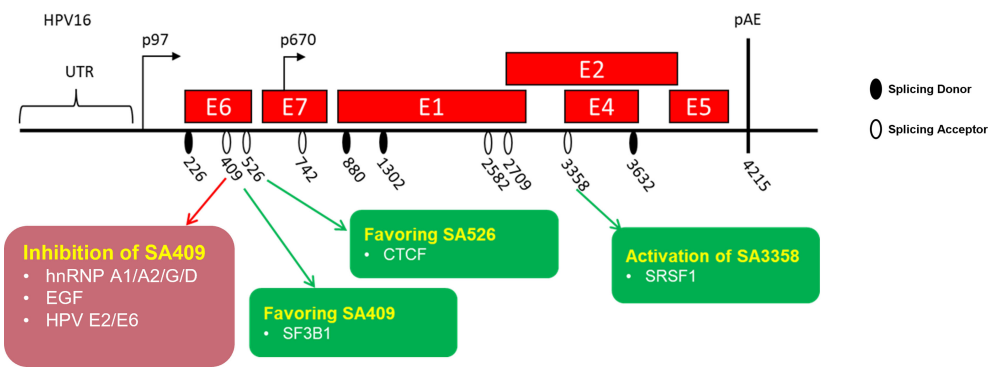


FIGURE 4 | Schematic presentation of identified splicing factors involved in the regulation of human papillomaviruses early oncoprotein E6/E7 RNA splicing. HPV16-encoded viral genes are represented by a "red" box. The splice sites are indicated.

dependent HPV16 E6/E7 splicing in cervical cancer-derived cells is particularly efficient and supposed to be conducted by interacting with cap binding complex. Furthermore, the E6/E7 splicing efficiency from 5′ ss SD226 is impacted by the distance of the cap-proximal intron to the 5′ cap. Further studies identified that the optimal distance is less than 307 nucleotides, which promotes the association of 5′ splice site with U1 snRNP. Once the distance of the 5′ cap to the SD226 increases, splicing is inhibited, promoting unspliced E6 mRNA expression and decreasing E6*I production (Zheng et al., 2004).

SRSF1/SRSF2

The HPV16 3′ ss SA3358 is the most commonly used splice site in the HPV16 genome. SA3358 generates early mRNAs encoding early proteins, including E6 and E7. The splicing factor serine/arginine-rich (SR) protein SRSF1 has been reported to enhance splicing to 3′ ss SA3358 by interacting with a splicing enhancer (ACCGAAGAA) located downstream of the splice site SA3358. Since SA3358 is used by the majority of the E6 and E7 mRNAs, high levels of SRSF1 function as proto-oncoprotein to upregulate the expression of E6/E7, further resulting in hyperproliferation and, possibly, cancer occurrence (Somberg and Schwartz, 2010). Another SR protein SRSF2 has been shown to be required for E6E7 mRNA production only in cervical cancer-derived cells. The role of SRSF2 is to maintain E6/E7 mRNA stability and inhibit their decay. The knock-down of SRSF2 in cervical tumor cells results in cell apoptosis, suggesting its oncogenic role in cervical cancer progression (McFarlane et al., 2015).

HPVs E2 and E6 Proteins

HPV16 early E2 protein and early oncoprotein E6 have been shown to have an inhibitory effect on HPV16 E6/E7 pre-mRNA splicing (Gomes and Espinosa, 2010). HPV16 E2 and E6 directly bind to the intron located in between SD226 and SA409 through their RNA interaction domain in the C-terminus, resulting in the reduction of the E6*I isoform in HPV16-infected cells (Gomes and Espinosa, 2010). This inhibitory effect may also be due to E2 and E6 interfering with several SR proteins, including SRSF9, SRSF6, and SRSF5. Taken together, these findings demonstrate that HPV16 viral proteins E2 and E6 function as splicing regulatory proteins which affect E6/E7 pre-mRNA splicing through their interaction with SR proteins (Bodaghi et al., 2009).

CTCF and SF3B1

The CCCTC-binding factor known as CTCF has been firstly identified as a zinc finger DNA-binding transcription factor. Over the past decade, the new tricks of CTCF has been elucidated, especially its new role as a RNA-binding alternative splicing factor (Gomes and Espinosa, 2010; Shukla et al., 2011; Guo et al., 2012; Monahan et al., 2012; Marina et al., 2016). A previous study has shown that CTCF binds to the HPV18 E2-coding region to increase the production of the E6*II isoform. The mutation of CTCF results in the increased production of unspliced E6 mRNA and spliced E6*I mRNA (Paris et al., 2015). It is reasonable to speculate that the increase of E6*II at the expense of unspliced E6 and E6*I mRNA has interrupted the balance of the E6/E7 ratio, causing infected cell apoptosis. However, the mechanism of CTCF in the regulation of E6/E7 expression was not well understood until recently. Researchers have revealed that the inhibitory effect of CTCF is dependent on CTCF and YY-1 loop formation in the HPV18 genome. The downregulation of YY-1 causes an interruption of loop formation and a reversal of epigenetic silencing (Pentland et al., 2018).

The splicing factor 3B1 (SF3B1) was reported to modulate HPV16 E6/E7 pre-mRNA splicing, resulting in increased E6*I in HPV16-positive head and neck cancer cells. The treatment of SF3B1 inhibitor meayamycin B causes the downregulation of the transcript E6*I but promotes the level of the unspliced E6 mRNAs (Gao et al., 2014). These findings further suggest the anti-oncogenic role of SF3B1 in HPV16-derived tumorigenesis. However, the exact mechanism of how SF3B1 affects HPV16 splicing has not been elucidated.

CONCLUSION

At present, there is no effective treatment for HPV infection and HPV-related cancer. Although there are vaccines against several high-risk HPV types, the vaccines have no effect on infected patients, and vaccines are not universally accessible due to the high price. Therefore, continuous research on the carcinogenic mechanism of HPV is of great significance for the prevention and treatment of HPV infections and their related cancers. As to the oncogenic proteins expressed in the early replication cycle of high-risk HPVs, the splicing regulation mechanism of E6 and E7 pre-mRNA has a profound impact on the entire replication cycle of HPVs. Moreover, the differentiation, proliferation, and immortalization transformation of HPV-infected cells are tightly related to the level of E6 and E7. Transcripts encoding E6 and E7 proteins are derived from the same pre-mRNA; the selection of splice site SA409 in HPV16 yields the most abundant isoform E6*I, which is further translated into the oncoprotein E7; the unspliced mRNA is translated into the oncoprotein E6 since it contains the complete E6-coding region. Such a mechanism indicates that a perfect balance between E6 and E7needs is required. In HPV16, if the splicing from SA409 becomes too efficient, the production of E6*I will be upregulated, resulting in increased E7 level, but E6 level is decreased, thereby causing apoptosis of the infected cells. Conversely, if the unspliced E6 mRNA is upregulated, the level of E7 will decrease, and the fate of the infected cells will go toward apoptosis. This balancing mechanism between E6 and E7 is regulated by the interaction of cellular splicing factors with cis-regulatory RNA elements on HPV pre-mRNA. A series of RNA regulatory elements and splicing factors are identified

to regulate E6/E7 mRNA splicing, indicating that the E6/E7 expression levels are tightly regulated during posttranscription. The role of various splicing factors in HPV-related cancer progression has also been characterized. Overexpression or knockdown of splicing factors will interrupt the ratio of E6/E7 proteins, thereby causing infected cell apoptosis. This weak point of high-risk HPV provides us with an opportunity to identify small molecules that can be used as a splicing factor inhibitor, paving the way to explore more novel antiviral and anticancer drugs.

AUTHOR CONTRIBUTIONS

Manuscript writing: YZ and XL. Figure and table making: YJ. Comments and final review: CW. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (grant 82072287).

REFERENCES

Ajiro, M., Jia, R., Zhang, L., Liu, X., and Zheng, Z. M. (2012). Intron Definition and a Branch Site Adenosine at Nt 385 Control RNA Splicing of HPV16 E6*I and E7 Expression. *PloS One* 7 (10), e46412. doi: 10.1371/journal.pone.0046412

Ajiro, M., Tang, S., Doorbar, J., and Zheng, Z. M. (2016). Serine/Arginine-Rich Splicing Factor 3 and Heterogeneous Nuclear Ribonucleoprotein A1 Regulate Alternative RNA Splicing and Gene Expression of Human Papillomavirus 18 Through Two Functionally Distinguishable Cis Elements. *J. Virol.* 90 (20), 9138–9152. doi: 10.1128/JVI.00965-16

Ajiro, M., and Zheng, Z. M. (2014). Oncogenes and RNA Splicing of Human Tumor Viruses. *Emerg. Microbes Infect.* 3 (9), e63. doi: 10.1038/emi.2014.62

Ajiro, M., and Zheng, Z. M. (2015). E6^E7, a Novel Splice Isoform Protein of Human Papillomavirus 16, Stabilizes Viral E6 and E7 Oncoproteins *via* HSP90 and GRP78. *mBio.* 6 (1), e02068–e02014. doi: 10.1128/mBio.02068-14

Arbyn, M., de Sanjose, S., Saraiya, M., Sideri, M., Palefsky, J., Lacey, C., et al. (2012). EUROGIN 2011 Roadmap on Prevention and Treatment of HPV-Related Disease. *Int. J. Cancer.* 131 (9), 1969–1982. doi: 10.1002/ijc.27650

Beaudoing, E., Freier, S., Wyatt, J. R., Claverie, J. M., and Gautheret, D. (2000). Patterns of Variant Polyadenylation Signal Usage in Human Genes. *Genome Res.* 10 (7), 1001–1010. doi: 10.1101/gr.10.7.1001

Bernard, H. U., Oltersdorf, T., and Seedorf, K. (1987). Expression of the Human Papillomavirus Type 18 E7 Gene by a Cassette-Vector System for the Transcription and Translation of Open Reading Frames in Eukaryotic Cells. *EMBO J.* 6 (1), 133–138. doi: 10.1002/j.1460-2075.1987.tb04730.x

Bienroth, S., Keller, W., and Wahle, E. (1993). Assembly of a Processive Messenger RNA Polyadenylation Complex. *EMBO J.* 12 (2), 585–594. doi: 10.1002/j.1460-2075.1993.tb05690.x

Black, D. L. (2003). Mechanisms of Alternative Pre-Messenger RNA Splicing. *Annu. Rev. Biochem.* 72, 291–336. doi: 10.1146/annurev.biochem.72.121801.161720

Bodaghi, S., Jia, R., and Zheng, Z. M. (2009). Human Papillomavirus Type 16 E2 and E6 are RNA-Binding Proteins and Inhibit *In Vitro* Splicing of pre-mRNAs With Suboptimal Splice Sites. *Virology.* 386 (1), 32–43. doi: 10.1016/j.virol.2008.12.037

Borgogna, C., Zavattaro, E., De Andrea, M., Griffin, H. M., Dell'Oste, V., Azzimonti, B., et al. (2012). Characterization of Beta Papillomavirus E4 Expression in Tumours From Epidermodysplasia Verruciformis Patients and in Experimental Models. *Virology.* 423 (2), 195–204. doi: 10.1016/j.virol.2011.11.029

Brant, A. C., Majerciak, V., Moreira, M. A. M., and Zheng, Z. M. (2019). HPV18 Utilizes Two Alternative Branch Sites for E6*I Splicing to Produce E7 Protein. *Virol. Sin.* 34 (2), 211–221. doi: 10.1007/s12250-019-00098-0

Bzhalava, D., Guan, P., Franceschi, S., Dillner, J., and Clifford, G. (2013). A Systematic Review of the Prevalence of Mucosal and Cutaneous Human Papillomavirus Types. *Virology.* 445 (1-2), 224–231. doi: 10.1016/j.virol.2013.07.015

Cerasuolo, A., Annunziata, C., Tortora, M., Starita, N., Stellato, G., Greggi, S., et al. (2017). Comparative Analysis of HPV16 Gene Expression Profiles in Cervical and in Oropharyngeal Squamous Cell Carcinoma. *Oncotarget* 8 (21), 34070–34081. doi: 10.18632/oncotarget.15977

Cerasuolo, A., Buonaguro, L., Buonaguro, F. M., and Tornesello, M. L. (2020). The Role of RNA Splicing Factors in Cancer: Regulation of Viral and Human Gene Expression in Human Papillomavirus-Related Cervical Cancer. *Front. Cell Dev. Biol.* 8, 474. doi: 10.3389/fcell.2020.00474

Cui, X., Hao, C., Gong, L., Kajitani, N., and Schwartz, S. (2022). HnRNP D Activates Production of HPV16 E1 and E6 mRNAs by Promoting Intron Retention. *Nucleic Acids Res.* 50 (5), 2782–2806. doi: 10.1093/nar/gkac132

De la Rosa-Rios, M. A., Martinez-Salazar, M., Martinez-Garcia, M., Gonzalez-Bonilla, C., and Villegas-Sepulveda, N. (2006). The Intron 1 of HPV 16 has a Suboptimal Branch Point at a Guanosine. *Virus Res.* 118 (1-2), 46–54. doi: 10.1016/j.virusres.2005.11.010

de Sanjose, S., Quint, W. G., Alemany, L., Geraets, D. T., Klaustermeier, J. E., Lloveras, B., et al. (2010). Human Papillomavirus Genotype Attribution in Invasive Cervical Cancer: A Retrospective Cross-Sectional Worldwide Study. *Lancet Oncol.* 11 (11), 1048–1056. doi: 10.1016/S1470-2045(10)70230-8

de Villiers, E. M. (2013). Cross-Roads in the Classification of Papillomaviruses. *Virology* 445 (1-2), 2–10. doi: 10.1016/j.virol.2013.04.023

de Villiers, E. M., Fauquet, C., Broker, T. R., Bernard, H. U., and Zur Hausen, H. (2004). Classification of papillomaviruses. *Virology* 324 (1), 17–27. doi: 10.1016/j.virol.2004.03.033

Doorbar, J. (2005). The Papillomavirus Life Cycle. *J. Clin. Virol.* 32 Suppl 1, S7–15. doi: 10.1016/j.jcv.2004.12.006

Doorbar, J., Parton, A., Hartley, K., Banks, L., Crook, T., Stanley, M., et al. (1990). Detection of Novel Splicing Patterns in a HPV16-Containing Keratinocyte Cell Line. *Virology* 178 (1), 254–262. doi: 10.1016/0042-6822(90)90401-C

Egawa, K. (2003). Do Human Papillomaviruses Target Epidermal Stem Cells? *Dermatology* 207 (3), 251–254. doi: 10.1159/000073085

Elliott, D. J., Oghene, K., Makarov, G., Makarova, O., Hargreave, T. B., Chandley, A. C., et al. (1998). Dynamic Changes in the Subnuclear Organisation of pre-mRNA Splicing Proteins and RBM During Human Germ Cell Development. *J. Cell Sci.* 111 (Pt 9), 1255–1265. doi: 10.1242/jcs.111.9.1255

Filippova, M., Evans, W., Aragon, R., Filippov, V., Williams, V. M., Hong, L., et al. (2014). The Small Splice Variant of HPV16 E6, E6, Reduces Tumor Formation in Cervical Carcinoma Xenografts. *Virology* 450-451, 153–164. doi: 10.1016/j.virol.2013.12.011

Filippova, M., Johnson, M. M., Bautista, M., Filippov, V., Fodor, N., Tungteakkhun, S. S., et al. (2007). The Large and Small Isoforms of Human Papillomavirus Type 16 E6 Bind to and Differentially Affect Procaspase 8 Stability and Activity. *J. Virol.* 81 (8), 4116–4129. doi: 10.1128/JVI.01924-06

Gao, Y., Trivedi, S., Ferris, R. L., and Koide, K. (2014). Regulation of HPV16 E6 and MCL1 by SF3B1 Inhibitor in Head and Neck Cancer Cells. *Sci. Rep.* 4, 6098. doi: 10.1038/srep06098

Garland, S. M., Hernandez-Avila, M., Wheeler, C. M., Perez, G., Harper, D. M., Leodolter, S., et al. (2007). Quadrivalent Vaccine Against Human Papillomavirus to Prevent Anogenital Diseases. *N. Engl. J. Med.* 356 (19), 1928–1943. doi: 10.1056/NEJMoa061760

Gomes, N. P., and Espinosa, J. M. (2010). Gene-Specific Repression of the P53 Target Gene PUMA *via* Intragenic CTCF-Cohesin Binding. *Genes Dev.* 24 (10), 1022–1034. doi: 10.1101/gad.1881010

Goodwin, E. C., and DiMaio, D. (2000). Repression of Human Papillomavirus Oncogenes in HeLa Cervical Carcinoma Cells Causes the Orderly Reactivation of Dormant Tumor Suppressor Pathways. *Proc. Natl. Acad. Sci. U.S.A.* 97 (23), 12513–12518. doi: 10.1073/pnas.97.23.12513

Graham, S. V., and Faizo, A. A. A. (2017). Control of Human Papillomavirus Gene Expression by Alternative Splicing. *Virus Res.* 231, 83–95. doi: 10.1016/j.virusres.2016.11.016

Graveley, B. R., Hertel, K. J., and Maniatis, T. (1999). SR Proteins are 'Locators' of the RNA Splicing Machinery. *Curr. Biol.* 9 (1), R6–R7. doi: 10.1016/S0960-9822(99)80032-3

Groves, I. J., and Coleman, N. (2015). Pathogenesis of Human Papillomavirus-Associated Mucosal Disease. *J. Pathol.* 235 (4), 527–538. doi: 10.1002/path.4496

Guccione, E., Pim, D., and Banks, L. (2004). HPV-18 E6*I Modulates HPV-18 Full-Length E6 Functions in a Cell Cycle Dependent Manner. *Int. J. Cancer.* 110 (6), 928–933. doi: 10.1002/ijc.20184

Guo, Y., Monahan, K., Wu, H., Gertz, J., Varley, K. E., Li, W., et al. (2012). CTCF/Cohesin-Mediated DNA Looping is Required for Protocadherin Alpha Promoter Choice. *Proc. Natl. Acad. Sci. U.S.A.* 109 (51), 21081–21086. doi: 10.1073/pnas.1219280110

Halec, G., Schmitt, M., Dondog, B., Sharkhuu, E., Wentzensen, N., Gheit, T., et al. (2013). Biological Activity of Probable/Possible High-Risk Human Papillomavirus Types in Cervical Cancer. *Int. J. Cancer.* 132 (1), 63–71. doi: 10.1002/ijc.27605

Hallegger, M., Llorian, M., and Smith, C. W. (2010). Alternative Splicing: Global Insights. *FEBS J.* 277 (4), 856–866. doi: 10.1111/j.1742-4658.2009.07521.x

Hanahan, D., and Weinberg, R. A. (2000). The Hallmarks of Cancer. *Cell.* 100 (1), 57–70. doi: 10.1016/S0092-8674(00)81683-9

Han, S. P., Tang, Y. H., and Smith, R. (2010). Functional Diversity of the hnRNPs: Past, Present and Perspectives. *Biochem. J.* 430 (3), 379–392. doi: 10.1042/BJ20100396

Hao, C., Zheng, Y., Jonsson, J., Cui, X., Yu, H., Wu, C., et al. (2022). hnRNP G/RBMX Enhances HPV16 E2 mRNA Splicing Through a Novel Splicing Enhancer and Inhibits Production of Spliced E7 Oncogene mRNAs. *Nucleic Acids Res.* 50 (7), 3867–3891. doi: 10.1093/nar/gkac213

Hertel, K. J., and Maniatis, T. (1999). Serine-Arginine (SR)-Rich Splicing Factors Have an Exon-Independent Function in pre-mRNA Splicing. *Proc. Natl. Acad. Sci. U.S.A.* 96 (6), 2651–2655. doi: 10.1073/pnas.96.6.2651

Howie, H. L., Katzenellenbogen, R. A., and Galloway, D. A. (2009). Papillomavirus E6 Proteins. *Virology* 384 (2), 324–334. doi: 10.1016/j.virol.2008.11.017

Howley, P. M., and Pfister, H. J. (2015). Beta Genus Papillomaviruses and Skin Cancer. *Virology.* 479-480, 290–296. doi: 10.1016/j.virol.2015.02.004

Hummel, M., Hudson, J. B., and Laimins, L. A. (1992). Differentiation-Induced and Constitutive Transcription of Human Papillomavirus Type 31b in Cell Lines Containing Viral Episomes. *J. Virol.* 66 (10), 6070–6080. doi: 10.1128/jvi.66.10.6070-6080.1992

Islam, S., Dasgupta, H., Roychowdhury, A., Bhattacharya, R., Mukherjee, N., Roy, A., et al. (2017). Study of Association and Molecular Analysis of Human Papillomavirus in Breast Cancer of Indian Patients: Clinical and Prognostic Implication. *PloS One* 12 (2), e0172760. doi: 10.1371/journal.pone.0172760

Johansson, C., and Schwartz, S. (2013). Regulation of Human Papillomavirus Gene Expression by Splicing and Polyadenylation. *Nat. Rev. Microbiol.* 11 (4), 239–251. doi: 10.1038/nrmicro2984

Kato, J., Matsushime, H., Hiebert, S. W., Ewen, M. E., and Sherr, C. J. (1993). Direct Binding of Cyclin D to the Retinoblastoma Gene Product (pRb) and pRb Phosphorylation by the Cyclin D-Dependent Kinase CDK4. *Genes Dev.* 7 (3), 331–342. doi: 10.1101/gad.7.3.331

Li, X., and Coffino, P. (1996). High-Risk Human Papillomavirus E6 Protein has Two Distinct Binding Sites Within P53, of Which Only One Determines Degradation. *J. Virol.* 70 (7), 4509–4516. doi: 10.1128/jvi.70.7.4509-4516.1996

Li, N., Franceschi, S., Howell-Jones, R., Snijders, P. J., and Clifford, G. M. (2011). Human Papillomavirus Type Distribution in 30,848 Invasive Cervical Cancers Worldwide: Variation by Geographical Region, Histological Type and Year of Publication. *Int. J. Cancer.* 128 (4), 927–935. doi: 10.1002/ijc.25396

Lim, K. H., Ferraris, L., Filloux, M. E., Raphael, B. J., and Fairbrother, W. G. (2011). Using Positional Distribution to Identify Splicing Elements and Predict pre-mRNA Processing Defects in Human Genes. *Proc. Natl. Acad. Sci. U.S.A.* 108 (27), 11093–11098. doi: 10.1073/pnas.1101135108

Liu, N., and Pan, T. (2016). N6-Methyladenosine-Encoded Epitranscriptomics. *Nat. Struct. Mol. Biol.* 23 (2), 98–102. doi: 10.1038/nsmb.3162

Li, Y., Wang, X., Ni, T., Wang, F., Lu, W., Zhu, J., et al. (2013). Human Papillomavirus Type 58 Genome Variations and RNA Expression in Cervical Lesions. *J. Virol.* 87 (16), 9313–9322. doi: 10.1128/JVI.01154-13

Lorincz, A. T., Reid, R., Jenson, A. B., Greenberg, M. D., Lancaster, W., and Kurman, R. J. (1992). Human Papillomavirus Infection of the Cervix: Relative Risk Associations of 15 Common Anogenital Types. *Obstet Gynecol.* 79 (3), 328–337. doi: 10.1097/00006250-199203000-00002

Manley, J. L., and Tacke, R. (1996). SR Proteins and Splicing Control. *Genes Dev.* 10 (13), 1569–1579. doi: 10.1101/gad.10.13.1569

Manzo-Merino, J., Massimi, P., Lizano, M., and Banks, L. (2014). The Human Papillomavirus (HPV) E6 Oncoproteins Promotes Nuclear Localization of Active Caspase 8. *Virology.* 450-451, 146–152. doi: 10.1016/j.virol.2013.12.013

Marina, R. J., Sturgill, D., Bailly, M. A., Thenoz, M., Varma, G., Prigge, M. F., et al. (2016). TET-Catalyzed Oxidation of Intragenic 5-Methylcytosine Regulates CTCF-Dependent Alternative Splicing. *EMBO J.* 35 (3), 335–355. doi: 10.15252/embj.201593235

Martinez-Zapien, D., Ruiz, F. X., Poirson, J., Mitschler, A., Ramirez, J., Forster, A., et al. (2016). Structure of the E6/E6AP/p53 Complex Required for HPV-Mediated Degradation of P53. *Nature.* 529 (7587), 541–545. doi: 10.1038/nature16481

Matsukura, T., Koi, S., and Sugase, M. (1989). Both Episomal and Integrated Forms of Human Papillomavirus Type 16 are Involved in Invasive Cervical Cancers. *Virology.* 172 (1), 63–72. doi: 10.1016/0042-6822(89)90107-4

McFarlane, M., MacDonald, A. I., Stevenson, A., and Graham, S. V. (2015). Human Papillomavirus 16 Oncoprotein Expression Is Controlled by the Cellular Splicing Factor SRSF2 (Sc35). *J. Virol.* 89 (10), 5276–5287. doi: 10.1128/JVI.03434-14

Mesplede, T., Gagnon, D., Bergeron-Labrecque, F., Azar, I., Senechal, H., Coutlee, F., et al. (2012). P53 Degradation Activity, Expression, and Subcellular Localization of E6 Proteins From 29 Human Papillomavirus Genotypes. *J. Virol.* 86 (1), 94–107. doi: 10.1128/JVI.00751-11

Monahan, K., Rudnick, N. D., Kehayova, P. D., Pauli, F., Newberry, K. M., Myers, R. M., et al. (2012). Role of CCCTC Binding Factor (CTCF) and Cohesin in the Generation of Single-Cell Diversity of Protocadherin-Alpha Gene Expression. *Proc. Natl. Acad. Sci. U S A.* 109 (23), 9125–9130. doi: 10.1073/pnas.1205074109

Munger, K., Basile, J. R., Duensing, S., Eichten, A., Gonzalez, S. L., Grace, M., et al. (2001). Biological Activities and Molecular Targets of the Human Papillomavirus E7 Oncoprotein. *Oncogene.* 20 (54), 7888–7898. doi: 10.1038/sj.onc.1204860

Olmedo-Nieva, L., Munoz-Bello, J. O., Contreras-Paredes, A., and Lizano, M. (2018). The Role of E6 Spliced Isoforms (E6*) in Human Papillomavirus-Induced Carcinogenesis. *Viruses.* 10 (1), 1–20. doi: 10.3390/v10010045

Ozbun, M. A. (2002). Human Papillomavirus Type 31b Infection of Human Keratinocytes and the Onset of Early Transcription. *J. Virol.* 76 (22), 11291–11300. doi: 10.1128/JVI.76.22.11291-11300.2002

Paris, C., Pentland, I., Groves, I., Roberts, D. C., Powis, S. J., Coleman, N., et al. (2015). CCCTC-Binding Factor Recruitment to the Early Region of the Human Papillomavirus 18 Genome Regulates Viral Oncogene Expression. *J. Virol.* 89 (9), 4770–4785. doi: 10.1128/JVI.00097-15

Parkin, D. M., and Bray, F. (2006). Chapter 2: The Burden of HPV-Related Cancers. *Vaccine.* 24 Suppl 3, S3/11–S3/25. doi: 10.1016/j.vaccine.2006.05.111

Patel, D., Huang, S. M., Baglia, L. A., and McCance, D. J. (1999). The E6 Protein of Human Papillomavirus Type 16 Binds to and Inhibits Co-Activation by CBP and P300. *EMBO J.* 18 (18), 5061–5072. doi: 10.1093/emboj/18.18.5061

Pentland, I., Campos-Leon, K., Cotic, M., Davies, K. J., Wood, C. D., Groves, I. J., et al. (2018). Disruption of CTCF-YY1-Dependent Looping of the Human Papillomavirus Genome Activates Differentiation-Induced Viral Oncogene Transcription. *PloS Biol.* 16 (10), e2005752. doi: 10.1371/journal.pbio.2005752

Pim, D., Massimi, P., and Banks, L. (1997). Alternatively Spliced HPV-18 E6* Protein Inhibits E6 Mediated Degradation of P53 and Suppresses Transformed Cell Growth. *Oncogene.* 5 (3), 257–264. doi: 10.1038/sj.onc.1201202

Pyeon, D., Pearce, S. M., Lank, S. M., Ahlquist, P., and Lambert, P. F. (2009). Establishment of Human Papillomavirus Infection Requires Cell Cycle Progression. *PloS Pathog.* 5 (2), e1000318. doi: 10.1371/journal.ppat.1000318

Richardson, H., Kelsall, G., Tellier, P., Voyer, H., Abrahamowicz, M., Ferenczy, A., et al. (2003). The Natural History of Type-Specific Human Papillomavirus Infections in Female University Students. *Cancer Epidemiol. Biomarkers Prev.* 12 (6), 485–490.

Rosenberger, S., De-Castro Arce, J., Langbein, L., Steenbergen, R. D., and Rosl, F. (2010). Alternative Splicing of Human Papillomavirus Type-16 E6/E6* Early mRNA is Coupled to EGF Signaling *via* Erk1/2 Activation. *Proc. Natl. Acad. Sci. U.S.A.* 107 (15), 7006–7011. doi: 10.1073/pnas.1002620107

Schafer, G., Blumenthal, M. J., and Katz, A. A. (2015). Interaction of Human Tumor Viruses With Host Cell Surface Receptors and Cell Entry. *Viruses.* 7 (5), 2592–2617. doi: 10.3390/v7052592

Scheffner, M., Huibregtse, J. M., Vierstra, R. D., and Howley, P. M. (1993). The HPV-16 E6 and E6-AP Complex Functions as a Ubiquitin-Protein Ligase in the Ubiquitination of P53. *Cell.* 75 (3), 495–505. doi: 10.1016/0092-8674(93)90384-3

Schmitt, M., Dalstein, V., Waterboer, T., Clavel, C., Gissmann, L., and Pawlita, M. (2010). Diagnosing Cervical Cancer and High-Grade Precursors by HPV16 Transcription Patterns. *Cancer Res.* 70 (1), 249–256. doi: 10.1158/0008-5472.CAN-09-2514

Schmitt, M., and Pawlita, M. (2011). The HPV Transcriptome in HPV16 Positive Cell Lines. *Mol. Cell Probes* 25 (2-3), 108–113. doi: 10.1016/j.mcp.2011.03.003

Sedman, S. A., Barbosa, M. S., Vass, W. C., Hubbert, N. L., Haas, J. A., Lowy, D. R., et al. (1991). The Full-Length E6 Protein of Human Papillomavirus Type 16 has Transforming and Trans-Activating Activities and Cooperates With E7 to Immortalize Keratinocytes in Culture. *J. Virol.* 65 (9), 4860–4866. doi: 10.1128/jvi.65.9.4860-4866.1991

Shin, K. H., Kim, R. H., Kang, M. K., Kim, R. H., Kim, S. G., Lim, P. K., et al. (2007). P53 Promotes the Fidelity of DNA End-Joining Activity by, in Part, Enhancing the Expression of Heterogeneous Nuclear Ribonucleoprotein G. *DNA Repair (Amst).* 6 (6), 830–840. doi: 10.1016/j.dnarep.2007.01.013

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., et al. (2011). CTCF-Promoted RNA Polymerase II Pausing Links DNA Methylation to Splicing. *Nature.* 479 (7371), 74–79. doi: 10.1038/nature10442

Snijders, P. J., van den Brule, A. J., Schrijnemakers, H. F., Raaphorst, P. M., Meijer, C. J., and Walboomers, J. M. (1992). Human Papillomavirus Type 33 in a Tonsillar Carcinoma Generates its Putative E7 mRNA *via* Two E6* Transcript Species Which are Terminated at Different Early Region Poly(A) Sites. *J. Virol.* 66 (5), 3172–3178. doi: 10.1128/jvi.66.5.3172-3178.1992

Somberg, M., and Schwartz, S. (2010). Multiple ASF/SF2 Sites in the Human Papillomavirus Type 16 (HPV-16) E4-Coding Region Promote Splicing to the Most Commonly Used 3'-Splice Site on the HPV-16 Genome. *J. Virol.* 84 (16), 8219–8230. doi: 10.1128/JVI.00462-10

Stanley, M. (2008). Immunobiology of HPV and HPV Vaccines. *Gynecol Oncol.* 109 (2 Suppl), S15–S21. doi: 10.1016/j.ygyno.2008.02.003

Taggart, A. J., DeSimone, A. M., Shih, J. S., Filloux, M. E., and Fairbrother, W. G. (2012). Large-Scale Mapping of Branchpoints in Human pre-mRNA Transcripts *In Vivo*. *Nat. Struct. Mol. Biol.* 19 (7), 719–721. doi: 10.1038/nsmb.2327

Toots, M., Mannik, A., Kivi, G., Ustav, M.Jr., Ustav, E., and Ustav, M. (2014). The Transcription Map of Human Papillomavirus Type 18 During Genome Replication in U2OS Cells. *PloS One* 9 (12), e116151. doi: 10.1371/journal.pone.0116151

Topisirovic, I., Svitkin, Y. V., Sonenberg, N., and Shatkin, A. J. (2011). Cap and Cap-Binding Proteins in the Control of Gene Expression. *Wiley Interdiscip Rev. RNA.* 2 (2), 277–298. doi: 10.1002/wrna.52

Vazquez-Vega, S., Sanchez-Suarez, L. P., Andrade-Cruz, R., Castellanos-Juarez, E., Contreras-Paredes, A., Lizano-Soberon, M., et al. (2013). Regulation of P14arf Expression by HPV-18 E6 Variants. *J. Med. Virol.* 85 (7), 1215–1221. doi: 10.1002/jmv.23568

Vinokurova, S., Wentzensen, N., Kraus, I., Klaes, R., Driesch, C., Melsheimer, P., et al. (2008). Type-Dependent Integration Frequency of Human Papillomavirus Genomes in Cervical Lesions. *Cancer Res.* 68 (1), 307–313. doi: 10.1158/0008-5472.CAN-07-2754

Wang, Z., Xiao, X., Van Nostrand, E., and Burge, C. B. (2006). General and Specific Functions of Exonic Splicing Silencers in Splicing Control. *Mol. Cell.* 23 (1), 61–70. doi: 10.1016/j.molcel.2006.05.018

Zheng, Z. M., and Baker, C. C. (2006). Papillomavirus Genome Structure, Expression, and Post-Transcriptional Regulation. *Front. Biosci.* 11, 2286–2302. doi: 10.2741/1971

Zheng, Y., Jonsson, J., Hao, C., Shoja Chaghervand, S., Cui, X., Kajitani, N., et al. (2020). Heterogeneous Nuclear Ribonucleoprotein A1 (hnRNP A1) and hnRNP A2 Inhibit Splicing to Human Papillomavirus 16 Splice Site SA409 Through a UAG-Containing Sequence in the E7 Coding Region. *J. Virol.* 94 (20), e01509–20. doi: 10.1128/JVI.01509-20

Zheng, Z. M., Tao, M., Yamanegi, K., Bodaghi, S., and Xiao, W. (2004). Splicing of a Cap-Proximal Human Papillomavirus 16 E6E7 Intron Promotes E7 Expression, But can be Restrained by Distance of the Intron From its RNA 5' Cap. *J. Mol. Biol.* 337 (5), 1091–1108. doi: 10.1016/j.jmb.2004.02.023

Zhou, Z., and Fu, X. D. (2013). Regulation of Splicing by SR Proteins and SR Protein-Specific Kinases. *Chromosoma.* 122 (3), 191–207. doi: 10.1007/s00412-013-0407-z

Zimmermann, H., Koh, C. H., Degenkolbe, R., O'Connor, M. J., Muller, A., Steger, G., et al. (2000). Interaction With CBP/p300 Enables the Bovine Papillomavirus Type 1 E6 Oncoprotein to Downregulate CBP/p300-Mediated Transactivation by P53. *J. Gen. Virol.* 81 (Pt 11), 2617–2623. doi: 10.1099/0022-1317-81-11-2617

zur Hausen, H. (2002). Papillomaviruses and Cancer: From Basic Studies to Clinical Application. *Nat. Rev. Cancer.* 2 (5), 342–350. doi: 10.1038/nrc798

zur Hausen, H. (2009). Papillomaviruses in the Causation of Human Cancers - a Brief Historical Account. *Virology*. 384 (2), 260–265. doi: 10.1016/j.virol.2008.11.046

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Cellular and Infection Microbiology

PERSPECTIVE
published: 13 July 2022
doi: 10.3389/fcimb.2022.916506

Small Cell (Neuroendocrine) Carcinoma of the Cervix: An Analysis for 19 Cases and Literature Review

*JunLing Lu, Ya Li and Jun Wang**

Department of Gynecology and Obstetrics, The Second Affiliated Hospital of Dalian Medical University, Dalian, China

OPEN ACCESS

Edited by:
Chengjun Wu,
Dalian University of Technology, China

Reviewed by:
Xiaojun Liu,
Shanghai Changzheng Hospital, China
Yi Zhang,
The First Affiliated Hospital of China Medical University, China
Qing Yang,
ShengJing Hospital of China Medical University, China

***Correspondence:**
Jun Wang
wj202fck@163.com

Specialty section:
This article was submitted to Clinical Microbiology, a section of the journal Frontiers in Cellular and Infection Microbiology

Received: *09 April 2022*
Accepted: *18 May 2022*
Published: *13 July 2022*

Citation:
Lu J, Li Y and Wang J (2022) Small Cell (Neuroendocrine) Carcinoma of the Cervix: An Analysis for 19 Cases and Literature Review. Front. Cell. Infect. Microbiol. 12:916506. doi: 10.3389/fcimb.2022.916506

Cervical SCNEC is a rare and highly malignant invasive tumor. The incidence is low, at less than 5% of all cervical cancers. Moreover, most patients with small cell carcinoma are interrelated with high risk HPV (more familiar HPV 18). Compared to squamous cell carcinoma or adenocarcinoma, patients of cevical SCNEC are more prone to lymph node invasion early, so the clinical manifestation is usually local or distant metastasis. We summarized the clinical features of 19 patients with cervical small cell carcinoma in the Second Affiliated Hospital of Dalian Medical University from 2012 to 2021, and retrospectively analyzed data from 1576 patients in 20 related studies and more than 50 pieces of literature in recent years by searching PubMed, Google schalor, Cochrane Library, Clinicalkey, and other databases. The collected patient data included age, clinical manifestation, TCT, HPV detection, the size and morphology of the tumor, local invasion depth, stage, lymph node status, initial treatment method, tumor-free survival, and so on. The positive rates of CGA, SYN, and CD56 in our cases were high, and NSE was a moderately sensitive index. P16 and Ki67 were the most sensitive, and all patients were positive. We found that multimodal treatment can indeed improve tumor-free survival (DFS), but the prognosis of patients is still very poor. For the early stages, our treatment principles refer to the guidelines of SGO, international gynecological cancer Cooperation (GCIG), and NCCN. We suggest a combination of surgery, radiotherapy, and chemotherapy. However, the general state of advanced patients is poor, whether they can tolerate the operation after neoadjuvant chemotherapy, whether the operation area can remain tumor-free, and whether this treatment will prolong the survival time of patients still need to be further discussed. In order to better prolong the tumor-free survival and prognosis of patients, we need to find gene changes suitable for targeted therapy, so as to complete the clinical application of these treatment methods. Further works are needed to explore more effective therapy for cervical SCNEC.

Keywords: neuroendocrine carcinoma, small cell, human papillomavirus, diagnosis, therapy

INTRODUCTION

Cervical cancer is the second most prominentcause of cancer-related death among women worldwide (Siegel et al., 2014). With the advent of the vaccine era and the widespread application of cervical screening, the incidence rate has declined, but tends to skew younger. A neuroendocrine tumor is a seldom-seen disease of the female reproductive system, and the cervix is the most common primary location (Salvo et al., 2019). In 2014, the World Health Organization divided cervical neuroendocrine tumors into two categories, which were low level (formerly known as carcinoid and atypical carcinoid) and high level (formerly known as small cell cancer or large cell neuroendocrine cancer) (Kurman et al., 2014). Cervical small cell neuroendocrine carcinomas are highly invasive and progress rapidly, their incidence is less than 5% of all cervical cancers. About 80% cervical SCNEC concerned high risky HPV (more familiar with HPV 18). Compared to squamous cell carcinoma or adenocarcinoma, patients are more prone to lymph node invasion early, which usually shows as local and distant metastases in clinic (Park et al). By searching PubMed, Google scholar, Cochrane Library, clinicalkey, and other databases, this paper retrospectively analyzed data from 1576 patients in 20 related studies and more than 50 papers in recent years of small cell carcinoma of the cervix in recent years and combined with 19 cases of this disease treated in our hospital from 2012 to 2021. We summarize the disease characteristics, pathogenesis, and relationship with HPV infection, which is in order to explore the timely diagnosis before canceration and improve the treatment, so as to improve the prognosis and survival rate of patients.

METHODS

The clinical data of small cell cancer treated in our hospital from 2012 to 2021 were analyzed retrospectively. The criteria for selecting cases were patients with a pathological diagnosis of small cell neuroendocrine carcinoma. The collected patient data included age, clinical manifestation, TCT, HPV detection, the size and morphology of the tumor, local invasion depth, stage, lymph node status, lymphatic vascular space involvement (LVSI), initial treatment method, tumor-free survival, follow-up interval, and mortality. All information was obtained through access to medical records. The research was approved by the ethics committee of the Second Affiliated Hospital of Dalian Medical University.

RESULTS

We found 19 cases of cervical small cell carcinoma in our hospital between 2012 to 2021. The total incidence rate was 1.5%. **Table 1** shows the general statistics of 19 patients. The median age was 48 years (with a range from 26 years to 70 years) which is five years younger than that of worldwide cervical carcinoma (Arbyn et al., 2020). The main clinical manifestation was irregular vaginal bleeding. In total 52.63% (10/19) of patients had postcoital bleeding before irregular vaginal spotting. Also, 36.84% (7/19) of patients had aqueous secretion or persistent vaginitis before abnormal bleeding, while 10.52% (2/19) of patients had no obvious symptoms before irregular vaginal bleeding. Only 47.37% (9/19) had TCT and HPV tests, the others were directly diagnosed by biopsy due to gynecological examination of suspected cervical cancer. While 55.56% (5/9) of the total had abnormal results with TCT. Among the patients with abnormal results, one patient was ASCUS(Atypical squamous cells of unknown significance), three patients were AGC(Atypical glandular cells), and one patient was ASC-H(Atypical squamous cells, not excluding high-grade squamous intraepithelial lesion). Simultaneously, we calculated that the HPV infection rate was 77.77% (7/9). In our study, we found that HPV infection was mainly high-risk 18 HPV positive. Five cases were found to be infected with HPV-18,with one being infected with HPV-18 and HPV-58 at the same time. In addition, one person was HPV-16 positive and one person was HPV-52 positive. In our cases, the median disease-free survival (DFS) of patients accepting multiple combination therapy was 15.5 months. **Table 2** shows the stage distribution, recurrence, treatment modalities, and prognosis of our cases. Nine patients with stage IB accepted radical hysterectomy and pelvic lymphadenectomy. Because patients before 2018 did not have stage IB3, we reassessed them according to their case data. All patients received adjuvant therapy after their operation, including pelvic external irradiation (EBRT) and chemotherapy combined with Etoposide and Cisplatin monthly. Three patients with stage IIB received pelvic three-dimensional transparent radiotherapy combined with synchronous DDP low-dose chemotherapy. Two patients with stage IIIB were given combined chemotherapy (EP+COA). Five stage IV patients mainly received combined chemotherapy with Etoposide and Cisplatin and some symptomatic treatment. **Table 3** describes the evaluation of tumor characteristics in nine surgically treated patients in detail. Two patients (2/9) had coarse tumors with a diameter greater than or equal to 4cm at diagnosis, and seven (7/9) patients had deep interstitial infiltration (≥ 1/2 of the matrix). All patients had intravascular tumor thrombus. Only one patient found pelvic lymph node metastasis, but there were seven patients in stage IB of recurrence.

Table 4 describes details of the immunohistochemistry related to small cell carcinoma in our case. We can see that SYN, CD56, CGA, p16, and Ki67 are relatively sensitive.

DISCUSSION

Cervical SCNEC and HPV

Cervical small cell neuroendocrine carcinoma (SCNEC) is a rare invasive cancer, accounting for no more than 3% of cervical cancer. Due to the high incidence of early lymph nodes and distant metastasis, it is usually found in the late stage (Chen et al., 2008; Cohen et al., 2010; Wharton et al., 2020). In our case, the proportion of cervical SCNEC is 1.5%, which is basically consistent with formerly reports. It is reported that the five year survival rate if diagnosed in the early stages is 30-46%, but only 0 – 15% at the advanced stage (Dongol et al., 2014). The early stage according to 2022 NCCN Guidelines Version 1 refers to stages IA1-IB2 and IIA1, while the advanced stage includes stages IB3, IIA2, and IIB-IV respectively (Abu-rustum et al., 2022).

TABLE 1 | Ordinary statistics of patients.

Variable	Figure	Literature data (52 articles,1576 patients)
Median Age, [range], years	48 (26-70)	the average age at diagnosis of worldwide cervical cancer was 53 years (40)
Irregular bleeding	100%	
postcoital bleeding before irregular vaginal bleeding	52.63% (10/19)	
abnormal results with TCT	55.56% (5/9)	
Median tumor-free survival(months)	15.5	Of the 179 eligible patients 104 were stage I, 19 stage IIA, 23 stage IIB, 9 stage III, 24 stage IV. The median tumor-free survival was16 months (Wang et al., 2012)
HPV infection rate	77.77% (7/9)	55% of cases were HPV16 positive; 41% were positive for HPV18 and 4% were positive for other types (Alejo et al., 2018)
HPV-18	71.42% (5/7)	
HPV-16	14.29% (1/7)	15% of cases were HPV16 positive; 70% were positive for HPV18 and 15% were positive for other types (Stoler et al., 1991)
HPV-52	14.29% (1/7)	
Tumor metastasis site		
pelvic lymph node	6	
Lung	3	
Distant metastasis such as brain or bone	8 (brain3, bone5)	

(Distant metastasis was found by imaging examination).

Moreover, a study has shown that despite the increase in comprehensive therapy, the prognosis of advanced patients remains worse (Lorusso et al., 2014). In our cases, small cell carcinoma patients have the symptoms of irregular vaginal bleeding. Before vaginal bleeding, 52.63% of patients experienced post-sexual bleeding. Among the patients screened for TCT and HPV, 55.56% have abnormal TCT and 77.77% have an HPV infection. This once again confirms the importance of cervical precancerous screening. Numerous clinical studies have shown that HPV infection is the main cause of cervical cancer (Small et al., 2017), and some studies have also confirmed that HPV infection is closely related to cervical SCNEC (Atienza-Amores et al., 2014; Dores et al., 2015). In the study by Alejo, 86% of neuroendocrine tumors were infected with HPV, mainly subtypes 16 and 18, and the positive rate of other subtypes was only 4%. Their data confirmed the correlation between cervical SCNEC, HPV, and $p16^{INK4a}$ overexpression (Alejo et al., 2018). Peifeng Li et al. retrospectively analyzed the clinical data of 30 cases in the primary period of cervical SCNEC. They concluded that infection with multiple HPV subtypes was not associated with disease severity or prognosis (Li et al., 2018). Xuan Pei et al. detected seven high-risk HPV types by quantitative multiplex PCR. They found that cervical SCNEC was particularly related to HPV 18 subtypes. The genetic changes of patients were caused by high risk HPV and mutations in frequent neuroendocrine cancers (Pei et al., 2021). Moreover, several studies have also demonstrated that HPV18 was specifically related to cervical SCNEC (Stoler et al., 1991; Ishida et al., 2004). In our cases, patients were mainly infected with HPV 18 (71.42%). This is consistent with the results of many studies. How to detect and block the occurrence of small cell carcinoma in time before its occurrence is the main difficulty we must overcome in the future.

Immunohistochemistry and Gene Detection

It is reported that commonly used neuroendocrine staining agents are chromogranin A (CGA), synaptophysin (SYN), CD56, and neuron-specific enolase (NSE). However, cervical SCNEC has not always responded to the neuroendocrine staining agents including CGA, SYN, and NSE (Stoler et al., 1991; Abeler et al., 1994; Conner et al., 2002). In our cases, the positive rates of the above indicators were 78.57% (CGA), 92.86% (SYN), 84.62% (CD56), and 50% (NSE). The sensitivity of p16 and Ki67 was high, and the patients were all positive. Inzani et al. found that 87.5% of patients had high expression of p16 protein, and more than half of the patients expressed SST5, SST2, and CDX2, while P63 and P40 were negative (Inzani et al., 2020). Changes in gene sequences are very important for the occurrence and progression of tumors. Deyin Xing et al. demonstrated the recurring gene changes referring to MAPK, PI3K/Akt/mTOR, and p53/BRCA pathways in cervical SCNEC

TABLE 2 | Treatment and prognosis of different stages.

Stage	Number	Number and interval of recurrence (average)	Treatment modalities
IB	IB1:3	IB1:2,36months	Radical hysterectomy and pelvic lymphadenectomy + EBRT and chemotherapy (Etoposide + Cisplatin) monthly.
	IB2:4	IB2:3,28months	
	IB3:2	IB3:2,18months	
IIB	3	3,8months	pelvic three-dimensional transparent radiotherapy combined with synchronous DDP low-dose chemotherapy.
IIIB	2	2,6 months	combined chemotherapy.
IV	5	5,3 months	combined chemotherapy with Etoposide and Cisplatin and some symptomatic treatment.
Total	19	17	

TABLE 3 | Tumor features (Number=9).

Classification	Number
Tumor size >4 cm	2
Exogenous	7
Endogenous	2
Depth of invasion	
More than 1/2 cervical thickness	7
Vascular tumor thrombus	9
Pelvic LN(+)	2

by using target next-generation gene sequencing technology (Xing et al., 2018). In the future, the existence of alterations in gene sequences may be used for targeted therapy.

Therapy of Cervical SCNEC

Due to the features of cervical SCNEC being low morbidity, highly malignant, powerful invasiveness, and poor prognosis, there is no accordant standard therapeutic regime now. In addition, given the insufficiency of prospective studies on therapeutic effects, the results of retrospective research are inconsistent, and the treatment methods are still controversial. In light of its special pathology, the therapeutic schedule can consult the methods of cervical cancers or other neuroendocrine cancers, such as small-cell carcinoma of the lung (SCLC), Merkel cell carcinoma (MCC), gastrointestinal neuroendocrine cancer, and so on (Becker et al., 2017; Sabari et al., 2017; Coggshall et al., 2018). Some studies have shown that the main treatment methods for cervical SCNEC include surgery, chemotherapy, and radiotherapy (Bermudez et al., 2001; Cohen et al., 2010; Gadducci et al., 2017). Some scholars demonstrated that the overall survival rate of patients with primary radical surgery was better than those who received direct radiotherapy. They preferred radical surgery combined with chemotherapy for the early stage. They found that patients receiving adjuvant chemotherapy had improved disease-free survival and reduced pelvic recurrence, and the overall survival rate tended to improve (Ishikawa et al., 2018). But Wang et al. indicated that the progression-free survival rate of radical surgery was worse (Wang et al., 2012). In our case, nine patients with stage IB accepted radical hysterectomy, and five of them received combined treatment after operation. Although seven of them recurred, we found a longer recurrence interval in patients who received early surgery combined with radiotherapy and chemotherapy. We summarized the case characteristics of non-recurrent patients and found that they all received satisfactory radical surgery and comprehensive treatment with radiotherapy and chemotherapy. Therefore, we still tend to recommend systemic treatment of early patients with surgery combined with concurrent radiotherapy and chemotherapy. For stages IA1-IB2 and IIA1, our treatment principles refer to the guidelines of SGO, international gynecological cancer Cooperation (GCIG), and NCCN, which all suggested radical surgery and adjuvant chemotherapy (Gardner et al., 2011; Stecklein et al., 2016). In our case, we often choose radiotherapy and chemotherapy to cure advanced-stage or recurrent patients. Caruso et al. retrospectively analyzed the patients who underwent a complete operation after neoadjuvant chemotherapy (NACT) in the later period. They found that compared with radiotherapy and chemotherapy, the surgical treatment may produce similar results. The possible benefit is to retain radiotherapy and chemotherapy as treatment options for late recurrence (Caruso et al., 2021). However, the general state of advanced patients is poor, whether they can tolerate the operation after neoadjuvant chemotherapy, whether the operation area can achieve tumor-free status, and whether the treatment will prolong the survival time of patients still need to be further discussed. At present, it is reported that angiogenesis inhibitors, cell signaling pathway inhibitors, immunotherapy, and apoptosis promoters can be used to detect lung SCNEC (Abidin et al., 2010). Therefore, we need to find gene changes suitable for targeted therapy, so as to complete the application of these treatments in cervical SCNEC. Salani et al. treated recurrent small cell neural cell carcinoma with trametinib. The patient was accompanied by a-kras mutation and achieved complete remission after 3 cycles of treatment (Salani et al., 2011). A recent study using next-generation sequencing technology to test cancer-related genes has shown that several targeted mutations (KRAS, PIK3CA, IRS2, Sox2, and HRR genes) may potentially become effective targeted treatment sites for patients (Pei et al., 2021). Eskander et al. analyzed the whole genome of 97 patients with cervical small cell carcinoma. They suggested that drugs targeting PIK3CA and other genes may become a potential treatment (Eskander et al., 2020). Furthermore, many previous studies have shown that PD-L1 is a notable manifestation in cervical carcinoma (Reddy et al., 2017; Chinn et al., 2019), and some scholars have also confirmed that treatment of immune check-point is working in patients with PD-L1 positive and MMRS deficiency (Ji et al., 2021).

CONCLUSIONS

The characteristics of cervical SCNEC are high malignancy, high invasiveness, and high mortality. It is interrelated with high risk HPV while HPV 18 carries a higher risk of cevical SCNEC. We found that multimodal treatment can indeed improve tumor-free survival (DFS), but the prognosis is poor. We commonly used neuroendocrine staining agents are CGA, SYN, CD 56, P16, Ki67, and neuron-specific NSE. For the early stages, our treatment principles refer to the guidelines of SGO,

TABLE 4 | Immunohistochemistry.

	SYN	CD56	NSE	CGA	CK8/18	CK5/6	p16	Ki67	EMA	Vim	LCA	CK7	P40
positive	13(14)	11(13)	2(4)	11(14)	3(3)	2(5)	5(5)	8(8)	1(3)	1(4)	0(5)	2(3)	2(3)
negative	1(14)	3(13)	2(4)	3(14)	0(3)	3(5)	0(5)	0(8)	2(3)	3(4)	5(5)	1(3)	1(3)

international gynecological cancer Cooperation (GCIG), and NCCN. We suggest a combination of surgery, radiotherapy, and chemotherapy. In order to better prolong tumor-free survival of patients, we need to find gene changes suitable for targeted therapy, so as to complete the application of these treatments in cervical SCNEC. Further studies are needed to explore more effective therapy for cervical SCNEC.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Written informed consent from the patients/participants was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

JW dealt with the case and drafted the manuscript. JL assisted in the collection and statistics of case data, and carried out all the documentary and article work out. All authors read and approved the final manuscript.

REFERENCES

Abeler, V. M., Holm, R., Nesland, J. M., and Kjorstad, K. E.. (1994). Small Cell Carcinoma of the Cervix. A Clinicopathologic Study of 26 Patients. *Cancer* 73, 672–677. doi: 10.1002/1097-0142(19940201)73:3<672::AID-CNCR2820730328>3.0.CO;2-R

Abidin, A. Z., Garassino, M. C., Califano, R., Harle, A., and Blackhall, F. (2010). Targeted therapies in Small Cell Lung Cancer: A Review. *Ther. Adv. Med. Oncol.* 2, 25–37. doi: 10.1177/1758834009356014

Alejo, M., Alemany, L., Clavero, O., Quiros, B., Vighi, S., Seoud, M., et al. (2018). Contribution of Human Papillomavirus in Neuroendocrine Tumors From a Series of 10, 575 Invasive Cervical Cancer Cases. *Papillomavirus Res.* 5, 134–142. doi: 10.1016/j.pvr.2018.03.005

Arbyn, M., Weiderpass, E., Bruni, L., Saraiya, M., Ferlay, J., Bray, F., et al. (2020). Estimates of Incidence and Mortality of Cervical Cancer in 2018: A World- Wide Analysis. *Lancet Global Health* 8 (2), 30482–6. doi: 10.1016/S2214-109X(19)30482-6

Atienza-Amores, M., Guerini-Rocco, E., Soslow, R. A., Park, K. J., and Weigelt, B. (2014). Small Cell Carcinoma of the Gynecologic Tract: A Multifaceted Spectrum of Lesions. *Gynecol Oncol.* 134, 410–418. doi: 10.1016/j.ygyno.2014.05.017

Abu-rustum, N. R., Yashar, C. M., Bradley, K., Brooks, R., Campos, S. M., Chino, J., et al (2022). *NCCN Clinical Practice Guidelines in Oncology, Cervical Cancer, Version 1.*

Becker, J. C., Stang, A., DeCaprio, J. A., Cerroni, L., Lebbe, C., Veness, M., Nghiem, P., et al. (2017). Merkel Cell Carcinoma. *Nat. Rev. Dis. primers.* 3, 17077. doi: 10.1038/nrdp.2017.77

Bermudez, A., Vighi, S., Garcia, A., and Sardi, J. (2001). Neuroendocrine Cervical Carcinoma: A Diagnostic and Therapeutic Challenge. *Gynecol Oncol.* 82, 32–39. doi: 10.1006/gyno.2001.6201

Caruso, G., Palaia, I., Di Donato, V., Pernazza, A., Gallo, R., Perniola, G., et al. (2021). Radical Surgery After Neoadjuvant Chemotherapy for Locally Advanced Neuroendocrine Cancer of the Cervix. *Anticancer Res.* 41(9), 4431–4438. doi: 10.21873/anticanres.15250

Chen, J., Macdonald, O. K., and Gaffney, D. K. (2008). Incidence, Mortality, and Prognostic Factors of Small Cell Carcinoma of the Cervix. *Obstet Gynecol.* 111 (6), 1394–1402. doi: 10.1097/AOG.0b013e318173570b

Chinn, Z., Stoler, M. H., and Mills, A. M. (2019). PD-L1 and IDO Expression in Cervical and Vulvar Invasive and Intraepithelial Squamous Neoplasias: Implications for Combination Immunotherapy. *Histopathol* 74 (2), 256–268. doi: 10.1111/his.13723

Coggshall, K., Tello, T. L., North, J. P., and Yu, S. S. (2018). Merkel Cell Carcinoma: An Update and Review: Pathogenesis, Diagnosis, and Staging. *J. Am. Acad. Dermatol.* 78 (3), 433–442. doi: 10.1016/j.jaad.2017.12.001

Cohen, J. G., Kapp, D. S., Shin, J. Y., Urban, R., Sherman, A. E., Chen, L. M., et al. (2010). Small Cell Carcinoma of the Cervix: Treatment and Survival Outcomes of 188 Patients. *Am. J. Obstetrics Gynecol.* 203 (4), 347. doi: 10.1016/j.ajog.2010.04.019

Conner, M. G., Richter, H., Moran, C. A., Hameed, A., and Albores-Saavedra, J. (2002). Small Cell Carcinoma of the Cervix: A Clinicopathologic and Immunohistochemical Study of 23 Cases. *Ann. Diagn. Pathol.* 6, 345–348. doi: 10.1053/adpa.2002.36661

Dongol, S., Tai, Y., Shao, Y., Jiang, J., and Kong, B. (2014). A Retrospective Clinicopathological Analysis of Small-Cell Carcinoma of the Uterine Cervix. *Mol. Clin. Oncol.* 2 (1), 71–75. doi: 10.3892/mco.2013.193

Dores, G. M., Qubaiah, O., Mody, A., Ghabach, B., and Devesa, S. S. (2015). A Population-Based Study of Incidence and Patient Survival of Small Cell Carcinoma in the United States, 1992-2010. *BMC Cancer.* 15, 185. doi: 10.1186/s12885-015-1188-y

Eskander, R. N., Elvin, J., Gay, L., Ross, J. S., Miller, V. A., Kurzrock, R., et al. (2020). Unique Genomic Landscape of High-Grade Neuroendocrine Cervical Carcinoma:Implications for Rethinking Current Treatment Paradigms. *JCO Precis. Oncol.* 4, 972–987. doi: 10.1200/PO.19.00248

Gadducci, A., Carinelli, S., and Aletti, G. (2017). Neuroendrocrine Tumors of the Uterine Cervix: A Therapeutic Challenge for Gynecologic Oncologists. *Gynecol Oncol.* 144, 637–646. doi: 10.1016/j.ygyno.2016.12.003

Gardner, G. J., Reidy-Lagunes, D., and Gehrig, P. A. (2011). Neuroendocrine Tumors of the Gynecologic Tract:a Society of Gynecologic Oncology (SGO) Clinical Document. *Gynecol Oncol.* 122(1), 190–198. doi: 10.1016/j.ygyno.2011.04.011

Inzani, F., Santoro, A., Angelico, G., Feraco, A., Spadola, S., Arciuolo, D., et al. (2020). Neuroendocrine Carcinoma of the Uterine Cervix:a Clinicopathologic and Immunohistochemical Study With Focus on Novel Markers (Sst2-Sst5). *Cancers* 12(5), 1211. doi: 10.3390/cancers12051211

Ishida, G. M., Kato, N., Hayasaka, T., Saito, M., Kobayashi, H., Katayama, Y., et al. (2004). Small Cell Neuroendocrine Carcinomas of the Uterine Cervix: A Histological, Immunohistochemical, and Molecular Genetic Study. *Int. J. Gynecolog Pathol.* 23 (4), 366–372. doi: 10.1097/01.pgp.0000139637.01977.61

Ishikawa, M., Kasamatsu, T., Tsuda, H., Fukunaga, M., Sakamoto, A., Kaku, T., et al. (2018). Prognostic Factors and Optimal Therapy for Stages I-II Neuroendocrine Carcinomas of the Uterine Cervix:a Multi-Center Retrospective Study. *Gynecol Oncol.* 148(1), 139–146. doi: 10.1016/j.ygyno.2017.10.027

Ji, X., Sui, L., Song, K., Lv, T., Zhao, H., and Yao, Q. (2021). PD-L1, PARP1, and MMRs as Potential Therapeutic Biomarkers for Neuroendocrine Cervical Cancer. *Cancer Med.* 10 (14), 4743–4751. doi: 10.1002/cam4.4034

Kurman, R. J., Carcangiu, M. L., and Herrington C. S, Y. R. (2014). *WHO Classification of Tumours of the Female Reproductive Organs* Bosman, FTJE Lakhani, SR Ohgaki, H eds. (Lyon: Lyon IARC Press), 307.

Li, P., Ma, J., Zhang, X., Guo, Y., Liu, Y., Li, X., et al. (2018). Cervical Small Cell Carcinoma Frequently Presented in Multiple High Risk HPV Infection and Often Associated With Other Type of Epithelial Tumors. *Diagn. Pathol.* 13 (1), 31. doi: 10.1186/s13000-018-0709-9

Lorusso, D., Petrelli, F., Coinu, A., Raspagliesi, F., and Barni, S. (2014). A Systematic Review Comparing Cisplatin and Carboplatin Plus Paclitaxel-Based Chemotherapy for Recurrent or Metastatic Cervical Cancer. *Gynecol Oncol.* 133 (1), 117–123. doi: 10.1016/j.ygyno.2014.01.042

Park, K. J., and Soslow, R. A.. Neoplastic Lesions of the Cervix. *Surgical Pathology Clinics* 4, 17–86. doi: 10.1016/B978-0-323-35909-2.00008-4

Pei, X., Xiang, L., Chen, W., Jiang, W., Yin, L., Shen, X., et al. (2021). The Next Generation Sequencing of Cancer-Related Genes in Small Cell Neuroendocrine Carcinoma of the Cervix[J]. *Gynecol Oncol.* 161, 779–786. doi: 10.1016/j.ygyno.2021.04.019

Reddy, O. L., Shintaku, P. I., and Moatamed, N. A. (2017). Programmed Death-Ligand 1 (PD-L1) is Expressed in a Significant Number of the Uterine Cervical Carcinomas. *Diagn. pathol* 12 (1), 45. doi: 10.1186/s13000-017-0631-6

Sabari, J. K., Lok, B. H., Laird, J. H., Poirier, J. T., and Rudin, C. M. (2017). Unravelling the Biology of SCLC: Implications for Therapy. *Nat. Rev. Clin. Oncol.* 14 (9), 549–561. doi: 10.1038/nrclinonc.2017.71

Salani, R., Backes, F. J., and Fung, M. F.. (2011). Posttreatment Surveillance and Diagnosis of Recurrence in Women With Gynecologic Malignancies:Society of Gynecologic Oncologists Recommendations. *Am. J. Obstet Gynecol* 204(6), 466–478. doi: 10.1016/j.ajog.2011.03.008

Salvo, G., Gonzalez Martin, A., NR, G., and Frumovitz, M. (2019). Updates and Management Algorithm for Neuroendocrine Tumors of the Uterine Cervix. *Int. J. Gynecol Cancer* 29 (6), 986–995. doi: 10.1136/ijgc-2019-000504

Siegel, R., Ma, J., Zou, Z., and Jemal, A. (2014). Cancer Statistics, 2014. *CA Cancer J. Clin.* 64, 9–29. doi: 10.3322/caac.21208

Small, W.Jr, Bacon, M. A., Bajaj, A., Chuang, L. T., Fisher, B. J., Harkenrider, M. M., et al. (2017). Cervical Cancer: A Global Health Crisis. *Cancer.* 123, 2404–2412. doi: 10.1002/cncr.30667

Stecklein, S. R., Jhingran, A., Burzawa, J., Ramalingam, P., Klopp, A. H., Eifel, P. J., et al. (2016). Patterns of Recurrence and Survival in Neuroendocrine Cervical Cancer. *Gynecol Oncol.* 143(3), 552–557. doi: 10.1016/j.ygyno.2016.09.011

Stoler, M. H., Mills, S. E., Gersell, D. J., and Walker, AN. (1991). Small-Cell Neuroendocrine Carcinoma of the Cervix. A Human Papillomavirus Type 18-Associated Cancer. *Am. J. Surg. Pathol.* 15, 28–32. doi: 10.1097/00000478-199101000-00003

Wang, K. L., Chang, T. C., Jung, S. M., Chen, C.-H., Cheng, Y.-M., Wu, H.-H., et al. (2012). Primary Treatment and Prognostic Factors of Small Cell Neuroendocrine Carcinoma of the Uterine Cervix:a Taiwanese Gynecologic Oncology Group Study. *Eur. J. Cancer* 48(10), 1484–1494. doi: 10.1016/j.ejca.2011.12.014

Wharton, D., Kim, E., Pagan, J., Small, W., Jaboin, J., and Ayala-Peacock, D. (2020). Patterns of Care and Outcomes for Small Cell Carcinoma of the Cervix: A National Retrospective Analysis of 542 Cases. *Adv. Radiat. Oncol.* 5 (3), 412–418. doi: 10.1016/j.adro.2019.08.008

Xing, D., Zheng, G., Kenneth Schoolmeester, J., Li, Z., Pallavajjala, A., Haley, L., et al. (2018). Next-Generation Sequencing Reveals Recurrent Somatic Mutations in Small Cell Neuroendocrine Carcinoma of the Uterine Cervix. *Am. J. Surg. Pathol.* 42 (6), 750–760. doi: 10.1097/PAS.0000000000001042

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 25 July 2022
DOI 10.3389/fcimb.2022.956801

OPEN ACCESS

EDITED BY
Ming Hu,
Qingdao University, China

REVIEWED BY
Jingkai Ji,
Shandong First Medical University, China
Lei Zhang,
Huazhong University of Science and Technology, China

*CORRESPONDENCE
Xiaofeng Yin
phwdwer@163.com;
15154258721@163.com

SPECIALTY SECTION
This article was submitted to Virus and Host, a section of the journal Frontiers in Cellular and Infection Microbiology

RECEIVED 30 May 2022
ACCEPTED 28 June 2022
PUBLISHED 25 July 2022

CITATION
Li J, Zhang Y, Qu Z, Ding R and Yin X (2022) ABCD3 is a prognostic biomarker for glioma and associated with immune infiltration: A study based on oncolysis of gliomas. *Front. Cell. Infect. Microbiol.* 12:956801. doi: 10.3389/fcimb.2022.956801

ABCD3 is a prognostic biomarker for glioma and associated with immune infiltration: A study based on oncolysis of gliomas

Jinchuan Li[1], Yi Zhang[1], Zhizhao Qu[2], Rui Ding[1] and Xiaofeng Yin[1]*

[1]Department of Neurosurgery, Second Hospital of Shanxi Medical University, Taiyuan, China, [2]Department of Neurosurgery, First Hospital of Shanxi Medical University, Taiyuan, China

Background: Gliomas are the most lethal primary brain tumors and are still a major therapeutic challenge. Oncolytic virus therapy is a novel and effective means for glioma. However, little is known about gene expression changes during this process and their biological functions on glioma clinical characteristics and immunity.

Methods: The RNA-seq data after oncolytic virus EV-A71 infection on glioma cells were analyzed to screen significantly downregulated genes. Once ABCD3 was selected, The Cancer Genome Atlas (TCGA), Chinese Glioma Genome Atlas (CGGA), Genotype-Tissue Expression (GTEx), Clinical Proteomic Tumor Analysis Consortium (CPTAC), and Human Protein Atlas (HPA) data were used to analyze the relationship between ABCD3 expression and clinical characteristics in glioma. We also evaluated the influence of ABCD3 on the survival of glioma patients. CIBERSORT and Tumor Immune Estimation Resource (TIMER) were also used to investigate the correlation between ABCD3 and cancer immune infiltrates. Gene set enrichment analysis (GSEA) was performed to functionally annotate the potential functions or signaling pathways related to ABCD3 expression.

Results: ABCD3 was among the top 5 downregulated genes in glioma cells after oncolytic virus EV-A71 infection and was significantly enriched in several GO categories. Both the mRNA and protein expression levels of ABCD3 were upregulated in glioma samples and associated with the prognosis and grades of glioma patients. The Kaplan–Meier (K-M) curve analysis revealed that patients with high ABCD3 expression had shorter disease-specific survival (DSS) and overall survival (OS) than those with low ABCD3 expression. Moreover, ABCD3 expression could affect the immune infiltration levels and diverse immune marker sets in glioma. A positive correlation was found between ABCD3 and macrophages and active dendritic cells in the microenvironment of both the GBM and LGG. Gene sets including the plk1 pathway, tyrobp causal network,

ir-damage and cellular response, and interleukin-10 signaling showed significant differential enrichment in the high ABCD3 expression phenotype.

Conclusion: Our results suggested that ABCD3 could be a potential biomarker for glioma prognosis and immunotherapy response and also further enriched the theoretical and molecular mechanisms of oncolytic virus treatment for malignant gliomas.

KEYWORDS

glioma, ABCD3, biomarker, oncolytic virus, immune infiltrates

Introduction

Gliomas are the most common form of primary brain tumors. They are generally divided into glioblastoma multiforme (GBM) and low-grade glioma (LGG) (Alifieris and Trafalis, 2015; Chen et al., 2017; Hu et al., 2021b). GBM is classified by the WHO as a grade IV tumor that is highly invasive and rapidly growing and has massive metastases (Hu et al., 2021a). Despite the currently available strategies of surgical resection, chemotherapy, and irradiation, the median survival for patients is only 12–15 months (Batash et al., 2017; Witthayanuwat et al., 2018). LGG (grades II and III) is a more inert precursor to glioblastoma and the prognosis is relatively encouraging, despite the possibility of evolving into a more aggressive GBM (Wang and Mehta, 2019). Usually, the disease remains unchanged for a long period of time. However, patients have few effective treatment options, and the prognosis for these patients today is poor (Alexander and Cloughesy, 2017). So, the search for new biomarkers in glioma patients is particularly urgent in order to provide a highly accurate prediction of patient survival and/or response to individualized treatment.

Several studies have confirmed the safe and effective use of oncolytic viruses in the treatment of glioma (Foreman et al., 2017; Suryawanshi and Schulze, 2021). Oncolytic virus therapy is an immunotherapy treatment that is characterized by virus-specific infection of glioma cells and apoptosis induction through the release of viral progeny (Zeng et al., 2021). The oncolytic virus treatment on glioma often results in the downregulation of genes that promote glioma malignancy, as well as changes in molecules associated with the immune response (Rius-Rocabert et al., 2020). A recent study shows that enterovirus A71 (EV-A71) has potent oncolytic activity in the treatment of glioma (Zhang et al., 2020a). Malignant glioma-derived cell lines are capable of being infected and killed by EV-A71 (Zhang et al., 2020a). An EV-A71 treatment can significantly slow down the progression of tumors in nude mice bearing glioma xenografts (Zhang et al., 2020a). Oncolytic viruses are capable of exhibiting an antitumor effect that is connected to both their intrinsic oncolytic properties and the way the immune system reacts (Hemminki et al., 2020).

In the treatment of glioma with EV-A71, we screened a significantly downregulated ATP-binding cassette (ABC) transporter subfamily D member 3 (ABCD3) by high-throughput sequencing data. Based on structural organization and amino acid homology, ABC transporters are classified into seven subfamilies (A to G) in humans (Dean et al., 2001). ABCD3, also known as the 70-kDa peroxisomal membrane protein (PMP70), is localized to the monolayer of the peroxisome, and the expression of the membrane protein increases as the peroxisome proliferates (Braiterman et al., 1998). The encoded protein is involved in the transport of fatty acids or fatty acyl coenzyme A in the peroxisome (Imanaka et al., 2000). Many studies have confirmed that ABCD3 plays an important role in the development and progression of many tumors and involves several regulatory mechanisms in tumor development (Reams et al., 2015; Zhang et al., 2020b). However, ABCD3 has hardly been evaluated in glioma.

In this study, our data were collected from the European Nucleotide Archive (ENA), The Cancer Genome Atlas (TCGA), Chinese Glioma Genome Atlas (CGGA), Genotype-Tissue Expression (GTEx), and Human Protein Atlas (HPA). Moreover, the Gene Expression Profiling Interactive Analysis (GEPIA) and R package identified the differences and correlations of ABCD3 with glioma. We evaluated the correlations between clinical characteristics and ABCD3 in glioma. Subsequently, the density of tumor-infiltrating immune cells (TIICs) in tumor microenvironments was studied using a recent metagene approach called CIBERSORT, as well as the Tumor Immune Estimate Resource (TIMER). We also examined the association between ABCD3 and infiltrating immune cells in tumors using these methods. According to our results, ABCD3 proves to be an effective diagnostic and prognostic biomarker for gliomas and to have a possible correlation with tumor immune interactions.

Materials and methods

Data acquisition and bioinformatics

The RNA-seq data after EV-A71 infection on glioma cells were obtained from the ENA with accession number PRJNA562271. In that study, glioma cells were infected with EV71 BrCr strain or mock infection at a multiplicity of infection (MOI) of 1 for 60 min at room temperature. The total RNA samples from CCF glioma cells were used for RNA-seq analysis by using the Illumina HiSeq™ 2000 System (Illumina, San Diego, USA). After getting the raw data, we utilized FastQC for quality testing. The differentially expressed genes (DEGs) were calculated by Gfold version 1.1.4, (Shanghai, China) and a GFOLD value ≥1 or ≤−1 was considered.

GBM and LGG patient datasets were downloaded from the TCGA, CGGA, and GTEx with ABCD3 gene expression and the corresponding clinical information data. The protein expression data were evaluated using the Human Protein Atlas and Clinical Proteomic Tumor Analysis Consortium (CPTAC) databases.

Survival analysis

For survival analysis, we compared the overall survival (OS) and disease-specific survival (DSS). Both GBM and LGG patients were divided separately into high and low ABCD3 groups based on the mean value of TPM. Survival analysis and visual representations were executed using R version 3.4.3, including survival (for survival analysis and Cox) and survminer (for visual representations) R packages. Kaplan–Meier (K-M) curves were carried out to compare the survival time differences. *P*-values from log-rank tests were calculated, and less than 0.05 was considered statistically significant.

ABCD3 expression and tumor-infiltrating immune cell analysis

To analyze the ABCD3 RNA sequencing expression data from tumors and normal samples in the GTEx and TCGA projects, a webserver called GEPIA was used as a standard processing pipeline. Plotting of ABCD3 expression was performed using ggplot2.

To assess the influence of ABCD3, we utilized two methods to estimate the infiltration cells. Firstly, we employed the single-sample gene set enrichment analysis (ssGSEA) using the R gene set variation analysis (GSVA) package (v1.34.0) to calculate the infiltration levels of 24 immune cells. The marker genes for these 24 immune cells come from this article (Bindea et al., 2013). To compare two groups of values, the Wilcoxon signed-rank tests or the two-sided non-parametric Wilcoxon rank-sum tests were used. The second method is TIMER (cistrome.shinyapps.io/timer), which is a web-based evaluation of the amount of tumor-infiltrating immune cells (Li et al., 2017). All GBM and LGG data were downloaded from the TCGA. The scatterplots displayed the purity-corrected partial Spearman's rho value and statistical significance.

Gene set enrichment analysis

To explore the potential signaling pathways and biological functions related to ABCD3 expression, we conducted gene set enrichment analysis (GSEA) based on MSigDB Collections using the clusterProfiler R package (v3.14.3). The GBM and LGG samples were sorted into high and low expression groups, respectively, according to ABCD3 expression. The different transcriptional profiles were identified and sorted by log2FC with DESeq2. Then, we ran the GSEA on these transcriptional profiles and assessed the enrichment using the normalized enrichment score (NES) as a metric of enrichment. False discovery rate (FDR) <0.25 and p.adjust value <0.05 were considered to be of statistical significance.

Statistical analysis

Statistical analyses from the TCGA were merged and performed using R version 3.4.3. The correlations between ABCD3 expression and clinical information were examined using logistic and Cox regressions. GraphPad Prism was used to perform statistical analysis. A two-tailed $P < 0.05$ was considered statistically significant.

Results

ABCD3 is significantly decreased during oncolytic virus-mediated apoptosis in glioma cells

It has been reported that oncolytic viruses EV-A71 have shown oncolytic efficacy in glioma (Zhang et al., 2020a). Based on this finding, we conducted the transcriptome analysis to further analyze the impact of EV-A71 infection on glioma cells. The RNA-seq data (PRJNA562271) were obtained from CCF glioma cells after EV71 or mock infection at an MOI of 1. After getting the raw data, we utilized FastQC for quality testing. The results confirmed a high level of quality and sufficient quantity for further gene functional analysis (Supplementary Figure 1). Using the TPM method (the number of transcripts per million clean tags), the gene expression level was calculated and normalized by Gfold. The distribution of DEGs from EV-A71 groups was obtained relative to the mock infection.

A total number of 441 genes was screened as upregulation (GFOLD value ≥ 1) and 320 genes as downregulation (GFOLD value ≤ −1) (Tables S1, S2). Since oncolytic viruses have a natural propensity to infect and kill tumor cells, the downregulated

genes' function after infection may contribute to enhance the malignant phenotype of gliomas. We performed Gene Ontology (GO) analysis for all downregulated genes and found many enriched GO categories (Figure 1). Furthermore, we found that ABCD3 was involved in several top and significantly enriched GO categories: cytosol, membrane, mitochondrion, protein binding, and ATP binding. Also, ABCD3 was among the top 5 downregulated genes in glioma cells after oncolytic virus EV-A71 infection (Table S1). However, the expression and correlation of ABCD3 with glioma have not been reported before.

The mRNA and protein expression levels of ABCD3 are upregulated in glioma samples

To examine the expression of ABCD3 in glioma progression, the data from publicly available datasets—TCGA, CGGA, GTEx, CPTAC, and HPA—were used to investigate the ABCD3 mRNA or protein expression patterns in glioma samples or normal tissue samples. Firstly, we compared the mRNA expression of ABCD3 in clinical brain tumors with different WHO-classified degrees in the TCGA database. In the TCGA database, brain tumors are divided into two cohorts: GBM (grade IV) and LGG (grades II and III). In general, ABCD3 was upregulated in both GBM and LGG compared with matched control samples (TCGA normal and GTEx data) ($P < 0.01$, Figure 2A). Analysis of the data from the CGGA showed similar results (Supplementary Figure 2A). In comparison among gliomas with WHO grades, ABCD3 was higher in grade IV than in grades II and III, and there was no statistical difference between grades II and III (Figure 2B).

We further checked ABCD3 protein expression levels. The protein expression level was also upregulated in glioblastoma multiforme in comparison with normal tissues based on the data from the CPTAC (Figure 2C), indicating that ABCD3 protein and mRNA expression levels were similar in different databases. Immunohistochemistry showed positive reactions of ABCD3 protein in cytoplasmic and membranous staining of both high- and low-grade gliomas. Meanwhile, ABCD3 displayed a strong intensity and over 75% quantity in high-grade glioma compared with a moderate intensity and around 50% quantity in low-grade glioma. In normal cerebral cortex tissue, ABCD3 has not been detected in the neuropil, endothelial cells, and neuronal cells and only showed low staining and less than 25% quantity in glial cells (Figure 2D).

ABCD3 expression is associated with clinical characteristics of glioma patients

Glioma harboring the co-deletion status of 1p and 19q chromosome arms (1p/19q) shows favorable prognostic and

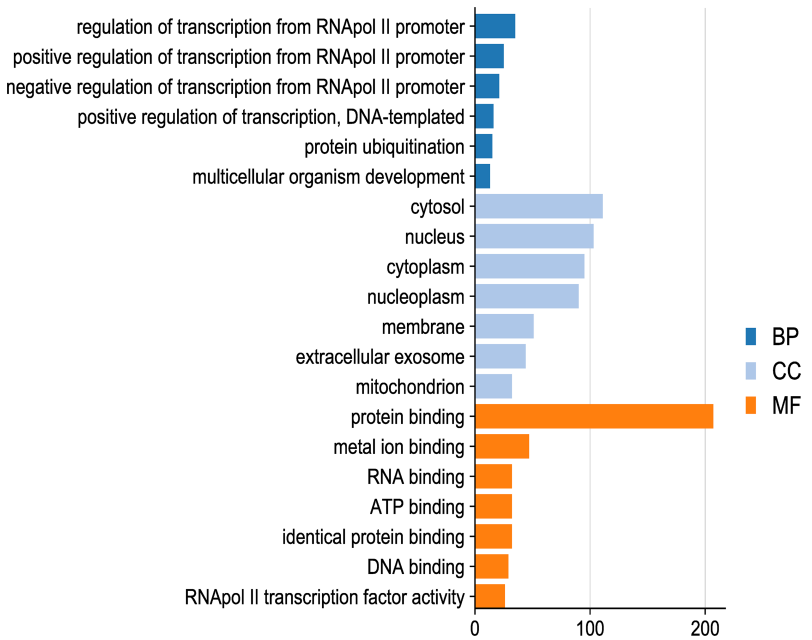


FIGURE 1
Gene Ontology (GO) analysis for all downregulated genes after EV-A71 infection on glioma cells. The distribution of downregulated genes from EV-A71 groups was obtained relative to the mock infection. GO analysis for all downregulated genes was performed and many enriched GO categories were found.

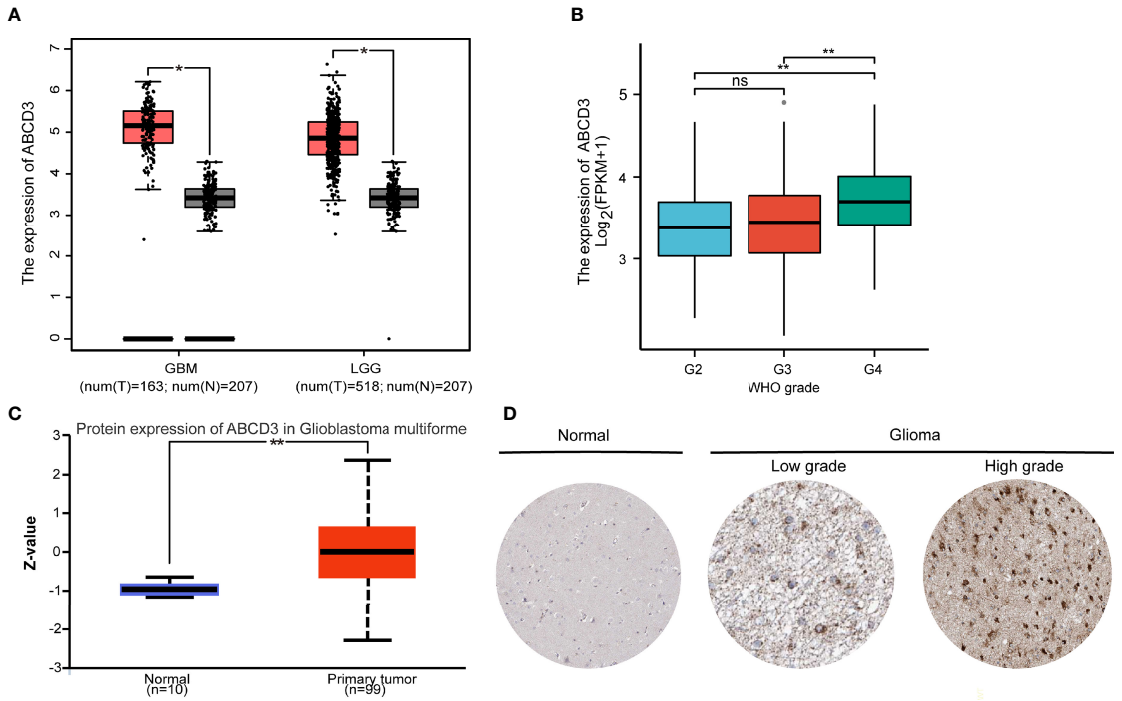


FIGURE 2
The mRNA and protein expression levels of ABCD3 are upregulated in glioma samples. **(A)** Differential expression of ABCD3 in both GBM and LGG compared with matched control samples. **(B)** Differential expression of ABCD3 in different glioma grades. **(C)** Differential ABCD3 protein expression levels in glioblastoma compared with normal tissues based on the data from the CPTAC. **(D)** Immunohistochemistry images of ABCD3 protein expression in glioma tissues and their normal controls. ns, no significance. *, $P < 0.05$; **, $P < 0.01$.

predictive values. In the subgroup analysis, we noticed that ABCD3 expression is significantly lower in the co-deletion (Codel) group than in the no co-deletion groups in the WHO II and III grades (Figure 3A). However, ABCD3 did not display a statistically different expression in other glioma subgroups in terms of gender and age (Supplementary Figure 3).

For survival analysis, both GBM and LGG patients were divided separately into high and low ABCD3 groups based on the mean value of TPM. The cutoff value of the GBM group is 5.10 and that of the LGG group is 5.01. The survival curve showed that high ABCD3 expression significantly decreased the OS and DSS. For GBM, the median survival time of the low ABCD3 expression group was 76.1 versus 40.3 months of the high ABCD3 expression group. Meanwhile, for LGG, the median survival time of the low ABCD3 expression group was 135.6 versus 68.4 months of the high ABCD3 expression group. Patients with high ABCD3 expression had a worse prognosis.

The hazard ratio (HR) for the high ABCD3 expression on OS of GBM patients was 1.60 (95% CI, 1.25 to 2.04; $P < 0.001$) and that of LGG patients was 1.46 (95% CI, 1.04 to 2.06; $P = 0.03$). Additionally, the HR for high ABCD3 expression on DSS of GBM patients was 1.71 (95% CI, 1.32 to 2.22; $P < 0.001$), and that of LGG patients was 1.65 (95% CI, 1.14 to 2.37; $P = 0.008$) (Figures 3B–E). Analysis of the data from the CGGA showed similar results (Supplementary Figure 2B). Moreover, Cox regression analysis also revealed that age and IDH status as well as ABCD3 and 1p/19q are significantly associated with overall survival (Figure 3F).

Relationship between ABCD3 expression and tumor-infiltrating immune cells

As tumor-infiltrating immune cells (IC) are associated with glioma prognosis (Sokratous et al., 2017), we conducted an analysis to find out if ABCD3 expression was associated with immune infiltration in GBM and LGG. All GBM and LGG samples were separately divided into high and low groups based on the mean value of ABCD3. We used an established computational resource ssGSEA to estimate the differing concentrations of 24 immune cell types in the high and low ABCD3 expression groups with CIBERSORT. The proportion of 24 subpopulations of immune cells appears clearly in Figure 4. The Wilcoxon signed-rank test was used to show the significant difference between the two groups. For GBM, macrophages, neutrophils, and Th2 cells were apparently increased in the high ABCD3 expression group. In contrast, DC, pDC, Th17 cells, and Treg cells were decreased in the high ABCD3 expression group (Figure 4A). The immune infiltration status of

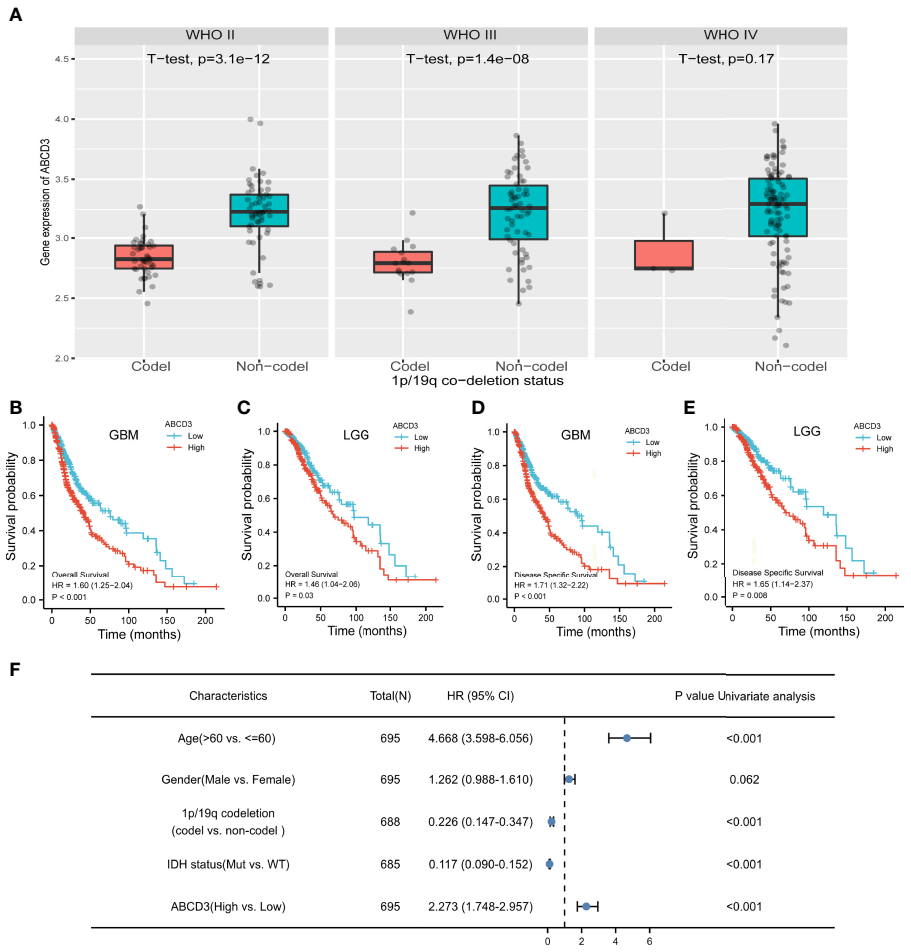


FIGURE 3
ABCD3 expression is associated with the clinical characteristics of glioma patients. **(A)** ABCD3 expression is lower in the 1p/19q co-deletion (Codel) groups than in the no co-deletion groups in all three WHO grade gliomas. **(B–E)** Overall survival and disease-specific survival curve of GBM and LGG patients based on differential ABCD3 expression. **(F)** Univariate Cox analysis of ABCD3 expression and other clinical pathological factors for GBM and LGG.

LGG displayed a similar pattern along with ABCD3 expression compared to GBM. However, the Th17 proportion in LGG was not associated with ABCD3 expression (Figure 4B). For macrophages that were significantly correlated with ABCD3, we provided a way to further display them in a scatter plot. The macrophage infiltrate correlated with both GBM and LGG, and correlations and *P*-values are included in Figures 4C, D.

A correlation exists between ABCD3 expression and immune infiltration levels in glioma obtained from TIMER

To ensure the accuracy of infiltration levels associated with ABCD3, we also assessed how ABCD3 expression correlates with immune infiltration levels in different grades of glioma using TIMER. The results showed a positive correlation between the high expression of ABCD3 and the high levels of all immune infiltration in LGG (Figure 5B). Compared with LGG, there was a weak significant correlation between immune infiltration and ABCD3 expression from TIMER. Moreover, a negative correlation exists between ABCD3 expression level and infiltrating levels of dendritic cells (r = −0.121, P = 1.33e−02) in GBM (Figure 5A).

Gene sets enriched in the ABCD3 expression phenotype

In order to functionally annotate the potential functions or signaling pathways related to ABCD3 expression, pathway enrichment was assessed by GSEA. GBM and LGG samples

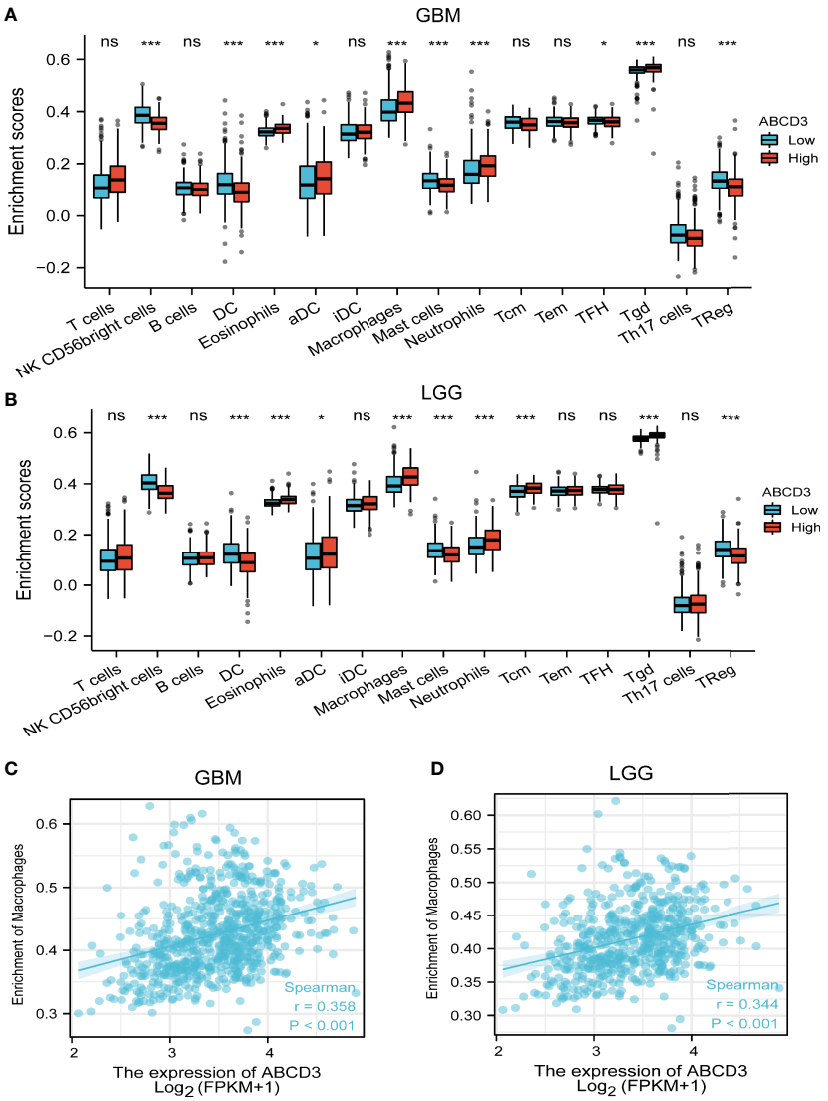


FIGURE 4
Relationship between ABCD3 expression and tumor-infiltrating immune cells. The proportion of 24 subpopulations of immune cell infiltration between the high and low expression of ABCD3 in GBM **(A)** or LGG **(B)** samples. Macrophages were mostly increased in the high ABCD3 expression group compared with the low ABCD3 expression group, while pDC cells were apparently decreased in the high expression group in both GBM and LGG. The macrophage infiltrate correlated with both GBM **(C)** and LGG **(D)**. ns, no significance; *, $P < 0.05$; ***, $P < 0.001$.

were sorted into high and low expression groups, respectively, according to ABCD3 expression. The different transcriptional profiles were identified and sorted by log2FC with DESeq2. Then, we ran GSEA on these transcriptional profiles and assessed the enrichment using the NES as a metric of enrichment. The top 5 ABCD3-related items are listed and visualized in each group (Table 1).

For the LGG samples with high ABCD3 expression, GSEA confirmed the strong enrichment for inflammatory genes among the differentially expressed pathways including interferon-gamma response, epithelial–mesenchymal transition (EMT), TNF-α signaling, and interferon-α response (Table 1 and Figure 6A). The inflammatory response and EMT are the hallmarks of many cancer types. K-ras signaling, oxidative phosphorylation, and cholesterol homeostasis gene sets were significantly enriched in the low ABCD3 expression phenotype for LGG (Figure 6B). These indicated the potential role of ABCD3 in the development of LGG.

Gene sets including the plk1 pathway, tyrobp causal network, ir-damage and cellular response, and interleukin-10 signaling showed significant differential enrichment in the high ABCD3 expression GBM phenotype (Figure 6C). Meanwhile,

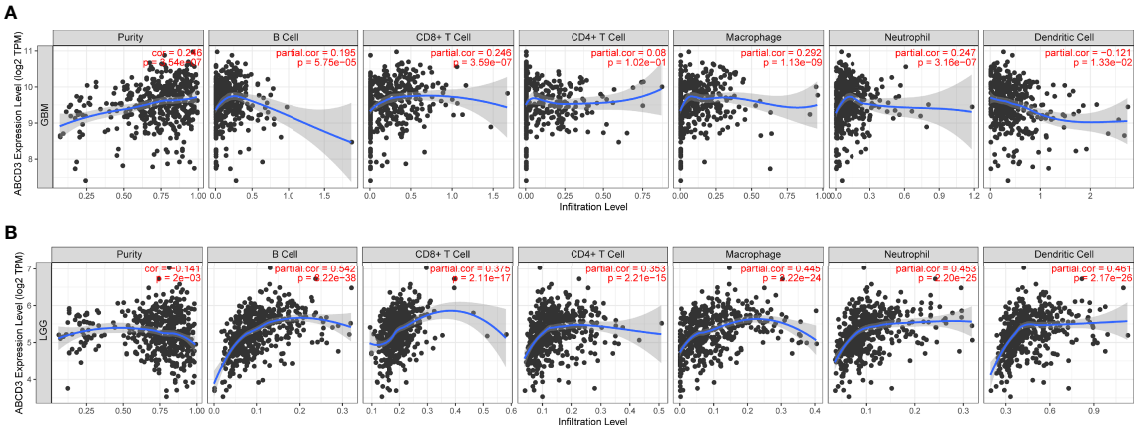


FIGURE 5
The ABCD3 expression level has a correlation with the tumor-infiltrating levels from TIMER. **(A)** ABCD3 expression has positive correlations with infiltrating levels of B cell, $CD8^+$ T cell, macrophage, and neutrophil, while it has a negative correlation with dendritic cells in GBM. **(B)** ABCD3 expression has a significant positive correlation with all immune infiltration cells in LGG.

the neuronal system, transmission across chemical synapses, synaptic vesicle pathway, and potassium channels showed significant differential enrichment in the low ABCD3 expression GBM group based on NES (Table 1 and Figure 6D).

Discussion

ABCD3 is involved in the transport of fatty acids in the peroxisome and may play an important role in tumorigenesis

TABLE 1 Gene sets enriched in different ABCD3 expression phenotypes.

Gene set name	NES	*P*-values	*q*-values
LGG high ABCD3 expression			
HALLMARK_G2M_CHECKPOINT	2.67	0.0034	0.0066
HALLMARK_INTERFERON_GAMMA_RESPONSE	2.16	0.0034	0.0066
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	2.11	0.0034	0.0066
HALLMARK_TNFA_SIGNALING_VIA_NFKB	2.01	0.0034	0.0066
HALLMARK_INTERFERON_ALPHA_RESPONSE	2.00	0.0027	0.0066
LGG low ABCD3 expression			
HALLMARK_KRAS_SIGNALING_DN	−1.71	0.0014	0.0066
HALLMARK_MYOGENESIS	−1.50	0.0084	0.0119
HALLMARK_OXIDATIVE_PHOSPHORYLATION	−1.40	0.0167	0.0207
HALLMARK_SPERMATOGENESIS	−1.24	0.0897	0.0983
HALLMARK_CHOLESTEROL_HOMEOSTASIS	−1.24	0.0250	0.1268
GBM high ABCD3 expression			
WP_TYROBP_CAUSAL_NETWORK	2.73	0.0027	0.0197
PID_AURORA_B_PATHWAY	2.44	0.0025	0.0197
PID_PLK1_PATHWAY	2.43	0.0026	0.0197
WP_DNA_IRDAMAGE_AND_CELLULAR_RESPONSE_VIA_ATR	2.39	0.0027	0.0197
REACTOME_INTERLEUKIN_10_SIGNALING	2.38	0.0026	0.0197
GBM low ABCD3 expression			
REACTOME_NEURONAL_SYSTEM	−2.62	0.0013	0.0197
REACTOME_TRANSMISSION_ACROSS_CHEMICAL_SYNAPSES	−2.49	0.0013	0.0197
WP_SYNAPTIC_VESICLE_PATHWAY	−2.45	0.0015	0.0197
REACTOME_POTASSIUM_CHANNELS	−2.41	0.0015	0.0197
REACTOME_NEUROTRANSMITTER_RELEASE_CYCLE	−2.38	0.0015	0.0197

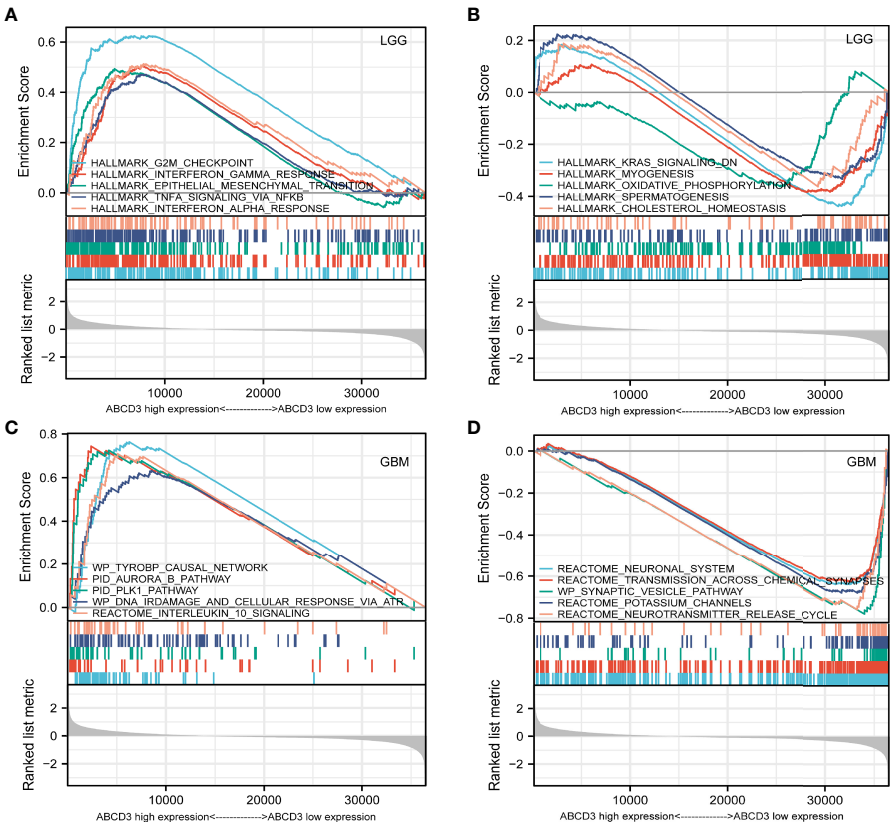


FIGURE 6
GSEA analysis of LGG or GBM samples with high or low ABCD3 expression phenotype. **(A)** GSEA results showing the top 5 enriched gene sets in the LGG samples with high ABCD3 expression based on NES. **(B)** GSEA results showing the top 5 enriched gene sets in the LGG samples with low ABCD3 expression based on NES. **(C)** GSEA results showing the top 5 enriched gene sets in the GBM samples with high ABCD3 expression based on NES. **(D)** GSEA results showing the top 5 enriched gene sets in the GBM samples with low ABCD3 expression based on NES.

(Reams et al., 2015). ABCD3 has been shown to be a biogenetic marker for malignancies such as human prostate cancer (Zhang et al., 2020b), ovarian cancer (Lima et al., 2015; Elsnerova et al., 2016), and non-small cell lung cancer (Tran, 2013) but has hardly been evaluated in glioma. Here, we screened ABCD3 as one of the most dramatically downregulated genes during oncolytic virus EV-A71 treatment. It is believed that oncolytic viruses mediate antitumor effects mainly through the following mechanisms: 1) specifically replicate in tumor cells with direct lysis and 2) release viral particles from lysed tumor cells to stimulate systemic immunity, such as promoting tumor antigen presentation, increasing the tumor microenvironment immune cell infiltration, modulating the tumor microenvironment, activating the immune cells, and activating the body's immune system through the immunomodulatory factors carried by them. In addition, it has been shown that some viruses can indirectly inhibit the immune system by infecting tumor-related vascular endothelial cells and also inhibit tumor angiogenesis. Therefore, during the lysis of gliomas by oncolytic viruses, significantly downregulated genes during the lysis of gliomas induced by oncolytic viruses are often necessary for the malignant proliferation of gliomas.

In the GO analysis for all downregulated genes, we also found that ABCD3 was involved in several top and significantly enriched GO categories. Oncolytic viruses cause glioma cell death through different mechanisms, including pyroptosis, necroptosis, and apoptosis. Genes that were downregulated in this process are often associated with glioma proliferation and malignancy. Oncolytic viruses also trigger an inflammatory immune response in the tumor microenvironment (TME). Gliomas that metastasize to the central nervous system (CNS) are characterized by the highly immunosuppressive TME. Therefore, ABCD3 is likely to be a diagnostic and prognostic biomarker associated with the clinical features and immune infiltration of gliomas.

Consequently, we conducted the first-ever analysis of ABCD3 expression in large cohorts of human glioma patients, which unveiled the potential role of ABCD3 in glioma. Based on RNA-seq analysis, protein analysis, and clinical data, we performed a retrospective analysis on glioma patients with

histological confirmation. We found that ABCD3 expression is associated with tumor grade in glioma patients as ABCD3 is higher in grade IV than in grades II and III. In gliomas, the co-deletion status of 1p and 19q chromosome arms is associated with the oligodendrocyte component of the tumor, and co-deletion abnormalities of both are detectable in 70% of low-grade gliomas and 60% of mesenchymal oligodendrogliomas. The 1p/19q co-deletion patients respond well to radiotherapy and chemotherapy, suggesting a longer survival and a relatively good prognosis. In the subgroup analysis, we noticed that ABCD3 expression was significantly lower in the co-deletion group than in the no co-deletion groups in all three WHO grades. Similarly, we found that ABCD3 expression levels in gliomas were related to prognosis. In gliomas, high ABCD3 expression was also associated with a poor prognosis. ABCD3 expression was found to be an independent prognostic factor of glioma patient prognosis in multivariate analyses. Cox regression analysis also revealed that age and IDH status as well as ABCD3 were significantly associated with overall survival.

Multiple diseases of the CNS, including malignant diseases, are associated with dysregulation of inflammatory responses (Pedemonte et al., 2006). Since gliomas are largely protected from immune cells infiltrating them, they have long been considered immune-inaccessible to antitumor-immune (Roesch et al., 2018). Given that oncolytic virus treatment could trigger an inflammatory immune response in the glioma microenvironment (Walton et al., 2018), a systematic analysis of the correlation between ABCD3 and tumor immunity was performed. Our studies showed that ABCD3 expression could affect the immune infiltration levels and diverse immune marker sets in glioma. A positive correlation was found between ABCD3 and macrophages and active dendritic cells in the microenvironment of both GBM and LGG. They provided a supportive stroma for the proliferation and invasion of neoplastic cells. These results may explain why high ABCD3 gliomas progressed malignantly and had adverse outcomes. ABCD3 was also closely related to glioma purity and immune score in LGG.

A lower purity indicates a poor prognosis and enhances the malignancy progression phenotype. High ABCD3 gliomas are more likely to occur in tissues with more complex microenvironments, based on this result.

GSEA also confirmed the strong enrichment for inflammatory genes among the differentially expressed pathways including the EMT, interferon-gamma response, TNF-α signaling, and interferon-α response. The inflammatory response and EMT are the hallmarks of many cancer types (Suarez-Carmona et al., 2017; Cho et al., 2019). K-ras signaling, oxidative phosphorylation, and cholesterol homeostasis gene sets were significantly enriched in the low ABCD3 expression phenotype for LGG. These indicated the potential role of ABCD3 in the development of LGG.

In conclusion, we screened an unreported ABCD3 gene by high-throughput bioinformatics analysis of genes significantly altered during treatment of gliomas with the oncolytic viruses EV-A71. Using multiple databases, we comprehensively analyzed the correlation between ABCD3 mRNA and protein expression levels and clinical glioma patient characteristics, immune infiltration, gene enrichment analysis, etc. We conclude that ABCD3 is highly expressed in different degrees of gliomas and is associated with poor prognosis and many clinical characteristics. We also found that ABCD3 expression could affect the immune infiltration levels and diverse immune marker sets in glioma and, especially, could enhance the malignancy progression phenotype of LGG. Our results suggested that ABCD3 could be a potential biomarker for glioma prognosis and immunotherapy response. Our results also further enriched the theoretical and molecular mechanisms of oncolytic virus treatment for malignant gliomas.

Data availability statement

Publicly available datasets were analyzed in this study. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material

Author contributions

Conceptualization: JL and XY. Resources: XY. Data curation: JL, YZ, and XY. Software: ZQ. Formal analysis: JL, YZ, RD, and XY. Supervision: RD. Funding acquisition: XY. Validation: XY. Investigation: JL and YZ. Visualization: JL and XY. Methodology: RD and XY. Writing—original draft: JL. Project administration: XY. Writing—review and editing: XY. All authors read and approved the final manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the Applied Basic Research Project of Shanxi Province (20210302123266).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2022.956801/full#supplementary-material

SUPPLEMENTARY FIGURE 1
A high level of quality and sufficient quantity of data was tested by FastQC. The FastQC results confirmed a high level of quality and sufficient quantity for further gene functional analysis.

SUPPLEMENTARY FIGURE 2
ABCD3 expression is associated with glioma grades and survival rate from CGGA. **(A)** ABCD3 expression is associated with glioma grades. **(B)** Overall survival curve of all grade primary glioma based on differential ABCD3 expression.

SUPPLEMENTARY FIGURE 3
ABCD3 different expression with glioma subgroups. ABCD3 did not display a statistical different expression in patient Gender **(A)** and Age **(B)**.

References

Alexander, B. M., and Cloughesy, T. F. (2017). Adult glioblastoma. *J. Clin. Oncol.* 35 (21), 2402–2409. doi: 10.1200/JCO.2017.73.0119

Alifieris, C., and Trafalis, D. T. (2015). Glioblastoma multiforme: Pathogenesis and treatment. *Pharmacol. Ther.* 152, 63–82. doi: 10.1016/j.pharmthera.2015.05.005

Batash, R., Asna, N., Schaffer, P., Francis, N., and Schaffer, M. (2017). Glioblastoma multiforme, diagnosis and treatment; recent literature review. *Curr. Med. Chem.* 24 (27), 3002–3009. doi: 10.2174/0929867324666170516123206

Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A. C., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 39 (4), 782–795. doi: 10.1016/j.immuni.2013.10.003

Braiterman, L. T., Zheng, S., Watkins, P. A., Geraghty, M. T., Johnson, G., McGuinness, M. C., et al. (1998). Suppression of peroxisomal membrane protein defects by peroxisomal ATP binding cassette (ABC) proteins. *Hum. Mol. Genet.* 7 (2), 239–247. doi: 10.1093/hmg/7.2.239

Chen, R., Smith-Cohn, M., Cohen, A. L., and Colman, H. (2017). Glioma subclassifications and their clinical significance. *Neurotherapeutics* 14 (2), 284–297. doi: 10.1007/s13311-017-0519-x

Cho, E. S., Kang, H. E., Kim, N. H., and Yook, J. I. (2019). Therapeutic implications of cancer epithelial-mesenchymal transition (EMT). *Arch. Pharm. Res.* 42 (1), 14–24. doi: 10.1007/s12272-018-01108-7

Dean, M., Rzhetsky, A., and Allikmets, R. (2001). The human ATP-binding cassette (ABC) transporter superfamily. *Genome Res.* 11 (7), 1156–1166. doi: 10.1101/gr.184901

Elsnerova, K., Mohelnikova-Duchonova, B., Cerovska, E., Ehrlichova, M., Gut, I., Rob, L., et al. (2016). Gene expression of membrane transporters: Importance for prognosis and progression of ovarian carcinoma. *Oncol. Rep.* 35 (4), 2159–2170. doi: 10.3892/or.2016.4599

Foreman, P. M., Friedman, G. K., Cassady, K. A., and Markert, J. M. (2017). Oncolytic virotherapy for the treatment of malignant glioma. *Neurotherapeutics* 14 (2), 333–344. doi: 10.1007/s13311-017-0516-0

Hemminki, O., Dos Santos, J. M., and Hemminki, A. (2020). Oncolytic viruses for cancer immunotherapy. *J. Hematol. Oncol.* 13 (1), 020–00922. doi: 10.1186/s13045-020-00922-1

Hu, M., Li, H., Xie, H., Fan, M., Wang, J., Zhang, N., et al. (2021a). ELF1 transcription factor enhances the progression of glioma *via* ATF5 promoter. *ACS Chem. Neurosci.* 12 (7), 1252–1261. doi: 10.1021/acschemneuro.1c00070

Hu, M., Yu, B., Zhang, B., Wang, B., Qian, D., Li, H., et al. (2021b). Human cytomegalovirus infection activates glioma activating transcription factor 5 *via* microRNA in a stress-induced manner. *ACS Chem. Neurosci.* 12 (20), 3947–3956. doi: 10.1021/acschemneuro.1c00576

Imanaka, T., Aihara, K., Suzuki, Y., Yokota, S., and Osumi, T. (2000). The 70-kDa peroxisomal membrane protein (PMP70), an ATP-binding cassette transporter. *Cell Biochem. Biophys.* 32, 1–3. doi: 10.1385/CBB:32:1-3:131

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 77 (21), e108–e110. doi: 10.1158/0008-5472.CAN-17-0307

Lima, R. A., Cândido, E. B., de Melo, F. P., Piedade, J. B., Vidigal, P. V., Silva, L. M., et al. (2015). Gene expression profile of ABC transporters and cytotoxic effect of ibuprofen and acetaminophen in an epithelial ovarian cancer cell line *in vitro*. *Rev. Bras. Ginecol. Obstet.* 37 (6), 283–290. doi: 10.1590/SO100-720320150005292

Pedemonte, E., Mancardi, G., Giunti, D., Corcione, A., Benvenuto, F., Pistoia, V., et al. (2006). Mechanisms of the adaptive immune response inside the central nervous system during inflammatory and autoimmune diseases. *Pharmacol. Ther.* 111 (3), 555–566. doi: 10.1016/j.pharmthera.2005.11.007

Reams, R. R., Jones-Triche, J., Chan, O. T., Hernandez, B. Y., Soliman, K. F., and Yates, C. (2015). Immunohistological analysis of ABCD3 expression in Caucasian and African American prostate tumors. *BioMed. Res. Int.* 132981 (10), 31. doi: 10.1155/2015/132981

Rius-Rocabert, S., García-Romero, N., García, A., Ayuso-Sacido, A., and Nistal-Villan, E. (2020). Oncolytic virotherapy in glioma tumors. *Int. J. Mol. Sci.* 21 (20), 7604. doi: 10.3390/ijms21207604

Roesch, S., Rapp, C., Dettling, S., and Herold-Mende, C. (2018). When immune cells turn bad-Tumor-Associated Microglia/Macrophages in glioma. *Int. J. Mol. Sci.* 19 (2), 436. doi: 10.3390/ijms19020436

Sokratous, G., Polyzoidis, S., and Ashkan, K. (2017). Immune infiltration of tumor microenvironment following immunotherapy for glioblastoma multiforme. *Hum. Vaccin. Immunother.* 13 (11), 2575–2582. doi: 10.1080/21645515.2017.1303582

Suarez-Carmona, M., Lesage, J., Cataldo, D., and Gilles, C. (2017). EMT and inflammation: inseparable actors of cancer progression. *Mol. Oncol.* 11 (7), 805–823. doi: 10.1002/1878-0261.12095

Suryawanshi, Y. R., and Schulze, A. J. (2021). Oncolytic viruses for malignant glioma: On the verge of success? *Viruses* 13 (7), 1294. doi: 10.3390/v13071294

Tran, Q. N. (2013). A novel method for finding non-small cell lung cancer diagnosis biomarkers. *BMC Med. Genomics* 1 (Suppl 1), 1755–8794. doi: 10.1186/1755-8794-6-S1-S11

Walton, R. W., Brown, M. C., Sacco, M. T., and Gromeier, M. (2018). Engineered oncolytic poliovirus PVSRIPO subverts MDA5-dependent innate immune responses in cancer cells. *J. Virol.* 92 (19), 00879–00818. doi: 10.1128/JVI.00879-18

Wang, T. J. C., and Mehta, M. P. (2019). Low-grade glioma radiotherapy treatment and trials. *Neurosurg. Clin. N. Am.* 30 (1), 111–118. doi: 10.1016/j.nec.2018.08.008

Witthayanuwat, S., Pesee, M., Supaadirek, C., Supakalin, N., Thamronganantasakul, K., and Krusun, S. (2018). Survival analysis of glioblastoma multiforme. *Asian Pac. J. Cancer Prev.* 19 (9), 2613–2617. doi: 10.22034/APJCP.2018.19.9.2613

Zeng, J., Li, X., Sander, M., Zhang, H., Yan, G., and Lin, Y. (2021). Oncolytic viro-immunotherapy: An emerging option in the treatment of gliomas. *Front. Immunol.* 12 (721830). doi: 10.3389/fimmu.2021.721830

Zhang, X., Wang, H., Sun, Y., Qi, M., Li, W., Zhang, Z., et al. (2020a). Enterovirus A71 oncolysis of malignant gliomas. *Mol. Ther.* 28 (6), 1533–1546. doi: 10.1016/j.ymthe.2020.04.005

Zhang, Y., Wang, J., Yang, J., and Yang, G. (2020b). Abnormal expression of ABCD3 is an independent prognostic factor for colorectal cancer. *Oncol. Lett.* 19 (5), 3567–3577. doi: 10.3892/ol.2020.11463

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Mini Review
PUBLISHED 28 July 2022
DOI 10.3389/fcimb.2022.955847

OPEN ACCESS

EDITED BY
Ming Hu,
Qingdao University, China

REVIEWED BY
Matthew Lawson Scott,
Louisiana State University Health Shreveport, United States
Marcela Lizano,
National Institute of Cancerology (INCAN), Mexico

*CORRESPONDENCE
Parvapan Bhattarakosol
parvapan.b@chula.ac.th;bhparvapan@gmail.com

SPECIALTY SECTION
This article was submitted to Virus and Host, a section of the journal Frontiers in Cellular and Infection Microbiology

RECEIVED 29 May 2022
ACCEPTED 06 July 2022
PUBLISHED 28 July 2022

CITATION
Baedyananda F, Sasivimolrattana T, Chaiwongkot A, Varadarajan S and Bhattarakosol P (2022) Role of HPV16 E1 in cervical carcinogenesis. *Front. Cell. Infect. Microbiol.* 12:955847. doi: 10.3389/fcimb.2022.955847

Role of HPV16 E1 in cervical carcinogenesis

Fern Baedyananda[1,2], Thanayod Sasivimolrattana[3,4], Arkom Chaiwongkot[1,4], Shankar Varadarajan[2] and Parvapan Bhattarakosol[1,4]*

[1]Division of Virology, Department of Microbiology, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand, [2]Department of Molecular and Clinical Cancer Medicine, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, United Kingdom, [3]Medical Microbiology Interdisciplinary Program, Graduate School, Chulalongkorn University, Bangkok, Thailand, [4]Center of Excellence in Applied Medical Virology, Department of Microbiology, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

Cervical cancer is the fourth most common cancer in women worldwide. More than 90% of cases are caused by the human papillomavirus (HPV). Vaccines developed only guard against a few HPV types and do not protect people who have already been infected. HPV is a small DNA virus that infects the basal layer of the stratified epithelium of the skin and mucosa through small breaks and replicates as the cells differentiate. The mucosal types of HPV can be classified into low-risk and high-risk groups, based on their association with cancer. Among HPV types in high-risk group, HPV type 16 (HPV-16) is the most common, causing 50% of all cancer cases. HPV infection can occur as transient or persistent infections, based on the ability of immune system to clear the virus. Persistent infection is characterized by the integration of HPV genome. HPV-16 exhibits a different integration pattern, with only 50% reported to be integrated at the carcinoma stage. Replication of the HPV genome depends on protein E1, an ATP-dependent helicase. E1 is essential for the amplification of the viral episome in infected cells. Previous studies have shown that E1 does not only act as a helicase protein but is also involved in recruiting and interacting with other host proteins. E1 has also been deemed to drive host cell proliferation. Recent studies have emphasized the emerging role of HPV E1 in cervical carcinogenesis. In this review, a possible mechanism by which E1 drives cell proliferation and oncogenesis will be discussed.

KEYWORDS

HPV, HPV16, HPV16 E1, cervical carcinogenesis, cervical cancer

Introduction

Cervical cancer (CC) is one of the most frequent common malignancies in women worldwide. Persistent infection with some types of human papillomaviruses (HPVs) is the major factor contributing to the development of cervical carcinogenesis (Schiffman et al., 2007; Petry, 2014). To observe the histological abnormality of the patients, colposcopy with biopsy, endocervical scraping, and cone biopsies are performed. The abnormalities detected in biopsy samples are termed cervical intraepithelial neoplasia (CIN) or dysplasia. These CINs are divided into 3 major stages, i.e., CIN1, CIN2, and CIN3, including carcinoma *in situ* (Waxman et al., 2012). Moreover, if cancerous cells are observed, they will be identified as squamous cell carcinoma (SCC) or adenocarcinoma.

HPVs are epitheliotropic viruses with an 8 kbp double-stranded circular DNA genome contained in the naked icosahedral capsid. The HPV's genomic DNA is packaged as a minichromosome with cellular nucleosomal histone (Acheson, 2011). HPV belongs to family *Papillomaviridae* and currently divided into 5 genera, namely *Alpha-*, *Beta-*, *Gamma-*, *Mupa-* and *Nupapapillomavirus* containing members, all of which can infect humans (Brooks et al., 2013). To this date, more than 200 HPV types have been identified (McBride, 2017). In the *Alphapapillomavirus* genus, HPVs are divided into two major groups - high-risk HPVs (Hr-HPVs) and low-risk HPVs (Lr-HPVs) - based on the cancer risk associated with their infection. Hr-HPVs include types 16, 18, 31 and 33 (Ghittoni et al., 2015), whereas Lr-HPVs include types 6 and11 (Brooks et al., 2013).

Papillomaviruses can infect a wide range of animal species; however, each type of papillomavirus is highly host- and tissue-specific. HPV infects cells in the basal layer of the epithelium, following wounds or breaks in the epithelium. Upon infection, HPVs maintain their genome as an extrachromosomal element, or episome, in the nucleus of infected cells. At this infection stage, and while the cells are in the lower strata of the epithelium, only the early genes are expressed (Graham, 2017). When the infected cell proliferates, the HPV genome replicates and increases the episomal copy numbers in the cell. The viral genome is replicated along with host cell DNA replication; after cell division occurs, the daughter cell contains a copy of the HPV genome. There is no new virion progeny produced in this initial phase. When the cell proliferates and differentiates, HPV DNA replication increases, resulting in a high episomal viral genome copy number (Burd, 2003). Finally, in the upper strata of the epithelium, the late proteins are encoded for capsid formation and released in the upper strata of the epithelium (Longworth and Laimins, 2004). Most HPV infections are transient and are cleared within approximately 2 years; nevertheless, if the host immune system is unable to clear the infection, a persistent infection occurs, possibly leading to viral genome integration into the host cell (Plummer et al., 2007).

The HPV genome encodes about 8 open-reading frames (ORFs), of which are divided into three functional regions - early (E) region, late (L) region, and noncoding part or long control region (LCR) (Yousefi et al., 2021). The early region encodes 6 early proteins (E1, E2, E4, E5, E6, and E7), whereas the late region encodes only 2 structural proteins, L1, and L2, that compose the capsid of HPVs (Chan et al., 2019). Whilst E1 has been shown to encode the primary protein responsible for viral replication (Berg and Stenlund, 1997), E2 is involved in transcriptional regulation (Chojnacki and Melendy, 2018), and E4 regulates virion release (Doorbar et al., 1996). The late proteins L1 and L2 are the major (80%) and minor (20%) capsid proteins, respectively. In contrast to these proteins, the functional roles of E5, E6 and E7 have been extensively characterized in the context of cancer. For instance, HPV E5 plays a role in immune evasion whereas E6 and E7 causes hyper-proliferation of the cell and cancer progression (Crook et al., 1991; Ganguly, 2012). The major role of the HPV E6 oncoprotein is to immortalize the cells *via* the ubiquitin-dependent proteasome degradation of p53, a cellular tumor suppressor protein, thus evading cancer cell death (Nguyen et al., 2002; Kelley et al., 2005). In addition, E6 can also degrade other apoptotic signaling cascade molecules (Filippova et al., 2002). HPV E7 oncoprotein plays a critical role in cervical carcinogenesis through dysregulation of cell cycle. This protein inactivates the retinoblastoma tumor suppressor protein (pRb) and downregulates E2F (Moody and Laimins, 2010). It is therefore evident that the role of HPV proteins, other than E5, E6, and E7, in carcinogenesis is under-investigated. Recently, it has been shown that HPV E1 may be involved in carcinogenesis. Here, we will review the possible role of E1 in cervical carcinogenesis.

The structure of E1 protein

The E1 protein serves as the primary replication protein of HPV and is an ATP-dependent helicase that binds to the viral origin of replication and unwinds the viral DNA to initiate replication (Hughes and Romanos, 1993). Across all papillomaviruses, the E1 protein is the most conserved, primarily owing to its helicase function, which is essential for the viral episome replication. It is possible that all phases of the viral replication cycle namely, establishment, maintenance, and amplification, require E1 function. The E1 protein consists of three main domains, each with a distinct and important function, i.e., The N-terminal regulatory domain, the DNA binding domain, and the helicase domain (Figure 1A) (Bergvall et al., 2013). The N-terminal of the protein contains the nuclear localization signal and the nuclear export signal,

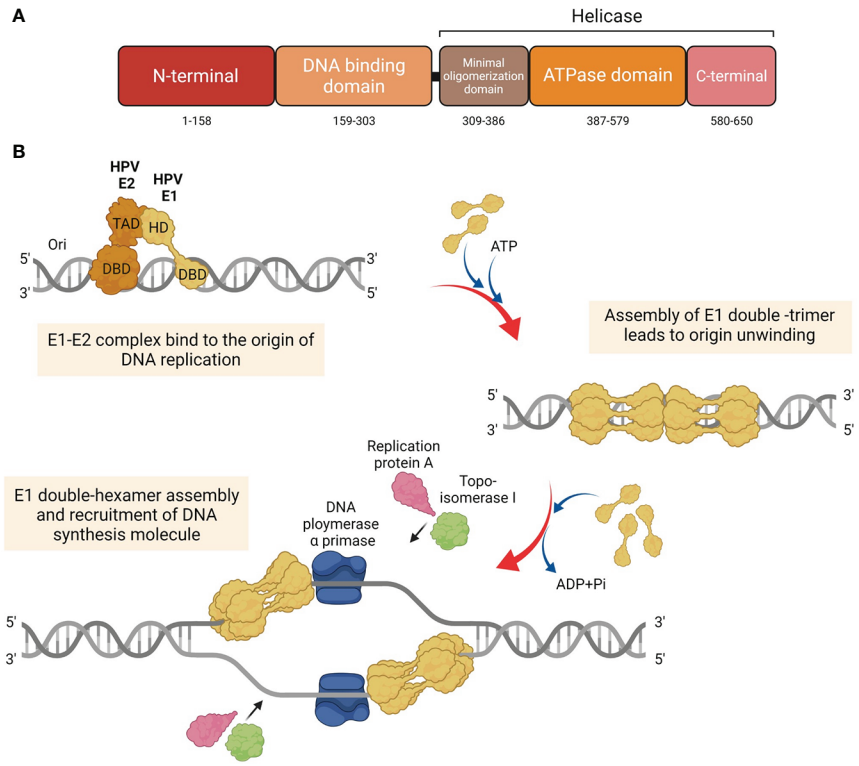


FIGURE 1
Structure and function of HPV E1. **(A)** Diagram of the HPV E1 protein domains. Three major functional domains of HPV E1, i.e., N-terminal, DNA binding domain (DBD), and the domain which construct the helicase domain: minimal oligomerization domain, ATPase domain, and C-terminal brace. **(B)** Schematic representation of the initiation of the HPV DNA replication associated with E1 protein. HPV E1 proteins are recruited to bind to the E1 DNA binding site at the origin by HPV E2. E1 and E2 are assemble as E1-E2-ori ternary complex. Additional of E1 proteins are recruited to assemble as E1 double-trimer and double-hexamer, respectively. Then, the ds-DNA are unwound, and the DNA replication are initiated by the recruitment of host DNA replication factors.

which functions to transport E1 between the nucleus and cytoplasm. The DNA binding domain recognizes specific sequences near the viral origin of replication, which is bound by the helicase domain to form a doughnut shaped complex around viral DNA (Enemark and Joshua-Tor, 2006).

Mechanisms of HPV E1 in viral DNA replication

To initiate HPV DNA replication, E1 interacts with E2 (Berg and Stenlund, 1997; Bergvall et al., 2013; Porter et al., 2017), which aids in its recruitment to the origin of replication (Frattini and Laimins, 1994) (Figure 1B). It must be noted that E1 can also establish viral replication in an E2 independent manner, albeit with lower efficiency (Bonne-Andrea et al., 1997). Upon binding to the origin of replication as an E1-E2-ori ternary complex, additional E1 molecules are recruited to assemble E1 double-trimers and E1 double-hexamers, which whilst unwinding DNA also recruits several DNA replication factors, such as DNA polymerase α primase (Pol α-prim), topoisomerase I (Topo I), and replication protein A (RPA) (Loo and Melendy, 2004; Bergvall et al., 2013).

Relationship between HPV16 E1 expression and cervical cancer progression

The most well-characterized function of HPV in cervical carcinogenesis is the overexpression of E6 and E7 oncoproteins, which mainly target p53 and pRb tumor suppressor proteins (Munger et al., 1992; Mantovani and Banks, 2001). However, a relationship between HPV E1 expression and cervical cancer progression has been reported. Interestingly, HPV16 E1 mRNA expression is positively correlated with cervical cancer progression (Baedyananda et al., 2017). This finding was in agreement with

another independent study that reported lower E1 expression in patients with low-grade squamous intraepithelial lesion (LSIL) compared to high-grade squamous intraepithelial lesion (HSIL) and cancer (Schmitt et al., 2011).

The physical state of the HPV genome, namely integrated, episomal, or mixed, is correlated with cervical carcinogenesis (Pirami et al., 1997; Williams et al., 2011). HPV integration is associated with overexpression of E6 and E7 (Jeon et al., 1995), which in turn have been shown to enhance the pathogenicity of HPV (Munger et al., 1992; Burke et al., 2012; Burke et al., 2014). However, recent findings have reported that E1 mRNA expression is neither impacted by the number of copies of the HPV16 genome, nor its physical state (Baedyananda et al., 2017). Similar observations have also been made in relation to E6 and E7 (Wang-Johanning et al., 2002), which implicates other factors, including epigenetic modifications in the control of HPV mRNA expression. Moreover, HPV genome has been reported to be highly methylated in cervical cancer samples with high copy numbers of integrated HPV (Chaiwongkot et al., 2013). In agreement, E1 promoters p97 and p670 are hypermethylated in cervical cancer samples when compared to normal samples (Baedyananda et al., 2017), thereby suggesting the possibility of these samples possessing high numbers of HPV genome copies.

The possible functions of HPV16 E1 in cervical carcinogenesis

HPV16 E1 dysregulates the expression of genes involved in cell survival

Failure to induce apoptosis could drive tumorigenesis through multiple mechanisms (Ichim and Tait, 2016). In HPV16 E1 overexpressing cells, several host genes that are involved in protein synthesis (RPL36A), metabolism (ALDOC), immune response (ISG20), DNA damage (ATR, BRCA1, and CHK1), and cell proliferation (CREB5, HIF1A, NFKB1, PIK3CA, JMJDIC, TSC22D3, FOXO3), have been shown to be significantly downregulated (Baedyananda et al., 2021). Of the transcriptional factors, CREB5, HIF1A, NFKB1, and PIK3CA, downregulation of NFKB1 and PIK3CA supports tumor growth and survival (Xia et al., 2014; Masoud and Li, 2015). JMJD1C (Jumonji Domain Containing 1C) is a histone demethylase, tumor suppressor protein that is usually found to be reduced or lost in breast cancer (Watanabe et al., 2013). TSC22D3 is a gene that encodes Glucocorticoid-induced leucine zipper (GILZ), which in turn activates FOXO3a-mediated transcription of the pro-apoptotic protein, Bim (Joha et al., 2012). Loss of FOXO3a activity has previously been linked to disease progression in carcinogen-induced lung adenocarcinoma (Blake et al., 2010), neck cancer (Shou et al., 2012) and urothelial cancer (Shiota et al., 2010). How HPV16 E1 promotes cell survival *via* regulation of FOXO3 in cervical carcinogenesis is still unclear.

HPV16 E1 mutations associated with cervical cancer

A variety of mutations in E1/E2 genes have been identified in cervical cancer samples. The clinical stage of patients with the discovered HPV E1 mutations have also been noted in various studies. Many of these variants were linked with lower grade lesions such as noted in previous studies where a 63-bp duplication variant was associated with lower disease progression (Sabol et al., 2008; Baedyananda et al., 2017). Other mutations have demonstrated reduced abilities to support HPV replication (Yao et al., 2019) and also failed to suppress the viral early promoter (Yao et al., 2019; Hirose et al., 2020). When examining within-host variations of HPV16 it was noted that non-synonymous substitutions occurred more frequently in invasive carcinoma specimens. These mutations also occurred more frequently in E1/E2 regions than in other regions of the viral genome (Hirose et al., 2020). Many factors determine the functional impact of mutations i.e., type of mutation, location of the mutation. Further studies are required to understand how information gained from this could be therapeutically exploited.

Helicase function of HPV E1 and possible roles in cancer

Genome instability, a hallmark of cancer (Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011), is caused by several factors that induce DNA damage (Langie et al., 2015), which could be exacerbated by defective DNA damage repair response and tumor suppressor molecules (Negrini et al., 2010). Inactivating mutations in DNA helicases have detrimental effects, such as in Werner syndrome, characterized by the appearance of premature aging features and early onset of age-related diseases such as cardiovascular diseases, diabetes mellitus, and carcinoma (Martin, 1985). Patients with Bloom's syndrome are also predisposed to carcinogenesis (German, 1997). In addition, cells transformed by the Epstein-Barr virus (EBV) and simian virus 40 (SV40) have exhibited upregulated helicases expression. In the context of HPV infection, E1 helicase facilitates HPV genome replication, *via* the host cells' DNA damage repair pathways and contributes to carcinogenesis (Moody and Laimins, 2009; King et al., 2010; Fradet-Turcotte et al., 2011; Sakakibara et al., 2011). HPV E1 proteins are able to directly cause double strand DNA breaks in the host genome.

The ATR-dependent DNA damage response pathway is activated by HPV replication or presence of HPV replication proteins, E1 and E2. This has been shown by the accumulation of γH2AX, ATR-interacting protein (ATRIP), and topoisomerase IIβ-binding protein 1 (TopBP1) in replication centers. Conversely, the viral oncoproteins E6 and E7 did not play a role in the accumulation of γH2AX, ATRIP and TopBP1 (Reinson et al., 2013).

Moreover, the homeobox transcription factor HOXC13 has been shown to upregulate the expression of HPV16/18 E1 and E2 (Ishii et al., 2020), which may enhance the helicase activity of E1 to result in DNA damage and genome instability. HPV E1's ability to directly cause host DNA damage and the implications on carcinogenesis should be further explored to better define the role of HPV E1 in carcinogenesis.

Role of HPV16 E1 in antiviral immune evasion

Type I interferons (IFN), such as, IFN-α and IFN-β, form integral components of the innate immune response, playing important roles in antiviral, anti-proliferative and immunomodulatory functions (Seth et al., 2006). HPV18 E1 modulates the expression of the genes involved in toll-like receptor, interferons and apoptosis pathways, and antiviral interferon-stimulated gene set (Castillo et al., 2014). Similarly, HPV16/18 E1 protein has been shown to enhance the expression of immune response genes, i.e., IFNβ1, IFNλ1, and interferon-stimulated gene (ISG) (Castro-Muñoz et al., 2019), thus implicating roles for HPV E1 in immune modulation, as evasion of the host immune response is key to persistent infection and HPV-related carcinogenesis.

Conclusion

Several studies have explored possible roles of HPV16 E1 in cervical carcinogenesis and have attributed numerous cellular mechanisms, including increased expression of genes involved in cell survival and apoptosis, inhibition of the antiviral immune response, viral genome maintenance, helicase activity and inactivating mutations (Figure 2). However, there are no functional evidence of HPV16 E1 directly on cancer development. It is possible that E1 is the only protein with both DNA helicase and ATPase activity, which may or may not be relevant for carcinogenesis. Unfortunately, only few studies worked on HPV16 E1 related to cancer development. For further study, the function of E1 protein in cervical carcinogenesis should be deeply explored. More functional assays related to the hallmarks of cancer, e.g., cell proliferation, apoptosis, cell cycle arrested, migration/invasion, wound healing, and colony formation, should be done in E1 overexpressed/knockdown/knockout cells. Moreover, many genes such as BCL2L1, CSP2, FOXO3a, JMJDIC, and TSC22D3, were dysregulated in either HPV18 or HPV16 E1 overexpressed/knockdown cells (Castillo et al., 2014; Baedyananda et al., 2021). To better understand the E1 associated cellular pathways, the relationship between E1 and those genes as well as E1 protein-protein interaction should be investigated. Since HPV E1 proteins are helicases, studying cells over-expressing HPV16 E1 and

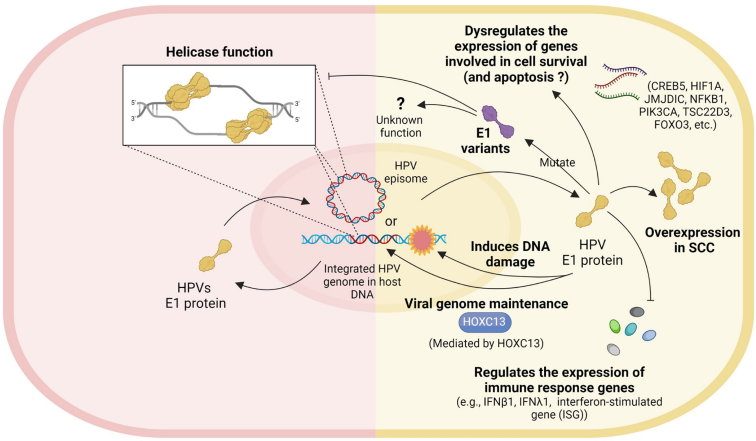


FIGURE 2
Role of HPV E1. Left panel: the well-known helicase function of HPV. Right panel: the possible roles of HPV E1 in cervical carcinogenesis include dysregulating the expression of genes involved in cell survival, E1 mutation, E1 overexpression, inducing DNA damage, viral genome maintenance, and regulating the expression of immune response genes.

observing the impact on genome instability could expand our understanding of the carcinogenic role of HPV16 E1. In addition, the association between the expression of some cellular genes mentioned above and various stage of cervical specimens should be determined to confirm the phenomenon truly occur in human. Taken together, uncovering the cellular and molecular mechanisms underlying HPV16 E1-mediated cervical carcinogenesis is therefore critical to identify novel biomarkers and/or druggable targets and/or a possible forth HPV oncoprotein beside E5, E6 and E7.

Author contributions

Conceptualization, FB, TS, and PB; Writing – Original Draft Preparation, FB, and TS; Writing – Review & Editing, PB, SV, and AC; Supervision, PB; Project Administration, PB and SV; Funding Acquisition, FB, TS, PB, SV and AC; Figure Illustration, TS. All authors contributed to the article and approved the submitted version.

Funding

The 100th Anniversary Chulalongkorn University Fund for Doctoral Scholarship, The Overseas Research Experience Scholarship for Graduate Student, Rachadaphiseksomphot Endowment Fund, and Thailand Science Research and Innovation Fund (TSRI), Chulalongkorn University, Bangkok 10330, Thailand and University of Liverpool, UK.

Acknowledgments

We would like to acknowledge that all figures were created with BioRender.com platform.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Acheson, N. H. (2011). Fundamentals of Molecular Virology. (Hoboken, New Jersey), *USA: John Wiley & Sons, INC.* (John Wiley & Sons, INC).

Baedyananda, F., Chaiwongkot, A., and Bhattarakosol, P. (2017). Elevated HPV16 E1 expression is associated with cervical cancer progression. *Intervirology* 60 (5), 171–180. doi: 10.1159/000487048

Baedyananda, F., Chaiwongkot, A., Varadarajan, S., and Bhattarakosol, P. (2021). HPV16 E1 dysregulated cellular genes involved in cell proliferation and host DNA damage: A possible role in cervical carcinogenesis. *PloS One* 16 (12), e0260841. doi: 10.1371/journal.pone.0260841

Berg, M., and Stenlund, A. (1997). Functional interactions between papillomavirus E1 and E2 proteins. *J. Virol.* 71 (5), 3853–3863. doi: 10.1128/JVI.71.5.3853-3863.1997

Bergvall, M., Melendy, T., and Archambault, J. (2013). The E1 proteins. *Virology* 445 (1-2), 35–56. doi: 10.1016/j.virol.2013.07.020

Blake, D. C.Jr., Mikse, O. R., Freeman, W. M., and Herzog, C. R. (2010). FOXO3a elicits a pro-apoptotic transcription program and cellular response to human lung carcinogen nicotine-derived nitrosaminoketone (NNK). *Lung Cancer* 67 (1), 37–47. doi: 10.1016/j.lungcan.2009.03.013

Bonne-Andrea, C., Tillier, F., McShan, G. D., Wilson, V. G., and Clertant, P. (1997). Bovine papillomavirus type 1 DNA replication: the transcriptional activator E2 acts *in vitro* as a specificity factor. *J. Virol.* 71 (9), 6805–6815.

Brooks, G. F., Carroll, K. C., Butel, J. S., Morse, S. A., and Mietzner, T. A. (2013). *Medical Microbiology* (United States: McGraw-Hill Education).

Burd, E. M. (2003). Human papillomavirus and cervical cancer. *Clin. Microbiol. Rev.* 16 (1), 1–17.

Burke, J. R., Hura, G. L., and Rubin, S. M. (2012). Structures of inactive retinoblastoma protein reveal multiple mechanisms for cell cycle control. *Genes Dev.* 26 (11), 1156–1166. doi: 10.1101/gad.189837.112

Burke, J. R., Liban, T. J., Restrepo, T., Lee, H. W., and Rubin, S. M. (2014). Multiple mechanisms for E2F binding inhibition by phosphorylation of the retinoblastoma protein c-terminal domain. *J. Mol. Biol.* 426 (1), 245–255. doi: 10.1016/j.jmb.2013.09.031

Castillo, A., Wang, L., Koriyama, C., Eizuru, Y., Jordan, K., and Akiba, S. (2014). A systems biology analysis of the changes in gene expression *via* silencing of HPV-18 E1 expression in HeLa cells. *Open Biol.* 4 (10), 1-9. doi: 10.1098/rsob.130119

Castro-Muñoz, L. J., Manzo-Merino, J., Muñoz-Bello, J. O., Olmedo-Nieva, L., Cedro-Tanda, A., Alfaro-Ruiz, L. A., et al. (2019). The human papillomavirus (HPV) E1 protein regulates the expression of cellular genes involved in immune response. *Sci. Rep.* 9 (1), 13620. doi: 10.1038/s41598-019-49886-4

Chaiwongkot, A., Vinokurova, S., Pientong, C., Ekalaksananan, T., Kongyingyoes, B., Kleebkaow, P., et al. (2013). Differential methylation of E2 binding sites in episomal and integrated HPV 16 genomes in preinvasive and invasive cervical lesions. *Int. J. Cancer* 132 (9), 2087–2094. doi: 10.1002/ijc.27906

Chan, C. K., Aimagambetova, G., Ukybassova, T., Kongrtay, K., and Azizan, A. (2019). Human papillomavirus infection and cervical cancer: Epidemiology, screening, and vaccination-review of current perspectives. *J. Oncol.* 2019, 3257939. doi: 10.1155/2019/3257939

Chojnacki, M., and Melendy, T. (2018). The human papillomavirus DNA helicase E1 binds, stimulates, and confers processivity to cellular DNA polymerase epsilon. *Nucleic Acids Res.* 46 (1), 229–241. doi: 10.1093/nar/gkx1103

Crook, T., Tidy, J. A., and Vousden, K. H. (1991). Degradation of p53 can be targeted by HPV E6 sequences distinct from those required for p53 binding and trans-activation. *Cell* 67 (3), 547–556. doi: 10.1016/0092-8674(91)90529-8

Doorbar, J., Medcalf, E., and Napthine, S. (1996). Analysis of HPV1 E4 complexes and their association with keratins *in vivo*. *Virology* 218 (1), 114–126. doi: 10.1006/viro.1996.0171

Enemark, E. J., and Joshua-Tor, L. (2006). Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* 442 (7100), 270–275. doi: 10.1038/nature04943

Filippova, M., Song, H., Connolly, J. L., Dermody, T. S., and Duerksen-Hughes, P. J. (2002). The human papillomavirus 16 E6 protein binds to tumor necrosis factor (TNF) R1 and protects cells from TNF-induced apoptosis. *J. Biol. Chem.* 277 (24), 21730–21739. doi: 10.1074/jbc.M200113200

Fradet-Turcotte, A., Bergeron-Labrecque, F., Moody, C. A., Lehoux, M., Laimins, L. A., and Archambault, J. (2011). Nuclear accumulation of the

papillomavirus E1 helicase blocks s-phase progression and triggers an ATM-dependent DNA damage response. *J. Virol.* 85 (17), 8996–9012. doi: 10.1128/JVI.00542-11

Frattini, M. G., and Laimins, L. A. (1994). Binding of the human papillomavirus E1 origin-recognition protein is regulated through complex formation with the E2 enhancer-binding protein. *Proc. Natl. Acad. Sci. U.S.A.* 91 (26), 12398–12402.

Ganguly, N. (2012). Human papillomavirus-16 E5 protein: oncogenic role and therapeutic value. *Cell Oncol. (Dordr)* 35 (2), 67–76. doi: 10.1007/s13402-011-0069-x

German, J. (1997). Bloom's syndrome. XX. the first 100 cancers. *Cancer Genet. Cytogenet.* 93 (1), 100–106.

Ghittoni, R., Accardi, R., Chiocca, S., and Tommasino, M. (2015). Role of human papillomaviruses in carcinogenesis. *Ecancermedicalscience* 9, 526. doi: 10.3332/ecancer.2015.526

Graham, S. V. (2017). Keratinocyte differentiation-dependent human papillomavirus gene regulation. *Viruses* 9 (9), 1-18. doi: 10.3390/v9090245

Hanahan, D., and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell* 100 (1), 57–70. doi: 10.1016/s0092-8674(00)81683-9

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144 (5), 646–674. doi: 10.1016/j.cell.2011.02.013

Hirose, Y., Yamaguchi-Naka, M., Onuki, M., Tenjimbayashi, Y., Tasaka, N., Satoh, T., et al. (2020). High levels of within-host variations of human papillomavirus 16 E1/E2 genes in invasive cervical cancer. *Front. Microbiol.* 11. doi: 10.3389/fmicb.2020.596334

Hughes, F. J., and Romanos, M. A. (1993). E1 protein of human papillomavirus is a DNA helicase/ATPase. *Nucleic Acids Res.* 21 (25), 5817–5823.

Ichim, G., and Tait, S. W. (2016). A fate worse than death: apoptosis as an oncogenic process. *Nat. Rev. Cancer* 16 (8), 539–548. doi: 10.1038/nrc.2016.58

Ishii, Y., Taguchi, A., and Kukimoto, I. (2020). The homeobox transcription factor HOXC13 upregulates human papillomavirus E1 gene expression and contributes to viral genome maintenance. *FEBS Lett.* 594 (4), 751–762. doi: 10.1002/1873-3468.13646

Jeon, S., Allen-Hoffmann, B. L., and Lambert, P. F. (1995). Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. *J. Virol.* 69 (5), 2989–2997. doi: 10.1128/JVI.69.5.2989-2997.1995

Joha, S., Nugues, A. L., Hetuin, D., Berthon, C., Dezitter, X., Dauphin, V., et al. (2012). GILZ inhibits the mTORC2/AKT pathway in BCR-ABL(+) cells. *Oncogene* 31 (11), 1419–1430. doi: 10.1038/onc.2011.328

Kelley, M. L., Keiger, K. E., Lee, C. J., and Huibregtse, J. M. (2005). The global transcriptional effects of the human papillomavirus E6 protein in cervical carcinoma cell lines are mediated by the E6AP ubiquitin ligase. *J. Virol.* 79 (6), 3737–3747. doi: 10.1128/JVI.79.6.3737-3747.2005

King, L. E., Fisk, J. C., Dornan, E. S., Donaldson, M. M., Melendy, T., and Morgan, I. M. (2010). Human papillomavirus E1 and E2 mediated DNA replication is not arrested by DNA damage signalling. *Virology* 406 (1), 95–102. doi: 10.1016/j.virol.2010.06.033

Langie, S. A., Koppen, G., Desaulniers, D., Al-Mulla, F., Al-Temaimi, R., Amedei, A., et al. (2015). Causes of genome instability: the effect of low dose chemical exposures in modern society. *Carcinogenesis* 36 Suppl 1, S61–S88. doi: 10.1093/carcin/bgv031

Longworth, M. S., and Laimins, L. A. (2004). Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol. Mol. Biol. Rev.* 68 (2), 362–372. doi: 10.1128/MMBR.68.2.362-372.2004

Loo, Y. M., and Melendy, T. (2004). Recruitment of replication protein a by the papillomavirus E1 protein and modulation by single-stranded DNA. *J. Virol.* 78 (4), 1605–1615.

Mantovani, F., and Banks, L. (2001). The human papillomavirus E6 protein and its contribution to malignant progression. *Oncogene* 20 (54), 7874–7887. doi: 10.1038/sj.onc.1204869

Martin, G. M. (1985). Genetics and aging; the Werner syndrome as a segmental progeroid syndrome. *Adv. Exp. Med. Biol.* 190, 161–170.

Masoud, G. N., and Li, W. (2015). HIF-1α pathway: role, regulation and intervention for cancer therapy. *Acta Pharm. Sin. B* 5 (5), 378–389. doi: 10.1016/j.apsb.2015.05.007

McBride, A. A. (2017). Oncogenic human papillomaviruses. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 372 (1732), 1-9. doi: 10.1098/rstb.2016.0273

Moody, C. A., and Laimins, L. A. (2009). Human papillomaviruses activate the ATM DNA damage pathway for viral genome amplification upon differentiation. *PloS Pathog.* 5 (10), e1000605. doi: 10.1371/journal.ppat.1000605

Moody, C. A., and Laimins, L. A. (2010). Human papillomavirus oncoproteins: pathways to transformation. *Nat. Rev. Cancer* 10 (8), 550–560. doi: 10.1038/nrc2886

Munger, K., Scheffner, M., Huibregtse, J. M., and Howley, P. M. (1992). Interactions of HPV E6 and E7 oncoproteins with tumour suppressor gene products. *Cancer Surv.* 12, 197–217.

Negrini, S., Gorgoulis, V. G., and Halazonetis, T. D. (2010). Genomic instability–an evolving hallmark of cancer. *Nat. Rev. Mol. Cell Biol.* 11 (3), 220–228. doi: 10.1038/nrm2858

Nguyen, M., Song, S., Liem, A., Androphy, E., Liu, Y., and Lambert, P. F. (2002). A mutant of human papillomavirus type 16 E6 deficient in binding alpha-helix partners displays reduced oncogenic potential *in vivo*. *J. Virol.* 76 (24), 13039–13048. doi: 10.1128/jvi.76.24.13039-13048.2002

Petry, K. U. (2014). HPV and cervical cancer. *Scand. J. Clin. Lab. Invest. Suppl.* 244, 59–62. doi: 10.3109/00365513.2014.936683

Pirami, L., Giache, V., and Becciolini, A. (1997). Analysis of HPV16, 18, 31, and 35 DNA in pre-invasive and invasive lesions of the uterine cervix. *J. Clin. Pathol.* 50 (7), 600–604. doi: 10.1136/jcp.50.7.600

Plummer, M., Schiffman, M., Castle, P. E., Maucort-Boulch, D., Wheeler, C. M., and Group, A. (2007). A 2-year prospective study of human papillomavirus persistence among women with a cytological diagnosis of atypical squamous cells of undetermined significance or low-grade squamous intraepithelial lesion. *J. Infect. Dis.* 195 (11), 1582–1589. doi: 10.1086/516784

Porter, S. S., Stepp, W. H., Stamos, J. D., and McBride, A. A. (2017). Host cell restriction factors that limit transcription and replication of human papillomavirus. *Virus Res.* 231, 10–20. doi: 10.1016/j.virusres.2016.11.014

Reinson, T., Toots, M., Kadaja, M., Pipitch, R., Allik, M., Ustav, E., et al. (2013). Engagement of the ATR-dependent DNA damage response at the human papillomavirus 18 replication centers during the initial amplification. *J. Virol.* 87 (2), 951–964. doi: 10.1128/JVI.01943-12

Sabol, I., Matovina, M., Gasperov, N. M., and Grce, M. (2008). Identification of a novel human papillomavirus type 16 E1 gene variant with potentially reduced oncogenicity. *J. Med. Virol.* 80 (12), 2134–2140. doi: 10.1002/jmv.21304

Sakakibara, N., Mitra, R., and McBride, A. A. (2011). The papillomavirus E1 helicase activates a cellular DNA damage response in viral replication foci. *J. Virol.* 85 (17), 8981–8995. doi: 10.1128/JVI.00541-11

Schiffman, M., Castle, P. E., Jeronimo, J., Rodriguez, A. C., and Wacholder, S. (2007). Human papillomavirus and cervical cancer. *Lancet* 370 (9590), 890–907. doi: 10.1016/S0140-6736(07)61416-0

Schmitt, M., Dalstein, V., Waterboer, T., Clavel, C., Gissmann, L., and Pawlita, M. (2011). The HPV16 transcriptome in cervical lesions of different grades. *Mol. Cell Probes* 25 (5-6), 260–265. doi: 10.1016/j.mcp.2011.05.003

Seth, R. B., Sun, L., and Chen, Z. J. (2006). Antiviral innate immunity pathways. *Cell Res.* 16 (2), 141–147. doi: 10.1038/sj.cr.7310019

Shiota, M., Song, Y., Yokomizo, A., Kiyoshima, K., Tada, Y., Uchino, H., et al. (2010). Foxo3a suppression of urothelial cancer invasiveness through Twist1, y-box-binding protein 1, and e-cadherin regulation. *Clin. Cancer Res.* 16 (23), 5654–5663. doi: 10.1158/1078-0432.CCR-10-0376

Shou, Z., Lin, L., Liang, J., Li, J. L., and Chen, H. Y. (2012). Expression and prognosis of FOXO3a and HIF-1alpha in nasopharyngeal carcinoma. *J. Cancer Res. Clin. Oncol.* 138 (4), 585–593. doi: 10.1007/s00432-011-1125-7

Wang-Johanning, F., Lu, D. W., Wang, Y., Johnson, M. R., and Johanning, G. L. (2002). Quantitation of human papillomavirus 16 E6 and E7 DNA and RNA in residual material from ThinPrep papanicolaou tests using real-time polymerase chain reaction analysis. *Cancer* 94 (8), 2199–2210. doi: 10.1002/cncr.10439

Watanabe, S., Watanabe, K., Akimov, V., Bartkova, J., Blagoev, B., Lukas, J., et al. (2013). JMJD1C demethylates MDC1 to regulate the RNF8 and BRCA1-mediated chromatin response to DNA breaks. *Nat. Struct. Mol. Biol.* 20 (12), 1425–1433. doi: 10.1038/nsmb.2702

Waxman, A. G., Chelmow, D., Darragh, T. M., Lawson, H., and Moscicki, A. B. (2012). Revised terminology for cervical histopathology and its implications for management of high-grade squamous intraepithelial lesions of the cervix. *Obstet. Gynecol.* 120 (6), 1465–1471. doi: 10.1097/AOG.0b013e31827001d5

Williams, V. M., Filippova, M., Soto, U., and Duerksen-Hughes, P. J. (2011). HPV-DNA integration and carcinogenesis: putative roles for inflammation and oxidative stress. *Future Virol.* 6 (1), 45–57. doi: 10.2217/fvl.10.73

Xia, Y., Shen, S., and Verma, I. M. (2014). NF-κB, an active player in human cancers. *Cancer Immunol. Res.* 2 (9), 823–830. doi: 10.1158/2326-6066.Cir-14-0112

Yao, Y., Yan, Z., Dai, S., Li, C., Yang, L., Liu, S., et al. (2019). Human papillomavirus type 16 E1 mutations associated with cervical cancer in a han Chinese population. *Int. J. Med. Sci.* 16 (7), 1042–1049. doi: 10.7150/ijms.34279

Yousefi, Z., Aria, H., Ghaedrahmati, F., Bakhtiari, T., Azizi, M., Bastan, R., et al. (2021). An update on human papilloma virus vaccines: History, types, protection, and efficacy. *Front. Immunol.* 12. doi: 10.3389/fimmu.2021.805695

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 15 August 2022
DOI 10.3389/fcimb.2022.964469

OPEN ACCESS

EDITED BY
Ming Hu,
Qingdao University, China

REVIEWED BY
Xiao Wu,
Nanjing University of Chinese Medicine, China
Lei Zhang,
Huazhong University of Science and Technology, China
Xuan Tang,
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, China
Song Yang,
Beijing Ditan Hospital, Capital Medical University, China

*CORRESPONDENCE
Xiaofeng Yin
yinxiaofeng@sxmu.edu.cn
Yanan Qiao
yananqiao@sxmu.edu.cn

SPECIALTY SECTION
This article was submitted to Virus and Host, a section of the journal Frontiers in Cellular and Infection Microbiology

RECEIVED 08 June 2022
ACCEPTED 27 July 2022
PUBLISHED 15 August 2022

CITATION
Yin X, Li J, Hao Z, Ding R and Qiao Y (2022) A systematic study of traditional Chinese medicine treating hepatitis B virus-related hepatocellular carcinoma based on target-driven reverse network pharmacology.
Front. Cell. Infect. Microbiol. 12:964469.
doi: 10.3389/fcimb.2022.964469

A systematic study of traditional Chinese medicine treating hepatitis B virus-related hepatocellular carcinoma based on target-driven reverse network pharmacology

Xiaofeng Yin[1]*, Jinchuan Li[1], Zheng Hao[1], Rui Ding[1] and Yanan Qiao[2]*

[1]Department of Neurosurgery, Second Hospital of Shanxi Medical University, Taiyuan, China, [2]Department of Pharmacy, Second Hospital of Shanxi Medical University, Taiyuan, China

Hepatocellular carcinoma (HCC) is a serious global health problem, and hepatitis B virus (HBV) infection remains the leading cause of HCC. It is standard care to administer antiviral treatment for HBV-related HCC patients with concurrent anti-cancer therapy. However, a drug with repressive effects on both HBV infection and HCC has not been discovered yet. In addition, drug resistance and side effects have made existing therapeutic regimens suboptimal. Traditional Chinese medicine (TCM) has multi-ingredient and multi-target advantages in dealing with multifactorial HBV infection and HCC. TCM has long been served as a valuable source and inspiration for discovering new drugs. In present study, a target-driven reverse network pharmacology was applied for the first time to systematically study the therapeutic potential of TCM in treating HBV-related HCC. Firstly, 47 shared targets between HBV and HCC were screened as HBV-related HCC targets. Next, starting from 47 targets, the relevant chemical components and herbs were matched. A network containing 47 targets, 913 chemical components and 469 herbs was established. Then, the validated results showed that almost 80% of the herbs listed in chronic hepatitis B guidelines and primary liver cancer guidelines were included in the 469 herbs. Furthermore, functional analysis was conducted to understand the biological processes and pathways regulated by these 47 targets. The docking results indicated that the top 50 chemical components bound well to targets. Finally, the frequency statistical analysis results showed the 469 herbs against HBV-related HCC were mainly warm in property, bitter in taste, and distributed to the liver meridians. Taken together, a small library of 913 chemical components and 469 herbs against HBV-related HCC were obtained with a target-driven approach, thus paving the way for the development of therapeutic modalities to treat HBV-related HCC.

KEYWORDS

traditional Chinese medicine, hepatitis B virus-related hepatocellular carcinoma, target-driven, reverse network pharmacology, a systematic study

Introduction

Hepatocellular carcinoma (HCC) is a serious health problem worldwide and the second most common cause of death from all malignancies (Bray et al., 2018). Hepatitis B virus (HBV) infection is still the most common etiological factor of HCC worldwide, especially in Asia (Catherine de Martel et al., 2015). Currently, essentially all HCC management guidelines recommend routine antiviral treatment to avoid HBV reactivation during treatment for HCC and reduce the recurrence of HCC after curative treatment (Sarin et al., 2016; European Association for the Study of the Liver, 2017; Terrault et al., 2018). However, a drug with repression effects on both HBV infection and HCC is not yet marketed. In addition, the current therapeutic regimens are far from optimal because of drug resistance, adverse, and toxic effects (Qiu et al., 2020). Novel multi-target medications with anti-HBV and anti-HCC activities are therefore urgently needed.

Traditional Chinese medicine (TCM) has always held a privileged position as an essential source of inspiration for discovering innovative drugs. Because they have multi-ingredient, multi-target properties that result in pharmacological synergism, it is possible that TCM had distinct advantages in dealing with multifactorial HBV infection and HCC (Kim and Kim, 2020; Zhang et al., 2020). In addition, TCM has time-honored theories about the diagnosis and treatment of liver diseases (Wang et al., 2012). In real clinical practice in China, TCM is part of the treatment regimen for HBV infection based on the syndrome of Chinese medicine. In hepatoma cancer therapy, TCM is mainly used to improve the anticancer drugs' efficacy and reduce their toxicity.

Network pharmacology integrates system bioinformatics, multi-directional pharmacology and omics to develop new strategies and study drugs' action mechanisms (Zhang et al., 2021). It can reveal the roles of pharmacological interventions, especially multi-target drugs. The overall concept of network pharmacology concurs with the "multi-ingredient and multi-target" theory of TCM. In general, based on the affirmative curative effects of TCM in treating a certain disease, network pharmacology was commonly used to predict the potential targets and underlying mechanisms of TCM interventions (Chen et al., 2021). In contrast, reverse network pharmacology is to use diverse public databases for initial disease target selection. Starting from the screened targets, effective medicine against this disease are subsequently explored (Dan et al., 2020; Dan et al., 2022). This specific target-driven assay allows us to identify medicine with a strong link to a disease. The strengths of target-driven approaches is that they are often simpler to execute than traditional phenotypic assays, operating with understanding of a drug's specific biological hypothesis from an earlier stage, thus identifying more highly selective medicine (Croston, 2017). In addition, target-driven approaches are often higher in throughput than phenotypic assays, facilitating faster large-scale screening.

The design of this study is shown in Figure 1. Firstly, the intersections of HBV and HCC targets were taken as HBV-related HCC targets. Next, starting from acquired targets, the corresponding chemical components and herbs were matched. A target-chemical component-herb network was established. The coverage of the herbs indicated in the guidelines was then used to evaluate the findings. Furthermore, functional analysis was performed for acquired HBV-related HCC targets. Molecular docking was applied to investigate the binding of obtained chemical components and targets. Finally, the frequency of herbs' properties, tastes, and meridian tropisms was assessed. Consequently, the aim of this research was to apply a target-driven reverse network pharmacology strategy to gain systematic insight into the therapeutic potential of TCM against HBV-related HCC.

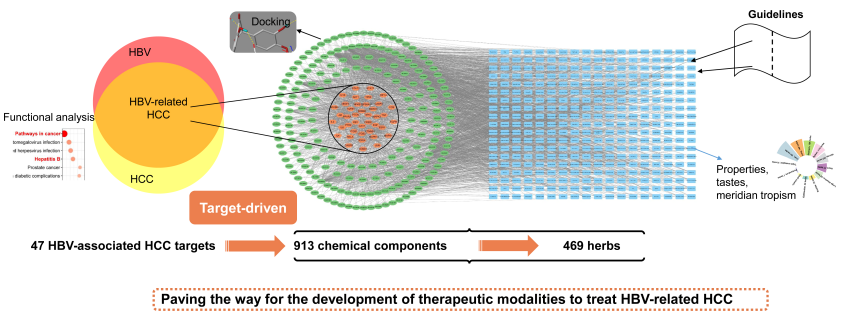


FIGURE 1
Workflow of this study.

Materials and methods

Searching for HBV-related HCC targets

To obtain HBV targets and HCC targets, the Genecards database (Safran et al., 2010), and Therapeutic target database (Wang et al., 2020) were queried. The information on these targets was standardized using the Uniprot database (The UniProt Consortium, 2020). PPI data were obtained from the STRING database (Szklarczyk et al., 2021), with the term "Homo sapiens" restricted. A confidence level of 0.9 was selected. The CytoNCA plug-in of Cytoscape was applied to calculate the topological parameters (Shannon et al., 2003).

Looking for relevant chemical components that act on HBV-related HCC targets

To find chemical components acting on relevant HBV-related HCC targets, the Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform (TCMSP) was searched (Ru et al., 2014). The screening conditions were as follows: oral bioavailability (OB) ≥ 30% and drug–likeness (DL) ≥ 0.18, half-life > 4 hours, molecular weight (MW) ≤ 500 Da, polar surface area (PSA) ≤ 140 Å2 and number of rotatable bonds (NBR) ≤ 10 (Lipinski et al., 2001; Huang et al., 2020). Given that the information provided by TCMSP was predicted by a computer, chemical components discarded in the initial screening were checked one by one to ensure that no relevant active chemical components were missed. These rechecked chemical components were selected as chemical components. The lists of the obtained chemical components and targets were introduced into Cytoscape to construct the target-chemical component network. A topological analysis was performed.

Searching for herbs containing obtained chemical components

The TCMSP was searched to find herbs containing obtained chemical components. Herbs containing obtained chemical components were collected and an herb-chemical component network was constructed. Cytoscape software was utilized to construct the network of target-chemical component-herb.

Enrichment analysis

Enriched gene ontology (GO) terms and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway of 47 targets were carried out by Metascape (Zhou et al., 2019). P value < 0.01 was considered to indicate significant enrichment.

Molecular docking

The crystal structures of the proteins were downloaded from the Protein Data Bank database (wwPDB Consortium, 2019). The 2D structures of chemical components were downloaded from the Pubchem database (https://pubchem.ncbi.nlm.nih.gov/). The PDB IDs for 46 targets are listed in

TABLE 1 The PDB IDs for 46 targets.

Gene symbol	PDB ID	Gene symbol	PDB ID	Gene symbol	PDB ID
AR	4OHA	JUN	6Y3V	CTNNB1	3FQN
ESR1	7BAA	CASP8	4JJ7	ERBB2	5MY6
RXRA	6LB4	FOS	1S9K	NFKB1	7LFC
MAPK14	3LFF	IL2	5LQB	PPARA	6KAX
RELA	6NV2	HIF1A	4H6J	TGFB1	6OM2
CASP3	4QUJ	EGFR	5UG9	EGF	1NQL
TNF	5UUI	STAT3	6NJS	ITGB3	3T3P
TP53	3D06	STAT1	3wwt	SMAD3	5OD6
AKT1	4GV1	CREB1	5ZK1	CDKN1B	6ATH
IL6	1ALU	RB1	2R7G	CXCL12	4UAI
VEGFA	4GLS	MYC	6G6K	EP300	3BIY
IL1B	5R8Q	PTEN	7PC7	ITGB1	4WK0
MAPK1	4ZZN	HSP90AA1	5J2X	KRAS	6P0Z
IL4	4YDY	MAPK3	4QTB	RHOA	6V6U
NFKBIA	6Y1J	MAPK8	2XRW	–	–
CCND1	2W96	PTK2	6YOJ	–	–

Table 1. The molecular docking software Sybyl X-2.0 (Tripos, St. Louis, USA) was used. The Surflex-Dock program was used for the docking calculations.

Herbal characteristics

A frequency analysis was done to draw the rules of herbs against HBV-related HCC. Various characteristics like channel tropism, flavor, and property were examined.

Results

Looking for and screening for HBV-related HCC targets

An illustration of the screening process for HBV-related HCC targets is shown in Figure 2. Firstly, 10116 HBV target information was collected by taking the union of the results from TTD and GeneCards databases. 17016 HCC targets were obtained in a similar manner. Secondly, targets with a relevance score ≥ 10 were selected for research. There were still 1644 HBV targets and 1921 HCC targets. Thirdly, after taking the intersection of 1644 HBV targets and 1921 HCC targets, 927 HBV-related HCC targets were derived. The protein-protein interactions (PPI) network was an important tool for learning about cellular regulation and function (Valgardson et al., 2019). Fourthly, the PPI network of 927 targets was constructed by STRING and visualized by Cytoscape (Shannon et al., 2003). The topological parameters, including betweenness centrality (BC), closeness centrality (CC) and degree centrality (DC), were calculated for further screening. Only the nodes with higher values of DC (above twofold the median degree of all nodes), BC and CC (above the median value of all nodes) were identified as hub nodes which played a pivotal role within biological networks (Yu et al., 2018). After calculation, the thresholds for primarily screening were DC ≥ 22, BC ≥ 394, CC ≥ 0.041736 and 204 targets were derived. The thresholds for re-screening were DC ≥ 36, BC ≥ 2544, CC ≥ 0.04245, and this reduced the targets number to 64. Lastly, any targets that could not map to active chemical components on TCMSP and HIT website were discarded. Of these 64 targets, 47 were selected as HBV-related HCC targets (Table 2). The PPI network of these 47 HBV-related HCC targets are depicted in dashed boxed section in Figure 2.

Screening for relevant chemical components and constructing the target-chemical component network

Forty-seven HBV-related HCC targets were mapped to 2013 chemical components. To make the present study more close to the real world, chemical components which met ADME and Lipinskir rules criteria (OB ≥ 30%, DL ≥ 0.18) were left. Other research-worthy chemical components were derived based on the official "Chinese Pharmacopoeia" (2020 edition). Eventually, 942 chemical components were identified. Later, a target-chemical component network was developed (Figure 3). The target-chemical component network included 960 nodes and 1854 edges (Zhu et al., 2019). Orange nodes represented the targets, while green nodes represented chemical components. The edges indicated the interaction between targets and chemical components. The larger the node size, the greater the degree of connectivity.

Screening for relevant herbs and constructing the target-chemical component-herb network

We not only focus on the chemical components, but also combined with the corresponding herbs. Among the above-

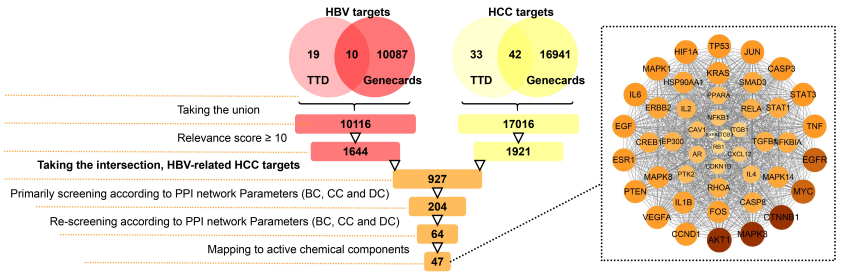


FIGURE 2
Procedure of searching and screening for HBV-associated HCC targets, which were derived by taking the intersection of HBV targets and HCC targets. Protein-protein interactions amongst the 47 targets were in the dashed box section.

TABLE 2 Forty-seven HBV-associated HCC targets [a].

Gene symbol	Uniprot ID	Protein name
AR	P10275	Androgen receptor
ESR1	P03372	Estrogen receptor
RXRA	P19793	Retinoic acid receptor RXR-alpha
MAPK14	Q16539	Mitogen-activated protein kinase 14
RELA	Q04206	Transcription factor p65
CASP3	P42574	Caspase-3
TNF	P01375	Tumor necrosis factor
TP53	P04637	Cellular tumor antigen p53
AKT1	P31749	RAC-alpha serine/threonine-protein kinase
IL6	P05231	Interleukin-6
VEGFA	P15692	Vascular endothelial growth factor A
IL1B	P01584	Interleukin-1 beta
MAPK1	P28482	Mitogen-activated protein kinase 1
IL4	P05112	Interleukin-4
NFKBIA	P25963	NF-kappa-B inhibitor alpha
CCND1	P24385	G1/S-specific cyclin-D1
JUN	P05412	Transcription factor AP-1
CASP8	Q14790	Caspase-8
FOS	P01100	Proto-oncogene c-Fos
IL2	P60568	Interleukin-2
HIF1A	Q16665	Hypoxia-inducible factor 1-alpha
EGFR	P00533	Epidermal growth factor receptor
STAT3	P40763	Signal transducer and activator of transcription 3
STAT1	P42224	Signal transducer and activator of transcription 1-alpha/beta
CREB1	P16220	Cyclic AMP-responsive element-binding protein 1
RB1	P06400	Retinoblastoma-associated protein
MYC	P01106	Myc proto-oncogene protein
PTEN	P60484	Phosphatidylinositol 3
CAV1	Q03135	Caveolin-1
HSP90AA1	P07900	Heat shock protein HSP 90-alpha
MAPK3	P27361	Mitogen-activated protein kinase 3
MAPK8	P45983	Mitogen-activated protein kinase 8
PTK2	Q05397	Focal adhesion kinase 1
CTNNB1	P35222	Catenin beta-1
ERBB2	P04626	Receptor tyrosine-protein kinase erbB-2
NFKB1	P19838	Nuclear factor NF-kappa-B p105 subunit
PPARA	Q07869	Peroxisome proliferator-activated receptor alpha
TGFB1	P01137	Transforming growth factor beta-1 proprotein
EGF	P01133	Pro-epidermal growth factor
ITGB3	P05106	Integrin beta-3
SMAD3	P84022	Mothers against decapentaplegic homolog 3
CDKN1B	P46527	Cyclin-dependent kinase inhibitor 1B
CXCL12	P48061	Stromal cell-derived factor 1
EP300	Q09472	Histone acetyltransferase p300
ITGB1	P05556	Integrin beta-1
KRAS	P01116	GTPase KRas
RHOA	P61586	Transforming protein RhoA

[a]The targets are sorted in decreasing order of degree value.

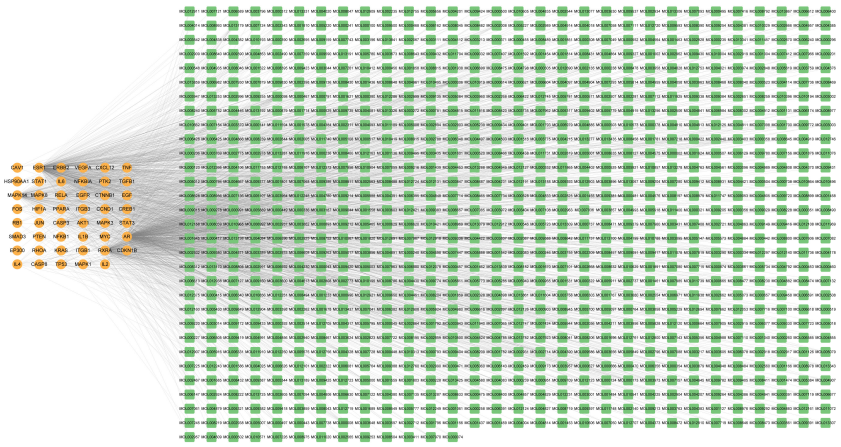


FIGURE 3
The 47 HBV-associated HCC targets-913 chemical components network. Orange nodes represented the targets, while green nodes represented chemical components. The edges indicated the interaction between targets and chemical components.

identified 942 chemical components (components in herbs), no corresponding herbs for 29 chemical components could be found. The remained 913 chemical components were provided in Supplementary Material and mapped to 469 herbs. A full list of 469 herbs was available from the Supplementary Material. That is to say, 469 herbs were associated with 47 targets by linking 913 chemical components. The network of 47 targets-913 chemical components-469 herbs was established by combining target-chemical component network and herb-chemical component network (Figure 4). To make the network more concise and more intuitive, chemical components and herbs with degree values less than six were hidden. The target-chemical component-herb network consisted of 1345 nodes and 4619 edges. Like Figure 3, orange and green nodes represented targets and chemical components. Blue nodes in Figure 4 represented mapped 469 herbs.

To assess the reliability of the final results, the recommended herbs in the guidelines for diagnosis and treatment of primary liver cancer in China (2020 Edition) and the clinical guidelines of diagnosis and treatment of chronic hepatitis B with traditional Chinese medicine (2018 Edition) were summarized. After removing duplicates, there were 112 herbs listed in these two guidelines. After comparison, 86 out of 112 herbs (almost 80%) were included in the 469 herbs, suggesting that the herbs that resulted from target-driven reverse network pharmacology were not divorced from clinical practice. Through design, new herbal formulae against HBV-related HCC could be developed based on the individual synergistic nature of each herb and the "Jun-

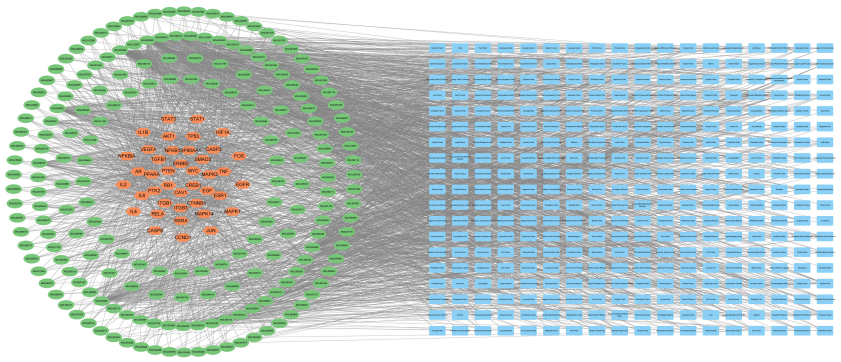


FIGURE 4
The 47 targets-913 chemical components-469 herbs network. Orange, green, and blue nodes represented the targets, chemical components, and herbs, respectively.

Chen-Zuo-Shi" (also known as "sovereign-minister-assistant-courier") rule.

The mechanisms of the 47 targets

To determine the potential mechanisms that the 47 HBV-related HCC targets were involved in, functional enrichment analysis was conducted. Pathway enrichment from the KEGG database showed that 174 pathways passed the filtering criteria, and the top 20 pathways according to p value were in Figure 5A. The most significantly enriched pathway was "pathways in cancer" with 40 targets clustered. This pathway was reported playing important tasks in HCC progression (Zhou et al., 2015). Three pathways including "hepatitis B" pathway were selected as the second group of most-significantly enriched pathways with quite similar p values and gene numbers. KEGG analyses highlighted the 47 targets that were not only topological connectors but also functional connectors in the crosstalk network of HBV and HCC. In addition, "human cytomegalovirus infection" and "kaposi sarcoma-associated herpesvirus infection" were also enriched in the second group. The human cytomegalovirus was a DNA virus that belonged to the herpes virus family. Therefore, the application of the matched chemical components and herbs based on these 47 targets may not be limited to HBV disease but also extend to human cytomegalovirus and kaposi sarcoma-associated herpesvirus infection.

The 47 targets were assigned to 1416 Gene Ontology (GO) terms, including 1280 biological process terms, 77 molecular function terms, and 59 cellular component terms. The top 20 GO terms ranked by p value are in Figure 5B. The types of the top 20 GO terms were biological processes and molecular functions instead of cellular components. The most significantly enriched terms were closely related to cell differentiation, proliferation, and migration processes, which were strongly associated with cancer occurrence, development, and metastasis (Sack et al., 2018). These enriched terms were also observed in previous HBV-related HCC literature (Wang et al., 2017; Fan et al., 2017; Zhang et al., 2021). Two molecular function terms most significantly enriched were associated with DNA-binding transcription factor binding, which was also closely linked to cancer (Bai et al., 2020).

Binding affinities between top 50 chemical components and targets

Molecular docking can foresee binding modes of small molecules with target proteins and predict molecular interactions (Saikia and Bordoloi, 2019). According to the rule of Sybyl, the total score shows the binding affinity between the small molecule and a potential target. A higher total score suggests a closer interaction between small molecule and potential target (Li et al., 2022; Zhang et al., 2019). A total score ≥ 4.0 indicates that the small molecule binds well to the targets (Guo et al., 2022). Considering that 3D structure for CAV1 was not available in the PDB database, docking studies were performed between 46 targets and the top 50 chemical components ranked by node degree value. As shown in Figure 6, 2300 scores were obtained, and the obtained total scores were presented on a color scale (red, higher than 4.0; blue, less than

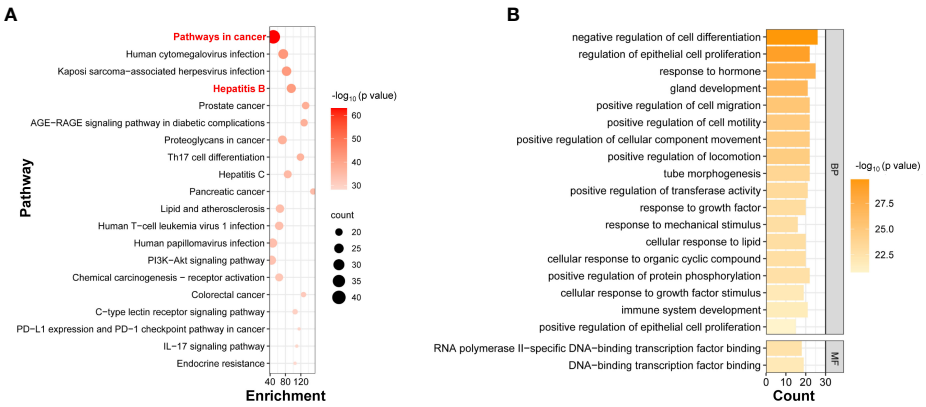


FIGURE 5
The functional enrichment analysis of 47 targets. **(A)** The top 20 KEGG pathways by p value. The bubble size indicated the number of targets clustered, and the color indicated the p value of the enrichment analysis. **(B)** The GO terms were sorted according to p value, with the most significantly enriched terms at the top. BP, biological processes. MF, molecular functions. The length of each bar indicated the number of targets clustered, and the color indicated the p value of the enrichment analysis.

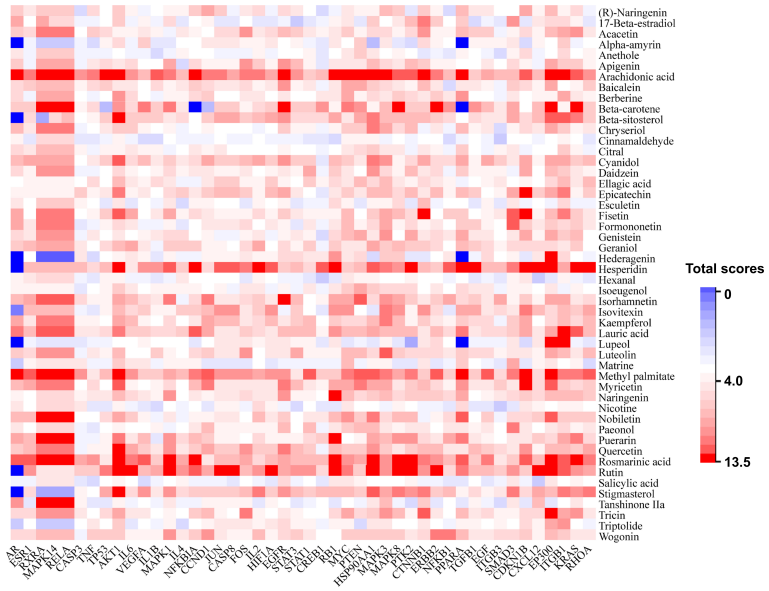


FIGURE 6
The binding scores between 46 targets (x axis) and the top 50 chemical components ranked by node degree value (y axis). Colors indicated different scores. (red, higher than 4.0; blue, less than 4.0).

4.0). The average score of 2300 sets of receptor-ligand docking results was 5.8, indicating good binding affinities between these 50 chemical components and the 46 targets in the target-chemical component network.

The top three binding score values were yielded by the binding of AKT1-hesperidin (score: 13.6), CXCL12-hesperidin (score: 12.4) and AKT1-rutin (score: 11.9). The corresponding hydrogen bonding plots are shown in Figure 7. It can be seen from Figure 7 that each ligand (chemical component) interacted with multiple residues of the target through at least one hydrogen bond. On the other hand, all these three combinations were not discovered in the target-chemical component network. That is, the potential chemical component-target combinations may be far more diverse than that included in the TCMSP database. There were still large amounts of interactions between active TCM chemical components and HBV-related HCC targets waiting to be further mined. The docking results could provide rapid and inexpensive technique support for future laboratory screening of related chemical component and herbs.

Properties, tastes, and meridian tropisms of herbs against HBV-related HCC

Figure 8 represented the frequency analysis results, which summarized the rules of 381 herbs that could be found in the Chinese Pharmacopoeia (2020 edition). In TCM, the properties of herbs mean their effects, which are classified into cold, cool, even, hot, and warm. The properties of Chinese medicinal herbs served as the foundation for herb analysis and clinical application (Huang et al., 2018). As shown in Figure 8A, the properties of 381 herbs were mainly warm (26.4%), followed by cold (23.6%). The total proportion of these two types of medicines accounted for more than 50% of all herbs. According to TCM theory, warm medicine can be warming and nourishing, thus reinforcing healthy Qi and helping to eliminate pathogenic factors (Liu et al., 2019). Cold medicine generally has a clearing action to get rid of the body of excess substances to regulate the balance of the body (Xia et al., 2020).

Herbs are also categorized according to their flavors, including bitter, pungent, salty, sour, and sweet. In terms of tastes, bitter accounted for the largest number (Figure 8B). The second was pungent. Bitter medicines have downward effects such as clearing away dampness and purging, whereas pungent medicines have outward and upward effects of dispersing (Xia et al., 2020). Using bitter medicines and pungent medicines in combination can dissolve stagnation of blood. This makes blood circulate smoothly, which can relieve pain.

Meridian tropisms are TCM organ systems. Namely, the target organs of herbs, like the heart, spleen, and liver (Xu et al., 2019). The theory of meridian tropism plays an important role in the clinical selection of TCM. Concerning meridian

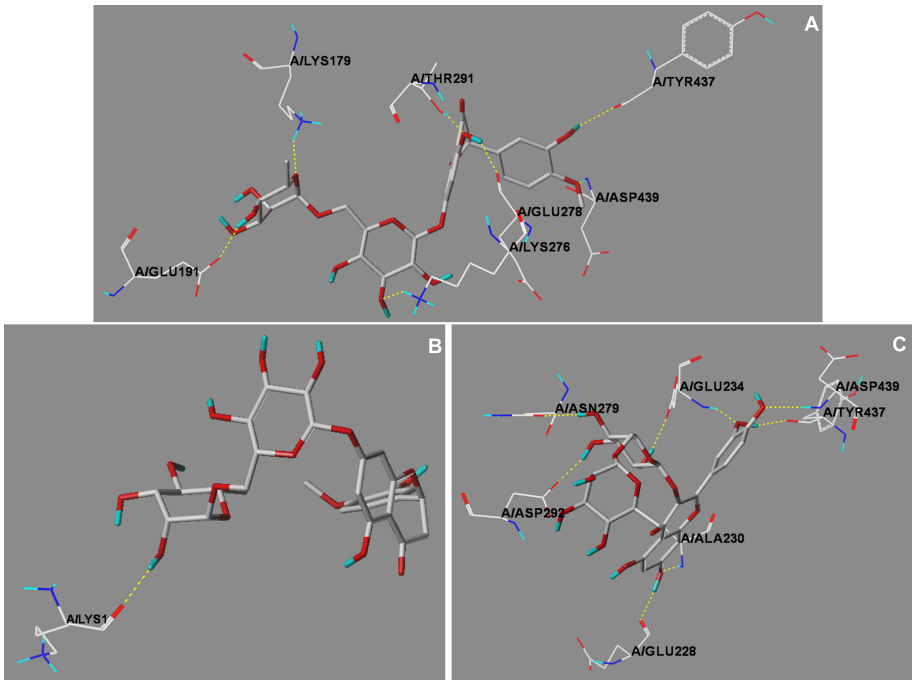


FIGURE 7
The hydrogen bonding plots of **(A)** AKT1-hesperidin, **(B)** CXCL12-hesperidi, and **(C)** AKT1-rutin. All ligands were depicted in a capped stick representation, while the interacting residues were shown as lines. The hydrogen bonds were yellow dashed lines.

tropisms, half of the herbs were liver meridian, which matched the disease region of HBV and HCC (Figure 8C). The second most matched meridian was the lung meridian. The liver governs regulating and ascending. The lung governs purification and descending. Both organs complement each other to harmonize qi, blood, and body fluid and to restore the non-pathological state (He et al., 2019).

Discussion

The diverse and growing knowledge of genomic information, multilevel biological interactions, and disease mechanisms facilitates the elucidation and discovery of new potential targets at which novel treatment development could be directed (Hoehe and Morris-Rosendahl, 2018). Target-driven reverse network pharmacology is one such strategy to use geneomewide target-ligand interaction networks in TCM to link genetic and drug data together (Dan et al., 2020).

Before the introduction of target-driven approaches, drug discovery was based primarily on phenotypic assays, often with limited information about the molecular mechanisms of disease. It is often necessary to characterize the mechanisms of active molecules identified in phenotypic screens to assist with the optimization of a candidate. Moreover, phenotypic assays exhibited lower throughput than target-driven approaches. In contrast, target-driven approaches allow significantly faster drug discovery and development than conventional phenotypic approaches.

Docking results suggested that the three strongest binding were AKT1-hesperidin, CXCL12-hesperidin and AKT1-rutin. Interestingly, these three sets of target-chemical component interaction were in line with previous experimental data. Hesperidin is a kind of citrus flavonoids and numerous studies have delineated anti-HBV and anti-HCC activities of hesperidin (Li et al., 2018; Parvez et al., 2019; Khuanphram et al., 2021). It has been demonstrated that AKT1 phosphorylation in RBL-2H3 cells and male rats can be suppressed by hesperidin (Kobayashi and Tanabe, 2006; Li et al., 2018). Hesperidin has been also reported attenuating the secretion of CXCL12 from A549 cells in a dose-dependent manner by ELISA method (Xia et al., 2018). Rutin, a flavonoid widely found in plants, exhibits anti-HCC activities in Wistar rats (Pandey et al., 2021). Rutin can regulate phosphorylated AKT1 expression in different tissues (Li et al., 2022; Liang et al., 2018; Fei et al., 2019). With the advent of increasing large-scale data acquisition, the application of network

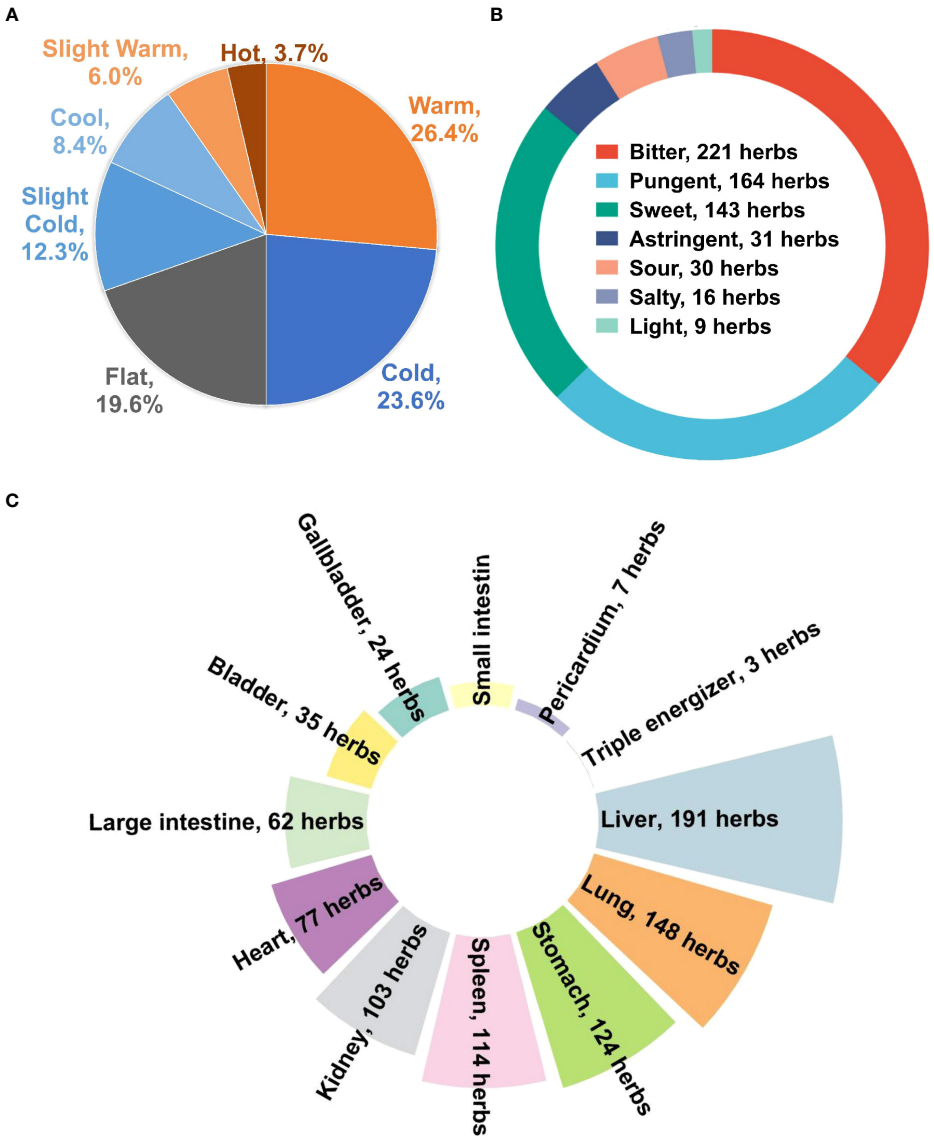


FIGURE 8
(A) Properties, **(B)** tastes, and **(C)** meridian tropism of herbs against HBV-associated HCC.

pharmacology in herbal formulas has provided a new horizon in the study of related domains such as active compound discovery, mechanism research, quality control, and others. Network pharmacology is expected to help traditional herbal medicine transition from experience-based medicine to evidence-based medicine (Li et al., 2022). However, it has inherent limitations, such as a lack of clinical data (Vogt and Mestres, 2019; Wu et al., 2021). In this respect, the major limitation of this work is a lack of *in vivo* and *in vitro* experiments. Experimental validations are needed to further verify our findings in later studies.

Conclusion

To our knowledge, this was the first attempt to systematically study HBV-related HCC treatment with TCM using target-driven reverse network pharmacology. A small library of 913 chemical components and 469 herbs against HBV-related HCC were acquired, with the hope of providing theoretical reference for more therapeutic options and may eventually benefiting clinical practice. Moreover, our studies promise to greatly expand the previous understanding of combined use of TCM-derived and western medicine.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

Conceptualization and writing, XY and YQ. Software, ZH, JL, and RD. Validation, YQ. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Shanxi Research Program of Application Foundation [No. 20210302123266]; Shanxi Medical University Second Hospital Doctoral Sustentation Fund [No. 201501-4].

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2022.964469/full#supplementary-material

References

Bai, K., He, S., Shu, L., Wang, W., Lin, S., Zhang, Q., et al. (2020). Identification of cancer stem cell characteristics in liver hepatocellular carcinoma by WGCNA analysis of transcriptome stemness index. *Cancer Med.* 9 (12), 4290–4298. doi: 10.1002/cam4.3047

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68 (6), 394–424. doi: 10.3322/caac.21492

Catherine de Martel, C., Maucort-Boulch, D., Plummer, M., and Franceschi, S. (2015). World-wide relative contribution of hepatitis b and c viruses in hepatocellular carcinoma. *Hepatology* 62 (4), 1190–1200. doi: 10.1002/hep.27969

Chen, J., Dou, P., Xiao, H., Dou, D., Han, X., and Kuang, H. (2021). Application of proteomics and metabonomics to reveal the molecular basis of atractylodis macrocephalae rhizome for ameliorating hypothyroidism instead of hyperthyroidism. *Front. Pharmacol.* 12, 664319. doi: 10.3389/fphar.2021.664319

Croston, G. E. (2017). The utility of target-based discovery. *Expert Opin. Drug Discovery* 12 (5), 427–429. doi: 10.1080/17460441.2017.1308351

Dan, W., Liu, J., Guo, X., Zhang, B., Qu, Y., and He, Q. (2020). Study on medication rules of traditional Chinese medicine against antineoplastic drug-induced cardiotoxicity based on network pharmacology and data mining. *Evid. Based Complement. Alternat. Med.* 2020, 7498525. doi: 10.1155/2020/7498525

Dan, W., Wu, C., and Xue, C. (2022). Rules of Chinese herbal intervention of radiation pneumonia based on network pharmacology and data mining. *Evidence-Based Complement Altern. Med: eCAM* 25), 7313864. doi: 10.1155/2022/7313864

European Association for the Study of the Liver (2017). EASL 2017 clinical practice guidelines on the management of hepatitis b virus infection. *J. Hepatol.* 67 (2), 370–398. doi: 10.1016/j.jhep.2017.03.021

Fan, H., Zhang, Q., Zhao, X., Lv, P., Liu, M., and Tang, H. (2017). Transcriptomic profiling of long non-coding RNAs in hepatitis b virus-related hepatocellular carcinoma. *Oncotarget* 8 (39), 65421–65434. doi: 10.18632/oncotarget.18897

Fei, J., Sun, Y., Duan, Y., Xia, J., Yu, S., Ouyang, P., et al. (2019). Low concentration of rutin treatment might alleviate the cardiotoxicity effect of pirarubicin on cardiomyocytes via activation of PI3K/AKT/mTOR signaling pathway. *Biosci. Rep.* 39 (6), BSR20190546. doi: 10.1042/BSR20190546

Guo, Y., Ning, B., Zhang, Q., Ma, J., Zhao, L., Lu, Q., et al. (2022). Identification of hub diagnostic biomarkers and candidate therapeutic drugs in heart failure. *Int. J. Gen. Med.* 15, 623–635. doi: 10.2147/IJGM.S349235

He, S., Guo, H., Zhao, T., Meng, Y., Chen, R., Ren, J., et al. (2019). A defined combination of four active principles from the danhong injection is necessary and sufficient to accelerate EPC-mediated vascular repair and local angiogenesis. *Front. Pharmacol.* 10, 1080. doi: 10.3389/fphar.2019.01080

Hoehe, M. R., and Morris-Rosendahl, D. J. (2018). The role of genetics and genomics in clinical psychiatry. *Dialogues Clin. Neurosci.* 20 (3), 169–177. doi: 10.31887/DCNS.2018.20.3/mhoehe

Huang, Z., Shi, X., Li, X., Zhang, L., Wu, P., Mao, J., et al. (2020). Network pharmacology approach to uncover the mechanism governing the effect of simiao powder on knee osteoarthritis. *BioMed. Res. Int.* 2020, 6971503. doi: 10.1155/2020/6971503

Huang, Y., Yao, P., Leung, K. W., Wang, H., Kong, X. P., Wang, L., et al. (2018). The yin-yang property of Chinese medicinal herbs relates to chemical composition but not anti-oxidative activity: An illustration using spleen-meridian herbs. *Front. Pharmacol.* 9, 1304. doi: 10.3389/fphar.2018.01304

Khuanphram, N., Taya, S., Kongtawelert, P., and Wongpoomchai, R. (2021). Sesame extract promotes chemopreventive effect of hesperidin on early phase of diethylnitrosamine-initiated hepatocarcinogenesis in rats. *Pharmaceutics* 13 (10), 1687. doi: 10.3390/pharmaceutics13101687

Kim, M., and Kim, Y. B. (2020). A network-based pharmacology study of active compounds and targets of fritillaria thunbergii against influenza. *Comput. Biol. Chem.* 89, 107375. doi: 10.1016/j.compbiolchem.2020.107375

Kobayashi, S., and Tanabe, S. (2006). Evaluation of the anti-allergic activity of citrus unshiu using rat basophilic leukemia RBL-2H3 cells as well as basophils of patients with seasonal allergic rhinitis to pollen. *Int. J. Mol. Med.* 17 (3), 511–515. doi: 10.3892/ijmm.17.3.511

Liang, W., Zhang, D., Kang, J., Meng, X., Yang, J., Yang, L., et al. (2018). Protective effects of rutin on liver injury in type 2 diabetic db/db mice. *Biomed. Pharmacother.* 107, 721–728. doi: 10.1016/j.biopha.2018.08.046

Li, M., Huang, W., Jie, F., and Wang, M. (2017). Discovery of Keap1-Nrf2 small-molecule inhibitors from phytochemicals based on molecular docking. *Food Chem. Toxicol.* 133, 110758. doi: 10.1016/j.fct.2019.110758

Li, M., Hu, S., Chen, X., Wang, R., and Bai, X. (2018). Research on major antitumor active components in zi-Cao-Cheng-Qi decoction based on hollow fiber cell fishing with high performance liquid chromatography. *J. Pharm. Biomed. Anal.* 149, 9–15. doi: 10.1016/j.jpba.2017.10.026

Li, X., Hu, X., Wang, J., Xu, W., Yi, C., Ma, R., et al. (2018). Inhibition of autophagy *via* activation of PI3K/Akt/mTOR pathway contributes to the protection of hesperidin against myocardial ischemia/reperfusion injury. *Int. J. Mol. Med.* 42 (4), 1917–1924. doi: 10.3892/ijmm.2018.3794

Li, Y., Li, Y., Qin, L., Ying, L., Li, Y., Qin, L., et al. (2021). Rutin prevents retinal ganglion cell death and exerts protective effects by regulating transforming growth

factor-β2/Smad2/3Akt/PTEN signaling in experimental rat glaucoma. *Trop. J. Pharm. Res.* 18 (5), 985–993. doi: 10.4314/tjpr.v18i5.11

Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* 46 (1–3), 3–26. doi: 10.1016/S0169-409X(00)00129-0

Li, Z., Qu, B., Wu, X., Chen, H., Wang, J., Zhou, L., et al. (2022). Methodology improvement for network pharmacology to correct the deviation of deduced medicinal constituents and mechanism: Xian-Ling-Gu-Bao as an example. *J. Ethnopharmacol.* 289, 115058. doi: 10.1016/j.jep.2022.115058

Liu, L., Yang, F., Jing, Y., and Xin, L. (2019). Data mining in xu runsan's traditional Chinese medicine practice: treatment of chronic pelvic pain caused by pelvic inflammatory disease. *J. Tradit. Chin. Med.* 39 (3), 440–450. doi: 10.19852/j.cnki.jtcm.2019.03.018

Pandey, P., Khan, F., Qari, H. A., and Oves, M. (2021). Rutin (Bioflavonoid) as cell signaling pathway modulator: Prospects in treatment and chemoprevention. *Pharm. (Basel)* 14 (11), 1069. doi: 10.3390/ph14111069

Parvez, M. K., Tabish Rehman, M., Alam, P., Al-Dosari, M. S., Alqasoumi, S. I., and Alajmi, M. F. (2019). Plant-derived antiviral drugs as novel hepatitis b virus inhibitors: Cell culture and molecular docking study. *Saudi Pharm. J.* 27 (3), 389–400. doi: 10.1016/j.jsps.2018.12.008

Qiu, J., Zhou, Q., Zhang, Y., Guan, M., Li, X., Zou, Y., et al. (2020). Discovery of novel quinazolinone derivatives as potential anti-HBV and anti-HCC agents. *Eur. J. Med. Chem.* 205, 112581. doi: 10.1016/j.ejmech.2020.112581

Ru, J., Li, P., Wang, J., Zhou, W., Li, B., Huang, C., et al. (2014). TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. *J. Cheminform* 6, 13. doi: 10.1186/1758-2946-6-13

Sack, L. M., Davoli, T., Li, M. Z., Li, Y., Xu, Q., Naxerova, K., et al. (2018). Profound tissue specificity in proliferation control underlies cancer drivers and aneuploidy patterns. *Cell* 173 (2), 499–514.e23. doi: 10.1016/j.cell.2018.02.037

Safran, M., Dalah, I., Alexander, J., Rosen, N., Iny Stein, T., Shmoish, M., et al. (2010). GeneCards version 3: the human gene integrator. *Database (Oxford)* 2010baq020. doi: 10.1093/database/baq020

Saikia, S., and Bordoloi, M. (2019). Molecular docking: Challenges, advances and its use in drug discovery perspective. *Curr. Drug Targets* 20 (5), 501–521. doi: 10.2174/1389450119666181022153016

Sarin, S. K., Kumar, M., Lau, G. K., Abbas, Z., Chan, H. L. Y., Chen, C. J., et al. (2016). Asian-Pacific clinical practice guidelines on the management of hepatitis b: a 2015 update. *Hepatol. Int.* 10 (1), 1–98. doi: 10.1007/s12072-015-9675-4

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. doi: 10.1101/gr.1239303

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49 (D1), D605–D612. doi: 10.1093/nar/gkab835

Terrault, N. A., Lok, A. S. F., McMahon, B. J., Chang, K. M., Hwang, J. P., Jonas, M. M., et al. (2018). Update on prevention, diagnosis, and treatment of chronic hepatitis b: AASLD 2018 hepatitis b guidance. *Hepatology* 67 (4), 1560–1599. doi: 10.1002/hep.29800

The UniProt Consortium. (2020). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* 49 (D1), D480–D489. doi: 10.1093/nar/gkaa1100

Valgardson, J., Cosbey, R., Houser, P., Rupp, M., Van Bronkhorst, R., Lee, M., et al. (2019). MotifAnalyzer-PDZ: A computational program to investigate the evolution of PDZ-binding target specificity. *Protein Sci.* 28 (12), 2127–2143. doi: 10.1002/pro.3741

Vogt, I., and Mestres, J. (2019). Information loss in network pharmacology. *Mol. Inform* 38 (7), e1900032. doi: 10.1002/minf.201900032

Wang, G., Dong, F., Xu, Z., Sharma, S., Hu, X., Chen, D., et al. (2017). MicroRNA profile in HBV-induced infection and hepatocellular carcinoma. *BMC Cancer* 17 (1), 805. doi: 10.1186/s12885-017-3816-1

Wang, G., Zhang, L., and Bonkovsky, H. L. (2012). Chinese Medicine for treatment of chronic hepatitis b. *Chin. J. Integr. Med.* 18 (4), 253–255. doi: 10.1007/s11655-012-1064-4

Wang, Y., Zhang, S., Li, F., Zhou, Y., Zhang, Y., Wang, Z., et al. (2020). Therapeutic target database 2020: enriched resource for facilitating research and early development of targeted therapeutics. *Nucleic Acids Res.* 48 (D1), D1031–D1041.

Wu, C., Huang, Z. H., Meng, Z. Q., Fan, X. T., Lu, S., Tan, Y. Y., et al. (2021). A network pharmacology approach to reveal the pharmacological targets and biological mechanism of compound kushen injection for treating pancreatic cancer based on WGCNA and *in vitro* experiment validation. *Chin. Med.* 16, 121. doi: 10.1186/s13020-021-00534-y

wwPDB Consortium. (2019). Protein data bank: The single global archive for 3D macromolecular structure data. *Nucleic Acids Res.* 47 (Database issue), D520–D528. doi: 10.1093/nar/gky949

Xia, P., Gao, K., Xie, J., Sun, W., Shi, M., Li, W., et al. (2020). Data mining-based analysis of Chinese medicinal herb formulae in chronic kidney disease treatment. *Evid. Based Complement. Alternat. Med.* 2020, 9719872. doi: 10.1155/2020/9719872

Xia, R., Xu, G., Huang, Y., Sheng, X., Xu, X., and Lu, H. (2018). Hesperidin suppresses the migration and invasion of non-small cell lung cancer cells by inhibiting the SDF-1/CXCR-4 pathway. *Life Sci.* 201, 111–120. doi: 10.1016/j.lfs.2018.03.046

Xu, H. Y., Zhang, Y. Q., Liu, Z. M., Chen, T., Lv, C. Y., Tang, S. H., et al. (2019). ETCM: an encyclopaedia of traditional Chinese medicine. *Nucleic Acids Res.* 47 (D1), D976–D982. doi: 10.1093/nar/gky987

Yu, G., Wang, W., Wang, X., Xu, M., Zhang, L., Ding, L., et al. (2018). Network pharmacology-based strategy to investigate pharmacological mechanisms of zuojinwan for treatment of gastritis. *BMC Complement Altern. Med.* 18 (1), 292. doi: 10.1186/s12906-018-2356-9

Zhang, J., Bi, R., Meng, Q., Wang, C., and Huo, X. (2019). Catalpol alleviates adriamycin-induced nephropathy by activating the SIRT1 signalling pathway *in vivo* and *in vitro*. *Br. J. Pharmacol.* 176 (23), 4558–4573. doi: 10.1111/bph.14822

Zhang, P., Feng, J., Wu, X., Chu, W., Zhang, Y., and Li, P. (2021). Bioinformatics analysis of candidate genes and pathways related to hepatocellular carcinoma in China: A study based on public databases. *Pathol. Oncol. Res.* 27, 588532. doi: 10.3389/pore.2021.588532

Zhang, X., Wang, M., and Zhou, S. (2020). Advances in clinical research on traditional Chinese medicine treatment of chronic fatigue syndrome. *Evid. Based Complement. Alternat. Med.* 2020, 4715679. doi: 10.1155/2020/4715679

Zhang, L., Yang, K., Wang, M., Zeng, L., Sun, E., Zhang, F., et al. (2021). Exploring the mechanism of cremastra appendiculata (SUANPANQI) against breast cancer by network pharmacology and molecular docking. *Comput. Biol. Chem.* 94, 107396. doi: 10.1016/j.compbiolchem.2020.107396

Zhou, X., Zheng, R., Zhang, H., and He, T. (2015). Pathway crosstalk analysis of microarray gene expression profile in human hepatocellular carcinoma. *Pathol. Oncol. Res.* 21 (3), 563–569. doi: 10.1007/s12253-014-9855-x

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10 (1), 1523. doi: 10.1038/s41467-019-09234-6

Zhu, N., Hou, J., Ma, G., and Liu, J. (2019). Network pharmacology identifies the mechanisms of action of shaoyao gancao decoction in the treatment of osteoarthritis. *Med. Sci. Monit.* 25, 6051–6073. doi: 10.12659/MSM.915821

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 20 September 2022
DOI 10.3389/fcimb.2022.1025900

OPEN ACCESS

EDITED BY
Ming Hu,
Qingdao University, China

REVIEWED BY
Lei Zhang,
Huazhong University of Science and Technology, China
Zhefu Li,
Qingdao Binhai University, China
Qibin Tang,
Sun Yat-sen University, China

*CORRESPONDENCE
Jun Ma
majun8536636@126.com;
chhopspmaj@163.com

SPECIALTY SECTION
This article was submitted to Virus and Host, a section of the journal Frontiers in Cellular and Infection Microbiology

RECEIVED 23 August 2022
ACCEPTED 02 September 2022
PUBLISHED 20 September 2022

CITATION
Li G, Wang Z, Chen D, Yin J, Mo Z, Sun B, Yang T, Zhang X, Zhai Z, Li Y, Chen P, Dai Y, Wang Z and Ma J (2022) Comprehensive analysis of a TPX2-related TRHDE-AS1/PKIA ceRNA network involving prognostic signatures in Hepatitis B virus-infected hepatocellular carcinoma. *Front. Cell. Infect. Microbiol.* 12:1025900. doi: 10.3389/fcimb.2022.1025900

Comprehensive analysis of a TPX2-related TRHDE-AS1/PKIA ceRNA network involving prognostic signatures in Hepatitis B virus-infected hepatocellular carcinoma

Gaopeng Li, Zhuangqiang Wang, Dong Chen, Jun Yin, Zhiyuan Mo, Bianyin Sun, Tao Yang, Xinning Zhang, Zhensheng Zhai, Yaoxuan Li, Pinggui Chen, Yunyan Dai, Zhiming Wang and Jun Ma*

General Surgery Department, Shanxi Bethune Hospital, Taiyuan, China

Hepatitis B virus (HBV) infection is a main carcinogenic factor of hepatocellular carcinoma (HCC). TPX2 microtubule nucleation factor is recently recommended as a novel prognostic biomarker in HBV-infected HCC tissues. This study aimed to explore a TPX2-related ceRNA regulatory network in HBV-infected HCC and the potential impact on HCC prognosis. We comprehensively identified 541 differential expressed lncRNAs (DElncRNAs), 37 DEmiRNAs and 439 DEmRNAs from HBV-related TCGA-HCC cohorts in $TPX2^{low}$ and $TPX2^{high}$ groups. Based on their RNA-RNA interaction and expression analysis, four DElncRNAs (TRHDE-AS1, DLX6-AS1, SNHG14, HOXA11-AS), four DEmiRNAs (miR-23b, miR-320a, miR-589, miR-126) and five DEmRNAs (PKIA, PCDHA2, SHCBP1, PRSS16, KIF18A) in HCC tumor vs normal groups were subjected to the hub regulatory networks analysis and further prognostic value analysis. Importantly, the TRHDE-AS1/miR-23b/PKIA ceRNA network was associated with HCC prognosis. Furthermore, cellular location analysis and base-base interaction analysis indicated that the cytoplasmic lncRNA TRHDE-AS1 was regarded as a ceRNA to sponging miR-23b and then regulating PKIA. Interestingly, correlation analysis suggested the expression correlation between TRHDE-AS1 and PKIA in HCC. Finally, we further performed the methylation and immune infiltration analysis to explore the functional process of PKIA in HCC. We proposed a ceRNA regulatory network may help elucidate the mechanism by which TPX2 contributes to the prognosis of HBV-related HCC.

KEYWORDS

hepatitis B virus, immune infiltration, ceRNA, prognosis, hepatocellular carcinoma

Introduction

Hepatocellular carcinoma (HCC) is a common primary liver cancer with high incidence and high mortality, accounting for 85% of liver cancer types (Sia et al., 2017). A 5-year survival rate of HCC patients is approximately 50%-60% (Llovet et al., 2003). Surgical partial hepatectomy is the first choice for treatment of resectable HCC cases (Lu and Dong, 2014). However, the 5-year recurrence rate of patients undergoing hepatectomy is still as high as 50-70% (Marasco et al., 2019). Hepatitis B virus (HBV), a member of the hepadnaviridae family, is a partially double-stranded DNA virus with a full length of 3200 bp (Tsukuda and Watashi, 2020). HBV infection is a major cause of serious liver-related diseases such as cirrhosis, liver failure and HCC (Seto et al., 2018). Elevated HBV DNA level is closely related to postoperative recurrence of HCC. A study indicated that antiviral therapy can reduce the activity of HBV and improve the overall survival rate of patients after radical resection of HCC (Wong et al., 2020). Hence, it is crucial to explore new biomarkers for the development and prognosis of HBV-related HCC.

Long non-coding RNAs (lncRNAs) participate in biological processes by controlling gene expression at the epigenetic, transcriptional and post-transcriptional levels, instead of translating into proteins (Mercer et al., 2009). Cytoplasmic lncRNAs can stabilize ribonucleoprotein complexes and act as competing endogenous RNAs (ceRNAs) for binding to miRNAs (Tay et al., 2014). Previous researches have reported the function of lncRNAs in different diseases, such as cancer, immune system diseases, and virus infection (Qi et al., 2015; Lu et al., 2020; Li et al., 2021). For example, lncRNA SNHG11/miR-184/AGO2 axis regulated cellular malignant processes, including proliferation, migration, and autophagy in HCC (Huang et al., 2020). Li et al. proposed that HULC/miR-200a-3p/ZEB1 network played an essential regulatory role in HCC epithelial-mesenchymal transition (EMT) process and tumor metastasis (Li et al., 2016). However, there is still no bioinformatics evidence on the regulation of ceRNA network in HBV-related HCC.

TPX2 microtubule nucleation factor is a microtubule-associated protein required for the formation of microtubules during cell mitosis (Neumayer et al., 2014). TPX2 recruits Aurora-A kinase to microtubules and drives the activation of Aurora-A kinase during mitosis (Asteriti et al., 2010). Accumulating analysis has shown that high expression of TPX2 is involved in worse prognosis of patients with HBV-related HCC (Zeng et al., 2020). In this research, we utilized the GEO RNA-seq data and the Cancer Genome Atlas (TCGA) HCC cohort to confirm the difference of TPX2 expression levels between HBV-related HCC tissues and normal samples, and screen differential expressed lncRNAs (DElncRNAs), DEmicroRNAs and DEmRNAs in $TPX2^{high}$ vs $TPX2^{low}$ expression groups (Figure 1). Next, by analyzing expression levels, overall survival, base-complementation of lncRNA-miRNA-mRNA and gene expression correlation, we comprehensively explored a potential TRHDE-AS1/PKIA ceRNA network by which TPX2 modulates in HBV-related HCC. Furthermore, we carried out the methylation and immune infiltration analysis to discuss the potential mechanism of PKIA regulation on HBV-HCC occurrence and prognosis. The findings about PTX2-related ceRNA patterns are helpful to discover the underlying mechanisms and to provide novel biomarkers for the prognosis of HBV-related HCC.

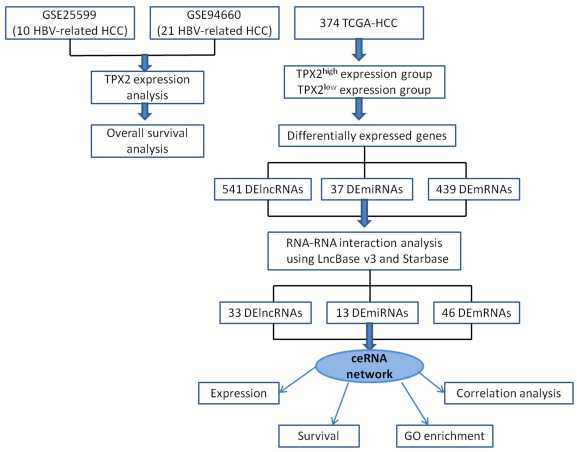


FIGURE 1
Flowchart of ceRNA network analysis in HBV-related HCC.

Methods

Data collection and gene expression analysis

The expression profiling by high throughput sequencing datasets (GSE25599 and GSE94660) were collected from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/gds/?term=). GSE25599 includes 10 match-paired HBV-related HCC and non-cancerous adjacent samples (Huang et al., 2011). GSE94660 includes 21 pairs of tumor and non-neoplastic liver tissues of HBV-HCC patients (Yoo et al., 2017). Raw RNA-seq data was obtained from TCGA-LIHC level 3 HTSeq-counts project (https://portal.gdc.cancer.gov/), which contains 50 cases with paracancerous tissues and 374 cases with HCC. The obtained counts data was converted into transcripts per million (TPM) format by using R software (version 3.6.3). TPX2 expression in pan-cancer was analyzed by using TCGA-All level 3 HTSeq-FPKM project and visualized by using ggplot2 (version 3.3.3). DElncRNAs, DEmiRNAs and DEmRNAs between $TPX2^{high}$-HCC samples and $TPX2^{low}$-HCC samples were identified with critical values of |LogFC| > 1.5, |

LogFC| > 0.5 and |LogFC| > 2, respectively (all adjusted p value < 0.05). Volcano plots and heatmap were visualized by ggplot2. The immunohistochemical (IHC) results of TPX2 protein expression was obtained from the Human Protein Atlas (https://www.proteinatlas.org/). The antibody HPA005487 against TPX2 protein was used for IHC assay.

Analysis of TPX2 and PKIA genetic alteration

cBioPortal online tool (https://www.cbioportal.org/) (Gao et al., 2013) was utilized to analyze TPX2 or PKIA mutations across Live Hepatocellular Carcinoma type. We selected "Mutations" and "Putative copy-number alterations from GISTIC" Genomic Profiles, and then entered gene "TPX2" and "PKIA". OncoPrint module showed a schematic diagram including the percentage of genetic alternation. Plots module presented a scatter diagram containing the correlation of mRNA vs Copy number alteration (CNA).

Analysis of overall survival and gene ontology (GO) enrichment

The survival R package (version 3.6.3) and survminer package (version 0.4.9) was downloaded to identify prognostic genes among DElncRNAs, DEmiRNAs and DEmRNAs. TCGA-HCC samples were divide into high and low expression groups. Survival package (version 3.2-10) was utilized for statistic analysis of overall survival. DEmRNAs were subjected to GO enrichment analysis using Database for annotation, visualization, and integrated discovery (DAVID, https://david.ncifcrf.gov/). The term counts and –log10pvalue for BP (Biological process), CC (Cellular component), and MF (Molecular function) were showed with box plot.

Analysis of lncRNA-miRNA-mRNA network interaction

Using DIANA-LncBase v3 (https://diana.e-ce.uth.gr/lncbasev3) online tool, we explored the RNAs-RNAs interaction between DElncRNAs and DEmiRNAs. In addition, the DEmRNA targeted by DEmiRNAs was predicted using Starbase (https://starbase.sysu.edu.cn/). The primary ceRNA network was visualized by Cytoscape. The hub genes with score > 2 was identified by Cytohubba plug-in module. To explore the role of lncRNA TRHDE-AS1 in HCC, we obtained the sequence of TRHDE-AS1 in LNCipedia (https://lncipedia.org/) and confirmed the subcellular location in LncLocator (http://www.csbio.sjtu.edu.cn/bioinf/lncLocator/). RNAhybrid (https://bibiserv.cebitec.uni-bielefeld.de/rnahybrid) was utilized to calculate a minimal free energy (mfe) and predict the binding sites between RNAs and RNAs.

Analysis of PKIA methylation

The expression levels DNA methyltransferases (DNMT1, DNMT3A, and DNMT3B) was investigated in $PKIA^{high}$ and $PKIA^{low}$-TCGA-HCC samples. UALCAN (http://ualcan.path.uab.edu/) online data analysis was performed to detect DNA methylation statue in HCC vs paracancerous samples. Furthermore, methylation sites in PKIA was analyzed using the MEXPRESS visualization tool (https://mexpress.be/).

Immune infiltration analysis

The correlation between GLA expression level and infiltration level in multiple immune cells (B cell, myeloid dendritic cell, macrophage, monocyte, neutrophil and T cell) was analyzed on the Tumor Immune Estimation Resource (TIMER, (http://cistrome.org/TIMER/). In addition, we explored the clinical relevance of tumor immune subsets with PKIA expression levels in HCC using immure outcome module. The correlation between PKIA level with multiple immune markers was explored in Gene_Corr module. The degree of correlation was calculated with purity-adjusted partial spearman's rho value.

Results

TPX2 is a differentially expressed gene (DEG) and prognostic biomarker in HBV-related HCC tissues

Previous literature identified TPX2 as a DEG in HBV-infected HCC specimens (Zeng et al., 2020). The raw data (GSE25599 and GSE94660) downloaded from SRA (SRP004768 and SRP099053) was subjected to analysis with limma R package. Firstly, we analyzed the expression level of TPX2 in HBV-related HCC and non-neoplastic liver tissues. As shown in Figures 2A, B, based on the GEO datasets, TPX2 shows a higher mRNA level in the HBV-HCC tissues than that in non-tumor groups ($p<0.001$). In addition, the Human Protein Atlas showed that TPX2 protein expression was significantly higher in HCC tissues than that in normal liver tissues (Figure 2C). Since TPX2 is upregulated in HCC tissues, we then screened the expression levels in TCGA pan-cancer RNA sequencing data. The upregulated TPX2 expression patterns in TCGA data are

shown in Figure 2D. We did obtain a significant difference of TPX2 level for HCC ($p<0.001$, Figure 2D). Next, we investigated the copy numbers of TPX2 in TCGA-HCC dataset using cBioPortal online tool to explore the mechanism of TPX2 upregulation in HCC. Oncoprint module shows that 2% HCC specimens (3/348) have putative copy-number amplification of TPX2 gene (Figure 2E). Consistently, low-level gene amplification event (gain copy-number) happens on 30% HCC samples (104/348) compared with TPX2 Diploid group (Figure 2F). Notably, TPX2 mRNA expression was significantly correlated with TPX2 copy number values (Spearman Correlation Coefficient = 0.31, $p<0.001$, Figure 2G). The 374 TCGA-HCC cases were divided into low TPX2 and high TPX2 expression groups and the relation of TPX2 expression with the overall survival (OS) was investigated. As shown in Figure 2H, HCC cases with higher TPX2 mRNA expression have shown a poor OS ($p<0.001$). The results indicated that TPX2 was upregulated in HBV-related HCC tissues and associated with poor prognosis of HCC.

Identification of DElncRNAs, DEmiRNAs and DEmRNAs in $TPX2^{high}$ and $TPX2^{low}$ groups

A posttranscriptional ceRNA regulatory network contributes to the occurrence and development of several cancers (Qi et al., 2015). Since TPX2 is a worse prognostic biomarker in HBV-related HCC, we aim to investigate the potential downstream network under TPX2 regulation. TCGA-HCC cohort was divided as two group ($TPX2^{high}$ and $TPX2^{low}$) according to the expression level of TPX2. The DElnRNAs between $TPX2^{low}$ and $TPX2^{high}$ samples were screened with the criteria of $|LogFC| > 1.5$ and adjusted p value < 0.05. We identified 541 DElncRNAs (439 upregulated and 102 down-regulated) between $TPX2^{low}$ and $TPX2^{high}$ groups in the TCGA-HCC dataset (Figure 3A). In addition, 37 DEmicroRNAs (27 up-regulated and 10 downregulated) were obtained using a criteria $|LogFC| > 0.5$ and adjusted $p < 0.05$ (Figure 3B). Applying a cut-off of $|LogFC| > 2$ and an adjusted p-value < 0.05 for the $TPX2^{high}$-HCC tissues compared with the $TPX2^{low}$-HCC samples, we identified 439 DEmRNAs (381 upregulated and 58 down-regulated) (Figure 3C). The heatmap of the top 20 DElncRNAs, DEmicroRNAs and DEmRNAs between $TPX2^{low}$ and $TPX2^{high}$ groups was showed in Figures 3D–F.

Prediction of lncRNA-miRNA-mRNA axis in TPX2-related HCC

To investigate the ceRNA network regulated by TPX2 in HBV-HCC, we utilized the DIANA-LncBase v3 (https://diana.e-ce.uth.gr/lncbasev3) online tool to identify miRNAs targeting the 541 DElncRNAs. The results predicted that 33 out of the 541 DElncRNAs interacted with 13 DEmicroRNAs. Next, potential mRNAs targeting the 13 DEmicroRNAs were predicated by using Starbase. Five DEmicroRNAs were predicated to target 46 DEmRNAs. In general, a total of 33 lncRNAs, 13 miRNAs and 46 mRNAs were subjected to Cytoscape software to visualize the lncRNA-miRNA-mRNA network (Figure 4B). The top 25 hub genes with nodes degree > 2 were identified and visualized by using CytoHubba module (Figure 4C). Finally, the hub regulatory network contained 9 lncRNAs (HOXA11-AS, SNHG14, H19, MEG3, PART1, TRHDE-AS1, LIN28B, MEG8, DLX6-AS1), 9 miRNAs (miR-589, miR-146b, miR-210, miR-4701, miR-320a, miR-429, miR-23b, miR-641, miR-126) and 6 mRNAs (KIF18A, SHCBP1, PCDHA2, PKIA, PCDHA3, PRSS16). Moreover, GO enrichment analysis was performed to investigate the functional annotation of DEGmRNAs and potential cellular function of ceRNA network. Figure 4A indicated that the DEmRNAs were significantly enriched in Biological Process terms, including "regulation of transcription from RNA polymerase II promoter", "cell differentiation", "cell proliferation".

Expression and prognostic value of the hub regulatory network in HCC

To further screen the candidate ceRNA network, we analyzed the expression of the top 25 hub targets in $TPX2^{high}$-HCC vs $TPX2^{low}$-HCC groups and HCC tumor vs normal groups. As shown in Figures 5A–C, the expression levels of six DElncRNAs (TRHDE-AS1, DLX6-AS1, HOXA11-AS, PART1, SNHG14, LIN28B), three DEmiRNAs (miR-146b, miR-320a, miR-589) and five DEmRNAs (KIF18A, SHCBP1, PCDHA2, PKIA, PRSS16) were significantly higher in TCGA-HCC tissues than in normal liver tissues ($p < 0.001$). In addition, the expression of H19, miR-23b, miR-126, miR-429 and miR-210 was downregulated in HCC tissues as compared to normal tissue ($p < 0.01$ and $p < 0.001$). Considering the results in $TPX2^{high}$-HCC vs $TPX2^{low}$-HCC groups, Table 1 elucidated a consistent finding with the expression analysis of four DElncRNAs (TRHDE-AS1, DLX6-AS1, SNHG14, HOXA11-AS), four DEmiRNAs (miR-23b, miR-320a, miR-589, miR-126) and five DEmRNAs (PKIA, PCDHA2, SHCBP1, PRSS16, KIF18A) in HCC tumor vs normal groups.

Then we investigated the correlation of above lncRNAs/miRNAs/mRNAs expression with the prognosis of patients with HCC, using the dataset of TCGA. As shown in Figure 6A, the Kaplan–Meier curves presented a correlation between high expression TRHDE-AS1 and worse overall survival ($p < 0.05$). However, low expression of miR-23b and miR-589 was

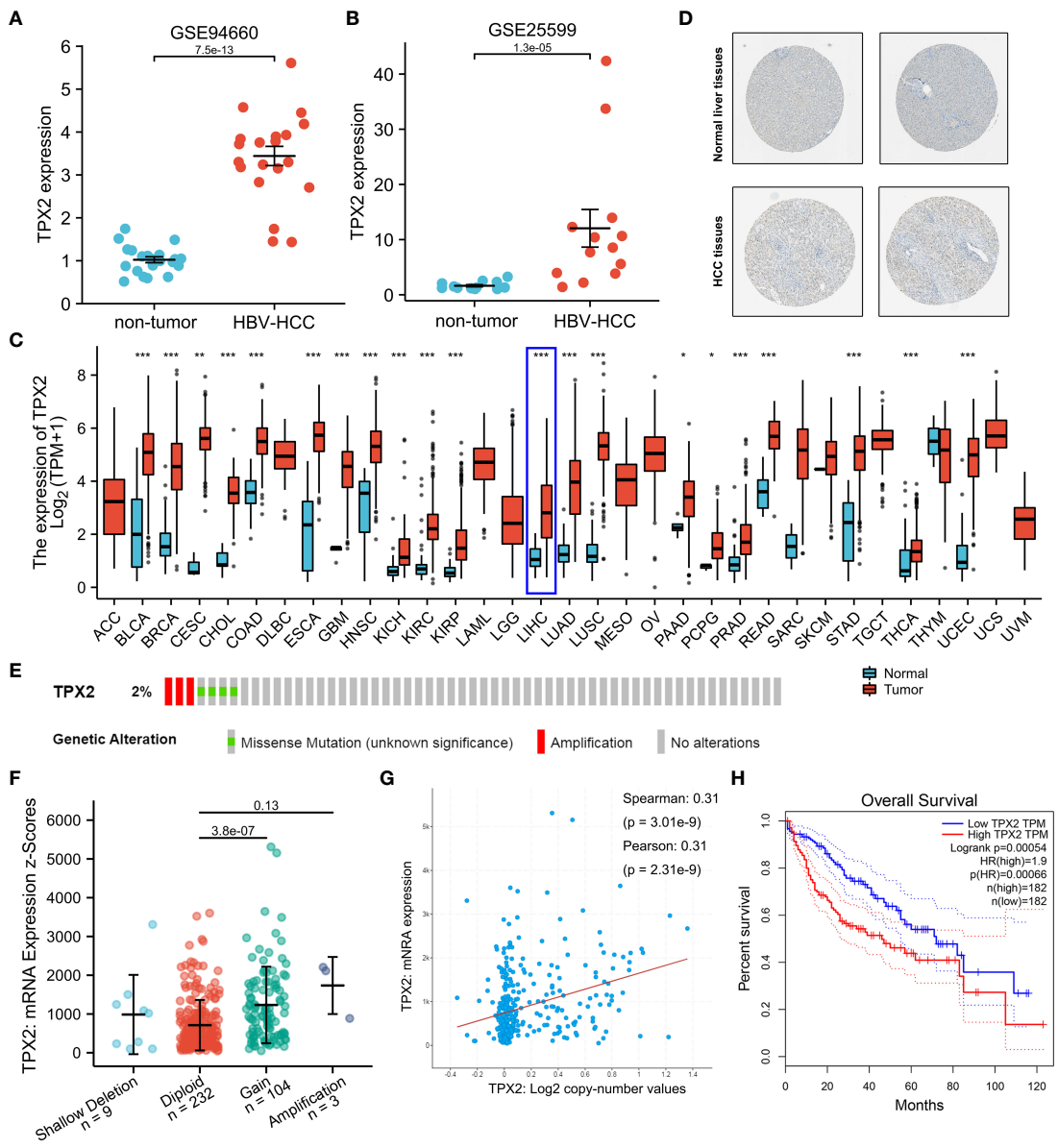


FIGURE 2
The role of TPX2 in HBV-related HCC. **(A, B)** GEO analysis (GSE25599, GSE94660) of the TPX2 gene expression in HBV-related HCC vs non-tumor groups. **(C)** TCGA analysis of the TPX2 gene expression levels in different tumor samples. **(D)** A validation of TPX2 protein expression in the Human Protein Atlas. **(E)** Heatmap downloaded from the cBioPortal online tool showing TPX2 genetic alternation in HCC samples. **(F, G)** cBioPortal plot analysis of TPX2 mRNA level vs. log2 copy-number value of the TCGA-HCC cohort. **(H)** GEPIA2 analysis of the TCGA-HCC cohort was performed for overall survival between PKIA high expression (50%) group and PKIA low expression (50%) group. *P < 0.05, **P < 0.01, ***P < 0.001.

associated with poor prognosis ($p < 0.05$, Figure 6B). Additionally, the high expression of PKIA, SHCBP1, PRSS16 and KIF18A worsened the overall survival in HCC ($p < 0.05$, $p < 0.001$, Figure 6C). In brief, one DElncRNA (TRHDE-AS1), two DEmiRNAs (miR-23b and miR-589), and four DEmRNAs (PKIA, SHCBP1, PRSS16 and KIF18A) were found to be prognostic biomarkers in HCC.

Validation of TRHDE-AS1//PKIA network in HCC

To determine a specific ceRNA network, the candidate lncRNA TRHDE-AS1 was subjected to a further analysis. Firstly, a lncRNA subcellular localization predictor (LncLocator online tool) was utilized to predict the location of

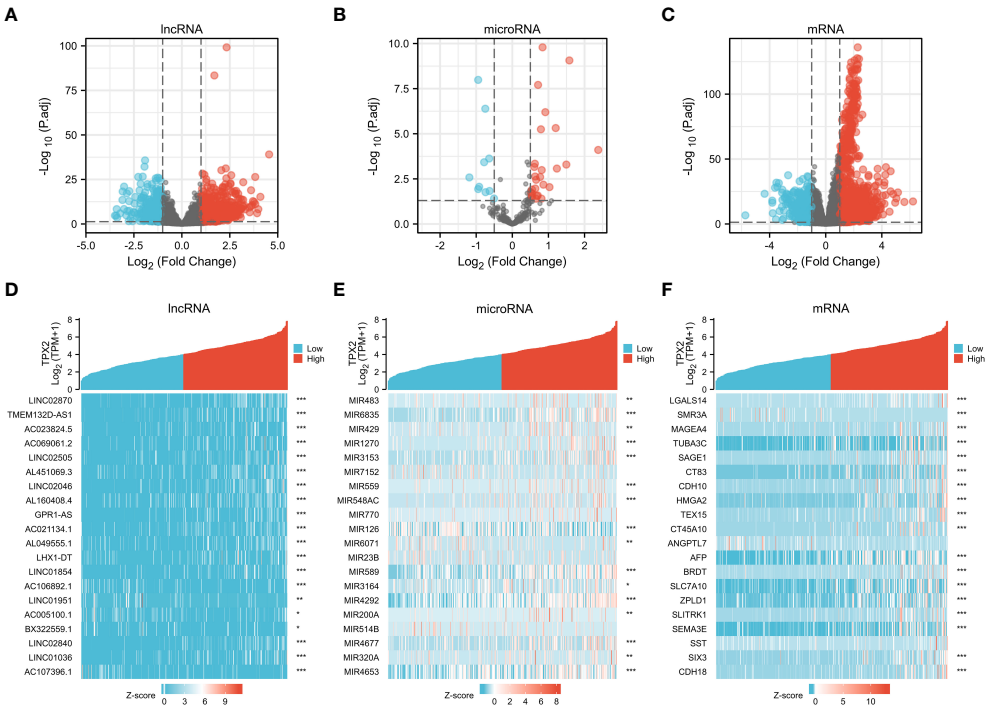


FIGURE 3
Identification of DElncRNAs, DEmiRNAs and DEmRNAs in TPX2^{high} vs TPX2^{low} groups. **(A–C)** Volcano plot of DElncRNA, DEmiRNAs and DEmRNA in TCGA-TPX2high vs TCGA-TPX2low groups. **(D–F)** Heatmap of the top 20 DElncRNAs, DEmiRNAs and DEmRNAs in TCGA-TPX2^{high} vs TCGA-TPX2^{low} samples. *P < 0.05, **P < 0.01, ***P < 0.001.

TRHDE-AS1 in human cells. As shown in Figure 7A, lncRNA TRHDE-AS1 accumulates in the cytoplasm and may act as a ceRNA to sponge miRNA. Next, RNAhybrid was used to predict the binding sites between TRHDE-AS1 and miRNA. A putative miR-23b binding site was identified in the regions of TRHDE-AS1 with an minimum free energy (mfe) -25.9 kcal/mol (Figure 7B). However, there is not a binding site between TRHDE-AS1 and miR-589. RNAhybrid algorithm predicted the interaction between miR-23b and PKIA, not other three prognostic DEmRNAs (Figure 7C). Schematic diagram showed the predicted binding sites between miR-23b and TRHDE-AS1/PKIA (Figures 7D, E). Correlation analysis showed that TRHDE-AS1 expression is positively correlated with PKIA in TCGA-HCC cohort ($r = 0.249$, $p < 0.001$). However, there was not significance between miR-23b expression and PKIA levels. These results indicated that TRHDE-AS1/PKIA axis is a potential ceRNA network in HCC.

Validation of PKIA methylation in HCC

Next, we further explored the potential mechanism regulating PKIA expression in HCC. As shown in Figure 8A, the genetic alteration status of PKIA did not appear in HCC samples of the TCGA cohorts. Moreover, we observed that there is not a potential association between copy-number of PKIA and expression level of PKIA mRNA (Figure 8B). Thus we hypothesized that the abnormal expression of PKIA might be related to DNA methylation in HCC. UALCAN online data analysis platform (http://ualcan.path.uab.edu/) evaluated epigenetic regulation of PKIA expression by promoter methylation. The results showed a methylation difference between primary tumor and normal tissues (Figure 8C). In addition, the data analysis in PKIA^{high} and PKIA^{low} groups showed that the levels of methylation-related genes (DNMT1, DNMT3A, DNMT3B) were downregulated in PKIA^{high}-TCGA samples, compared with PKIA^{low}-TCGA group ($p < 0.001$, Table 2). Furthermore, the MEXPRESS visualization tool (https://mexpress.be/) was utilized to investigate the potential PKIA DNA methylation. A negative correlation of PKIA level and DNA methylation at the promoter region (probe ID: cg09043127) was shown in Figure 8D ($p < 0.01$). These results suggested that the DNA methylation of PKIA is responsible for PKIA upregulation, which provided a potential mechanism to understand the ceRNA network regulation in HCC.

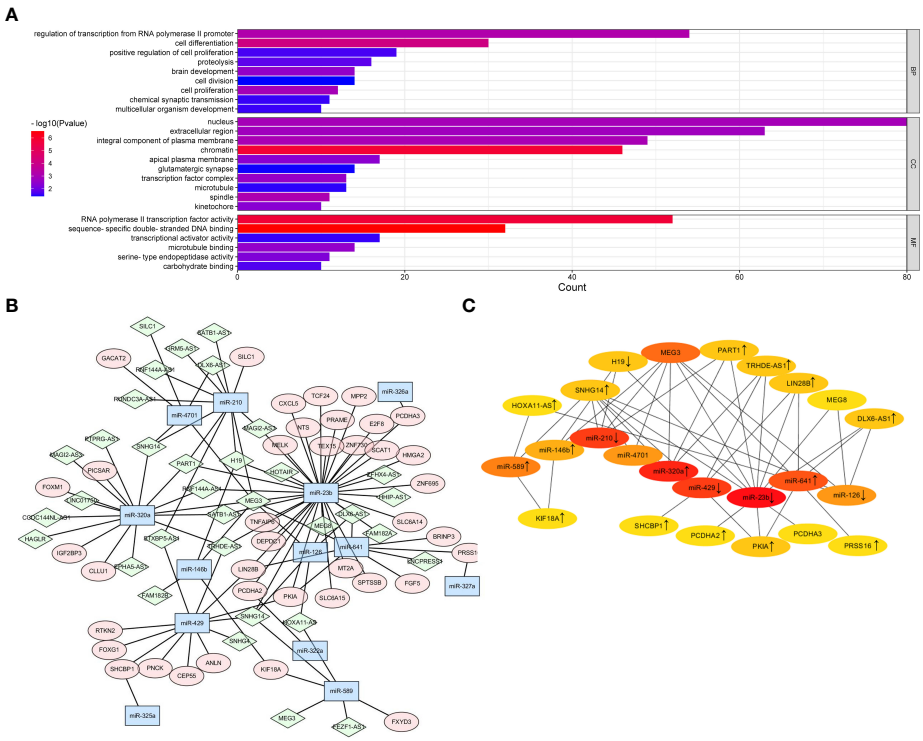


FIGURE 4
Prediction of lncRNA-miRNA-mRNA network in TPX2-related HCC. **(A)** Construction of ceRNA network for DElncRNAs, DEmiRNAs and DEmRNAs was visualized by Cytoscape. Green diamond indicates lncRNAs. Blue hexagon indicates miRNAs. Pink ellipse indicates mRNAs. **(B)** A CytoHubba module constructed a ceRNA network containing 25 hub genes. **(C)** The GO enrichment plot of DEmRNAs in the TCGA-TPX2^{high} vs TCGA-TPX2^{low} samples.

Analysis of relationship between PKIA and immune infiltration in HCC

As is well-known that immune cell infiltrating is an essential factor for cancer patient' survival. The above analysis indicated that TRHDE-AS1/PKIA axis can be important prognostic role in HCC. We then explored the effect of PKIA expression on immune infiltration in HCC. The results from TIMER showed that PKIA level had a obvious negative correlation with tumor purity in HCC (Figure 9A). After adjusting with tumor purity, PKIA expression levels were significantly correlated with the infiltration levels of B cells (cor = 0.424, p = 2.02e-16), CD8+ T cells (cor = 0.445, p = 4.65e-18), CD4+ T cells (cor = 0.499, p = 4.53e-23), macrophages (cor = 0.565, p = 3.60e-30), neutrophils (cor = 0.419, p = 4.05e-16), and dendritic cells (cor = 0.555, p = 6.59e-29). In addition, PKIA gene expression was significantly correlated with the infiltration levels of 49 immune cell markers in HCC (Table 3). To explore the clinical relevance of several tumor immune subsets in a Cox proportional hazard model, TIMER algorithm was utilized for estimating the abundances of immune infiltrates and survival in HCC. As shown in Figure 9B, the high abundances of CD4+ T cell ($p < 0.05$), macrophage ($p < 0.01$), neutrophil ($p < 0.01$). The results indicated that immune infiltrates in HCC was associated with worse prognosis.

Discussion

Emerging publications have reported that TPX2 overexpression is closely related to the development of various malignant tumors, including cervical cancer, esophageal squamous cell carcinoma and pancreatic cancer (Chang et al., 2012; Hsu et al., 2014; Gomes-Filho et al., 2020). In this study, our GEO and TCGA analysis results also confirmed the overexpression of TPX2 in HBV-related HCC tissues and an amplification of TPX2 DNA copy-number alterations (CNA) in HCC, suggesting that CNA might contribute to upregulation of TPX2 among HCC samples. Interestingly, the clinical RNA-seq-based evidence supports the role of TPX2 expression in prognosis of HCC. However, the regulatory mechanism which TPX2 plays a role in tumorigenesis of HBV-HCC remains to be

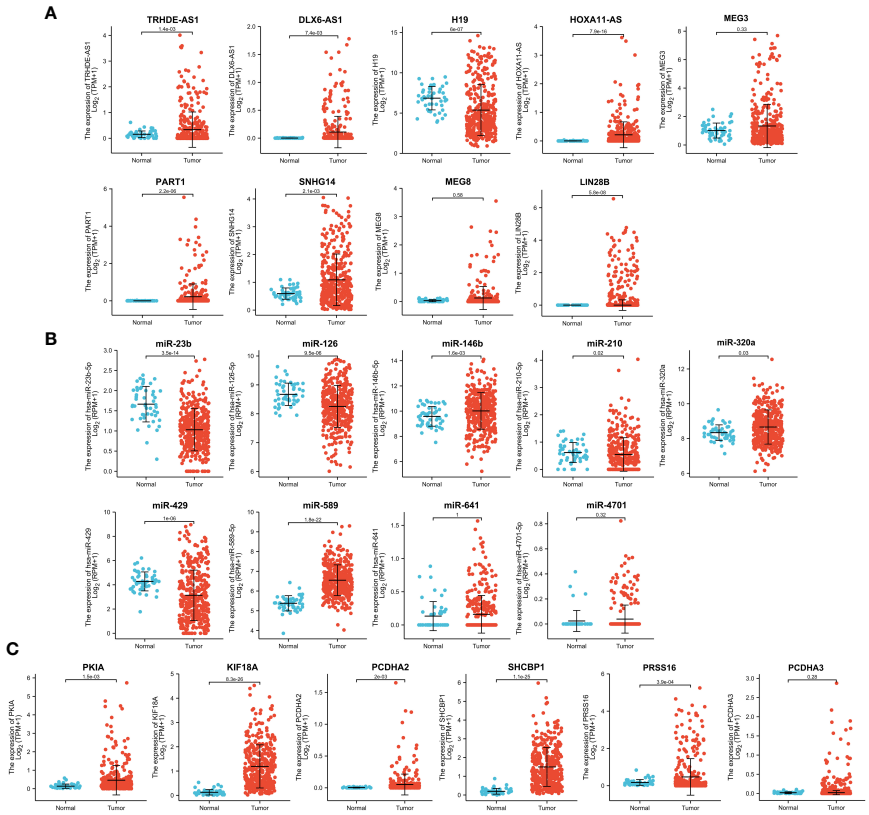


FIGURE 5
Expression analysis of 25 hub genes in TCGA-HCC. Analysis of TCGA-HCC cohorts showed the expression levels of **(A)** DElncRNAs, **(B)** DEmiRNAs and **(C)** DEmRNAs in normal and tumor groups.

TABLE 1 Hub gene expression in TPX2^{low} and TPX2^{high} groups.

gene_name	gene_id	gene_type	log2FC	padj	up/down
MEG3	ENSG00000214548	lncRNA	2.266005503	4.85E-23	up
TRHDE-AS1	ENSG00000236333	lncRNA	1.463752992	0.000743	up
DLX6-AS1	ENSG00000231764	lncRNA	1.61313044	0.000124	up
H19	ENSG00000130600	lncRNA	1.654348779	5.52E-08	up
PART1	ENSG00000152931	lncRNA	-1.178797661	0.011164	down
SNHG14	ENSG00000224078	lncRNA	1.158558875	2.11E-12	up
MEG8	ENSG00000225746	lncRNA	2.576108478	6.94E-13	up
HOXA11-AS	ENSG00000240990	lncRNA	1.334528937	3.4E-06	up
miR-23b	ENSG00000207563	miRNA	-0.91784113	0.008692	down
miR-320a	ENSG00000208037	miRNA	0.807026004	0.039311	up
miR-429	ENSG00000198976	miRNA	1.495517577	0.00051	up
miR-210	ENSG00000199038	miRNA	0.77019038	0.002489	up
miR-641	ENSG00000207631	miRNA	0.680429209	0.02815	up
miR-589	ENSG00000207973	miRNA	0.91402209	6.36E-07	up
miR-4701	ENSG00000264201	miRNA	0.542796982	0.033826	up
miR-126	ENSG00000199161	miRNA	-0.945521126	1.04E-08	down
miR-146b	ENSG00000202569	miRNA	-0.741363943	0.017269	down

(Continued)

TABLE 1 Continued

gene_name	gene_id	gene_type	log2FC	padj	up/down
PKIA	ENSG00000171033	mRNA	2.03482396	2.58E-20	up
PCDHA2	ENSG00000204969	mRNA	2.20983771	3.74E-10	up
SHCBP1	ENSG00000171241	mRNA	2.068436367	1.53E-77	up
PRSS16	ENSG00000112812	mRNA	2.058319355	5.38E-09	up
KIF18A	ENSG00000121621	mRNA	2.231162037	2E-123	up

answered. We comprehensively examined DElncRNAs, DEmiRNAs and DEmRNAs based on the expression level of TPX2 in HBV-HCC to explore a potential ceRNA network.

The function of lncRNAs in the occurrence and development of HCC is closely related to its structural characteristics and its role in the biological process (Zhang et al., 2020). LncRNAs participate in majority processes of gene regulation, including genome transcription, mRNA splicing and RNA attenuation (Peng et al., 2017; Gil and Ulitsky, 2020). In the study, TCGA-HCC database was utilized to identify DElncRNA, DEmiRNAs and DEmRNAs between $TPX2^{high}$ vs $TPX2^{low}$ groups, so as to screen a TPX2-related ceRNA network. Based on the RNA-RNA interaction, the DEmiRNAs obtained from DIANA-LncBase v3 and Starbase

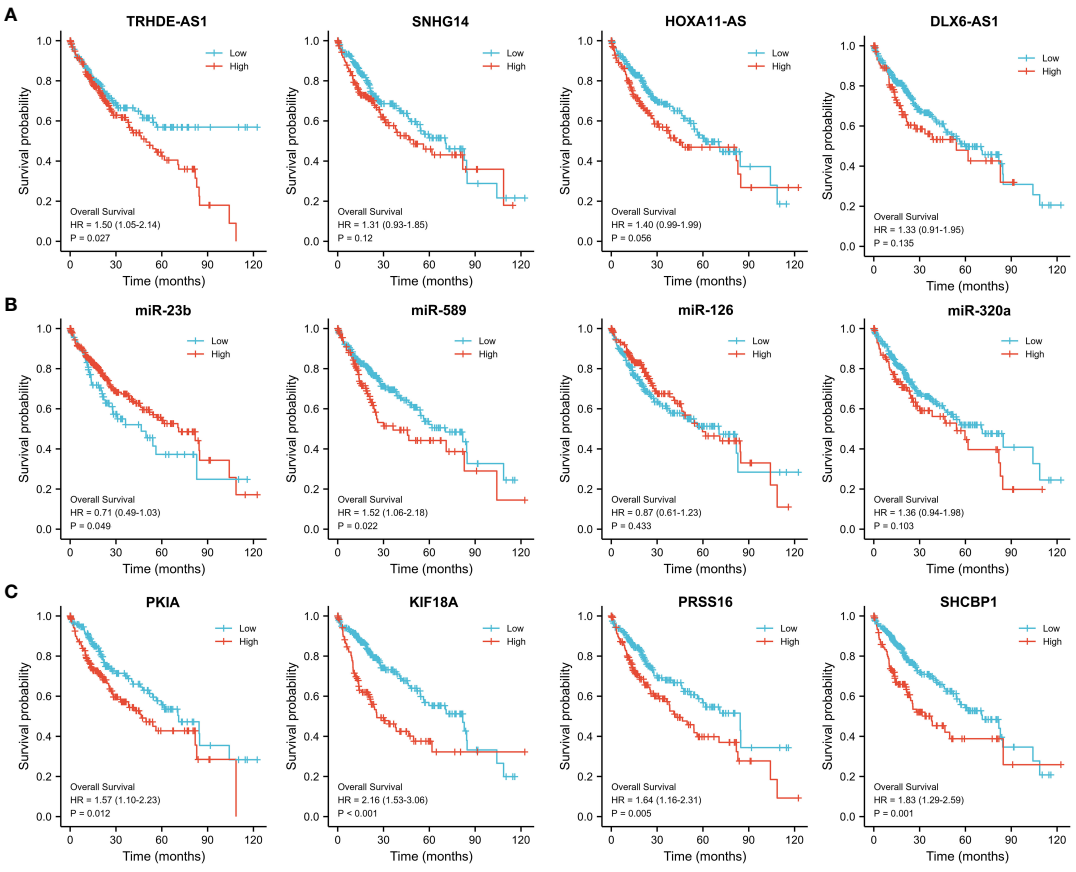


FIGURE 6
Prognostic value of the hub regulatory network in TCGA-HCC. Overall survival analysis of **(A)** DElncRNAs, **(B)** DEmiRNAs and **(C)** DEmRNAs using Kaplan–Meier survival curves.

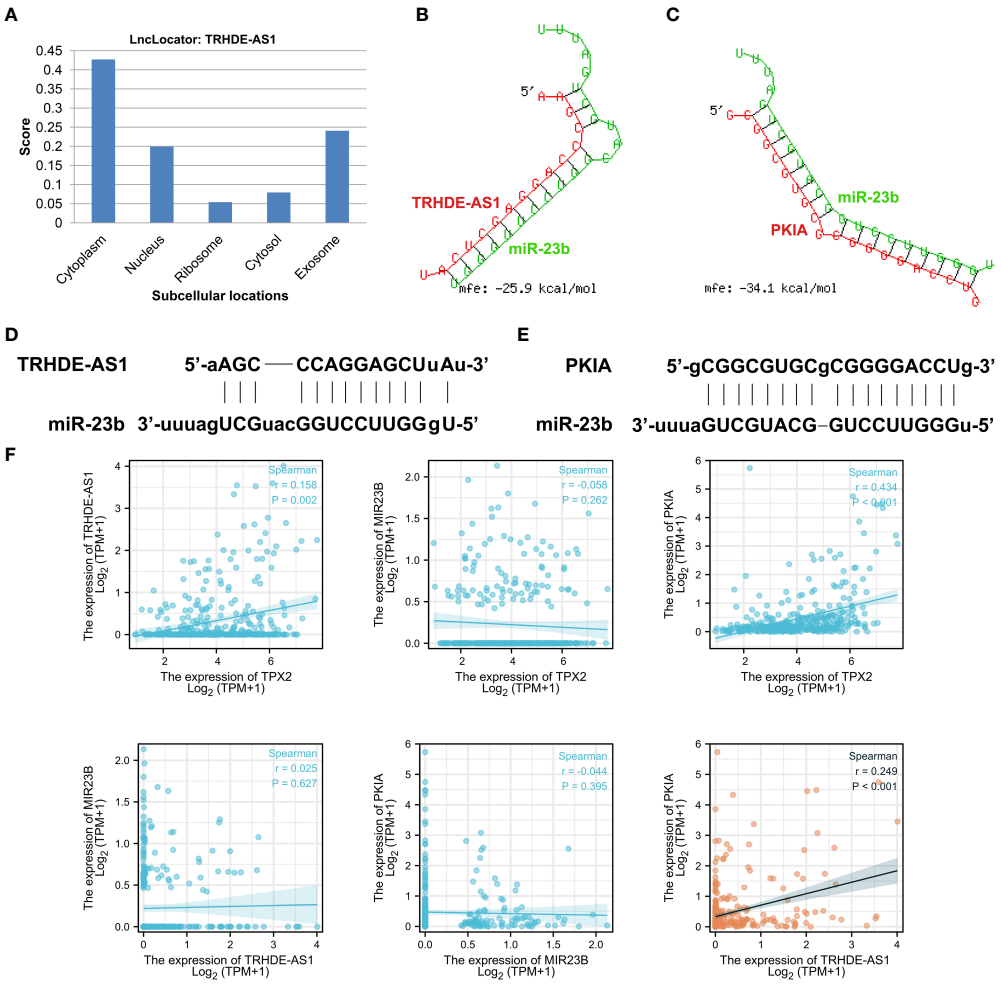


FIGURE 7
Validation of TRHDE-AS1/miR-23b/PKIA network in HCC. **(A)** Subcellular locations analysis of lncRNA TRHDE-AS1 using LncLocator. **(B, C)** RNAhybrid was utilized to predict the binding between TRHDE-AS1/PKIA and miR-23b. **(D, E)** Schematic diagram showing the predicted binding sites between miR-23b and TRHDE-AS1/PKIA. **(F)** Correlation analysis between RNAs and RNAs or between TPX2 and RNA.

tools are intersected to obtain the granularity common miRNAs. We constructed a primary lncRNA-miRNA-mRNA network which contains a total of 33 lncRNAs, 13 miRNAs and 46 mRNAs. Considering the results from Cytohubba hub genes, common DEGs in tumor vs normal groups and DEGs in $TPX2^{high}$ vs $TPX2^{low}$ groups was subjected to prognosis analysis in HCC. The common DEGs included four DElncRNAs (TRHDE-AS1, DLX6-AS1, SNHG14, HOXA11-AS), four DEmiRNAs (miR-23b, miR-320a, miR-589, miR-126) and five DEmRNAs (PKIA, PCDHA2, SHCBP1, PRSS16, KIF18A. Several reports have indicated the roles of these DElncRNAs in tumor pathogenesis. LncRNA DLX6-AS1 acting as a miR-513c sponge was found to promote the tumor malignancy progression of HCC through modulating Cul4A/ANXA10 (Liu et al., 2021). LncRNA SNHG14 regulated PTEN through activating PABPC1, which can aggravate cell proliferation and angiogenesis of HCC cells (Zhang et al., 2020). LncRNA HOXA11-AS promoted the HCC cell proliferation and EMT process with a mechanism in which HOXA11-AS may act as a ceRNA by directly sponging miR-506-3p8 to regulate the Slug expression (Liu et al., 2020). In this comprehensive research, we first presented evidence of TRHDE-AS1 upregulation in HCC and the potential correlation between TRHDE-AS1 expression and prognosis across TCGA-HCC cohort and TPX2-related HCC.

LncRNAs can act as miRNA sponges, recognizing cytosolic miRNAs and thereby regulating the expression of miRNA targeting gene (Paraskevopoulou and Hatzigeorgiou, 2016). In addition to the function as ceRNAs, lncRNAs also can compete with miRNAs for binding to target mRNAs and control a variety

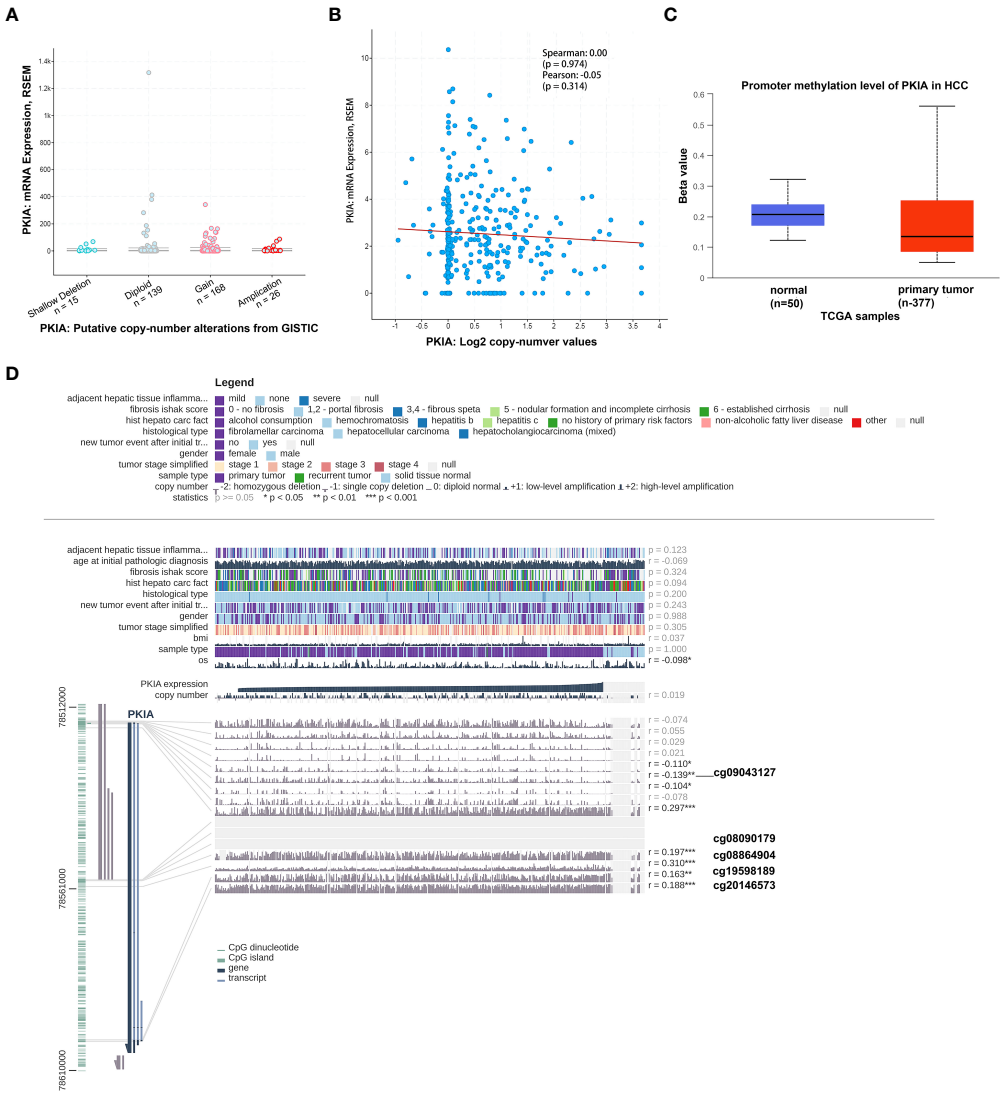


FIGURE 8
Validation of PKIA methylation in HCC. **(A, B)** cBioPortal plot analysis of PKIA mRNA level vs. log2 copy-number value of the TCGA-HCC cohort. **(C)** UALCAN online data analysis platform analyzing the methylation on PKIA promoter. **(D)** MEXPRESS visualization tool showing the potential PKIA DNA methylation sites.

TABLE 2 Expression level of methylation markers in PKIA^{low} and PKIA^{high} groups.

gene_name	gene_id	log2FC	padj	significance
DNMT1	ENSG00000130816	-0.572	7.07E-11	***
DNMT3A	ENSG00000119772	-0.417	3.6E-06	***
DNMT3B	ENSG00000088305	-0.737	1.13E-10	***

***p < 0.001.

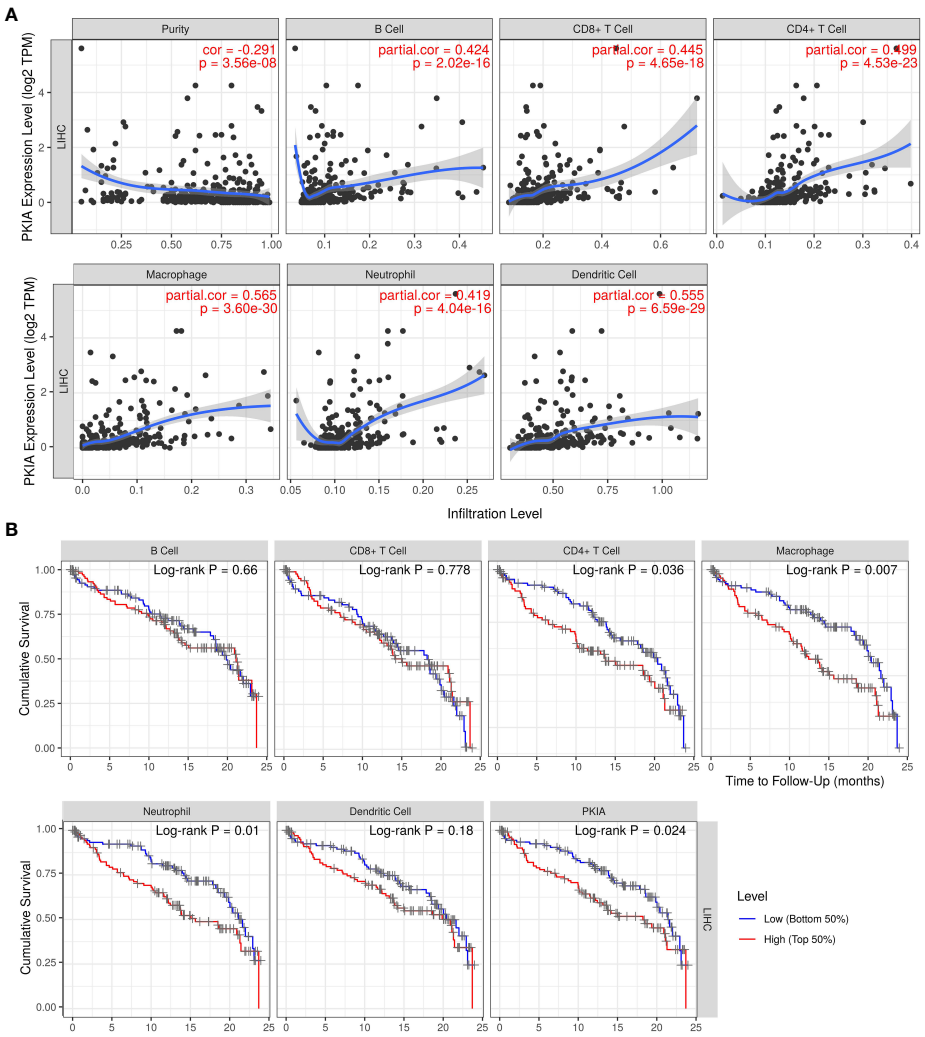


FIGURE 9
Analysis of relationship between PKIA and immune infiltration in HCC. **(A)** Correlation analysis between PKIA expression and infiltrating levels of B cells, CD8+ T cells, CD4+ T cells, macrophages, neutrophils and dendritic cells in HCC. **(B)** Survival analysis between PKIA expression and immune cells in HCC.

of cellular processes, including cell proliferation, differentiation, immune responses, angiogenesis and inflammation (Schmitz et al., 2016). In this study, we confirmed the abnormal expression levels of miR-23b and miR-589 in HCC, as well as their prognostic patterns. miR-589, likely expressed in HCC tissues, played a functional important role in prognostic signature and was recently proved to involve in HCC cell growth (Xu et al., 2018). Previous report suggested that miR-23b was down-regulated obviously in clinical HCC tissues and its exogenous upregulation inhibited EMT process of HCC cells (Cao et al., 2017; He et al., 2018). In this research, we show a schematic diagram indicating the bind sites between TRHDE-AS1 and miR-23b. However, no significant correlation was found between TRHDE-AS1 expression and miR-23b level in TCGA-HCC database.

Moreover, the interaction between miR-23b and downstream target mRNA was also found by using Starbase online tool. The mRNA 3'UTR- of cAMP-dependent protein kinase inhibitor alpha (PKIA) is predicted to bind with miR-23b. Although osteosarcoma patients with low PKIA expression was reported to have poor prognosis (Lin et al., 2022), our TCGA-based expression and survival analysis results indicated a correlation between PKIA high expression and worse prognosis for HCC patients. A report has suggested the role of over-expressed PKIA in the tumor growth of prostate cancer (Hoy et al., 2020) and proposed gene amplifications of PKIA

10.3389/fcimb.2022.1025900

TABLE 3 Correlation analysis between PKIA and immune infiltration biomarkers in HCC.

Description	Gene markers	LIHC		
		correlation	p value	significance
CD8+ T cell	CD8A	0.488	1.29E-23	***
	CD8B	0.448	9.36E-20	***
T cell	CD3D	0.570	2.72E-33	***
	CD2	0.557	1.21E-31	***
B cell	CD19	0.470	8.31E-22	***
	CD79A	0.471	6.82E-22	***
Monocyte	CD86	0.603	4.62E-38	***
	CSF1R	0.508	1.01E-25	***
	CD86	0.603	4.62E-38	***
TAM	CCL2	0.439	6.26E-19	***
	CD68	0.451	5.16E-20	***
	IL10	0.467	1.59E-21	***
M1 macrophage	NOS2	0.061	0.242514208	ns
	IRF5	0.312	8.02E-10	***
	PTGS2	0.440	1.50E-36	***
M2 macrophage	VSIG4	0.410	1.76E-16	***
	MS4A4A	0.405	4.75E-16	***
Neutrophils	CEACAM8	0.127	0.014402811	*
	ITGAM	0.455	2.54E-20	***
	CCR7	0.438	8.32E-19	***
Natural killer cell	KIR2DL1	0.046	0.373942946	ns
	KIR2DL3	0.231	6.87E-06	***
	KIR2DL4	0.269	1.40E-07	***
	KIR3DL1	0.076	0.143874889	ns
	KIR3DL2	0.241	2.56E-06	***
	KIR3DL3	0.096	0.064199312	ns
	KIR2DS4	0.166	0.001346004	**
Dendritic cell	HLA-DPB1	0.487	1.74E-23	***
	HLA-DQB1	0.415	7.07E-17	***
	HLA-DRA	0.445	1.81E-19	***
	HLA-DPA1	0.469	1.03E-21	***
	CD1C	0.421	2.44E-17	***
	NRP1	0.382	2.63E-14	***
	ITGAX	0.561	4.18E-32	***
Th1	TBX21	0.386	1.24E-14	***
	STAT4	0.371	1.52E-13	***
	STAT1	0.430	4.32E-18	***
	IFNG	0.400	1.11E-15	***
	TNF	0.514	2.23E-26	***
Th2	GATA3	0.583	3.33E-35	***
	STAT6	0.154	0.003022523	**
	STAT5A	0.484	3.23E-23	***
	IL13	0.099	0.056678881	ns
Tfh	BCL6	0.159	0.002096074	**
	IL21	0.170	0.001009247	**
Th17	STAT3	0.265	2.34E-07	***
	IL17A	0.089	0.088102091	ns

(Continued)

TABLE 3 Continued

Description	Gene markers	LIHC		
		correlation	p value	significance
Treg	FOXP3	0.254	7.17E-07	***
	CCR8	0.526	9.66E-28	***
	STAT5B	0.166	5.61E-06	***
T cell exhaustion	PDCD1	0.553	4.13E-31	***
	CTLA4	0.570	2.17E-33	***
	LAG3	0.359	1.07E-12	***
	HAVCR2	0.613	1.35E-39	***
	GZMB	0.315	5.73E-10	***

ns, not significance, *P < 0.05, **P < 0.01, ***P < 0.001.

across in prostate and lung cancers (Hoy et al., 2020). However, our analysis showed a minor alterations in HCC and no correlation between PKIA copy-number and PKIA mRNA expression. Interestingly, for HCC patients of TCGA, we observed a reduced DNA methylation status in primary HCC samples, which give a evidence for the upregulation of PKIA in HCC. Moreover, we found that TRHDE-AS1 expression level was positively correlated with PKIA level in HCC.

Next, we further interpreted the function of TRHDE-AS1/PKIA network in TPX2-related HCC and elucidated the potential mechanism of TRHDE-AS1/PKIA network in HCC patient prognosis. Meanwhile, HBV replication activates the chronic inflammation, immune response mechanisms and related inflammatory pathways, leading to the recruitment of NK cells, NKT cells, and cytotoxic T cells (Tang et al., 2018). Hepatocytes infiltrate into the microenvironment of related inflammatory cytokines, trigger liver cell damage and indirectly induce HCC (Lim et al., 2019; Jia et al., 2020). We utilized TIMER online tool which contains multiple immune deconvolution methods to analyze the correlation between PKIA expression and the immune infiltration levels. TAMs can promote tumor malignancy by producing cathepsins and facilitate cancer cell migration and invasion (Pathria et al., 2019). In this research, the expression levels of tumor-associated macrophages (TAMs)-related markers (CCL2, CD68 and IL10) were a positive correlated PKIA expression, indicating an ontogenetic role of PKIA in HCC.

Taken together, this study constructed a ceRNA network containing TRHDE-AS1 and PKIA, which related to TPX2 regulation in the HBV-related HCC. In addition, TRHDE-AS1/PKIA network also participated in prognosis of HCC patients. However, more research on the function of TRHDE-AS1 in HBV-infected HCC cells needs to be undertaken before the association between TRHDE-AS1 and PKIA is more clearly understood. Moreover, further experimental verification will be carried out to confirm the binding sites between TRHDE-AS1/PKIA and miR-23b in HCC.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Author contributions

Conceptualization, GL, ZQW, DC, and JM. Resources, GL, ZQW, JY, ZM, BS, and JM. Data curation, GL, TY, XZ, ZZ, ZQW, and JM. Software, GL, ZQW, DC, YL, PC, and YD. Formal analysis, GL, ZQW, DC, and JM. Supervision, JY, ZM, BS, YL, PC, and YD. Funding acquisition, JM. Validation, GL and JM. Visualization, GL, ZMW, DC, JY, and ZM. Methodology, TY, XZ, ZZ, ZMW, and JM. Writing—original draft, GL. Writing—review and editing, JM. All authors read and approved the final manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was funded by Shanxi Scholarship Council of China (Grant No: 2021-165), Shanxi Province Science Foundation for Distinguished Young Scholar (Grant No: 201901D211547), science and research fund of Shanxi Health Commission(Grant No: 2019059, 2022042, 2022043), Shanxi Province"136 Revitalization Medical Project Construction Funds", National Natural Science Foundation of China for Young Scholars (Grant No: 81201810), the doctor project of

Shanxi Cancer Hospital, China (2017A06), Natural Science Foundation of Guangdong Province, China (2015A030313057).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Asteriti, I. A., Rensen, W. M., Lindon, C., Lavia, P., and Guarguaglini, G. (2010). The aurora-A/TPX2 complex: a novel oncogenic holoenzyme? *Biochim. Biophys. Acta* 2, 230–239. doi: 10.1016/j.bbcan.2010.08.001

Cao, J., Liu, J., Long, J., Fu, J., Huang, L., Li, J., et al. (2017). microRNA-23b suppresses epithelial-mesenchymal transition (EMT) and metastasis in hepatocellular carcinoma *via* targeting Pyk2. *BioMed. Pharmacother* 89, 642–650. doi: 10.1016/j.biopha.2017.02.030

Chang, H., Wang, J., Tian, Y., Xu, J., Gou, X., and Cheng, J. (2012). The TPX2 gene is a promising diagnostic and therapeutic target for cervical cancer. *Oncol. Rep.* 27, 1353–1359. doi: 10.3892/or.2012.1668

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* 6, 2004088. doi: 10.1126/scisignal.2004088

Gil, N., and Ulitsky, I. (2020). Regulation of gene expression by cis-acting long non-coding RNAs. *Nat. Rev. Genet.* 21, 102–117. doi: 10.1038/s41576-019-0184-5

Gomes-Filho, S. M., Dos Santos, E. O., Bertoldi, E. R. M., Scalabrini, L. C., Heidrich, V., Dazzani, B., et al. (2020). Aurora a kinase and its activator TPX2 are potential therapeutic targets in KRAS-induced pancreatic cancer. *Cell Oncol.* 43, 445–460. doi: 10.1007/s13402-020-00498-5

He, R. Q., Wu, P. R., Xiang, X. L., Yang, X., Liang, H. W., Qiu, X. H., et al. (2018). Downregulated miR-23b-3p expression acts as a predictor of hepatocellular carcinoma progression: A study based on public data and RT-qPCR verification. *Int. J. Mol. Med.* 41, 2813–2831. doi: 10.3892/ijmm.2018.3513

Hoy, J. J., Salinas Parra, N., Park, J., Kuhn, S., and Iglesias-Bartolome, R. (2020). Protein kinase a inhibitor proteins (PKIs) divert GPCR-Gαs-cAMP signaling toward EPAC and ERK activation and are involved in tumor growth. *FASEB J.* 34, 13900–13917. doi: 10.1096/fj.202001515R

Hsu, P. K., Chen, H. Y., Yeh, Y. C., Yen, C. C., Wu, Y. C., Hsu, C. P., et al. (2014). TPX2 expression is associated with cell proliferation and patient outcome in esophageal squamous cell carcinoma. *J. Gastroenterol.* 49, 1231–1240. doi: 10.1007/s00535-013-0870-6

Huang, W., Huang, F., Lei, Z., and Luo, H. (2020). LncRNA SNHG11 promotes proliferation, migration, apoptosis, and autophagy by regulating hsa-miR-184/AGO2 in HCC. *Onco Targets Ther.* 13, 413–421. doi: 10.2147/OTT.S237161

Huang, Q., Lin, B., Liu, H., Ma, X., Mo, F., Yu, W., et al. (2011). RNA-Seq analyses generate comprehensive transcriptomic landscape and reveal complex transcript patterns in hepatocellular carcinoma. *PLoS One* 6, 17. doi: 10.1371/journal.pone.0026168

Jia, L., Gao, Y., He, Y., Hooper, J. D., and Yang, P. (2020). HBV induced hepatocellular carcinoma and related potential immunotherapy. *Pharmacol. Res.* 159, 4. doi: 10.1016/j.phrs.2020.104992

Lim, C. J., Lee, Y. H., Pan, L., Lai, L., Chua, C., Wasser, M., et al. (2019). Multidimensional analyses reveal distinct immune microenvironment in hepatitis b virus-related hepatocellular carcinoma. *Gut.* 68, 916–927. doi: 10.1136/gutjnl-2018-316510

Lin, C., Miao, J., He, J., Feng, W., Chen, X., Jiang, X., et al. (2022). The regulatory mechanism of LncRNA-mediated ceRNA network in osteosarcoma. *Sci. Rep.* 12, 022–11371. doi: 10.1038/s41598-022-11371-w

Liu, X., Peng, D., Cao, Y., Zhu, Y., Yin, J., Zhang, G., et al. (2021). Upregulated lncRNA DLX6-AS1 underpins hepatocellular carcinoma progression *via* the miR-513c/Cul4A/ANXA10 axis. *Cancer Gene Ther.* 28, 486–501. doi: 10.1038/s41417-020-00233-0

Liu, Y., Yan, W., Zhou, D., Jin, G., and Cheng, X. (2020). Long non–coding RNA HOXA11–AS accelerates cell proliferation and epithelial–mesenchymal transition in hepatocellular carcinoma by modulating the miR–506–3p/Slug axis. *Int. J. Mol. Med.* 46, 1805–1815. doi: 10.3892/ijmm.2020.4715

Li, C., Wei, B., and Zhao, J. (2021). Competing endogenous RNA network analysis explores the key lncRNAs, miRNAs, and mRNAs in type 1 diabetes. *BMC Med. Genomics* 14, 021–00877. doi: 10.1186/s12920-021-00877-3

Li, S. P., Xu, H. X., Yu, Y., He, J. D., Wang, Z., Xu, Y. J., et al. (2016). LncRNA HULC enhances epithelial-mesenchymal transition to promote tumorigenesis and metastasis of hepatocellular carcinoma *via* the miR-200a-3p/ZEB1 signaling pathway. *Oncotarget.* 7, 42431–42446. doi: 10.18632/oncotarget.9883

Llovet, J. M., Burroughs, A., and Bruix, J. (2003). Hepatocellular carcinoma. *Lancet.* 362, 1907–1917. doi: 10.1016/S0140-6736(03)14964-1

Lu, W. P., and Dong, J. H. (2014). Hepatectomy for hepatocellular carcinoma in the era of liver transplantation. *World J. Gastroenterol.* 20, 9237–9244. doi: 10.3748/wjg.v20.i28.9237

Lu, S., Zhu, N., Guo, W., Wang, X., Li, K., Yan, J., et al. (2020). RNA-Seq revealed a circular RNA-microRNA-mRNA regulatory network in hantaan virus infection. *Front. Cell Infect. Microbiol.* 10. doi: 10.3389/fcimb.2020.00097

Marasco, G., Colecchia, A., Colli, A., Ravaioli, F., Casazza, G., Bacchi Reggiani, M. L., et al. (2019). Role of liver and spleen stiffness in predicting the recurrence of hepatocellular carcinoma after resection. *J. Hepatol.* 70, 440–448. doi: 10.1016/j.jhep.2018.10.022

Mercer, T. R., Dinger, M. E., and Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.* 10, 155–159. doi: 10.1038/nrg2521

Neumayer, G., Belzil, C., Gruss, O. J., and Nguyen, M. D. (2014). TPX2: of spindle assembly, DNA damage response, and cancer. *Cell Mol. Life Sci.* 71, 3027–3047. doi: 10.1007/s00018-014-1582-7

Paraskevopoulou, M. D., and Hatzigeorgiou, A. G. (2016). Analyzing MiRNA-LncRNA interactions. *Methods Mol. Biol.* 1402, 3378–3375_3321. doi: 10.1007/978-1-4939-3378-5_21

Pathria, P., Louis, T. L., and Varner, J. A. (2019). Targeting tumor-associated macrophages in cancer. *Trends Immunol.* 40, 310–327. doi: 10.1016/j.it.2019.02.003

Peng, W. X., Koirala, P., and Mo, Y. Y. (2017). LncRNA-mediated regulation of cell signaling in cancer. *Oncogene.* 36, 5661–5667. doi: 10.1038/onc.2017.184

Qi, X., Zhang, D. H., Wu, N., Xiao, J. H., Wang, X., and Ma, W. (2015). ceRNA in cancer: possible functions and clinical implications. *J. Med. Genet.* 52, 710–718. doi: 10.1136/jmedgenet-2015-103334

Schmitz, S. U., Grote, P., and Herrmann, B. G. (2016). Mechanisms of long noncoding RNA function in development and disease. *Cell Mol. Life Sci.* 73, 2491–2509. doi: 10.1007/s00018-016-2174-5

Seto, W. K., Lo, Y. R., Pawlotsky, J. M., and Yuen, M. F. (2018). Chronic hepatitis b virus infection. *Lancet.* 392, 2313–2324. doi: 10.1016/S0140-6736(18)31865-8

Sia, D., Villanueva, A., Friedman, S. L., and Llovet, J. M. (2017). Liver cancer cell of origin, molecular class, and effects on patient prognosis. *Gastroenterology.* 152, 745–761. doi: 10.1053/j.gastro.2016.11.048

Tang, L. S. Y., Covert, E., Wilson, E., and Kottilil, S. (2018). Chronic hepatitis b infection: A review. *JAMA* 319, 1802–1813. doi: 10.1001/jama.2018.3795

Tay, Y., Rinn, J., and Pandolfi, P. P. (2014). The multilayered complexity of ceRNA crosstalk and competition. *Nature.* 505, 344–352. doi: 10.1038/nature12986

Tsukuda, S., and Watashi, K. (2020). Hepatitis b virus biology and life cycle. *Antiviral Res.* 182, 28. doi: 10.1016/j.antiviral.2020.104925

Wong, D. K., Cheng, S. C. Y., Mak, L. L., To, E. W., Lo, R. C., Cheung, T. T., et al. (2020). Among patients with undetectable hepatitis b surface antigen and hepatocellular carcinoma, a high proportion has integration of HBV DNA into hepatocyte DNA and no cirrhosis. *Clin. Gastroenterol. Hepatol.* 18, 449–456. doi: 10.1016/j.cgh.2019.06.029

Xu, M., Wang, Y., He, H. T., and Yang, Q. (2018). MiR-589-5p is a potential prognostic marker of hepatocellular carcinoma and regulates tumor cell growth by targeting MIG-6. *Neoplasma*. 65, 753–761. doi: 10.4149/neo_2018_171125N762

Yoo, S., Wang, W., Wang, Q., Fiel, M. I., Lee, E., Hiotis, S. P., et al. (2017). A pilot systematic genomic comparison of recurrence risks of hepatitis b virus-associated hepatocellular carcinoma with low- and high-degree liver fibrosis. *BMC Med.* 15, 017–0973. doi: 10.1186/s12916-017-0973-7

Zeng, X. C., Zhang, L., Liao, W. J., Ao, L., Lin, Z. M., Kang, W., et al. (2020). Screening and identification of potential biomarkers in hepatitis b virus-related hepatocellular carcinoma by bioinformatics analysis. *Front. Genet.* 11. doi: 10.3389/fgene.2020.555537

Zhang, H., Chen, X., Zhang, J., Wang, X., Chen, H., Liu, L., et al. (2020). Long non−coding RNAs in HBV−related hepatocellular carcinoma (Review). *Int. J. Oncol.* 56, 18–32. doi: 10.3892/ijo.2019.4909

Zhang, H., Xu, H. B., Kurban, E., and Luo, H. W. (2020). LncRNA SNHG14 promotes hepatocellular carcinoma progression *via* H3K27 acetylation activated PABPC1 by PTEN signaling. *Cell Death Dis.* 11, 020–02808. doi: 10.1038/s41419-020-02808-z

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Brief Research Report
PUBLISHED 13 October 2022
DOI 10.3389/fcimb.2022.917009

OPEN ACCESS

EDITED BY
Chengjun Wu,
Dalian University of Technology, China

REVIEWED BY
Wenjing Shen,
The First Hospital of China Medical University, China
Pu Li,
Tianjin Central Hospital for Gynecology and Obstetrics, China
Guangwei Wang,
ShengJing Hospital of China Medical University, China

*CORRESPONDENCE
Jun Wang
wj202fck@163.com
Shichao Han
hscuperman@126.com

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to Clinical Microbiology, a section of the journal Frontiers in Cellular and Infection Microbiology

RECEIVED 10 April 2022
ACCEPTED 20 July 2022
PUBLISHED 13 October 2022

CITATION
Lu J, Na J, Li Y, Wang X, Wang J and Han S (2022) Gastric-type mucinous endocervical adenocarcinomas: A case report and literature review.
Front. Cell. Infect. Microbiol. 12:917009.
doi: 10.3389/fcimb.2022.917009

Gastric-type mucinous endocervical adenocarcinomas: A case report and literature review

Junling Lu†, Jing Na†, Ya Li, Xinyou Wang, Jun Wang* and Shichao Han*

Department of Gynecology and Obstetrics, Second Affiliated Hospital of Dalian Medical University, Dalian, China

Gastric-type mucinous endocervical adenocarcinomas (GAS) are new variant types of cervical adenocarcinomas according to the 2014 World Health Organization (WHO) classification. GAS is a unique disease that can be differentiated from typical adenocarcinomas—it is less common and more aggressive and likely to have deep invasion and horizontal diffusion, invasion of the uterus and vagina, early distant metastases, and a lower 5-year survival rate compared to the usual-type cervical cancer. At present, initial treatment and postoperative adjuvant therapy are not conclusive, but early detection and early treatment are a consensus that can improve prognosis. Most of its occurrence has nothing to do with human papillomavirus (HPV) infection. Whether it is only negative for the subtypes that can be detected at present and whether it may be an unknown subtype of infection need to be further explored in the future. The clinical symptoms commonly include aqueous secretion, lower abdominal pain, and elevated serum carbohydrate antigen-19-9 (CA19-9) levels, which may be helpful for diagnosis. MRI and PET-CT can help to describe the characteristics of lesions and judge the state of the systemic metastasis. We believe that early detection and surgical treatment will give patients more benefits. Looking for potential gene and molecular changes and establishing biomarkers to identify molecular targets will be the key to early identification and target therapy.

KEYWORDS

human papillomavirus, gastric type, adenocarcinoma, mucinous, diagnosis, therapy

Introduction

Cervical cancers are the fourth most frequent cancers among women in the world (Ferlay et al., 2015), which are the most common reasons for cancer deaths in women globally. In 2020, the World Health Organization (WHO) proposed a global strategy, which is expediting the elimination of cervical cancer as a community health issue

(WHO, 2020). As we all know, infection with high-risk human papillomavirus (HPV) has been implicated as the primary cause of cervical carcinogenesis (Schiffman et al., 2007). However, we found that few patients with cervical cancer are HPV negative. What is terrible is that these patients are easy to be forgotten and missed. We will report a case of specific tumor stage IIIc1p gastric-type mucinous endocervical adenocarcinoma (GAS) that is HPV negative. Meanwhile, by searching PubMed, ClinicalKey, and other databases, we try to retrospectively analyze the related research of GAS in the recent 10 years, which is expected to summarize the experience of diagnosis and treatment of GAS.

Case report

A 42-year-old woman was transferred to our hospital in January 2022. Approximately 5 months before hospitalization, the patient began to have no obvious inducement to discharge a large amount of transparent thin mucus from the vagina without peculiar smell. The symptoms were relieved after vaginal medication. Two weeks before hospitalization, she developed a dull pain in her lower abdomen also without an apparent cause. At that time, she had no fever, dizziness, fatigue, and other discomfort. She accepted the ultrasound examination at the local community clinic that showed that the local echo of the cervix was reduced and the blood flow at the external orifice of the cervix was abundant. Subsequently, she accepted ThinPrep cytologic test (TCT) and HPV genotype test in the local hospital. The results of the TCT showed atypical glandular cells (AGCs), and the HPV DNA test was negative. Then, she underwent colposcopy, cervical multipoint biopsy, and endocervical curettage. The pathology showed mucinous adenocarcinoma and gastric adenocarcinoma. Finally, she came to our hospital in order to seek further diagnosis and treatment. She is in good health and has no chronic and infectious diseases in the past. She had never been screened for cervical cancer or vaccinated against cervical cancer before. She has been pregnant eight times, including two cesarean sections and six abortions. There is no history of cancer in her family. Her mother and sister are all suffering from hypertension, and her father is suffering from diabetes.

By bimanual rectovaginal examination, we found that the cervical surface was smooth but barrel thickened, which is easy to bleed when touched. At the same time, we did not find parametrial invasion. MRI of the pelvis showed that the cervical mass was about 40 mm * 36 mm * 48 mm, invading the lower part of the uterine body and pelvic vessels, but without parametrial, vaginal extension and enlarged pelvic lymph nodes (Figure 1). PET-CT showed that there were a large number of radioactive foci in the cervix and lower uterine segment, which were suspected to be malignant tumors, but there was no obvious radioactive foci, suggesting lymph node metastasis in the pelvic cavity and distant metastasis (Figure 2). Carbohydrate antigen 125 (CA125), carcinoembryonic antigen

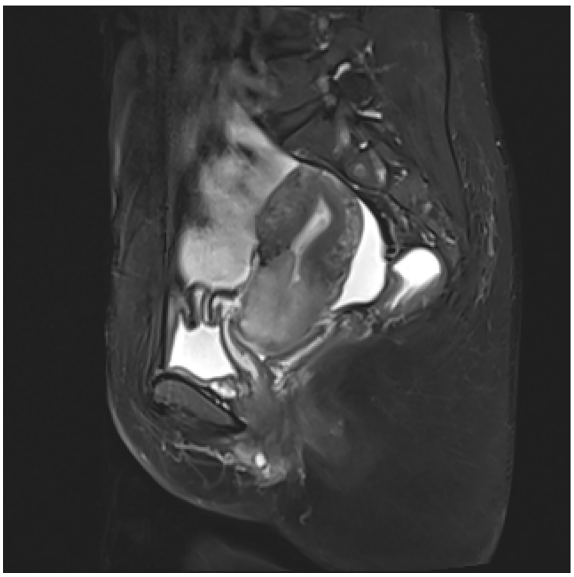


FIGURE 1
Pelvic magnetic resonance imaging (MRI) showed a cervical tumor 40 mm * 36 mm * 48 mm in size, without invasion of the uterus and vagina, and no pelvic lymph node metastasis.

(CEA), alpha fetoprotein (AFP), and human epididymis protein 4 (HE4) were all within normal range, but carbohydrate antigen-19-9 (CA19-9) and the antigen of squamous cell carcinoma (SCC) rose; the corresponding results were 1,303.61 U/ml and 2.89 ng/ml, especially CA19-9 increased significantly. No obvious abnormality was found in other examinations including colonoscopy and gastroscopy. Finally, according to the clinical staging standard of the International Federation of Obstetrics and Gynecology (FIGO, 2018), we set the preoperative staging of the patient as IB3. Then, we performed cystoscopy, bilateral ureteral D-J tube placement, extensive hysterectomy (type C), bilateral salpingo-oophorectomy, and bilateral pelvic lymphadenectomy (level 2). After surgery, the isolated uterus was dissected. The morphology of the uterus was regular. No space-occupying lesions were found in the uterine cavity. The endometrium was thin. Tumors could be seen between the circumferential walls of the cervix. The size was about 4 cm * 4 cm * 3.5 cm and did not invade the vagina. According to the naked eye examination of the sample, the deep cervical interstitial infiltration was full-thickness. The width of the parauterine tissues on both sides was measuredto be about 3 cm, the length of the vaginal resection was about 3 cm, and there was no significant increase in pelvic lymph nodes on both sides (Figure 3). The pathology revealed that the cervical adenocarcinoma (non-HPV-related, gastric-type), with an infiltration depth of 1.7 cm, infiltrates the whole layer of the cervical muscle wall. Multiple intravascular tumor thrombi and nerve invasion can be seen, involving the endometrium and the muscle layer, without involving the vaginal fornix. No cancer tissue was found in the broken end of the vagina

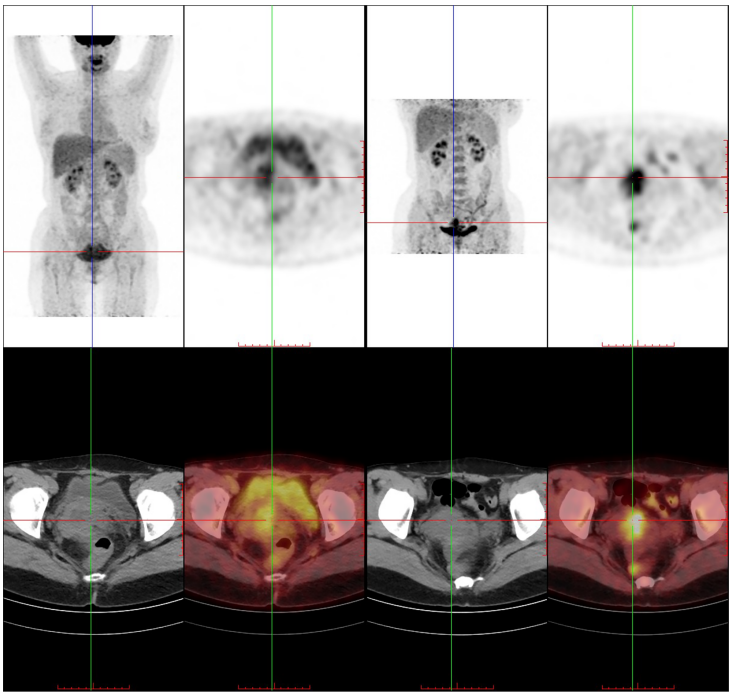


FIGURE 2
The cervix is widened with mass radioactive concentration foci, with a range of about 38 mm * 24 mm, SUVmax 4.0, delay 5.2. Radioactive concentration can be seen in the lower part of the uterine body, SUVmax 3.9, delay 5.6. Uterine effusion.

and parauterine soft tissue. Two of the 20 pelvic lymph nodes showed metastasis and invasion (Figure 4). Immunohistochemical stains showed that tumor cells were MLH1 (+), MSH (+),PMS2 (+), HER2 (1+), ER (<1% weak +), BRCA1 (+), p53 (85% medium intensity +), ERCC1 (+), PRMI (–), TopoII grade. During molecular detection, no mutation of BRAF gene V600E was detected. On the first day after the operation, CA19-9 decreased rapidly (662.16 μ/ml) and fell to the normal range (CA19-9: 12.72 μ/ml) after 1 month. Because she has some prognostic factors, such as tumor body 4 cm and involves the whole layer of the cervix, multiple intravascular tumor thrombi and nerve invasion, partial uterine body involvement and pelvic lymph node metastasis, she needs to receive a combination of chemotherapy and radiotherapy. We plan that she will receive three cycles of chemotherapy first, then pelvic radiotherapy, and finally three cycles of chemotherapy.

Discussion

Endocervical adenocarcinoma (ECA) is a group of tumors with higher heterogeneity, and about 15% is not associated with HPV infection (Stolnicu et al., 2018). In 2014, WHO classified cervical adenocarcinoma according to morphological and cellular characteristics and proposed a new disease category of GAS for the first time (Kurman et al., 2014). In 2018, the International Endocervical Adenocarcinoma Criteria and Classification (IECC) system attempted to categorize ECA by etiology (HPV status) and morphology with molecular and outcome evidence (Stolnicu et al., 2018). Obviously, the latter is more comprehensive and clinically relevant. It is reported that the incidence rate of GAS is 20%–25% of all cervical adenocarcinomas (Mikami and McCluggage, 2013).

HPV infection is directly related to cancer. High-risk HPV is highly associated with genital cancers including cervical cancer and head and neck squamous cell carcinoma (HNSCC) (Antonsson et al., 2015; Schiffman et al., 2016). Data showed that HPV was detected in 99.7% of cervical cancer patients, in whom HPV16 alone accounted for about 55% (de Sanjosé et al., 2007). The incidence rate of HNSCC caused by HPV infection has been increasing in recent years (Morgan et al., 2017). Several studies on cervical adenocarcinoma have shown that the prevalence of HPV infection is lower. About 5%–25% of cervical adenocarcinoma is independent of HPV, which depends on many factors such as geographical region, histological subtype, and HPV detection method (Pirog et al., 2014; Molijn et al., 2016). Several studies have indicated that GAS is mainly not related to HPV infection (Houghton et al., 2010; Park et al., 2011). In addition, whether it is only negative for the subtypes that can be detected at present and whether it may be an unknown subtype of infection need to be further explored in the future. In the era of HPV vaccination, the

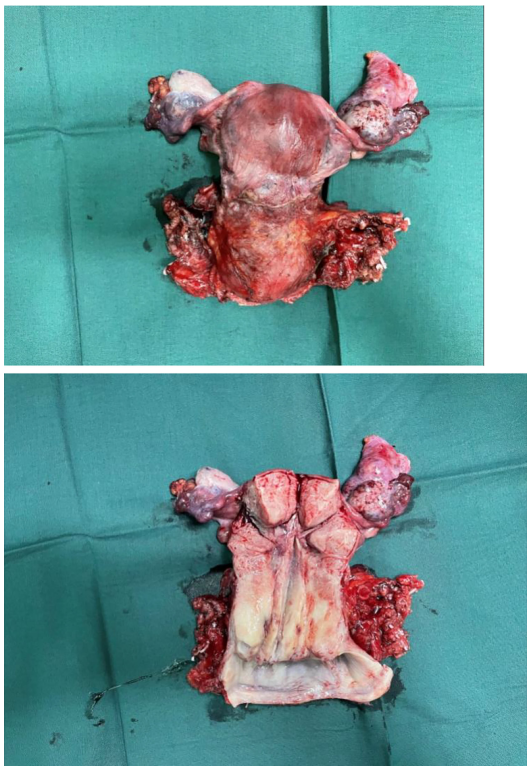


FIGURE 3
Isolated uterus after the operation.

prevention and screening of cervical cancer have played a vital role. For HPV-negative cervical gastric adenocarcinoma, we still need to further explore its pathogenesis and methods of prevention and screening.

Early detection and timely treatment of cervical precancerous lesions are very important for cancer prevention and early detection. The high-risk HPV test of patients with GAS is usually negative, while the positive rate of routine cervical exfoliative cytology is low, which poses a great challenge for us to find its precursor lesions. The precursor lesions of GAS include lobar endocervical glandular hyperplasia (LEGH), atypical LEGH, and endocervical adenocarcinoma *in situ* (AIS). The latter two have common genetic characteristics, such as 1p deletion and 3q acquisition. In addition, GAS is also closely related to Peutz-Jeghers syndrome (PJS) (Kawauchi et al., 2008; Beggs et al., 2010). About 10% of GAS patients have PJS, which may be related to STK11 gene mutation (Peng et al., 2015). These may bring some help to our early discovery of GAS.

As early as 1988, a case–control study of 39 cases and 409 controls in the greater Milan region showed that the relative risk of cervical adenocarcinoma increased with the increase of the number of deliveries and abortions, the age of birth, and the age of first sexual intercourse (Parazzini et al., 1988). The current analyses were based on data from the multicenter TeQaZ study, which shows that currently living with a partner, having at least four children, having had more than one sexual partner, ever smoking, and obesity were significantly associated with cervical cancer (Tanaka et al., 2021). In our report, the patient has been

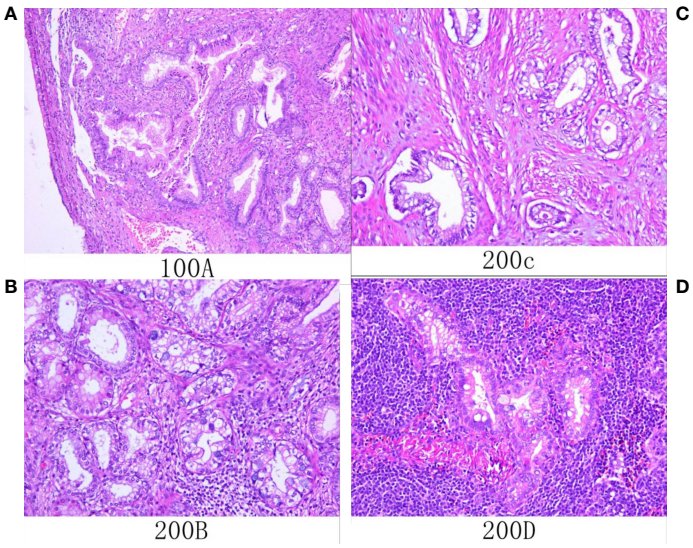


FIGURE 4
Histological (**A–D**; hematoxylin and eosin staining) results. **(A)** Microscopic examination showed that atypical glands are irregular and expand with gastric differentiation. **(B)** The tumor showed gastric differentiation, including an atypical nucleus and abundant eosinophilic cytoplasm that was transparent or pale. **(C)** The tumor invaded the myometrium. **(D)** Tumor cells metastasized to local lymph nodes.

pregnant eight times, including two cesarean sections and six abortions. We think that multiple pregnancies may be associated with cervical cancer, but whether it is a high-risk factor for the occurrence of GAS still needs further research to confirm.

Nakamura et al. (2018) reviewed 322 cases of cervical cancer and found that it is hard to diagnose GAS by preoperative biopsy, but some specific clinical manifestations, such as aqueous secretion and lower abdominal pain, high serum CA19-9 level, and immunohistochemistry (HIK1083 and MUC6 staining), often help the diagnosis. Ihara et al. (1988) showed that CA125 and CA19-9 are useful tumor markers for the treatment of cervical adenocarcinoma and endometrial cancer, especially in advanced or recurrent cases. They found that CA125 or CA19-9 immunohistochemistry was localized in cancer tissue, and in most recurrent cases, serum CA125 and CA19-9 increased significantly in the early stage (Ihara et al., 1988). Maruyama et al. (1988) used histochemical methods to detect the immunohistochemical localization of epithelial mucin and carbohydrate antigen (CA125, CA19-9) in normal cervical glands and cervical adenocarcinoma. They found that CA19-9, CA125, and CEA were localized in cervical adenocarcinoma and intestinal adenocarcinoma groups (Maruyama et al., 1988). The patient in our case also had more typical symptoms of vaginal fluid and lower abdominal pain. The character of vaginal discharge in patients with gastric adenocarcinoma is the same as that of gastric secretion, and the mechanism remains to be further explored. In our case, CA19-9 increased significantly before the operation, decreased to one-half on the first day after the operation, and returned to normal level 1 month after the operation. Combined with previous studies, we believe that CA19-9 may be used as a reference index to monitor the efficacy and recurrence of gastric cervical adenocarcinoma. At the same time, MRI is a good and accessible tool for diagnosis, which can describe the characteristics of lesions and evaluate local regional expansion. PET-CT can help judge the state of the whole body and whether there is distant metastasis. In our case, the tumor cells were with much transparent and light eosinophilic cytoplasm, clear cell boundaries, low nuclear plasma ratio, and irregularly distributed nuclei at the base of the gland. It is consistent with the pathomorphological characteristics of cervical gastric adenocarcinoma.

Several recent studies have revealed diversities in the pathological mechanism and molecular features between GAS and usual-type endocervical adenocarcinoma (UEA). They found that GAS showed obviously deeper infiltration depth, greater horizontal diffusion, later stage, and more common distant metastasis. In other words, it was more likely to invade the parauterine and vagina. At the same time, P53 protein was always immunostained. Moreover, the disease-free survival time of GAS was obviously shorter than that of UEA (Nishio et al., 2019; Jung et al., 2020). In our case, the patient was diagnosed as IB3 before the operation, but postoperative pathology confirmed pelvic lymph node metastasis, which is basically consistent with their findings. Garg et al. (2019) elucidate the molecular characteristics of gastric cervical adenocarcinoma by using next-generation sequencing (NGS) to evaluate 161 unique cancer-driver genes. Their study showed the genetic heterogeneity of gastric cervical adenocarcinoma and found that there may be some potentially operable molecular changes. These molecular characteristics are very important for better discerning this rare tumor and better treatment (Garg et al., 2019). Lu et al. (2021) also used NGS technology to identify the unique genomic changes of GAS, mainly involving cell cycle and PI3K/AKT (Phosphatidylinositde-3-Kinase/Protein Kinase B) signaling pathways. If we can further identify the molecular characterization of GAS and find out the potential operable molecular changes, it will bring new hope for tumor-targeted therapy. By finding genetic changes of the disease, we can understand the pathogenesis better, and it is possible to get potential targets and individualized treatment strategies in the future.

As GAS is relatively rare, there is still lack of prospective clinical research. The treatment of GAS is controversial, and there is no unified standard. The Chinese Expert Consensus on the Clinical Diagnosis and Treatment of Cervical and Gastric Adenocarcinoma (2021 Edition) pointed out that in the absence of standard treatment methods, the treatment of GAS can refer to the 2022 NCCN Clinical Practice Guidelines for Cervical Cancer (First Edition) for the treatment of small cell carcinoma (Gynecological oncology professional committee (Study Group) of gynecologist and obstetrician branch of Chinese Medical Association, 2021). Therefore, the selection of our case treatment plan is based on NCCN guidelines. This guideline points out that for patients with tumors ≤4 cm, radical hysterectomy + pelvic lymphadenectomy ± para-aortic lymph node sampling is the first choice for surgery. Chemotherapy or concurrent radiotherapy and chemotherapy can be selected after the operation. However, you can also choose concurrent radiotherapy and chemotherapy + brachytherapy and then consider combining other systemic treatments (NCCN, 2022). Whether the scope of surgery is similar to that of ovarian cancer staging is still controversial. We believe that whether bilateral appendectomy, omentum resection, and appendectomy can prolong the survival time and improve the quality of life of patients still needs further study. It is reported that gastric cervical adenocarcinoma has chemoresistance. Therefore, its prognosis is poorer than the other types of cervical adenocarcinomas (Kawakami et al., 2010). Li et al. (2010) emphasized the importance of surgical treatment, followed for advanced stages by radiochemotherapy. Therefore, we believe that early detection and surgical treatment will give patients more benefits.

The recurrence rate of GAS is higher, approximately 40%, and the 5-year survival rate is only about 30% (Kojima et al., 2007; Nishio et al., 2019). The prognosis is related to tumor stage, parauterine invasion, resection margin, metastasis, and treatment. A study found that patients with programmed death-ligand 1 (PD-L1)-positive GAS had a worse prognosis (including

disease progression-free survival and overall survival) than those with PD-L1-negative GAS. It is suggested that PD-L1 can be used as a marker of poor prognosis of GAS (Chen et al., 2021). We will closely follow up this patient and continue to track her prognosis. In the future, we need to do more work to find indicators to monitor the prognosis of GAS.

Conclusions

GAS is usually HPV negative, which is difficult to find in cervical precancer screening. Although the current literature reports that GAS is not related to high-risk HPV (Kojima et al., 2007; Kurman et al., 2014; Mikami et al., 2020), whether it is only negative for the currently detectable subtypes and whether it may be an unknown infection subtype are worth further exploration in the future. Combined with previous studies, we think that CA19-9 may be used as a reference index to monitor the efficacy and recurrence of gastric cervical adenocarcinoma. It is not sensitive to chemotherapy drugs, and the prognosis is worse than those of SCC and cervical adenocarcinoma. Early radical surgery is very important for patients. Looking for potential gene and molecular changes and establishing biomarkers to identify molecular targets will be the key to early detection and targeted therapy.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Ethics statement

Written informed consent was obtained from the individual (s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

JL and SH dealt with the case and drafted the manuscript. JN and YL assisted collected case data and literature and carried out all the documentary and article work out. JW has made great efforts and has given many constructive suggestions for this paper during the revision and production period. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Antonsson, A., Neale, R. E., Boros, S., Lampe, G, Coman, W. B., Pryor, D. I., et al. (2015). Human papillomavirus status and p16 (ink4a) expression in patients with mucosal squamous cell carcinoma of the head and neck in queensland, australia. *Cancer Epidemiol.* 39, 174–181. doi: 10.1016/j.canep.2015.01.010

Beggs, A. D., Latchford, A. R., Vasen, H. F. A., Moslein, G., Alonso, A., Aretz, S., et al. (2010). Peutz-jeghers syndrome: A systematic review and recommendations for management. *Gut* 59 (7), 975–986. doi: 10.1136/gut.2009.198499

Chen, L., Lucas, E., Zhang, X., Liu, Q., Zhuang, Y., Lin, W., et al. (2021). PD-L1 expression in HPV-independent cervical adenocarcinoma and its prognostic significance. *Histopathology* 80 (2), 338–47. doi: 10.1111/his.14552

de Sanjosé, S., Diaz, M., Castellsagué, X., Clifford, G., Bruni, L., Muñoz, N., et al. (2007). Worldwide prevalence and genotype distribution of cervical human papillomavirus DNA in women with normal cytology: A meta-analysis. *Lancet Infect. Dis.* 7 (7), 453–459. doi: 10.1016/S1473-3099(07)70158-5

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., et al. (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* 136 (5), E359–E386. doi: 10.1002/ijc.29210

Garg, S., Nagaria, T. S., Clarke, B., Freedman, O., Khan, Z., Schwock, J., et al. (2019). Molecular characterization of gastric-type endocervical adenocarcinoma using next-generation sequencing. *Modern. Pathol* 32 (12), 1823–33. doi: 10.1038/s41379-019-0305-x

Gynecological oncology professional committee (Study Group) of gynecologist and obstetrician branch of Chinese Medical Association (2021). The Chinese expert consensus on the clinical diagnosis and treatment of cervical and gastric adenocarcinoma (2021 edition). *Chin. J. Pract. Gynecol. Obstet.* 37(11), 1133–38. doi: 10.19538/j.fk2021110112

Houghton, O., Jamison, J., Wilson, R., Carson, J., and McCluggage, W. G. (2010). p16 immunoreactivity in unusual types of cervical adenocarcinoma does not reflect human papillomavirus infection. *Histopathology* 57, 342–350. doi: 10.1111/j.1365-2559.2010.03632.x

Ihara, Y., Shimizu, T., Kawaguchi, K., Yomura, W., Fujiwara, T., and Inoue, K. (1988). Serum CA125 and CA19-9 levels in adenocarcinoma of the uterine cervix and endometrial carcinoma. *Comp. Study. Nihon. Sanka. Fujinka. Gakkai. Zasshi.* 40 (11), 1711–1718.

Jung, H., Bae, G. E., Kim, H. M., and Kim, H.-S. (2020). Clinicopathological and molecular differences between gastric-type mucinous carcinoma and usual-type endocervical adenocarcinoma of the uterine cervix. *Cancer Genomics - Proteomics* 17 (5), 627–641. doi: 10.21873/cgp.20219

Kawakami, F., Mikami, Y., Kojima, A., Ito, M., Nishimura, R., and Manabe, T. (2010). Diagnostic reproducibility in gastric-type mucinous adenocarcinoma of the

uterine cervix: validation of novel diagnostic criteria. *Histopathology* 56, 551–553. doi: 10.1111/j.1365-2559.2010.03500.x

Kawauchi, S., Kusuda, T., Liu, X. P., Suehiro, Y., Kaku, T., Mikami, Y., et al. (2008). Is lobular endocervical glandular hyperplasia a cancerous precursor of minimal deviation adenocarcinoma? A comparative molecular-genetic and immunohistochemical study. *Am. J. Surg. Pathol.* 32 (12), 1807–1815. doi: 10.1097/PAS.0b013e3181883722

Kojima, A., Mikami, Y., Sudo, T., Yamaguchi, S., Kusanagi, Y., Ito, M., et al. (2007). Gastric morphology and immunophenotype predict poor outcome in mucinous adenocarcinoma of the uterine cervix. *Am. J. Surg. Pathol.* 31 (5), 664–672. doi: 10.1097/01.pas.0000213434.91868.b0

Kurman, R., Carcangiu, M. L., Herrington, C. S., and Young, R. H. (2014). *WHO classification of tumours of female reproductive organs* (Lyon: IARC Publications).

Li, G., Jiang, W., Gui, S., and Xu, C. (2010). Minimal deviation adenocarcinoma of the uterine cervix. *Int. J. Gynaecol. Obstet.* 110, 89–92. doi: 10.1016/j.ijgo.2010.03.016

Lu, S., Shi, J., Zhang, X., Kong, F., Liu, L., Dong, X., et al. (2021). Comprehensive genomic profiling and prognostic analysis of cervical gastric-type mucinous adenocarcinoma. *Virchows. Archiv.* 479 (5), 893–903. doi: 10.1007/s00428-021-03080-y

Maruyama, C., Yamamichi, N., and Konishi, F. (1988). Histochemical analysis of uterine cervical adenocarcinoma with reference to mucosubstances and distribution of CEA, CA125 and CA19-9. *Nihon. Sanka. Fujinka. Gakkai. Zasshi.* 40 (4), 429–436.

Mikami, Y.. (2020). Gastric-type mucinous carcinoma of the cervix and its precursors - historical overview. *Histopathology* 76 (1), 102–111. doi: 10.1111/his.13993

Mikami, Y., and McCluggage, W. G. (2013). Endocervical glandular lesions exhibiting gastric differentiation: An emerging spectrum of benign, premalignant, and malignant lesions. *Adv. Anat. Pathol.* 20, 227–237. doi: 10.1097/PAP.0b013e31829c2d66

Molijn, A., Jenkins, D., Chen, W., Zhang, X., Pirog, E., Enqi, W., et al. (2016). The complex relationship between human papillomavirus and cervical adenocarcinoma. *Int. J. Cancer* 138 (2), 409–416. doi: 10.1002/ijc.29722

Morgan, I. M., DiNardo, L. J., and Windle, B. (2017). Integration of human papillomavirus genomes in head and neck cancer: Is it time to consider a paradigm shift? *Viruses* 9 (8), 208. doi: 10.3390/v9080208

Nakamura, A., Yamaguchi, K., Minamiguchi, S., Murakami, R., Abiko, K., Hamanishi, J., et al. (2018). Mucinous adenocarcinoma, gastric type of the uterine cervix: clinical features and HER2 amplification. *Japanese. Soc. Clin. Mol. Morphol.*

Abu-rustum, N. R., Yashar, C.M., Bradley, K., Brooks, R., Campos, S. M., Chino, J., et al (2022). "Clinical practice guidelines in oncology," in *Cervical cancer, version 1*.

Nishio, S., Mikami, Y., Tokunaga, H., Yaegashi, N., Satoh, T., Saito, M., et al. (2019). Analysis of gastric-type mucinous carcinoma of the uterine cervix - An aggressive tumor with a poor prognosis: a multi-institutional study. *Gynecol. Oncol.* 153 (1), 13–19. doi: 10.1016/j.ygyno.2019.01.022

Parazzini, F., La Vecchia, C., Negri, E., Fasoli, M., and Cecchetti, G. (1988). Risk factors for adenocarcinoma of the cervix: A case-control study. *Comp. Study. Br. J. Cancer.* 57 (2), 201–204. doi: 10.1038/bjc.1988.43

Park, K. J., Kiyokawa, T., Soslow, R. A., Lamb, C. A., Oliva, E., Zivanovic, O., et al. (2011). Unusual endocervical adenocarcinomas: An immunohistochemical analysis with molecular detection of human papillomavirus. *Am. J. Surg. Pathol.* 35, 633–646. doi: 10.1097/PAS.0b013e31821534b9

Peng, W. X., Kure, S., Ishino, K., Kurose, K., Yoneyama, K., Wada, R., et al. (2015). P16-positive continuous minimal deviation adenocarcinoma and gastric type adenocarcinoma in a patient with peutz-jeghers syndrome. *Int. J. Clin. Exp. Pathol.* 8 (5), 5877–5882.

Pirog, E. C., Lloveras, B., Molijn, A., Tous, S., Guimerà, N., Alejo, M., et al. (2014). HPV prevalence and genotypes in different histological subtypes of cervical adenocarcinoma, a worldwide analysis of 760 cases. *Mod. Pathol.* 27 (12), 1559–1567. doi: 10.1038/modpathol.2014.55

Schiffman, M., Doorbar, J., Wentzensen, N., de Sanjosé, S., Fakhry, C., Monk, B. J., et al. (2016). Carcinogenic human papillomavirus infection. *Nat. Rev. Dis. Primers.* 2, 16086. doi: 10.1038/nrdp.2016.86

Schiffman, M., Castle, P. E., Jeronimo, J., Rodriguez, A. C., Wacholder, S., et al. (2007). Human papillomavirus and cervical cancer. *Lancet* 370 (9590), 890–907.

Stolnicu, S., Barsan, I., Hoang, L., Patel, P., Terinte, C., Pesci, A., et al. (2018). International endocervical adenocarcinoma criteria and classification (IECC): A new pathogenetic classification for invasive adenocarcinomas of the endocervix. *Am. J. Surg. Pathol.* 42, 214–226. doi: 10.1097/PAS.0000000000000986

Tanaka, L. F., Schriefer, D., Radde, K., Schauberger, G., and Klug, S. J. (2021). Impact of opportunistic screening on squamous cell and adenocarcinoma of the cervix in Germany: A population-based case-control study. *PloS One* 16 (7), e0253801. doi: 10.1371/journal.pone.0253801

WHO (2020). *Global strategy to accelerate the elimination of cervical cancer as a public health problem* (Geneva: World Health Organization).

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 08 November 2022
DOI 10.3389/fcimb.2022.1007950

OPEN ACCESS

EDITED BY
Yongfen Xu,
Institut Pasteur of Shanghai (CAS), China

REVIEWED BY
Wencheng Zhu,
Chinese Academy of Sciences (CAS), China
Wuchang Zhang,
Shanghai Jiao Tong University, China
Minjun Yang,
Lund University, Sweden

*CORRESPONDENCE
Jiao Wei
drweijiao@hotmail.com
Honglei Zhang
alali626@163.com
Chengjun Wu
wcj5532@dlut.edu.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Virus and Host, a section of the journal Frontiers in Cellular and Infection Microbiology

RECEIVED 31 July 2022
ACCEPTED 10 October 2022
PUBLISHED 08 November 2022

CITATION
Zhou Q, Yuan O, Cui H, Hu T, Xiao GG, Wei J, Zhang H and Wu C (2022) Bioinformatic analysis identifies HPV-related tumor microenvironment remodeling prognostic biomarkers in head and neck squamous cell carcinoma.
Front. Cell. Infect. Microbiol. 12:1007950.
doi: 10.3389/fcimb.2022.1007950

Bioinformatic analysis identifies HPV-related tumor microenvironment remodeling prognostic biomarkers in head and neck squamous cell carcinoma

Qimin Zhou[1†], Ouyang Yuan[2†], Hongtu Cui[3†], Tao Hu[3], Gary Guishan Xiao[4], Jiao Wei[1*], Honglei Zhang[5*] and Chengjun Wu[1*]

[1]Department of Plastic and Reconstructive Surgery, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [2]Department of Oncology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [3]School of Biomedical Engineering, Dalian University of Technology, Dalian, China, [4]School of Pharmaceutical Science and Technology, Dalian University of Technology, Dalian, China, [5]Department of Otolaryngology Head and Neck Surgery, Air Force Medical Centre, People's Liberation Army (PLA), Beijing, China

Head and neck squamous cell carcinomas (HNSCCs) are highly aggressive tumors with rapid progression and poor prognosis. Human papillomavirus (HPV) infection has been identified as one of the most important carcinogens for HNSCC. As an early event in HNSCC, infection with HPV leads to altered immune profiles in the tumor microenvironment (TME). The TME plays a key role in the progression and transformation of HNSCC. However, the TME in HNSCC is a complex and heterogeneous mix of tumor cells, fibroblasts, different types of infiltrating immune cells, and extracellular matrix. Biomarkers relevant to the TME, and the biological role of these biomarkers, remain poorly understood. To this end, we performed comprehensive analysis of the RNA sequencing (RNA-Seq) data from tumor tissue of 502 patients with HNSCC and healthy tissue of 44 control samples. In total, we identified 4,237 differentially expressed genes, including 2,062 upregulated and 2,175 downregulated genes. Further in-depth bioinformatic analysis suggested 19 HNSCC tumor tissue-specific genes. In the subsequent analysis, we focused on the biomarker candidates shown to be significantly associated with unfavorable patient survival: *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, *TGFB1*, and *VEGFC*. We found that the expression of these genes was negatively regulated by DNA methylation. Strikingly, all of these potential biomarkers are profoundly involved in the activation of the epithelial–mesenchymal transition (EMT) pathway in HNSCCs. In addition, these targets were found to be positively correlated with the immune invasion levels of $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells, but negatively correlated with B-cell infiltration and $CD8^+$ T-cell invasion. Notably, our data showed that the expression levels

of *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, and *TGFB1* were significantly overexpressed in HPV-positive HNSCCs compared to normal controls, indicating the potential role of these biomarkers as transformation and/or malignant progression markers for HNSCCs in patients with HPV infection. Taken together, the results of our study propose *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, and *TGFB1* as potential prognostic biomarkers for HNSCCs, which might be involved in the HPV-related TME remodeling of HNSCC. Our findings provide important implications for the development and/or improvement of patient stratification and customized immunotherapies in HNSCC.

KEYWORDS

head and neck squamous cell carcinoma, HPV, biomarkers, tumor microenvironment, immune evasion, immunotherapy

Introduction

Head and neck cancers rank as the sixth most frequent malignancy worldwide, with more than 90% of this disease comprising head and neck squamous cell carcinoma (HNSCC), including different squamous epithelium-derived tumors of the oral cavity, oropharynx, larynx, and hypopharynx (Schmitz and Machiels, 2010; Bray et al., 2018; Chow, 2020). Recent statistics have pointed out that around 890,000 new cases and more than 450,000 deaths from HNSCC occur worldwide per year, accounting for 1%–2% of all malignant tumors, which is projected to exceed 30% by the year 2030 (Ferlay et al., 2019).

Aside from alcohol consumption and smoking, infection with human papillomavirus (HPV) is an increasingly common risk factor for HNSCC. HPV infection is associated with most oropharyngeal cancers (>70%) and a small minority of cancers at other anatomical sites in the head and neck (Leemans et al., 2018; Johnson et al., 2020; Lechner et al., 2022). Depending on the HPV infection status, HNSCCs are subdivided into HPV-negative (HPV^-) and HPV-positive (HPV^+) types, which have been recognized as distinct entities due to these tumors displaying a plethora of molecular and clinicopathological differences (Leemans et al., 2018). According to the latest reports, HPV^+ HNSCCs represent at least 25% of the HNSCC cases worldwide (Lechner et al., 2022). Alarmingly, the global frequency of HPV^+ HNSCCs has increased dramatically during the last few decades (D'Souza et al., 2007). For instance, the epidemic proportions of HPV^+ oropharyngeal cancers have risen from 7.2% (1990–1994) to 32.7% (2010–2012), with a similar epidemic boost of HPV^+ tonsillar cancers from 25% (1993–1999) to 62% (2006–2011) also being reported (D'Souza et al., 2007; Nichols et al., 2013; Nichols et al., 2013).

Unfortunately, despite previous studies suggesting an indicative role of the HPV status in the prognosis of HNSCC patients, its eligibility for guiding specific treatments is, to date, unclear (Sun et al., 2021). Therefore, regardless of HPV^+ and HPV^- HNSCC patients presenting remarkable differences in characterization, they are still provided the same therapeutic strategies (Ghiani and Chiocca, 2022). Adopting a combination of surgery, chemotherapy, and radiotherapy for patients with different HPV status has resulted in unignorable side effects, as well as a tremendous economic burden to society (Schaaij-Visser et al., 2010; Braakhuis et al., 2012). As a result, a 5-year survival rate of only a less than 50% has been achieved (Schaaij-Visser et al., 2010; Braakhuis et al., 2012). Thus, it is a necessary and urgent requirement to identify new biomarkers for HNSCC that would benefit patient stratification and survival improvement.

During the past few years, immunotherapies have dramatically changed the treatment algorithm of various cancer types (Hamid et al., 2013; Powles et al., 2020). Recently, for recurrent and metastatic HNSCC patients, programmed death-1 (PD-1) targeted checkpoint inhibitors (CPIs) have been approved as the first- and second-line approaches (Ferris et al., 2016; Burtness et al., 2019; Cohen et al., 2019). However, the treatment effects of CPIs rely mainly on the infiltration, reinvigoration, and activation of tumor-oriented antigen-specific T cells (Gubin et al., 2014), and HPV-infected malignant cells can change the tumor microenvironment (TME), affecting antigen presentation and immune cell infiltration (particularly the $CD4^+$/$CD8^+$ T-cell ratio), which have a significant impact on HNSCC immunotherapy (Lechner et al., 2022). Unfortunately, only 15%–20% of patients with HNSCC ultimately benefit from CPI treatment, highlighting the need for effective biomarkers for patient selection for these interventions (Ferris et al., 2016; Burtness et al., 2019). Principally, understanding the biological functions and the underpinning mechanisms of these biomarkers could vastly improve the efficacy of immunotherapies (Chen et al., 2019).

Of note is that the combination of high-throughput sequencing technology with bioinformatic analysis provides an efficient approach for the development of biomarkers and the discovery of their underlying mechanisms in different cancers (Langfelder and Horvath, 2008). For instance, the weighted gene co-expression network analysis (WGCNA) algorithm, used for analyzing relationships between gene sets and clinical features, has been widely applied to identify targeted modules and hub genes in cancer networks, such as breast cancer and gastric cancer (Tian et al., 2020; Rezaei et al., 2022). In this study, the RNA sequencing (RNA-Seq) datasets of patients with HNSCC were analyzed and the key genes involved in disease occurrence and development identified by subsequent bioinformatic analysis. To unravel the clinical utility of those potential biomarkers, sufficient diagnostic and prognostic models have been constructed. To examine the underlying biological mechanisms of these potential biomarkers in HNSCC, their expression levels, methylation status, and the degree of immune infiltration were verified by multiple databases and experiments. This study reveals a novel panel of potential prognostic biomarkers that may be involved in HPV-related inflammatory TME remodeling and immune evasion. Our findings could provide a basis for further exploration of the clinical significance and molecular mechanism and for facilitating immunotherapy development for HNSCCs.

Materials and methods

Data acquisition and pre-processing

The methodological workflow of the current study is illustrated in Supplementary Figure S1, which included data acquisition, processing, and validation. The RNA-Seq data were downloaded from the cohort of The Cancer Genome Atlas (TCGA) Head and Neck Cancer (HNSC) in the UCSC database (https://xenabrowser.net/datapages/), including 502 tumor tissue samples (500 primary site tumors and 2 metastatic tumors) and 44 normal tissue samples. The referenced gene chip dataset GSE138206 (HG-U133_Plus_2; Affymetrix Human Genome U133 Plus 2.0 Array) was downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/), including six oral squamous cell carcinoma (OSCC) tumor tissue samples and six paired normal tissue samples. The filterByExpr function of the edgeR package was performed to eliminate low-expression samples from the RNA-Seq data (Robinson et al., 2010). The RMA algorithm of the Affy package was used for background correction and standardization (log2 transformation) of the chip datasets (Gautier et al., 2004). Raw data (counts) from the TCGA dataset were standardized by TPM (transcript per million) before WGCNA.

Identification of differentially expressed genes

This study used DESeq2, edgeR, and limma to screen the RNA-Seq data and microarray datasets and subsequently fish out the differentially expressed genes (DEGs) between the HNSCC and normal tissue samples (Love et al., 2014; Ritchie et al., 2015). A $|\log2\text{foldchange}| > 1$ and $p_{.\text{adjust}} < 0.01$ were adopted as the analysis thresholds.

Construction of the weighted gene co-expression network in HNSCC

Based on the variability of genes as determined using median absolute deviation (MAD), a weighted gene co-expression network was constructed, which included the top 5,000 genes selected using the R software "WGCNA" package (Langfelder and Horvath, 2008). Initially, a scale-free network was constructed based on the Pearson's correlation coefficients between genes and the selected appropriate thresholds. Furthermore, a topological overlapping matrix (TOM) was transformed *via* an adjacency matrix, and the system clustering tree was obtained by hierarchical clustering. Finally, gene modules were established using the dynamic pruning algorithm, and the correlation between these modules and the clinical features was determined.

Functional enrichment analysis

In order to further understand the function of the selected genes, the R software "ClusterProfilter" package was used for Gene Ontology (GO) enrichment analysis. In addition, Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was performed for the pathway enrichment analysis of DEGs and the modular genes in the co-expression network of HNSCC (Yu et al., 2012). A $p_{.\text{adjust}} < 0.01$ [false discovery rate (FDR) corrected] was taken as the threshold to determine significant GO biological processes and biologically important KEGG pathways. The oncogene analysis platform GSCAlite was used for further pathway enrichment analysis of the selected key genes (Liu et al., 2018).

Protein interaction network analysis, survival analysis, and screening of targeted genes

Protein–protein interaction (PPI) networks can facilitate the identification of hub genes in a network. In this study, the STRING database was referenced to analyze overlapping genes

of the DEGs and key module genes. Thereafter, the key candidate genes were segregated using the Degree, MCC (Maximal Clique Centrality), MNC (Maximum Neighborhood Component), and EPC (Edge Percolated Component) algorithms of the CytoHubba plug-in in Cytoscape software (Chin et al., 2014). Finally, the R software "Survival" package was utilized for survival analysis of the candidate genes, and a $p < 0.05$ was chosen as the cutoff for potential biomarkers with prognostic value in HNSCC.

Methylation of key genes and immune cell infiltration analysis

The DNA methylation data were downloaded from the head and neck cancer (HNSC) in the DiseaseMeth 3.0 database, including 530 tumor tissue samples and 50 normal tissue samples. The methylation of key genes was analyzed using the DiseaseMeth 2.0 database, in line with Pearson's correlation mapping generated by R software (Hinshaw and Shevde, 2019). Moreover, TIMER is a comprehensive resource database for systematic analysis of immune infiltrations, which applies RNA-Seq profile data to detect the immune cell infiltration in different tumor tissues (Li et al., 2017). Therefore, we employed this method to analyze the correlation between the expression levels of key genes and the abundance of immune infiltrates, including B cells, $CD8^+$ T cells, $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells, in HNSCC ($n = 522$). The gene expression levels representing tumor purity are shown at the far left of the panel.

Prognostic analysis of key genes in HNSCC

To analyze the prognostic value of the key genes in HNSCC, univariate and multivariate Cox proportional risk regression models were formulated. We then predicted the risk score for each HNSCC patient by stepwise multivariate Cox proportional risk regression analysis. The risk scoring model was constructed using the R software package "Survival" and the "SurvMiner" package (Zhang, 2016). The risk score was calculated as follows: expGene1 *… + expGenen * coefficient (expGene represents the expression level of each key gene, and the coefficient is the corresponding regression coefficient generated by the risk model). Consequently, an overall survival (OS) time curve was plotted, which was based on HNSCC patients with a high or a low risk according to the median risk score. In addition, receiver operating characteristic (ROC) curves for 1-, 3-, and 5-year survival were plotted to verify the specificity and sensitivity of the model (Heagerty et al., 2000). In the prognostic analysis, statistical significance was defined as $p < 0.05$.

Expression verification of key genes

Multiple databases were searched to verify the expression of the key genes. Of these, GEPIA (Gene Expression Profiling Interactive Analysis) is a cancer data mining tool with resources from the TCGA database, the GTEx (Genotype-Tissue Expression) database, and the HPA (Human Protein Atlas) database that include protein expression information in tumor tissue based on immunoassay techniques [Western blotting, immunohistochemistry (IHC), and immunofluorescence] (Tang et al., 2017). In this study, we verified the expression levels of the key genes in tumor and normal tissues *via* the GEO and GEPIA databases. The correlations between the expression levels of key genes and the prognosis of patients with HNSCC were measured in the GEPIA database, and the corresponding protein expression levels of the key genes were verified using the HPA database. A $p < 0.05$ was considered statistically significant.

Results

Identification of the differentially expressed genes and enriched biological processes in HNSCC

To identify the DEGs between HNSCC and normal tissues, the transcriptome data obtained from the TCGA database, which included 502 HNSCC and 44 normal tissues, were analyzed (Figure 1A). A total of 4,237 DEGs were screened out, including 2,062 upregulated and 2,175 downregulated genes (Figure 1B). The top 200 DEGs were used to generate a clustering heat map, in which clear segregation of the gene expression profiles was observed between HNSCC and normal samples (Figure 1C). Subsequently, we conducted GO enrichment analysis on the 2,062 upregulated genes to identify the enriched biological processes in HNSCC. Biological processes such as external encapsulating structure organization, external matrix, extracellular matrix, and extracellular structure were found to be significantly enriched in HNSCC (Figure 1D). These findings suggest that these DEGs may be involved in TME formation during the occurrence and development of HNSCC.

Establishing co-expression network modules for HNSCC occurrence and progression

Based on the median deviation of the gene expression levels in HNSCC, the top 5,000 genes were selected depending on their ranking of absolute median difference in gene expression levels for the construction of a weighted gene co-expression network. A total of 546 individuals were included to construct the

10.3389/fcimb.2022.1007950

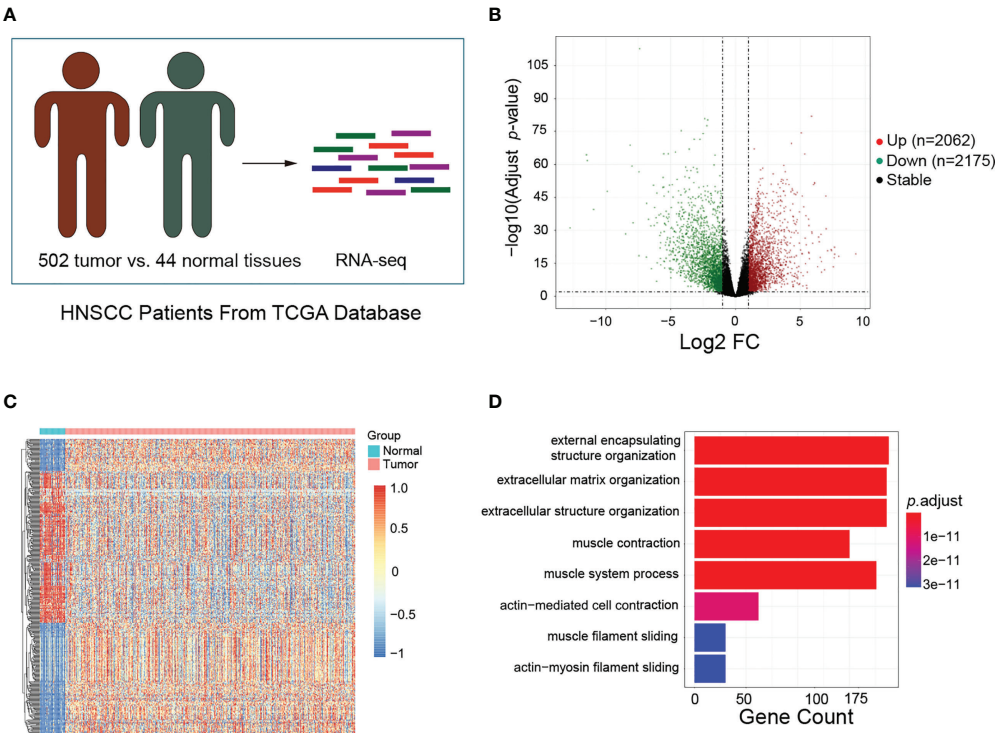


FIGURE 1
Identification of differentially expressed genes (DEGs) and enriched biological processes in head and neck squamous cell carcinoma (HNSCC). **(A)** Schematic of the patient samples included in the RNA sequencing (RNA-Seq). **(B)** Volcano plots showing the fold change (FC) and *p*-values for the comparisons of tumor tissues and normal tissues. Upregulated (log2FC ≥ 1, $p < 0.05$) and downregulated (log2FC ≤ −1, $p < 0.05$) genes are depicted in *red* and *green*, respectively. **(C)** Top 200 DEGs between tumor and normal tissues visualized as heatmaps from RNA-Seq. Data were *z*-score normalized. **(D)** Gene Ontology enrichment analysis for the top 200 DEGs between tumor and normal tissues. Only the significant enrichment values ($p < 0.05$) of the GO terms from the biological process category are listed.

hierarchical clustering map. As illustrated by the results in Supplementary Figure S2, there were no obvious outlier samples. With $\beta = 6$ as the soft threshold to establish the gene scale-free network, a strong topological fitting index was achieved with $R^2 > 0.9$ (Figures 2A, B). By adopting MEDissThres at 0.25 for merging similar gene modules, the co-expression network was successfully divided into 23 modules (Figure 2C). Finally, 19 gene modules were obtained and a module cluster dendrogram was drawn (Figure 2D), excluding the genes that could not be classified into other modules for subsequent analysis. Additionally, the association between these 19 modules and the clinical features of patients with HNSCC was assessed and a module–trait relationship diagram depicted (Figure 2E). The black module was positively associated with HNSCC, while the red module showed a negative association (correlation coefficients of 0.44 and −0.44 for the black and red modules, respectively, $p < 0.01$). Thus, the black and red modules were selected as the key modules in the occurrence and development of HNSCC. In the additional GO biological process enrichment analysis, the black module mainly included extracellular matrix tissue, extracellular structure tissue, external matrix tissue, and positive regulation of cell adhesion, while the red module corresponded to epidermal development, epidermal cell differentiation, skin development, and keratinocyte differentiation (Figures 2F, G). These results indicate that the HNSCC-associated DEGs were mainly concerned with TME remodeling, and overlapping results of the DEGs and key module genes were selected as candidate genes for further analysis (Figure 2H).

Identification and validation of the biomarkers in HNSCC

The above candidate genes were subsequently investigated using the STRING online database. A total of 275 candidates (143 upregulated and 132 downregulated genes) with interaction scores >0.4 were screened out to construct a PPI network, which included 275 nodes and 209,665 connections (Supplementary Figure S3). The PPI files were then imported into Cytoscape software, and the CytoHubba plug-in was used to further analyze potential biomarkers in these candidate genes. For reduction of

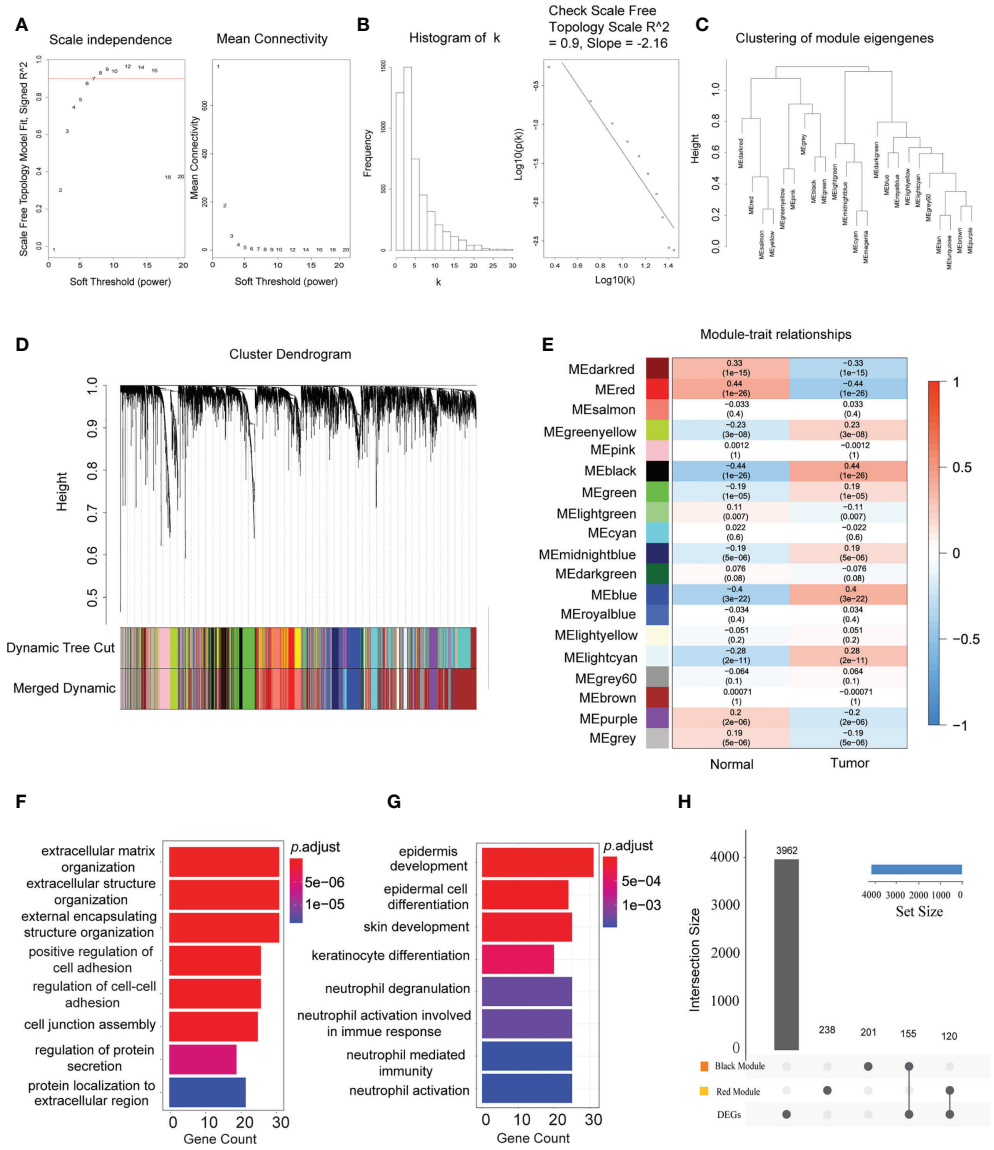


FIGURE 2
Establishment of the co-expression network modules for the occurrence and progression of head and neck squamous cell carcinoma (HNSCC). **(A)** Scale independence and mean connectivity of the network in different statistical thresholds. The panel displays the strong correlation of the statistical thresholds with a scale-free fit index. **(B)** Robust influence of the statistical threshold on mean connectivity. **(C)** Clustering of the eigengene modules by weighted gene co-expression network analysis (WGCNA) and their correlation with the occurrence and progression of HNSCC. **(D)** Cluster dendrogram based on the dissimilarity of the topological overlap matrix. *Different colours* correspond to the co-expression modules in HNSCC. **(E)** Heatmap of the correlation between the identified modules and clinical features (n = 44 in normal tissues and n = 502 in tumor tissues). The x-axis corresponds to the clinical features, while the y-axis represents the identified modules. The *colour scale* (*blue* to *red*) indicates correlation. Data are presented as the correlation coefficient in the *top row* and the p-value in the *bottom row in parentheses*. **(F)** Significant enriched Gene Ontology (GO) biological process terms for the HNSCC positively associated (*black*) module. **(G)** Significant enriched GO biological process terms for the HNSCC negatively associated (*red*) module. **(H)** Intersection of the differentially expressed genes in the different modules. *Bar charts* indicate the number of involved genes in the *black* or *red* modules.

error, the MCC, MNC, EPC, and Degree algorithms were employed. The top 30 candidates with the highest scores were eventually selected (Supplementary Table S1), and the overlapping nodes from the different methods used were identified as potential biomarkers of HNSCC. A total of 19 potential biomarkers were identified, namely, *MMP13*, *LCN2*, *COL4A5*, *ITGA3*, *CXCL8*, *TGFB1*, *ITGA5*, *IL1A*, *COL7A1*, *MMP3*, *SERPINE1*, *IL1B*, *SPP1*, *SERPINB2*, *PLAU*, *VEGFC*, *MMP1*, *COL17A1*, and *PLAUR* (Figure 3A). To verify the robustness of these screening results, an additional GO

10.3389/fcimb.2022.1007950

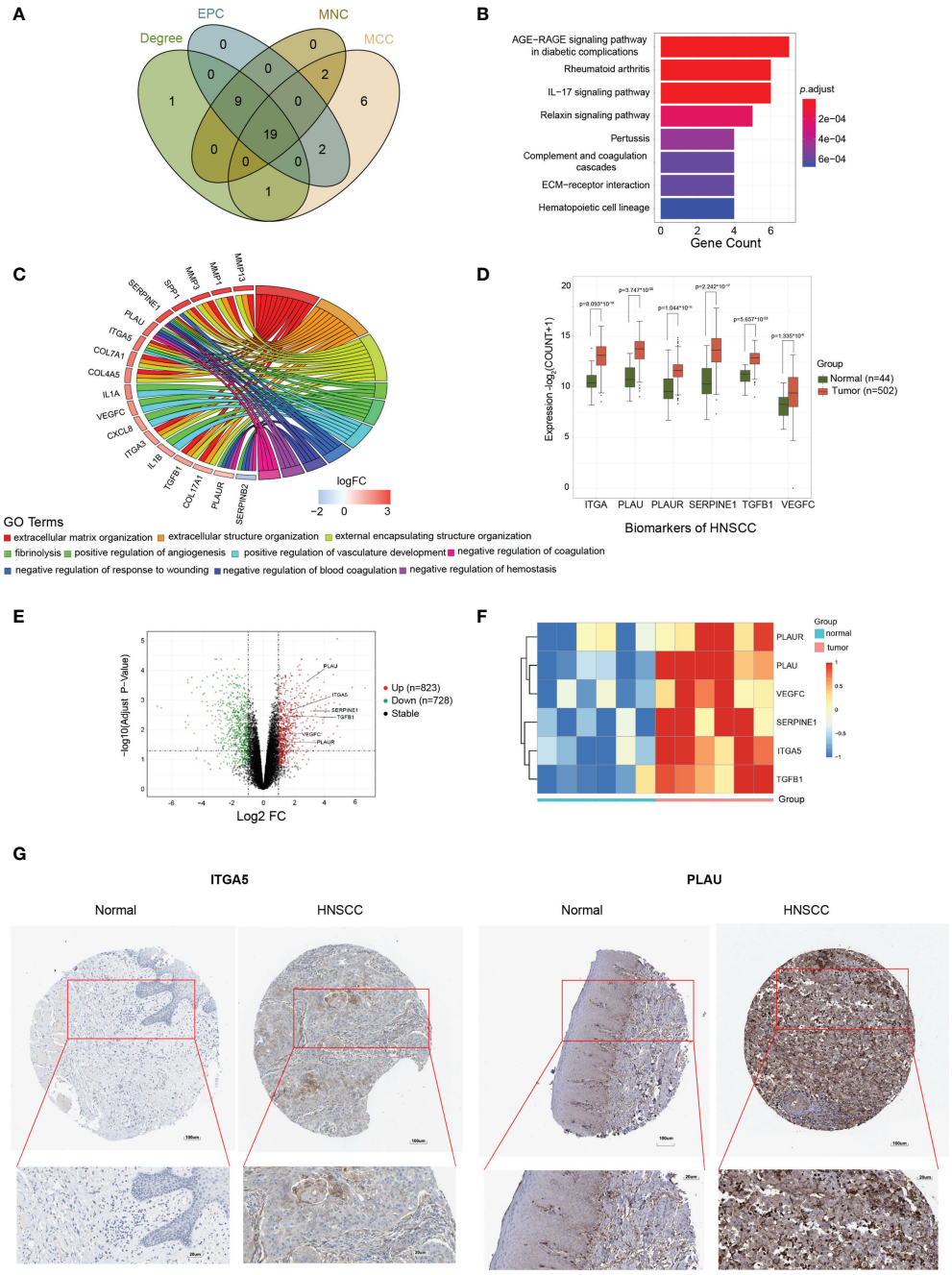


FIGURE 3
Identification and validation of the biomarkers for head and neck squamous cell carcinoma (HNSCC). **(A)** Venn diagram of the top 30 HNSCC-associated candidates selected by Cytoscape analysis, depending on the MCC (Maximal Clique Centrality), MNC (Maximum Neighborhood Component), EPC (Edge Percolated Component), and Degree algorithms. **(B)** Significant enriched Gene Ontology (GO) pathway signaling for the 19 potential biomarkers. **(C)** Significant enriched GO biological process terms for the 19 potential biomarkers. **(D)** Expression levels of *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, *TGFB1*, and *VEGFC* in tumor *versus* normal tissues. **(E)** Volcano plots showing the fold change (FC) and *p*-value of the biomarker candidates for the comparisons of tumor and normal tissues from the GSE138206 dataset. Upregulated (log2FC $\geq$ 2, $p < 0.05$) and downregulated (log2FC ≤ -2, $p < 0.05$) genes are depicted in *red* and *green*, respectively. **(F)** Expression patterns of *PLAUR*, *PLAU*, *VEGFC*, *SERPINE1*, *ITGA5*, and *TGFB1* in oral squamous cell carcinoma (OSCC) tumor ($n = 6$) and normal ($n = 6$) tissues visualized as heatmaps from the microarray. Data were *z*-score normalized. **(G)** Representative immunohistochemistry staining images of the ITGA5 and PLAU protein expression levels in tumor and normal tissues. Data were from the Human Protein Atlas database.

enrichment analysis was executed. The results showed that these biomarkers were enriched in the extracellular matrix, extracellular structure, and external matrix tissues in terms of biological processes (Figures 3B, C), which are consistent with the enrichment analysis results of the DEGs and the black module. Of these genes, *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, *TGFB1*, and *VEGFC* were significantly highly expressed in tumor compared to normal tissues (Figure 3D). We validated these potential biomarkers in an independent dataset from the GEO database, which recruited tumor tissues from OSCC that accounts for the high mortality and morbidity rates in HNSCC and paired normal tissues ($n = 6$ for each group) (Figure 3E). As expected, *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, *TGFB1*, and *VEGFC* were consistently overexpressed in tumor tissues compared to normal tissues ($\log2FC > |1|$ and $p \leq 0.01$) (Figure 3F). In addition, increased expressions of the PLAU and ITGA5 proteins in HNSCC tissues compared to normal tissues were observed in the immunohistochemical staining data found in the HPA database (Figure 3G). To address the localization of those markers in HNSCC, we explored the HPA database and analyzed the immunohistochemical staining data available for the other markers. The results showed that the ITGA5 and SERPINE1 proteins were mainly expressed in HNSCC tumor cells (with medium staining), while the PLAU protein was enriched in the extracellular matrix of HNSCC (with strong positive reaction). In contrast, the PLAUR and TGFB1 proteins were expressed in tumor cells, but only at a very weak level (data not shown). Unfortunately, there was no valid IHC imaging for the VEGFC protein in HNSCC. Overall, these results further confirmed the validity of the expression levels of the biomarker panel from the perspective of multiple databases.

Prognostic value of the biomarkers in HNSCC

To explore the clinical value of potential biomarkers, Kaplan–Meier analysis was conducted to compare the impact on OS between the high (above median) and the low (below median) expression level of each marker. The results highlighted that *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, *TGFB1*, and *VEGFC* presented significant differences on patients' survival time, revealing that a high expression, rather than a low expression, of these markers correlated with a significantly poorer survival of patients with HNSCC (Figure 4). Additionally, these results were verified in the GEPIA database (Supplementary Figures S4, S5). Therefore, these six tumor-related genes (*ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, *TGFB1*, and *VEGFC*) were considered as prognostic biomarkers for HNSCC.

Prognostic and prediction model construction of the biomarkers in HNSCC

To determine the impact of these biomarker candidates on the prognosis of patients, 457 patients were recruited for prognostic analysis. An OS curve was plotted for the high-risk and low-risk groups (Figure 5A), which showed that patients' survival time was significantly lower in the high-risk group than that in the low-risk group ($p < 0.001$). Furthermore, to validate the usefulness and exemplify the advantage of our prediction models in HNSCC patients' 1-, 3-, and 5-year outcomes, diagnostic ROC curves were plotted according to their risk scores. Surprisingly, the corresponding area under the curve (AUC) values were 0.644, 0.638, and 0.577, respectively (Figure 5B), which were significantly higher than those of the TNM clinical staging system ($p = 0.0016$) and the pathological tumor staging models ($p = 0.0056$) (Figures 5C, D, respectively). These results suggest that the current established patient score model has better performance in predicting the OS of patients with HNSCC. Finally, we also examined the distribution of the risk scores and OS of HNSCC patients, which also confirmed the accuracy of this model (Figures 5E, F).

The expression levels of the biomarker candidates are regulated by DNA methylation in HNSCC

To explore the potential mechanism regulating the expression of these biomarker candidates, we analyzed the methylation status of these encoding genes in HNSCC by applying data derived from the Illumina Human Methylation 450 platform. Interestingly, the results demonstrated that, with the exception of *TGFB1*, the DNA methylation levels of the CpG sites in the other five biomarker candidates in tumor tissues were significantly lower in HNSCC tissues (Figure 6A). On the other hand, the expression levels of these genes were negatively correlated with DNA methylation (Figure 6B), indicating a correlation between the increased gene expression and DNA hypomethylation.

Biomarkers are involved in TME remodeling and immune infiltration in HNSCC

Strikingly, all of these biomarkers are profoundly involved in the activation of the epithelial–mesenchymal transition (EMT) pathway (Figures 7A, B). These novel findings indicate that this

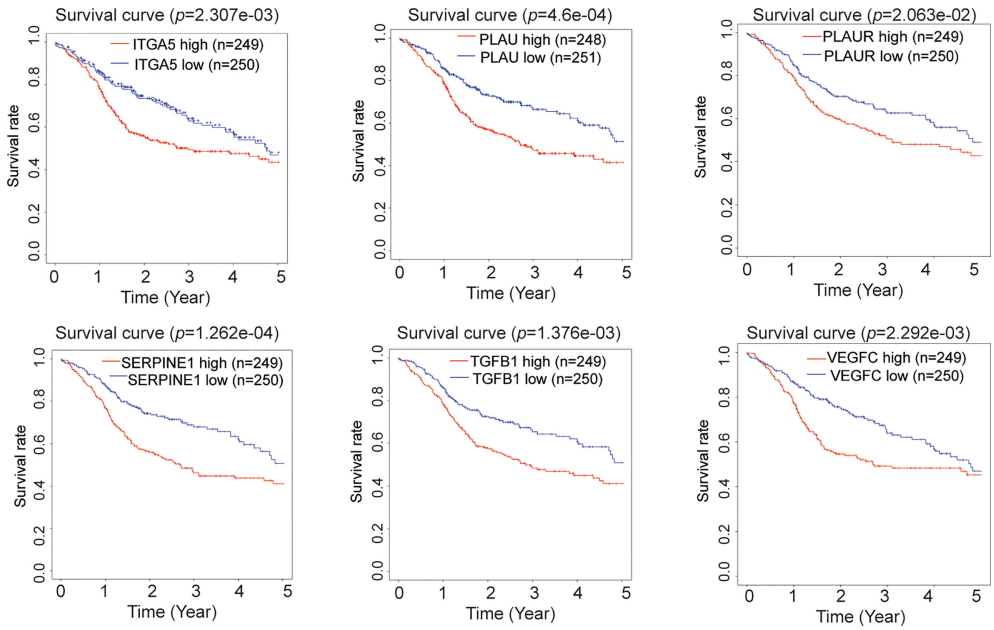


FIGURE 4
Prognostic value of the biomarkers for head and neck squamous cell carcinoma (HNSCC). Kaplan–Meier survival curves of patients with HNSCC based on the high or low expression levels (above or below median expression) of *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, *TGFB1*, and *VEGFC*.

panel of biomarkers plays a crucial role in TME remodeling and may also contribute to the immune evasion process in the occurrence and development of HNSCC. In order to characterize the underlying biological function of these markers in TME remodeling and immune infiltration, we performed the TIMER algorithm to analyze the correlation between their expression levels and the immune cell infiltration status in HNSCC. Our data showed that the expression levels of these biomarkers in HNSCC were positively correlated with the immune invasion levels of $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells, but negatively correlated with the infiltration of B cells. However, for $CD8^+$ T cells, with the exception of *PLAUR*, the expression levels of the other five genes were negatively correlated with the infiltrating levels (Figure 7C).

Biomarkers are involved in the transformation and progression of HNSCCs in patients with HPV infection

To elucidate the role of these potential biomarkers in HPV-related HNSCC, we analyzed the expression levels of 19 DEGs and 6 biomarker candidates in 34 HPV^+ HNSCC patients and 44 normal controls. A clear separation of the expression patterns of the 19 DEGs between the two groups (Figure 8A) was observed, in which 12 DEGs—*ITGA5*, *SPP1*, *PLAU*, *PLAUR*, *MMP13*, *LCN2*, *COL4A5*, *TGFB1*, *COL7A1*, *SERPINE1*, *SERPINB2*, and *MMP1*—presented statistical significance. More importantly, five out of six biomarker candidates, namely, *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, and *TGFB1*, were significantly overexpressed in HPV^+ HNSCCs compared to normal controls (Figure 8B). These results indicate that these biomarker candidates are potentially involved in HPV infection-related transformation and/or malignant progression of HNSCCs, which might be used as early indicators for the development and progression of HNSCCs in patients with HPV infection.

Discussion

Head and neck cancer is the sixth most common cancer type worldwide, and its incidence and mortality are on a rising trend (Bray et al., 2018; Chow, 2020). Due to the lack of an effective method for early diagnosis, most patients with HNSCC are in the mid to the advanced stages when they present in clinics (Mydlarz et al., 2010); moreover, precise prognosis and predictive biomarkers are also lacking at present (Pignon et al., 2009). Notably, the double-stranded DNA virus HPV infection is one of the major oncogenic factors for HNSCC, and the incidence of HPV-associated HNSCCs has significantly increased in the last few decades (Simard et al., 2012; Ellington et al., 2020). Despite their histological similarities, HPV^+ and HPV^- HNSCCs are now considered as two distinct cancers,

10.3389/fcimb.2022.1007950

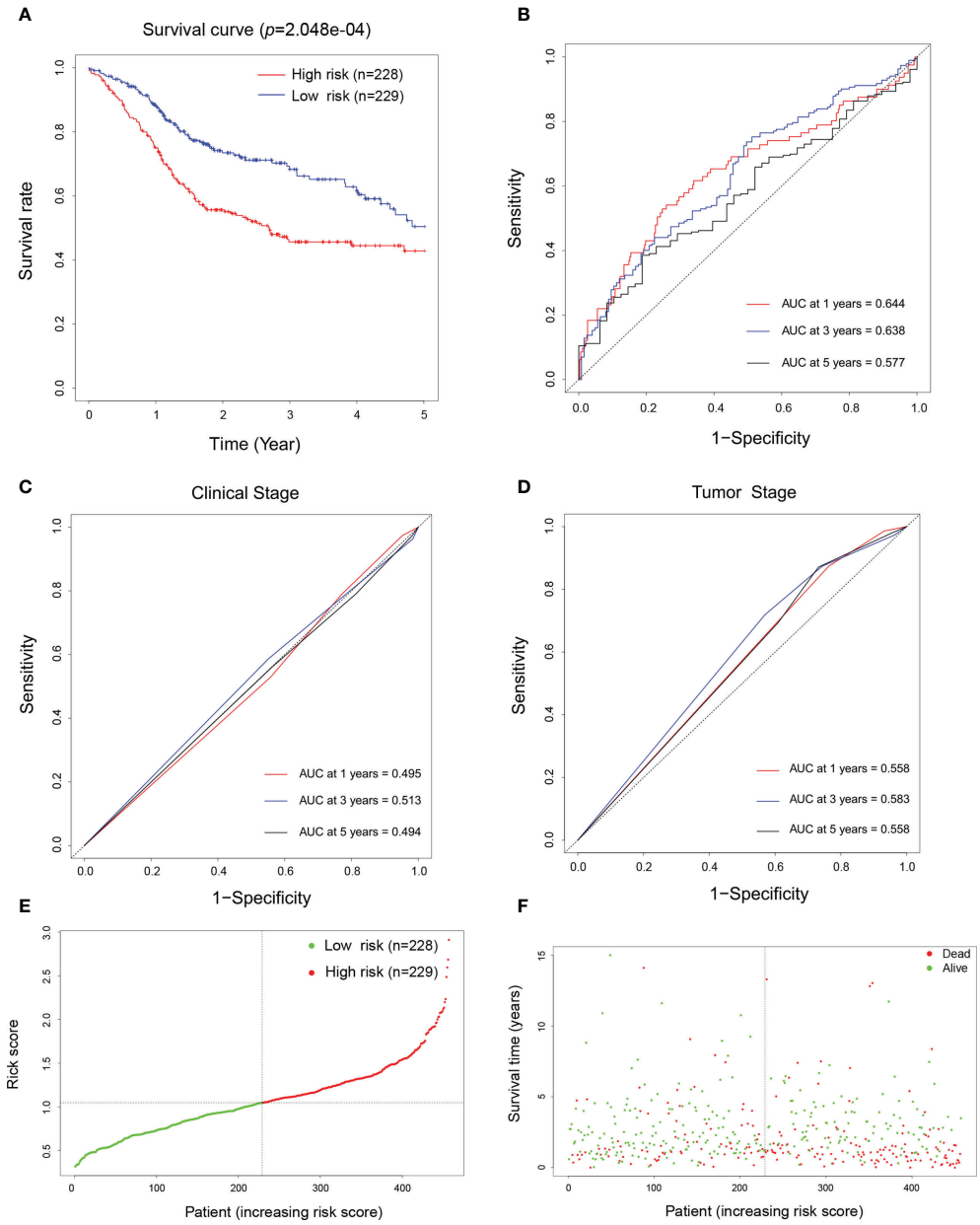


FIGURE 5
Prognostic and prediction model construction of the biomarkers for head and neck squamous cell carcinoma (HNSCC). **(A)** Overall survival curve of patients with HNSCC depending on their high (*n* = 228) and low (*n* = 229) risk scores defined by the prognostic model. **(B–D)** Diagnostic receiver operating characteristic (ROC) curves generated using the risk scores **(B)**, the TNM clinical staging system **(C)**, and pathological staging system **(D)** aiming to predict the 1-, 3-, and 5-year survival of patients with HNSCC. **(E)** Normalized distribution of the risk scores and the cutoff value for classifying the high- (*red curve*, *n* = 228) and low-risk (*green curve*, *n* = 229) groups. **(F)** Distribution of the survival status of individual HNSCC patients (*red dots* represent dead and *green dots* represent alive) according to their risk scores (*dashed line* represents the cutoff value).

owing to their substantial molecular and clinical differences (Leemans et al., 2018). However, due to a lack of effective biomarkers, the clinical management of these diseases is still the same, which leads to unfavorable outcomes and various side effects (Johnson et al., 2020; Lechner et al., 2022). Thus, exploration of clinically applicable biomarkers for HNSCC and an understanding of their underlying molecular mechanisms are urgently needed.

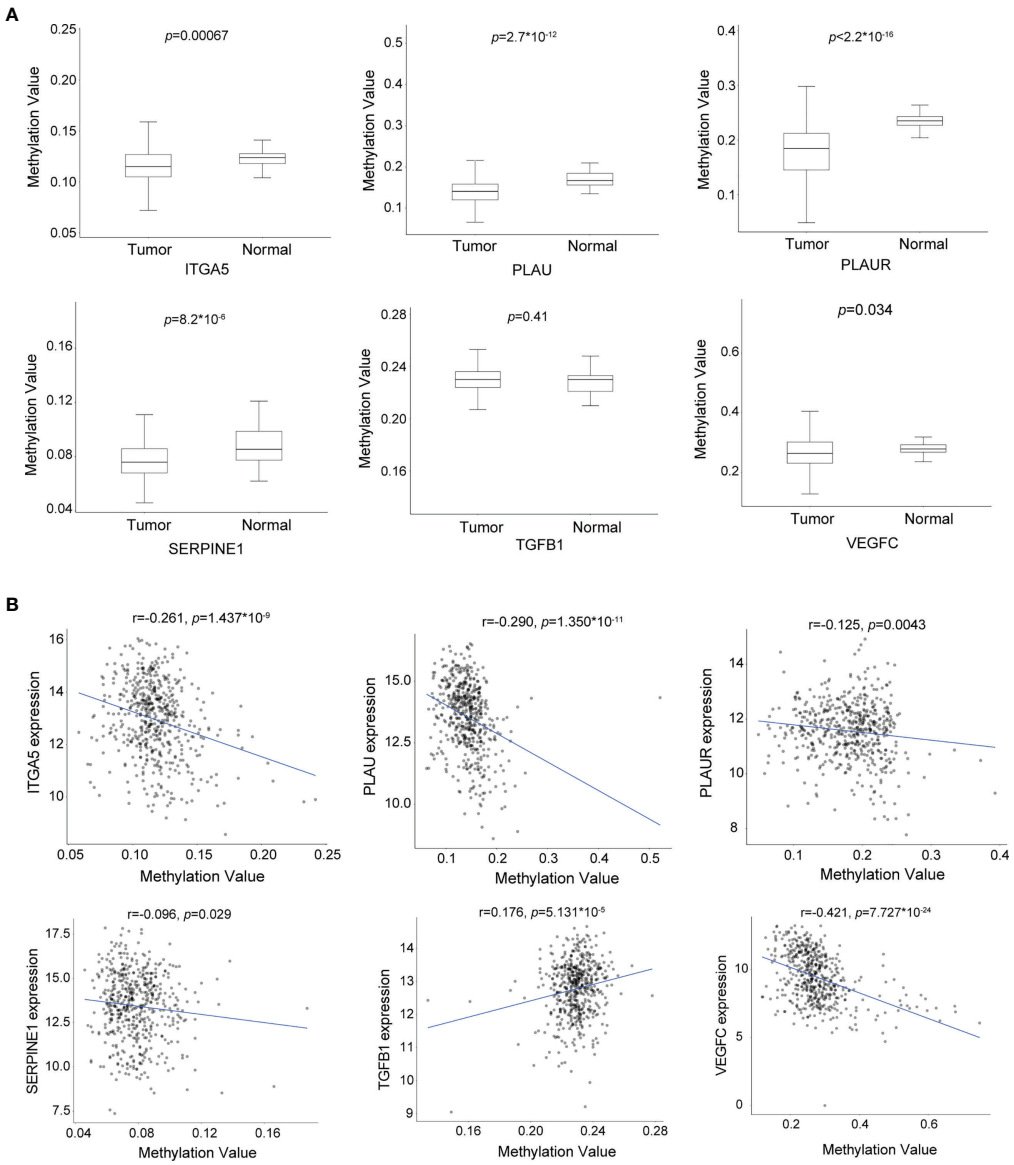


FIGURE 6
The expression levels of the biomarker candidates are regulated by DNA methylation in head and neck squamous cell carcinoma (HNSCC).
(A) DNA methylation levels of *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, *TGFB1*, and *VEGFC* in HNSCC (n = 530) and normal (n = 50) tissues.
(B) Correlation between the DNA methylation levels and gene expression levels of *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, *TGFB1*, and *VEGFC* (n = 502 and n = 22 for HNSCC and normal tissues, respectively). * means multiplication.

With the continuous development of high-throughput sequencing, gene microarray technologies, and bioinformatics methods, the association between genomic alterations and disease development can be examined at the molecular level, providing great advantages for studying the pathogenesis, diagnosis, and prognosis of diseases (Li et al., 2017; Tang et al., 2017; Liu et al., 2018; Rezaei et al., 2022). Nonetheless, previous studies have tended to focus on an individual gene associated with HNSCC, but cross-validations between different databases have been inadequate (Arroyo-Solera et al., 2019; Fan et al., 2019). In this study, based on HNSCC datasets from the TCGA database, we screened out a group of 19 candidate key genes using different bioinformatic analysis approaches. Subsequent PPI analysis was conducted to build interaction networks among them, followed by GO analysis of the key candidate genes, which showed that they were mainly related to the extracellular matrix, extracellular structure, and external matrix. These microenvironmental components are essential in regulating cellular function and maintaining normal homeostasis (Theocharis et al., 2016). Dysregulation of these

10.3389/fcimb.2022.1007950

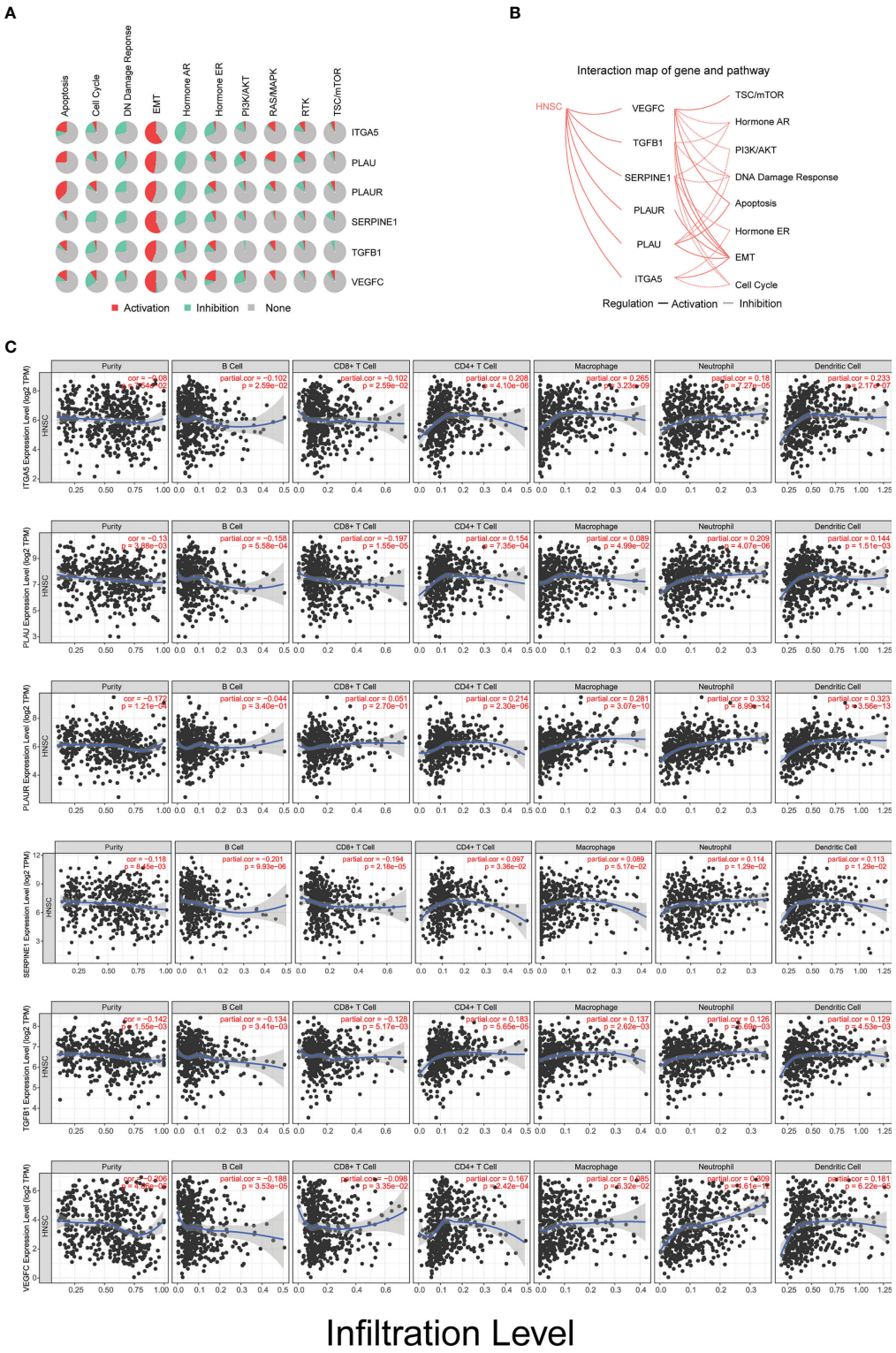


FIGURE 7
Biomarkers are involved in tumor microenvironment (TME) remodeling and immune infiltration in head and neck squamous cell carcinoma (HNSCC). **(A)** Activation/inhibition functionality of *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, *TGFB1*, and *VEGFC* on different biological signaling pathways. **(B)** Interaction map of *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, *TGFB1*, and *VEGFC* with potential biological signaling pathways in HNSCC. **(C)** Relationship between the infiltration levels of six types of immune cells and the expression level of the six biomarkers in patients with HNSCC (n = 522).

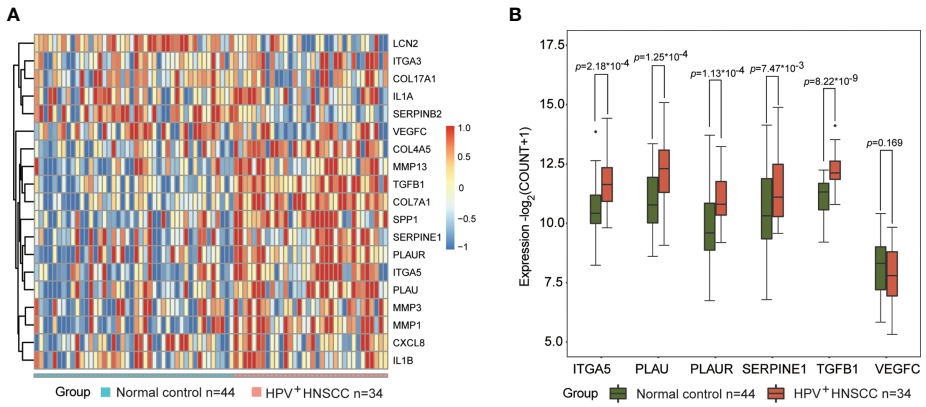


FIGURE 8
Biomarkers are involved in the transformation and progression of head and neck squamous cell carcinoma (HNSCCs) in patients with human papillomavirus (HPV) infection. **(A)** Expression patterns of 19 differentially expressed genes (DEGs) in HPV-positive HNSCCs (n = 34) and normal controls (n = 44) visualized as heatmaps. **(B)** Expression levels of *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, *TGFB1*, and *VEGFC* in HPV-positive HNSCCs (n = 34) and normal tissues (n = 44). * means multiplication.

components causes microenvironment remodeling, leading to abnormal cell proliferation, invasion, and differentiation, which play important roles during the occurrence and development of many cancers (Lu et al., 2011; Caon et al., 2020).

Indeed, there are two predominant reasons contributing to HPV$^+$ HNSCC development: i) genetic alterations of host cells are exerted by the integration of viral DNA into the host genome, which causes genomic instability, as well as a transcriptional shift by oncoviral proteins E6 and E7 that interacts and destabilizes a large number of host proteins (Leemans et al., 2018); ii) HPV infection breaks the dynamic and complex network of intercellular communication, which leads to an oncogenic TME remodeling (Gameiro et al., 2022). However, the key genes involved in the aforementioned processes and their biological functions are not well defined. Here, we finally identified five candidates—*ITGA5*, *TGFB1*, *PLAU*, *PLAUR*, and *SERPINE1*—as potential biomarkers for the prognosis and diagnosis of HPV-related HNSCC, which may play essential roles in the initiation and development of HNSCC. Additional pathway analysis highlighted that these biomarker candidates are particularly involved in EMT transformation, which is a cancer-specific biological process in which epithelial cells lose their cellular stickiness and polarity and transform from a non-motile epithelial cell to a mesenchymal cell signature, acquiring characteristics of migration and invasion (Thiery, 2002; Tsai and Yang, 2013).

Of note is that the current study identified a novel panel of biomarker candidates for HNSCC diagnosis and prognosis, which have yet to be sophisticatedly studied in relation to EMT and immunomodulation, especially with HPV-related HNSCC. To the best of our knowledge, with the exception of *TGFB1*, which has been proposed to be correlated with the clinical outcomes of HPV$^+$ oropharynx squamous cell carcinoma patients after radiotherapy (Tao et al., 2018), none of the other genes have been previously reported with regard to HNSCC development and patient prognosis. Several studies have suggested that carcinoma cells activate SNAIL signaling and secrete large quantities of TGFB1 protein, together with other immunoregulatory cytokines and chemokines to accelerate EMT transformation (Wrzesinski et al., 2007; Teicher, 2007; Scheel et al., 2011), which act as immunosuppressive factors that induce regulatory T cells (Tregs) and attenuate the cytolytic activities of natural killer (NK) cells (Bellone et al., 1995; Viel et al., 2016). Thus, *TGFB1* blockades have been considered to have great promise for enhanced antitumor activity; hence, anti-*TGFB1* therapeutics, such as antisense oligonucleotide-, ligand-, and receptor-targeted neutralizing antibodies, and small molecule inhibitors were broadly tested in different cancer types (Akhurst and Hata, 2012; Colak and ten Dijke, 2017). Recently, a preclinical study has demonstrated that the combination of galunisertib (a small molecule inhibitor of TGFB1 protein receptor I kinase) with an anti-PD-L1 (programmed death-ligand 1) regime presented a thoroughly improved antitumor effect compared to either galunisertib or anti-PD-L1 monotherapy, which may be attributed to the induction of enhanced antitumor T-cell activities by galunisertib (Holmgaard et al., 2018). Therefore, examination of the potential function in TME remodeling and immunomodulation of the rest of the markers in the panel were of paramount interest.

Intriguingly, the expression levels of the biomarker candidates, including *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, and

VEGFC, were positively correlated with the infiltration of $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells, but negatively correlated with the infiltration of B cells in HNSCC. However, for the infiltrating levels of $CD8^+$ T cells, except for no significant association found for *PLAUR*, the other markers showed negative correlation. Markedly, in patients with HPV^+ HNSCC, lymphocyte infiltration is generally associated with improved prognosis (Solomon et al., 2018). After adjusting the effect of HPV infection, HNSCC tumors with higher T cell (especially for $CD8^+$ T cells) infiltrates were found to be associated with significantly better survival (Keck et al., 2015; Matlung et al., 2016). In contrast, the infiltration of $CD4^+$ T cells (particularly Tregs) plays an immunosuppressive role, which led to decreased $CD8^+$/Treg ratios within tumors in spite of the high $CD8^+$ T-cell infiltrates, and is correlated with poor prognosis of patients with HNSCC (Russell et al., 2013; Mandal et al., 2016). In addition, macrophages recruited to the immunosuppressive TME would transform into tumor-associated macrophages, which secrete the immunosuppressive cytokines, including IL-1beta, IL-6, IL-10, TGFB1, and the PD-L1 protein and other checkpoint ligands, which are correlated with unfavorable outcomes for patients with HNSCC (Marcus et al., 2004; Costa et al., 2013). Of note is that the TGFB1 protein expression has been shown to be positively correlated with HPV infection in cervical lesions (Iancu et al., 2010). To date, the association between other immune cell infiltration status and HNSCC patient survival remains to be determined, and further studies are warranted to understand the function of this panel. High-risk HPV infection has been established as a risk factor of developing HNSCC. In general, HPV-related HNSCC is believed to exhibit increased immune infiltrates and improved response to anti-PD-1 therapy compared to HPV-unrelated tumors (Partlová et al., 2015; Mandal et al., 2016; Chen et al., 2018; Wang et al., 2019). However, the composition of immune cells in the tumor immune microenvironment (TIME) is quite complex and heterogeneous. Thus, more specific biomarkers that can be used to characterize the TIME in HNSCCs are needed. More importantly, exploring the biological interactions of HPV infection and these biomarker candidates will help researchers have a better understanding of HPV-related HNSCC development. In this study, we identified five genes, namely, *ITGA5*, *TGFB1*, *PLAU*, *PLAUR*, and *SERPINE1*, as biomarker candidates involved in TIME remodeling for HPV-related HNSCCs. To this end, attenuating the expression of these identified candidate genes in tumors can potentially enhance the infiltration of $CD8^+$ T cells and B cells and decrease the infiltration of $CD4^+$ T cells and macrophages, which would be a promising direction for improving patient outcomes and developing targeted therapeutics for HNSCC.

In conclusion, with sophisticated bioinformatic analysis of transcriptome data from cancer and normal tissues, we have identified a novel panel of DEGs for HPV-related HNSCC, which included *ITGA5*, *PLAU*, *PLAUR*, *SERPINE1*, and *TGFB1*. Following verification and validation of the expression levels and clinical impact of these candidate genes in multiple databases, we propose that this panel of candidate genes could potentially be used as a biomarker panel in the prognostic evaluation and early detection of HPV-related HNSCC. Furthermore, these key players are likely to be involved in TME remodeling and immune evasion, which affect the response to immune therapy for HNSCC. Our work reveals predictors/biomarkers for improved immune therapy response in HPV-related HNSCC.

Data availability statement

Publicly available datasets were analyzed in this study. The data analyzed during this study are available at The Cancer Genome Atlas (TCGA; https://cancergenome.nih.gov/) and the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/).

Author contributions

QZ, CW, and HC conceived the original idea and conceptualized the study design, which were refined by OY, GX, and HZ. HC and TH collected the data for the study, which were analyzed and interpretated together with QZ, OY, and JW. QZ, OY, and HC drafted the manuscript, which was revised by JW, HZ, and CW. JW, HZ, and CW supervised the study. All authors critically reviewed and finalized the manuscript and gave their approval for submission.

Funding

This study was supported by grants from the National Natural Science Foundation of China (No.82103658 and No.82072287) and the Shanghai Pujiang Program (No.22PJD040).

Acknowledgments

We thank Peihan Guo for valuable scientific discussions. Aaron Chen-Xiao, from Peninsula High School at Rancho Palos Verdes, CA, USA, is acknowledged for data analysis. Andrew C. McCourt is acknowledged for revising the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer WuZ declared a shared affiliation with authors QZ and OY, to the handling editor at the time of review.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2022.1007950/full#supplementary-material

References

Akhurst, R. J., and Hata, A. (2012). Targeting the TGFbeta signalling pathway in disease. *Nat. Rev. Drug Discov* 11, 790–811. doi: 10.1038/nrd3810

Arroyo-Solera, I., Pavon, M. A., Leon, X., Lopez, M., Gallardo, A., Cespedes, M. V., et al. (2019). Effect of serpinE1 overexpression on the primary tumor and lymph node, and lung metastases in head and neck squamous cell carcinoma. *Head Neck* 41, 429–439. doi: 10.1002/hed.25437

Bellone, G., Aste-Amezaga, M., Trinchieri, G., and Rodeck, U. (1995). Regulation of NK cell functions by TGF-β1. *J. Immunol.* 155, 1066–1073.

Braakhuis, B. J., Brakenhoff, R. H., and Leemans, C. R. (2012). Treatment choice for locally advanced head and neck cancers on the basis of risk factors: biological risk factors. *Ann. Oncol.* 23 Suppl 10, x173–x177. doi: 10.1093/annonc/mds299

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492

Burtness, B., Harrington, K. J., Greil, R., Soulières, D., Tahara, M., de Castro, G.Jr., et al. (2019). Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-048): A randomised, open-label, phase 3 study. *Lancet* 394, 1915–1928. doi: 10.1016/S0140-6736(19)32591-7

Caon, I., Bartolini, B., Parnigoni, A., Carava, E., Moretto, P., Viola, M., et al. (2020). Revisiting the hallmarks of cancer: The role of hyaluronan. *Semin. Cancer Biol.* 62, 9–19. doi: 10.1016/j.semcancer.2019.07.007

Chen, Y. P., Wang, Y. Q., Lv, J. W., Li, Y. Q., Chua, M. L. K., Le, Q. T., et al. (2019). Identification and validation of novel microenvironment-based immune molecular subgroups of head and neck squamous cell carcinoma: implications for immunotherapy. *Ann. Oncol.* 30, 68–75. doi: 10.1093/annonc/mdy470

Chen, X., Yan, B., Lou, H., Shen, Z., Tong, F., Zhai, A., et al. (2018). Immunological network analysis in HPV associated head and neck squamous cancer and implications for disease prognosis. *Mol. Immunol.* 96, 28–36. doi: 10.1016/j.molimm.2018.02.005

Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., and Lin, C. Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* 8 Suppl 4, S11. doi: 10.1186/1752-0509-8-S4-S11

Chow, L. Q. M. (2020). Head and neck cancer. *N Engl. J. Med.* 382, 60–72. doi: 10.1056/NEJMra1715715

Cohen, E. E. W., Soulières, D., Le Tourneau, C., Dinis, J., Licitra, L., Ahn, M. J., et al. (2019). Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study. *Lancet* 393, 156–167. doi: 10.1016/S0140-6736(18)31999-8

Colak, S., and ten Dijke, P. (2017). Targeting TGF-β signaling in cancer. *Trends Cancer* 3, 56–71. doi: 10.1016/j.trecan.2016.11.008

Costa, N. L., Valadares, M. C., Souza, P. P., Mendonca, E. F., Oliveira, J. C., Silva, T. A., et al. (2013). Tumor-associated macrophages and the profile of inflammatory cytokines in oral squamous cell carcinoma. *Oral. Oncol.* 49, 216–223. doi: 10.1016/j.oraloncology.2012.09.012

D'Souza, G., Kreimer, A. R., Viscidi, R., Pawlita, M., Fakhry, C., Koch, W. M., et al. (2007). Case-control study of human papillomavirus and oropharyngeal cancer. *N Engl. J. Med.* 356, 1944–1956. doi: 10.1056/NEJMoa065497

Ellington, T. D., Henley, S. J., Senkomago, V., O'Neil, M. E., Wilson, R. J., Singh, S., et al. (2020). Trends in incidence of cancers of the oral cavity and pharynx - united states 2007-2016. *MMWR Morb Mortal Wkly Rep.* 69, 433–438. doi: 10.15585/mmwr.mm6915a1

Fan, Q. C., Tian, H., Wang, Y., and Liu, X. B. (2019). Integrin-alpha5 promoted the progression of oral squamous cell carcinoma and modulated PI3K/AKT signaling pathway. *Arch. Oral. Biol.* 101, 85–91. doi: 10.1016/j.archoralbio.2019.03.007

Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D. M., Pineros, M., et al. (2019). Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int. J. Cancer* 144, 1941–1953. doi: 10.1002/ijc.31937

Ferris, R. L., Blumenschein, G.Jr., Fayette, J., Guigay, J., Colevas, A. D., Licitra, L., et al. (2016). Nivolumab for recurrent squamous-cell carcinoma of the head and neck. *N Engl. J. Med.* 375, 1856–1867. doi: 10.1056/NEJMoa1602252

Gameiro, S. F., Evans, A. M., and Mymryk, J. S. (2022). The tumor immune microenvironments of HPV(+) and HPV(-) head and neck cancers. *WIREs Mech. Dis.* 14, e1539. doi: 10.1002/wsbm.1539

Gautier, L., Cope, L., Bolstad, B. M., and Irizarry, R. A. (2004). Affy–analysis of affymetrix GeneChip data at the probe level. *Bioinformatics* 20, 307–315. doi: 10.1093/bioinformatics/btg405

Ghiani, L., and Chiocca, S. (2022). High risk-human papillomavirus in HNSCC: Present and future challenges for epigenetic therapies. *Int. J. Mol. Sci.* 23, 3483. doi: 10.3390/ijms23073483

Gubin, M. M., Zhang, X., Schuster, H., Caron, E., Ward, J. P., Noguchi, T., et al. (2014). Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* 515, 577–581. doi: 10.1038/nature13988

Hamid, O., Robert, C., Daud, A., Hodi, F. S., Hwu, W. J., Kefford, R., et al. (2013). Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. *N Engl. J. Med.* 369, 134–144. doi: 10.1056/NEJMoa1305133

Heagerty, P. J., Lumley, T., and Pepe, M. S. (2000). Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 56, 337–344. doi: 10.1111/j.0006-341X.2000.00337.x

Hinshaw, D. C., and Shevde, L. A. (2019). The tumor microenvironment innately modulates cancer progression. *Cancer Res.* 79, 4557–4566. doi: 10.1158/0008-5472.CAN-18-3962

Holmgaard, R. B., Schaer, D. A., Li, Y., Castaneda, S. P., Murphy, M. Y., Xu, X., et al. (2018). Targeting the TGFβ pathway with galunisertib, a TGFβRI small molecule inhibitor, promotes anti-tumor immunity leading to durable, complete responses, as monotherapy and in combination with checkpoint blockade. *J. ImmunoTher. Cancer* 6, 1–15. doi: 10.1186/s40425-018-0356-4

Iancu, I. V., Botezatu, A., Goia-Rusanu, C., Stănescu, A., Huică, I., Nistor, E., et al. (2010). TGF-beta signalling pathway factors in HPV-induced cervical lesions. *Roum Arch. Microbiol. Immunol.* 69, 113–118.

Johnson, D. E., Burtness, B., Leemans, C. R., Lui, V. W. Y., Bauman, J. E., and Grandis, J. R. (2020). Head and neck squamous cell carcinoma. *Nat. Rev. Dis. Primers* 6, 92. doi: 10.1038/s41572-020-00224-3

Keck, M. K., Zuo, Z., Khattri, A., Stricker, T. P., Brown, C. D., Imanguli, M., et al. (2015). Integrative analysis of head and neck cancer identifies two biologically distinct HPV and three non-HPV subtypes. *Clin. Cancer Res.* 21, 870–881. doi: 10.1158/1078-0432.CCR-14-2481

Langfelder, P., and Horvath, S. (2008). WGCNA: an r package for weighted correlation network analysis. *BMC Bioinf.* 9, 559. doi: 10.1186/1471-2105-9-559

Lechner, M., Liu, J., Masterson, L., and Fenton, T. R. (2022). HPV-associated oropharyngeal cancer: epidemiology, molecular biology and clinical management. *Nat. Rev. Clin. Oncol.* 19, 306–327. doi: 10.1038/s41571-022-00603-7

Leemans, C. R., Snijders, P. J. F., and Brakenhoff, R. H. (2018). The molecular landscape of head and neck cancer. *Nat. Rev. Cancer* 18, 269–282. doi: 10.1038/nrc.2018.11

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 77, e108–e110. doi: 10.1158/0008-5472.CAN-17-0307

Liu, C. J., Hu, F. F., Xia, M. X., Han, L., Zhang, Q., and Guo, A. Y. (2018). GSCALite: a web server for gene set cancer analysis. *Bioinformatics* 34, 3771–3772. doi: 10.1093/bioinformatics/bty411

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8

Lu, P., Takai, K., Weaver, V. M., and Werb, Z. (2011). Extracellular matrix degradation and remodeling in development and disease. *Cold Spring Harb. Perspect. Biol.* 3, a005058. doi: 10.1101/cshperspect.a005058

Mandal, R., Şenbabaoğlu, Y., Desrichard, A., Havel, J. J., Dalin, M. G., Riaz, N., et al. (2016). The head and neck cancer immune landscape and its immunotherapeutic implications. *JCI Insight* 1, e89829–e89829. doi: 10.1172/jci.insight.89829

Marcus, B., Arenberg, D., Lee, J., Kleer, C., Chepeha, D. B., Schmalbach, C. E., et al. (2004). Prognostic factors in oral cavity and oropharyngeal squamous cell carcinoma. *Cancer* 101, 2779–2787. doi: 10.1002/cncr.20701

Matlung, S. E., Wilhelmina Van Kempen, P. M., Bovenschen, N., Baarle, D. V., and Willems, S. M. (2016). Differences in T-cell infiltrates and survival between HPV+ and HPV- oropharyngeal squamous cell carcinoma. *Future Sci. OA* 2, FSO88. doi: 10.4155/fso.15.88

Mydlarz, W. K., Hennessey, P. T., and Califano, J. A. (2010). Advances and perspectives in the molecular diagnosis of head and neck cancer. *Expert Opin. Med. Diagn.* 4, 53–65. doi: 10.1517/17530050903338068

Nichols, A. C., Dhaliwal, S. S., Palma, D. A., Basmaji, J., Chapeskie, C., Dowthwaite, S., et al. (2013). Does HPV type affect outcome in oropharyngeal cancer? *J. Otolaryngol. - Head Neck Surg.* 42, 9. doi: 10.1186/1916-0216-42-9

Nichols, A. C., Palma, D. A., Dhaliwal, S. S., Tan, S., Theuer, J., Chow, W., et al. (2013). The epidemic of human papillomavirus and oropharyngeal cancer in a Canadian population. *Curr. Oncol.* 20, 212–219. doi: 10.3747/co.20.1375

Partlová, S., Bouček, J., Kloudová, K., Lukešová, E., Zábrodský, M., Grega, M., et al. (2015). Špíšek, distinct patterns of intratumoral immune cell infiltrates in patients with HPV-associated compared to non-virally induced head and neck squamous cell carcinoma. *Oncoimmunology* 4, e965570. doi: 10.4161/21624011.2014.965570

Pignon, J. P., le Maître, A., Maillard, E., and Bourhis, J. (2009). Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. *Radiother Oncol.* 92, 4–14. doi: 10.1016/j.radonc.2009.04.014

Powles, T., Park, S. H., Voog, E., Caserta, C., Valderrama, B. P., Gurney, H., et al. (2020). Avelumab maintenance therapy for advanced or metastatic urothelial carcinoma. *N Engl. J. Med.* 383, 1218–1230. doi: 10.1056/NEJMoa2002788

Rezaei, Z., Ranjbaran, J., Safarpour, H., Nomiri, S., Salmani, F., Chamani, E., et al. (2022). Identification of early diagnostic biomarkers *via* WGCNA in gastric cancer. *BioMed. Pharmacother.* 145, 112477. doi: 10.1016/j.biopha.2021.112477

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. doi: 10.1093/nar/gkv007

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Russell, S., Angell, T., Lechner, M., Liebertz, D., Correa, A., Sinha, U., et al. (2013). Immune cell infiltration patterns and survival in head and neck squamous cell carcinoma. *Head Neck Oncol.* 5, 24.

Schaaij-Visser, T. B., Brakenhoff, R. H., Leemans, C. R., Heck, A. J., and Slijper, M. (2010). Protein biomarker discovery for head and neck cancer. *J. Proteomics* 73, 1790–1803. doi: 10.1016/j.jprot.2010.01.013

Scheel, C., Eaton, E. N., Li, S. H. J., Chaffer, C. L., Reinhardt, F., Kah, K. J., et al. (2011). Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell* 145, 926–940. doi: 10.1016/j.cell.2011.04.029

Schmitz, S., and Machiels, J. P. (2010). Molecular biology of squamous cell carcinoma of the head and neck: relevance and therapeutic implications. *Expert Rev. Anticancer Ther.* 10, 1471–1484. doi: 10.1586/era.10.115

Simard, E. P., Ward, E. M., Siegel, R., and Jemal, A. (2012). Cancers with increasing incidence trends in the united states: 1999 through 2008. *CA Cancer J. Clin.* 62, 118–128. doi: 10.3322/caac.20141

Solomon, B., Young, R. J., and Rischin, D. (2018). Head and neck squamous cell carcinoma: Genomics and emerging biomarkers for immunomodulatory cancer treatments. *Semin. Cancer Biol.* 52, 228–240. doi: 10.1016/j.semcancer.2018.01.008

Sun, Y., Wang, Z., Qiu, S., and Wang, R. (2021). Therapeutic strategies of different HPV status in head and neck squamous cell carcinoma. *Int. J. Biol. Sci.* 17, 1104–1118. doi: 10.7150/ijbs.58077

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 45, W98–W102. doi: 10.1093/nar/gkx247

Tao, Y., Sturgis, E. M., Huang, Z., Wang, Y., Wei, P., Wang, J. R., et al. (2018). TGFβ1 genetic variants predict clinical outcomes of HPV-positive oropharyngeal cancer patients after definitive radiotherapy. *Clin. Cancer Res.* 24, 2225–2233. doi: 10.1158/1078-0432.CCR-17-1904

Teicher, B. A. (2007). Transforming growth factor-β and the immune response to malignant disease. *Clin. Cancer Res.* 13, 6247–6251. doi: 10.1158/1078-0432.CCR-07-1654

Theocharis, A. D., Skandalis, S. S., Gialeli, C., and Karamanos, N. K. (2016). Extracellular matrix structure. *Adv. Drug Delivery Rev.* 97, 4–27. doi: 10.1016/j.addr.2015.11.001

Thiery, J. P. (2002). Epithelial-mesenchymal transitions in tumour progression. *Nat. Rev. Cancer* 2, 442–454. doi: 10.1038/nrc822

Tian, Z., He, W., Tang, J., Liao, X., Yang, Q., Wu, Y., et al. (2020). Identification of important modules and biomarkers in breast cancer based on WGCNA. *Onco Targets Ther.* 13, 6805–6817. doi: 10.2147/OTT.S258439

Tsai, J. H., and Yang, J. (2013). Epithelial-mesenchymal plasticity in carcinoma metastasis. *Genes Dev.* 27, 2192–2206. doi: 10.1101/gad.225334.113

Viel, S., Marçais, A., Guimaraes, F. S. F., Loftus, R., Rabilloud, J., Grau, M., et al. (2016). TGF-β inhibits the activation and functions of NK cells by repressing the mTOR pathway. *Sci. Signaling* 9, ra19. doi: 10.1126/scisignal.aad1884

Wang, J., Sun, H., Zeng, Q., Guo, X.-J., Wang, H., Liu, H.-H., et al. (2019). HPV-positive status associated with inflamed immune microenvironment and improved response to anti-PD-1 therapy in head and neck squamous cell carcinoma. *Sci. Rep.* 9, 1–10. doi: 10.1038/s41598-019-49771-0

Wrzesinski, S. H., Wan, Y. Y., and Flavell, R. A. (2007). Transforming growth factor-β and the immune response: Implications for anticancer therapy. *Clin. Cancer Res.* 13, 5262–5270. doi: 10.1158/1078-0432.CCR-07-1157

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an r package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

Zhang, Z. (2016). Semi-parametric regression model for survival data: graphical visualization with r. *Ann. Transl. Med.* 4, 461. doi: 10.21037/atm.2016.08.61

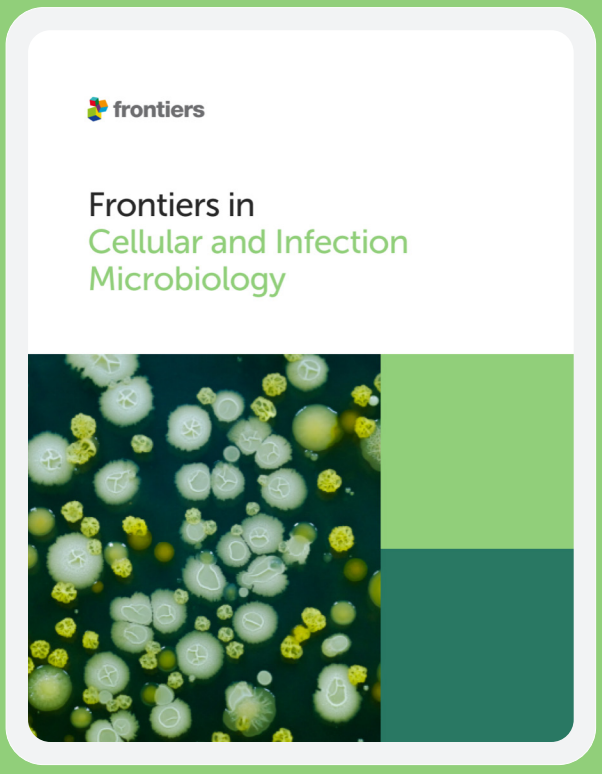
Frontiers in
Cellular and Infection Microbiology

Investigates how microorganisms interact with their hosts

Explores bacteria, fungi, parasites, viruses, endosymbionts, prions and all microbial pathogens as well as the microbiota and its effect on health and disease in various hosts.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact