Am J Transl Res 2022;14(12):8843-8861
www.ajtr.org /ISSN:1943-8141/AJTR0145356

Original Article

Deciphering cervical cancer-associated biomarkers by integrated multi-omics approach

Xuhong Li[1], Mostafa A Abdel-Maksoud[2], Iqra Iqbal[3], Ayman Mubarak[2], Mohamed A Farrag[2], Muhammad Haris[4], Sumaiah Alghamdi[2], Qurat Ul Ain[5], Sally Almekhlafi[2]

[1]Department of Gynaecology and Obstetrics, Shanghai Eighth People's Hospital, Shanghai, China; [2]Department of Botany and Microbiology, College of Science, King Saud University, P.O. 2455, Riyadh 11451, Saudi Arabia; [3]Azra Naheed Medical College, Lahore, Pakistan; [4]Department of Anatomy, Institute of Basic Medical Sciences, Khyber Medical University, Peshawar, Pakistan; [5]Anhui Provincial Hospital, Division of Life Science and Medicine, University of Science and Technology China, Hefei, Anhui, China

Received July 15, 2022; Accepted November 13, 2022; Epub December 15, 2022; Published December 30, 2022

Abstract: Objectives: Cervical Squamous Cell Carcinoma (CESC) is one of the most fatal female malignancies, and the underlying molecular mechanisms governing this disease have not been fully explored. In this research, we planned to conduct the analysis of Gene Expression Omnibus (GEO) cervical squamous cell carcinoma microarray datasets by a detailed *in silico* approach and to explore some novel biomarkers of CESC. Methods: The top commonly differentially expressed genes (DEGs) from the GSE138080 and GSE113942 datasets were analyzed by Limma package-based GEO2R tool. The protein-protein interaction (PPI) network of the DEGs was drawn through Search Tool for the Retrieval of Interacting Genes (STRING), and top 6 hub genes were obtained from Cytoscape. Expression analysis and validation of hub genes expression in CESC samples and cell lines were done using UALCAN, OncoDB, GENT2, and HPA. Additionally, cBioPortal, Gene set enrichment analysis (GSEA) tool, Kaplan-Meier (KM) plotter, ShinyGO, and DGIdb databases were also used to check some important values of hub genes in CESC. Results: Out of 79 DEGs, the minichromosome maintenance complex component 4 (MCM4), nucleolar and spindle-associated protein 1 (NUSAP1), cell division cycle associated 5 (CDCA5), cell division cycle 45 (CDC45), denticleless E3 ubiquitin protein ligase homolog (DTL), and chromatin licensing and DNA replication factor 1 (CDT1) genes were regarded as hub genes in CESC. Further analysis revealed that the expressions of all these hub genes were significantly elevated in CESC cell lines and samples of diverse clinical attributes. In this study, we also documented some important correlations between hub genes and some other diverse measures, including DNA methylation, genetic alterations, and Overall Survival (OS). Last, we also identify hub genes associated ceRNA network and 31 important chemotherapeutic drugs. Conclusion: Through detailed *in silico* methodology, we identified 6 hub genes, including MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1, which are likely to be associated with CESC development and diagnosis.

Keywords: Cervical cancer, hub genes, biomarkers, chemotherapeutic drugs

Introduction

Cervical cancer is the 4th most common female malignancy around the globe [1, 2]. The early diagnosis and prognosis of cervical cancer can enable a 5-year survival rate of 90%. However, the 5-year survival rate of patients with advanced cervical cancer, especially those with metastasis, has notably decreased [3, 4]. In addition, traditional treatment methods for cervical cancer continue to have many drawbacks. For instance, surgical treatment of cervical cancer is possible only if the patient is diagnosed early, and postoperative radiotherapy for cervical cancer can cause irreversible damage to the healthy cells of the ovaries and uterus [5, 6]. Therefore, emerging therapeutic options for cervical cancer such as targeted and immunotherapy are getting attention in cervical cancer treatment research [3]. Moreover, sensitive and reliable molecular biomarkers and therapeutic targets for cervical cancer have yet to be discovered.

The exploration of tumor-associated genes is a key step for studying tumor pathogenesis and

designing appropriate treatment strategies [7]. Exploring cervical cancer at the molecular level can thus help to show its etiology and identify possible sensitive and reliable molecular biomarkers for its diagnosis and treatment [8]. Currently, high throughput technology is the most reliable method for studying human genomics, and in this regard, recently, RNA-sequencing (RNA-seq) and microarray technologies have been widely utilized worldwide for discovering molecular biomarkers and underlying mechanisms in cancer studies [9, 10]. Open access to microarray and RNA-sequencing-based cancer expression data stored in the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases allows researchers to easily compare cancer patients' expression profiles with those of normal controls to discover novel molecular biomarkers and therapeutic targets [11, 12].

In cervical cancer research, the expression profiles of cervical cancer patients and healthy individuals from the GEO database have been previously explored using bioinformatic approaches to identify molecular biomarkers and underlying signaling pathways of cervical cancer pathophysiology [13, 14]. However, available literature is limited. Therefore, further in-depth analysis of the GEO cervical cancer expression data may help explore sensitive and novel biomarkers of tumor development and progression.

In this study, two microarray datasets (GSE138080 and GSE113942) were integrated from the GEO database and further screened by *in silico* methodology to identify crucial hub genes as molecular biomarkers, which may provide new insights into cervical cancer pathophysiology and treatment.

Materials and methods

Microarray datasets

By searching the GEO database with the "cervical cancer" keyword, two microarray expression datasets (GSE138080 and GSE113942) were acquired for analysis in the current research. The GSE138080 dataset consisted of 10 normal cervical tissue samples and 10 cervical cancer tissue samples, whereas the GSE113942 dataset included 24 normal cervical tissue samples and 28 cervical cancer tissue samples.

Identification of DEGs

Using the GEO2R tool (https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE40435) [15], the gene expression profiles of normal cervical tissues and cancer tissues were compared. The false discovery rate was determined using the Benjamini-Hochberg method and the DEGs selection criteria were adjusted based on *P*-value (< 0.05) and |log fold change (FC)| ≥ 1.5.

A PPI network of DEGs was determined

The PPI network of the selected DEGs was drawn with the help of the STRING database (http://www.bork.embl-heidelberg.de/STRING/) [16]. This database is a powerful computerized resource for determining interactions among proteins of interest. During STRING analysis, a combined interaction score (≥ 0.7) was marked as a cut-off criteria. Then, the drawn PPI of DEGs was imported to Cytoscape (version 3.6.0) [17] for further processing.

Hub gene identification

A Molecular Complex Detection (MCODE) application [18] of Cytoscape was utilized to select the most significant module in the obtained PPI network of the DEGs, using cutoff = 2, node score cutoff = 0.2, k-core = 2, and maximum depth = 100 as the selection criteria. Based on the degree method, the top six hub genes were identified from the selected module using the Cytohubba application in the PPI network.

Expression analysis of hub genes

For expression analysis of the hub genes, UALCAN was performed [19]. UALCAN is a recently developed web-based tool for customized and integrated multi-omics analyses of the gene(s) of interest across various The Cancer Genome Atlas (TCGA) datasets. In our study, the selected Cervical Squamous Cell Carcinoma (CESC) dataset for hub genes' expression analysis consisted of 308 cervical tissue samples (305 cancer samples and 3 normal tissue samples). Moreover, we also used this database to evaluate the clinical variable-wise expression of hub genes. For defining the dif-

ferential expression of the hub genes, a $P < 0.05$ value was used as the cutoff.

Verification of hub gene expressions

Next, we also verified hub gene expression in CESC cell lines and tissues relative to normal tissues using OncoDB, GEPIA, Human Protein Atlas (HPA), and GENT2 online databases [20-22]. These databases are newly developed for customized and integrated multi-omics analyses of the gene(s) of interest across cancer datasets. For defining a differential expression of the hub genes, a $P < 0.05$ value was used as a cutoff.

DNA methylation analysis

OncoDB is a user friendly database that is available online to carry out gene expression and methylation analysis across thousands of cancer patients suffering from 33 major types of cancer [20]. This resource was used to analyze the DNA methylation of hub genes across CESC patients.

cBioPortal data analysis

cBioPortal (https://www.cbioportal.org/) provides easy access to cancer genomic data for multidimensional *in silico* analysis [23]. We used this database to evaluate the genomic mutations of hub genes in CESC.

Gene set enrichment analysis (GSEA)

We carried out GSEA of the hub genes using normalized RNA-seq data from TCGA projects in the ShinyGo database [24]. This analysis helped us to understand the biologic functions and pathways of the hub genes. A P value < 0.05 and FDR < 0.05 were considered significant.

ceRNA network construction

The ENCORI database (https://starbase.sysu.edu.cn/) [25] was used in this study to predict the hub genes' targeted microRNAs (miRNAs), long non-coding RNAs (lncRNAs), and propose a ceRNA network.

Survival analysis

The overall survival (OS) analysis of hub genes across CESC patients was carried out with the help of oncoDB and GEPIA tools [20]. These tools obtained survival-associated vital information from the TCGA project. During the analysis, log-rank P < 0.05 was considered significant.

Hub gene associated drugs

The identified hub genes may be promising therapeutic targets, thus we conducted Drug-Gene Interaction Database (DGIdb) analysis to identify hub gene-associated drugs. This database provides details on drugs targeting hub genes from various reliable databases and medical literature [26].

Results

Overlapping DEG screening, hub gene identification, and expression analysis

Initially, we used two datasets (GSE138080 and GSE113942) to screen overlapping DEGs with the filtering criteria mentioned in the Methods section. As shown in the drawn PPI network using STRING in **Figure 1A**, containing 79 nodes, 318 edge nodes, and an average node degree score of 7.1, a total of 79 overlapping DEGs were identified between these two datasets. Next, MCODE analysis revealed a 31 gene-based most significant sub-module in the PPI of the DEGs (**Figure 1B**). Later, based on the degree method, the identified sub-module was further analyzed by Cytohubba analysis to identify the top 6 hub genes. Results revealed that MCM4 (mini-chromosome maintenance proteins), NUSAP1 (Nucleolar and Spindle Associated Protein 1), CDCA5 (Cell Division Cycle Associated 5), CDC45 (Cell Division Cycle 45), DTL (Denticleless), and CDT1 (Chromatin Licensing and DNA Replication Factor 1) were the top 6 hub genes in the identified sub-module of the DEGs (**Figure 1C**). Ultimately, the hub genes' expression in the normal cervical and CESC tissues was explored through UALCAN. A total of 3 normal and 305 CESC tissues were used for this purpose in UALCAN. As highlighted in **Figure 1D** and **1E**, all 6 hub genes (MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1) were up-regulated in CESC samples relative to normal cervical samples (**Figure 6**).

Hub gene expression is correlated with clinical variables in CESC patients

Next, we evaluated correlations between hub gene expression and different diverse clinical

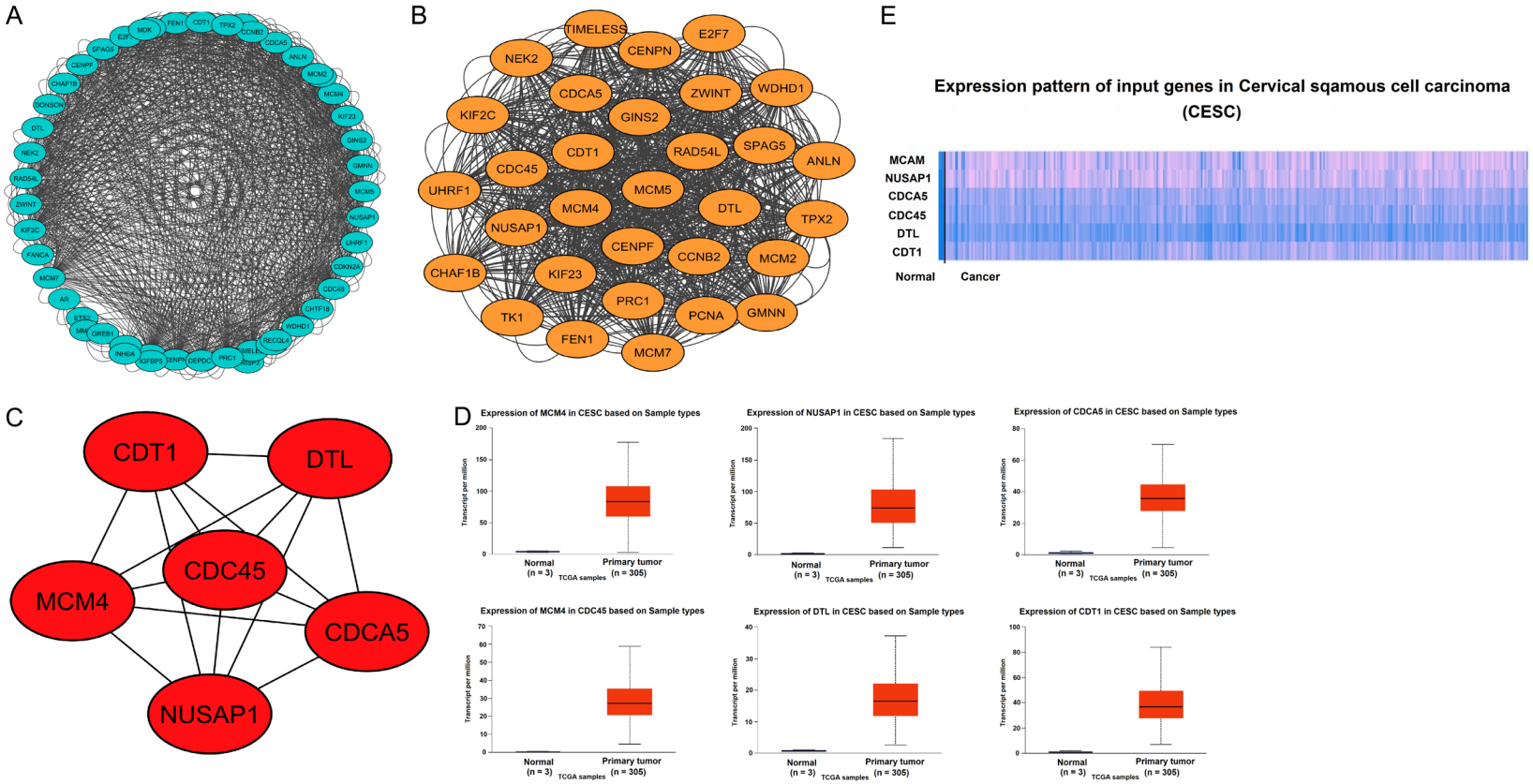

Figure 1. Screening of overlapping differentially expressed genes (DEGs), sub-module identification, hub genes identification, and expression analysis in cervical squamous cell carcinoma (CESC) samples. (A) A PPI network of identified 61 overlapping DEGs between GSE138080 and GSE113942 datasets, (B) A PPI of the identified sub-module via Molecular Complex Detection (MCODE) analysis, (C) Top 6 identified hub gens in the sub-module based on degree method, and (D and E) Expression analysis of hub genes across CESC and normal cervical samples via UALCAN.

variables of Cervical Squamous Cell Carcinoma (CESC) patients through the UALCAN database. The correlation analysis showed that hub genes expression was notably corrected with different analyzed clinical values, including cancer stage (**Figure 2A**), race (**Figure 2B**), body weight (**Figure 2C**), age (**Figure 2D**), and nodal metastasis status (**Figure 2E**). These results suggest that higher expression of the identified hub genes (MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1) is not restricted to any specific cancer stage, race, body weight, age, or nodal metastasis status, but they are more likely to be involved in advanced cancer stage and nodal metastasis status in CESC patients.

Verification of hub gene expression

To verify hub gene expression in CESC patients from another cohort, OncoDB, GEPIA, and GENT2 enabled detection of hub gene expression levels across CESC patients and cell lines, adjacent to normal cervical tissues and cell lines. Results of the analysis showed that all hub genes (MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1) were significantly upregulated in CESC samples and cell lines relative to controls (**Figure 3**). These results are consistent with the UALCAN analysis results.

Protein levels of hub genes

Furthermore, the proteomic expression of MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1 in CESC and normal tissues was measured through immunohistochemistry (IHC) by utilizing the Human Protein Atlas (HPA) database. As highlighted in **Figure 4**, across CESC tissues, the proteomic expressions of MCM4 (Staining: High), NUSAP1 (Staining: High), CDCA5 (Staining: High), CDC45 (Staining: High), DTL (Staining: High), and CDT1 (Staining: High) were higher than those in the control tissues, i.e. MCM4 (Staining: Medium), NUSAP1 (Staining: Low), CDCA5 (Staining: Low), CDC45 (Staining: Medium), DTL (Staining: Low), and CDT1 (Staining: Low). Moreover, we also documented variations in the proteomic expression of different hub gene-associated isoforms across CESC samples in the GEPIA database. As shown in **Figure 5**, proteomic expression of different hub gene isoforms varied from isoform to isoform. Therefore, it was concluded that hub gene overexpression in CESC samples also varied from isoform to isoform.

DNA methylation analysis

We queried whether the identified highly expressed 6 hub genes (MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1) had any association with DNA promoter methylation across CESC samples relative to controls. To observe this, hub gene methylation levels in CESC patients and normal controls were detected through the oncoDB database. Results showed that the promoter regions of MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1 hub genes had lower methylation levels ($P < 0.05$) in CESC samples relative to controls (**Figure 6**).

Genomics changes in the hub genes

As the result of cBioPortal analysis, around 4.1% (NUSAP1, CDCA5, and CDT1), 1.7% (MCM4), and 1.5% (CDC45 and DTL) CESC samples stored in cBioPortal had genomic changes (**Figure 7A**). Mutations were the major genomic alterations in MCM4 and CDCA5 hub genes, while deep amplification and deletion were the major genomic changes in CDC45, DTL, NUSAP1 and CDT1 hub genes across CESC patients (**Figure 7A**). Moreover, cBioPortal analysis showed that mutations across hub genes can change amino acids in different domains of the encoded proteins (**Figure 7B**). Thus, hub genes were genetically altered in a small proportion of the CESC patients.

GO functions and KEGG pathways of hub genes

GO functions of the identified hub genes are classified into biological processes (BP), cell component (CC), and molecular function (MF). Concerning BP, the identified hub genes are significantly involved in DNA replication preintiation complex assembly, Double-strand break repair via break-induced replication, and Mitotic DNA replication among others (**Figure 8**). Concerning CC, the identified hub genes were significantly correlated with Cul4b-RING E3 ubiquitin ligase complex, DNA replication preintiation complex, CMG complex, and MCM complex, and others (**Figure 8**). The MF results showed that hub genes were enriched in DNA replication origin binding, DNA polymerase binding, single stranded DNA binding, and chromatin binding, and others (**Figure 8**). Ultimately, Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis revealed that hub genes are

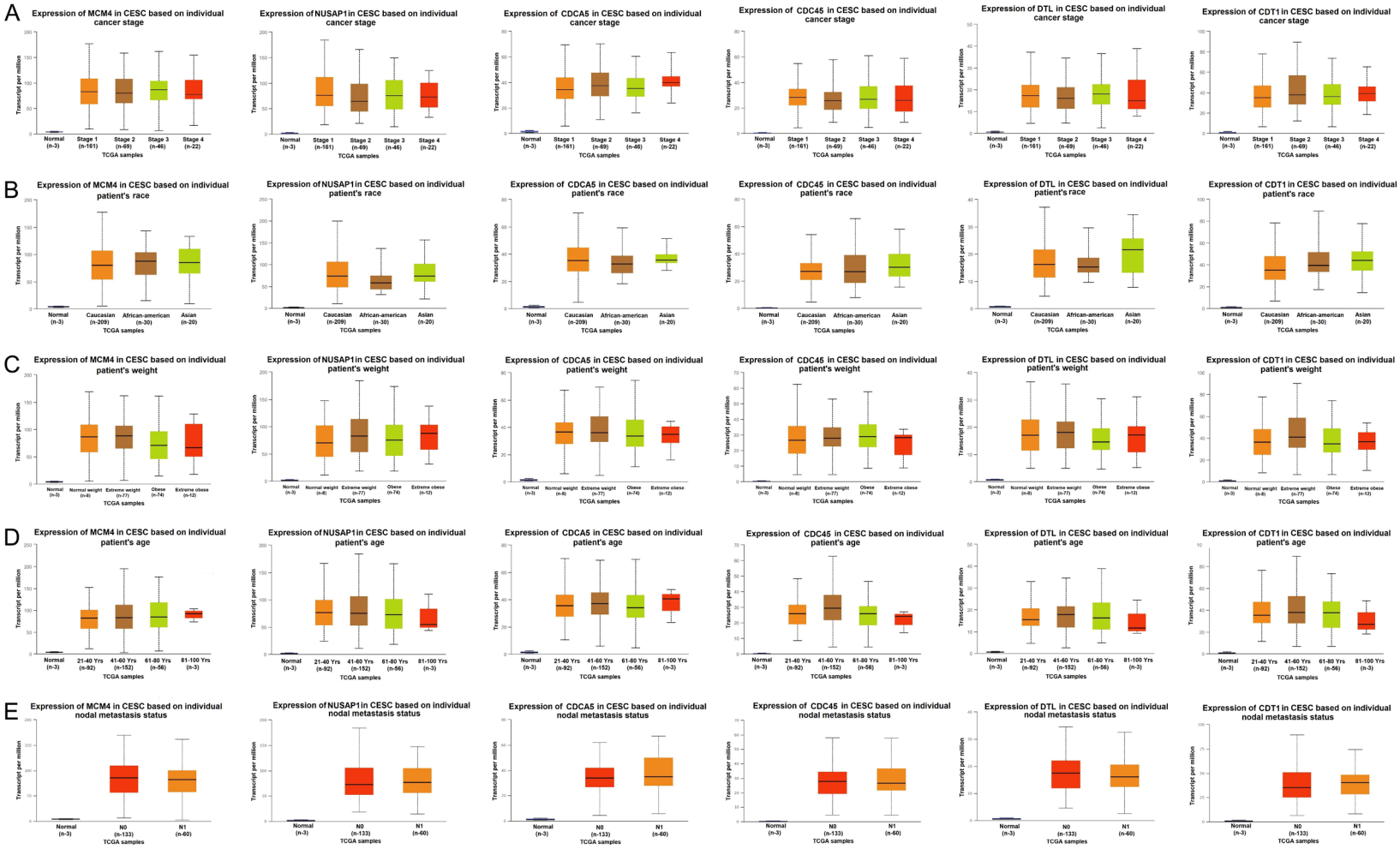

Figure 2. Hub genes expression evaluation across diverse clinical variables of CESC patients. (A) Hub genes expression across different cancer stages, (B) Hub genes expression across different races, (C) Hub genes expression across different body weights, (D) Hub genes expression across different ages, and (E) Hub genes expression across different nodal metastasis status.

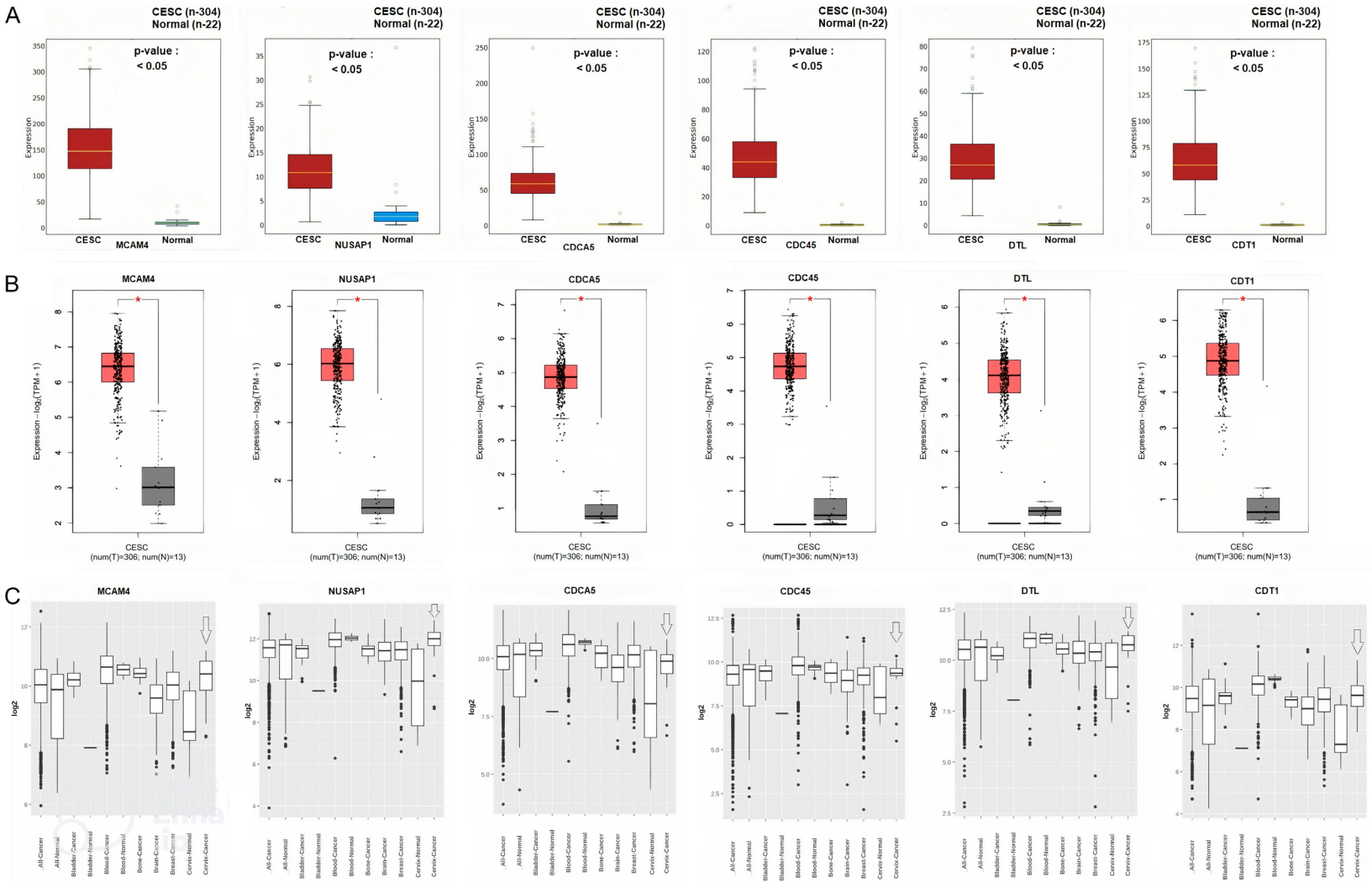

Figure 3. Results of the expression level of hub genes across CESC samples and cell lines. (A) GENT2-based expression of hub genes across CESC patients, (B) GEPIA-based expression of hub genes across CESC patients, and (C) GENT2-based expression of hub genes across CESC cell lines.

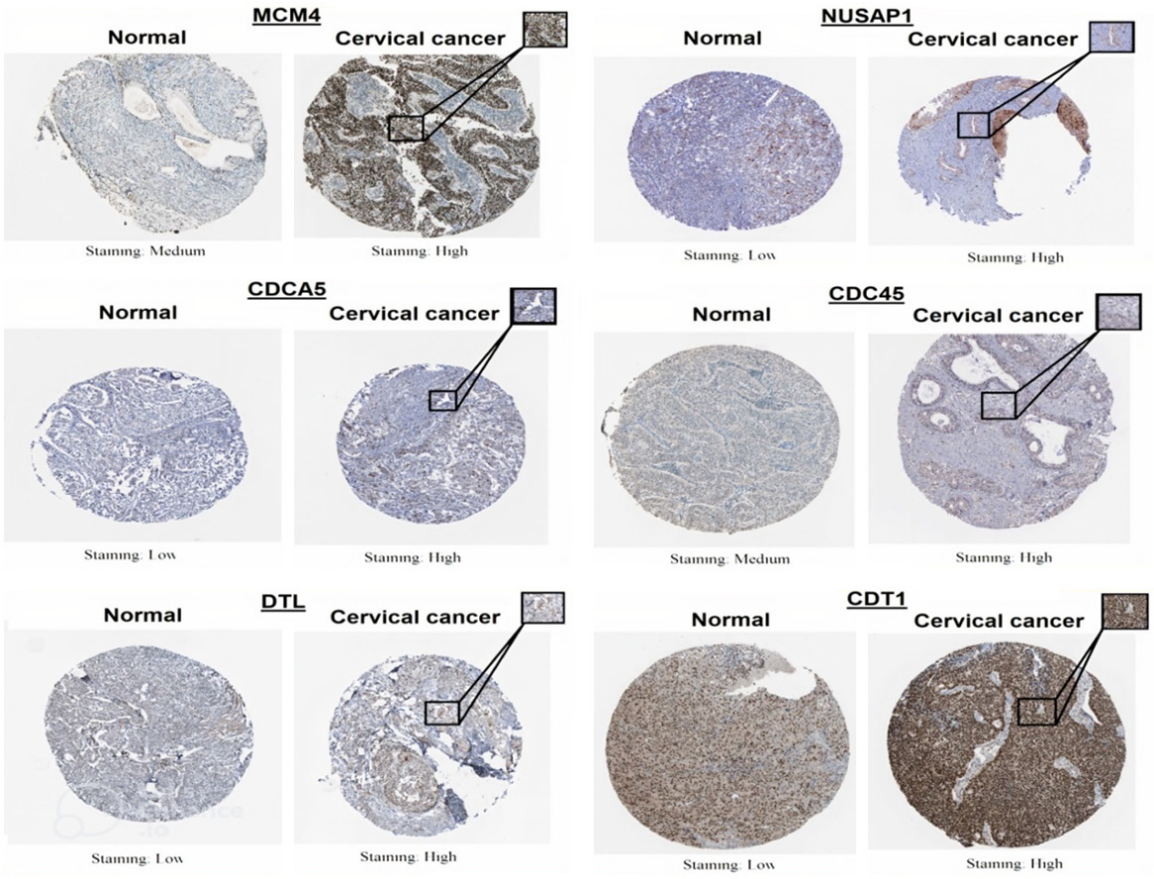

Figure 4. IHC-based proteomic expression of hub genes in CESC patients and normal individual samples in the Human Protein Atlas (HPA).

involved in diverse pathways, including DNA replication, Cell cycle, and MicroRNAs in cancer (**Figure 8**).

Survival analysis of the hub genes

Hub genes (MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1) prognostic information was obtained by the Kaplan-Meier (KM) Plotter and GEPIA tools. Results of both tools showed that CESC patients having a higher expression of the identified hub genes (MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1) did not have any negative impact on their OS duration (**Figure 9**).

Screening of ceRNA network involving hub genes

Next, we screened the ceRNA network involving hub genes by the different online databases. For this purpose, we obtained 288 miRNAs targeting hub genes in the ENCORI database (**Figure 10A**). Then, from these 288 miRNAs, we identified a single miRNA (has-mir-34a-5p) that was targeting all 6 identified hub genes (**Figure 10B**). Therefore, this miRNA is considered the most useful one. After this, we also identified 56 has-mir-34a-5p regulatory lncRNAs by Starbase (**Figure 10C**). Unfortunately, our results of ceRNA network screening involving identified hub genes that were based on a small sample size, which might limit the reliability of the ceRNA network.

Drug analysis of the hub genes

Considering hub genes as therapeutic targets, information on several chemotherapeutic drugs was retrieved from the DGIdb database in this study. Results of the analysis revealed a total of 27 candidate chemotherapeutic drugs capable of inhibiting hub gene expression including Aflatoxin B1, Afuresertib, Azathioprine, Dasatinib, Calcitriol, Decamethrin, Azathioprine, Pravastatin, and Geraniol (**Table 1**). However,

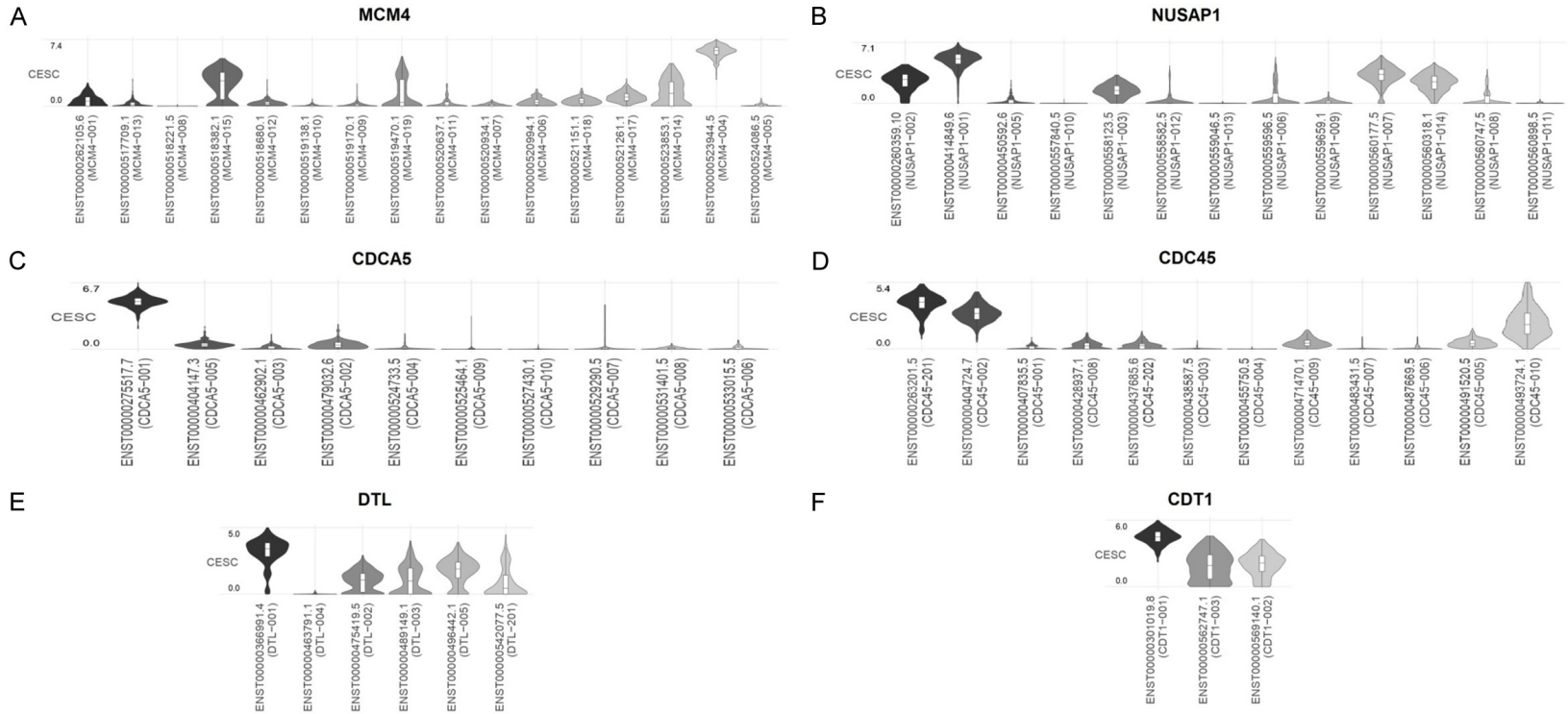

Figure 5. Isoform-wise proteomic expression of hub genes in CESC patients and normal individual samples in GEPIA.

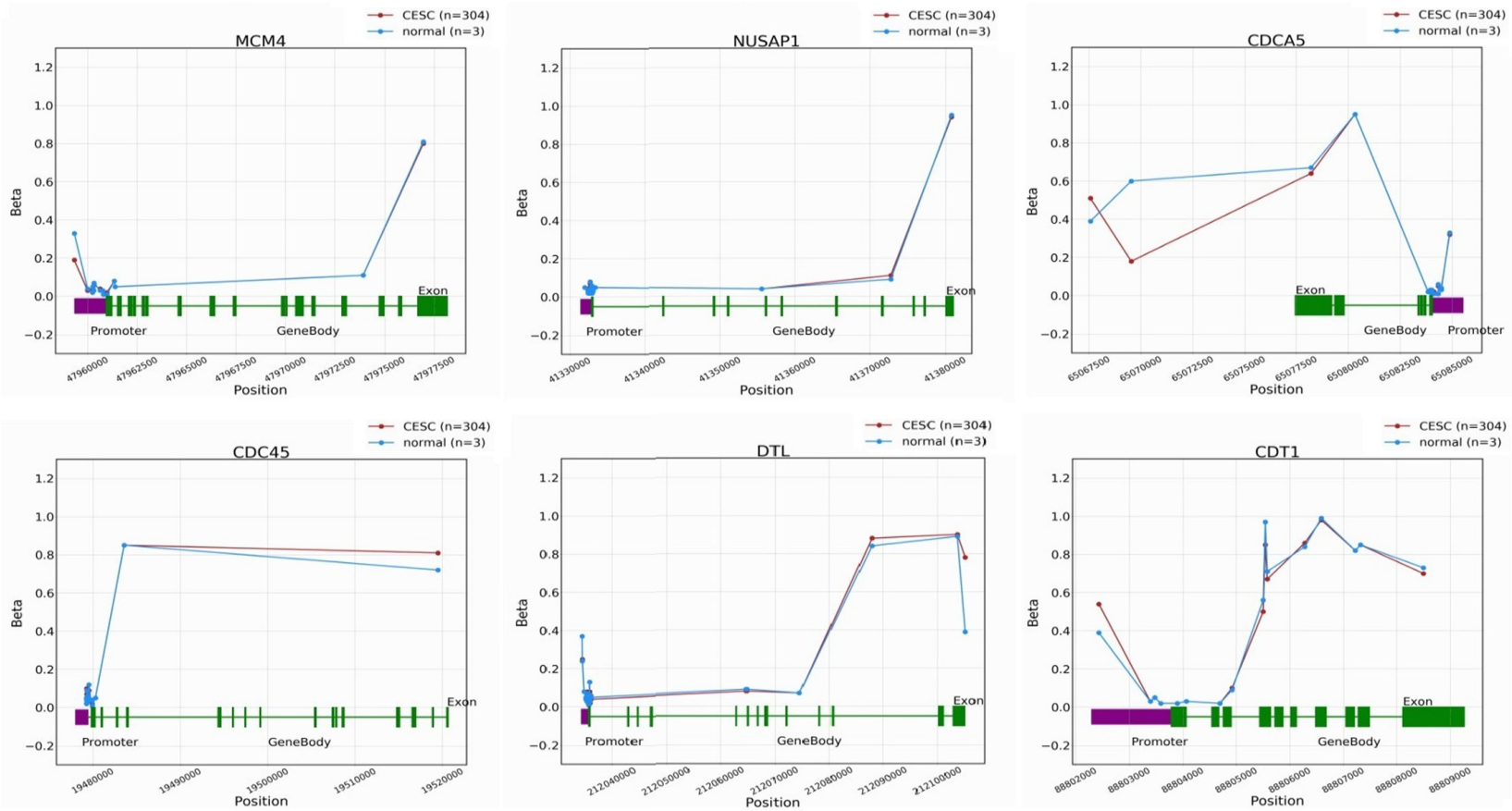

Figure 6. Visualization of DNA promoter methylation levels of MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1 hub genes across CESC samples and normal controls.

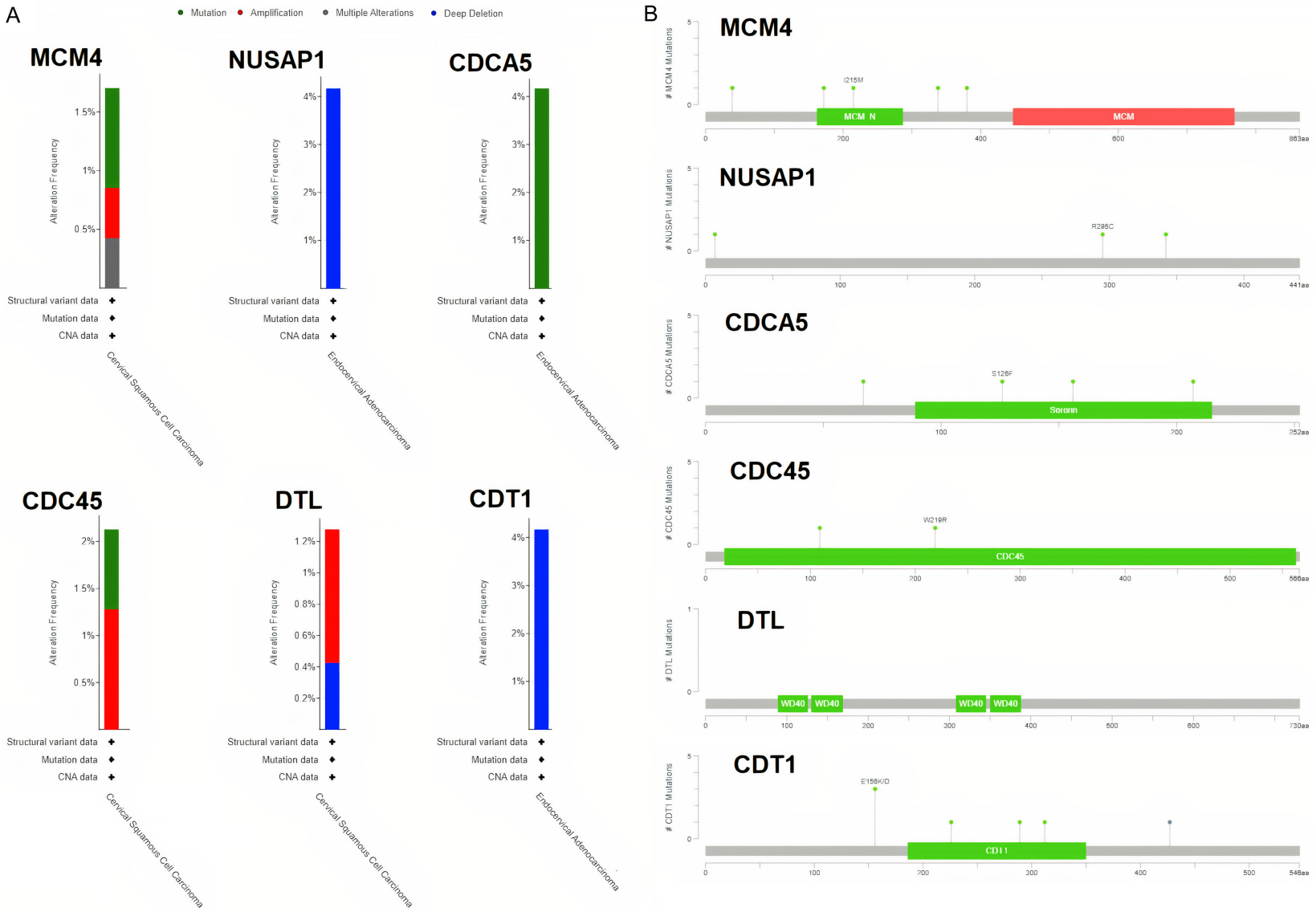

Figure 7. Genomic change analysis of identified hub genes across CESC samples based on cBioPortal. (A) Percentages of genetically altered CESC samples for hub genes and the types of genomic alterations, (B) Amino acid changes due to genomic mutations in the proteins encoded by hub genes.

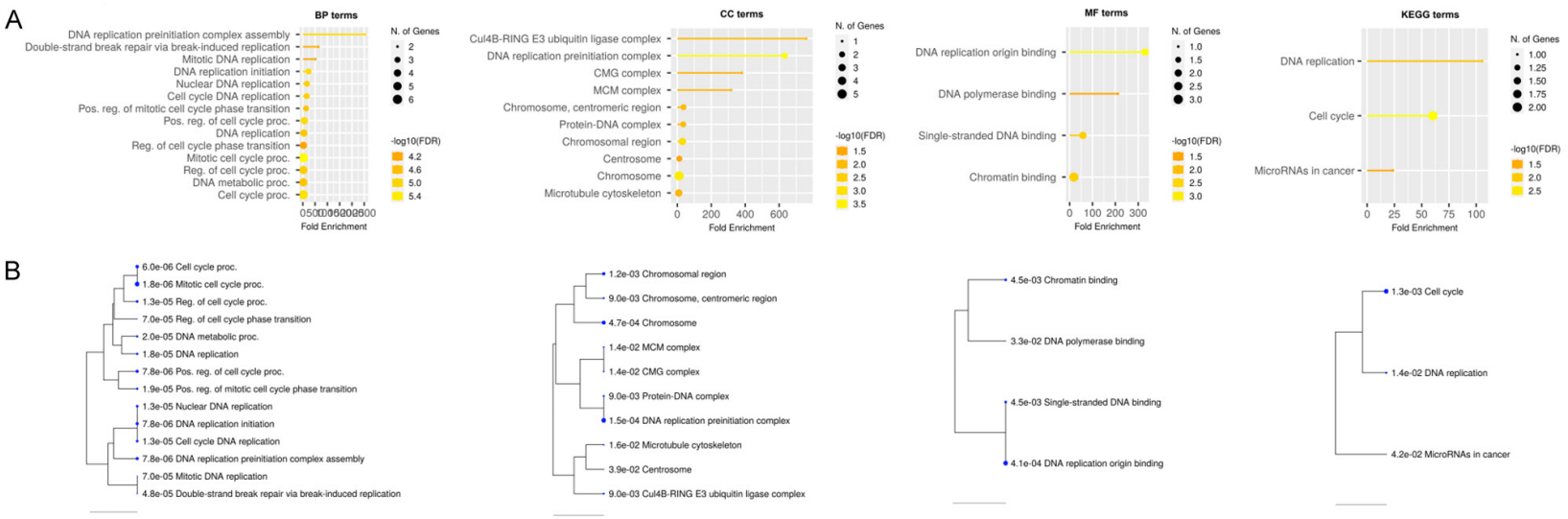

Figure 8. Gene Ontology (GO) functions and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of hub genes through ShinyGO. (A) Dotplots of GO functions and KEGG pathways of hub genes and (B) Tree representation of GO functions and KEGG pathways of hub genes.

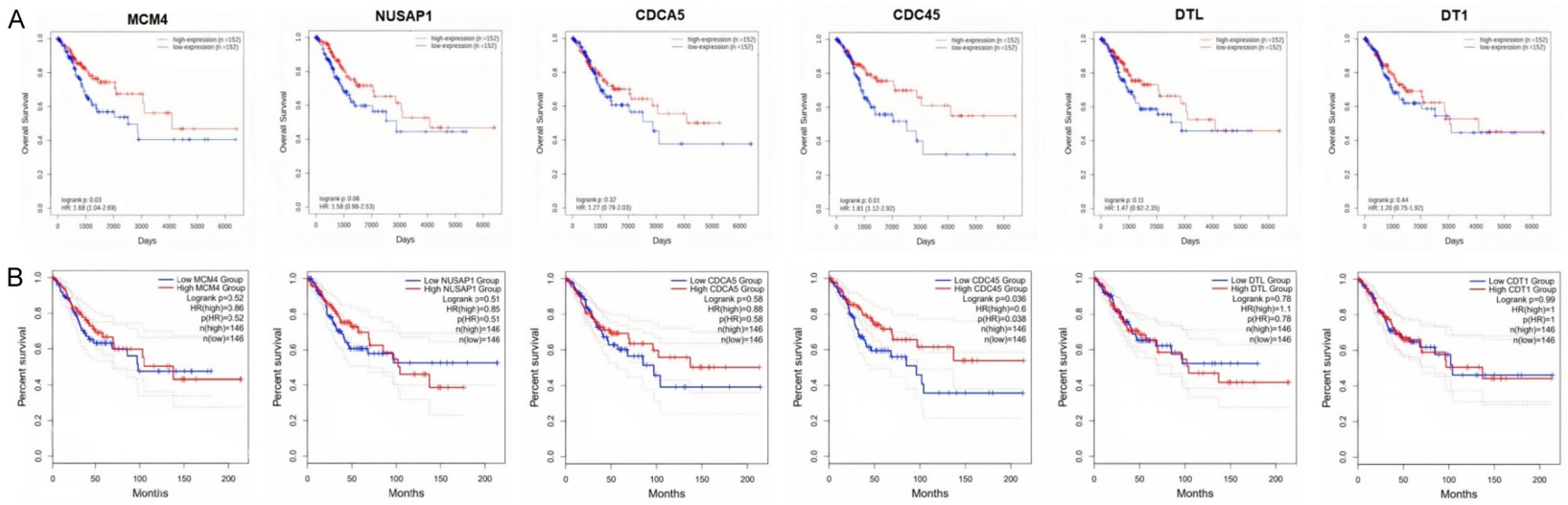

Figure 9. Overall survival (OS) analysis of hub genes. (A) OS analysis via KM plotter and (B) OS analysis via GEPIA.

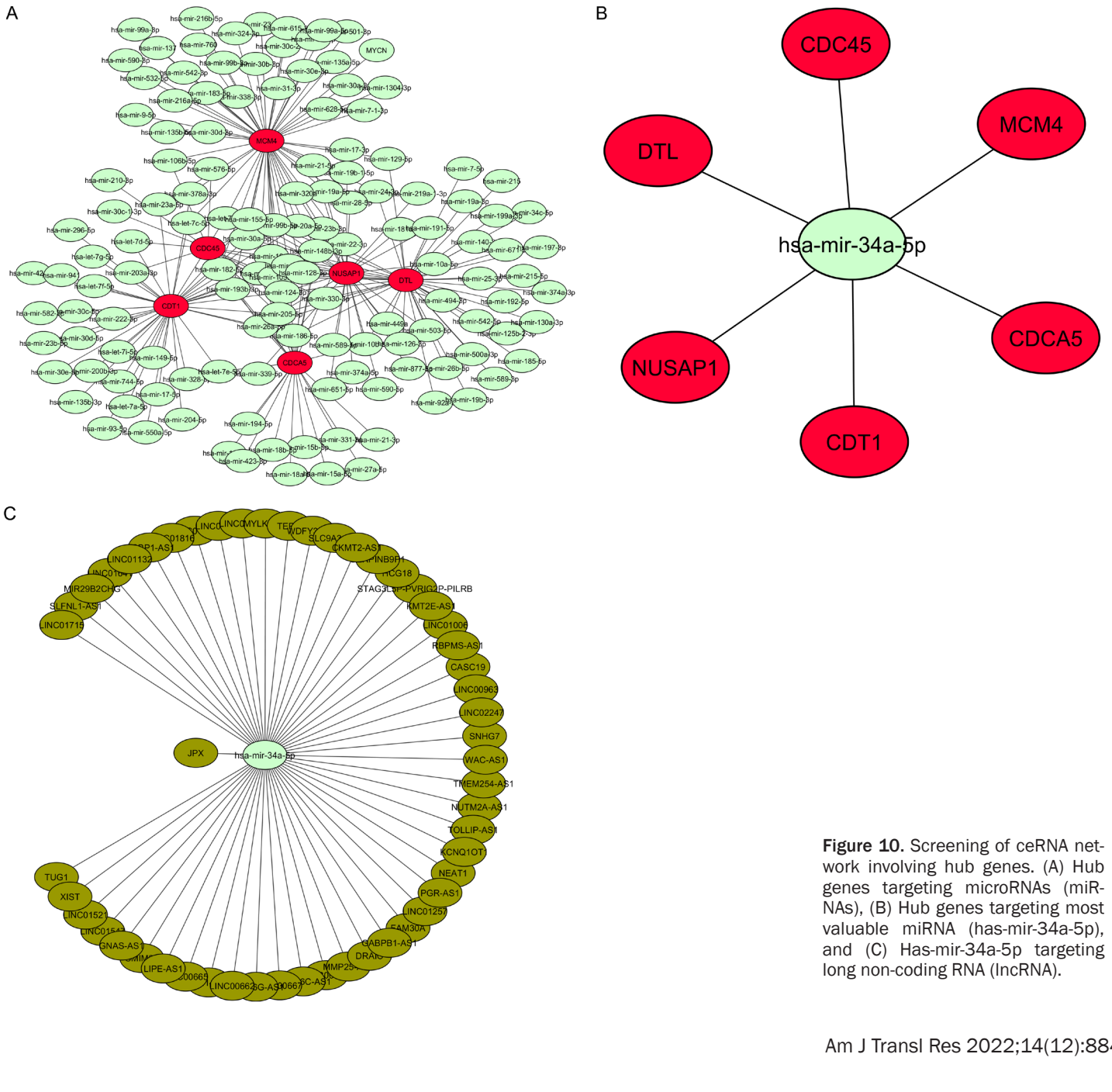

Figure 10. Screening of ceRNA network involving hub genes. (A) Hub genes targeting microRNAs (miRNAs), (B) Hub genes targeting most valuable miRNA (has-mir-34a-5p), and (C) Has-mir-34a-5p targeting long non-coding RNA (lncRNA).

Table 1. DGIdb-based hub genes associated drugs

Sr. No	Gene	Drug	Interaction	Mechanism	Reference count
1	MCM4	Aflatoxin B1	inhibitory	decrease MCM4 expression	2
2	MCM4	Afuresertib	inhibitory	decrease MCM4 expression	6
3	MCM4	Corticosterone	inhibitory	decrease MCM4 expression	4
4	MCM4	Dibutyl Phthalate	inhibitory	decrease MCM4 expression	3
5	MCM4	Diethylnitrosamine	inhibitory	decrease MCM4 expression	6
6	MCM4	Estradiol	inhibitory	decrease NUSAP1 expression	4
7	NUSAP1	Azathioprine	inhibitory	decrease NUSAP1 expression	7
8	NUSAP1	Bisphenol A	inhibitory	decrease NUSAP1 expression	9
9	NUSAP1	Dasatinib	inhibitory	decrease NUSAP1 expression	5
10	CDCA5	Calcitriol	inhibitory	decrease CDC5A expression	2
11	CDCA5	Cisplatin	inhibitory	decrease CDC5A expression	1
12	CDCA5	Decamethrin	inhibitory	decrease CDC5A expression	1
13	CDCA5	Fipronil	inhibitory	decrease CDC5A expression	4
14	CDCA5	Ivermectin	inhibitory	decrease CDC5A expression	3
15	CDC45	Azathioprine	inhibitory	decrease CDC45 expression	5
16	CDC45	Calcitriol	inhibitory	decrease CDC45 expression	4
17	CDC45	Deguelin	inhibitory	decrease CDC45 expression	7
18	CDC45	Etoposide	inhibitory	decrease CDC45 expression	6
19	DTL	Demecolcine	inhibitory	decrease DTL expression	11
20	DTL	Doxorubicin	inhibitory	decrease DTL expression	4
21	DTL	Lucanthone	inhibitory	decrease DTL expression	5
22	DTL	Pravastatin	inhibitory	decrease DTL expression	7
23	CDT1	Calcitriol	inhibitory	decrease CDT1 expression	8
24	CDT1	Fenthion	inhibitory	decrease CDT1 expression	2
25	CDT1	Geraniol	inhibitory	decrease CDT1 expression	2
26	CDT1	Oxaliplatin	inhibitory	decrease CDT1 expression	4
27	CDT1	Palbociclib	inhibitory	decrease CDT1 expression	4

Reference count = count of the studies in medical literature which supported the specific hub gene-drug interaction.

further experimental validation of the drug analysis results is necessary in CESC samples.

Discussion

Cervical Squamous Cell Carcinoma (CESC) is one of the most prevalent female cancers. Recent significant improvements in cancer research have made it possible to prevent and treat it [27]. However, the exact nature of the molecular mechanisms governing CESC development and progression is notfully understood. The advancement in high throughput technologies has helped researchers and clinicians to explore novel molecular biomarkers for cancer development and progression by analyzing microarray datasets.

In this research, we performed the analysis of two microarray datasets (GSE138080 and GSE113942) by a detailed *in silico* methodology. First, in total, top 79 common DEGs were retrieved from both these datasets using GEO2R analysis. Then, the PPI network of the retrieved DEGs was constructed and visualized by Cytoscape, which consisted of 79 nodes and 380 edges. The MCODE analysis of the retrieved DEGs identified a top sub-module of 31 DEGs in the PPI network. Ultimately, based on the degree method, the Cytohubba analysis revealed MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1 as hub genes. The expression profiling of the identified hub genes by UALCAN revealed that these hub genes were significantly up-regulated in CESC patients relative to normal controls. Furthermore, the OS analysis showed that higher expression of MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1 hub genes did not correlate with the shorter surviv-

al duration of the CESC patients. In view of the above results, this suggests MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1 may be associated with the diagnosis but not the prognosis of CESC patients.

MCM4 (Mini-chromosome maintenance complex component 4) is an important subunit of the CMG replicative helicase, which is a crucial factor for maintaining different diverse processes, including DNA recombination, replication, and repair [28]. Overexpression of MCM4 was earlier reported in gastric cancer (GC) metastatic tumor samples relative to normal controls [28]. MCM4 dysregulation is also known to contribute to the development and progression of colorectal cancer (CRC). In CRC, MCM4 upregulation is linked with metastasis, proliferative capacity, vascular invasion, and tumors' histological grade [29]. Contrary to this, MCM4 knockdown in HCT116 CRC cell lines resulted in cell cycle arrest, a decrease in cell growth, and an increase in apoptosis [30]. Moreover, the MCM4 expression in CRC patients' stool samples was found up-regulated compared to healthy individuals, suggesting that MCM4 expression may be used as a non-invasive diagnostic biomarker across CRC patients [31]. In the current study, our results showed a significant higher expression of MCM4 across CESC patients, suggesting it as a possible novel molecular biomarker. However, MCM4 dysregulation must be investigated in a larger CESC cohort to confirm its role in CESC development and progression.

Nucleolar and spindle-associated protein 1 (NUSAP1) is a key gene and its activity is crucial for different cellular events in mitosis, such as in spindle assembly and cytokinesis [32]. Mitosis dysregulation is a very common event in cancer cells. Earlier studies reported that NUSAP1 down-regulation leads to various abnormal events in cells, such as growth 2 (G2)/mitotic (M) cell cycle phase arrest, abnormal chromosomal segregation, abnormalities in interphase nuclei, chromosomal misalignment, and aberrant spindle assembly, etc. [32]. Although dysregulation of the NUSAP1 gene is earlier reported in breast cancer [33], pancreatic cancer [34], oral cancer [35], and hepatocellular carcinoma [36]. However, to our knowledge, there is a lack of studies documenting expression variation of the NUSAP1 gene in CESC patients. In the current study, through detailed *in silico* analysis we revealed that the higher expression of NUSAP1 is associated with CESC. However, detailed studies based on large CESC cohorts are still required for verification of our findings.

The cell division cycle-associated 5 (CDCA5) gene is involved in regulating sister chromatid segregation and cohesion [37]. This activity is carried out by CDCA5 through the stabilization of a cohesive complex to assure correct chromosome separation during mitosis and meiosis processes. In addition to this, CDCA5 is involved in the DNA repair process [38] and regulates the expression of various cell cycle-linked proteins, thereby being associated with promoting cell proliferation and inducing apoptosis in cancer cells [39]. According to previous studies, the up-regulation of CDCA5 is closely associated with the development and progression of different cancers, including liver [40, 41], bladder [42], colorectal [43], and lung [44]. Moreover, through bioinformatic analysis, CDCA5 overexpression proved to be significantly involved in the development of prostate cancer (PCa) [45, 46]. However, the underlying molecular mechanisms and pathways of CDCA5 in the development of PCa are not clear. Although our bioinformatic analysis indicated a higher expression of CDCA5 in CESC patients, future studies based on large CESC cohorts are warranted to validate the diagnostic and prognostic value of CDCA5 in CESC.

A protein encoded by the Cell Division Cycle 45 (CDC45) gene acts as the initiation factor for DNA replication [47]. The dysregulation of CDC45 was earlier reported to be involved in the initiation and progression of several human cancers and thus is used as a reliable therapeutic target. For example, *Huang et al.* showed that the down-regulation of CDC45 is responsible for suppressing uncontrolled proliferation of the cells in non-small cell lung cancer (NSCLC), ultimately resulting in G2/M phase cell cycle arrest [48]. The results of this study favored a carcinogenic effect of CDC45. Furthermore, *Sun et al.* revealed an up-regulation of CDC45 expression in papillary thyroid cancer (PTC), which was ultimately linked to promoting proliferation of cancer cells [49]. In this study, we showed that CDC45 was significantly upregulated across CESC samples, indicating that

CDC45 may play an oncogenic role in CESC initiation and progression and may be a reliable therapeutic target.

DTL (denticleless protein homolog) protein is an important regulator of CDT1 protein degradation after DNA damage [50]. It is reported that the CRL4 CDT2 complex, along with Rad6/18 and monoubiquitinated PCNA is crucial for DTL functioning in regulating the DNA replication process [51]. Previous studies reported that DTL is overexpressed in cutaneous melanoma patients and reflects an unfavorable prognosis [52]. Moreover, Vanderdys *et al.* in their latest study revealed that DTL is overexpressed in head and neck squamous cell carcinoma (HNSCC) patients, and down-regulation of DTL by short interfering RNA (siRNA) results in the attenuation of HNSCC cell growth [53]. In the current study, we noticed a significant overexpression of DTL in CESC patients, suggesting it as a novel diagnostic biomarker. However, more studies based on CESC cell lines are needed to confirm that.

Chromatin licensing and DNA replication factor 1 (CDT1) is a key factor involved in the initiation of the DNA replication process [54]. Studies suggested that abnormalities in the DNA replication process are linked with cancer development and progression [55]. CDT1 protein helps to coordinate the cell cycle with proliferation inside the cells by making a pre-RC complex at the beginning of the cell cycle, which is further responsible for loading the MCMs family of proteins onto chromatin [56]. So far in the medical literature, limited studies have reported the higher expression of CDT1 across different cancers and linked it to cancer development and metastasis [57-59]. However, the exact molecular pathways involving CDT1 are not fully explored in human cancers. In the current study, we characterized a higher expression of the CDT1 gene in CECS patients using a detailed *in silico* approach and suggested it as a novel diagnostic biomarker. However, further wet-lab experiments are needed to confirm these results.

Promoter methylation and genetic alteration analyses of the hub genes showed that promoter hypomethylation was linked only with the overexpression of MCM4, CDCA5, DTL and CDT1 hub genes. Since promoter methylation can be reversed, targeted treatment therapies based on the promoter methylation of these hub genes can help control the expression of these genes in CESC patients. Moreover, hub genes were not found genetically altered in too many CESC samples, therefore, we speculate that genetic alterations do not participate in the dysregulation of hub genes. We further showed by OS analysis that identified hub genes were not good prognostic biomarkers due to their irrelevancy for the survival durations of CESC patients. In addition to this, in the present study, hub gene-related experimentally validated lncRNAs and miRNAs were also explored in order to construct a lncRNA-miRNA-mRNA network of the mRNA, lncRNAs, and miRNAs that could help to understand the development of CESC at the molecular level in more depth.

GSEA analysis of the identified hub genes revealed that these hub genes were enriched in DNA replication preintiation complex assembly, Double-strand break repair via break-induced replication, and Mitotic DNA replication, BP terms. The identified hub genes were significantly correlated with Cul4b-RING E3 ubiquitin ligase complex, DNA replication preintiation complex, CMG complex, and MCM complex as CC terms, DNA replication origin binding, DNA polymerase binding, single stranded DNA binding, and chromatin binding as MF terms, and DNA replication, Cell cycle, and MicroRNAs as cancer pathway terms. Finally, we identified 31 chemotherapeutic drugs against hub genes through the DGIdb database. Based on that finding, the expression of identified hub genes can be controlled using a variety of drugs. Therefore, looking at the hub gene-drug interaction, we speculate that CESC patients could be treated with these drugs in the future. However, further *in vitro* and *in vivo* studies are needed.

Conclusion

Through an integrated bioinformatic approach, our study has revealed 6 hub genes (MCM4, NUSAP1, CDCA5, CDC45, DTL, and CDT1) that may play pathogenic roles in CESC development and can also be used as molecular biomarkers for CESC patients. However, more studies are needed before clinical application.

Acknowledgements

The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry

of Education in Saudi Arabia for funding this research work through the project no. (IFKSURG-2-1118).

Disclosure of conflict of interest

None.

Address correspondence to: Mostafa A Abdel-Maksoud, Department of Botany and Microbiology, College of Science, King Saud University, P.O. 2455, Riyadh 11451, Saudi Arabia. E-mail: mabdelmaksoud@ksu.edu.sa; Muhammad Haris, Department of Anatomy, Institute of Basic Medical Sciences, Khyber Medical University, Peshawar, Pakistan. E-mail: dx_harris@hotmail.com

References

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394-424.

[2] Fu K, Zhang L, Liu R, Shi Q, Li X and Wang M. MiR-125 inhibited cervical cancer progression by regulating VEGF and PI3K/AKT signaling pathway. World J Surg Oncol 2020; 18: 115.

[3] Cohen PA, Jhingran A, Oaknin A and Denny L. Cervical cancer. Lancet 2019; 393: 169-182.

[4] Ji Y and Wang H. Prognostic prediction of systemic immune-inflammation index for patients with gynecological and breast cancers: a meta-analysis. World J Surg Oncol 2020; 18: 197.

[5] Andrae B, Andersson TM, Lambert PC, Kemetli L, Silfverdal L, Strander B, Ryd W, Dillner J, Törnberg S and Sparén P. Screening and cervical cancer cure: population based cohort study. BMJ 2012; 344: e900.

[6] Brucker SY and Ulrich UA. Surgical treatment of early-stage cervical cancer. Oncol Res Treat 2016; 39: 508-514.

[7] Tamura G. Alterations of tumor suppressor and tumor-related genes in the development and progression of gastric cancer. World J Gastroenterol 2006; 12: 192-8.

[8] Menderes G, Black J, Schwab CL and Santin AD. Immunotherapy and targeted therapy for cervical cancer: an update. Expert Rev Anticancer Ther 2016; 16: 83-98.

[9] Merrick BA, London RE, Bushel PR, Grissom SF and Paules RS. Platforms for biomarker analysis using high-throughput approaches in genomics, transcriptomics, proteomics, metabolomics, and bioinformatics. IARC Sci Publ 2011; 121-142.

[10] Huang X, Liu S, Wu L, Jiang M and Hou Y. High throughput single cell RNA sequencing, bioinformatics analysis and applications. Adv Exp Med Biol 2018; 1068: 33-43.

[11] Tomczak K, Czerwińska P and Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) 2015; 19: A68-77.

[12] Wang Z, Lachmann A and Ma'ayan A. Mining data and metadata from the gene expression omnibus. Biophys Rev 2019; 11: 103-110.

[13] Xue H, Sun Z, Wu W, Du D and Liao S. Identification of hub genes as potential prognostic biomarkers in cervical cancer using comprehensive bioinformatics analysis and validation studies. Cancer Manag Res 2021; 13: 117-131.

[14] Liu Y, Yi Y, Wu W, Wu K and Zhang W. Bioinformatics prediction and analysis of hub genes and pathways of three types of gynecological cancer. Oncol Lett 2019; 18: 617-628.

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

[16] Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ and Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019; 47: D607-D613.

[17] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.

[18] Bader GD and Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 2003; 4: 2.

[19] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK and Varambally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017; 19: 649-658.

[20] Tang G, Cho M and Wang X. OncoDB: an interactive online database for analysis of gene expression and viral infection in cancer. Nucleic Acids Res 2022; 50: D1334-D1339.

[21] Park SJ, Yoon BH, Kim SK and Kim SY. GENT2: an updated gene expression database for normal and tumor tissues. BMC Med Genomics 2019; 12 Suppl 5: 101.

[22] Tang Z, Li C, Kang B, Gao G, Li C and Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017; 45: W98-W102.

[23] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML,

Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C and Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012; 2: 401-404.
[24] Ge SX, Jung D and Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics 2020; 36: 2628-2629.
[25] Huang DP, Zeng YH, Yuan WQ, Huang XF, Chen SQ, Wang MY, Qiu YJ and Tong GD. Bioinformatics analyses of potential miRNA-mRNA regulatory axis in HBV-related hepatocellular carcinoma. Int J Med Sci 2021; 18: 335-346.
[26] Wagner AH, Coffman AC, Ainscough BJ, Spies NC, Skidmore ZL, Campbell KM, Krysiak K, Pan D, McMichael JF, Eldred JM, Walker JR, Wilson RK, Mardis ER, Griffith M and Griffith OL. DGIdb 2.0: mining clinically relevant drug-gene interactions. Nucleic Acids Res 2016; 44: D1036-1044.
[27] Šarenac T and Mikov M. Cervical cancer, different treatments and importance of bile acids as therapeutic agents in this disease. Front Pharmacol 2019; 10: 484.
[28] Bell SD and Botchan MR. The minichromosome maintenance replicative helicase. Cold Spring Harb Perspect Biol 2013; 5: a012807.
[29] Giaginis C, Georgiadou M, Dimakopoulou K, Tsourouflis G, Gatzidou E, Kouraklis G and Theocharis S. Clinical significance of MCM-2 and MCM-5 expression in colon cancer: association with clinicopathological parameters and tumor proliferative capacity. Dig Dis Sci 2009; 54: 282-291.
[30] Liu Y, He G, Wang Y, Guan X, Pang X and Zhang B. MCM-2 is a therapeutic target of Trichostatin A in colon cancer cells. Toxicol Lett 2013; 221: 23-30.
[31] Davies RJ, Freeman A, Morris LS, Bingham S, Dilworth S, Scott I, Laskey RA, Miller R and Coleman N. Analysis of minichromosome maintenance proteins as a novel method for detection of colorectal cancer in stool. Lancet 2002; 359: 1917-1919.
[32] Iyer J, Moghe S, Furukawa M and Tsai MY. What's Nu (SAP) in mitosis and cancer? Cell Signal 2011; 23: 991-998.
[33] Chen DT, Nasir A, Venkataramu C, Fulp W, Gruidl M and Yeatman T. Evaluation of malignancy-risk gene signature in breast cancer patients. Breast Cancer Res Treat 2010; 120: 25-34.
[34] Kokkinakis DM, Liu X and Neuner RD. Modulation of cell cycle and gene expression in pancreatic tumor cell lines by methionine deprivation (methionine stress): implications to the therapy of pancreatic adenocarcinoma. Mol Cancer Ther 2005; 4: 1338-1348.
[35] Okamoto A, Higo M, Shiiba M, Nakashima D, Koyama T, Miyamoto I, Kasama H, Kasamatsu A, Ogawara K, Yokoe H, Tanzawa H and Uzawa K. Down-regulation of nucleolar and spindle-associated protein 1 (NUSAP1) expression suppresses tumor and cell proliferation and enhances anti-tumor effect of paclitaxel in oral squamous cell carcinoma. PLoS One 2015; 10: e0142252.
[36] Satow R, Shitashige M, Kanai Y, Takeshita F, Ojima H, Jigami T, Honda K, Kosuge T, Ochiya T, Hirohashi S and Yamada T. Combined functional genome survey of therapeutic targets for hepatocellular carcinoma. Clin Cancer Res 2010; 16: 2518-2528.
[37] Zhang N and Pati D. Sororin is a master regulator of sister chromatid cohesion and separation. Cell Cycle 2012; 11: 2073-2083.
[38] Zhang N and Pati D. Handcuff for sisters: a new model for sister chromatid cohesion. Cell Cycle 2009; 8: 399-402.
[39] Zhang N, Panigrahi AK, Mao Q and Pati D. Interaction of Sororin protein with polo-like kinase 1 mediates resolution of chromosomal arm cohesion. J Biol Chem 2011; 286: 41826-41837.
[40] Chen H, Chen J, Zhao L, Song W, Xuan Z, Chen J, Li Z, Song G, Hong L, Song P and Zheng S. CDCA5, transcribed by E2F1, promotes oncogenesis by enhancing cell proliferation and inhibiting apoptosis via the AKT pathway in hepatocellular carcinoma. J Cancer 2019; 10: 1846-1854.
[41] Shen Z, Yu X, Zheng Y, Lai X, Li J, Hong Y, Zhang H, Chen C, Su Z and Guo R. CDCA5 regulates proliferation in hepatocellular carcinoma and has potential as a negative prognostic marker. Onco Targets Ther 2018; 11: 891-901.
[42] Yeh CR, Hsu I, Song W, Chang H, Miyamoto H, Xiao GQ, Li L and Yeh S. Fibroblast ERα promotes bladder cancer invasion via increasing the CCL1 and IL-6 signals in the tumor microenvironment. Am J Cancer Res 2015; 5: 1146-1157.
[43] Shen A, Liu L, Chen H, Qi F, Huang Y, Lin J, Sferra TJ, Sankararaman S, Wei L, Chu J, Chen Y and Peng J. Cell division cycle associated 5 promotes colorectal cancer progression by activating the ERK signaling pathway. Oncogenesis 2019; 8: 19.
[44] Nguyen MH, Koinuma J, Ueda K, Ito T, Tsuchiya E, Nakamura Y and Daigo Y. Phosphorylation and activation of cell division cycle associated 5 by mitogen-activated protein kinase play a crucial role in human lung carcinogenesis. Cancer Res 2010; 70: 5337-5347.
[45] Wang L, Tang H, Thayanithy V, Subramanian S, Oberg AL, Cunningham JM, Cerhan JR, Steer CJ and Thibodeau SN. Gene networks and microRNAs implicated in aggressive prostate cancer. Cancer Res 2009; 69: 9490-9497.

[46] Li HY, Jin N, Han YP and Jin XF. Pathway cross-talk analysis in prostate cancer based on protein-protein network data. Neoplasma 2017; 64: 22-31.
[47] Hennessy KM, Lee A, Chen E and Botstein D. A group of interacting yeast DNA replication genes. Genes Dev 1991; 5: 958-969.
[48] Huang J, Li Y, Lu Z, Che Y, Sun S, Mao S, Lei Y, Zang R, Li N, Zheng S, Liu C, Wang X, Sun N and He J. Analysis of functional hub genes identifies CDC45 as an oncogene in non-small cell lung cancer-a short report. Cell Oncol (Dordr) 2019; 42: 571-578.
[49] Sun J, Shi R, Zhao S, Li X, Lu S, Bu H and Ma X. Cell division cycle 45 promotes papillary thyroid cancer progression via regulating cell cycle. Tumour Biol 2017; 39: 1010428317705342.
[50] Jin J, Arias EE, Chen J, Harper JW and Walter JC. A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. Mol Cell 2006; 23: 709-721.
[51] Evans DL, Zhang H, Ham H, Pei H, Lee S, Kim J, Billadeau DD and Lou Z. MMSET is dynamically regulated during cell-cycle progression and promotes normal DNA replication. Cell Cycle 2016; 15: 95-105.
[52] Li Z, Wang R, Qiu C, Cao C, Zhang J, Ge J and Shi Y. Role of DTL in hepatocellular carcinoma and its impact on the tumor microenvironment. Front Immunol 2022; 13: 834606.
[53] Vanderdys V, Allak A, Guessous F, Benamar M, Read PW, Jameson MJ and Abbas T. The neddylation inhibitor pevonedistat (MLN4924) suppresses and radiosensitizes head and neck squamous carcinoma cells and tumors. Mol Cancer Ther 2018; 17: 368-380.
[54] Yang CC, Kato H, Shindo M and Masai H. Cdc7 activates replication checkpoint by phosphorylating the Chk1-binding domain of Claspin in human cells. Elife 2019; 8: e50796.
[55] Zheng L, Meng Y, Campbell JL and Shen B. Multiple roles of DNA2 nuclease/helicase in DNA metabolism, genome stability and human diseases. Nucleic Acids Res 2020; 48: 16-35.
[56] Mahadevappa R, Neves H, Yuen SM, Jameel M, Bai Y, Yuen HF, Zhang SD, Zhu Y, Lin Y and Kwok HF. DNA replication licensing protein MCM10 promotes tumor progression and is a novel prognostic biomarker and potential therapeutic target in breast cancer. Cancers (Basel) 2018; 10: 282.
[57] Mahadevappa R, Neves H, Yuen SM, Bai Y, McCrudden CM, Yuen HF, Wen Q, Zhang SD and Kwok HF. The prognostic significance of Cdc6 and Cdt1 in breast cancer. Sci Rep 2017; 7: 985.
[58] Bravou V, Nishitani H, Song SY, Taraviras S and Varakis J. Expression of the licensing factors, Cdt1 and Geminin, in human colon cancer. Int J Oncol 2005; 27: 1511-1518.
[59] Karakaidos P, Taraviras S, Vassiliou LV, Zacharatos P, Kastrinakis NG, Kougiou D, Kouloukoussa M, Nishitani H, Papavassiliou AG, Lygerou Z and Gorgoulis VG. Overexpression of the replication licensing regulators hCdt1 and hCdc6 characterizes a subset of non-small-cell lung carcinomas: synergistic effect with mutant p53 on tumor growth and chromosomal instability-evidence of E2F-1 transcriptional control over hCdt1. Am J Pathol 2004; 165: 1351-1365.