Heliyon 10 (2024) e40179

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Novel bioinformatic approaches show the role of driver genes in the progression of cervical cancer: An in-silico study

Amir Hossein Yari [a,b,1], Parisa Shiri Aghbash [b,c,1], Mobina Bayat [c,d], Shiva Lahouti [b], Nazila Jalilzadeh [b], Leila Nariman Zadeh [b], Amir Mohammad Yari [a], Parinaz Tabrizi-Nezhadi [a], Javid Sadri Nahand [d], Habib MotieGhader [a,e,**], Hossein Bannazadeh Baghi [b,c,d,*]

[a] Department of Biology, Tabriz Branch, Islamic Azad University, Tabriz, Iran
[b] Immunology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran
[c] Department of Virology, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
[d] Infectious and Tropical Diseases Research Center, Tabriz University of Medical Sciences, Tabriz, Iran
[e] Department of Software Engineering, Engineering Faculty, Istanbul Topkapi University, Istanbul, Turkey

ARTICLE INFO

Keywords:
Bioinformatics
Systems biology
Transcription factors
Driver genes
WGCNA R package
Cervical cancer

ABSTRACT

Background: The goal of this bioinformatics research is to get a comprehensive understanding of the driver genes and their function in the development, progression, and treatment of cervical cancer. This study constitutes a pioneering attempt, adding to our knowledge of genetic diversity and its ramifications.

Material and methods: In this project, we use bioinformatics and systems biology methods to identify candidate transcription factors and the genes they regulate in order to identify micro-RNAs and LncRNAs that regulate these transcription factors and lead to the discovery of new medicines for the treatment of cervical cancer. From the differentially expressed genes available via GEO's GSE63514 accession, we use driver genes to choose these candidates. We then used the WGCNA tool in R to rebuild the co-expression network and its modules. The hub genes of each module were determined using CytoHubba, a Cystoscope plugin. The biomarker potential of hub genes was analyzed using the UCSC Xena browser and the GraphPad prism program. The TRRUST database is used to locate the TFs that regulate the expression of these genes. In order to learn how drugs, MicroRNAs, and LncRNAs interact with transcription factors, we consulted the Drug Target Information Database (DGIDB), the miRWalk database, and the LncHub database. Finally, the online database Enrichr is utilized to analyze the enrichment of Gene Ontology and KEGG pathways.

Results: By combining the mRNA expression levels of 2041 driver genes from 14 early-stage Cervical cancer and 24 control samples, a co-expression network was built. The cluster analysis shows that the collection of shared genes may be broken down into seven distinct groups, or "modules." According to the average linkage hierarchical clustering and S_{ummary} smaller than 2, we found five modules (represented by the colors blue, brown, red, green, and grey) in our research. Then, we identify 5 high-degree genes from these modules that may serve as diagnostic

* Corresponding author. Immunology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran.
** Corresponding author. Department of Biology, Tabriz Branch, Islamic Azad University, Tabriz, Iran.
E-mail addresses: habib_moti@ut.ac.ir (H. MotieGhader), hb.zadeh@gmail.com (H.B. Baghi).
[1] Amir Hossein Yari and Parisa Shiri Aghbash have equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e40179
Received 12 February 2024; Received in revised form 31 October 2024; Accepted 5 November 2024
Available online 14 November 2024

biomarkers (ZBBX, PLCH1, TTC7B, DNAH7, and ZMYND10). In addition, we identify four transcription factors (SRF, RELA, NFKB1, and SP1) that regulate the expression of genes in the co-expression module. Drugs, microRNAs, and long noncoding RNAs are then shown to cooperate with transcription factors. At last, the KEGG database's pathways were mined for information on how the co-expression module fits within them. More clinical trials are required for more trustworthy outcomes, and we collected this data using bioinformatics methods.
Conclusion: The major goal of this research was to identify diagnostic and therapeutic targets for cervical cancer by learning more about the involvement of driver genes in cancer's earliest stages.

1. Introduction

Cervical cancer is the fourth most frequent disease among women, with an estimated 604,000 cases and 342,000 deaths by 2020 [1]. Persistent infection with the human papillomaviruses (HPV) is linked to about 100 % of cervical malignancies [2]. The non-enveloped single-stranded DNA genome of human papillomaviruses, a member of the *papillomaviridae* family, is categorized into high-risk, likely high-risk, and low-risk types [3]. HPVs −16 and −18 are responsible for over 70 % of cervical malignancies among high-risk types [4,5]. Indeed, infection of mucosal squamous epithelial cells with one of the high-risk types of papillomaviruses results in precancerous lesions and cervical cancer [6,7]. The HPV virus genome has eight ORFs that encode early genes such as E1, E2, E4, E5, E6, and E7, and late genes such as L1 and L2 [8]. Oncoproteins E5, E6, and E7 are the early genes that cause persistent infection and malignancy in infected cervical cells [9,10]. For example, oncoprotein E7 attaches to the retinoblastoma (Rb) protein and destroys it, activating the transcription factor E2F1, interrupting the cell cycle, and causing unrestrained cell proliferation [1]. The p53 tumor suppressor protein controls these cell cycle abnormalities, ubiquitinated and degraded by the oncoprotein E6 by creating complexes with E6AP (E6-related protein) [11,12]. As a result, infected cervical cells become cancerous. Cancer cells move throughout the body through the circulation or lymphatic system, resulting in secondary tumors in other organs [1]. In this respect, the identification of cervical cancer mechanisms in treatment strategies for primary prevention, such as HPV vaccinations, and secondary prevention, such as screening and diagnosis of precancerous lesions, is critical [13]. Unfortunately, cervical cancer is often identified in advanced stages in developing countries, leading to inadequate screening and immunization programs and poor therapeutic efficacy. Therefore, identifying relevant treatment strategies is essential for early diagnosis, prognosis, and therapy prediction [14]. Unregulated transcription factors (TFs) play a crucial role in dysfunction tumor suppressors and cancer genes in malignancies such as cervical cancer [15]. Indeed, TFs bind to DNA promoter or gene enhancer regions as significant regulators of gene expression. As a result, following TFs' mutations and dysfunction, abnormal expression of genes, such as inhibition of expression of genes associated with differentiation and cell death, is induced [16,17]. Moreover, TFs, the primary regulators in embryonic development and early morphogenesis containing homeoboxes, are impaired in cervical cancer malignancy [15]. In this regard, transcription factors may be considered potential therapeutic targets.

On the other hand, genomic alternations are involved in the onset and progression of cervical cancer. Disorders such as copy number alterations (CNAs) [18], DNA methylation [19], and dysfunctional microRNA (miRNA) have been shown to occur during cervical cancer [20]. In addition, genetic and epigenetic mutations lead to increased malignancy and therapeutic resistance. As a result, tumor cells can maintain the ability to proliferate and metastasize to other tissues and adjacent organs [21]. These mutations occur in a set of genes called cancer driver genes, leading to disruption and impacting the hemostatic growth of crucial cell functions. Cancer driver genes can be divided into tumor suppressor genes (TSGs) or oncogenes (OGs) based on their role in disease formation. Unharmed TSGs inhibit the onset or progression of malignancy; in contrast, OGs cause malignancies due to mutations and genomic changes [22]. Significant advances have been made in surgery, chemotherapy, and radiotherapy; however, molecular biomarkers such as miRNAs have been considered for potential therapeutic purposes [23]. MicroRNAs, as small non-coding RNAs, regulate post-transcriptional gene expression. Abnormal expression of miRNAs has been reported to be directly related to cervical cancer progression, metastasis, and unfavorable treatment outcomes [14]. These small molecules, with a length of 18–22 nucleotides, can control the expression of various oncogenes or tumor suppressor genes. In addition, miRNAs can regulate 60 % of the genes encoding proteins in the human genome; in this way, they are known as the central modulators of the human genome [24,25]. The role of miRNAs has been studied in many cancers such as ovarian cancer, lung cancer, breast cancer, and cervical cancer [14]. In cervical cancer tissues and cell lines, aberrant expression of miRNA causes genetic alteration such as deletion, amplification, point mutations, and epigenetic changes such as histone changes and irregular DNA methylation [26]. However, to plan effective treatment and intervention and discover potential biomarkers, further studies on the mechanisms involved in cervical cancer and recurrence of malignancy are needed [21]. Due to the rapid growth of high throughput technologies such as microarray and RNA-seq, new insights into gene expression profiles have been proposed. As a comprehensive method, systems biology can be a new approach to identifying biomarkers and finding genes and cellular pathways involved in diseases and malignancies [27]. As a result, reconstructing a co-expression network from a complete transcriptome profile can identify the essential transcripts involved and their network interactions. The weighted gene co-expression network analysis (WGCNA) algorithm is the most appropriate method of inferring the expression gene network to identify essential genes associated with a unique character involved in malignancy and disease [28]. Hub genes, the most corresponding genes in a module, are highly functional and can play a potential role in many diseases and biomarkers for therapeutic purposes [29].

Bioinformatics and systems biology approaches are used in this project to determine prospective transcription factors and the genes

they regulate to discover new medicines, MicroRNAs and LncRNAs that govern these transcription factors concerning cervical cancer. For this purpose, we use driver genes to choose these genes from the differentially expressed genes shipped from GEO with accession number GSE63514. Next, we reconstructed the co-expression network and modules using the WGCNA package in R. CytoHubba, a plugin for Cytoscape, was used to identify the hub genes of each module. TRRUST database is used to identify the transcription factors (TFs) that target these genes. UCSC Xena browser and GraphPad prism software were used to evaluate the biomarker capability of hub genes. The DGIDB, miRWalk and LncHub online databases were used to determine the relationship between medicines, MicroRNAs and LncRNAs with transcription factors, finally Enrichr, the online database, was used to examine the enrichment of Gene Ontology and KEGG pathways (Fig. 1).

2. Materials and methods

2.1. Dataset and pre-processing

A search of the Gene Expression Omnibus database (GEO, https://www.ncbi.nlm.nih.gov/geo/) for the keyword "early-stage cervical cancer" and "cervical cancer" yielded datasets including samples from different pathological stages as well as normal controls. Finally, the dataset GSE63514 was found to meet our requirements and downloaded. It was GSE63514 [30] that assembled microarray profiles, including gene expression profiles of healthy controls and tumors from patients with cervical cancer. GEO2R was used to evaluate and normalize expression data from each of these samples. The driver gene lists related to cervical cancer were exported from the DriverDBv3 dataset, containing information on driver genes [31]. The establishment of a co-expression network relied on the

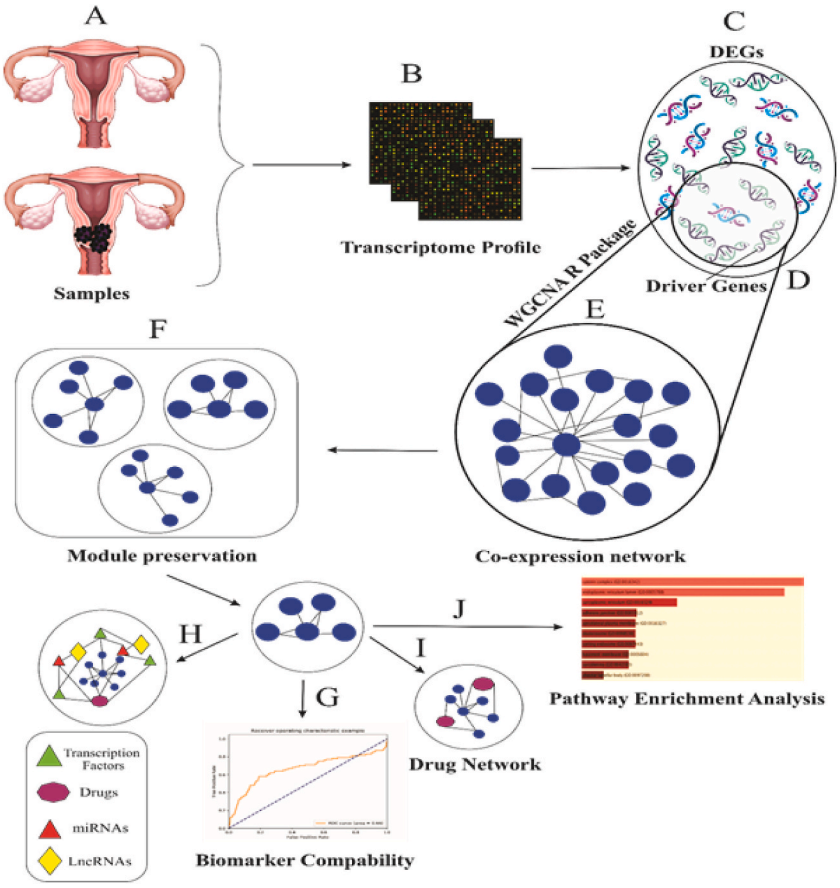

Fig. 1. The flowchart shows the overall procedure of the proposed method. At first, the transcriptome profile for healthy (control) and early-stage cervical cancer samples were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) with the accession number GSE63514 then the differentially expressed genes (DEGs) were exported with the aid of GEO2R tool (A, B and C). From http://driverdb.tms.cmu.edu.tw, we have extracted a list of driver genes known to be involved in cervical cancer (D). The WGCNA package in R was used to rebuild the networks of genes that interact with one another and Essential gene modules were generated from this network (E and F). The transcription factors (TFs) that regulate the expression of these genes were found using the TRRUST database. The biomarker potential of hub genes was analyzed using the UCSC Xena browser and the GraphPad prism program (G). Drug-MicroRNA-LncRNA-Transcription Factor Associations Determined Using Online Databases DGIDB, miRWalk, and LncHub (H). Consequently, the mRNA-drug network was rebuilt using the DGIDB database (I). Enrichr, a web-based database, was then used to investigate the enrichment of Gene Ontology and KEGG pathways (J).

selection of driver genes. We utilized DriverDBv3's (http://driverdb.tms.cmu.edu.tw/) gene symbols to select driver genes from GEO's differentially expressed genes after exporting a list of DriverDBv3's driver genes. It is possible to preserve and share high-performance functional genomic data contributed by the scientific community on the Gene Expression Omnibus (GEO), a global resource that is freely accessible [32]. DriverDBv3 is a cancer omics database that incorporates somatic mutation, RNA expression, miRNA expression, methylation, copy number variation, and clinical data using annotation bases and well-established bioinformatics techniques. To locate specific driver genes and mutations, DriverDB uses openly accessible bioinformatics techniques, which were previously featured in articles from 2014 to 2016. In this updated version of DriverDB, they aim to display cancer omics' cutting-edge data understandably [33].

2.2. Weighted gene co-expression network analysis (WGCNA)

An algorithm called WGCNA was employed to determine gene expression levels. Using the flashClust tool in the R environment, cluster analysis was utilized to discover data with outliers [34]. After no further grouping of the outlier data, Descriptive and inferential statistics between all differentially expressed genes were calculated. β (soft thresholding power) parameter based on the network's size was devised to rebuild the network for strongly connected genes and eliminate weakly linked ones. In other words, scale-free topology [35]. A topological overlap matrix (TOM) and dissimilarity measure were used to identify the modules. Hierarchical clustering was used to create gene dendrograms, and the dynamic tree-cut approach was used to identify modules of co-expressed genes as branches of the gene dendrogram [36]. DeepSplit was set to 3, and the minimum module size was 50. Combining modules with comparable gene expression patterns is possible because of their proximity. As the first and most crucial element in each module, the module eigengene was created to represent a gene expression profile and serve as a summary for each module. These were then divided into subgroups depending on how closely they were linked. Dissimilarity was set at 0.14 between the modules and the aggregated, highly expressed modules. The module-trait relationship was utilized to quantify the links between the identified modules and each early-stage cervical cancer subtype (MTR) [37].

2.3. Module preservation

"Module Preservation" was used to find modules shared across the two datasets. In addition, the Z-score was summarized using this method. The $Z_{Summary}$ score is what we used to evaluate module preservation in this investigation. No conservation, poor to moderate preservation, or important preservation are considered for modules with $Z_{summary}$ values less than 2, 2–10, or higher, respectively [38]. Modules are having a "$Z_{summary}$" value less than 2 were the ones we selected.

2.4. Identifying hub genes of module

The Maximal Clique Centrality (MCC) algorithm was the most efficient way for locating hub nodes in co-expression network modules [39]. CytoHubba [40], a plugin in Cytoscape [41], was used to compute the MCC of each node The genes with the top five MCC scores were deemed hub genes in this research. Hub genes are characterized as genes having a strong link between potential modules. Connectivity that is in the top 10 percent is deemed to be high. For instance, if the size of the module was 1000, the top 100 genes would be considered hub genes.

2.5. Identifying transcription factors (TFs)

Library of search tools for obtaining Transcription Factors (TRRUST (https://www.grnpedia.org/trrust/)) was utilized to establish the TFs and the genes they regulated the TFs-gene interactions were studied using Trust database. We may learn more about how TFs affect the module genes with the help of this investigation. In the current bioinformatic study, we predicted the TFs for the module genes that had been shown to be involved. The TFs-genes regulatory findings were produced and we found high degree TFs in the network. TRRUST (version 2) is a manually curated database of human and mouse transcriptional regulatory networks [42].

2.6. TCGA validation, biomarker capability of hub genes

RNAseq gene expression data were obtained and interpreted using the UCSC Xena Functional Genomics Explorer. Using the TCGA, a receiver operating characteristic (ROC) curve analysis was also conducted [43] to determine the potential of the hub genes expression pattern as a diagnostic biomarker for early-stage cervical cancer. Then, patient and control values were calculated based on the expression of these genes in tumor and normal samples obtained from TCGA-CESC. The area under the curve (AUC) was then determined using GraphPad Prism8 software to analyze the ROC curve. The UCSC Xena Functional Genomics Browser is an open-source, web-based program for analyzing and visualizing public data hubs. In addition, Xena provides the secure analysis and presentation of your functional genomics data collection using publicly accessible and shareable genomic/phenotypic data sets [44].

2.7. Data enrichment

The modules' genes were chosen for gene ontology (GO) and pathway enrichment analysis. Enrichr (https://maayanlab.cloud/Enrichr/), the web-based program, was used for pathway enrichment analysis to establish which biological pathways are

influenced by each module's genes. Then, using the GO operation, we gathered information on biological processes (BP), cellular components (CC), and molecular activities (MF).

Enrichr is a comprehensive resource for curated gene sets and a search engine that accumulates biological knowledge for further biological discoveries [45].

2.8. Drug, MicroRNA and LncRNA-TF network

After identifying the most critical TFs, we constructed Drug, MicroRNA and LncRNA-TF network. For this purpose, the DGIdb (https://maayanlab.cloud/Enrichr/), miRWalk (http://mirwalk.umm.uni-heidelberg.de/) and lncHUB (https://maayanlab.cloud/lnchub/) databases were used respectively. As a potential treatment for cervical cancer, the results may suggest MicroRNAs and LncRNAs as options for potential drugs. Using the Drug Gene Interaction database, putative new pharmaceuticals were suggested (DGIdb). This database gathers drug-gene interaction information from six separate sources (My Cancer Genome 39, TALC 40, TEND 41, PharmGKB 42, TTD43, and DrugBank 44) [46]. miRWalk is an open-source software giving a simple interface that provides predicted and confirmed miRNA-binding sites of known genes of human, mouse, rat, dog and cow [47]. lncHUB is online tool for Gene-lncRNA correlation from the 11,284 TCGA samples [48].

2.9. Drug and module genes network

The DGIDB database was used to generate a drug-gene network for the module's retrieved genes, and the cytoscape program was used to view the network. Finally, the high degree drugs were introduced.

2.10. Statistical analysis

Statistical techniques were used in order to ensure the resilience of the results. The gene expression levels of cervical cancer samples and normal controls were compared using the student's t-test, with a significance threshold of $p < 0.05$. Utilizing GraphPad Prism software, a Receiver Operating Characteristic (ROC) curve analysis was performed to assess the diagnostic capability of the discovered hub genes by computing the area under the curve (AUC) for each gene. The False Discovery Rate (FDR) approach was used to apply numerous testing adjustments once selection criteria were defined. Prior to the *t*-test, normality and homogeneity of variance were evaluated; non-parametric alternatives were taken into account when assumptions were broken. R (version 4.2.3) and GraphPad Prism (version 8) were used for all analyses.

3. Results

3.1. Dataset and preprocessing

To discover genes associated with various stages of cervical cancer, we analyzed the GSE63514 microarray data set. The characteristics of this dataset are detailed in Table 1. This dataset consists of various groups but we select twenty-four Normal samples and fourteen early stage cervical cancer samples groups to further investigations. After removing nameless genes and duplicate genes we identified 22186 genes that expression levels were significantly different (Supplementary file 1). The genes related with cervical cancer are shown in supplementary file 2. There are 2041 driver genes related with cervical cancer, and we choose the driver genes from the Differentially Expressed genes.

3.2. Weighted gene co-expression network analysis

A co-expression network was constructed using the mRNA expression levels of 2041 driver genes in 14 early-stage cervical cancer samples and 24 healthy controls. Using WGCNA with a screen-out power of 1–20 and a scale-independent topology index (R2) of 10 as the ideal value, we were able to increase the network's average connectedness and reach an R2 of 0.8550. Fig. 2 uses R2, the mean

Table 1
Specifications of the GSE63514 dataset.

Groups	Source name	Number of samples	Expression Array:
Normal samples (Healthy control samples)	Normal cervical epithelium	24 samples	**GPL570:** [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
Cancer	Cervical squamous epithelial cancer	28 samples	
CIN1 lesions	Cervical intraepithelial neoplasm, low grade lesion	14 samples	
CIN2 lesions	Cervical intraepithelial neoplasm, moderate grade lesion	22 samples	
CIN3 lesions	Cervical intraepithelial neoplasm, high grade lesion	40 samples	

connection, and a range of soft threshold values to illustrate this scale dependency (Fig. 2). After that, an adjacency matrix of expression data was generated. That's why we put it to use in constructing our topological overlap matrix (TOM).

3.3. Module preservation

To check whether the modules were shared across the two groups, we next created two gene dendrograms, where the modules were shown in distinct colors (Fig. 3). Contiguous gene sets have been organized into 5 different modules. Next, we built a comprehensive clustering tree with the help of the WGCNA software tool. For example, in Fig. 2A, where each line represents a different gene, clusters of genes with similar expression patterns are shown by lines of the same color. All the genes that are grayed out cannot be mixed. You could see seven different modules. Our analysis of the data using average linkage hierarchical clustering (Fig. 3a) revealed five modules represented by the hues blue, brown, green, grey, and red. The number of genes in each module is shown in Table 2. The grey module contains genes that do not fit into the other categories. Based on the link between the components, a cluster tree diagram (Fig. 3b) was generated. Additionally, the module Preservation function supplied a Z-score summary for every module. We found one module containing $Z_{summery}<2$. This module is colored red. The very well-preserved modules were collected for further examination (Fig. 4).

3.4. Identifying hub genes of red module and biomarkers

CytoHubba, a plugin in Cytoscape, was used to calculate the MCC of each node. This study considered the genes with the top 5 MCC values as hub genes of module. With the help of this plugin the hub genes of red module were explored (Fig. 5).

3.5. Identifying transcription factors (TFs)

The TFs-gene interactions were studied using TRRUST database. We may learn more about how TFs affect the module genes with the help of this investigation. In the current bioinformatic study, we predicted the TFs for the module genes that had been shown to be involved. The TFs-genes regulatory findings were produced and we found high degree TFs in the network (Table 3 and Fig. 6).

3.6. TCGA validation, biomarker capability and overall survival of hub genes

Additionally, the TCGA database was used to validate the results of the GEO datasets and co-expression network. The TCGA dataset consists of 306 cervical cancer samples and 21 normal samples, all screened via RNA sequencing (RNA-seq). This comprehensive dataset provides a robust foundation for our analysis, as RNA-seq allows for a detailed examination of gene expression profiles at a high resolution. The expression of red cluster genes with a high degree of expression was examined in cervical cancer and normal tissues. TCGA results demonstrate that compared to normal specimens, these genes are significantly dysregulated in cervical cancer samples (Supplementary file3). In addition, as shown in Supplementary file3, TTC7B and DNAH7 expressions have the ability to act as diagnostic markers for cervical cancer and normal samples, as measured by their AUC. Table 4 displays the p-value, AUC, and

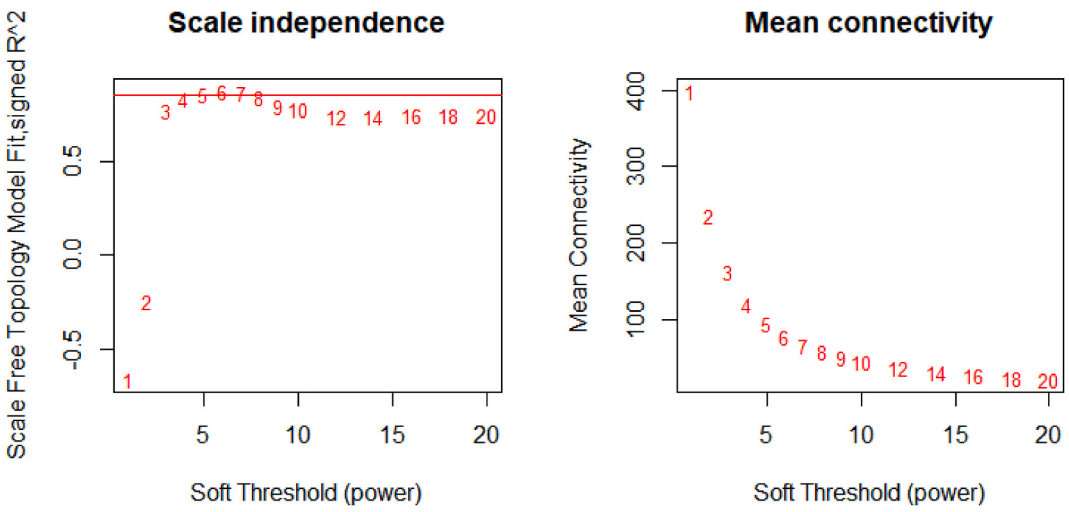

Fig. 2. A: R2 demonstrates how the power value affects the scale independence of co-expression of the modules. B: The average co-expression network connection for different amounts of soft-thresholding power. To determine the significance of the relationships between genes, WGCNA uses weighted gene correlation network analysis (WGCNA), which involves calculating initial gene correlations and then applying a power function to them. After that, modules are created to organize genes with comparable expression patterns. The network with the fewest nodes and maximum degree is the most scale-independent one, and we can quantify this property with a power value.

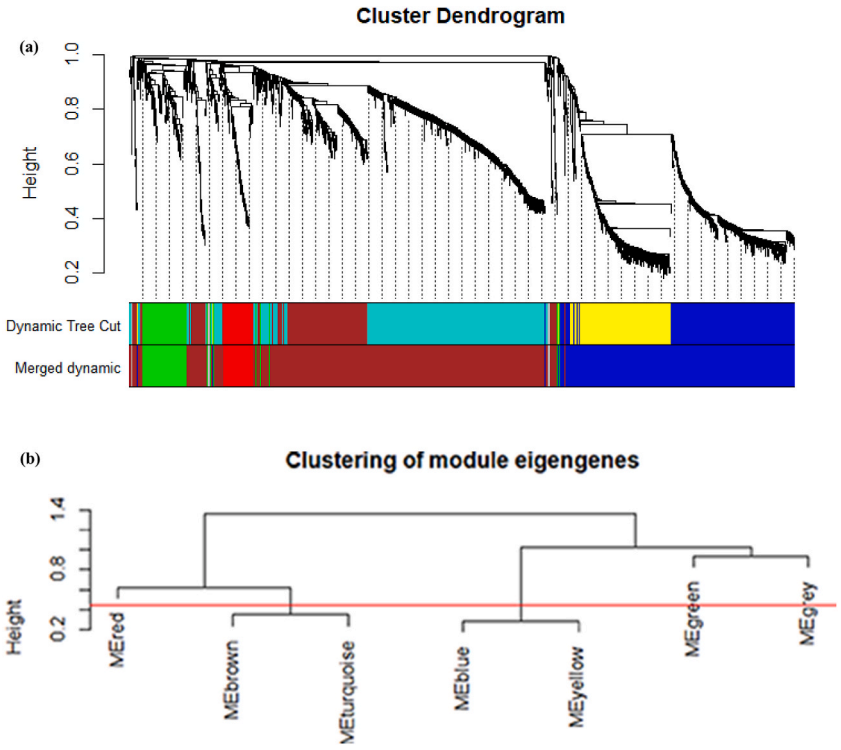

Fig. 3. Gene co-expression is being explored. Analyses of gene expression networks have shown distinct modules of co-expressed genes. The vertical axis represents the value of gene expression, while the horizontal axis represents the number of genes in the sample. In a dendrogram, a vertical line represents a gene, and a branch denotes a module of highly co-expressed genes (one color). Modules that have been detected appear in the first color band, whereas modules that have been combined appear in the second. (b) Module eigengenes clustered. The components of Part A were integrated in this manner. The threshold for module eigengene dissimilarity, shown in red, is the modules' merged level. In a single main branch, modules below a specific threshold will be merged into a single module. 28 distinct modules were discovered and condensed into 27 primary modules.

Table 2
This table indicates the $Z_{summery}$ and high degree genes of each module.

Module name	Module size	$Z_{summary}$
Blue	392	30
Brown	328	16
Red	91	1.80
Green	144	2.70
Grey	10	−0.95

biomarker potential of these genes. The expression levels, ROC curve, and overall survival of hub genes are shown in Additional file3.

3.7. Data enrichment

We used the Enrichr database for the gene ontology and pathways enrichment of the genes employed in this experiment. The ontology indices for each of the three routes, Biological Process (BP), Cellular component (CC), Molecular function (MF), and KEGG Pathway, are analyzed independently for the genes of the red module (Fig. 7, Table 5 and supplemental file 4 has the entire data).

3.8. Drug, MicroRNA and LncRNA -TF network

After identifying the most important TFs, a Drug, MicroRNA, and LncRNA-TF network was created. The DGIdb, miRWalk, and lncHUB databases were used for this purpose. As a potential treatment for cervical cancer, the results may suggest MicroRNAs and LncRNAs as options for potential drugs. Using the Drug Gene Interaction database, putative new pharmaceuticals were suggested (DGIdb). The co-expression genes, transcription factors that control these genes, LncRNAs, MicroRNAs, and medicines that target these transcription factors (Fig. 8). Additional file 5 and Table 6 describe the network in detail.

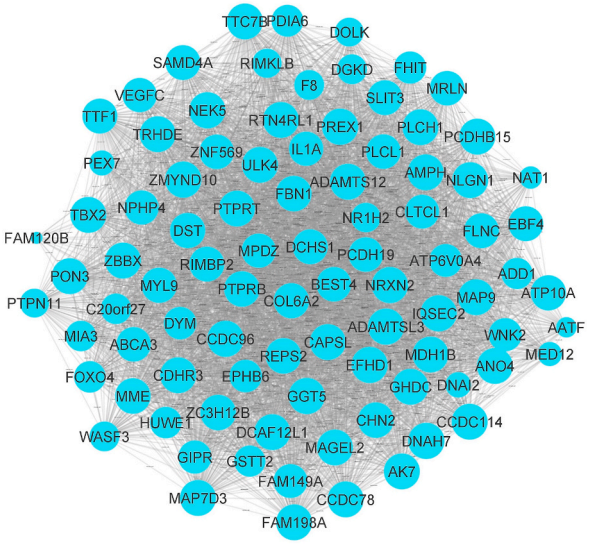

Fig. 4. Red module co expression network. Blue nodes represent the genes and edges represent the co expression of these genes.

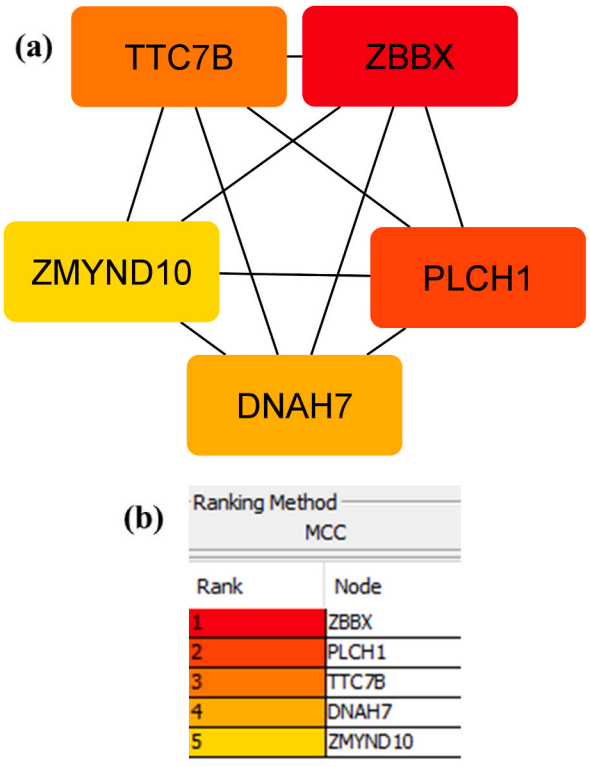

Fig. 5. The hub genes of red module and their rank.

3.9. Drug-target genes (TGs) network

DGIDB database was also used to construct a drug-gene network for module's retrieved genes and cytoscape software was used to visualize the network. (Table 7, Fig. 9, and supplementary file 6).

Table 3
TFs and their target genes.

Transcription Factors (TFs)	Description	Genes
SRF	serum response factor (c-fos serum response element-binding transcription factor)	MYL9, SAMD4A
RELA	v-rel reticuloendotheliosis viral oncogene homolog A (avian)	REPS2, F8, IL1A
NFKB1	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	REPS2, IL1A, F8
SP1	Sp1 transcription factor	GIPR, MME, FBN1

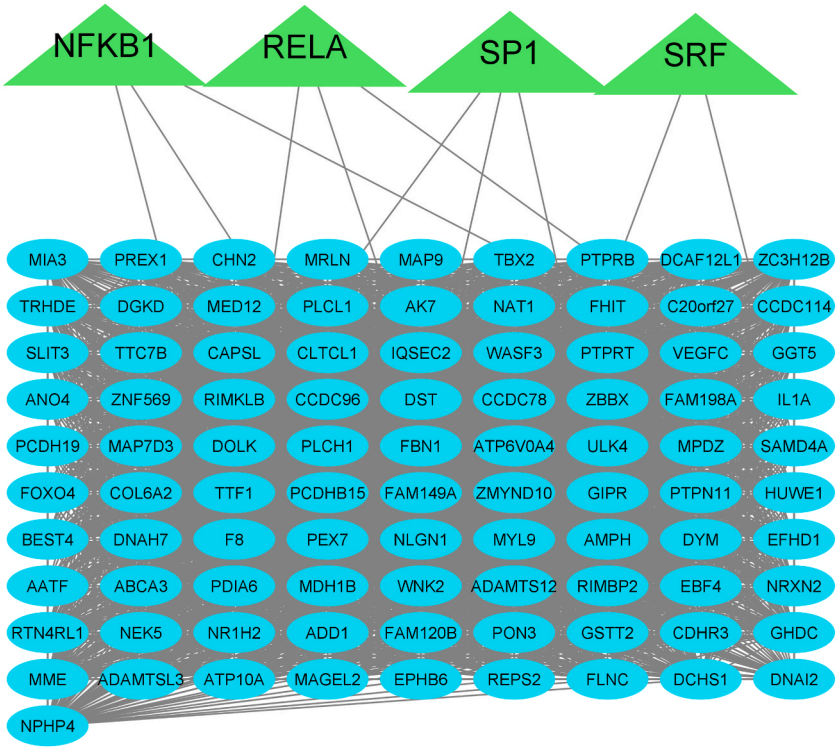

Fig. 6. TF-genes network. Blue nodes represent the genes and edges represent the co expression of these genes and green triangles show the transcription factors.

Table 4
The p-value, AUC, and biomarker capability of these genes.

Gene	P-value	AUC	Biomarker capability
ZBBX	0.5843	0.6126	✘
PLCH1	0.3414	0.6385	✘
TTC7B	0.0047	0.9737	✓
DNAH7	0.0539	0.7805	✓
ZMYND10	0.9817	0.5038	✘

4. Discussion

In the field of cervical cancer prevention, screening, and treatment, many advances have been made, such as modern radiotherapy and targeted treatment; however, the treatment of cervical cancer has not improved significantly [49–51]. Additionally, the death rate from cervical cancer is remains high in low-income countries despite the usefulness of vaccination and regular medical tests in avoiding the disease [52]. In this regard, considering the need to search for potential regulators involved in the development of cervical cancer, our study was conducted to identify biomarkers related to cervical cancer and potential therapeutic targets.

According to reports, there are several variables and different stages involved in the development of cervical cancer [53,54]. Also, in addition to human papilloma virus (hr-HPV) infection, other factors such as epigenetic abnormalities, changes in DNA methylation, as well as environmental factors such as smoking, alcohol consumption, immune status, number of sexual partners, and genetic background, plays an important role in causing cervical cancer [55,56].

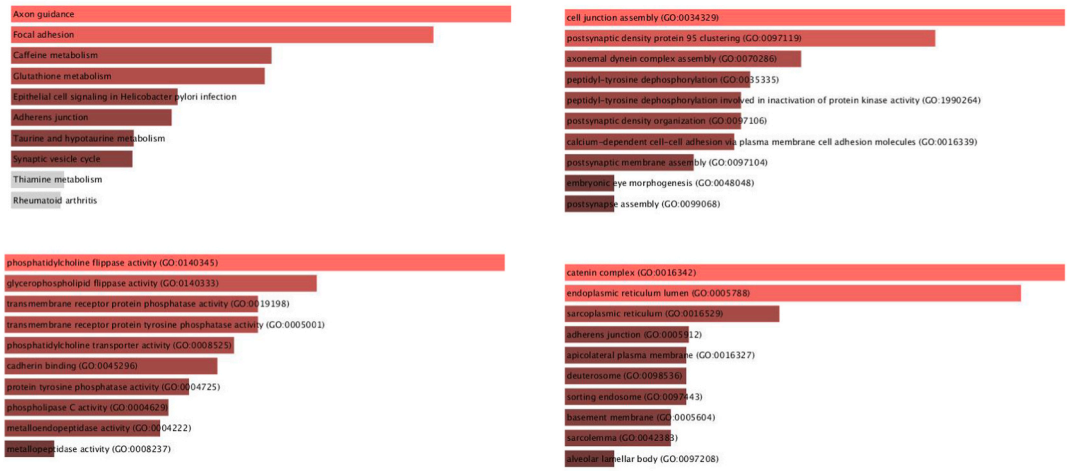

Fig. 7. Samples of Gene ontology and pathway enrichment.

Table 5
samples of Gene ontology and pathway enrichment.

Biological process (BP)		Cellular component (CC)		Molecular function (MF)		KEGG PATHWAY	
Term	P-Value	Term	P-Value	Term	P-Value	Term	P-Value
cell junction assembly (GO:0034329)	1.03E-04	catenin complex (GO:0016342)	0.008739	phosphatidylcholine flippase activity (GO:0140345)	2.03E-04	Axon guidance	**0.009593**
postsynaptic density protein 95 clustering (GO:0097119)	2.03E-04	endoplasmic reticulum lumen (GO:0005788)	0.009751	glycerophospholipid flippase activity (GO:0140333)	0.001312	Focal adhesion	**0.013418**
axonemal dynein complex assembly (GO:0070286)	4.11E-04	sarcoplasmic reticulum (GO:0016529)	0.017856	transmembrane receptor protein phosphatase activity (GO:0019198)	0.002357	Caffeine metabolism	**0.026994**
peptidyl-tyrosine dephosphorylation (GO:0035335)	5.37E-04	adherens junction (GO:0005912)	0.022404	transmembrane receptor protein tyrosine phosphatase activity (GO:0005001)	0.002357	Glutathione metabolism	**0.027794**
peptidyl-tyrosine dephosphorylation involved in inactivation of protein kinase activity (GO:1990264)	5.63E-04	apicolateral plasma membrane (GO:0016327)	0.022546	phosphatidylcholine transporter activity (GO:0008525)	0.002988	Epithelial cell signaling in Helicobacter pylori infection	**0.040496**

In this regard, with the development of high-throughput technologies, new biomarkers were identified for the early detection of this type of cancer through the analysis of microarray datasets. In addition, in order to identify biomarkers and discover the molecular mechanism involved in cancer, microarray and sequencing technologies are widely used [57].

Impairment in the immune function of people with cervical cancer, in turn, can lead to the development of effective treatment strategies. In this regard, Cheriyan et al. investigated the anergy of T cells through anti-CD3 capacity in order to extract cell signaling proteins in people with cervical cancer [58]. Also, they showed that the decrease in the expression and function of downstream proteins or defects in TCR, lead to disruption in the transfer of NF-κB to the nucleus and the levels of other proteins such as TCR-z, CD3-e, zap 70, p56lck, NF-κB-P50, Rel-A, Rel-B, Rel-C [58]. Levey and Srivastava also reported that the reduction of TCR-z, zap-70 or p56lck protein expression, as well as disruption of Rel-A transfer from cytoplasm to nucleus, cause T cell dysfunction in response to antigenic stimulation [59]. In addition, the reduction of TCR-z chain activation and the inability to transfer components of the NF-κB pathway from the cytoplasm to the nucleus, and subsequently the disruption of IL-2 secretion, can be one of the causes of immune response disorders [58]. In this regard, our study also showed that the level of expression of NF-κB as well as Rel-A decreases significantly in the early stages of cervical cancer compared to normal samples. As a result, it can be stated that disruption in the signaling pathway of T lymphocytes and subsequently transcription factors such as NF-κB and Rel-A in cervical cancer patients leads to disruption in their anti-tumor responses.

It has been reported that NF-κB has an important role in cell growth, survival, and apoptosis. Also, inhibition or induction of apoptosis following NF-κB activation depends on different factors and cell types. In addition, rel/nuclear factor kappa B (NF-κB) proteins have two protein groups, including proteins that require proteolytic processing and proteins that do not require proteolytic processing [60]. The first group proteins are Rel-A (known as p65), Rel-c and Rel-B, and the second group proteins are NF-κB1 (known as p105) and NF-κB2 (known as p100). Also, the proteins of the second group produce p50 and p52 proteins, respectively [60]. Besides,

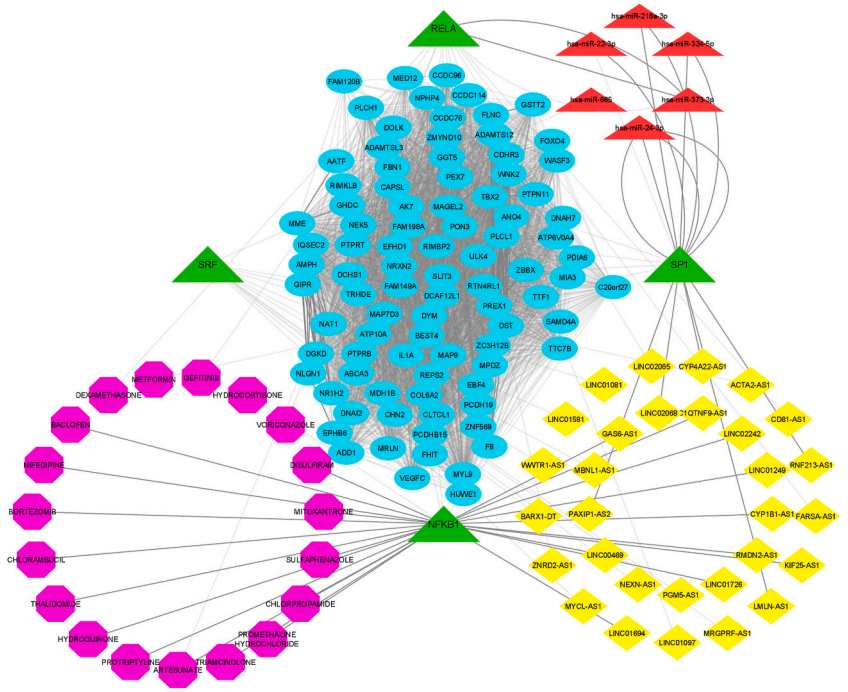

Fig. 8. The co-expression genes and the transcription factors that regulate these genes and the LncRNAs, MicroRNAs and drugs that target these transcription factors. Genes are represented by blue ellipses, green triangles show transcription factors, red triangles indicate the MicroRNAs whereas yellow diamonds show LncRNAs and purple hexagons represent medications.

Table 6
genes and the transcription factors that regulate these genes and the LncRNAs, MicroRNAs and drugs that target these transcription factors.

TF	Genes	Drugs	MicroRNAs	LncRNAs
SRF	MYL9, SAMD4A	Not reported	hsa-miR-665	MBNL1-AS1, LINC01581, BARX1-DT MYCL-AS1, MRGPRF-AS1, ACTA2-AS1 PGM5-AS1, NEXN-AS1, LINC01081 WWTR1-AS1
RELA	REPS2, F8, IL1A	ARTESUNATE, VORICONAZOLE, GEFITINIB, DEXAMETHASONE	hsa-miR-373-3p hsa-miR-373-3p hsa-miR-373-3p	FARSA-AS1, CD81-AS1, LINC02065 LINC01097, ZNRD2-AS1
NFKB1	REPS2, IL1A, F8	BACLOFEN, NIFEDIPINE BORTEZOMIB, ARTESUNATE CHLORAMBUCIL, THALIDOMIDE HYDROQUINONE, PROTRIPTYLINE TRIAMCINOLONE, PROMETHAZINE HYDROCHLORIDE, CHLORPROPAMIDE SULFAPHENAZOLE, MITOXANTRONE DISULFIRAM, HYDROCORTISONE	Not reported	LINC01726, CYP1B1-AS1, LINC02242 LINC00469, KIF25-AS1, RMDN2-AS1 LINC01249, C1QTNF9-AS1, GAS6-AS1 LINC01694
SP1	GIPR, MME, FBN1	METFORMIN	hsa-miR-24-3p hsa-miR-24-3p hsa-miR-324-5p hsa-miR-24-3p hsa-miR-216a-3p hsa-miR-22-3p hsa-miR-24-3p hsa-miR-216a-3p hsa-miR-22-3p hsa-miR-324-5p	LMLN-AS1, CYP4A22-AS1, LINC02068 RNF213-AS1, PAXIP1-AS2

Table 7
Drugs and their target genes.

TGs	Drugs
NAT1	DESVENLAFAXINE, DESVENLAFAXINE SUCCINATE, DEXMETHYLPHENIDATE, HYDROCHLORIDE, DEXTROAMPHETAMINE, DEXTROAMPHETAMINE ADIPATE, DEXTROAMPHETAMINE SULFATE, DIETHYLPROPION HYDROCHLORIDE, DOXEPIN HYDROCHLORIDE, DULOXETINE, DULOXETINE HYDROCHLORIDE, FLUOROURACIL, FRUCTOSE, FUROSEMIDE, GEFITINIB, GUANETHIDINE, MONOSULFATE, HYDROCHLOROTHIAZIDE, HYDROXYCHLOROQUINE, IMIPRAMINE, HYDROCHLORIDE, LENALIDOMIDE, LEVOMILNACIPRAN HYDROCHLORIDE, LISDEXAMFETAMINE DIMESYLATE, MAPROTILINE HYDROCHLORIDE, MAZINDOL, MESALAMINE METHYLPHENIDATE HYDROCHLORIDE, NEFAZODONE, HYDROCHLORIDE, NORTRIPTYLINE, OCRIPLASMIN, OLANZAPINE, ORPHENADRINE CITRATE, ORPHENADRINE HYDROCHLORIDE, PHENMETRAZINE HYDROCHLORIDE PHENTERMINE, PHENTERMINE HYDROCHLORIDE, PROTRIPTYLINE HYDROCHLORIDE PSEUDOEPHEDRINE HYDROCHLORIDE, PSEUDOEPHEDRINE POLISTIREX, RANITIDINE RILONACEPT, SACUBITRIL, SORBITOL, SPIRONOLACTONE, SUNITINIB
ADD1	AFLIBERCEPT AMITRIPTYLINE HYDROCHLORIDE AMOXAPINE AMPHETAMINE AMPHETAMINE ADIPATE
VEGFC	TORSEMIDE TRIMIPRAMINE MALEATE VANDETANIB VENLAFAXINE HYDROCHLORIDE
IL1A	BUPROPION HYDROCHLORIDE COLLAGENASE CLOSTRIDIUM HISTOLYTICUM
F8	BENZPHETAMINE HYDROCHLORIDE BETHANIDINE SULFATE BEVACIZUMAB

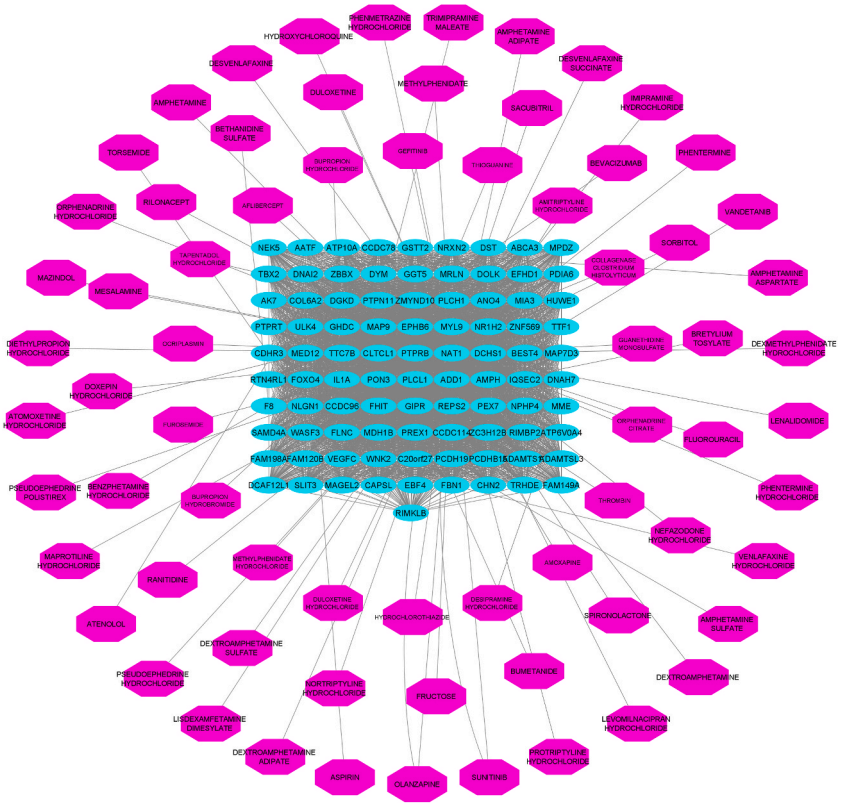

Fig. 9. Drugs and their target genes. Genes are represented by purple ellipses whereas yellow hexagons show drugs.

it has been reported that the transcriptional activity of NF-κB is mostly regulated by Rel-A; in addition, p50-Rel-c dimers have the lowest frequency [60]. It is worth mentioning that one of the causes of drug resistance in cancers can be mentioned the nuclear factor kappa B (NF-κB) transcription factor. In this regard, Garg et al., showed that in patients with cervical cancer treated with chemotherapy radiation therapy (CRT), there is an inverse relationship between the transcription factor NF-κB and its outcomes [61]. In addition, Altoos et al., also reported that poor clinical results in patients treated with CRT could be caused by increased cytoplasmic expression of NF-κB-p65 [62].

Moreover, P65 subunit is involved in different types of human cancers, and in malignancies such as esophagus, breast, liver cells, digestive system, cervix, pancreas, and lung, there are P65 gene expression changes [63]. Xue Pan stated in 2008 that NF-kB p65 induces apoptosis in malignancy of pancreatic cancer cells [64]. However, Sergio Vaira in 2008 reported that NF-κBp65 leads to inhibition of apoptosis in cancer cells and malignant progression and has an oncogenic role [65].

Sp1 is another transcription factor involved in cervical cancer and adjacent tissues, and it is a clinical biomarker and therapeutic target in patients' resistant to radiotherapy [66]. Besides, studies indicate that Sp1 as a critical mediator leads to the control of HPV18 function in cervical cancer, and is closely related to metastasis, cancer recurrence, and malignant progression [67–69]. Since Sp1 is a type of cell cycle regulator, Deng et al., in 2019, has been demonstrated that its excessive decrease and increase lead to the acceleration and inhibition of the cell cycle in the G2/M phase of cervical cancer cells, respectively [66]. Also, considering that cells are sensitive to radiation therapy in G2/M phase, so increasing Sp1 transcription factor causes inhibition of cells in this phase and increase sensitivity to radiation therapy [66]. In fact, this transcription factor leads to increased expression of CDK1, the main target molecule of the G2/M cell cycle checkpoint, in cervical cancer cells [66]. Moreover, Deng et al., showed that the number of changes in CDK1 expression is directly related to the level of Sp1, especially in patients' resistant to radiation therapy. Also, our study showed that in the early stages of cervical cancer, the expression of Sp1 increases significantly compared to normal people, which can be concluded that this increased expression increases cell growth and also resistance to radiation therapy in patients with cervical cancer [66]. Furthermore, transcription factors and epigenetic regulators can lead to the modulation of the expression level of lncRNAs [70,71]. It has also been reported that the expression level of lncRNAs changes in tumor cells such as cervical cancer cells [71]. For example, the transcription factor Sp1 leads to the overexpression of lncRNA SNHG14 and subsequently to the development of renal cell malignancy [72]. Also, in gastric carcinoma, Sp1, through activating EZH2, leads to increased transcription of lncRNA UCA1 and increased growth and progression of malignancy [73]. In addition, Jang et al., showed that in cervical cancer, the Sp1 transcription factor interacts with the LUCAT1 promoter to regulate the expression of this gene [70]. In other words, the low expression of Sp1 is directly related to the expression of LUCAT1 and leads to the suppression of its expression. It is worth mentioning that LUCAT1 is an oncogene identified in cervical cancer and leads to the development of cervical cancer [70]. In addition, it has reported that in normal cells the transcription factor Sp1 binds to the GC box in the MALAT1 promoter [74]; while wang et al., in 2021, illustrated that in cells infected with hr-HPV, E7 oncogene with over activation of Sp1 leads to increased activity and transcription of MALAT1 and malignant progression [75]. Also, Chou et al., reported in 2022 that in the malignancy of osteosarcoma cells, the overexpression of transcription factors Sp1 and NF-κB (Rel-A) may play an important role in increasing metastasis [76]. As a member of the MADS box transcription factor family, SRF plays an important role in important cellular activities [77]. Also, studies indicate that SRF, together with MRTF and TCF, lead to the regulation of migration, invasion, and growth mechanisms through the signaling and cytoskeleton-dependent pathway [78]. However, the mechanism of action of SRF in the proliferation and invasion of cells involved in cervical malignancy is not clear. In this regard, in the study conducted by MA et al., in 2019, on cervical cancer cells, it was reported that inhibiting the expression of SRF through the control of the early growth response pathway-1 (Egr-1), leads to a decrease in the growth and proliferation of cervical cancerous cells [79]. In this regard, our study also showed that the level of SRF expression in normal samples has decreased significantly compared to the initial stage of cervical cancer. In addition, Silverman and Collins showed in their study that SRF increases the transcription of Egr-1 gene [80]. In other words, increased expression of SRF in cancer cells leads to decreased expression of E-cadherin and increased expression of N-cadherin in EMT [81–83]. It has been shown that EMT as an important process in cell metastasis leads to the loss of polarity of epithelial cells and subsequently increases cell migration [84,85]. It has also been reported that following the inhibition of SRF expression due to the transfer of SRF siRNA to cervical cancer cells, the expression levels of N-cadherin and E-cadherin decrease and increase, respectively; and subsequently leads to the inhibition of proliferation and invasion of cells [86]. Also, MA et al., observed that following the inhibition of SRF expression in cervical cancer cells, the expression level of Egr-1 decreases, which subsequently affects the expression level of E-cadherin and N-cadherin and cell metastasis in cervical cancer cells [79]. As a result, they reported that in normal tissue compared to cancer tissue in the early stages, the expression level of SRF is significantly low [79]. In addition to transcription factors, hub genes are also involved in the process of malignant progression. For example, axonal dynein heavy chain (DNAH) genes, as a family of genes encoding axonal dynein heavy chain, are involved in cell motility; in such a way that changes in DNAH gene expression play a role in the process of malignancy [87]. Reports indicate that sensitivity to chemotherapy increases in patients with gastric carcinoma. In other words, mutated DNAH in patients with gastric carcinoma leads to increased sensitivity to chemotherapy in these patients. However, resistance to chemotherapy has been reported in some of these patients [87]. In fact, functionally, the activity of ATPases is affected by DNAH proteins; and they play a role in biological processes such as microtubule movement, cilia assembly, cilia movement and internal/external dynein arm assembly [88–90]. ZMYND10, as another type of hub gene, is a 50kd protein that has a DNA binding domain on the MYND type at the C-terminus [91]. This region is mostly found in transcription repressors. This gene is involved in many malignancies such as lung cancer, glioma tumors, ovarian cancer, liver cancer, esophageal squamous cell carcinoma, neuroblastoma, and myelodysplastic syndrome, and it is reduced due to genetic and epigenetic changes [92]. In addition, ZMYND10 as a tumor suppressor leads to induction of apoptosis, cell cycle arrest, and inhibition of angiogenesis [92]. Also, reports indicate that increased expression of this gene leads to increased sensitivity to chemotherapy in malignancies [93,94]. ZMYND10 is deleted following promoter hypermethylation in multiple cancer types. In other words, ZMYND10

is a tumor suppressor gene and has epigenetic regulatory activity, whose hypermethylation is associated with poor prognosis in all types of cancer [92]. In this regard, during their study on breast cancer, Wang et al., in 2019, reported that the inhibition of NEDD9 oncogene expression through the reregulation of miR-145-5p by ZMYND10 leads to the inhibition of tumorigenesis, cell metastasis, and malignant progression [92]. In addition, they showed that following ZMYND10 demethylation, its expression increases and thus leads to inhibition of malignant progression. In other words, ZMYND10 through miR145-5p/NEDD9 signaling storm leads to inhibition of cell migration and spread of breast malignancy [92]. Concerning the novelty of this investigation, it is possible to assert that this research constitutes a pioneering undertaking, given that to date, no study has examined the particular role of driver genes in progression of cervical cancer. The transcription factors and driver genes that may be essential to the development of cervical cancer have been clarified by this work. Nonetheless, the intrinsic difficulties and constraints linked to our in-silico methodology need careful deliberation. The Gene Expression Omnibus (GEO) and DriverDBv3, two freely available databases, provided the gene expression data used in this investigation. Although these sites provide insightful information, there might be wide variations in the data's representativeness and quality. Variations in the procedures used to collect the samples, prepare them, or run the experiments might create biases and compromise the validity of our findings. Furthermore, the findings' generalizability may be limited by the fact that the dataset (GSE63514) may not fully capture the genetic variety of cervical cancer patients across other ethnicities. Experimental confirmation is necessary since our predictions are computational in nature and were produced using sophisticated bioinformatics tools such as CytoHubba and Weighted Gene Co-expression Network Analysis (WGCNA). The biological significance of the discovered genes is yet unknown given the absence of in vitro or in vivo investigations. Thorough experimental research is necessary to support the hub genes and transcription factors' promise as diagnostic and therapeutic tools. Despite their strength in processing complicated data, bioinformatics tools sometimes depend on mathematical models and assumptions that may not fully capture biological facts. For example, the scale-free network structure assumption made by WGCNA could not hold true in every situation. Moreover, the identification of hub genes via network connection does not always correspond to their functional significance in the advancement of cervical cancer. Such computational restrictions may result in the omission of important genes linked to the illness or false positive results. Numerous variables, including as environmental influences, genetic background, and epigenetic alterations, affect the levels of gene expression. Although insightful, our attention on mRNA expression levels does not fully capture the regulatory environment, which includes activity realities, protein expression, and post-transcriptional regulation. As a result, firm judgments on the functions of discovered genes at the protein level or in more general cellular settings cannot be made. As useful as in-silico techniques are in producing ideas and possible treatment targets, their conversion into clinically applicable biomarkers or approaches requires careful experimental confirmation. As a result of patient heterogeneity, illness stage, and treatment history, the discovered transcription factors and driver genes may vary in their usefulness as biomarkers among clinical samples. In order to evaluate the therapeutic significance of our results, clinical trials and patient-based research are essential. Based on bioinformatics resources like DGIDB, miRWalk, and LncHub, we have predicted the interactions between transcription factors and non-coding RNAs (miRNAs and LncRNAs). However, because non-coding RNAs' regulatory functions are so context-dependent, these links need to be confirmed through experimentation. Further studies are needed to validate these interactions specifically in cervical cancer. With the purpose of finding effective drugs drug screening will be employed, utilizing databases like the Drug-Gene Interaction Database (DGIDB) to explore interactions between existing drugs and the identified genes. DGIDB has suggested potential drug repurposing options for cervical cancer treatment. Additionally, virtual screening of drug libraries will be conducted to identify novel compounds that can interact with these driver genes. Although preliminary interactions have been identified, further experimental validation through in vitro assays is necessary. Future studies will focus on assessing the efficacy of these drugs, aiming to refine therapeutic strategies for cervical cancer based on bioinformatics findings. In conclusion, the limits of in-silico research must be recognized even if this study provides a thorough overview of prospective molecular targets in cervical cancer through bioinformatics and systems biology methodologies. Confirmation of these results and assessment of their potential in therapeutic and diagnostic applications necessitate experimental validation and clinical investigations.

5. Conclusions

The results of this study highlight the significance that driver genes play in the development of cervical cancer, highlighting them as potential targets for diagnosis and treatment. With this research, the authors hoped to learn more about the function of driver genes in cervical cancer, both as diagnostic indicators and as targets for treatment. We employed bioinformatics methods to do this task. We mapped the co-expression network for 2041 driver genes to identify the key early-stage cervical cancer genes. We next use bioinformatics techniques to the data to identify candidate diagnostic biomarkers and treatment drugs. Computational approaches provide the backbone of our investigation. Therefore, clinical studies both in vitro and in vivo are required. We expect the findings will inform efforts to refine methods of diagnosis, prevention, and treatment.

CRediT authorship contribution statement

Amir Hossein Yari: Writing – review & editing, Writing – original draft, Investigation, Formal analysis. **Parisa Shiri Aghbash:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis. **Mobina Bayat:** Writing – review & editing, Investigation. **Shiva Lahouti:** Writing – review & editing. **Nazila Jalilzadeh:** Writing – review & editing. **Leila Nariman Zadeh:** Writing – review & editing. **Amir Mohammad Yari:** Writing – review & editing. **Parinaz Tabrizi-Nezhadi:** Writing – review & editing. **Javid Sadri Nahand:** Writing – review & editing. **Habib MotieGhader:** Writing – review & editing, Supervision, Project administration, Investigation. **Hossein Bannazadeh Baghi:** Writing – review & editing, Supervision, Project administration,

Investigation.

Consent to participate

None.

Consent to publish

None.

Ethical approval

None.

Data availability

The data supporting this study's findings are available from the corresponding author upon request.

Funding

This study didn't receive any funding.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This project was supported by the Immunology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e40179.

References

[1] Y. Huang, et al., Rab31 promotes the invasion and metastasis of cervical cancer cells by inhibiting MAPK6 degradation, Int. J. Biol. Sci. 18 (1) (2022) 112.
[2] A. Kumar, E. Rathi, S.G. Kini, Drug repurposing approach for the identification and designing of potential E6 inhibitors against cervical cancer: an in silico investigation, Struct. Chem. 31 (1) (2020) 141–153.
[3] K. Van Doorslaer, et al., The Papillomavirus Episteme: a major update to the papillomavirus sequence database, Nucleic Acids Res. 45 (D1) (2017) D499–D506.
[4] N. Muñoz, et al., Epidemiologic classification of human papillomavirus types associated with cervical cancer, N. Engl. J. Med. 348 (6) (2003) 518–527.
[5] P.S. Aghbash, N. Hemmat, H. Fathi, H.B. Baghi, Monoclonal antibodies in cervical malignancy-related HPV, Front. Oncol. 12 (2022) 904790.
[6] E.M. Burd, Human papillomavirus and cervical cancer, Clin. Microbiol. Rev. 16 (1) (2003) 1–17.
[7] A.M. Almeida, J.A. Queiroz, F. Sousa, Â. Sousa, Cervical cancer and HPV infection: ongoing therapeutic research to counteract the action of E6 and E7 oncoproteins, Drug Discov. Today 24 (10) (2019) 2044–2057.
[8] A.K. Grabowska, A.B. Riemer, Suppl 2: the invisible enemy–how human papillomaviruses avoid recognition and clearance by the host immune system, Open Virol. J. 6 (2012) 249.
[9] C.A. Moody, L.A. Laimins, Human papillomavirus oncoproteins: pathways to transformation, Nat. Rev. Cancer 10 (8) (2010) 550–560.
[10] P.S. Aghbash, et al., The effect of Wnt/β-catenin signaling on PD-1/PDL-1 axis in HPV-related cervical cancer, Oncol. Res. 30 (3) (2022) 99–116.
[11] M. Scheffner, J.M. Huibregtse, P.M. Howley, Identification of a human ubiquitin-conjugating enzyme that mediates the E6-AP-dependent ubiquitination of p53, Proc. Natl. Acad. Sci. USA 91 (19) (1994) 8797–8801.
[12] M. Scheffner, J.M. Huibregtse, R.D. Vierstra, P.M. Howley, The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53, Cell 75 (3) (1993) 495–505.
[13] M. Arbyn, S. de Sanjose, E. Weiderpass, HPV-based cervical cancer screening, including self-sampling, versus screening with cytology in Argentina, Lancet Global Health 7 (6) (2019) e688–e689.
[14] M. Abbas, et al., Role of miRNAs in cervical cancer: a comprehensive novel approach from pathogenesis to therapy, Journal of Gynecology Obstetrics and Human Reproduction 50 (9) (2021) 102159.
[15] Y. Jiang, S. Chen, Functional new transcription factors (TFs) associated with cervical cancer, Journal of Healthcare Engineering 2022 (2022).
[16] J.H. Bushweller, Targeting transcription factors in cancer—from undruggable to reality, Nat. Rev. Cancer 19 (11) (2019) 611–624.
[17] M. Lambert, S. Jambon, S. Depauw, M.-H. David-Cordonnier, Targeting transcription factors for cancer treatment, Molecules 23 (6) (2018) 1479.
[18] A.J. Sigurdson, et al., Kin-cohort estimates for familial breast cancer risk in relation to variants in DNA base excision repair, BRCA1 interacting and growth factor genes, BMC Cancer 4 (1) (2004) 1–12.
[19] W. Verlaat, et al., Genome-wide DNA methylation profiling reveals methylation markers associated with 3q gain for detection of cervical precancer and cancer, Clin. Cancer Res. 23 (14) (2017) 3813–3822.

[20] T.-h. Cheung, et al., Dysregulated microRNAs in the pathogenesis and progression of cervical neoplasm, Cell Cycle 11 (15) (2012) 2876–2884.
[21] M. Lando, et al., Gene dosage, expression, and ontology analysis identifies driver genes in the carcinogenesis and chemoradioresistance of cervical cancer, PLoS Genet. 5 (11) (2009) e1000719.
[22] Z. Waks, et al., Driver gene classification reveals a substantial overrepresentation of tumor suppressors among very large chromatin-regulating proteins, Sci. Rep. 6 (1) (2016) 1–12.
[23] Y. Xu, H. Luo, Q. Hu, H. Zhu, Identification of potential driver genes based on multi-genomic data in cervical cancer, Front. Genet. 12 (2021) 43.
[24] B. Hrdlickova, R.C. de Almeida, Z. Borek, S. Withoff, Genetic variation in the non-coding genome: involvement of micro-RNAs and long non-coding RNAs in disease, Biochim. Biophys. Acta (BBA) - Mol. Basis Dis. 1842 (10) (2014) 1910–1922.
[25] Y. Wang, J. Luo, H. Zhang, J. Lu, microRNAs in the same clusters evolve to coordinately regulate functionally related genes, Mol. Biol. Evol. 33 (9) (2016) 2232–2247.
[26] S. Lin, R.I. Gregory, MicroRNA biogenesis pathways in cancer, Nat. Rev. Cancer 15 (6) (2015) 321–333.
[27] X. Yang, et al., High-throughput transcriptome profiling in drug and biomarker discovery, Front. Genet. 11 (2020) 19.
[28] M. Giulietti, G. Occhipinti, G. Principato, F. Piva, Weighted gene co-expression network analysis reveals key genes involved in pancreatic ductal adenocarcinoma development, Cell. Oncol. 39 (4) (2016) 379–388.
[29] N. Hemmat, et al., Neutrophils, crucial, or harmful immune cells involved in coronavirus infection: a bioinformatics study, Front. Genet. 11 (2020) 641.
[30] J.A. den Boon, et al., Molecular transitions from papillomavirus infection to cervical precancer and cancer: role of stromal estrogen receptor signaling, Proc Natl Acad Sci U S A 112 (25) (2015) E3255–E3264.
[31] S.H. Liu, et al., DriverDBv3: a multi-omics database for cancer driver gene research, Nucleic Acids Res. 48 (D1) (2020) D863–d870.
[32] T. Barrett, et al., NCBI GEO: mining tens of millions of expression profiles—database and tools update, Nucleic Acids Res. 35 (suppl_1) (2007) D760–D765.
[33] S.-H. Liu, et al., DriverDBv3: a multi-omics database for cancer driver gene research, Nucleic Acids Res. 48 (D1) (2019) D863–D870.
[34] R. Ihaka, R. Gentleman, R: a language for data analysis and graphics, J. Comput. Graph Stat. 5 (3) (1996) 299–314.
[35] B. Zhang, S. Horvath, A general framework for weighted gene co-expression network analysis, Stat. Appl. Genet. Mol. Biol. 4 (1) (2005).
[36] P. Langfelder, B. Zhang, S. Horvath, Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R, Bioinformatics 24 (5) (2008) 719–720.
[37] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinf. 9 (2008).
[38] P. Langfelder, R. Luo, M.C. Oldham, S. Horvath, Is my network module preserved and reproducible? PLoS Comput. Biol. 7 (1) (2011) e1001057.
[39] C.-H. Chin, et al., cytoHubba: identifying hub objects and sub-networks from complex interactome, BMC Syst. Biol. 8 (4) (2014) 1–7.
[40] C.H. Chin, et al., cytoHubba: identifying hub objects and sub-networks from complex interactome, BMC Syst. Biol. 8 (Suppl 4) (2014) S11.
[41] P. Shannon, et al., Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (11) (2003) 2498–2504.
[42] H. Han, et al., TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions, Nucleic Acids Res. 46 (D1) (2018) D380–d386.
[43] J. Li, et al., Identifying scientific project-generated data citation from full-text articles: an investigation of TCGA data citation, Journal of Data and Information Science 1 (2) (2016) 32–44.
[44] M.J. Goldman, et al., Visualizing and interpreting cancer genomics data via the Xena platform, Nat. Biotechnol. 38 (6) (2020) 675–678.
[45] M.V. Kuleshov, et al., Enrichr: a comprehensive gene set enrichment analysis web server 2016 update, Nucleic Acids Res. 44 (W1) (2016) W90–W97.
[46] S.L. Freshour, et al., Integration of the drug–gene interaction database (DGIdb 4.0) with open crowdsource efforts, Nucleic Acids Res. 49 (D1) (2020) D1144–D1151.
[47] C. Sticht, C. De La Torre, A. Parveen, N. Gretz, miRWalk: an online resource for prediction of microRNA binding sites, PLoS One 13 (10) (2018) e0206239.
[48] A. Lachmann, et al., Geneshot: search engine for ranking genes from arbitrary text queries, Nucleic Acids Res. 47 (W1) (2019) W571–W577.
[49] X. Liu, X. Liu, X. Han, FANCI may serve as a prognostic biomarker for cervical cancer: a systematic review and meta-analysis, Medicine 100 (51) (2021).
[50] H. Zeng, et al., Changing cancer survival in China during 2003–15: a pooled analysis of 17 population-based cancer registries, Lancet Global Health 6 (5) (2018) e555–e567.
[51] P. Shiri Aghbash, N. Hemmat, B. Baradaran, H. Bannazadeh Baghi, siRNA-E6 sensitizes HPV-16-related cervical cancer through Oxaliplatin: an in vitro study on anti-cancer combination therapy, Eur. J. Med. Res. 28 (1) (2023) 42.
[52] K.M. Abbas, K. Van Zandvoort, M. Brisson, M. Jit, Effects of updated demography, disability weights, and cervical cancer burden on estimates of human papillomavirus vaccination impact at the global, regional, and national levels: a PRIME modelling study, Lancet Global Health 8 (4) (2020) e536–e544.
[53] J. Yao, et al., Long noncoding RNA TOB1-AS1, an epigenetically silenced gene, functioned as a novel tumor suppressor by sponging miR-27b in cervical cancer, Am. J. Cancer Res. 8 (8) (2018) 1483.
[54] W. Zhang, et al., The identification and validation of EphA7 hypermethylation, a novel biomarker, in cervical cancer, BMC Cancer 22 (1) (2022) 1–12.
[55] M. Hinkula, et al., A population-based study on the risk of cervical cancer and cervical intraepithelial neoplasia among grand multiparous women in Finland, Br. J. Cancer 90 (5) (2004) 1025–1029.
[56] L. Kjellberg, et al., Smoking, diet, pregnancy and oral contraceptive use as risk factors for cervical intra-epithelial neoplasia in relation to human papillomavirus infection, Br. J. Cancer 82 (7) (2000) 1332–1338.
[57] H. Xue, et al., Identification of hub genes as potential prognostic biomarkers in cervical cancer using comprehensive bioinformatics analysis and validation studies, Cancer Manag. Res. 13 (2021) 117.
[58] V.T. Cheriyan, et al., Signaling defects and functional impairment in T-cells from cervical cancer patients, Cancer Biother. Rad. 24 (6) (2009) 667–673.
[59] D.L. Levey, P.K. Srivastava, Alterations in T cells of cancer-bearers: whence specificity? Immunol. today 17 (8) (1996) 365–368.
[60] M.F. Shehata, Rel/Nuclear factor-kappa B apoptosis pathways in human cervical cancer cells, Cancer Cell Int. 5 (1) (2005) 1–13.
[61] A.K. Garg, et al., Expression of nuclear transcription factor kappa B in locally advanced human cervical cancer treated with definitive chemoradiation, Int. J. Radiat. Oncol. Biol. Phys. 78 (5) (2010) 1331–1336.
[62] B. Altoos, et al., Prognostic significance of nuclear factor kappa b expression in locally advanced cervical cancer patients treated definitively with concurrent chemoradiation, Am. J. Clin. Oncol. 43 (1) (2020) 47–51.
[63] S. Rezapour, et al., STC1 and NF-κB p65 (Rel A) is constitutively activated in colorectal cancer, Clin. Lab. 62 (3) (2016) 463–469.
[64] X. Pan, et al., Nuclear factor-κB p65/relA silencing induces apoptosis and increases gemcitabine effectiveness in a subset of pancreatic cancer cells, Clin. Cancer Res. 14 (24) (2008) 8143–8151.
[65] S. Vaira, et al., RelA/p65 promotes osteoclast differentiation by blocking a RANKL-induced apoptotic JNK pathway in mice, J. Clin. Invest. 118 (6) (2008) 2088–2097.
[66] Y.-R. Deng, et al., Sp1 contributes to radioresistance of cervical cancer through targeting G2/M cell cycle checkpoint CDK1, Cancer Manag. Res. 11 (2019) 5835.
[67] F. Hoppe-Seyler, K. Butz, Activation of human papillomavirus type 18 E6–E7 oncogene expression by transcription factor Sp1, Nucleic Acids Res. 20 (24) (1992) 6701–6706.
[68] F. Wang, et al., miR-375 is down-regulated in squamous cervical cancer and inhibits cell migration and invasion via targeting transcription factor SP1, Am. J. Pathol. 179 (5) (2011) 2580–2588.
[69] E.-S. Choi, et al., Modulation of specificity protein 1 by mithramycin A as a novel therapeutic strategy for cervical cancer, Sci. Rep. 4 (1) (2014) 1–8.
[70] L. Zhang, S.-K. Liu, L. Song, H.-R. Yao, SP1-induced up-regulation of lncRNA LUCAT1 promotes proliferation, migration and invasion of cervical cancer by sponging miR-181a, Artif. Cell Nanomed. Biotechnol. 47 (1) (2019) 555–563.
[71] H. Jiang, et al., Long non-coding RNA TPT1-AS1 promotes cell growth and metastasis in cervical cancer via acting AS a sponge for miR-324-5p, J. Exp. Clin. Cancer Res. 37 (1) (2018) 1–12.
[72] G. Liu, Z. Ye, X. Zhao, Z. Ji, SP1-induced up-regulation of lncRNA SNHG14 as a ceRNA promotes migration and invasion of clear cell renal cell carcinoma by regulating N-WASP, Am. J. Cancer Res. 7 (12) (2017) 2515.

[73] Z.-Q. Wang, et al., Long noncoding RNA UCA1 induced by SP1 promotes cell proliferation via recruiting EZH2 and activating AKT pathway in gastric cancer, Cell Death Dis. 8 (6) (2017).
[74] S. Li, et al., Sp1-mediated transcriptional regulation of MALAT1 plays a critical role in tumor, J. Cancer Res. Clin. Oncol. 141 (11) (2015) 1909–1920.
[75] T. Wang, et al., LncRNA MALAT1 was regulated by HPV16 E7 independently of pRB in cervical cancer cells, J. Cancer 12 (21) (2021) 6344.
[76] C.-H. Chou, et al., Dihydromyricetin suppresses cell metastasis in human osteosarcoma through SP-1-and NF-κB-modulated urokinase plasminogen activator inhibition, Phytomedicine 90 (2021) 153642.
[77] Y. An, H. Duan, The role of m6A RNA methylation in cancer metabolism, Mol. Cancer 21 (1) (2022) 1–24.
[78] S. Ro, Multi-phenotypic role of serum response factor in the gastrointestinal system, Journal of neurogastroenterology and motility 22 (2) (2016) 193.
[79] L. Ma, Y. Yu, X. Qu, Suppressing serum response factor inhibits invasion in cervical cancer cell lines via regulating Egr-1 and epithelial-mesenchymal transition, Int. J. Mol. Med. 43 (1) (2019) 614–620.
[80] E.S. Silverman, T. Collins, Pathways of Egr-1-mediated gene transcription in vascular biology, Am. J. Pathol. 154 (3) (1999) 665–670.
[81] X. Zhao, et al., SRF expedites metastasis and modulates the epithelial to mesenchymal transition by regulating miR-199a-5p expression in human gastric cancer, Cell Death Differ. 21 (12) (2014) 1900–1913.
[82] Y.-J. Li, M. Dong, F.-M. Kong, J.-P. Zhou, Folate-decorated anticancer drug and magnetic nanoparticles encapsulated polymeric carrier for liver cancer therapeutics, Int. J. Pharm. 489 (1–2) (2015) 83–90.
[83] J.C. Evans, et al., Cyclodextrin mediated delivery of NF-κB and SRF siRNA reduces the invasion potential of prostate cancer cells in vitro, Gene Ther. 22 (10) (2015) 802–810.
[84] M. Burute, et al., Polarity reversal by centrosome repositioning primes cell scattering during epithelial-to-mesenchymal transition, Dev. Cell 40 (2) (2017) 168–184.
[85] R. Bhattacharya, T. Mitra, S. Ray Chaudhuri, S.S. Roy, Mesenchymal splice isoform of CD44 (CD44s) promotes EMT/invasion and imparts stem-like properties to ovarian cancer cells, J. Cell. Biochem. 119 (4) (2018) 3373–3383.
[86] A. Labernadie, et al., A mechanically active heterotypic E-cadherin/N-cadherin adhesion enables fibroblasts to drive cancer cell invasion, Nat. Cell Biol. 19 (3) (2017) 224–237.
[87] C. Zhu, et al., Somatic mutation of DNAH genes implicated higher chemotherapy response rate in gastric adenocarcinoma patients, J. Transl. Med. 17 (1) (2019) 1–11.
[88] A.J. Roberts, et al., ATP-driven remodeling of the linker domain in the dynein motor, Structure 20 (10) (2012) 1670–1680.
[89] J. Neesen, et al., Identification of dynein heavy chain genes expressed in human and mouse testis: chromosomal localization of an axonemal dynein gene, Gene 200 (1–2) (1997) 193–202.
[90] M. Fliegauf, T. Benzing, H. Omran, When cilia go bad: cilia defects and ciliopathies, Nat. Rev. Mol. Cell Biol. 8 (11) (2007) 880–893.
[91] W.L. Yau, et al., Functional studies of the chromosome 3p21. 3 candidate tumor suppressor gene BLU/ZMYND10 in nasopharyngeal carcinoma, Int. J. Cancer 119 (12) (2006) 2821–2826.
[92] Y. Wang, et al., ZMYND10, an epigenetically regulated tumor suppressor, exerts tumor-suppressive functions via miR145-5p/NEDD9 axis in breast cancer, Clin. Epigenet. 11 (1) (2019) 1–16.
[93] H.J. Yoo, et al., BLU enhances the effects of anti-angiogenic activity in combination with gemcitabine-based chemotherapeutic agents, Int. J. Biochem. Cell Biol. 45 (7) (2013) 1236–1245.
[94] S.T. Park, et al., Tumor suppressor BLU promotes paclitaxel antitumor activity by inducing apoptosis through the down-regulation of Bcl-2 expression in tumorigenesis, Biochem. Biophys. Res. Commun. 435 (1) (2013) 153–159.