ORIGINAL ARTICLE

Bioinformatic Analysis of miR-200b/429 and Hub Gene Network in Cervical Cancer

Vaibhav Shukla[1] · Sandeep Mallya[2] · Divya Adiga[1] · S. Sriharikrishnaa[1] · Sanjiban Chakrabarty[1,3] · Shama Prasada Kabekkodu[1,3]

Abstract

The miR-200b/429 located at 1p36 is a highly conserved miRNA cluster emerging as a critical regulator of cervical cancer. Using publicly available miRNA expression data from TCGA and GEO followed by independent validation, we aimed to identify the association between miR-200b/429 expression and cervical cancer. miR-200b/429 cluster was significantly overexpressed in cancer samples compared to normal samples. miR-200b/429 expression did not correlate with patient survival; however, its overexpression correlated with histological type. Protein–protein interaction analysis of 90 target genes of miR-200b/429 identified EZH2, FLT1, IGF2, IRS1, JUN, KDR, SOX2, MYB, ZEB1, and TIMP2 as the top ten hub genes. PI3K–AKT and MAPK signaling pathways emerged as major target pathways of miR-200b/429 and their hub genes. Kaplan–Meier survival analysis showed the expression of seven miR-200b/429 target genes (EZH2, FLT1, IGF2, IRS1, JUN, SOX2, and TIMP2) to influence the overall survival of patients. The miR-200a-3p and miR-200b-5p could help predict cervical cancer with metastatic potential. The cancer hallmark enrichment analysis showed hub genes to promote growth, sustained proliferation, resistance to apoptosis, induction of angiogenesis, activation of invasion, and metastasis, enabling replicative immortality, evading immune destruction, and tumor-promoting inflammation. The drug–gene interaction analysis identified 182 potential drugs to interact with 27 target genes of miR-200b/429 with paclitaxel, doxorubicin, dabrafenib, bortezomib, docetaxel, ABT-199, eribulin, vorinostat, etoposide, and mitoxantrone emerging as the top ten best candidate drugs. Taken together, miR-200b/429 and associated hub genes can be helpful for prognostic application and clinical management of cervical cancer.

Keywords Cervical cancer · miR-200b/429 cluster · Prognosis · Bioinformatic analysis · TCGA

✉ Shama Prasada Kabekkodu
    shama.prasada@manipal.edu

Extended author information available on the last page of the article

Introduction

Cervical cancer is a malignancy of the cervix, with 570 000 cases and 311 000 deaths globally in 2018 (Arbyn et al. 2020). High-risk human papillomavirus (HPV) infection and genetic and epigenetic changes are strongly linked to cervical cancer development (Mac and Moody 2020). It is the second most common cancer among women aged between 15 and 44, with an estimated age-standardized incidence of 13.1/100000 worldwide. Over one-third of the global cervical cancer load is contributed together by India and China (Arbyn et al. 2020). The relative 5-year survival rate for stage I, stage II, stage III, and stage IV cervical cancer were 83.5%, 80.6%, 66.0%, and 37.1%, respectively (Balasubramaniam et al. 2020). Early detection and understanding of the molecular complexity are proposed to reduce the high mortality rate of advanced stage and distant cervical cancer (Mishra et al. 2011). Despite the availability of sensitive and specific screening tools for early detection, many cervical cancer cases are detected at an advanced stage, contributing to high mortality and morbidity. Hence, there is a need to (i) identify the marker for early diagnosis, prognosis, and therapeutic applications and (ii) a molecular mechanism for better management of cervical cancer. Towards this, analyzing the epigenome may provide sensitive and specific markers for the clinical management of cervical cancer.

Epi-genome-wide studies using microarray and next-generation sequencing have identified the global epi(genomic) changes responsible for carcinogenesis and provided an opportunity to translate these molecular changes as markers for improved clinical management of cervical cancer (Varghese et al. 2018). Nevertheless, the findings from the genome-wide study required cross-validation as the markers identified may be inconsistent because of (i) tumor heterogeneity, (ii) techniques used, (iii) sample types and source, and (iv) algorithms used for data analysis (Hamid et al. 2009). Hence, reanalysis of big data may provide new insights into the regulatory factors, molecular mechanisms, and signaling pathways altered during cervical cancer. Resistance to treatment and lack of effective therapy contributes to a high mortality rate in the advanced stage of cervical cancer (Mehta et al. 2017). Previous in silico analyses have suggested that the reanalysis of big data can identify reliable diagnostic and prognostic markers for the clinical management of cervical cancer (Shukla et al. 2021).

The miR-200b/429 located at 1p36 is a conserved miRNA cluster encoding for three miRNAs, namely miR-200b, miR-200a, and miR-429. Abnormal expression of members of miR-200b/429 have been reported in cervical cancer tissue and cell lines (González-Quintana et al. 2016; Zeng et al. 2016; Fan et al. 2017). Several malignancies have shown dysregulated expression of the miR-200b/429 family (Nam et al. 2008; Mateescu et al. 2011). miR-200b/429 cluster has been demonstrated to target the distinct genes at different phases of cancer growth, even in the same tumor. When miR-200b/429 cluster targets zinc finger E-box-binding homeobox1/2 (ZEB1/2), they prevent local invasion in breast cancer, but when they target SEC23a, they encourage metastatic lung colonization (Korpal et al. 2011). The small RNA sequencing of normal and cervical cancer patients showed

miR-200b/429 as one of the most differentially expressed miRNA clusters as per our study (Shukla et al. 2019). It is a potential biomarker for locally advanced radioresistant cervical cancer (Nilsen et al. 2022). Besides, its expression was dysregulated in the The Cancer Genome Atlas-Cervical Squamous Cell Carcinoma (TCGA-CESC) datasets. Functional studies have suggested that the members of this cluster regulated proliferation, migration, and invasion and showed both tumor-suppressive and oncogenic functions in cervical cancer (Hu et al. 2010; González-Quintana et al. 2016; Zeng et al. 2016; Fan et al. 2017). Hence, miR-200b/429 cluster expressions have been proposed as a sensitive and specific marker for diagnostic and prognostic application in cervical cancer (Zeng et al. 2016; KZ et al. 2020). Although miR-200b/429 is a clustered miRNA, none of the previous cervical cancer studies have investigated the diagnostic, prognostic, and functional significance of these miRNA as a single cluster. The present study aimed to test the diagnostic and prognostic significance of miR-200b/429 in cervical cancer using miRNA expression data from TCGA-CESC and gene expression omnibus (GEO). Further, the in silico findings of miR-200b/429 expression were validated in an independent cohort of cervical cancer samples. We have constructed the interaction networks of miR-200b/429 with mRNA, lncRNA, and circRNA. We have predicted the metastatic and prognostic significance of miR-200b/429. Moreover, we have performed functional enrichment, protein–protein interaction network (PPIN) analysis, and hub genes (HG) identification. A network was constructed based on miR-200b/429 and its validated protein-coding targets, and drug–gene interaction was predicted. Thus, our integrated *in silico* analysis has identified the interactome of miR-200b/429 with potential diagnostic and prognostic functions. The novel pathways and interactome may provide insight into a deeper understanding of cervical cancer and an opportunity to translate these findings for improved cervical cancer management.

Materials and Methods

Data Source, Collection, and miR-200b/429 Cluster Expression Analysis

The expression of miR-200b, miR-200a, and miR-429 were extracted from TCGA-CESC (normal = 3 samples and tumor = 310 samples) using UALCAN (Chandrashekar et al. 2017) and GEO database (http://www.ncbi.nlm.nih.gov/gds/). We used GEO datasets to strengthen our data as TCGA has fewer normal samples. The GEO datasets analyzed included GSE86100 (6 normal and 6 tumors) (Gao et al. 2016) and GSE105409 (1 normal pool of 16 samples and 1 tumor pool of 30 samples) (Kawai et al. 2018). The differential expression of miR-200b/429 in TCGA-CESC and GEO datasets were analyzed using Transcriptome Alterations in Cancer Omnibus (Chou et al. 2019) and GEO2R tools (https://www.ncbi.nlm.nih.gov/geo/geo2r/), respectively. The associations between members of the miR-200b/429 cluster with clinical parameters were tested by MEXPRESS (Koch et al. 2015).
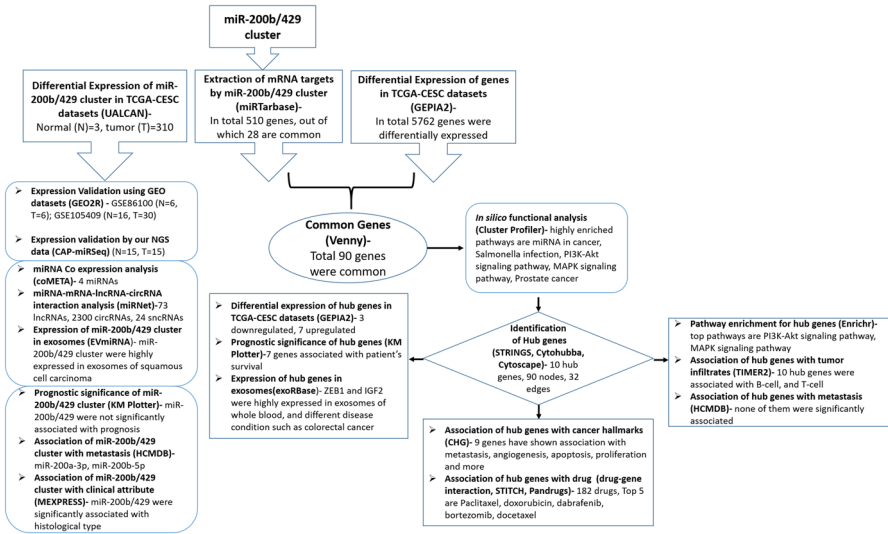
Fig. 1 The schematic diagram of in silico analysis of miR-200b/429 cluster, their targets, expression, and associated pathway in CC

The conservation of this miRNA cluster was predicted using the ECR browser (Ovcharenko et al. 2004). The flowchart for bioinformatic analysis is shown in Fig. 1.

miR-200b/429 Cluster Target Prediction and Expression Analysis

Co-Expression Meta-Analysis of miRNA target (COMETA) (Gennarino et al. 2012) tool was used to predict the co-expressing miRNAs of the miR-200b/429 cluster. miRTarBase (Huang et al. 2020a, b) was used to predict the miR-200b/429 targets. miRTarBase is a freely available online database consisting of well-annotated and experimentally confirmed miRNA–target interactions. The expression analysis was performed using GEPIA2 (Tang et al. 2019). The genes commonly differentially expressed in multiple datasets were subsequently considered as miR-200b/429 targets. The gene expression in exosomes was performed by exoRBase (Li et al. 2018).

Functional Enrichment Analysis and Protein–Protein Interaction Network Construction

The in silico functional enrichment (biological functions, gene ontology, and pathway) of miR-200b/429 targets were undertaken using Enrichr. Gene ontology analysis consisted of enrichment for biological process (BP), molecular function (MF), and cellular components (CC) (Xie et al. 2021). A protein–protein interaction network (PPIN) regulated by miR-200b/429 was constructed using STRING V11 (Szklarczyk et al. 2015). PPIN was constructed for differentially expressed targets

of miR-200b/429 using known interactions, predicted interactions, text mining, co-expression, and protein homology. The hub genes of the networks were predicted by the cytohubba tool (Chin et al. 2014), and Cytoscape to visualizes networks. Using Cluster Profiler, functional enrichment analysis of genes was performed (Yu et al. 2012). Hub genes-immune association is performed by TIMER2 (Li et al. 2021).

Systems Biology Analysis/Downstream Analysis

The miR-200b/429 centric lncRNA, circRNA, and sncRNA network was constructed using the miRNet tool (Chang et al. 2020). The parameters used for network construction includes Organism type: Human, ID type: miRbase ID, Tissue type: Cervix, Targets: lncRNA, circRNA, and sncRNA. DIANA-lncbase V3 was used to identify the experimentally validated miRNA–lncRNA interactions in cervical cancer (Karagkouni et al. 2020). The prognostic significance of miR-200b/429 and hub genes was evaluated using the Kaplan–Meier survival curve with the Log-rank method KM plotter tool (Nagy et al. 2021). Using the TACCO tool (Chou et al. 2019), a prognostic model was built using the expression of 10 hub genes for overall survival using the random forest algorithm. HCMDB (Human Cancer Metastasis Database) (Zheng et al. 2018) tool was used to identify the association between miR-200b/429 and hub genes with metastasis.

Prediction of Drug–Gene Interaction

The differentially expressed targets of miR-200b/429 were used for predicting the druggable genes using Drug-Gene Interaction Database 3.0 (Freshour et al. 2021). The drug–gene interaction network was generated using STITCH (Kuhn et al. 2008) and prioritized using PanDrugs (Piñeiro-Yáñez et al. 2018).

Validation by Small RNA Sequencing

We had previously performed small RNA sequencing in 15 normal and 15 cervical cancer samples using the P1™ chip using Ion Proton system™ (Shukla et al. 2019). The miRNA alignment and differential expression analysis were undertaken using the CAP-miRSeq pipeline (Sun et al. 2014). The expression of miR-200b/429 was cross-validated using our previously published data (Shukla et al. 2019).

Results

miR-200b/429 Cluster is Overexpressed in Cervical Cancer

The miR-200b/429 is located at 1p36 and is a highly conserved miRNA cluster (Supplementary Fig. 1a and 1b). The expression of individual members of the miR-200b/429 cluster was evaluated in TCGA-CESC, GSE86100, and GSE105409, datasets consisting of 24 normal and 346 cervical cancer samples. The miR-200b/429
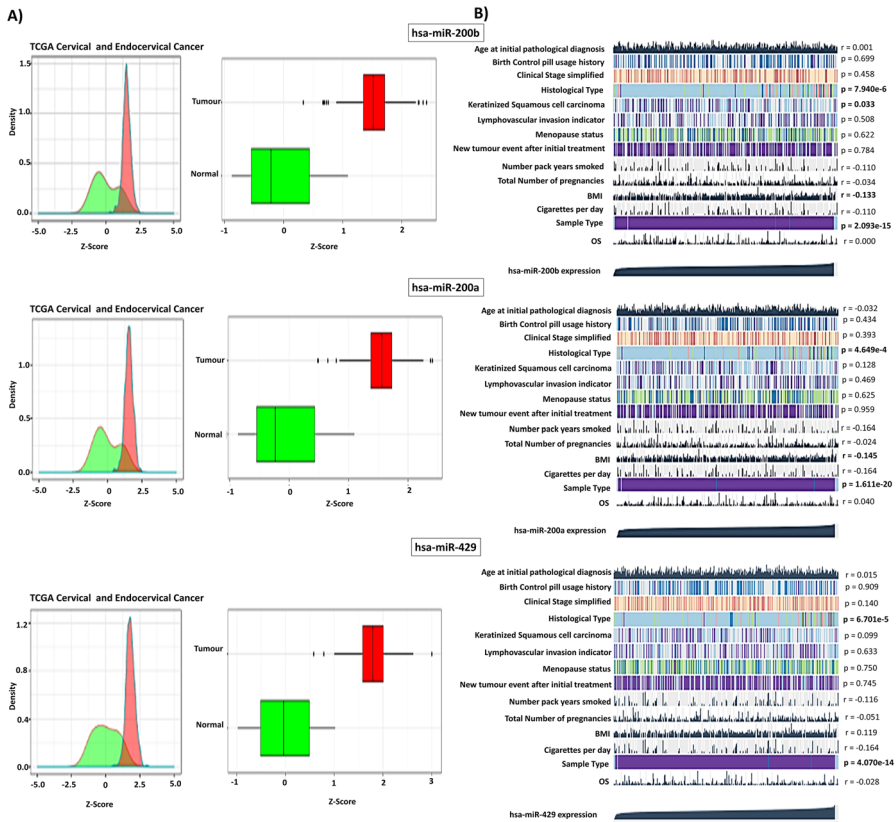
Fig. 2 **a** TCGA-CESC datasets have revealed the overexpression of miR-200b, miR-200a, and miR-429 in cervical cancer compared to healthy controls. The analysis was performed using miRNACANCER-MAP (Tong et al. 2018), **b** Correlation analysis between clinical attributes and miR-200b/429 expression. miR-200b, miR-200a, and miR-429 expression have shown to be significantly associated with histological type. The analysis was performed using MEXPRESS

cluster expression was significantly higher in tumor samples compared to normal samples in TCGA-CESC, GSE86100, and GSE105409 datasets (Figs. 2a, 3). The description for the datasets is shown in Supplementary Table 1 and 2. The MEX-PRESS analysis identified that miR-200b/429 expression was significantly correlated with the histological type ($P < 0.01$, Fig. 2b).

Validation of miR-200b/429 Expression

The miR-200b/429 was cross-validated using our previously published small RNA sequencing data. The expression of miR-200a (1.81-fold), miR-200b (1.53-fold), and miR-429 (1.60-fold) were significantly higher in our small cohort of cervical
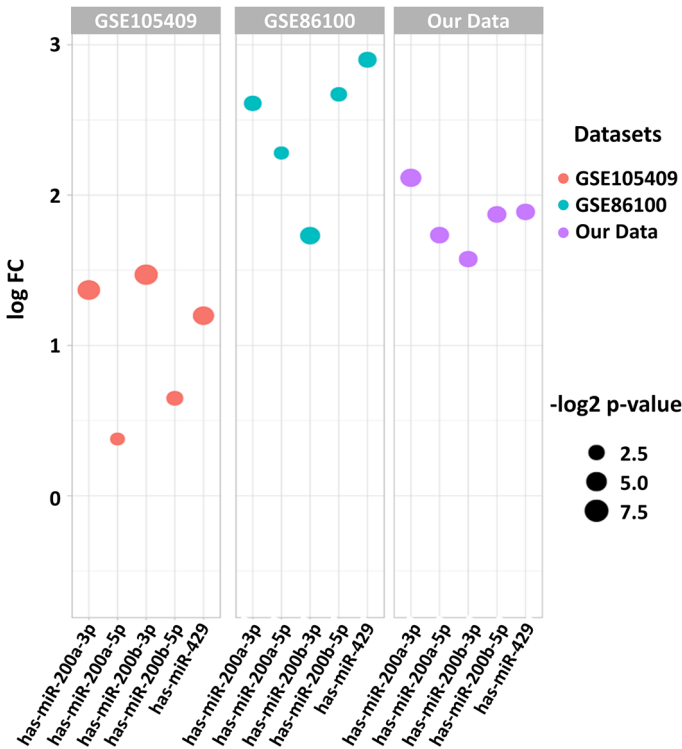
Fig. 3 Expression of miR-200b/429 cluster in GEO datasets and our NGS datasets

cancer samples when compared with the normal cervix (Fig. 3). These results further confirm our in silico analysis.

miR-200b/429 Cluster and Target Network Construction

The miRNA–miRNA co-expression network constructed using the COMETA tool identified miR-200a, miR-200b, miR-429, and miR-1278 as members of the co-expression as the immediate members of the interaction network (Fig. 4a). We next constructed a cervical cancer-specific miR-200b/429—lncRNA–circRNA–sncRNA network by a miRNA-centric network visual analytics platform (miRNet). The miR-200b/429 interacted with 73 lncRNA, 2300 circRNA, and 24 sncRNA consisting of 4333 edges (Fig. 4b). Further, we have identified 31 experimentally validated interactions between miR-200b/429 cluster and lncRNAs in cervical cancer using DIANA LncBase v3. Among 31 interactions, two were common with lncRNAs identified from miRNET analysis, namely, MALAT1 and ATP5F1AP2. The prediction using miRTarBase identified 510 protein-coding genes (miR-200a: 181, miR-200b:210, and miR-429: 119) as targets of miR-200b/429. Interestingly, 28 protein-coding genes were commonly targeted by all
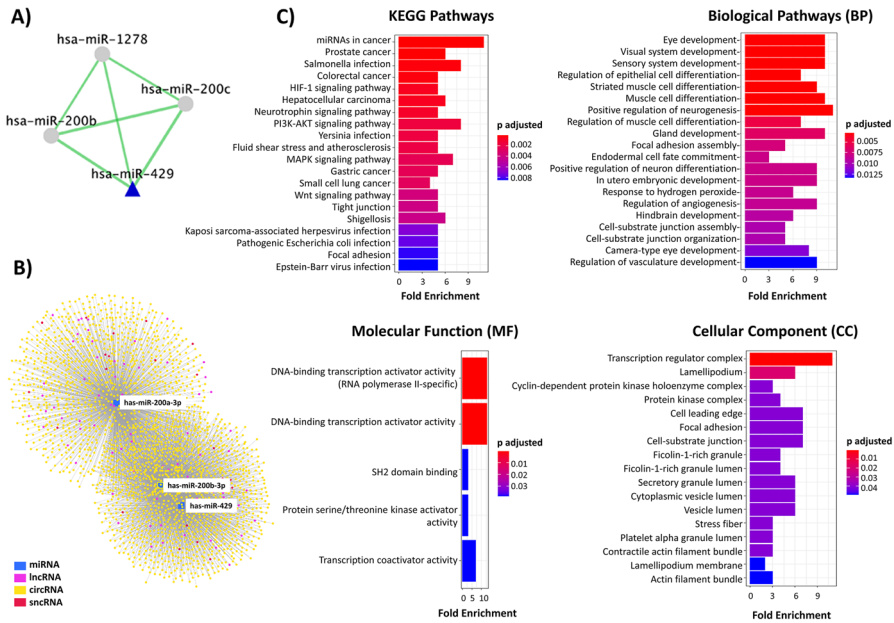
Fig. 4 miR-200b/429 cluster in silico analysis. **a** miRNA–miRNA interaction analysis has shown these clusters are interacting with each other along with miR-1278. **b** Construction of cervix specific miR-200b/429 cluster, mRNA, lncRNA, circRNA, and sncRNA network. **c** Functional enrichment analysis of 90 common genes targeted by miR-200b/429 cluster

three members of this cluster (Supplementary Fig. 1c, Supplementary Table 3 & 4). As per Log_2FC, Cutoff:1 and q-value Cutoff:0.01 have identified that 5762 genes are differentially expressed between normal and tumor samples in the TCGA-CESC dataset (Supplementary Table 5). The overlapping analysis between 5762 differentially expressed genes (1851 up to and 3911 down) with 510 targets of miR-200b/429 identified 90 differentially expressed genes as targets (Supplementary Fig. 1d & Supplementary Table 6).

Functional Enrichment Analysis of miR-200b/429

The functional enrichment analysis was undertaken for 90 differentially expressed targets of miR-200b/429. The pathway enrichment was performed by comparing the differentially expressed genes with KEGG annotations. The top ten pathways enriched included microRNAs in cancer, Salmonella infection, PI3K–AKT signaling pathway, MAPK signaling pathway, Prostate cancer, Hepatocellular carcinoma, Shigellosis, Colorectal cancer, HIF-1 signaling pathway, and Neurotrophin signaling pathway. The top ten BP enriched included focal adhesion assembly, gland development, regulation of muscle cell differentiation, positive regulation of neurogenesis, muscle cell differentiation, striated muscle cell differentiation, regulation of epithelial cell differentiation, sensory system development, visual system development,

and eye development. The top MF enriched are DNA-binding transcription activator activity, RNA polymerase II-specific DNA-binding transcription activator activity, SH2 domain-binding protein serine/threonine kinase activator activity, and transcription coactivator activity. The top CC enriched included secretory granule lumen, ficolin-1-rich granule lumen, ficolin-1-rich granule, cell–substrate junction, focal adhesion, cell leading edge, protein kinase complex, cyclin-dependent protein kinase holoenzyme complex, lamellipodium, and transcription regulator complex. Besides, the disease gene network construction and disease ontology predicted the differentially expressed genes to participate in cancer (Fig. 4c).

PPIN of miR-200b/429 Targets

The PPIN of 90 target genes of miR-200b/429 was constructed using STRING v11. The PPIN consisted of 90 nodes and 32 edges with a PPI enrichment p-value of 4.27e-05. The hub genes of the network predicted using cytohubba-identified Enhancer of zeste homolog 2 (EZH2), Fms-related receptor tyrosine kinase 1 (FLT1), Insulin-like growth factor 2 (IGF2), Insulin receptor substrate 1 (IRS1), Jun proto-oncogene, AP-1 transcription factor subunit (JUN), Kinase insert domain receptor (KDR), SRY-box transcription factor 2 (SOX2), MYB proto-oncogene transcription factor (MYB), Zinc finger E-box-binding homeobox 1 (ZEB1), and TIMP metallopeptidase inhibitor 2 (TIMP2) as the top ten hub genes of the target network consisting of 10 nodes, 25 edges with a PPI enrichment p-value of 3.42e-14 (Supplementary Fig. 2). We also performed the 10 hub genes systemic analysis of immune infiltrate across the TCGA-CESC cohort ($n = 306$). Our analysis has shown the association of hub genes with B-cell, T-cell CD4+, and T-cell CD8+ with a p-value < 0.05 (Supplementary Table 7). To determine the gene expression in the same tissue, we used TCGA-CESC datasets to analyze the gene expression. Next, we performed a Pearson correlation analysis of miRNA and mRNA expression from the TCGA datasets using TACCO (Chou et al. 2019) and identified that miR-200b/429 clusters primary targets, ZEB1, IRS1, FLT1 KDR, and TIMP2, as shown in Table 1.

Table 1 miRNA:mRNA correlation based on TCGA-CESC datasets

miRNA	Target gene	Pearson's r	Spearman's p
hsa-miR-200b-3p	IRS1	− 0.068	− 0.091
hsa-miR-200b-3p	FLT1	− 0.196	− 0.157
hsa-miR-200b-3p	KDR	− 0.167	− 0.126
hsa-miR-200b-3p	ZEB1	− 0.273	− 0.233
hsa-miR-200a-3p	ZEB1	− 0.253	− 0.239
hsa-miR-429	TIMP2	− 0.24	− 0.212
hsa-miR-429	ZEB1	− 0.277	− 0.261

Expression and Functional Enrichment Analysis of Hub Target Genes of miR-200b/429

The expression of 10 hub genes in normal and cervical cancer samples was tested using the GEPIA2 tool, 306 tumors, and 3 normal samples. EZH2, SOX2, and MYB were significantly upregulated among the hub genes, while FLT1, IGF2, IRS1, JUN, KDR, ZEB1, and TIMP2 were downregulated, respectively (Fig. 5).

The cluster profiler performed the enrichment of hub genes, which included analysis of gene ontology, pathway enrichment, transcription factor mapping, miRNA search, and others. PI3K–AKT signaling pathway and MAPK signaling pathway emerged as major target pathways. (Supplementary Fig. 3).

Prognostic Significance of miR-200b/429 and Hub Genes

The KM survival analysis showed that the miR-200b/429 cluster expression did not significantly affect patient survival. We next analyzed the prognostic significance of 10 hub genes in patient survival. Interestingly, the differential expression of 7 genes (EZH2, FLT1, IGF2, IRS1, JUN, SOX2, and TIMP2) was shown to influence patient survival as analyzed by the KM plotter (Fig. 6). Next, we analyzed the role of differential expression of miR-200b/429 and hub gene on cervical cancer metastasis. The expression of miR-200a-3p expression is significantly higher in primary tumors with metastasis than without, with head & neck, and lung being the major metastatic
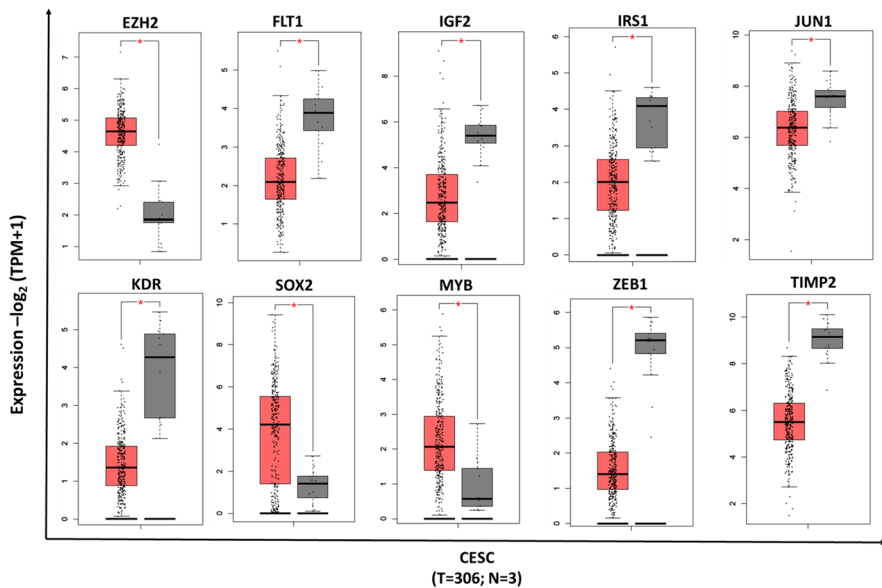

Fig. 5 The expression of hub genes in cervical cancer samples compared to normal. $p < 0.05$ were considered significant
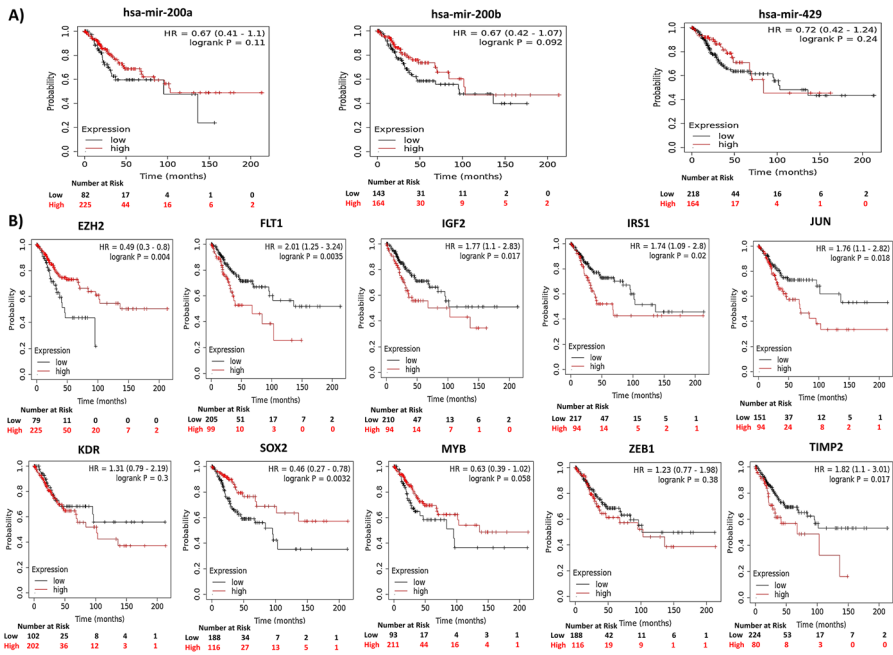
Fig. 6 The prognostic significance of miR-200b/429 and hub genes

sites. Interestingly, hsa-miR-200b-5p was significantly under-expressed in metastatic cervical cancer samples. However, hsa-miR-200a-5p, hsa-miR-200b-3p, and hsa-miR-429 expressions were not differentially expressed between primary tumors with metastasis compared to those without metastasis. None of the hub genes were significantly associated with metastasis (Supplementary Fig. 4).

miR-200b/429-Hub and Cancer Hallmarks

We next investigated the potential of ten hub genes in the induction of cancer hallmarks using CHG online tool. Among the ten hub genes, only data were available for 9 genes. The CHG analysis showed that abnormal expression of hub genes (EZH2, FLT1, IGF2, IRS1, JUN, KDR, SOX2, MYB, and ZEB1) might facilitate participating in pathways leading to the promotion of growth, sustained proliferation, resistance to apoptosis, induction of angiogenesis, activation of invasion, and metastasis, enabling replicative immortality, evading immune destruction, and tumor-promoting inflammation (Supplementary Table 8).

Drug–Gene Interaction Analysis

The potentially druggable genes of miR-200b/429 targets were predicted using the DGIdb database. Screening using DGIdb predicted 27 genes and 182 potential drugs. Furthermore, the interaction network of 27 genes using STITCH identified 34 nodes and 41 edges with a PPI enrichment p-value of 2.02e-05. Paclitaxel, doxorubicin, dabrafenib, bortezomib, docetaxel, ABT-199, eribulin, vorinostat, etoposide, and mitoxantrone as the top 10 best candidate drugs emerged out of Pandrug analysis (Supplementary tables 9 and 10).

Discussion

Despite the progress and improvement in diagnosis and prognosis, cervical cancer is one of the key public health problems in many developing and under-developed countries (Hull1 et al. 2020). Infection with HPV and genetic and epigenetic changes in both host and HPV play a critical role in the development and progression of cervical cancer (Senapati et al. 2016). Pap screening and HPV vaccines have helped to reduce the cervical cancer burden globally. However, several countries still have high incidence and mortality due to cervical cancer (Arbyn et al. 2020). The lack of effective screening programs, socioeconomic status, early cancer detection, the tendency to metastasis, and therapy resistance are some critical factors contributing to the high incidence and mortality of cervical cancer (Mishra et al. 2011). Thus, understanding this disease at the molecular level is critical for improved diagnosis, prognosis, and finding alternative drugs and targets for effectively managing cervical cancer. Recent progress in systems biology, genome-wide studies, and big data in public domains provide a platform to identify the key molecular changes that occur during tumorigenesis and provide an opportunity to understand the genes and pathways responsible for carcinogenesis and theragnostic applications. miRNAs are small non-coding RNAs critical for regulating the expression of protein-coding genes. We and others have shown that abnormal miRNA expression occurs during cervical carcinogenesis, and profiling can be used as a potential diagnostic and prognostic indicator (Pereira et al. 2010). By analyzing the publicly available miRNA datasets in TCGA and GEO, we aimed to identify the prognostic potential of the miR-200b/429 cluster in cervical cancer. Additionally, we have predicted the target genes, hub genes, pathways, interaction, and biological function using in silico approach. Our study demonstrated that miR-200b/429 overexpression might promote cervical cancer via activation of PI3K–AKT and MAPK signaling pathways. Besides, the targets of miR-200b/429 have prognostic value in cervical cancer. Furthermore, either the deregulated expression of miR-200b/429 or its target genes can predict metastasis in cervical cancer. Our study has identified several novel genes as targets of FDA-approved drugs.

Through integrated bioinformatic analysis using 4 miRNA expression datasets, we first showed that miR-200b/429 is overexpressed in cervical cancer compared to normal control samples and may act as an oncogene and promoter of cervical

cancer. Furthermore, our study showed that measuring miR-200b/429 could be helpful for the diagnosis and histological classification of cervical cancer. The three miRNAs of the cluster were differentially expressed in primary tumors with metastasis compared to those without metastasis. Thus, measuring the miR-200b/429 expression could be useful as a predictor of metastasis in cervical cancer. However, the miR-200b/429 cluster expression did not correlate with patient survival.

The hsa-miR-200b, the first member of the cluster (miR-200b/429), has been shown to promote cervical cancer cell proliferation and metastasis by inhibiting Forkhead box G1 (FOXG1) (Zeng et al. 2016). miR-200b was also shown to be hypomethylated and target Homeodomain-interacting protein kinase 3 (HIPK3) in cervical cancer (Varghese et al. 2018). Downregulation of miR-200b increased cell apoptosis and inhibited tumor growsth in vivo (Zeng et al. 2016). Hsa-miR-200a was upregulated in HPV-positive cervical tissues (Mandal et al. 2019; Vojtechova et al. 2016).

In contrast to its oncogenic function, other studies have shown its tumor-suppressive roles. For example, a study by Cheng et al. 2016 has shown miR-200b to suppress cell invasion and metastasis via inhibiting epithelial to mesenchymal transition (EMT) in cervical cancer (Cheng et al. 2016). Moreover, a study by He et al. 2020 has shown that overexpression of hsa-miR-200b increases apoptosis (He et al. 2020).

Hsa-miR-200a has been shown to downregulate metastatic genes such as ZEB1, and ZEB2, which is suggested to suppress the migration of cervical cancer cells (Hu et al. 2010). Similarly, hsa-miR-429 targets the inhibitor of nuclear factor Kappa B kinase subunit beta (IKKβ) and regulates NF-kB, suggesting its tumor-suppressive function (Fan et al. 2017). Hence, this cluster was shown to play an essential role in the progression of cervical cancer and may play an important role in the future as a biomarker. miR-200b/429 cluster target gene prediction was carried out for a deeper understanding of its function and led to the identification of 312 target genes. Further screening of 312 target genes in the TCGA-CESC dataset identified 90 target genes as differentially expressed and was subsequently used for downstream analysis. A PPIN of 90 genes identified FLT1, IGF2, IRS1, JUN, KDR, ZEB1, TIMP2 (downregulated), and EZH2, SOX2, and MYB (upregulated) in cervical cancer samples when compared with normal samples. Among the 10 hub genes, the expression of EZH2, FLT1, IGF2, IRS1, JUN, SOX2, and TIMP2 individually correlated with cervical cancer patient survival. Very interestingly, an overall survival model based on 10 hub genes using a random forest approach can be used to identify the high-risk and low-risk categories with a sensitivity and specificity of 0.92 and 0.93, respectively (Supplementary Fig. 5). Besides, the expression of 7 genes (EZH2, FLT1, IRS1, KDR2, MYB2, JUN, and TIMP2) was significantly different in metastatic tissue when compared with the primary tumor. Among the hub genes, except for TIMP2, the other 9 genes were associated with the induction of various cancer hallmarks. These data indicate that miR-200b/429 and the associated hub gene network may facilitate the acquisition of multiple cancer hallmarks to promote cervical cancer. We anticipate that measuring miR-200b/429 and hub gene expression may be useful for diagnosis, prognosis, histological evaluation, and prediction of metastasis in cervical cancer.

We next performed a functional enrichment analysis of miR-200b/429 for a deeper understanding of its role in cervical cancer. We showed that miR-200b/429 could regulate key cell signaling pathways, specifically the PI3K−AKT and MAPK signaling pathways. Activation of PI3K−AKT and MAPK signaling pathways are often activated in cervical cancer, and its role is well established in cervical cancer (Zhang et al. 2015). The enriched BP, MF, and CC terms are often associated with cancer.

Therapy resistance is the major factor associated with clinical outcomes and patient survival. Many previous studies have shown that both intrinsic and acquired resistance to conventional chemotherapeutic agents is one of the critical problems in cancer treatment and patient survival (Vasan et al. 2019). Hence, it is essential to understand the genes and pathways leading to therapy resistance for better management of cancer patients. Towards this, we have performed the drug–gene interaction analysis and identified 27 druggable genes and 182 potential drugs. Furthermore, prioritization identified 10 FDA-approved drugs as the best candidate for targeting miR-200b/429 and associated networks. However, more detailed clinical studies are required before further conclusions are drawn. Nevertheless, our study has identified genes, pathways, and new targets amenable for future studies. Thus, the new drugs predicted in our study may find new clinical use in cervical cancer. miRNAs have now also been shown to interact with other non-coding RNAs such as lncRNAs and circRNAs. For this reason, it is essential to understand the complex connections of the miR-200b/429 cluster in cervical cancer. With this, we have predicted miRNA interactions with lncRNAs and circRNAs. LncRNAs, which modulate gene expression, are now emerged as a new class of functional regulatory elements and have been associated with diverse human diseases. lncRNAs have been shown to contain miRNA-binding sites and exert their functions through titrating other miRNAs (Paraskevopoulou and Hatzigeorgiou 2016). These lncRNAs have been shown to act as sponges for miRNAs to reduce their activity. For example, lncRNA TPT1-AS1 sponges miR-324-5p and promotes cell growth and metastasis in cervical cancer (Jiang et al. 2018). Therefore, we have predicted miR-200b/429 cluster-binding sites on lncRNA and found XIST and MALAT1 to be common targets. Studies have shown that lncRNA XIST promotes cervical cancer progression by upregulating FUS via competitively binding to miR-200a (Zhu et al. 2018). lncRNA MALAT1 promotes cervical cancer through sponging miR-429 (Shen et al. 2019).

Similarly, circRNA can regulate gene expression by sponging miRNA. For example, circRNA cSMARCA5 acts as a sponge for miR-432 and regulates cervical cancer progression (Huang et al. 2020b). Towards this, we have predicted miRNA-binding sites on circRNA and identified six common circRNAs (MORC3_hsa-circRNA5148, CDK17_hsa-circRNA10073, TP53INP1_hsa-circRNA15552, USP9X_hsa-circRNA7984, XPO4_hsa-circRNA10219, and SEC23A_hsa-circRNA10433). The interaction of these circRNAs with miR-200b/429 needs further experimentation and validation. There are also reports showing the destruction of miRNA by lncRNA, which ultimately accumulated the circRNA. For exp, lncRNA Cyrano directs the destruction of miR-7 and accumulated circRNA Cdr1as in the cytoplasm of neurons (Kleaveland et al. 2018). In this study, we focused on miRNA–lncRNA and miRNA–circRNA interaction analysis based on

miRNA-binding sites on lncRNA and circRNA. Further, the relationship between miRNA–mRNA–lncRNA–circRNA needs to be established to explore the crosstalk between non-coding RNAs. The understanding of miRNA's complex networks will be the starting point to unraveling their role in tumorigenesis and disease progression. Given the significance of the mRNAs, lncRNAs, and circRNAs discovered in our work and other prior findings, more research into the involvement of these coding and non-coding RNAs in the pathophysiology of cervical cancer could be attempted. Our study made use of genome-wide data from public sources as well as freely available web tools; validation through experimental approach must be assessed to improve the miRNA-based therapy in the clinic.

Conclusion

Using the available miRNA and transcriptomics datasets, we have identified the prognostic value, biological function, and signaling pathways of miR-200b/429 and their target hub genes in cervical cancer. Altogether, we have identified the miR-200b/429 and associated gene regulatory networks that may offer a promising candidate for diagnosis, prognosis, and therapeutic application in cervical cancer. Our study did have certain limitations. Firstly, all the analysis was carried out using freely available online tools that analyzed the TCGA-CESC dataset. The TCGA-CESC dataset contained only three normal samples, is another limitation of the study. It was challenging to consider demographic aspects such as diverse age groups, ethnicities, geographical locations, tumor stage and classification for all the patients while assessing the DEGs due to the complexity of the datasets in our study. Secondly, all the analyses performed were in silico analysis that requires further validation. Therefore, further research is needed to determine their mechanistic basis. MiRNAs exert different functions in different tissue. Furthermore, miRNAs target genes show differential expression in different tissue types. Thus it is important to perform a correlation analysis between miRNA and its target gene expression from the same tissue. In our study, we have not used the miRNA–target gene expression from the same tissue, which is one of the study's limitations.

Regarding its function in cervical cancer, ZEB1, which has been identified as a common gene controlled by the miR-200b/429 cluster, requires more investigation. The bioinformatics analysis used in this work to examine the overall survival and expression levels of the ten hub genes may aid in developing a biomarker panel for this patient population. We still require larger, prospective studies to validate these hub genes to identify their sensitivity and specificity, particularly in cervical cancer. We have also discovered that hub genes are present in exosomes from healthy individuals, whole blood, and various disease conditions like colorectal cancer and pancreatic adenocarcinoma. However, to confirm the existence of a minimally invasive cervical cancer-specific biomarker, we must further validate the expression of these hub genes in cervical cancer serum or plasma samples.

Data Availability Not Applicable.

Code Availability Not Applicable.

Declarations

Conflict of interest The authors declare no conflict of interest.

Ethical Approval Ethical approval was not required for the present study. The study data were retrieved and synthesized from the already published studies (secondary data analysis).

Consent for Publication All the authors have read and approved the final draft of the manuscript.

References

Arbyn M, Weiderpass E, Bruni L et al (2020) Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. Lancet Glob Heal 8:e191–e203. https://doi.org/10.1016/S2214-109X(19)30482-6

Balasubramaniam G, Gaidhani RH, Khan A et al (2020) Survival rate of cervical cancer from a study conducted in India. Indian J Med Sci. https://doi.org/10.25259/IJMS_140_2020

Chandrashekar DS, Bashel B, Balasubramanya SAH et al (2017) UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 19:649–658. https://doi.org/10.1016/j.neo.2017.05.002

Chang L, Zhou G, Soufan O, Xia J (2020) miRNet 2.0: Network-based visual analytics for miRNA functional analysis and systems biology. Nucleic Acids Res 48:W244–W251. https://doi.org/10.1093/nar/gkaa467

Cheng YX, Zhang QF, Hong L et al (2016) MicroRNA-200b suppresses cell invasion and metastasis by inhibiting the epithelial-mesenchymal transition in cervical carcinoma. Mol Med Rep 13:3155–3160. https://doi.org/10.3892/mmr.2016.4911

Chin CH, Chen SH, Wu HH et al (2014) cytoHubba: Identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. https://doi.org/10.1186/1752-0509-8-S4-S11

Chou PH, Liao WC, Tsai KW, et al (2019) TACCO, a Database Connecting Transcriptome Alterations, Pathway Alterations and Clinical Outcomes in Cancers. Sci Rep. Doi: https://doi.org/10.1038/s41598-019-40629-z

Fan JY, Fan YJ, Wang XL et al (2017) miR-429 is involved in regulation of NF-κBactivity by targeting IKKβ and suppresses oncogenic activity in cervical cancer cells. FEBS Lett 591:118–128. https://doi.org/10.1002/1873-3468.12502

Freshour SL, Kiwala S, Cotto KC et al (2021) Integration of the drug-gene interaction database (DGIdb 4.0) with open crowdsource efforts. Nucleic Acids Res 49:D1144–D1151. https://doi.org/10.1093/nar/gkaa1084

Gao D, Zhang Y, Zhu M et al (2016) MiRNA expression profiles of HPV-infected patients with cervical cancer in the uyghur population in China. PLoS ONE. https://doi.org/10.1371/journal.pone.0164701

Gennarino VA, D'Angelo G, Dharmalingam G et al (2012) Identification of microRNA-regulated gene networks by expression analysis of target genes. Genome Res 22:1163–1172. https://doi.org/10.1101/gr.130435.111

González-Quintana V, Palma-Berré L, Campos-Parra AD et al (2016) MicroRNAs are involved in cervical cancer development, progression, clinical outcome and improvement treatment response (Review). Oncol Rep 35:3–12. https://doi.org/10.3892/or.2015.4369

Hamid JS, Hu P, Roslin NM et al (2009) Data integration in genetics and genomics: methods and challenges. Hum Genomics Proteomics. https://doi.org/10.4061/2009/869093

He L, Wang J, Chang D et al (2020) Effect of miRNA-200b on the proliferation and apoptosis of cervical cancer cells by targeting RhoA. Open Med 15:1019–1027. https://doi.org/10.1515/med-2020-0147

Hu X, Schwarz JK, Lewis JS et al (2010) A microRNA expression signature for cervical cancer prognosis. Cancer Res 70:1441–1448. https://doi.org/10.1158/0008-5472.CAN-09-3289

Huang HY, Lin YCD, Li J et al (2020a) MiRTarBase 2020: Updates to the experimentally validated microRNA-target interaction database. Nucleic Acids Res 48:D148–D154. https://doi.org/10.1093/nar/gkz896

Huang P, Qi B, Yao H et al (2020b) Circular RNA cSMARCA5 regulates the progression of cervical cancer by acting as a microRNA-432 sponge. Mol Med Rep 21:1217–1223. https://doi.org/10.3892/mmr.2020.10910

Hull R, Mbele M, Makhafola T et al (2020) Cervical cancer in low and middle.income countries (Review). Oncol Lett 20:2058–2074. https://doi.org/10.3892/ol.2020.11754

Jiang H, Huang G, Zhao N et al (2018) 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. https://doi.org/10.1186/s13046-018-0846-8

K Z, T M, Z H, et al (2020) Serum Level of microRNA-429 is as a Promising Diagnostic Biomarker for Cervical Cancer Patients Doi: https://doi.org/10.21203/RS.3.RS-47883/V1

Karagkouni D, Paraskevopoulou MD, Tastsoglou S et al (2020) DIANA-LncBase v3: indexing experimentally supported miRNA targets on non-coding transcripts. Nucleic Acids Res 48:D101–D110. https://doi.org/10.1093/NAR/GKZ1036

Kawai S, Fujii T, Kukimoto I et al (2018) Identification of miRNAs in cervical mucus as a novel diagnostic marker for cervical neoplasia. Sci Rep. https://doi.org/10.1038/s41598-018-25310-1

Kleaveland B, Shi CY, Stefano J, Bartel DP (2018) A network of noncoding regulatory RNAs acts in the mammalian brain. Cell 174:350-362.e17. https://doi.org/10.1016/j.cell.2018.05.022

Koch A, De Meyer T, Jeschke J, Van Criekinge W (2015) MEXPRESS: visualizing expression, DNA methylation and clinical TCGA data. BMC Genomics. https://doi.org/10.1186/s12864-015-1847-z

Korpal M, Ell BJ, Buffa FM et al (2011) Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization. Nat Med 17:1101–1109. https://doi.org/10.1038/NM.2401

Kuhn M, von Mering C, Campillos M et al (2008) STITCH: interaction networks of chemicals and proteins. Nucleic Acids Res. https://doi.org/10.1093/nar/gkm795

Li S, Li Y, Chen B et al (2018) ExoRBase: a database of circRNA, lncRNA and mRNA in human blood exosomes. Nucleic Acids Res 46:D106–D112. https://doi.org/10.1093/nar/gkx891

Li T, Fu J, Zeng Z et al (2021) TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res 48:W509–W514. https://doi.org/10.1093/NAR/GKAA407

Mac M, Moody CA (2020) Epigenetic regulation of the human papillomavirus life cycle. Pathogens 9:1–18

Mandal P, Saha SS, Sen S et al (2019) Cervical cancer subtypes harbouring integrated and/or episomal HPV16 portray distinct molecular phenotypes based on transcriptome profiling of mRNAs and miRNAs. Cell Death Discov. https://doi.org/10.1038/s41420-019-0154-x

Mateescu B, Batista L, Cardon M et al (2011) miR-141 and miR-200a act on ovarian tumorigenesis by controlling oxidative stress response. Nat Med 17:1627–1635. https://doi.org/10.1038/NM.2512

Mehta FF, Baik S, Chung SH (2017) Recurrence of cervical cancer and its resistance to progestin therapy in a mouse model. Oncotarget 8:2372–2380. https://doi.org/10.18632/oncotarget.13676

Mishra GA, Pimple SA, Shastri SS (2011) An overview of prevention and early detection of cervical cancers. Indian J Med Paediatr Oncol 32:125–132

Nagy Á, Munkácsy G, Győrffy B (2021) Pancancer survival analysis of cancer hallmark genes. Sci Rep. https://doi.org/10.1038/s41598-021-84787-5

Nam EJ, Yoon H, Kim SW et al (2008) MicroRNA expression profiles in serous ovarian carcinoma. Clin Cancer Res 14:2690–2695. https://doi.org/10.1158/1078-0432.CCR-07-1731

Nilsen A, Hillestad T, Skingen VE et al (2022) miR-200a/b/-429 downregulation is a candidate biomarker of tumor radioresistance and independent of hypoxia in locally advanced cervical cancer. Mol Oncol 16:1402–1419. https://doi.org/10.1002/1878-0261.13184

Ovcharenko I, Nobrega MA, Loots GG, Stubbs L (2004) ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. Nucleic Acids Res 32:W280. https://doi.org/10.1093/nar/gkh355

Paraskevopoulou MD, Hatzigeorgiou AG (2016) Analyzing MiRNA–LncRNA interactions. Methods in Molecular Biology. Humana Press Inc., Totowa, pp 271–286

Pereira PM, Marques JP, Soares AR et al (2010) Microrna expression variability in human cervical tissues. PLoS ONE. https://doi.org/10.1371/journal.pone.0011780

Piñeiro-Yáñez E, Reboiro-Jato M, Gómez-López G et al (2018) PanDrugs: a novel method to prioritize anticancer drug treatments according to individual genomic data. Genome Med. https://doi.org/10.1186/s13073-018-0546-1

Senapati R, Senapati NN, Dwibedi B (2016) Molecular mechanisms of HPV mediated neoplastic progression. Infect Agent Cancer. https://doi.org/10.1186/s13027-016-0107-4

Shen F, Zheng H, Zhou L et al (2019) Overexpression of MALAT1 contributes to cervical cancer progression by acting as a sponge of miR-429. J Cell Physiol 234:11219–11226. https://doi.org/10.1002/jcp.27772

Shukla V, Varghese VK, Kabekkodu SP et al (2019) Enumeration of deregulated miRNAs in liquid and tissue biopsies of cervical cancer. Gynecol Oncol 155:135–143. https://doi.org/10.1016/j.ygyno.2019.08.012

Sriharikrishnaa S, Shukla V, Khan GN et al (2021) Integrated bioinformatic analysis of miR-15a/16–1 cluster network in cervical cancer. Reprod Biol. https://doi.org/10.1016/j.repbio.2021.100482

Sun Z, Evans J, Bhagwate A et al (2014) CAP-miRSeq: a comprehensive analysis pipeline for microRNA sequencing data. BMC Genomics. https://doi.org/10.1186/1471-2164-15-423

Szklarczyk D, Franceschini A, Wyder S et al (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43:D447–D452. https://doi.org/10.1093/nar/gku1003

Tang Z, Kang B, Li C et al (2019) GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res 47:W556–W560. https://doi.org/10.1093/nar/gkz430

Tong Y, Ru B, Zhang J (2018) miRNACancerMAP: an integrative web server inferring miRNA regulation network for cancer. Bioinformatics 34:3211–3213. https://doi.org/10.1093/bioinformatics/bty320

Varghese VK, Shukla V, Kabekkodu SP et al (2018) DNA methylation regulated microRNAs in human cervical cancer. Mol Carcinog 57:370–382. https://doi.org/10.1002/mc.22761

Vasan N, Baselga J, Hyman DM (2019) A view on drug resistance in cancer. Nature 575:299–309

Vojtechova Z, Sabol I, Salakova M et al (2016) Comparison of the miRNA profiles in HPV-positive and HPV-negative tonsillar tumors and a model system of human keratinocyte clones. BMC Cancer. https://doi.org/10.1186/s12885-016-2430-y

Xie Z, Bailey A, Kuleshov MV et al (2021) Gene set knowledge discovery with enrichr. Curr Protoc. https://doi.org/10.1002/cpz1.90

Yu G, Wang LG, Han Y, He QY (2012) ClusterProfiler: An R package for comparing biological themes among gene clusters. Omi A J Integr Biol 16:284–287. https://doi.org/10.1089/omi.2011.0118

Zeng F, Xue M, Xiao T et al (2016) MiR-200b promotes the cell proliferation and metastasis of cervical cancer by inhibiting FOXG1. Biomed Pharmacother 79:294–301. https://doi.org/10.1016/j.biopha.2016.02.033

Zhang L, Wu J, Ling MT et al (2015) The role of the PI3K/Akt/mTOR signalling pathway in human cancers induced by infection with human papillomaviruses. Mol Cancer. https://doi.org/10.1186/s12943-015-0361-x

Zheng G, Ma Y, Zou Y et al (2018) HCMDB: the human cancer metastasis database. Nucleic Acids Res 46:D950–D955. https://doi.org/10.1093/nar/gkx1008

Zhu H, Zheng T, Yu J et al (2018) LncRNA XIST accelerates cervical cancer progression via upregulating Fus through competitively binding with miR-200a. Biomed Pharmacother 105:789–797. https://doi.org/10.1016/j.biopha.2018.05.053

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Vaibhav Shukla[1] · Sandeep Mallya[2] · Divya Adiga[1] · S. Sriharikrishnaa[1] · Sanjiban Chakrabarty[1,3] · Shama Prasada Kabekkodu[1,3]

[1] Department of Cell and Molecular Biology, Manipal School of Life Sciences, Manipal Academy of Higher Education, Manipal, India

[2] Department of Bioinformatics, Manipal School of Life Sciences, Manipal Academy of Higher Education, Manipal, India

[3] Center for DNA Repair and Genome Stability (CDRGS), Manipal Academy of Higher Education, Manipal, Karnataka, India