Life Sciences 233 (2019) 116708

Contents lists available at ScienceDirect

Life Sciences

journal homepage: www.elsevier.com/locate/lifescie

Hypomethylation of the lncRNA SOX21-AS1 has clinical prognostic value in cervical cancer

Ruijie Wang[a,b], Ya Li[c], Peipei Du[c], Xiaoan Zhang[d], Xiaofu Li[c,*], Guomei Cheng[a,**]

[a] Department of Obstetrics and Gynecology, The Third Affiliated Hospital of Zhengzhou University, Zhengzhou, China
[b] Department of Obstetrics and Gynecology, Academy of Medical Sciences of Zhengzhou University, Zhengzhou, China
[c] Department of Cytopathology, The Third Affiliated Hospital of Zhengzhou University, Zhengzhou, China
[d] Department of Imaging, The Third Affiliated Hospital of Zhengzhou University, Zhengzhou, China

ARTICLE INFO

Keywords:
lncRNA SOX21-AS1
DNA methylation
lncRNA
Cervical cancer
Prognosis

ABSTRACT

Aims: Cervical cancer seriously affects women's health. The function of methylated alterations in the long non-coding RNAs (lncRNAs) promote the progression and metastasis of cancer. Our study aims to identify the functional effects of lncRNA methylation in cervical carcinogenesis.

Main methods: Genome-wide DNA methylation of 6 samples was assessed using the Illumina Infinium MethylationEPIC BeadChip. RNA sequencing (RNA-seq) data and survival follow-up time of 307 samples from The Cancer Genome Atlas (TCGA) dataset were enrolled in this study. The statistical analysis and graphical work were mainly realized by R language.

Key findings: Methylation map identified 3962 hypermethylated CpG sites and 4484 hypomethylated CpG sites in cervical cancer ($|\Delta\beta| \geq 0.20$). Bioinformatic analysis of the lncRNA expression identified 363 upregulated and 664 downregulated lncRNAs with log2 (fold change) ≥ 1.00 in squamous cervical carcinoma (SCC) samples. Weighted gene co-expression network analysis (WGCNA) and Venn diagram revealed that lncRNA MAGI2 antisense RNA 3 (lncRNA MAGI2-AS3), lncRNA WT1 antisense RNA (lncRNA WT1-AS) and lncRNA SOX21 antisense divergent transcript 1 (lncRNA SOX21-AS1) were important methylation changed lncRNAs. Kaplan-Meier survival curves showed only lncRNA SOX21-AS1 had clinical prognostic value in cervical cancer. Gene set enrichment analysis (GSEA) suggest that lncRNA SOX21-AS1 involve in the multiple cellular processes and might significantly suppress cervical tumorigenesis.

Significance: These insights into the functional role of lncRNA SOX21-AS1 DNA methylome alterations in cervical cancer might promote clinically new applicable in diagnosis and prognosis.

1. Introduction

Cervical cancer is the fourth leading cause of cancer death in women globally, with an estimated 570,000 new cases and 311,000 deaths in 2018 worldwide [1]. The National Central Cancer Registry found that the 5-year survival rate of cervical cancer was 45.4% in China [2] and tended to be found in younger people compared to most other cancers [3]. Over the last few decades, persistent human papillomavirus (HPV) infection has been shown to be the main cause of cervical cancer [4]. The combined application of HPV testing and liquid-based thin layer cell detection has somewhat increased the rate of cervical cancer discovery [1,5], but the effect of screening remain limited [6]. Identification of cervical cancer biomarkers is great interest.

DNA methylation plays a role in genome stability and gene expression [7]. In particular, aberrant DNA promoter methylation is an important mechanism for loss of gene function in tumors [8,9]. Hypermethylation of the gene promoter region can result in gene silencing and an abnormal tumor microenvironment [10]. For example, the particular metabolic profile of hepatocellular carcinoma might be due to down-regulation of gene expression by gene hypermethylation [11]. Hypomethylation has been observed at early time points in tumor development, contributing to cellular transformation [12,13]. DNA methylation is stable and heritable and does not change the DNA sequence.

* Correspondence to: X. Li, Department of Cytopathology, The Third Affiliated Hospital of Zhengzhou University, No. 7 Front Kangfu Street, Er'qi District, Zhengzhou 450052, China.
** Correspondence to: G. Cheng, Department of Obstetrics and Gynecology, The Third Affiliated Hospital of Zhengzhou University, No. 7 Front Kangfu Street, Er'qi District, Zhengzhou 450052, China.
E-mail addresses: lixiaofu1964@hotmail.com (X. Li), guomei69@126.com (G. Cheng).

https://doi.org/10.1016/j.lfs.2019.116708
Received 9 June 2019; Received in revised form 27 July 2019; Accepted 28 July 2019
Available online 29 July 2019
0024-3205/

It is a highly sensitive marker for cancer [14] and has the potential to function as a primary screening tool [10].

Long non-coding RNAs (lncRNAs) are now known to play important roles in skin defects, brain damage, diabetes, hypertension and other diseases, especially cancer [15,16]. The lncRNAs are involved in multiple processes in cancer by interacting with RNA, DNA, and proteins through epigenetics, transcription, post-transcriptional regulation, and some pathological processes [17]. The reports observed that hypomethylation of EPIC1 (epigenetically-induced lncRNA1) was associated with poor prognosis in luminal B breast cancer [18], and found that during calcific aortic valve disease DNA hypomethylation in the promoter region of lncRNA H19 led to its overexpression, which promoted osteogenesis by decreasing the expression of Notch1 [19]. In mammals, the X-linked lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF (CCCTC binding factor) and maintains H3K27me3 methylation [20]. Another instance, the reexpression of DNMT1-associated colon cancer repressed lncRNA 1 (lncRNA DACOR1) results in a genome-wide increase of DNA methylation in human colon cancer cell lines [21]. LncRNA prostate cancer gene expression marker 1(lncRNA PCGEM1)promotes chemoresistance via inhibition of PARP cleavage and delaying the induction of tumor suppressors p53 and p21. The study demonstrated that lncRNA nuclear-enriched abundant transcript 1 (lncRNA NEAT1) knockdown affected the viability and morphology of Burkitt's lymphoma cells [22]. Therefore, we have sought novel biomarkers of cervical cancer in the cervical cancer methylated landscape in order to provide more effective diagnostic strategies.

2. Material and methods

2.1. Clinical samples

For the methylation array analysis, the inclusion criteria for the cancer group was being HPV16 positive, having cytological diagnosis of squamous cervical carcinoma (SCC), and having confirmed pathology. The normal group was negative for HPV and had only mild inflammatory reaction of cytological results. For this study, all cervical scrapings were randomly selected from a prospective collection of samples stored in our biological sample banks. All samples were collected at the Third Affiliated Hospital of Zhengzhou University from 2016 to 2018, and all specimens were from women with no history of prior dysplasia [23]. The clinical characteristics of the specimens are shown in Table 1. This study followed the ethical guidelines of the Ethics Committee of the Third Affiliated Hospital of Zhengzhou University.

2.2. HPV detection and genotyping

HPV DNA was isolated from the exfoliated cervical cells using the One-step Nucleic Acid Extraction Kit (HybriBio Ltd., Chaozhou, China). HPV genotyping was performed with the HPV GenoArray Diagnostic Kit (HybriBio Ltd., Chaozhou, China) that can differentiate between 37 HPV types using the HybriMax test [24].

Table 1
The clinical characteristics of the specimens detected by methylation microarray.

Sample	Age	Cytologic diagnosis	HPV genotyping	Histopathologic diagnosis	Group
x_893	34	Inflammation	–	–	N_1
x_130	28	Inflammation	–	–	N_2
x_158	28	Inflammation	–	–	N_3
x_461	60	SCC	16	SCC	SCC_1
x_572	40	SCC	16	SCC	SCC_2
x_250	35	SCC	16	SCC	SCC_3

2.3. DNA extraction

Genomic DNA was extracted from the cervical cell scrapings using the DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's instructions. After DNA extraction, the purified DNA was stored at −80 °C until further analysis. Each DNA sample was analyzed spectrophotometrically and by electrophoresis on 1% agarose gels [25].

2.4. Differential methylation locus analysis

The purity and concentration of the DNA were estimated using a Nanodrop 2000 (Thermo Scientific). Approximately 500 ng of genomic DNA from each sample was used for sodium bisulfite conversion using the EZ DNA Methylation Gold Kit (Zymo Research, USA) following the manufacturer's standard protocol. Genome-wide DNA methylation was assessed using the Illumina Infinium MethylationEPIC BeadChip (Illumina Inc., edition: V1.0B4, USA) according to the manufacturer's instructions. The raw data were analyzed using the ChAMP package in R for deriving the methylation level, and the methylation status of each probes was denoted as the β-value. CpG sites having $|\Delta\beta| \geq 0.20$ (in SCC vs. normal cervical cell scrapings) and $p \leq 0.01$ were considered to be the differentially methylated sites [26]. The CpG sites were identified as hypermethylated ($\Delta\beta > 0$) or hypomethylated ($\Delta\beta < 0$) [27]. All raw methylation array data were deposited in the Gene Expression Omnibus database under accession ID: GSE 134772.

2.5. RNA-sequencing data collection

Transcription-level data from the cervical samples and complete clinical datasets were obtained from TCGA (https://portal.gdc.cancer.gov/). In total, 307 samples of RNA sequencing data and survival follow-up time were enrolled in this study.

2.6. Screening of differentially expressed lncRNAs and mRNAs

Genome-wide transcriptional expression profiles were obtained from the RNA-seq count data from TCGA, which consists of mRNA and lncRNA expression data. The edgeR package of R version 3.5.2 (http://www.r-project.org) was used to screen for differentially expressed lncRNAs and mRNAs. A log2 (fold change) ≥ 1.00 and a false discovery rate (FDR) < 0.05 were considered to be indicative of differential expression.

2.7. Co-expression analysis

Weighted gene co-expression network analysis (WGCNA) is a systems biology method for constructing relationship patterns [28] that uses the soft threshold method to provide a more extensive and exact correlation between transcripts compared to more general methods such as Pearson's correlation coefficient [29]. Therefore, we used the WGCNA package in R 3.5.2 to build a co-expression network in order to analyze the potential mRNA targets of lncRNAs. Differentially expressed lncRNAs and mRNAs with log2 (fold change) ≥ 1.00 and FDR < 0.05 were analyzed. In order to screen long non-coding genes that met the requirements as far as possible, the soft threshold was set to a minimum of 0.61, and this yielded an mRNA cluster that was co-expressed with the lncRNAs that was large enough for further functional analysis of lncRNAs [30].

2.8. Screening of methylation-modified long non-coding genes in the promoter region and target lncRNAs' survival analysis

The overlapping set of long non-coding genes with methylated CpG sites in the promoter region and co-expression screening-related lncRNAs was used for further analysis. In order to reflect the filtering process, we produced a Venn diagram in R 3.5.2 that showed all

possible logical relations among a finite collection of the set of all methylated CpG loci, the set of promoter-region CpG loci, the set of RNA-seq lncRNAs, and the set of lncRNAs co-expressed with mRNAs. At the intersection of these sets were lncRNA MAGI2 antisense RNA 3 (lncRNA MAGI2-AS3), lncRNA WT1 antisense RNA (lncRNA WT1-AS) and lncRNA SOX21 antisense divergent transcript 1 (lncRNA SOX21-AS1) with common features in the promoter region and similar changes in methylation sites. We then used log10 (gene expression) data to create a column chart in the ggpubr package from R 3.5.2. The heatmap showed that lncRNA MAGI2-AS3 and lncRNA WT1-AS had hypermethylated expression in the promoter, while lncRNA SOX21-AS1 was hypomethylated in the promoter. Clinical data were extracted from the TCGA database. We also studied the association between key lncRNAs and the survival time of cervical cancer patients using Kaplan-Meier survival curves.

2.9. Gene set enrichment analysis

LncRNA SOX21-AS1's chromosome localization was obtained from NCBI (https://www.ncbi. nlm.nih.gov/gene/). GSEA was performed using the GSEA software, and the gene sets used in this work were downloaded from the Molecular Signatures Database (http://software.broadinstitute.org/gsea/msigdb).

2.10. Statistical analysis

Normalization of the methylation data was performed using the peak-based correction method [31]. The changes in expression of lncRNA SOX21-AS1 and SOX21 were analyzed by Pearson correlation in GraphPad Prism 7. The co-expression network diagram was drawn using the Cytoscape software (version 3.7.0). The prognostic values of lncRNA MAGI2-AS3, lncRNA WT1-AS, lncRNA SOX21-AS1, ZEB1, DDR2, and WT1 expression in cervical cancer patients were estimated by Kaplan-Meier analysis. Other statistical computations and figure drawing were performed with several packages (ChAMP, edgeR, WGCNA, survival, ggpubr, pheatmap, and ggplot2) in the statistical software environment R version 3.5.2 (https://www.r-project.org/). For all statistical methods, $p < 0.05$ was considered to be a significant difference.

3. Results

3.1. Differentially methylated CpG loci distribution in cervical cancer and healthy tissues

Methylation sites were interrogated on the Illumina Infinium MethylationEPIC BeadChip testing 853,307 CpG sites located in the gene promoter, gene coding region, CpG islands (CGIs), and enhancer regions found in the ENCODE and FANTOM5 project [32]. We identified 3962 hypermethylated CpG sites and 4484 hypomethylated CpG sites in the cervical cancer samples ($|\Delta\beta| \geq 0.20$). Hierarchical clustering analysis showed systematic variations in the level of methylation among the samples. The data suggested that the levels of methylation in cervical cancer samples differed from the levels in normal controls (Fig. 1A). To find the majority of events in cervical cancer and normal tissue samples, the distribution of CpG sites in the gene context was further analyzed. In the methylation distance to transcription start sites (TSS), cervical cancer samples compared to normal samples had greater differences in the methylation β-value between TSS to approximately TSS1500 (Fig. 1B). The gene promoter region contained TSS1500, TSS200, the 5′-UTR, and exon 1 [27]. We also analyzed the location of the hypermethylated and hypomethylated CpG loci in the CGIs, CGI shore, CGI shelf, and other regions. The hypermethylated CpG sites were mainly concentrated in the CGIs in contrast to the hypomethylated CpG sites, which were distributed in the CGI shore or in sites farther away (Fig. 1C). Overall, the number of CpG sites was greatest in the CGIs. A CGI is a region with relatively high content of CpG dinucleotides, and approximately 70% of the annotated gene promoters are associated with a CGI [33]. In addition, genes whose promoters are especially rich in CpG sequences tend to be expressed in most tissues [34]. Therefore, the gene promoter region contained core regions with significant changes in methylated CpG sites. We then selected probes with methylation site changes in the promoter region of lncRNA genes for further analysis.

3.2. Differentially expressed lncRNAs and mRNAs in cervical cancer and normal tissues

Volcano plots showed the significantly changed lncRNAs and mRNAs with log2 (fold change) ≥ 1.00 and FDR < 0.05 in cervical cancer compared with normal tissues (Fig. 2). In total, 1027 lncRNAs showed differential expression in cervical cancer, including 363 upregulated lncRNAs and 664 downregulated lncRNAs. Of the 3745 mRNAs that showed differential expression between the two groups, 1631 were upregulated and 2114 were downregulated.

3.3. Co-expression analysis of differentially expressed lncRNAs and mRNAs

At present, correlations between lncRNAs and mRNAs based on their expression profiles have reportedly pointed at co-regulation or functional relatedness, although the functions of most lncRNAs have not been well established. [35]. Therefore, by constructing a lncRNA-mRNA co-expression network, we can predict the role that lncRNAs might play in cervical cancer. The co-expression network was composed of 258 nodes and 1688 edges between 50 lncRNAs and 208 coding genes (Fig. 3).

3.4. DNA promoter methylation correlated with functional long non-coding genes and clinical prognosis

Co-expression results showed that 50 lncRNAs were connected to mRNAs. We used the Venn diagram to show the results of the overlap between the co-expression screening-related lncRNAs and the methylated CpG sites in the promoter regions of the long non-coding genes (Fig. 4A). LncRNA MAGI2-AS3, lncRNA WT1-AS, and lncRNA SOX21-AS1 were at the core of the intersection (Table 2). The expression of lncRNA MAGI2-AS3 was reduced in cervical cancer compared to normal tissues ($p < 0.01$), and its promoter region methylation probes (cg04652097, cg08641579, cg21784917) showed hypermethylation in the cervical cancer samples. The results for lncRNA WT1-AS ($p < 0.01$) were consistent with those of lncRNA MAGI2-AS3, and the lncRNA WT1-AS DNA methylation probes (cg12766736, cg07281879, cg25247290) were significantly higher in the cervical cancer samples. In contrast, expression of lncRNA SOX21-AS1 was increased in the cervical cancer samples ($p < 0.05$) (Fig. 4B), and the probe cg11007120 showed hypomethylation in the promoter gene (Fig. 4C). LncRNA MAGI2-AS3 was connected with ZEB1 and DDR2 mRNAs, lncRNA WT1-AS was connected with WT1 mRNA, and lncRNA SOX21-AS1 was connected with SOX21 mRNA (Table 3). Next, we evaluated the prognostic impact of lncRNA MAGI2-AS3, lncRNA WT1-AS, and lncRNA SOX21-AS1 expression in cervical cancer. The Kaplan-Meier survival analysis showed that patients in the TCGA dataset with higher lncRNA SOX21-AS1 levels in their tumors had better survival compared to patients with lower levels ($p < 0.05$, Fig. 4D). There were no significant differences between the expression levels of lncRNA MAGI2-AS3 and lncRNA WT1-AS and the survival time of the patients with cervical cancer ($p > 0.05$, Fig. 4D). The potential association between the expression level of coding genes (ZEB1, DDR2, WT1, and SOX21) and survival time of cervical cancer patients was examined. Interestingly, we found the same situation in which increased SOX21 in tumor patients was accompanied by improved overall survival ($p < 0.05$, Fig. 4D). The prognostic analysis of ZEB1, DDR2, and WT1 showed no

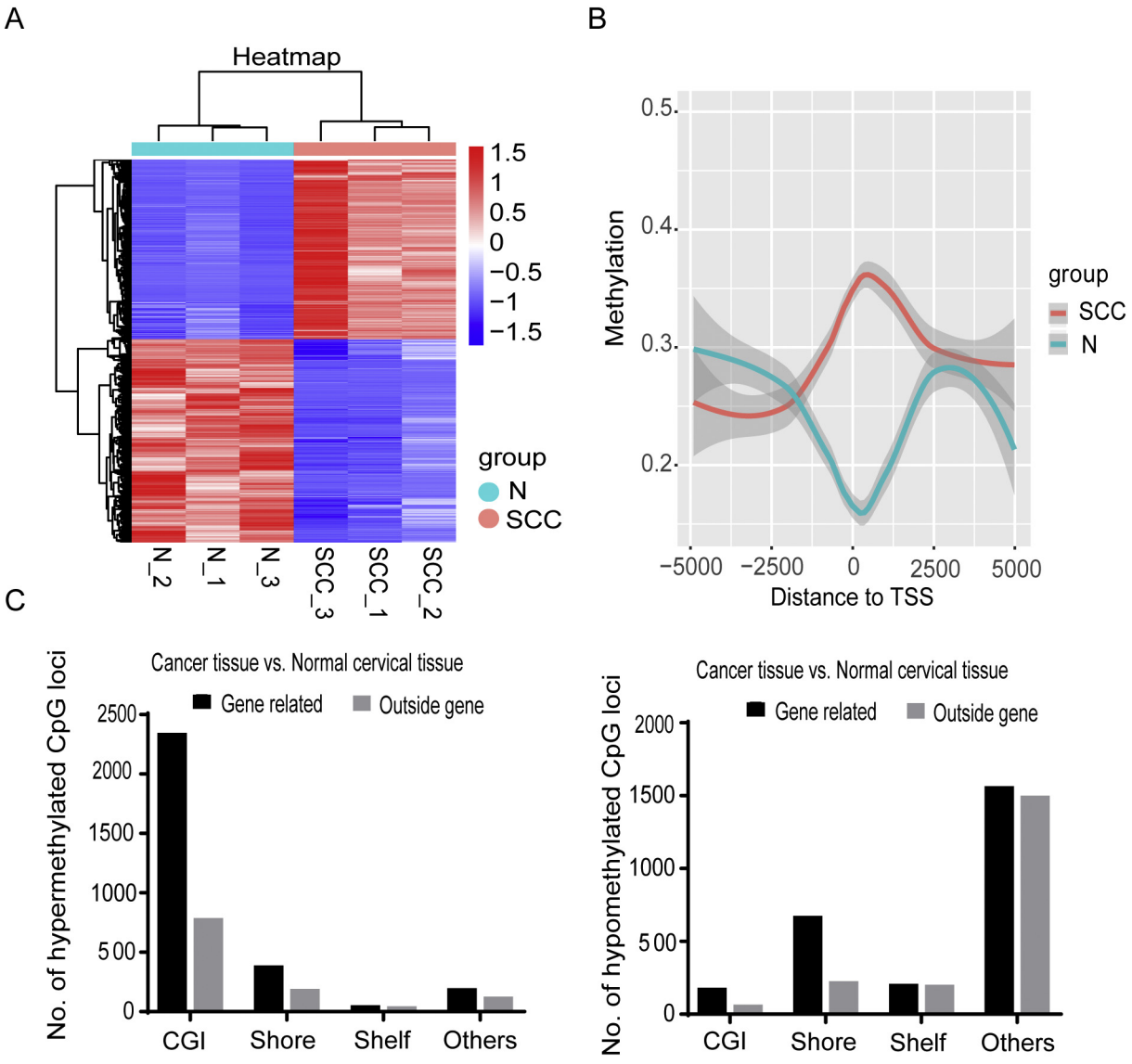


Fig. 1. Distribution of differentially methylated CpG loci in cervical cancer and normal tissues. (A) Heatmap showing the expression profiles of hyper- and hypo-methylated CpG sites in SCC. (B) The methylation distance to TSS. (C) The numbers of hypermethylated and hypomethylated CpG loci located in the CpG island (CGI), CpG island shore, CpG island shelf, and other regions.

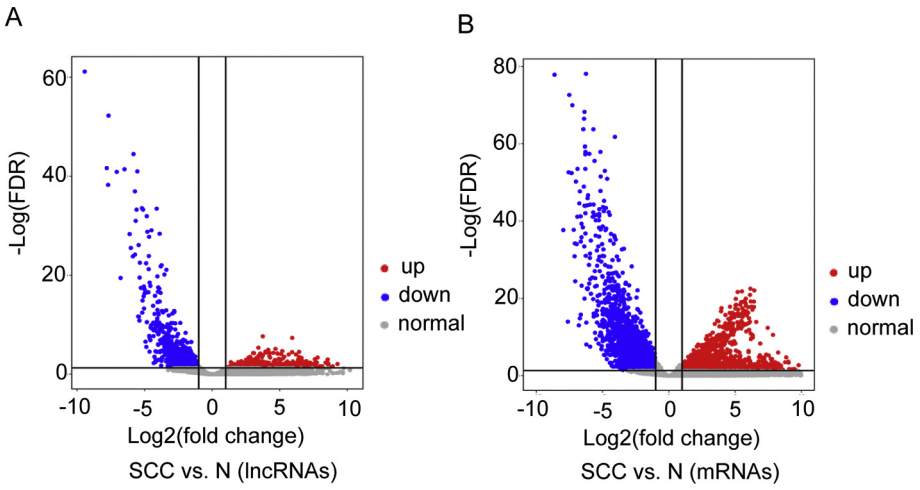


Fig. 2. Volcano plots showing significantly changed lncRNAs and mRNAs. (A) lncRNAs and (B) mRNAs with log2 (fold change) ≥ 1.00 and FDR < 0.05 in cervical cancer compared with normal tissues.

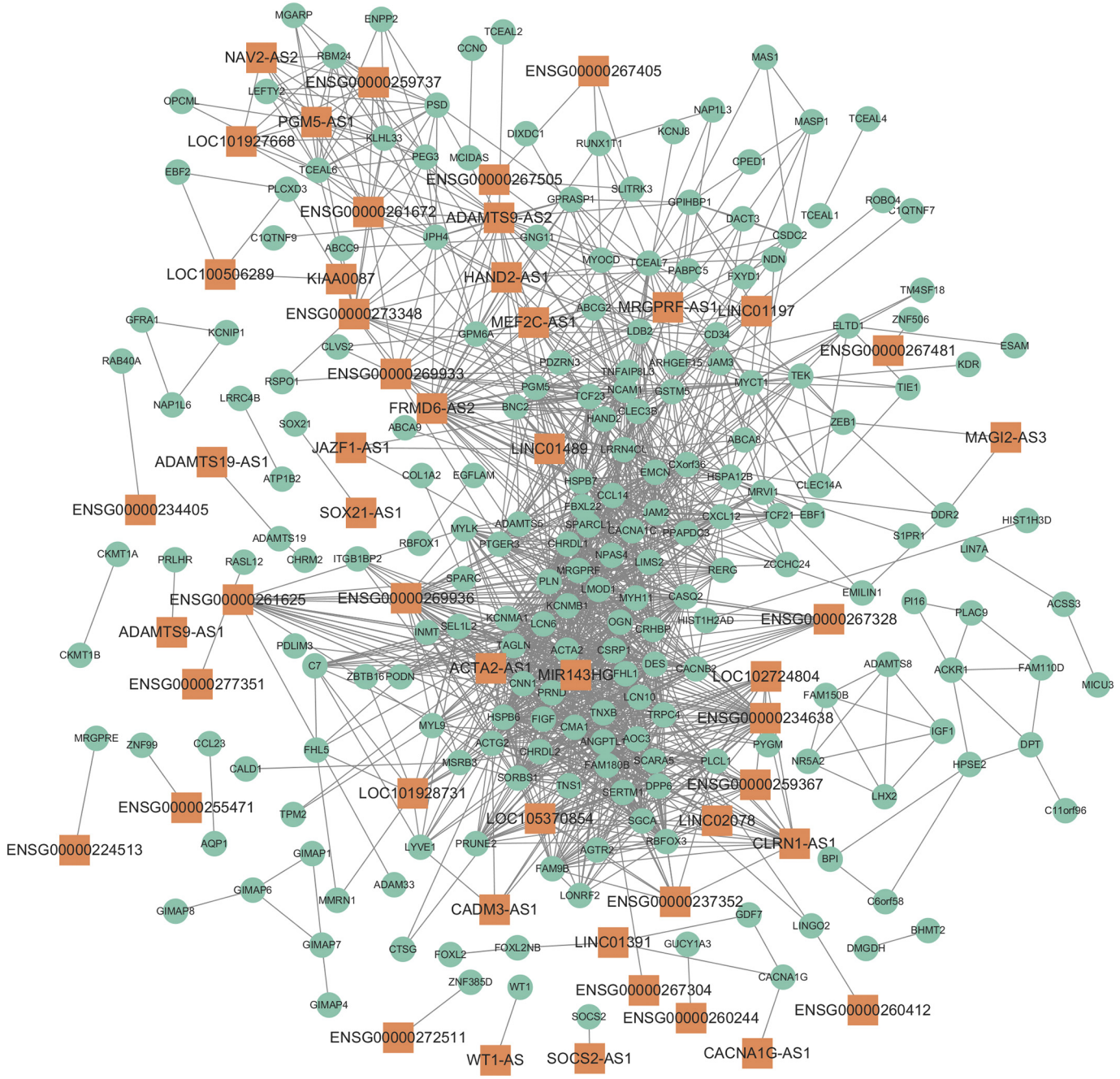
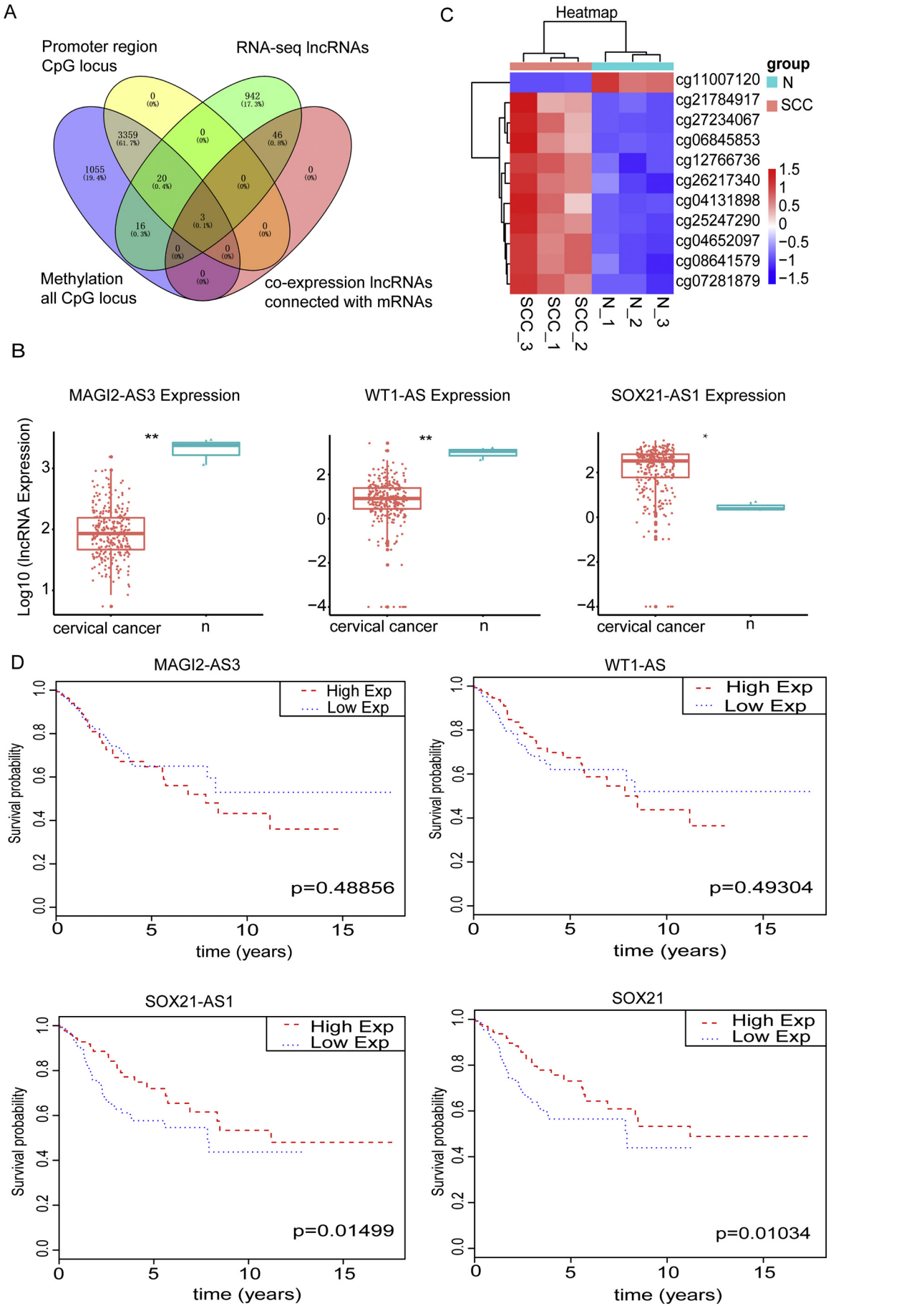


Fig. 3. Co-expression network constructed by weighted correlation network analysis. Green circles, mRNAs; orange rectangles, lncRNAs; lines, correlative relationships. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

statistical significance ($p > 0.05$, Table 3). Finally, we used SOX21's functional annotation in the RNA-seq data in the TCGA database to predict the potential functions of lncRNA SOX21-AS1 in cervical cancer.

3.5. LncRNA SOX21-AS1–related biological processes in cervical cancer

LncRNA SOX21-AS1 is located on human chromosome 13q32.1:94712716-94715945, a 3230 bp non-coding RNA expressed in the nucleus, and its direction of transcription is the opposite of SOX21 (Fig. 5A). Yang et al. found that lncRNA SOX21-AS1 shares a bidirectional promoter (head-to-head) with SOX21 on human chromosome 13q32.1 and that they are located upstream of a common CpG-rich region [36]. Therefore, we hypothesized that their transcriptional activity might be affected by DNA methylation. We found that SOX21 and lncRNA SOX21-AS1 were simultaneously upregulated in cervical cancer compared with normal tissues. SOX21 expression was significantly correlated with lncRNA SOX21-AS1 expression in 304 cervical cancer patients ($r = 0.9054$, $p < 0.0001$, Fig. 5B). GSEA analysis of SOX21 showed that lncRNA SOX21-AS1 was positively correlated with biological processes such as epidermis development, molting cycle, skin development, and skin epidermis development as well as with molecular functions in cervical cancer, including nucleic acid binding transcription factor activity and sequence-specific DNA binding of RNA polymerase II (Fig. 5C). Taken together, our results indicated that lncRNA SOX21-AS1 was a good prognostic indicator in cervical cancer due to its role as a tumor suppressor. LncRNA SOX21-AS1 was highly expressed in cervical cancer due to the hypomethylation of its promoter region. Furthermore, the predominantly nuclear localization of lncRNA



(caption on next page)

Fig. 4. Methylation-modified lncRNAs in the promoter region. (A) Venn diagram showing how the lncRNAs studied here were identified at the intersection of the sets of methylation at all CpG sites, the presence of CpG loci, RNA-seq–identified lncRNAs, and lncRNAs co-expressed with mRNAs. (B) LncRNA MAGI2-AS3, lncRNA WT1-AS, and lncRNA SOX21-AS1 were expressed in cervical cancer and normal tissues. * $p < 0.05$ and ** $p < 0.01$. (C) The changes in methylation sites of the promoter regions in lncRNA MAGI2-AS3, lncRNA SOX21-AS1, and lncRNA WT1-AS. (D) Survival analysis of lncRNA MAGI2-AS3, lncRNA WT1-AS, lncRNA SOX21-AS1, and SOX21 in cervical cancer.

Table 2
Overlapping relationships of hub long non-coding genes in the Venn diagram.

Methylation				RNA-seq	
Gene name	CpG ID	Region	Δβ-values	Gene name	Log (FC)
MAGI2-AS3	cg04652097	Exon1; 5′-UTR	0.431	MAGI2-AS3	−3.992
	cg08641579	Exon1; 5′-UTR	0.398		
	cg21784917	5′-UTR	0.244		
WT1-AS	cg12766736	5′-UTR	0.332	WT1-AS	−4.833
	cg07281879	5′-UTR	0.385		
	cg25247290	5′-UTR	0.487		
SOX21-AS1	cg11007120	5′-UTR	−0.276	SOX21-AS1	7.092

Log (FC):Log2 (Fold Change).

Table 3
Survival analysis of key lncRNAs and connected mRNAs.

LncRNA	Survival analysis p-value	mRNA	Survival analysis p-value
MAGI2-AS3	0.48856	ZEB1; DDR2	0.99665; 0.92081
WT1-AS	0.49304	WT1	0.24225
SOX21-AS1	0.01499	SOX21	0.01034

SOX21-AS1 might avoid the RNA interference-silencing machinery in the cytoplasm of the cells [36]. Finally, the biological function of lncRNA SOX21-AS1 was enhanced in cervical cancer.

4. Discussion

Entire genomic landscapes are accessible through high-throughput sequencing technology [37], and the big data being generated through genomics studies is one of the clearest ways to provide precise oncological interventions [38]. In the present study, we used Illumina Infinium MethylationEPIC BeadChip and RNA-seq data from the TCGA projects to identify the comprehensive character of the lncRNA DNA methylation landscape in cervical cancer. Our research strategy was consistent with that of Wang et al., who showed that integrating methylation microarray and RNA-seq data is a cost-effective strategy for studying the regulatory effect of DNA methylation of lncRNA genes [18]. Our analysis showed that 3962 hypermethylated CpG sites and 4484 hypomethylated CpG sites that were distributed throughout the coding and non-coding genes. Most of the methylation site changes were concentrated in the promoter region, which suggests that the promoter regions of genes are subject to control by epigenetic methylation.

LncRNAs have received much attention in the biological sciences, but their precise functions are still poorly understood. Previous studies, by repurposing multi-dimensional genomic and epigenetic data, have successfully identified differentially expressed lncRNAs in cancer [39] and cancer-related lncRNAs that are regulated by changes in DNA methylation [18]. To provide some insights into the biological functions of lncRNAs in cervical cancer, we integrated lncRNA and mRNA data from the TCGA projects. We identified 363 upregulated lncRNAs, 664 downregulated lncRNAs, and 3745 differentially expressed mRNAs, suggesting the extensive involvement of lncRNAs in cervical cancer. The functional annotations of lncRNAs are not yet complete [29], so in order to understand the functions of lncRNAs involved in the progression of cervical cancer we used WGCNA. The module analyzed by WGCNA is a group of genes with similar expression profiles, and these genes are considered to be functionally related [28], and Feng et al. used this method to predict the roles of lncRNAs in acute myeloid leukemia [29]. Similarly, Liu et al. successfully established the potential biological function network of lncRNAs in cancer development by observing the functions of co-expressed oncogenes [40]. In the same way, we identified lncRNAs that are associated with mRNAs. In our work, we sought to find lncRNAs that are epigenetically activated or inactivated in tumors through DNA methylation in the promoter region. Next, we carried out the intersection analysis between the methylation data set and co-expression analysis of the lncRNAs set, and this identified epigenetic regulation of lncRNA MAGI2-AS3, lncRNA WT1-AS, and lncRNA SOX21-AS1. Such expressed regulation of lncRNA was consistent with their somatic DNA methylation alterations identified in this study.

We found low expression of the hypermethylated genes, such as the lncRNA MAGI2-AS3 and lncRNA WT1-AS genes, but the down-regulation of lncRNA MAGI2-AS3 by hypermethylation was not observed in other studies. In a previous study, lncRNA MAGI2-AS3 was associated with poor prognosis in bladder cancer [41], while our study showed that the prognostic power of lncRNA MAGI2-AS3 and cervical cancer was not statistically significant. The ZEB1 and DDR2 genes were associated with lncRNA MAGI2-AS3 in the co-expression results, but because ZEB1 and DDR2 were not associated with the prognosis of cervical cancer, no further analysis was performed.

LncRNA WT1-AS is known as a cancer-related lncRNA [42], and in ovarian clear cell adenocarcinoma hypermethylation is an important mechanism of lncRNA WT1-AS gene inactivation [43]. LncRNA WT1-AS, as a tumor suppressor, promotes apoptosis of hepatocellular carcinoma cells [44] and inhibits gastric cancer cells' proliferation and invasion [45]. However, there are no reports on the prognostic value of lncRNA WT1-AS in cervical cancer. According to the analysis of the clinical data from the TCGA database, lncRNA WT1-AS was of no value in the prognosis of cervical cancer patients. In the co-expression results, WT1 had a connection to lncRNA WT1-AS, and WT1 also had little prognostic value in cervical cancer.

In contrast, the expression of hypo-methylation-modified lncRNA SOX21-AS1 was significantly higher in the cervical cancer group compared to the normal group. LncRNA SOX21-AS1 is the antisense of SOX21, and both are located on chr13q32.1. Our data showed synergistic effects on the increases of lncRNA SOX21-AS1 and SOX21 expression in cervical tumors ($r = 0.9054$, $p < 0.0001$). A study by Yang et al. in 2016 showed that lncRNA SOX21-AS1 shares a bidirectional promoter with SOX21, and the biological functions of lncRNA SOX21-AS1 were predicted by SOX21 functional annotation [36]. In the present study, GSEA analysis of SOX21 suggested that lncRNA SOX21-AS1 is involved in cellular biological processes and in sequence-specific DNA binding of RNA polymerase II transcription factor in cervical cancer. In a previous study, the enrichment of RNA polymerase II at enhancer elements was shown to be dependent on the transcription of lncRNAs in breast and prostate cancers [46]. Thus, lncRNA SOX21-AS1 might act as a molecular signal for transcriptional activity or might guide specific complexes to the site of transcription.

According to the survival analysis, lncRNA SOX21-AS1 was a good biomarker in cervical cancer and functioned as a tumor suppressor. It is worth noting that the prognostic value of lncRNA SOX21-AS1 was not the same as some have previously reported. LncRNA SOX21-AS1 has been verified as an oncogene in tumorigenesis, and it is increased in lung adenocarcinoma, promotes cancer cell proliferation, and is an indicator of poor prognosis [47]. In hepatocellular carcinoma, silencing of lncRNA SOX21-AS1 could suppress cell proliferation and metastasis

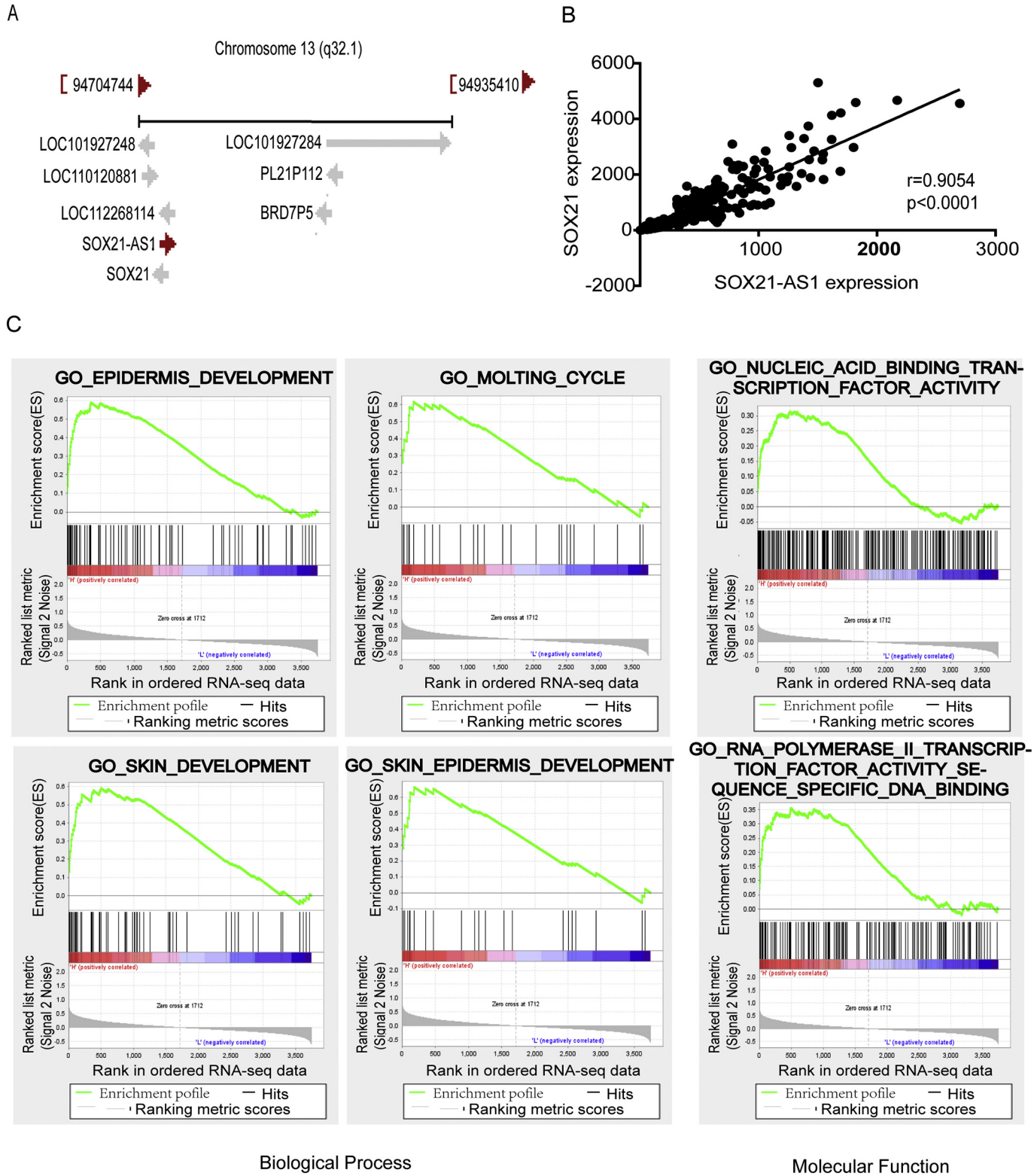


Fig. 5. LncRNA SOX21-AS1-induced functions. (A) Chromosome localization of lncRNA SOX21-AS1 and SOX21. (B) Correlation coefficients of lncRNA SOX21-AS1 and SOX21 expression in cervical cancer. (C) The functions of SOX21 using GSEA.

and promote cell cycle arrest [48]. LncRNA SOX21-AS1 was over-expressed in colorectal cancer, and the aberrant expression of lncRNA SOX21-AS1 indicated poor prognosis in these patients [49]. Furthermore, lncRNA SOX21-AS1 has been shown to be up-regulated in Alzheimer's disease, and silencing of lncRNA SOX21-AS1 inhibits apoptosis and reduces oxidative stress damage by activating the Wnt signaling pathway [50]. Unlike what has been shown above in relation to cancer, lncRNA SOX21-AS1 is not likely to be an oncogenic lncRNA in cervical cancer. LncRNA SOX21-AS1 showed good prognosis in glioblastoma, and it appeared to act as a protective lncRNA because its expression was higher in the low-risk group [51].

There is increasing evidence for the tumor suppressor activity of

lncRNA SOX21-AS1. For example, aberrant DNA hypermethylation-silenced lncRNA SOX21-AS1 is an indicator of poor prognosis in patients with oral cancer [36], and this is in line with our results. LncRNA SOX21-AS1 plays different roles in different cancers, implying that lncRNA SOX21-AS1, as a strong prognostic indicator, has obvious tumor heterogeneity and important clinical value. However, how the lncRNA SOX21-AS1 pathway is involved in the complex tumor environment is unclear. We found that lncRNA SOX21-AS1 might influence the occurrence of cervical cancer through the regulation of SOX2 by SOX21. A previous study showed that SOX21 overexpression inhibited the expression of SOX2 and induced apoptosis in glioblastoma [52,53]. Ferletta et al. reported that downregulation of SOX2 using siRNA reduced SOX21 expression at the protein and mRNA levels in glioma cells, suggesting that SOX2 can positively regulate the transcriptional activity of SOX21 [53]. There was a balance between SOX21 and SOX2 expression, and although SOX2 is an oncogene in cervical cancer [54], the interaction of SOX21 and SOX2 is more likely to inhibit cervical cancer progression. We inferred that overexpression of lncRNA SOX21-AS1 enhanced the inhibitory effect of SOX21 on SOX2 and suppressed the progression of cervical cancer because of the effect of hypomethylation in the promoter region of the lncRNA SOX21-AS1 gene. In addition, the predominantly nuclear localization of lncRNA SOX21-AS1 might avoid the RNA interference-silencing machinery in the cytoplasm [36].

5. Conclusion

In summary, we describe the strong prognostic association of lncRNA SOX21-AS1 in cervical cancer through the comprehensive application of high-throughput chip technology, transcript sequencing, and bioinformatics analysis. The tumor growth suppression in cervical cancer as a result of increased lncRNA SOX21-AS1 expression and the observation of a lncRNA SOX21-AS1–specific mechanism in cervical cancer supports future development of lncRNA-based cervical cancer therapies.

Acknowledgements

We thank Ruirui Wang, Yanfang Zhi and Fang Wang for excellent technical assistance and helpful discussion. The authors are indebted to all colleagues for their valuable suggestions and discussions.

Funding

This study was supported by project and fellowship grants from the Science and Technology Department of Henan (China, Grant No.22170005).

Author contribution statement

Ruijie Wang performed most of the experiments; Peipei Du and Ya Li contributed to their design; Ruijie Wang, Xiaofu Li and Guomei Cheng analyzed data and wrote the manuscript; Ruijie Wang and Guomei Cheng finalized the paper; Guomei Cheng supervised the study. All authors read and approved the final manuscript.

Declaration of competing interests

The authors declare no conflicts of interest.

References

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

[2] H. Zeng, R. Zheng, Y. Guo, S. Zhang, X. Zou, N. Wang, L. Zhang, J. Tang, J. Chen, K. Wei, et al., Cancer survival in China, 2003–2005: a population-based study, Int. J. Cancer 136 (8) (2015) 1921–1930, https://doi.org/10.1002/ijc.29227.

[3] X. Jiang, H. Tang, T. Chen, Epidemiology of gynecologic cancers in China, J. Gynecol. Oncol. 29 (1) (2018) e7, https://doi.org/10.3802/jgo.2018.29.e7.

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

[5] Q. Feng, A. Balasubramanian, S.E. Hawes, P. Toure, P.S. Sow, A. Dem, B. Dembele, C.W. Critchlow, L. Xi, H. Lu, et al., Detection of hypermethylated genes in women with and without cervical neoplasia, J. Natl. Cancer Inst. 97 (4) (2005) 273–282, https://doi.org/10.1093/jnci/dji041.

[6] X. Hao, H. Luo, M. Krawczyk, W. Wei, W. Wang, J. Wang, K. Flagg, J. Hou, H. Zhang, S. Yi, et al., DNA methylation markers for diagnosis and prognosis of common cancers, Proc. Natl. Acad. Sci. U. S. A. 114 (28) (2017) 7414–7419, https://doi.org/10.1073/pnas.1703577114.

[7] M.H.J. Esteller, Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours, J. Pathol. 196 (1) (2002) 1–7, https://doi.org/10.1002/path.1024.

[8] N. Ohtani-Fujita, T. Fujita, A. Aoike, N.E. Osifchin, P.D. Robbins, T. Sakai, CpG methylation inactivates the promoter activity of the human retinoblastoma tumor-suppressor gene, Oncogene 8 (4) (1993) 1063–1067, https://doi.org/10.1016/0145-21 26(93)90034-I.

[9] D.F. Jarrard, H. Kinoshita, Y. Shi, C. Sandefur, D. Hoff, L.F. Meisner, C. Chang, J.G. Herman, W.B. Isaacs, N. Nassif, Methylation of the androgen receptor promoter CpG island is associated with loss of androgen receptor expression in prostate cancer cells, Cancer Res. 58 (23) (1998) 5310–5314 (DOI: 10.1097/0 0005392-199904010-00253).

[10] R.D. Steenbergen, P.J. Snijders, D.A. Heideman, C.J. Meijer, Clinical implications of (epi)genetic changes in HPV-induced cervical precancerous lesions, Nat. Rev. Cancer 14 (6) (2014) 395–405, https://doi.org/10.1038/nrc3728.

[11] Cancer Genome Atlas Research Network, Electronic address wbe, Cancer genome atlas research N: comprehensive and integrative genomic characterization of hepatocellular carcinoma, Cell 169 (7) (2017) 1327–1341 e1323 https://doi.org/10.1016/j.cell.2017.05.046.

[12] P.A. Jones, S.B. Baylin, The fundamental role of epigenetic events in cancer, Nat Rev Genet 3 (6) (2002) 415–428, https://doi.org/10.1038/nrg816.

[13] H. Shen, P.W. Laird, Interplay between the cancer genome and epigenome, Cell 153 (1) (2013) 38–55, https://doi.org/10.1016/j.cell.2013.03.008.

[14] J.W. Kornfeld, J.C. Bruning, Regulation of metabolism by long, non-coding RNAs, Front. Genet. 5 (2014) 57, https://doi.org/10.3389/fgene.2014.00057.

[15] L. Li, H.Y. Chang, Physiological roles of long noncoding RNAs: insight from knockout mice, Trends Cell Biol. 24 (10) (2014) 594–602, https://doi.org/10.1016/j.tcb.2014.06.003.

[16] H.N. Zhang, Q.Q. Xu, A. Thakur, M.O. Alfred, M. Chakraborty, A. Ghosh, X.B. Yu, Endothelial dysfunction in diabetes and hypertension: role of microRNAs and long non-coding RNAs, Life Sci. 213 (2018) 258–268, https://doi.org/10.1016/j.lfs.2018.10.028.

[17] K.C. Wang, Y.W. Yang, B. Liu, A. Sanyal, R. Corces-Zimmerman, Y. Chen, B.R. Lajoie, A. Protacio, R.A. Flynn, R.A. Gupta, et al., A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression, Nature 472 (7341) (2011) 120–124, https://doi.org/10.1038/nature09819.

[18] Z. Wang, B. Yang, M. Zhang, W. Guo, Z. Wu, Y. Wang, L. Jia, S. Li, Cancer genome atlas research N, Xie W et al: lncRNA epigenetic landscape analysis identifies EPIC1 as an oncogenic lncRNA that interacts with MYC and promotes cell-cycle progression in cancer, Cancer Cell 33 (4) (2018) 706–720 e709 https://doi.org/10.1016/j.ccell.2018.03.006.

[19] F. Hadji, M.C. Boulanger, S.P. Guay, N. Gaudreault, S. Amellah, G. Mkannez, R. Bouchareb, J.T. Marchand, M.J. Nsaibia, S. Guauque-Olarte, et al., Altered DNA methylation of long noncoding RNA H19 in calcific aortic valve disease promotes mineralization by silencing NOTCH1, Circulation 134 (23) (2016) 1848–1862, https://doi.org/10.1161/CIRCULATIONAHA.116.023116.

[20] F. Yang, X. Deng, W. Ma, J.B. Berletch, N. Rabaia, G. Wei, J.M. Moore, G.N. Filippova, J. Xu, Y. Liu, et al., The lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation, Genome Biol. 16 (2015) 52, , https://doi.org/10.1186/s13059-015-0618-0.

[21] M.E. Forrest, A.M. Khalil, Review: regulation of the cancer epigenome by long noncoding RNAs, Cancer Lett. 407 (2017) 106–112, https://doi.org/10.1016/j.can let.20 17.03.040.

[22] A. Bhan, M. Soleimani, S.S. Mandal, Long noncoding RNA and cancer: a new paradigm, Cancer Res. 77 (15) (2017) 3965–3981, https://doi.org/10.1158/0008-5472.CA N-16-2634.

[23] D. Solomon, D. Davey, R. Kurman, A. Moriarty, D. O'Connor, M. Prey, S. Raab, M. Sherman, D. Wilbur, T. Wright Jr.et al., The 2001 Bethesda system: terminology for reporting results of cervical cytology, JAMA 287 (16) (2002) 2114–2119, https://doi.org/10.1001/jama.287.16.2114.

[24] K. Shigehara, T. Sasagawa, J. Doorbar, S. Kawaguchi, Y. Kobori, T. Nakashima, M. Shimamura, Y. Maeda, T. Miyagi, Y. Kitagawa, et al., Etiological role of human papillomavirus infection for inverted papilloma of the bladder, J. Med. Virol. 83 (2) (2011) 277–285, https://doi.org/10.1002/jmv.21966.

[25] T. Mamiatis, E.F. Fritsch, J. Sambrook, J. Engel, Molecular cloning-a laboratory manual. New York: cold Spring Harbor laboratory. 1982, 545 S. 42 $, Acta Biotechnol. (1985), https://doi.org/10.1002/abio.370050118.

[26] M. Bibikova, B. Barnes, C. Tsan, V. Ho, B. Klotzle, J.M. Le, D. Delano, L. Zhang, G.P. Schroth, K.L. Gunderson, et al., High density DNA methylation array with single CpG site resolution, Genomics 98 (4) (2011) 288–295, https://doi.org/10.

1016/j.ygeno.2011.07.007.
[27] S.A. Farkas, N. Milutin-Gasperov, M. Grce, T.K. Nilsson, Genome-wide DNA methylation assay reveals novel candidate biomarker genes in cervical cancer, Epigenetics 8 (11) (2013) 1213–1225, https://doi.org/10.4161/epi.26346.
[28] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinformatics 9 (2008) 559, , https://doi.org/10.1186/1471-210 5-9-559.
[29] Y. Feng, Y. Shen, H. Chen, X. Wang, R. Zhang, Y. Peng, X. Lei, T. Liu, J. Liu, L. Gu, et al., Expression profile analysis of long non-coding RNA in acute myeloid leukemia by microarray and bioinformatics, Cancer Sci. 109 (2) (2018) 340–353, https://doi.org/10.1111/cas.13465.
[30] Z. Dong, X. Zhu, Y. Li, L. Gan, H. Chen, W. Zhang, J. Sun, Oncogenomic analysis identifies novel biomarkers for tumor stage mycosis fungoides, Medicine (Baltimore) 97 (21) (2018) e10871, , https://doi.org/10.1097/MD.0000000000010871.
[31] S. Dedeurwaerder, M. Defrance, E. Calonne, H. Denis, C. Sotiriou, F. Fuks, Evaluation of the infinium methylation 450K technology, Epigenomics 3 (6) (2011) 771–784, https://doi.org/10.2217/epi.11.105.
[32] W. Zhou, P.W. Laird, H. Shen, Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes, Nucleic Acids Res. 45 (4) (2017) e22, https://doi.org/10.1093/nar/gkw967.
[33] S. Saxonov, P. Berg, D.L. Brutlag, A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters, Proc. Natl. Acad. Sci. U. S. A. 103 (5) (2006) 1412–1417, https://doi.org/10.1073/pnas.0510310103.
[34] A.M. Deaton, A. Bird, CpG islands and the regulation of transcription, Genes Dev. 25 (10) (2011) 1010–1022, https://doi.org/10.1101/gad.2037511.
[35] Q. Liao, C. Liu, X. Yuan, S. Kang, R. Miao, H. Xiao, G. Zhao, H. Luo, D. Bu, H. Zhao, et al., Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network, Nucleic Acids Res. 39 (9) (2011) 3864–3878, https://doi.org/10.1093/nar/gkq1348.
[36] C.M. Yang, T.H. Wang, H.C. Chen, S.C. Li, M.C. Lee, H.H. Liou, P.F. Liu, Y.K. Tseng, Y.L. Shiue, L.P. Ger, et al., Aberrant DNA hypermethylation-silenced SOX21-AS1 gene expression and its clinical importance in oral cancer, Clin. Epigenetics 8 (2016) 129, , https://doi.org/10.1186/s13148-016-0291-5.
[37] E.P. Consortium, An integrated encyclopedia of DNA elements in the human genome, Nature 489 (7414) (2012) 57–74, https://doi.org/10.1038/nature11247.
[38] J.L. Warner, D.A. Patt, Section editors for the IYSoCI: Cancer informatics in 2017: a new beginning and a bright future, Yearb Med Inform 27 (1) (2018) 223–226, https://doi.org/10.1055/s-0038-1667086.
[39] Z. Du, T. Fei, R.G. Verhaak, Z. Su, Y. Zhang, M. Brown, Y. Chen, X.S. Liu, Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer, Nat. Struct. Mol. Biol. 20 (7) (2013) 908–913, https://doi.org/10.1038/nsmb.2591.
[40] Y. Liu, M. Zhao, lnCaNet: pan-cancer co-expression network for human lncRNA and cancer genes, Bioinformatics 32 (10) (2016) 1595–1597, https://doi.org/10.1093/bioinformatics/btw017.
[41] F. Wang, Y. Zu, S. Zhu, Y. Yang, W. Huang, H. Xie, G. Li, Long noncoding RNA MAGI2-AS3 regulates CCDC19 expression by sponging miR-15b-5p and suppresses bladder cancer progression, Biochem. Biophys. Res. Commun. 507 (1–4) (2018) 231–235, https://doi.org/10.1016/j.bbrc.2018.11.013.
[42] A.L. Hancock, K.W. Brown, K. Moorwood, H. Moon, C. Holmgren, S.H. Mardikar, A.R. Dallosso, E. Klenova, D. Loukinov, R. Ohlsson, et al., A CTCF-binding silencer regulates the imprinted genes AWT1 and WT1-AS and exhibits sequential epigenetic defects during Wilms' tumourigenesis, Hum. Mol. Genet. 16 (3) (2007) 343–354, https://doi.org/10.1093/hmg/ddl478.
[43] M. Kaneuchi, M. Sasaki, Y. Tanaka, H. Shiina, H. Yamada, R. Yamamoto, N. Sakuragi, H. Enokida, M. Verma, R. Dahiya, WT1 and WT1-AS genes are inactivated by promoter methylation in ovarian clear cell adenocarcinoma, Cancer 104 (9) (2005) 1924–1930, https://doi.org/10.1002/cncr.21397.
[44] L. Lv, G. Chen, J. Zhou, J. Li, J. Gong, WT1-AS promotes cell apoptosis in hepatocellular carcinoma through down-regulating of WT1, J. Exp. Clin. Cancer Res. 34 (119) (2015), https://doi.org/10.1186/s13046-015-0233-7.
[45] T. Du, B. Zhang, S. Zhang, X. Jiang, P. Zheng, J. Li, M. Yan, Z. Zhu, B. Liu, Decreased expression of long non-coding RNA WT1-AS promotes cell proliferation and invasion in gastric cancer, Biochim. Biophys. Acta 1862 (1) (2016) 12–19, https://doi.org/10.1016/j.bbadis.2015.10.001.
[46] S. Tomita, M.O. Abdalla, S. Fujiwara, H. Matsumori, K. Maehara, Y. Ohkawa, H. Iwase, N. Saitoh, M. Nakao, A cluster of noncoding RNAs activates the ESR1 locus during breast cancer adaptation, Nat. Commun. 6 (2015) 6966, , https://doi.org/10.1038/ncomms7966.
[47] X. Lu, C. Huang, X. He, X. Liu, J. Ji, E. Zhang, W. Wang, R. Guo, A novel long non-coding RNA, SOX21-AS1, indicates a poor prognosis and promotes lung adenocarcinoma proliferation, Cell. Physiol. Biochem. 42 (5) (2017) 1857–1869, https://doi.org/10.1159/000479543.
[48] C. Wei, H. Wang, F. Xu, Z. Liu, R. Jiang, LncRNA SOX21-AS1 is associated with progression of hepatocellular carcinoma and predicts prognosis through epigenetically silencing p21, Biomed. Pharmacother. 104 (2018) 137–144, https://doi.org/10.1016/j.biopha.2018.05.010.
[49] A.W. Wei, L.F. Li, Long non-coding RNA SOX21-AS1 sponges miR-145 to promote the tumorigenesis of colorectal cancer by targeting MYO6, Biomed. Pharmacother. 96 (2017) 953–959, https://doi.org/10.1016/j.biopha.2017.11.145.
[50] L. Zhang, Y. Fang, X. Cheng, Y.J. Lian, H.L. Xu, Silencing of long noncoding RNA SOX21-AS1 relieves neuronal oxidative stress injury in mice with Alzheimer's disease by upregulating FZD3/5 via the Wnt signaling pathway, Mol. Neurobiol. (2018), https://doi.org/10.1007/s12035-018-1299-y.
[51] Y. Paul, S. Thomas, V. Patil, N. Kumar, B. Mondal, A.S. Hegde, A. Arivazhagan, V. Santosh, K. Mahalingam, K. Somasundaram, Genetic landscape of long non-coding RNA (lncRNAs) in glioblastoma: identification of complex lncRNA regulatory networks and clinically relevant lncRNAs in glioblastoma, Oncotarget 9 (51) (2018) 29548–29564, https://doi.org/10.18632/oncotarget.25434.
[52] M. Uchikawa, Y. Kamachi, H. Kondoh, Two distinct subgroups of group B sox genes for transcriptional activators and repressors: their expression during embryonic organogenesis of the chicken, Mech. Dev. 84 (1–2) (1999) 103–120, https://doi.org/10.1016/s0925-4773(99)00083-0.
[53] M. Ferletta, D. Caglayan, L. Mokvist, Y. Jiang, M. Kastemar, L. Uhrbom, B. Westermark, Forced expression of Sox21 inhibits Sox2 and induces apoptosis in human glioma cells, Int. J. Cancer 129 (1) (2011) 45–60, https://doi.org/10.1002/ijc.25647.
[54] X. Chang, J. Zhang, C. Huang, X. Pang, Q. Luo, H. Zhang, S. Zhang, Sex-determining region Y-related high mobility group box (SOX)-2 is overexpressed in cervical squamous cell carcinoma and contributes cervical cancer cell migration and invasion in vitro, Tumour Biol. 36 (10) (2015) 7725–7777 33 https://doi.org/10.1007/s13277-015-3450-y.