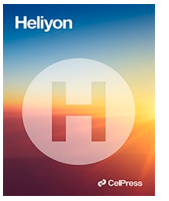
Heliyon 10 (2024) e32857

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Abnormal activation of genomic LINE1 elements caused by DNA demethylation contributes to lncRNA CASC9 overexpression in esophageal squamous cell carcinoma

Xuedan Chen [a,1], Juan Li [a,b,1], Xingying Guan [a], Yun Bai [a], Kai Wang [a,*]

[a] Department of Medical Genetics, Department of Basic Medicine, Army Medical University, Chongqing 400038, China
[b] Jinfeng Laboratory, Chongqing 400038, China

A B S T R A C T

Long noncoding RNA (lncRNA) cancer susceptibility 9 (CASC9) has been found to be overexpressed and functions as an oncogene in many cancer types. We investigated the molecular mechanism underlying CASC9 overexpression in esophageal squamous cell carcinoma (ESCC). Transcripts containing exons 2 and 6 and exons 4 and 6 showed the highest CASC9 expression levels in ESCC, no transcripts were detected in the normal esophageal epithelial Het1A cell line. The Long Interspersed Nuclear Element-1 (LINE1 or L1) element in the genome was found to participate in the evolution of lncRNA CASC9, the antisense promoter (ASP) of L1 provides the *cis*-regulatory elements necessary for CASC9 activation, and the antisense chain of L1 participates in the formation of exons of CASC9. The activation of the antisense promoter was due to the aberrant hypomethylation of L1 elements. An active enhancer element was identified in the downstream region of CASC9 gene by ChIP-seq and ChIP-qPCR. The interaction between ASP and the enhancer elements was confirmed by chromosome conformation capture (3C). Thus, our results suggest that the L1 ASP activation due to aberrant hypomethylation and downstream enhancer interaction plays a key role in the overexpression of lncRNA CASC9 in ESCC.

1. Introduction

Esophageal cancer is one of the most common cancers in the world and esophageal squamous cell carcinoma (ESCC) is the predominant type in China [1]. Despite the constant progress in ESCC treatment, only about 20 % of ESCC patients will survive for 5 years after diagnosis. It is very important to explore the molecular mechanism of ESCC tumorigenesis.

Recently, important roles of lncRNAs in tumorigenesis and cancer development have been elucidated. Many tumor-related lncRNA molecules have been identified including MALAT1, HOTAIR, H19, SOXOT, GAS5 and MEG3 [2–4]. They are abnormally expressed in lung, liver, colon, and breast cancer, as well as nasopharyngeal carcinoma and other tumor tissues, and are involved in the growth, apoptosis, proliferation and metastasis of many tumor cells [2,3,5]. Using a microarray assay, our group previously found that lncRNA CASC9 is overexpressed in esophageal squamous cell carcinoma, and CASC9 could promote ESCC growth by negatively regulating PDCD4 via EZH2 [6]. Overexpression of CASC9 correlates with the degree of malignancy and lymph node metastasis of ESCC and the pooled prognosis of ESCC patients [7]. Other groups have also demonstrated that CASC9 is overexpressed in liver cancer, gastric

Abbreviations: 3C, chromosome conformation capture; ASP, antisense promoter; BSP, Bisulfite sequencing PCR; CASC9, cancer susceptibility 9; ChIP, chromatin immunoprecipitation; ESCC, esophageal squamous cell carcinoma; Het-1A, human esophageal epithelial cell; LINE1, (Long Interspersed Nuclear Element-1, L1); rt-PCR, real-time quantitative reverse transcription polymerase chain reaction; STR, Short Tandem Repeat.

* Corresponding author.
E-mail address: kai.wang@tmmu.edu.cn (K. Wang).

[1] Xuedan Chen and Juan Li contributed equally to this study.

https://doi.org/10.1016/j.heliyon.2024.e32857
Received 12 September 2023; Received in revised form 22 May 2024; Accepted 11 June 2024
Available online 13 June 2024

cancer, nasopharyngeal carcinoma, lung adenocarcinoma, glioma, breast cancer and other tumors, and that CASC9 promotes tumor growth and metastasis and reduce the sensitivity to antitumor drugs [8–10]. Thus, lncRNA CASC9 is an important noncoding oncogene that is widely involved in the genesis and development of various tumors. Thus, CASC9 can be used as a potential tumor diagnostic and prognostic marker, making it important to study the mechanisms underlying its overexpression in cancer.

According to the RepeatMasker annotation in the University of California Santa Cruz (UCSC) Genome Browser, *cis*-regulatory elements for CASC9 transcription are located in two L1 elements [11]. *In silico* analyses revealed that L1 may participate in the regulation of CASC9. L1 elements are transposable reverse transcription elements with non-long terminal repeats (non-LTR). And it is the only active autonomous retrotransposable element in the human genome [12]. The full-length L1 element is approximately 6 Kb long and consists of a 5 ′untranslated region (UTR), two open reading frames (ORF1 and ORF2) and a 3′ untranslated region. A full-length L1 element has two promoters: a sense promoter that transcripts its two ORFs, and an ASP that is transcript in the opposite direction. The ASP can act as an alternative promoter for surrounding genes. Transposition of L1 rarely occurs under normal circumstances, but L1 may be activated in many tumors [13]. According to The Cancer Genome Atlas (TCGA) database, abnormal activation of L1 elements occurs in digestive tract tumors, especially in colon and esophageal cancers [14,15].

Until recently, many researches have been conducted on L1 elements in the human genome that can give rise to illegitimate transcripts. However, there are no reports on lncRNA transcription controlled by a promoter overlapping with the L1 ASP. In the present study, we found that the L1 element in the CASC9 region provides an active promoter that is hypomethylated during CASC9 transcription in ESCC. The promoter interacts with an enhancer downstream of CASC9. To our knowledge, this is the first study that to elucidate the relationship between the methylation status of L1 elements and the expression of alternative transcripts originating from L1 ASP in ESCC. In particular, our results suggest that L1 ASP acts as an active promoter, producing lncRNA CASC9, which is overexpressed in ESCC.

2. Materials and methods

2.1. Cell lines and tissue samples

Three esophageal squamous cancer cell lines (EC109, KYSE150, and KYSE450) were purchased from the Cell Bank of Chinese Academy of Science. Normal human esophageal epithelial cell line (Het-1A) was purchased from ATCC (American Type Culture Collection). The Short Tandem Repeat (STR) method was used to genotype all cells at the Laboratory of Birth Defects and Reproductive Health in Chongqing China.

EC109, KYSE150 and KYSE450 cells were cultured in RPMI-1640 medium (Hyclone, USA) supplemented with 10 % fetal bovine serum, 100U/ml penicillin and 0.1 mg/ml streptomycin. Het-1A cell was cultured in BEGM medium (Lonza/Clonetics Corporation). In order to investigate the methylation statue of L1 elements, the tissue samples were collected. Tumors and adjacent normal tissues were collected from patients who underwent surgery at the Thoracic and Cardiac Surgery of Southwest Hospital. The tissue samples was stored in a −80 °C refrigerator. This study was performed with the approval of the Ethics Committee of Army Medical University (Third Military Medical University), Chongqing, China and with the 1964 Helsinki Declaration.

2.2. Luciferase assay

L1 ASP and serial truncations in the CASC9 region were constructed in the pGL3-basic vector. The primers used for vector construction are listed in Supplementary Table S1. Approximately 1×10^5 cells were seed in 24-well plates. The next day, approximately 250 ng cloned reporter vector was transfected into cancer cells using Lipofectamine 3000 (Invitrogen). After 48 h, the cells were lysed using passive lysis buffer. The activity of firefly luciferase and Renilla were detected using Dual-Luciferase Reporter Assay System and the firefly luciferase activity was normalized to Renilla activity. The results are presented as fold-changes compared to the negative control.

2.3. RNA extraction and RT-qPCR

Total RNA was extracted using RNAiso reagent (Takara, China) according to the manufacturer' s instructions. Then RNA was reverse transcribed into complementary DNA (cDNA) for quantification using PrimerScriptTM RT reagent Kit with gDNA Eraser (Takara) according to the manufacturer's instructions. Quantitative PCR (qPCR) was performed using the SYBR Premix (Takara, China) on a Bio-Rad CFX96 system. The qPCR primers are listed in Supplementary Table S1. GAPDH was used as an internal control. The analysis was performed in triplicate.

2.4. Western blotting

Cells were harvested and lysised in RIPA lysis buffer (Beyotime, China) with freshly added PMSF. Total Proteins were separated by SDS-PAGE and electrotransferred to 0.22-μm PVDF membranes, after blocked with 5 % skimmed milk, membranes were incubated with primary antibodies overnight at 4 °C and then incubated with a horseradish peroxidase-conjugated secondary antibody. The antibodies used were anti-LINE-1 ORF1p (Abcam). The membrane was then incubated with ECL chromogenic substrate, and protein bands were visualized using the ChemiDoc Touch Imaging System (Bio-Rad, USA). GAPDH antibody (KangChen, China) was used as a control.

2.5. ChIP PCR and ChIP-Seq

Chromatin immunoprecipitation (ChIP) was performed using the EZ-Magna ChIP A/G Chromatin Immunoprecipitation Kit (Millipore, USA) according to the manufacturers' instructions. EC109, KYSE150 and Het1A cells were used for the ChIP analysis. Three replications of samples have been taken to perform ChIP-Seq. Approximately 2×10^6 cells were cross-linked with formaldehyde and then sonicated to obtain 200-500bp chromatin fragments. Fragmented chromatin was cleared and incubated with 2 mg of anti-H3K4me1 (Millipore, USA), anti-H3K4me3 (Millipore, USA) or anti-rabbit IgG (Millipore, USA) antibody at 4 °C overnight. The crosslinking was reversed, and the DNA was treated with Proteinase K and RNase A purified using the Qiaquick PCR-purification kit (Qiagen). Enrichment was tested using qPCR or sequencing. The qPCR primers are listed in Supplementary Table S1. For ChIP-Seq, the precipitated DNA was repaired by the PNK enzyme, and then ligated to adapters. PCR amplified fragments were sequenced using a Solexa 1G Genome Analyzer. First, we used FastQC to analyze the quality of the raw reads. Then the fastp software was used to trimming the raw reads. BWA (Burrows Wheeler Aligner) was used to align the reads to the reference genome accurately [16]. The ChIP-Seq and data analysis was performed by Novogene China. The data was then converted to UCSC Genome Browser Extensible Data files.

2.6. Chromosome conformation capture (3C)

To perform 3C, we used the TaqMan® 3C Chromosome Conformation Kits (Life technologies, USA) to analyze the three-dimensional organization of chromosomes in vivo in EC109, KYSE150, and Het1A cells as previously described [17]. Samples were crosslinked with 2 % formaldehyde for 10 min at room temperature. The reaction was quenched with glycine (0.125 M). Nuclei were resuspended using the buffer containing 0.3 % SDS and incubated for 1 h at 37 °C, with shaking. The crosslinked DNA was digested with the restriction enzyme *Dpn*II overnight. *Dpn*II was inactivated using 1.6 % SDS and incubating at 65 °C for 20 min. The DNA was ligated using T4 ligase at 16 °C for 4.5 h. Proteinase K was added to reverse the crosslinked DNA at 65 °C overnight. The samples were then incubated with RNaseA for 30 min at 37 °C, and the DNA was purified by phenol extraction. The PCR products spanning the enzyme cut sites of interest were gel-purified, and the DNA concentration was tested using Nanodrop (Thermo, USA). Next, qPCR was

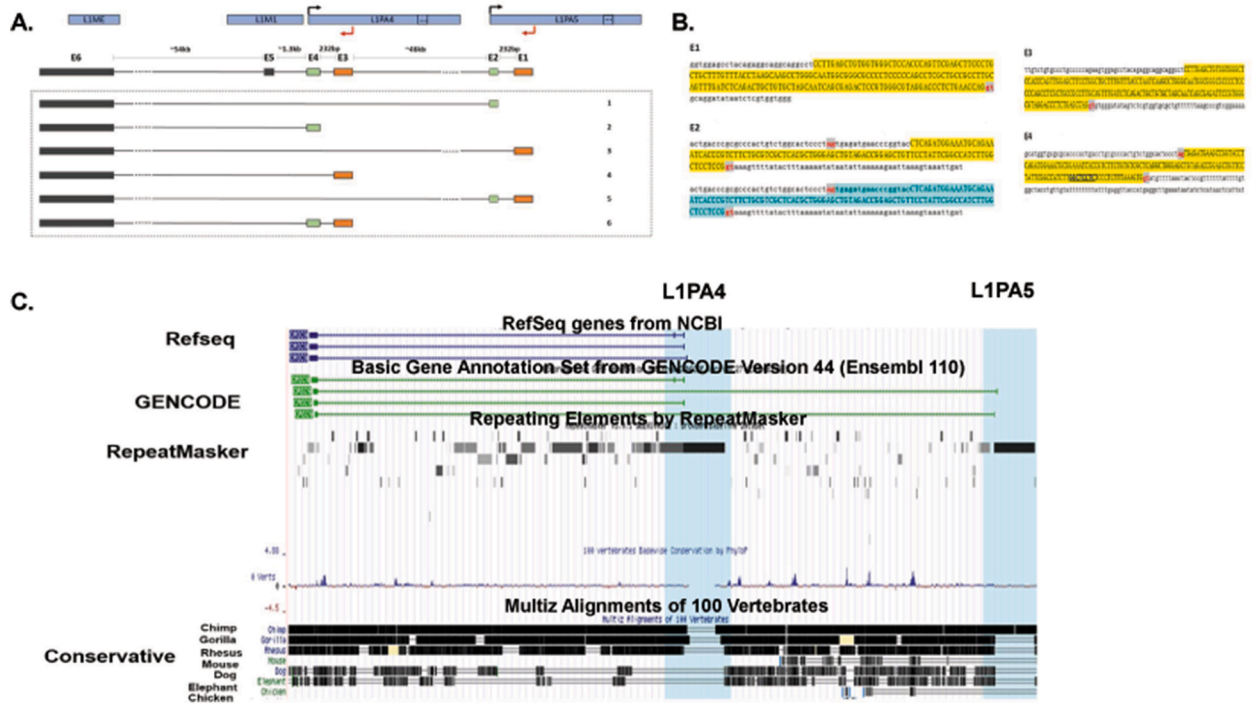


Fig. 1. Schematic diagram of lncRNA CASC9 gene structure. (A) L1 annotation information is displayed at the top. The black turning arrow represent the L1 promoter, and the red turning arrow represent the reverse promoter of L1 (ASP). The dashed box below is the six CASC9 transcripts identified by our group. E1, E2, E3, E4, E5, and E6 represent the corresponding exons. (B) We used BLAST program to compare nucleotide sequences which we identified in ESCC to sequence databases. The sequences of exon 1 and exon 3 are the same and the sequences of exon 2 and exon 4 are the same. The exon 2 which we identified is 113bp, which is 17 more bases long at the 5′ end than exon 2 (96bp) contained in the database. The annotated sequence in the database is marked by yellow. The sequences are identified by our group are marked by blue. Splice sites are in red. (C) Exon 6 is conserved in primates. Exon 1 and 2 are conserved in human and chimpanzee, while exon 3 and 4 only exist in human genome. Exon 1 and 2 of CASC9 are located in L1PA5, exon 3 and 4 are located in L1PA4, while exon 6 is located in L1 fragments, and the slightly larger one is located in L1ME. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

performed using the designed primers to quantify the frequency of crosslinking. All primers used for the 3C assay are listed in Supplementary Table S1.

2.7. Bisulfite sequencing PCR

Primers for BSP (bisulfite sequencing PCR) were designed using Methyl Primer Express Software. Bisulfite treatment was performed on 500 ng of genomic DNA. PCR products were purified by gel electrophoresis and inserted into PMD19-T vectors (Takara) by T-A cloning. The cloned vector was extracted and sequenced by Sangon Biotech Company, China. The primers are listed in Supplementary Table 1.

2.8. Pyrosequencing

To determine the methylation rate of the L1 ASP at each CpG site in the CASC9 gene region, a modified pyrosequencing procedure was performed. The PCR and sequencing primers were designed by using the PyroMark Assay Design software (QIAGEN). The primer sequences are listed in Supplementary Table S1. The L1 ASP in the CASC9 region was amplified, and the PCR products were biotinylated for downstream pyrosequencing because one of the PCR primers was biotin-labeled. The PCR products were pyrosequenced by the Sangon Biotech Company (China). Methylation levels at each CpG site were calculated.

2.9. Statistical analysis

All data were analyzed using the GraphPad software (version 7.0, GraphPad Software) and SPSS Statistics software (version 23.0, IBM). Differences between experimental groups and control groups were assessed by two-tailed Student' s *t*-test or one-way ANOVA as appropriate. All data are shown as the mean ± standard error of three independent experiments. A p-value <0.05 was considered to be statistically significant.

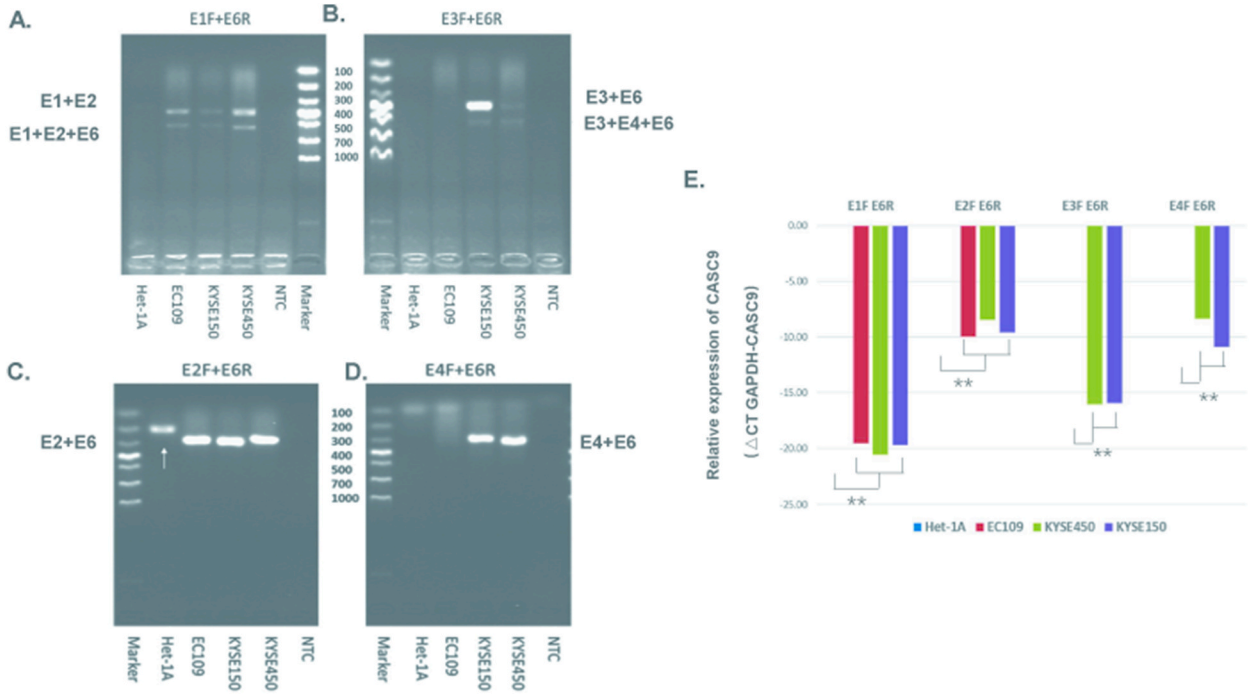


Fig. 2. The transcript forms of CASC9 in normal esophageal epithelium Het-1A and esophageal squamous cell carcinoma cells (EC109, KYSE450, KYSE150). (A) The forward and reverse RT-PCR primers were located in exon 1 and exon 6 respectively. The entire image of the gel is shown in Supplementary Figure S4 (B) The forward and reverse RT-PCR primers were located in exon 3 and exon 6 respectively. The entire image of the gel is shown in Supplementary Figure S4 (C) The forward and reverse RT-PCR primers were located in exon 2 and exon 6 respectively. Arrow indicates nonspecific amplification. The entire image of the gel is shown in Supplementary Figure S4 (D) The forward and reverse RT-PCR primers were located in exon 4 and exon 6 respectively. The entire image of the gel is shown in Supplementary Figure S4 (E) Primer E1F and E6R could detected the total amount of transcripts contain E1, E6 and E1, E2, E6. Primer E2F and E6R could detected the transcript contains E2, E6. Primer E3F and E6R could detect the total amount of transcripts contain E3, E6 and E3, E4, E6. Primer E4F and E6R could detected the transcript contains E4, E6. According to the dissolution curve, there is no any transcripts in Het1A. NTC: Non-Template Control.

3. Results

3.1. The structure and alternative splicing of CASC9 in ESCC

According to the UCSC database, CASC9 is located in the gene desert of 8q21, and there are no known protein-coding genes within 100 kb upstream and downstream of this gene. There are many long interspersed nuclear elements in this area, including two almost complete L1 elements with a length of approximately 6 kb. Exons 1 and 2 and exons 3 and 4 are respectively located in the 5′-untranslated region (5′ UTR) of two L1 elements. The two ASPs are located upstream of exons 2 and 4 of CASC9 (Fig. 1A). Therefore, the ASP of L1 may regulate the transcription of CASC9. We identified six CASC9 transcripts using RT-PCR and clone sequencing. Then we used the Basic Local Alignment Search Tool (BLAST) program to compare the nucleotide sequences identified in ESCC with sequence databases [18]. Transcripts containing exons 1, 2, and 6 and those containing exons 3, 4, and 6 were identified by our research group for the first time and were not previously included in the databases. Very high similarity (85 %) was found between exons 2 and 4. The exon 2 sequence we identified is 113 bp and is 17 bases longer at the 5′ end than exon 2 (96bp) contained in the databases, and the rest of the sequences are identical. Moreover, we identified exon-intron splicing sites located in the intron region, while non-classical splice sites were identified in the introns for the 96 bp variant of exon 2. Therefore, we indicated that the 5′ end of exon 2 containing the 96 bp sequence annotated in the database is not complete, and the 113 bp exon 2 and the 125 bp exon 4 sequences are the complete exons of CASC9 (Fig. 1B).

According to evolutionary conservation analysis, we found that the exons of CASC9 are conserved to different degrees. Exon 6 is conserved among primates, exons 1 and 2 are conserved in humans and chimpanzees, and exons 3 and 4 are conserved only in humans. Exons 1 and 2 of CASC9 are located in L1PA5, which is a subfamily of L1 elements, exons 3 and 4 are located in L1PA4, and exon 6 is located in L1 fragments; the slightly longer exon 6 sequence is located in L1ME. Therefore, CASC9 sequence conservation was consistent with the evolution of the L1 elements over time according to location (Fig. 1C). Therefore, we believe that the L1 element in the genome participated in the evolution of lncRNA CASC9, that the ASPs of L1 may provide the *cis*-regulatory elements necessary for CASC9 transcription, and that the antisense chain of L1 participates in the formation of the exons of CASC9.

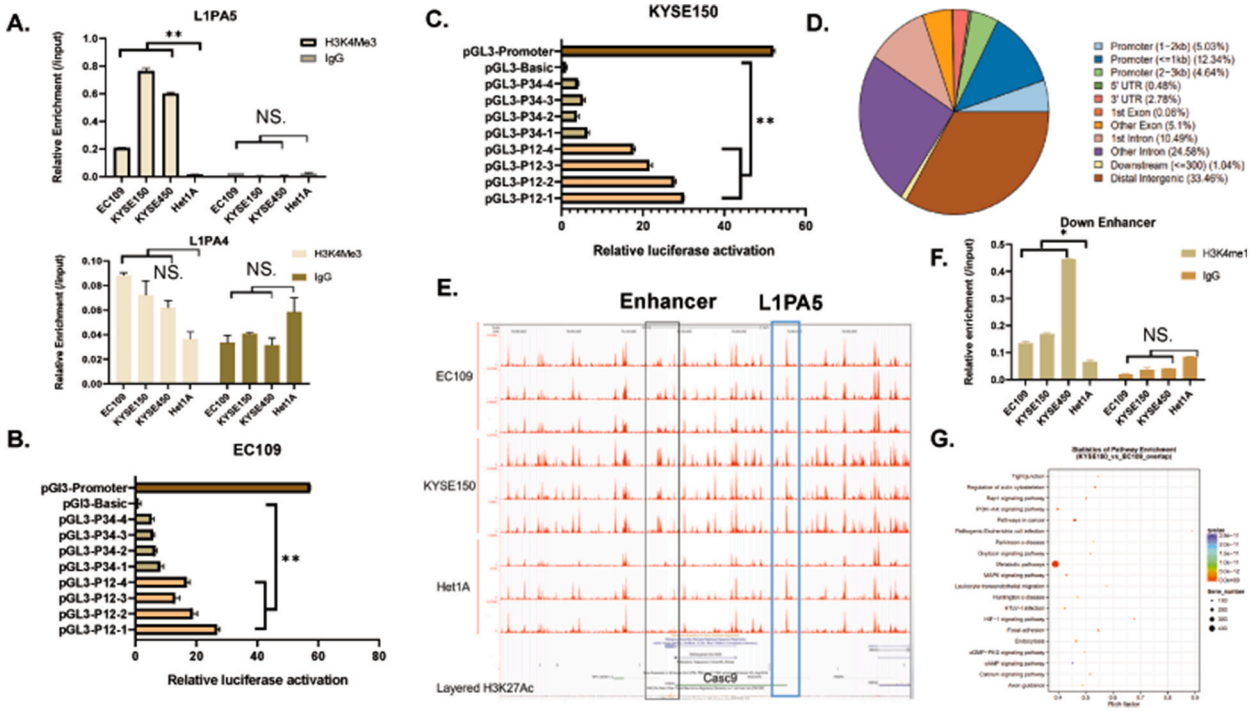


Fig. 3. Identification basal promoter and enhancer elements of CASC9 (A) ChIP-qPCR assays were performed in ESCC and Het1A cells using H3K4me3-specific or IgG antibody. Enrichment of L1 ASPs of CASC9 was measured by qPCR. Data shown are the mean fold increase ±SD from two independent experiments. ** means $p < 0.01$. (B–C) Luciferase reporter gene expression assays with the L1 ASP region of CASC9 and three truncations. The full length of L1 ASP regions of CASC9 (pGL3-P12-1 or pGL3-P34-1) and truncations (pGL3-P12-2, pGL3-P12-3, pGL3-P12-4, pGL3-P34-1, pGL3-P34-2, pGL3-P34-3, pGL3-P34-4) was cotransfected into EC109 (B) or KYSE150 (C) cell lines. (D) Distribution of H3K4me enriched peaks in ESCC specific cells. (E) The genome browser view showed H3K4me1 signals in EC109, KYSE150 and Het1A cell lines. The black boxes highlighted the downstream of CASC9 gene. (F) ChIP assays were performed in ESCC and Het1A cells using H3K4me-specific or IgG antibody. Enrichment of candidate enhancers of CASC9 was measured by qPCR. (G) Signaling pathways of ESCC cell lines specific enrichment reads analyzed by KEGG.

3.2. Transcripts and expression of CASC9 isoforms in ESCC

We used RT-PCR and gel electrophoresis to detect the transcript forms of CASC9 in normal esophageal epithelium Het-1A cells and esophageal squamous cell carcinoma cells (EC109, KYSE450, and KYSE150). EC109 cells had transcripts containing exons 1, 2, and 6, whereas KYSE150 and KYSE450 had transcripts containing exons 1, 2, 3, 4, and 6. There were no transcripts in the normal esophageal epithelium Het-1a cells (Fig. 2A–D). RT-qPCR was performed to determine the relative expression levels of CASC9 in ESCC. The results indicated that two kinds of transcripts were highly expressed in the ESCC cell lines, one wasa transcript consisted of exons 2, 6 detected by primer E2F and E6R, another was a transcript consisted of exons 4, and 6 detected by primer E4F and E6R Meanwhile, no transcripts were detected in the Het1A cells (Fig. 2E). So, these results indicated that ESCC cell lines exhibited significantly higher CASC9 level compared with normal cell line Het-1A.

3.3. Identification of basal promoter and enhancer elements in CASC9

H3K4me3 is a well-known epigenetic marker that occurs in the promoter regions of genes involved in transcriptional activation and plays an important role in regulating gene expression [19]. Therefore, a chromatin immunoprecipitation assay was performed to detect the enrichment of H3K4me3 in these cell lines. ChIP-qPCR results showed enrichment of H3K4me3 on the ASPs of L1PA5 in the EC109, KYSE150, and KYSE450 cell lines, whereas there was no significant enrichment thereof in the L1PA4 ASP region (Fig. 3A). Moreover, there was no significant enrichment of H3K4Me3 in Het1A cells (Fig. 3A). These results indicated that the overexpression of CASC9 in cancer cells is probably due to the increased activity of L1 ASP.

To further evaluate basal, proximal promoter activity of the L1 ASP, a series of 5′ deletions inserted to the pGL3-basic luciferase vector were constructed and transfected into EC109 and KYSE150 cells (Fig. S1). We found that luciferase activity significantly increased in the presence of pGL3-P12, which contains L1PA5, suggesting that there are promoter activities in this L1 ASP. A series of deletions resulted in a slight to moderate reduction in promoter activity; however, the activity remained greater than that observed with pGL3-basic (Fig. 3B–C). These findings suggest that the L1 ASPs in L1PA5 may function as basal promoters critical for CASC9 expression.

Enhancers regulate cell-type-specific gene expression by facilitating the transcription of target genes. Active enhancers are usually marked by H3K4me1 in a cell type-specific manner in mammalian cells [20]. Therefore, we performed ChIP–seq in ESCC (EC109 and KYSE150) and Het1A cells to determine the localization of candidate H3K4me1-binding sites. We compared the H3K4me1 enriched regions between ESCC and Het1A cell lines, and the coefficient of correlation is shown in Fig. S2. In addition, there were 21,233 H3K4me1 islands that overlapped in the EC109 and KYSE150 cell types (Fig. S2B). In addition to these overlapping islands, we also found that 7719 H3K4me1 islands were ESCC cell type-specific, and approximately 22 % of these were associated with promoters (Figs. S2B and 3D).

To map the candidate *cis*-regulatory elements on a genome-wide scale, we identified the genomic regions of H3K4me ChIP-seq reads downstream of the CASC9 gene regions in EC109 and KYSE150 cells, but not in Het1A cells (Fig. 3E). Similar results were

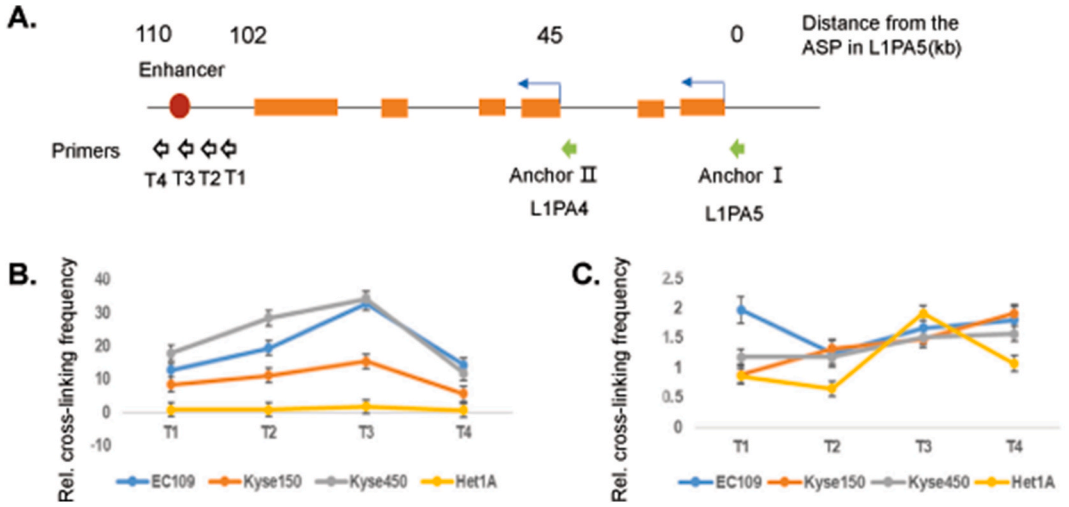


Fig. 4. Long-range interactions with endodermal enhancer at the Casc9 locus in ESCC and Het1A cell lines. A. Schematic representation of the human CASC9 locus. The exons of CASC9 (black boxes) are displayed together with the downstream endodermal enhancer (black oval). The ASPs of L1 elements are indicated by arrows, *Dpn*II restriction sites investigated in the 3C assays showed by white arrows (T1 to T4). The primers used as the anchors (close to the ASP of L1PA5 and L1PA4) is represented by black arrows (Anchor I and Anchor II). Three independent 3C-qPCR experiments were performed on the same biological samples. B. Relative crosslinking frequencies between fixed *Dpn*II fragment anchor I (ASP in L1PA5) and the genomic sites (sites T1, T2, T3 and T4) in the enhancer region. C. Relative crosslinking frequencies between fixed *Dpn*II fragment Anchor II (ASP in L1PA4) and the genomic sites (sites T1, T2, T3 and T4) in the enhancer region. Standard error of the mean is indicated.

obtained for H3K4me1 occupation using ChIP–qPCR (Fig. 3F). Furthermore, using Encyclopedia of DNA Elements (ENCODE), we also detected H3K27ac enrichment in this region, which has previously been shown to be associated with active promoters and enhancers [21].

Pathway analysis demonstrated that specific genes were enriched in ESCC cell lines and were primarily associated with proteoglycans, the PI3K−Akt signaling pathway, and the Rap1 signaling pathway in cancer (Fig. 3G). We previously found that the PI3K−Akt and Rap1 signaling pathways involve many well-known genes that function in cell motility and cancer metastasis in ESCC. CASC9 depletion could reduce the phosphorylation of FAK, PI3K, and Akt, which are downstream effectors of the integrin pathway [22]. In addition, RAP1 could promote esophageal cancer cell migration and invasion through MMP2. These results suggest that CASC9 overexpression in ESCC may be due to the H3K4me1 enrichment of ESCC cell type-specific gene regions.

3.4. Spatial interaction and looping between the enhancer and the ASPs in ESCC

In many cases, active enhancers appear to engage in direct physical interactions with nearby promoters and play an important role in the regulation of gene expression. Therefore, we performed 3C analysis on these cell lines to relate the spatial conformation of the CASC9 locus to its transcriptional status. 3C-qPCR assays were performed on EC109, KYSE150, KYSE450, and Het1A cell lines, which displayed differential CASC9 gene expression. Locus-wide crosslinking frequencies were determined between *Dpn*II-fixed fragments (Anchor I or II) and other *Dpn*II sites of the endodermal enhancer (Fig. 4A).

Fixed fragment anchor I, a small fragment containing the ASP of L1PA5, specifically interacted with four genomic sites (T1, T2, T3, and T4) in the ESCC cell lines (Fig. 4B). Interestingly, these interactions were lost in Het1A cells (Fig. 4B). However, no significant interaction peaks were observed between fragment anchor II and the *Dpn*II sites of the endodermal enhancer in any of the cell lines (Fig. 4C). Hence, we concluded that CASC9 overexpression in ESCC could result from proximal promoters by forming an enhancer-promoter interaction loop. In contrast, there were not any enhancer-promoter interaction loops in the normal esophageal epithelial cell line, Het1A.

3.5. The CpG island methylation status of CASC9 in ESCC and Het1A cell lines

The status of L1 promoter methylation may play important roles in distinguishing between expressed and unexpressed L1 loci on a broad scale. BSP and pyrosequencing were used to investigate the methylation status of the two L1 ASPs in the CASC9 gene region. The

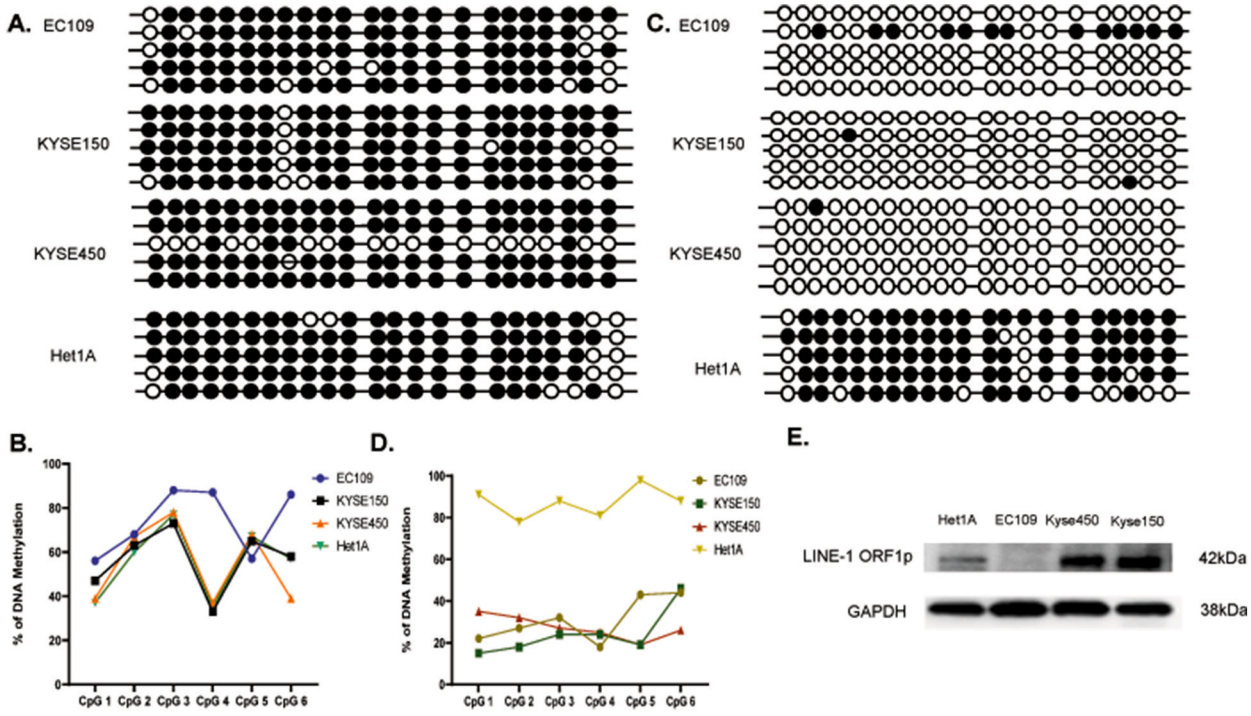


Fig. 5. Methylation levels of the CpG sites typed by BSP and pyrosequencing data. Representative direct sequencing of the BSP image and dot plots of methylation status of CpG sites in ESCC and Het1A cell lines in L1PA4 ASP region (A) or L1PA5 ASP region (C). White and black dots indicate unmethylated and methylated CpGs, respectively. B. Methylation levels of L1PA4 in different cells detected by pyrosequencing assay. D. Methylation Levels of L1PA5 in different cells detected by pyrosequencing assay. E. Protein expression of LINE-1 ORF1 in ESCC and Het1A cell lines detected by Western blotting. The entire image of western blots is shown in Supplementary Fig. S5.

sequence details of the BSP regions in the L1PA5 and L1PA4 ASPs are shown in Fig. S3.

The BSP results showed that the methylation ratios of L1PA4 ASP in EC109, KYSE150, KYSE450, and Het1A cells were 87.8 %, 90.4 %, 84.3 %, and 87.8 %, respectively ($p > 0.05$) (Fig. 5A). This was confirmed by pyrosequencing of the ESCC and Het1A cell lines ($p > 0.05$), and hypermethylation of the L1PA4 ASP region was observed in the ESCC and Het1A cell lines (Fig. 5B). The BSP results showed that the methylation ratio of L1PA5 ASP was 83.8 % in the normal esophageal epithelial cell line Het1A, which was higher than that in EC109 (12.4 %), KYSE150 (1.9 %), and KYSE450 (0.95 %) cell lines (Fig. 5C). This was confirmed by the pyrosequencing of ESCC and Het1A cell lines (Fig. 5D). Then BSP was used to detect the methylation ratios of L1PA5 ASP in 5 ESCC and paired normal specimens. As expected, the methylation ratio of L1PA5 ASP was higher in 80 % (4/5) non-tumor tissues (data not shown). We then used western blotting to detect LINE-1 retrotransposable element ORF1 protein expression in these cell lines. LINE1 ORF1 protein expression was upregulated in KYSE150 and KYSE450 cells compared to that in Het1A cells (Fig. 5E). These results indicate that CpG island methylation in L1PA5 may be involved in the activation of L1, resulting in the overexpression of CASC9 in ESCC.

4. Discussion

lncRNA CASC9 has been identified as a novel lncRNA in ESCC. In our previous study, we found that CASC9 was overexpressed in ESCC and promoted the cell growth of ESCC. And we demonstrate that lncRNA CASC9 functions as an oncogene by negatively regulating PDCD4 expression through recruiting EZH2 and subsequently altering H3K27me3 level in ESCC [6]. High CASC9 expression was associated with lymph node metastasis and advanced TNM staging. We found that CASC9 upregulates LAMC2 via CBP-mediated histone acetylation, and functions as a pro-metastatic gene through the FAK-PI3K/Akt signaling pathways [7]. Other studies have shown that CASC9 is overexpressed in many cancer types and promotes cell growth, differentiation, invasion, and metastasis [23]. However, few studies have investigated the mechanisms by which CASC9 is overexpressed in tumors. Cai found that NF-kB could active CASC9 and promote bladder cancer cell growth and metastasis [24]. In this study, we identified six CASC9 transcripts in ESCC. The transcripts containing exons 2 and 6 showed the highest expression in ESCC. There were no transcripts in normal esophageal epithelial Het1A cells. Interestingly, we found that the L1 element in the genome participated in the generation of the lncRNA CASC9. The ASP of L1 provides the *cis*-regulatory elements necessary for CASC9 transcription, and the antisense chain of L1 participates in the formation of CASC9 exons.

L1 elements have attracted broad research attention in recent years, because some L1s are still active in the human genome [12]. A full-length L1 element contains a 5′ UTR with an internal promoter for transcription, an ORF1 that encodes a protein for RNA binding, an ORF2 that encodes a protein for reverse transcription, a short 3′ UTR, and a poly(A) tail. In the 5' UTR there is not only a sense strand promoter for its own transcription but also an ASP. Speek was the first to demonstrate the activity of L1 ASP and discovered four chimeric L1 ASP transcripts [25], suggesting that the human genome contains many active sites directed by L1 ASP. They also found that active L1 ASP can drive transcription of chimeric transcripts in specific cell types in a few instances, including cancer [26]. Cruickshanks et al. found 18 new L1 ASP transcripts, some of which were only identified in breast and colon cancer tissues but not in the corresponding normal tissues [27]. They also found that DNA methylation in normal tissues represses the activity of L1 ASP, and 5-aza-cytidine, a nucleoside analog of cytidine that specifically inhibits DNA methylation, upregulates the expression of L1 ASP transcripts in tumor cell lines [28]. Thus, some human genes are transcribed from ASP of L1 in some instances [29]. In addition to mRNA, L1 can also reveal long noncoding RNAs transcription in stem cells [30]. Takafumi found that the L1 ASP could provide promoters for lncRNA transcription and contribute to the tissue-specific transcription of lncRNAs. This indicated that L1 can be reused as a functional domain or *cis*-element for lncRNA transcription [31]. In this study, we performed a luciferase reporter and H3K4me3 ChIP assays and found that there were promoter activities in L1 ASPs upstream of the CASC9 gene in ESCC cell lines. Therefore, we believe that the ASP of L1 provides the *cis*-regulatory elements necessary for CASC9 transcription and that the antisense chain of L1 participates in the formation of exons of CASC9.

Approximately 926,535 enhancers exist in the human genome, as estimated by the Encyclopedia of DNA Elements (ENCODE) project [32], indicating that human genes are regulated by approximately 40–50 enhancers. In the present study, we identified an enhancer downstream of CASC9. We used the 3C assay to demonstrate the direct interaction between L1 ASP and the enhancer downstream of CASC9, which we identified earlier in this study. Hence, the overexpression of CASC9 in ESCC could result from the proximal promoters by forming an enhancer-promoter interaction loop. However, there are no enhancer-promoter interaction loops in the normal esophageal epithelial cell line, Het1A. Enhancers function as transcription factor-binding platforms that increase transcription by directly stimulating target promoters. Using the ENCODE database, we identified many transcription factors, such as YY1, SMAD4, and TCF4, in the enhancer regions in stem and cancer cells. These transcription factors are involved in the development and progression of ESCC [33–35]. The interactions between the enhancers and these transcription factors need to be confirmed in further studies.

L1 ASPs are usually inactivated because the CpG islands are heavily methylated in healthy individuals. Hypomethylation of the L1 promoter has been reported in many cancers including hepatocellular cancer [36], prostate cancer [37], breast cancer [38], esophageal adenocarcinoma [39], and colon cancer [40]. Baba et al. found that L1 hypomethylation might cause genomic instability and result in genomic gains at various loci including CDK6 [41]. In this study, we found that the L1 ASP in the CASC9 gene region was activated by hypomethylation, and that the activated ASP provided a promoter for lncRNA CASC9 overexpression in ESCC. These results indicate that L1 hypomethylation may be a cause of high CASC9 expression in ESCC. Therefore, it has the potential to be used as a diagnostic and prognostic biomarker in cancer [15,41,42]. The hypomethylation status of L1 in cell free DNA (cfDNA) can also be used as a tumor surrogate marker [43,44].

5. Conclusion

Our data strongly suggest that aberrant promoter hypomethylation in the L1 elements could active the antisense promoter of L1 locate in the CASC9 gene region. And the active ASP of L1 could interact with an enhancer which was identified in this study result in the overexpression of CASC9 in ESCC. Furthermore, the hypomethylation status of L1 and the expression of CASC9 have the potential to serve as novel biomarkers for the diagnosis and treatment of ESCC.

Ethics statement

This study was performed with the approval of the Ethics Committee of Army Medical University (Third Military Medical University), Chongqing, China and with the 1964 Helsinki Declaration. The approval number is (B) KY2021165.

Data availability

All data generated or analyzed during this study are included in this published article.

CRediT authorship contribution statement

Xuedan Chen: Methodology, Data curation. **Juan Li:** Methodology, Investigation, Funding acquisition, Data curation. **Xingying Guan:** Methodology, Investigation. **Yun Bai:** Conceptualization. **Kai Wang:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation.

Declaration of competing interest

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

Acknowledgements and funding

This work was supported by the Natural Science Foundation Project of CQ CSTC [cstc2020jcyj-msxmX0909].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e32857.

References

[1] W.Q. Chen, R.S. Zheng, P.D. Baade, S.W. Zhang, H.M. Zeng, F. Bray, et al., Cancer Statistics in China, 2015, Ca-Cancer J Clin 66 (2016) 115–132.
[2] Y. Wu, X. Chen, Y. Liang, J. Li, K. Zhang, L. Dai, et al., Overexpression of long non-coding RNA SOX2OT promotes esophageal squamous cell carcinoma growth, Cancer Cell Int. 18 (2018) 76.
[3] K. Wang, J. Li, G. Xiong, G. He, X. Guan, K. Yang, et al., Negative regulation of lncRNA GAS5 by miR-196a inhibits esophageal squamous cell carcinoma growth, Biochem. Biophys. Res. Commun. 495 (2018) 1151–1157.
[4] J. Li, H. Meng, Y. Bai, K. Wang, Regulation of lncRNA and its role in cancer metastasis, Oncol. Res. 23 (2016) 205–217.
[5] E. Anastasiadou, L.S. Jacob, F.J. Slack, Non-coding RNA networks in cancer, Nat. Rev. Cancer 18 (2018) 5–18.
[6] Y. Wu, L. Hu, Y. Liang, J. Li, K. Wang, X. Chen, et al., Up-regulation of lncRNA CASC9 promotes esophageal squamous cell carcinoma growth by negatively regulating PDCD4 expression through EZH2, Mol. Cancer 16 (2017) 150.
[7] Y. Liang, X. Chen, Y. Wu, J. Li, S. Zhang, K. Wang, et al., LncRNA CASC9 promotes esophageal squamous cell carcinoma metastasis through upregulating LAMC2 expression by interacting with the CREB-binding protein, Cell Death Differ 25 (11) (2018) 1980–1995.
[8] Z. Bing, J. Han, Z. Zheng, N. Liang, FOXO3-induced oncogenic lncRNA CASC9 enhances gefitinib resistance of non-small-cell lung cancer through feedback loop, Life Sci. 287 (2021) 120012.
[9] Z. Chen, Q. Chen, Z. Cheng, J. Gu, W. Feng, T. Lei, et al., Long non-coding RNA CASC9 promotes gefitinib resistance in NSCLC by epigenetic repression of DUSP1, Cell Death Dis. 11 (2020) 858.
[10] J.H. Noh, M. Gorospe, AKTions by Cytoplasmic lncRNA CASC9 promote hepatocellular carcinoma survival, Hepatology 68 (2018) 1675–1677.
[11] J.D. Fernandes, A. Zamudio-Hurtado, H. Clawson, W.J. Kent, D. Haussler, S.R. Salama, et al., The UCSC Repeat Browser Allows Discovery and Visualization of Evolutionary Conflict across Repeat Families, 11, Mobile DNA-Uk, 2020.
[12] Pascarella G., Hon C.C., Hashimoto K., Busch A., Luginbuhl J., Parr C., et al., Recombination of repeat elements generates somatic complexity in human genomes, Cell 2022 08 4;185 (16), 3025-3040.e6.
[13] U. Miglio, E. Berrino, M. Panero, G. Ferrero, L. Coscujuela Tarrero, V. Miano, et al., The expression of LINE1-MET chimeric transcript identifies a subgroup of aggressive breast cancers, Int. J. Cancer 143 (2018) 2838–2848.
[14] K. Yamaguchi, A.O. Soares, L.A. Goff, A. Talasila, J.A. Choi, D. Ivenitsky, et al., Striking heterogeneity of somatic L1 retrotransposition in single normal and cancerous gastrointestinal cells, Proc Natl Acad Sci U S A. 117 (2020) 32215–32222.
[15] S. Hoshimoto, H. Takeuchi, S. Ono, M.S. Sim, J.L. Huynh, S.K. Huang, et al., Genome-wide hypomethylation and specific tumor-related gene hypermethylation are associated with esophageal squamous cell carcinoma outcome, J. Thorac. Oncol. 10 (2015) 509–517.
[16] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, et al., The sequence Alignment/map format and SAMtools, Bioinformatics 25 (2009) 2078–2079.

[17] H. Hagege, P. Klous, C. Braem, E. Splinter, J. Dekker, G. Cathala, et al., Quantitative analysis of chromosome conformation capture assays (3C-qPCR), Nat. Protoc. 2 (2007) 1722–1733.
[18] S. Lee, S. Oh, K. Kang, K. Han, BLAT-based Comparative Analysis for Transposable Elements: BLATCAT, 2014, Biomed Research International, 2014.
[19] J. Wysocka, T. Swigut, H. Xiao, T.A. Milne, S.Y. Kwon, J. Landry, et al., A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling, Nature 442 (2006) 86–90.
[20] A. Rada-Iglesias, R. Bajpai, T. Swigut, S.A. Brugmann, R.A. Flynn, J. Wysocka, A unique chromatin signature uncovers early developmental enhancers in humans, Nature 470 (2011), 279-+.
[21] N.D. Heintzman, R.K. Stuart, G. Hon, Y. Fu, C.W. Ching, R.D. Hawkins, et al., Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome, Nat. Genet. 39 (2007) 311–318.
[22] Y. Liang, X. Chen, Y. Wu, J. Li, S. Zhang, K. Wang, et al., LncRNA CASC9 promotes esophageal squamous cell carcinoma metastasis through upregulating LAMC2 expression by interacting with the CREB-binding protein, Cell Death Differ. 25 (2018) 1980–1995.
[23] Y. Qi, C. Song, J. Zhang, C. Guo, C. Yuan, Oncogenic LncRNA CASC9 in cancer progression, Curr. Pharmaceut. Des. 27 (2021) 575–582.
[24] C. Longjun, Z. Jianjun, P. Kun, H. Lin, S. Zhenduo, D. Yang, et al., NF-kappaB-Activated lncRNACASC9 promotes bladder cancer progression by regulating the TK1 expression, JAMA Oncol. 2022 (2022) 9905776.
[25] M. Speek, Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes, Mol. Cell Biol. 21 (2001) 1973–1985.
[26] K. Matlik, K. Redik, M. Speek, L1 antisense promoter drives tissue-specific transcription of human genes, J. Biomed. Biotechnol. 2006 (2006) 71753.
[27] H.A. Cruickshanks, N. Vafadar-Isfahani, D.S. Dunican, A. Lee, D. Sproul, J.N. Lund, et al., Expression of a large LINE-1-driven antisense RNA is linked to epigenetic silencing of the metastasis suppressor gene TFPI-2 in cancer, Nucleic Acids Res. 41 (2013) 6857–6869.
[28] H.A. Cruickshanks, C. Tufarelli, Isolation of cancer-specific chimeric transcripts induced by hypomethylation of the LINE-1 antisense promoter, Genomics 94 (2009) 397–406.
[29] P. Nigumann, K. Redik, K. Matlik, M. Speek, Many human genes are transcribed from the antisense promoter of L1 retrotransposon, Genomics 79 (2002) 628–634.
[30] D. Kelley, J. Rinn, Transposable elements reveal a stem cell-specific class of long noncoding RNAs, Genome Biol. 13 (2012) R107.
[31] T. Chishima, J. Iwakiri, M. Hamada, Identification of transposable elements contributing to tissue-specific expression of long non-coding RNAs, Genes 9 (2018).
[32] E.P. Consortium, J.E. Moore, M.J. Purcaro, H.E. Pratt, C.B. Epstein, N. Shoresh, et al., Expanded encyclopaedias of DNA elements in the human and mouse genomes, Nature 583 (2020) 699–710.
[33] J. Li, X. Han, Y. Gu, J. Wu, J. Song, Z. Shi, et al., LncRNA MTX2-6 suppresses cell proliferation by acting as ceRNA of miR-574-5p to accumulate SMAD4 in esophageal squamous cell carcinoma, Front. Cell Dev. Biol. 9 (2021) 654746.
[34] Y. Tang, Y. Chen, R. Liu, W. Li, B. Hua, Y. Bao, Wnt signaling pathways: a role in pain processing, NeuroMolecular Med. 24 (2022) 233–249.
[35] Q. Wu, P. Liu, L. Wang, Many facades of CTCF unified by its coding for three-dimensional genome architecture, J Genet Genomics 47 (2020) 407–424.
[36] C.H. Lin, S.Y. Hsieh, I.S. Sheen, W.C. Lee, T.C. Chen, W.C. Shyu, et al., Genome-wide hypomethylation in hepatocellular carcinogenesis, Cancer Res. 61 (2001) 4238–4243.
[37] W.A. Schulz, J.P. Elo, A.R. Florl, S. Pennanen, S. Santourlidis, R. Engers, et al., Genomewide DNA hypomethylation is associated with alterations on chromosome 8 in prostate carcinoma, Genes Chromosomes Cancer 35 (2002) 58–65.
[38] P.S. Yan, C. Venkataramu, A. Ibrahim, J.C. Liu, R.Z. Shen, N.M. Diaz, et al., Mapping geographic zones of cancer risk with epigenetic biomarkers in normal breast tissue, Clin. Cancer Res. 12 (2006) 6626–6636.
[39] T.T. Doucet-O'Hare, N. Rodic, R. Sharma, I. Darbari, G. Abril, J.A. Choi, et al., LINE-1 expression and retrotransposition in Barrett's esophagus and esophageal carcinoma, Proc Natl Acad Sci U S A. 112 (2015) E4894–E4900.
[40] C.M. Suter, D.I. Martin, R.L. Ward, Hypomethylation of L1 retrotransposons in colorectal cancer and adjacent normal tissue, Int. J. Colorectal Dis. 19 (2004) 95–101.
[41] Y. Baba, M. Watanabe, A. Murata, H. Shigaki, K. Miyake, T. Ishimoto, et al., LINE-1 hypomethylation, DNA copy number alterations, and CDK6 amplification in esophageal squamous cell carcinoma, Clin. Cancer Res. 20 (2014) 1114–1124.
[42] S. Iwagami, Y. Baba, M. Watanabe, H. Shigaki, K. Miyake, T. Ishimoto, et al., LINE-1 hypomethylation is associated with a poor prognosis among patients with curatively resected esophageal squamous cell carcinoma, Ann. Surg. 257 (2013) 449–455.
[43] E. Boldrin, M. Curtarello, M. Dallan, R. Alfieri, S. Realdon, M. Fassan, et al., Detection of LINE-1 hypomethylation in cfDNA of esophageal adenocarcinoma patients, Int. J. Mol. Sci. 21 (2020).
[44] M.S. Taylor, C. Wu, P.C. Fridy, Y. Senussi, W.C. Cheng, J. Heaps, et al., Ultrasensitive detection of circulating LINE-1 ORF1p as a specific multi-cancer biomarker, Cancer Discov 13 (12) (2023) 2532–2547.