Long noncoding RNA LINC01186, regulated by TGF-β/SMAD3, inhibits migration and invasion through Epithelial-Mesenchymal-Transition in lung cancer

Yajing Hao, Xinling Yang, Dongdong Zhang, Jianjun Luo, Runsheng Chen

Key Laboratory of RNA Biology, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China
Beijing Key Laboratory of Noncoding RNA, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China
University of Chinese Academy of Sciences, Beijing 100049, China
College of Lab Medicine, Hebei North University, Zhangjiakou 075000, China

Abstract

Accumulating evidence suggests that long noncoding RNAs (lncRNAs) are crucial regulators of the Epithelial-Mesenchymal-Transition (EMT). TGF-β signaling is a major inducer of EMT and can facilitate lung cancer metastasis. However, the role of lncRNAs in this process remains largely unknown. Here, we have identified 291 lncRNAs which were differentially expressed in lung cancer tissues compared with adjacent normal tissues. Of these, the gene body or vicinity of 19 transcripts were also bound by SMAD3. The expression of LINC01186 was significantly decreased in A549 cells treated with TGF-β1. Furthermore, LINC01186 was stably down-regulated in lung cancer tissues compared with normal tissues in TCGA data sets and another published lung cancer data sets. The bioinformatics analysis suggested that LINC01186 was associated with TGF-β signaling pathway involves TGF-β ligands, receptors and receptors-activated SMAD protein 2 and 3. TGF-β binds to type II and I receptors, which form a heteromeric complex of transmembrane serine/threonine kinases, resulting in phosphorylation of the receptor-regulated Smads (R-Smads), SMAD2 and SMAD3. Phosphorylated SMAD2/SMAD3 forms a heteromeric complex with SMAD4, translocates to the nucleus and regulates the transcriptional induction or repression of target genes (Xu et al., 2009). SMAD3 and SMAD4 bind directly to

1. Introduction

Lung cancer is a malignant lung tumor characterized by uncontrolled cell growth (Non-Small Cell Lung Cancer Treatment (PDQ(R)), 2002), and can spread beyond the lung by the process of metastasis (Mittal, 2016). Lung cancer is the most common cause of cancer-related death in men and second most common in women after breast cancer (McGuire, 2016). Despite of the rapid progress in lung cancer research, the survival of patients remains poor, only around 17% of patients could survive for 5 years or more (Siegel et al., 2012). Metastasis and recurrence are the major causes of death in patients with lung cancer (Sang et al., 2015). This emphasizes the urgency of gaining a better understanding of the mechanisms associated with metastasis in lung carcinogenesis.

Epithelial-Mesenchymal-Transition (EMT) is a biologic process that allows polarized epithelial cells to transform into fibroblast-like mesenchymal cells. During this transition, the cell markers shift from a predominately epithelial to a more mesenchymal type, and the cells become more motile and invasive, eventually leading to metastasis (Kalluri and Weinberg, 2008). A large number of cytokines and growth factors can induce and maintain EMT, among which Transforming growth factor β (TGF-β) is a potent driver (Arvelo et al., 2016). The canonical TGF-β signaling pathway involves TGF-β ligands, receptors and receptors-activated SMAD protein 2 and 3. TGF-β binds to type II and I receptors, which form a heteromeric complex of transmembrane serine/threonine kinases, resulting in phosphorylation of the receptor-regulated Smads (R-Smads), SMAD2 and SMAD3. Phosphorylated SMAD2/ SMAD3 forms a heteromeric complex with SMAD4, translocates to the nucleus and regulates the transcriptional induction or repression of target genes (Xu et al., 2009). SMAD3 and SMAD4 bind directly to
chromatin through their N-terminal MH1 domains, while SMAD2 does not bind directly to chromatin (Yagi et al., 1999; Dennler et al., 1998). Several groups have reported the pivotal roles of SMAD3 in EMT, e.g. renal tubular epithelial cells deficient in SMAD3 fail to undergo EMT in response to TGF-β or mechanical stress (Sato et al., 2003), and keratinocytes derived from SMAD3 knockout mice show reduced migration in response to TGF-β (Ashcroft et al., 1999). Recently, Song and colleagues demonstrated that increased SMAD2 and SMAD3 promoted lung cancer growth and metastasis (Tang et al., 2015).

The ENCODE Consortium has elucidated that the human genome is pervasively transcribed and that the protein-coding genes occupy only a small proportion (1–2%) of the genome (Birney et al., 2007). This indicates there are many non-protein-coding transcripts, including small noncoding RNAs and long noncoding RNAs (lncRNAs). Among the apparent vast and various noncoding transcripts, lncRNAs are commonly defined as a class of transcripts longer than 200 nucleotides but with no apparent protein-coding capability (Shi et al., 2013). Compared with protein-coding genes, lncRNAs are generally expressed at lower levels, are enriched in the chromatin and nucleus of the cell, and display less sequence conservation across related species (Derrien et al., 2012). Recently, many researches have demonstrated that lncRNAs are biologically functional and play important roles in the initiation and progression of various cancers (Wilusz et al., 2009; Ponting et al., 2009; Wapinski and Chang, 2011; Tsai et al., 2011), including lung cancer. Previous studies also showed that a number of lncRNAs, such as HOTAIR (Nakagawa et al., 2013), SPRY4-IT1 (Sun et al., 2014), H19 (Matouk et al., 2010), and MEG3 (Lu et al., 2013), function differently in lung cancer. However, the roles of lncRNAs in regulating migration and invasion through EMT in lung cancer are still not well studied.

In this study, we sought to identify lncRNAs that were regulated by SMAD3 and played a role in EMT process in lung cancer. At first, we identified 291 lncRNAs that were differentially expressed lncRNAs (DElncRNAs) between lung cancer tissues and adjacent normal tissues through a reanalysis of public microarray data. Combined with SMAD3 Chromatin Immunoprecipitation Sequencing (ChIP-Seq) data which was deposited to the Gene Expression Omnibus (GEO) repository by previous study (Iosogaya et al., 2014), we then obtained 19 potentially SMAD3-regulated lncRNAs. Finally, we focused on the analysis and characterization on the functional roles of LINC01186 in lung cancer. The results showed that LINC01186 inhibited cell migration and invasion. The expression of LINC01186 was repressed by SMAD3, and shRNA knocking-down of LINC01186 affected the expression levels of several EMT molecular markers. Our study proposed the existence of an aberrant TGF-β/SMAD3/LINC01186 signaling axis leading to lung cancer aggressiveness through EMT induction, and LINC01186 could serve as a new diagnostic biomarker and therapeutic target.

2. Materials and methods

2.1. Cell culture and TGF-β1 treatment

The human cell lines A549, H1299 and 293T were used in this study. A549 and H1299, which are lung cancer epithelial cell lines, were incubated in RPMI 1640 medium (Life Technologies, 16400-044). 293T was cultured in DMEM medium (Life Technologies, DMEM, 11995-065) supplemented with 10% FBS. A549 cells were starved in media without FBS for 24 h before the addition of 5 ng/ml recombinant human TGF-β1 (R&D Systems, 240-B-010) or vehicle.

2.2. RNA extraction and quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR)

Total RNA was extracted from cultured cells with TRIzol reagent (Invitrogen, 15596-026) according to the manufacturer’s instructions, and then subjected to DNasel (Ambion, AM2235) treatment. qRT-PCR was performed using the Transcript II Green One-Step qRT-PCR SuperMix kit (TransGen Biotech, AQ311-01) with 50 ng RNA as template in a 10 μl reaction volume using SYBR green master mixture on the Rotor-Gene® Q real-time cycle (Qiagen). The results were normalized to the expression of GAPDH, and the relative expression was calculated by the delta-delta Ct method. The primer sequences are listed in Table 1.

2.3. Microarray analysis and ChIP-seq data analysis

The transcriptome profiles for 20 pairs of lung cancer tissues and matched normal tissues were measured by the combined mRNA + lncRNA V2.0 microarray on an Agilent platform. The microarray data had previously been deposited on the Gene Expression Omnibus (GEO), under accession number GSE70880. Data normalization was carried out as previously described (Yuan et al., 2016). After obtaining the normalized expression value of the probes, we first re-annotated the probes by mapping the probe sequences to gene sequences from the Gencode database (Gencodev23) (Harrow et al., 2012) using the Blat software (Kent, 2002) with the parameter --minidensity = 100. Only probes mapping to the unique gene sequence without mismatch were retained. If more than one probe was mapped to a gene, the median expression values of these probes would be used to...
represent the expression value of this gene. Paired-t-test was used to obtain the differentially expressed genes (DEGs) using a Limma package in R (Smyth, 2004). p value ≤ 0.05 and 2-fold change in its expression level were chosen as the cut-off criteria.

The ChIP-Seq data of SMA3D in A549 cell line was downloaded from the GEO database with accession number GSE51509 (I3ogaya et al., 2014). We downloaded the locations of the peaks regions in human genome (h18) obtained from the dataset. We then converted the genomic coordinates to specific human genome versions (hg19) using the UCSC LiftOver Tool (Rosenblomn et al., 2015). We defined the upstream 5 k of the transcription start sites as the gene promoter region in order to obtain all the possible SMA3D-regulated candidate IncRNAs. Binding site locations of SMA3D intersected with the promoter region or the gene body region of the DEGs using the IntersectBed from BedTools (Quinlan and Hall, 2010) with the default parameters.

2.4. Gene set enrichment analysis (GSEA) analysis

The GseaPreRanked tool in GSEA v2.0 (Subramanian et al., 2005) was used to perform GSEA on LINC01186. We obtained the gene sets not only from the MSigDB database v5.0 but also from published gene signatures. Pearson correlation coefficients between expression profiles of mRNAs and LINC01186 across all lung tissues were calculated through our python script. Statistical significance was assessed by comparing the enrichment score to enrichment results generated from 1000 random permutations of the gene sets to obtain p values (nominal p value).

2.5. Weighted gene co-expression network analysis (WGCNA) analysis and gene ontology analysis

We used weighted correlation network analysis (WGCNA v1.51) (Langfelder and Horvath, 2008) to find modules of high correlated genes that were co-expressed with LINC01186. The LINC01186-mRNA pairs with the p value of Pearson correlation coefficients, calculated described below, below 0.01 were considered to be significantly co-expressed and finally 788 co-expression mRNAs were selected to construct the network. Firstly, a soft threshold was chosen to create network with a scale free topology, using the function pickSoft Threshold in the WGCNA. Secondly, we calculated the adjacencies using the soft thresholding power 12, transformed the adjacency into Topological Overlap Matrix (TOM) and calculated the corresponding dissimilarity. Then the networks were built, modules of transcripts with similar expression patterns were created and eigengenes for these modules were calculated with default settings, except that the minModuleSize was set to 20, and the mergeCutHeight was set to 0.25. Next, we obtained 6 modules and exported the network to Cytoscape (Shannon et al., 2003) for visualization. Finally, functional enrichment analysis of those WGCNA clusters of transcripts was performed with the default settings, except that the scale-free topology was used.

2.6. Subcellular fractionation

The nuclear and cytoplasmic fractions of A549 cells were isolated with the NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo, 78833). The nuclear location of LINC01186 was assessed by qRT-PCR on total RNA isolated from the subcellular fractions. GAPDH and ACTB were served as the cytosolic controls while U1 was used as the nuclear control.

2.7. Western blotting

Western blotting was performed as standard protocols. The primary antibodies used in this study were obtained as follows: GAPDH (Abnova), Vimentin, E-cadherin (Santa Cruz), and Fibronectin (BD Transduction).

2.8. ShRNA-mediated interference and lentivirus packaging

shRNAs against LINC01186 were designed using online software (http://www.clontech.com/). The nuclear location of LINC01186 was assessed by qRT-PCR. Five quadruplexes were designed, and the sequences of the most effective shRNAs were provided as follows: sh2-LINC01186: 5'-GAGAGAGAGATCCAGAAGA-3', sh5-LINC01186: 5'-GACTCAGCAGTCA CGTAA-3'. A non-targeting, scrambled silencing RNA was used as a negative control (NC). shRNAs against LINC01186 and the negative control hairpins were cloned into pScicR-puro lentiviral vector. Virus packaging was performed in 293T cell after cotransfection of recombinant lentivirus expression plasmids for LINC01186 together with the packaging plasmids (pPRE, pCMV-VSVG, and pR5-REV, Addgene) using Lipofectamine 2000 (Invitrogen, 11668-019). Cell medium containing virus was harvested twice 48 h and 72 h after transfection, and the viruses were concentrated with 5 × PEG8000. After A549 cells were infected with viruses for 24 h, thereafter fresh medium with puromycin was added to select for positive cells for additional 3 days incubation. The RNAi efficiency of LINC01186 was assessed by qRT-PCR analysis.

2.9. siRNA-mediated knockdown of SMA3D

siRNAs used to knock down SMA3D mRNAs were designed at the website of Integrated DNA Technologies (http://www.idtdna.com) and synthesized by the GenePharma company. A nonspecific siRNA used as the negative control (NC) was also purchased from the GenePharma Company. siRNA sequences (5′ to 3′) are listed as follows: si1-SMA3D: sense rGrUrGrArGrArCrUrGrUrCrUrCrArCrCUT, antisense rArUrGrGrUrGrUrGrArGrArCrUrGrUrCrCrArGrc; si2- SMA3D: sense rGrCrUrGrCrUrGrArGrArGrArGrUrGrArGrArGrArArCrArGAC, antisense rGrUrUrCrUrCrArCrUrUrCrArCrUrGrGrUrArGr CrcArA. A549 cells were seeded in 12-well plates at a desired concentration for 12 h and transfected with siRNA oligonucleotides with Lipofectamine 2000 according to the manufacturer’s instructions.

2.10. Cloning and overexpression

All PCR primers used for cloning are listed in Table 2. The DNA polymerase for cloning was ExPfu DNA Polymerase (GenStar, A161-01), and restriction endonucleases were from New England Biolabs (NEB). The full-length LINC01186 transcript and SMA3D open reading frame were amplified from A549 cDNA and cloned into the pcDNA3.1 (+) plasmid. Overexpression of recombinant plasmid was performed in A549 or H1299 cells by transient transfection using Lipofectamine 2000.

2.11. Cell migration and invasion assay

For knockdown and overexpression of LINC01186, A549 or H1299 cells were transfected with shRNAs or plasmids, trypsinized and counted after 48 h. A total of 7 × 10^4 cells (for the migration assay) or 10 × 10^4 cells (for the invasion assay) in 100 μl serum-free medium were seeded into the upper chamber of Transwell plates (Corning, 3422) with (for invasion) or without (for migration) BD Matrigel matrix (BD, 356230) coating. 800 μl medium containing 20% FBS was added into the lower chambers as the chemoattractant factor. After another 8 h culture at 37 °C and 5% CO2, the non-migrating or non-invading cells in upper chamber were gently removed, while the migrating or invading cells were counted.

Table 2

<table>
<thead>
<tr>
<th>Primer sequences for construct plasmid.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genes</td>
</tr>
<tr>
<td>LINC01186-Kpn I-F</td>
</tr>
<tr>
<td>LINC01186-Nol I-R</td>
</tr>
<tr>
<td>SMA3D-Kpn I-F</td>
</tr>
<tr>
<td>SMA3D-Nol I-R</td>
</tr>
</tbody>
</table>

Y. Hao et al. / Gene 608 (2017) 1–12
stained with 0.1% crystal violet for 20 min, washed with water, air dried and photographed using an Olympus microscope imaging systems. The migrating and invading cells were counted and averaged from images of six random fields (magnification was 200×) for differential analysis.

2.12. Colony formation assay

For colony formation assays, 3000 A549 cells (knockdown of LINC01186) or 1000 cells (overexpression of LINC01186) were seeded into 10 cm petri dishes and maintained in RPMI 1640 medium with 10% FBS at 37 °C. After one week, the cells grew to visible colonies. Washed twice with PBS and fixed with 100% absolute ethyl alcohol for 20 min, the colony cells were stained with 0.1% crystal violet for another 20 min. The numbers of colonies per plate were counted for statistical analysis.

2.13. Statistical analysis

Experimental data are shown as the means ± SD, all experiments were conducted at least three times. Significance was determined using the Student's t-test: *, p < 0.05; **, p < 0.01 and ***, p < 0.001.

3. Results

3.1. Nineteen lncRNAs were expressed differently in lung cancer tissues relative to adjacent normal tissues and potentially regulated by SMAD3

To identify lncRNAs that were regulated by SMAD3 and participated in the development of lung cancer, we re-analyzed the microarray data from pairs of lung cancer and matched surrounding normal tissues of 20 patients in combination with SMAD3 ChIP-Seq dataset. We re-annotated the probes in the original microarray data, and found that 12,322 lncRNAs were interrogated by one or more unique probes. To determine how many lncRNAs were differentially expressed and thus might have potential roles in cancer-related processes, we compared the expression levels of lncRNAs between lung cancer tissues and adjacent normal tissues, and found 133 up-regulated and 158 down-regulated lncRNAs in the tumor tissues (Fig. 1a). We then obtained the SMAD3 ChIP-Seq data of A549 cells from GEO and found that 19 of the DElncRNAs (Table S1) which gene body or vicinity were also bound by SMAD3 (Fig. 1b). Among the 19 lncRNAs, there were 3 known lncRNAs: PVT1 (Yang et al., 2014), CCAT1 (Luo et al., 2014) and LINC00472 (Shen et al., 2015a) (Fig. 1c), that were associated with lung cancer or other cancers in previously published papers. We discarded 3 reported lncRNAs (LINC00472, PVT1 and CCAT1) and STARD4-AS1 (5525 bp) whose length was oversize, and

![Fig. 1. Nineteen lncRNAs are expressed differently in lung cancer tissues and bound to SMAD3.](image-url)
examined the expression of the left 15 novel IncRNAs in A549 cells via qRT-PCR (Fig. 1d).

3.2. LINC01186 is significantly down-regulated in TGF-β1 treated cells

To decipher whether these 19 IncRNAs might be involved in TGF-β induced EMT, we treated A549 cells with TGF-β1. After incubation with TGF-β1 for 48 h, the A549 cells underwent morphological changes to a spindle-shaped cells (Fig. 2a), showed decreased expression of the epithelial marker CDH1, as well as enhanced expression of mesenchymal markers, including SNAIL, CDH2, VIM and FN (Fig. 2b-c), all changes compatible with an EMT transition. After examined expression abundance of the 15 candidate IncRNAs in A549 cells using qRT-PCR, we abandoned 5 IncRNAs whose expression were too low, including EVADR, AC022173.2 and so on. We examined the 10 IncRNAs remaining in the analysis for changes in their expression levels. The results showed that LINC-PINT3 was significantly up-regulated in the TGF-β treated cells compared with untreated cells, while LINC01186, RP11-8L2.1, TARID and TM4SF1-AS1 were significantly down-regulated (Fig. 2d–e). Given that the relative change in LINC01186 expression was relatively large, the absolute expression of LINC01186 was relative high in cells, and that its locus is found in an intergenic region of the human genome, we surmised that it might be a promising candidate IncRNA for independent functions.

3.3. GSEA and WGCNA analysis supported that LINC01186 regulated cell invasion and migration through the TGF-β pathway

To gain insights into the biological significance of LINC01186 in the lung cancer, GSEA was performed. Pearson correlation coefficients between expression profiles of mRNAs and LINC01186 across all lung

Fig. 2. LINC01186 is significantly down-regulated in TGF-β1 treated A549 cells. a The morphological change of A549 cells with (right) or without (left) 5 ng/ml TGF-β1 treatment for 48 h. b qRT-PCR analysis of EMT markers mRNA levels. c Western blotting analysis of EMT markers protein levels. d Heatmap of log2(Foldchange – values) in TGF-β1 treated cells compared with untreated cells for 10 novel candidate IncRNAs with 3 replicates. The expression values of genes are assessed by qRT-PCR in A549 cells. Red (blue) denotes positive (negative) log2(Foldchange – values). e qRT-PCR analysis of candidate IncRNAs level in A549 cells treated with TGF-β1. All these data show the mean ± SEM from three independent assays. Statistical significance is measured by Student’s t-test. *p < 0.05, **p < 0.01, ***p < 0.001.
tissues were also calculated, and based on these LINC01186-associated gene sets were identified. Interestingly, genes that were expressionally correlated with LINC01186 were enriched for functions related to TGF-β, and the functional gene sets were obtained from two independent studies (Calon et al., 2012; Sartor et al., 2010) (Fig. 3a). Particularly, the gene set that was down-regulated after TGF-β treatment of A549 cells was significantly enriched in genes whose expression was positively correlated with that of LINC01186 (CALON_2012_ENDTBRS and LUNG_2009_A549_Down). Correspondingly, the gene set that up-regulated after TGF-β treatment of A549 cells was significantly enriched in LINC01186-negatively-associated genes (CALON_2012_MATBRS and LUNG_2009_A549_Up). These results support the possibility that LINC01186 is negatively correlated with TGF-β in tumors.

Fig. 3. Functional analysis supports that LINC01186 regulates cell invasion and migration through the TGF-β pathway. a GSEA analysis of LINC01186 co-expressed mRNAs in previous published datasets. The co-expressed mRNAs are enriched in TGF-β-associated gene sets. The description of the gene sets and the nominal p values are shown. b Graphical view of the LINC01186 co-expressed mRNAs network where the nodes correspond to genes and the edges to the co-expression links. The clusters are the gene modules obtained from WGCNA analysis. Functional enrichment is done by DAVID for each module, respectively. The significant GO terms of biological processes with a p value < 0.05 are selected and depicted.
To identify the most relevant gene expression signatures for *LINC01186*, we generated a gene co-expression network using WGCNA. This approach allowed us to correlate gene modules with external variables (*LINC01186* associated gene sets in this case) and predicted the potential function of *LINC01186*. WGCNA groups genes into discrete modules based on their topological overlap within a network built from gene expression data. We used WGCNA to analyze the 786 genes whose expression were correlated with *LINC01186* (Pearson correlation, p value ≤ 0.01). These genes were segmented into six modules, and for each of these modules, gene ontology enrichment analysis was performed with DAVID online tools (Huang da et al., 2009a; Huang da et al., 2009b). We focused our attention on the module 1, 2, 3 and 4 that were significantly enriched in functions associated with cell migration, cell adhesion, angiogenesis and cell junction (Fig. 3b and Fig. S1), and which indicated *LINC01186* might regulate cell invasion and migration and participate in EMT process. All genes in different modules have been deposited in Table S3. These findings along with GSEA results supported the hypothesis that *LINC01186* might regulate cell invasion and migration through the TGF-β pathway.

3.4. Molecular characterization and expression pattern of *LINC01186*

In order to verify that *LINC01186* functions as RNA rather than protein, we examined the coding potential of *LINC01186* using in silico prediction. The in silico results obtained with three powerful computational tools (Sun et al., 2013; Wang et al., 2013; Kong et al., 2007) for distinguishing coding and noncoding transcripts consistently showed that *LINC01186* has no ability to encode a protein (Table S2). To confirm the expression patterns of *LINC01186* in tissues are in consistent with cell lines, we tested the expression level changes in other two previous published datasets (TCGA LUAD RNA-seq dataset; ENA archive: EPR001058). The results suggested that *LINC01186* was stably down-regulated in lung cancer tissues compared with normal tissues (Wilcoxon rank test: p < 0.05) (Fig. 4a). PCR analysis yielded a product which accorded with previous annotations of *LINC01186* as a ~600 nt long transcript, consisting of two exons, and originating from a locus of approximately 1.7 kb in the short arm of chromosome X (Fig. 4b–c). Next, comparison on the nuclear and cytosolic fractions of A549 cells showed that the *LINC01186* transcript was mainly located in the nucleus irrespective of TGF-β1 treatment (Fig. 4d).

The expression of IncRNAs is often tissue-specific, and only a few IncRNAs with a consistent expression pattern in different types of cancer have been reported. We therefore examined the *LINC01186* expression in 13 different types of cancers curated in TCGA and found *LINC01186* was differentially expressed in 9 types of cancers while down-regulated in 7 types of cancers in accordance with our results in lung cancer (Fig. 4e). The results suggest that *LINC01186* might serve as a common biomarker for multiple types of cancers.

3.5. *LINC01186* inhibits migration, invasion and colony formation

It is well-established that TGF-β could promote cancer metastasis through its effects on the tumor microenvironment. *LINC01186* was significantly down-regulated in TGF-β1 treated cells. GSEA and WGCNA
Fig. 5. **LINC01186** inhibits migration, invasion and colony formation. a Targeting **LINC01186** in A549 cells by two independent shRNAs led to depletion of the transcript as measured by qRT-PCR. b Representative images of transwell migration (top) and invasion (bottom) assays in shRNA-mediated knocking-down of **LINC01186**. c The barplot shows the summary of the number of migrated (left) and invaded (right) cells per field. d **LINC01186** levels are measured by qRT-PCR after transfection of H1299 cells with negative control or **LINC01186**-overexpressed vectors. e Representative images of transwell migration (top) and invasion (bottom) assays in H1299 cells overexpressing **LINC01186** or control vector. f The barplot shows the summary of the number of migrated (left) and invaded (right) cells. g and h Colony formation assays indicate that **LINC01186** knockdown increased A549 cells in vitro tumorigenesis ability (top) and **LINC01186** overexpression decrease the ability. All these data show the mean ± SEM from three independent assays. Each sample used three technical replicates. Statistical significance is measured by Student’s t-test. **p < 0.01, ***p < 0.001.
results indicated that LINC01186 might regulate cell invasion and migration and participate in EMT. To study this further, we performed the transwell migration assays on lung cancer cell lines. We knocked down LINC01186 using two shRNAs (Fig. 5a) and found that treatment of cells with LINC01186 specific shRNAs increased the migration and invasive capacity of A549 cells (Fig. 5b–c). We also constructed a plasmid for overexpression of LINC01186 in H1299 cells (Fig. 5d). H1299 cells overexpressing LINC01186 were subjected to the transwell assay and their migration ability and invasive capacity were determined as well. It turned out that ectopic expression of LINC01186 in H1299 cells reduced the migration and invasive capacity of the cells (Fig. 5e–f). Gain and loss of function experiments thus suggested that LINC01186 inhibited migration and invasion capacity of lung cancer cells.

Colony-formation assay was further performed to assess the effect of LINC01186 on in vitro tumorigenesis of the transfected lung cancer cells. As shown in Fig. 5g–h and Fig. S2, LINC01186 depleted cells formed significantly more colonies compared with control cells. Notably, sh5 construct showed much stronger effect than sh2 construct, probably because the knockdown efficiency of sh5 was better than sh2 or sh2 efficiency lost faster along with time. On the other hand, cells overexpressing LINC01186 formed markedly fewer colonies.

3.6. LINC01186 is regulated by SMAD3 and inhibits EMT

SMAD3 protein has been studied extensively as a pivotal intracellular effector of TGF-β. SMAD3 acts as a transcription factor and the SMAD3 ChIP-Seq data analysis above showed that SMAD3 binds to the gene body of LINC01186 (Fig. 6a), which suggests the possibility that LINC01186 may be regulated by SMAD3. The previous results also demonstrated that LINC01186 was significantly down-regulated in TGF-β1 treated cells. To explore the relationship between SMAD3 and LINC01186, SMAD3 knockdown and overexpression experiments were carried out in lung cancer cell lines. Compared with negative controls, the expression of LINC01186 was significantly increased in siRNA-SMAD3 transfected cells (Fig. 6b) and decreased in cells overexpressing SMAD3 (Fig. 6c). The results above thus suggested that LINC01186 was regulated by SMAD3.

To explore whether LINC01186 inhibit lung cancer cells migration and invasion through the EMT process, we detected the expression level changes of the EMT markers after stable LINC01186 knockdown by qRT-PCR and Western blotting. It was shown that knockdown of LINC01186 dramatically decreased the expression of epithelial markers and enhanced expressions of mesenchymal markers, including SNAI1 mRNA, E-cadherin, and Vimentin proteins (Fig. 6d–e). The above results indicated that LINC01186, regulated by SMAD3, is essential for TGF-β1 induced EMT.

4. Discussion

TGF-β1 can facilitate cancer progression by promoting tumor invasion and metastasis through up-regulation of EMT regulatory factors (Xu et al., 2009). Most previous studies on the effects of TGF-β1 have focused on the roles of protein-coding genes in the process of tumor migration and invasion (Yin et al., 2008; Ahmed et al., 2001; Sugimachi et al., 2014). Recent evidence highlights the crosstalk between IncRNAs and the TGF-β1 signaling pathway in many cancers including esophageal cancer (Chen et al., 2014), melanoma (PLoS One, 2015), breast cancer (Matouk et al., 2014), and HCC (Li and Kang, 2014). However, whether IncRNAs participate in processes involving TGF-β1 and EMT in lung cancer has not been known. We used microarray datasets of 20 pairs of lung cancer and adjacent normal tissues to identify DEIncRNAs. These DEIncRNAs may contribute to the development of lung cancer. SMA3 plays an important role in the canonical TGF-β1 signaling pathway. Reduced expression of SMAD3 has been shown to suppress metastasis of breast cancer cells (Tian et al., 2003). Then, we obtained SMAD3 ChIP-Seq data and detected 19 DEIncRNAs that were potentially regulated by SMAD3. Among the 19 DEIncRNAs, there are 3 previously reported cancer related IncRNAs: PVT1 (Yang et al., 2014), CCAT1 (Luo et al., 2014), and LINC00472 (Shen et al., 2015a) (Fig. 1c). PVT1 was up-regulated in NSCLC tissues in corresponding normal tissue (Yang et al., 2014) and might serve as a novel molecular marker for lymph node metastasis (Liu et al., 2016). CCAT1 promoted gastric cancer migration (Zhou et al., 2016) and serves as an oncogenic IncRNA in several cancers (Xin et al., 2016). LINC00472 functioned as a tumor suppressor in breast cancer (Shen et al., 2015b). All these previous findings are in accordance with the results from our current study, which further confirm that the microarray results and ChIP-Seq results are trustworthy.

In this study, we showed that LINC01186, one of the 19 DEIncRNAs, plays an important role in migration, invasion and colony formation in vitro of lung cancer cells through gain and loss of function experiments. Furthermore, we demonstrated that LINC01186 is a downstream target of SMAD3, since knockdown and overexpression SMAD3 significantly altered the expression of LINC01186. Moreover, LINC01186 influenced the expression of EMT markers at both transcript and protein level. This is the first study to establish a role of a previously uncharacterized IncRNA in TGF-β1 induced EMT. Based on these data, we proposed a model in which LINC01186 is regulated by SMAD3 and inhibits the lung cancer migration and invasion through influencing the EMT process. LINC01186 is thus a new noncoding molecule that can mediate the relationship between TGF-β1 and EMT process, and understanding how LINC01186 is involved in lung cancer metastasis may potentially contribute towards new prognostic models and alternative treatments.

Utilizing the TCGA data, we found LINC01186 was differentially expressed in several cancers. This suggests that LINC01186 may be a molecular marker in many different types of tumors. While this study has focused the function of LINC01186 in lung cancer, the potential functions of LINC01186 in other tumors are deserving of future research. In addition to LINC01186, 7 other novel IncRNAs were also good candidates for functional studies of TGF-β1 induced EMT. However, while LINC01186 can alter the expression of EMT markers, we have not been able to identify the molecules that interact directly with LINC01186, and the detailed mechanism by which LINC01186 functions in EMT and metastasis still remains to be elucidated further.

Table 1 lists the 19 DEIncRNAs in microarray and potentially regulated by SMAD3. Table 2 lists the coding potential of LINC01186. Table 3 lists genes in different modules obtained from WGCNA analysis. Fig. S1 is the high-definition map of the Fig. 3b. Fig. S2 is the high-definition map of the Fig. 5g (Fig. S2a: Negative control construct, Fig. S2b: sh2-LINC01186 construct, Fig. S2c: sh5-LINC01186 construct, Fig. S2d: pcDNA3.1(+) plasmid, Fig. S2e: pcDNA3.1(+) LINC01186 plasmid). Supplementary data associated with this article can be found in the online version, at doi: 10.1016/j.gene.2017.01.023.

Table 1 lists the 19 DEIncRNAs in microarray and potentially regulated by SMAD3. Table 2 lists the coding potential of LINC01186. Table 3 lists genes in different modules obtained from WGCNA analysis. Fig. S1 is the high-definition map of the Fig. 3b. Fig. S2 is the high-definition map of the Fig. 5g (Fig. S2a: Negative control construct, Fig. S2b: sh2-LINC01186 construct, Fig. S2c: sh5-LINC01186 construct, Fig. S2d: pcDNA3.1(+) plasmid, Fig. S2e: pcDNA3.1(+) LINC01186 plasmid). Supplementary data associated with this article can be found in the online version, at doi: 10.1016/j.gene.2017.01.023.

Funding

This work was supported by grants from the National Natural Science Foundation of China (31520103905 and 31401098) and the National High Technology Research and Development Program (“863” Program) of China (2015AA020108 and 2014AA021502).

Conflict of interest

The authors declare that they have no competing interests.
Acknowledgements

We are very grateful to Dr. Geir Skogerbø and Mr. Guofeng Zhao for helpful suggestions and critical reading of this manuscript. We also thank Dr. Xiaowei Chen and Dr. Zhen Fan for discussion on high-throughput sequencing.

References


学霸图书馆

www.xuebalib.com

本文献由“学霸图书馆-文献云下载”收集自网络，仅供学习交流使用。

学霸图书馆（www.xuebalib.com）是一个“整合众多图书馆数据库资源，
提供一站式文献检索和下载服务”的24小时在线不限IP图书馆。
图书馆致力于便利、促进学习与科研，提供最强文献下载服务。

图书馆导航：
图书馆首页 文献云下载 图书馆入口 外文数据库大全 疑难文献辅助工具