Gene 608 (2017) 1–12

Contents lists available at ScienceDirect

Gene

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

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

Yajing Hao [a,b,c,1], Xinling Yang [a,b,c,d,1], Dongdong Zhang [a,b], Jianjun Luo [a,b,*], Runsheng Chen [a,b,*]

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

ARTICLE INFO

Article history:
Received 25 September 2016
Received in revised form 5 January 2017
Accepted 20 January 2017
Available online 22 January 2017

Keywords:
Lung cancer
TGF-β
Long noncoding RNA
Epithelial-Mesenchymal-Transition
SMAD3

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 and might participate in EMT process. Moreover, knocking-down LINC01186 promoted cell migration and invasion, whereas, LINC01186 overexpression prevented cell metastasis. Importantly, LINC01186 expression was regulated by SMAD3. And LINC01186 affected several EMT markers expression. These findings suggest that LINC01186, a mediator of TGF-β signaling, can play a significant role in the regulation of EMT and lung cancer cell migration and invasion.

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 predominantly epithelial to a more mesenchymal type, and the cells become more motile and invasive, eventually leading to metastasis (Kalluri and Weinberg, 2009). 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

Abbreviations: lncRNAs, long noncoding RNAs; EMT, Epithelial-Mesenchymal-Transition; TGF-β, transforming growth factor β; ChIP-Seq, Chromatin Immunoprecipitation Sequencing; qRT-PCR, quantitative real-time reverse transcription polymerase chain reaction; GSEA, gene set enrichment analysis; WGCNA, weighted gene co-expression network analysis; FBS, fetal bovine serum; GEO, Gene Expression Omnibus; DEGs, differentially expressed genes; NC, negative control; DElncRNAs, differentially expressed lncRNAs; NEB, New England Biolabs.

☆ The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors. The authors declare no competing financial interest.

* Corresponding author at: Institute of Biophysics, Chinese Academy of Sciences, No.15, Datun Road, Chaoyang District, Beijing 100101, China.
E-mail addresses: luojianj@ibp.ac.cn (J. Luo), rschen@ibp.ac.cn (R. Chen).

[1] Equal contributors.

http://dx.doi.org/10.1016/j.gene.2017.01.023
0378-1119/

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 (Isogaya 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, 11875-093) supplemented with 10% fetal bovine serum (FBS) (Life Technologies, 16000-044). 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 DNaseI (Ambion, AM2235) treatment. qRT-PCR was performed using the TransScript 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 cycler (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 –minIdentity = 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

Table 1
Primer sequences for qRT-PCR.

Genes	Primer sequences (5′ to 3′)
LINC00511	Forward: AAAGGAAGAAATGACCGAGGG
	Reverse: GAGTCCTCATGCCTATAATCACG
LINC00862	Forward: CAGCGATTGGAGTGATGTAGT
	Reverse: AGAAGTCCCAAGTCCCAAATC
TARID	Forward: CGACTAGATCGTTTGGCTCTTT
	Reverse: GCGGTTGATACTCCCTGTATTT
AC022173.2	Forward: GGTGGATTGAAGCTGGTCG
	Reverse: CTTTCTTACAGGTGCGTTCTTTG
BBOX1-AS1	Forward: CAGACTCCTGCTTTGCTCTT
	Reverse: GGAAGCATCTTCTCAGCTTCT
LINC-PINT	Forward: GTGAAGCAGAATAAACCACTGAAC
	Reverse: AGATGGTTCCAGTCCCTCTT
RP11-8 L2.1	Forward: TGCTTCTCCATCCAGTTCTTC
	Reverse: CAGGAATGCTCTTGGGAGTT
MIR100HG	Forward: TTTGGAGTGTGGCAGAGTAAG
	Reverse: CTGCCCAGATAGACTGTTTCC
RP11-191 N8.2	Forward: TGCTGACATGTTACAGTCCTG
	Reverse: TTCAAAGTTTGCCAAGGTGC
RP11-443B7.2	Forward: CTGCTCAGCAAATTCCTTGATT
	Reverse: CAACACAGAGAGATGGAACATTTG
RP11-114H23.1	Forward: GGAGAGTTTCAGAGCAGAGAAG
	Reverse: CAGCTGGTCTCCCTGTTTC
TM4SF1-AS1	Forward: CTCTCCACTCCTGCTTTCATAC
	Reverse: AGCAAACCATCCTCAGTTCTT
LINC01186	Forward: CTAAAACGGGAAAGATGCCTTC
	Reverse: TCAAGATTAGGGAACATACTGGC
EVADR	Forward: ACCGATACAAGAACCTTCCAC
	Reverse: GCAAACTTTCCAGGACATTCAG
SNAI1	Forward: TCGGAAGCCTAACTACAGCGA
	Reverse: AGATGAGCATTGGCAGCGAG
CDH1	Forward: TGCCCAGAAAATGAAAAAGG
	Reverse: GTGTATGTGGCAATGCGTTC
CDH2	Forward: ACAGTGGCCACCTACAAAGG
	Reverse: CCGAGATGGGGTTGATAATG
VIM	Forward: GAGAACTTTGCCGTTGAAGC
	Reverse: GCTTCCTGTAGGTGGCAATC
FN	Forward: CAGTGGGAGACCTCGAGAAG
	Reverse: TCCCTCGGAACATCAGAAAC
SMAD3	Forward: CCTCAGTGACAGCGCTATTT
	Reverse: CGAACTCCTGGTTGTTGAAGA
U1	Forward: CCATGATCACGAAGGTGGTTT
	Reverse: ATGCAGTCGAGTTTCCCACAT
ACTB	Forward: CATGTACGTTGCTATCCAGGC
	Reverse: CTCCTTAATGTCACGCACGAT
GAPDH	Forward: GGAGCGAGATCCCTCCAAAAT
	Reverse: GGCTGTTGTCATACTTCTCATGG