Original Article
Cancer-associated methylated IncRNAs in patients with bladder cancer

Shufang Zhang1, Hui Cao1, Lili Ye2, Xiaohong Wen1, Shunlan Wang1, Wenwen Zheng3, Yingai Zhang1, Denggao Huang1, Yuanhui Gao1, Haifang Liu1, Haowei He1, Xin Gao1, Yinyi Chen1, Mei Chen1, Yang Xiang4, Fei Wang5

1Central Laboratory, Affiliated Haikou Hospital of Xiangya Medical College, Central South University, Haikou 570208, Hainan, China; 2Clinical Laboratory, Jilin Cancer Hospital, Changchun 130021, Jilin, China; 3Department of Clinical Laboratory, The Sixth Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510655, Guangdong, China; 4Department of Urology, Affiliated Haikou Hospital of Xiangya Medical College, Central South University, Haikou 570208, Hainan, China; 5Department of Urology, People’s Hospital of Hainan Province, Haikou 570311, Hainan, China

Received January 15, 2019; Accepted May 11, 2019; Epub June 15, 2019; Published June 30, 2019

Abstract: Epigenetic modifications via DNA methylation and long non-coding RNAs (IncRNAs) have been identified in bladder cancer (BC). However, DNA methylation of IncRNAs involved in BC has not been elucidated. Here, DNA immunoprecipitation-sequencing (MeDIP-seq) and RNA-sequencing (RNA-seq) were carried out using eight paired tumor and adjacent normal tissue samples from patients with BC. Differences in methylation patterns between tumors and controls were compared and the percentage of differentially methylated genes, including IncRNA genes, was calculated. RNA-seq data were subjected to gene ontology (GO), Kyoto encyclopedia of genes, and genomes (KEGG) analysis. The association between DNA methylation modification and IncRNA expression was determined by pairwise analysis of MeDIP-seq and RNA-seq data. The most enriched motifs in the promoter region, as well as the methylated density in the 3 kb region surrounding super-enhancers of IncRNA genes, were analyzed. A peak of 5mC methylation in the region 2 kb upstream of the transcription start site (TSS), with the lowest point in the TSS region, was observed. In total, 436 and 239 genes were identified to be hyper and hypomethylated, respectively, in BC tissue around the TSS region. RNA-seq revealed differentially expressed IncRNAs between tumor and normal tissues, many of which were cancer-associated IncRNAs based on GO and KEGG pathway analysis. Combined MeDIP-seq and RNA-seq analysis revealed that expression of 26 IncRNAs were candidates of 5mC controlled genes. The possible link between 5mC modification and differential IncRNAs may relate to enrichment of 5mC reads in the region surrounding super-enhancers of IncRNA. Survival analysis indicated that the methylated IncRNA, LINC00574, was associated with shorter overall survival time in patients with BC (HR = 1.7, p-value = 0.035). Taken together, these findings indicate that IncRNAs genes are under control of DNA methylation. Methylated IncRNA genes, which are transcripted to LINC00574, may serve as biomarkers for BC prognosis.

Keywords: DNA methylation, MeDIP-seq, bladder cancer, IncRNA

Introduction
Bladder cancer (BC) is one of the most common cancers of the urinary system. Risk factors for BC are known to include tobacco use, Schistosoma haematobium infection, industrial exposure, regular diet, and lifestyle trends. In 2012, approximately 430,000 new BC cases were diagnosed [1]. In western countries, it is estimated there are 118,000 newly diagnosed cases, which has led to a substantial health service burden [2, 3]. In China, rates of BC have increased rapidly over the five-year period 2003-2008, with a higher increase in women than men [4]. Despite surgery, dissection, and various adjuvant treatments for BC, the five-year survival rate remains low (60%) and there is a high risk of recurrence. It is reported that 30-70% of tumors reoccur [5], and 30% of tumors may develop into muscle-invasive disease [6, 7]. Therefore, better understanding of disease-causing molecules and mechanisms in BC is urgently needed to improve tumor prevention and control.
Epigenetics is the study of gene expression changes that do not involve alterations in the DNA sequence, such as DNA methylation, histone modifications, and non-coding RNA-mediated gene silencing. DNA methylation is a system of genetic text annotation, the most common form of which is to add a methyl group to the 5’ cytosine of C-G dinucleotides to form cytosine guanine dinucleotides (CpG), which provides information for gene transcription and mediating gene expression suppression or silencing [8]. Aberrant epigenetic modifications have been found in several types of cancers including BC [8].

Long non-coding RNAs (lncRNAs) are a class of RNA defined as transcripts > 200 nt in length transcribed from non-protein coding regions of the genome [9]. Multiple reports have demonstrated the expression of lncRNAs in mammals [10, 11]. Many lncRNAs are expressed specifically in cancer cells [12], and changes to their expression are critical for tumor initiation and progression [13].

Recently, growing evidence has focused on the relationship between DNA methylation and non-coding genes in human cancer [14]. Aberrant epigenetic modifications of lncRNAs have been reported in several types of cancers, including BC [15-18]. However, understanding of DNA methylation patterns of lncRNAs genes in BC remains limited. In the present study, we carried out methylated DNA immunoprecipitation-sequencing (MeDIP-seq) and RNA-sequencing for BC-associated methylated lncRNA screening in eight pairs of BC and matched normal adjacent tissue samples. We then further analyzed the association between differentially expressed lncRNA controlled by DNA methylation and BC prognosis.

Methods and materials

Clinical sample

Eight paired tumor and adjacent normal tissues from patients with BC were collected at the People’s Hospital of Hainan Province from March 2014 to July 2016 before microscopic examination. BC tumor tissues were confirmed to be invasive urothelial bladder cancer by pathologists, with grades ranging from 1 to 4 and sizes ranging from 0.6 cm$^3$ to 135 cm$^3$. All tissue samples were immediately transferred into liquid nitrogen and stored at -80°C until genomic DNA extraction. This experiment was approved by the Committee on the Ethics of People’s Hospital of Hainan Province. All patients were fully informed about the study and provided informed consent.

MeDIP-seq

Genomic DNA from the eight tumor tissues and paired tumor-adjacent normal tissues was isolated using the MagMedIP kit (Diagenode, Denville, NJ, USA) according to the manufacturer’s instructions. Next, methylated DNA was separated from unmethylated fragments by immunoprecipitation using a monoclonal antibody against 5-methylcytidine (Diagenode, Denville, NJ, USA), as previously described [4]. Illumina sequencing libraries of the enriched methylated DNA were prepared using NEBNext reagents according to the manufacturer’s recommendations (New England Biolabs, MA, USA). After quantitative evaluation of the size distribution and concentration of the library using an Agilent Bioanalyzer 2100 (Agilent Technologies, CA, USA), 100 nucleotide (nt) sequencing runs were conducted on an Illumina HiSeq 2000 platform (Illumina, CA, USA).

Differential methylated gene profile analysis

Raw MeDIP-seq data were processed according to the standard procedures of the Beijing Genome Institute (BGI; Shenzhen, China). Reads were aligned to the human reference genome (hg19) using Bowtie2 (version 2.2.6), with default parameters (max mismatches in seed alignment = 0). We then used MACS software (MACS 1.4.0) to scan for methylation peaks (Peak) genome-wide. The read counts of tumors and normal groups were converted into peak values. We then extracted the distribution of peak values across genes, as well as regions 2 kb upstream and downstream. A heat map was created using R (version 3.4.3) to identify genes that were either hyper or hypomethylated based on a fold change of ≥ 1.5 and P < 0.05 (Student’s t-test). Hyper and hypomethylated reads were counted and sorted for the following gene regions: upstream 2 kb, 5’UTR, CDS, intron, 3’UTR, and downstream 2 kb. The average signals of DNA methylations were compared among genes derived from the coding regions and lncRNA encoding regions.

RNA-seq for differential expression of lncRNAs

Total RNA was isolated from tumor and normal tissues using Trizol (Thermo Fisher Scientific,
IL, USA), according to standard protocols. Ribosomal RNA (rRNA) removal was performed using the RiboMinus Eukaryote Kit (Qiagen, Hilden, Germany). To prepare the RNA library for Illumina sequencing, we used the NEB Next Ultra Directional RNA Library Prep Kit for Illumina (NEB, MA, USA), according to the manufacturer’s recommendations.

Libraries were assessed for quality with an Agilent 2100 assay and quantified by quantitative PCR prior to Illumina sequencing. After quality control, clean reads were obtained, filtered, and mapped to the human reference genome hg19 using TopHat (version 2.1.0) [19] with the default parameters (max mismatches in seed alignment = 0). The expression level of each gene was calculated, and the results were normalized using the RPKM (Reads Per Kilobase of exon model per Million mapped reads) method. Differentially expressed genes were identified according to established methods [20]. To address the issue of multiple comparisons, the threshold p-value was determined by controlling the false discovery rate (FDR, Benjamini-Hochberg method). In our analysis, differentially expressed genes were defined as those with FDR ≤ 0.001 and a more than two-fold expression change. Cluster analysis of differentially expressed genes was performed using R. We also used gene ontology (GO) annotation and the KEGG pathway analysis, with FDR ≤ 0.001 considered to be significantly enriched.

Prediction of super-enhancer and transcription factor binding enrichment
To analyze methylation modifications in lncRNA genes near the enhancer, transcription factor enriched motifs were identified on the 5mC gain (Up) and loss (Down) regions using HO-MER (version 4.10), a de novo motif discovery algorithm. Hypo and hypermethylation enriched regions around the annotated transcription start sites (TSS) ± 3 kb with a signal density (normalized by the RPKM method) of reads above the background (P < 0.01, empirical test) were defined as candidate super-enhancers.

Survival analysis
Overall survival and disease-free survival time of patients with BC with different expression levels of LINC00574 and LINC00092 were analyzed using data obtained from the Cancer Genome Atlas (TCGA, https://cancergenome.nih.gov/). Survival analysis was analyzed on the GEPIA web server [21] (http://gepia.cancer-pku.cn/) with the default parameters. A log-rank test was used to compare differences in survival time between groups. P < 0.05 was considered to be statistically significant.

Results

Genome-wide distribution of 5mC in BC tissues
To visualize genome-wide changes in 5mC between tumor and normal tissues, MeDIP-seq was conducted. Sequencing reads matching tumor or normal groups were converted to density signals and depicted as a peak diagram on the basis of their genomic location. In total, MeDIP-seq identified 436 hypermethylated and 239 hypomethylated genes in BC tissues around the TSS region (Figure 1A). As shown in Figure 1B, the overall level of modification of intergenic regions was higher than the methylation level in regions 2 kb (Upstream 2 kb) upstream or downstream (Downstream 2 kb) of the transcriptional stop point. Notably, a sharp peak was observed in the 2 kb region upstream of the TSS (Figure 1B), which quickly dropped to its lowest point at the TSS. Determination of the distribution of 5mC methylation showed that 5mC was most enriched in the intron regions followed by CDS, but not in the UTR or gene body flanking regions (Figure 1D). Of the differentially 5mC methylated genes, >70% were coding genes, whereas few (~20%) were lncRNAs coding genes (Figure 1C) in both tumor and normal tissues.

Global change of lncRNA expression in BC tissue
Next, lncRNA expression profiles were analyzed in eight paired tumor and normal tissues by RNA-seq. As shown in Figure 2, 895 lncRNAs were differentially expressed in tumor tissue relative to normal tissue. Among these lncRNAs, 228 were upregulated and 667 were downregulated relative to normal tissue. The expression profile of tumor tissues thus differs from normal tissues (Figure 2A), as the expression signals in the heatmap show separate clustering of tumor and normal tissues. Furthermore, analysis of differentially expressed lncRNAs revealed that downregulated lncRNAs were distributed across all chromosomes in BC tissue, with the upregulated lncRNAs found on all but chromosomes 20 and 23 (Figure 2B).
To provide an overview of the potential role of the differential lncRNAs, GO and KEGG enriched analyses were conducted. As shown in Figure 3A, the four most enriched GO terms were cell, cell part, cellular process and binding, with 1201, 1200, 1125, and 1106 genes enriched in each category, respectively (Figure 3A). KEGG enrichment analysis revealed the top five enriched pathways were the pathways in cancer, cGMP-PKG signaling pathway, PI3K-Akt signaling pathway, focal adhesion, and Rap1 signaling pathway (Figure 3B).

Expression of lncRNA is regulated by 5mC and associated with BC prognosis

To explore the potential link between change in lncRNA expression and differential methylation status in BC, the preference of 5mC modification and profile of 5mC state on the location around TSS were analyzed. Figure 4A lists the five most enriched motifs of gain-5mC (UP) and loss-5mC (Down) loci. Comparison of locus-specific changes in 5mC in the super-enhancer region of lncRNA genes between tumor and normal control showed a reduced signal of 5mC modification in tumor tissue (Figure 4B).

The combination of MeDIP-seq and lncRNA sequencing identified 26 lncRNAs that were consistently hypermethylated and downregulated or hypomethylated and upregulated across all tumor tissues from the eight patients with BC (Table 1). To determine if there was an association between these lncRNAs and BC outcome, lncRNAs were subjected to GEPIA analysis. Indeed, two lncRNAs were found to be associated with BC survival. As shown in Figure 5A,
Cancer-associated methylated lncRNAs

Am J Transl Res 2019;11(6):3790-3800

Hypomethylated LINC00574 was associated with poor overall survival (HR = 1.7, \( p \)-value = 0.035) such that patients with BC with higher expression of LINC00574 displayed shorter overall survival time. However, no significant association was found between expression of LINC00574 and survival time in patients with BC (Figure 5A). In addition, hypomethylated LINC00092 showed no significant association with survival in patients with BC (Figure 5B). The association between lncRNA expression and progression of BC requires further verification.

Discussion

DNA methylation and lncRNA are two important epigenetic regulatory factors in human cancer. In the present study, we investigated the association between DNA methylation and BC-related lncRNAs expression using MeDIP-seq and RNA-seq data in eight patients with BC.

Intragenic DNA methylation has been shown to play a role in several gene regulation processes [22]. Regions with increased CpG content (known as CpG islands) are known to overlap with TSS in about 60-70% of genes in the human genome. By contrast, regions with low CpG modification are frequently located regions other than the TSS [22]. DNA methylation primarily occurs in CpGs, but may also occur in non-CpG regions [23]. Our MeDIP-seq revealed that the lowest level of methylation was at the TSS site. At the same time, there was a signal of a significant DNA methylation peak in the 2 kb region near the TSS.

The function of IncRNA in regulating tumor cell behavior - including cell proliferation, apoptosis, migration, and invasion - has been widely reported [24-28]. In BC, IncRNA is particularly involved in the regulation of proliferation, metastasis, and drug resistance [29-32]. In the present study, we demonstrated the overall expression profile of lncRNA in control tissues and eight patients with BC tumor. KEGG enrichment analysis of differentially expressed lncRNAs in tumor and control tissues revealed that these lncRNAs were significantly enriched in pathways in cancer, cGMP-PKG signaling pathway, PI3K-Akt signaling pathway, focal adhesion, and the Rap1 signaling pathway. Among these, the PI3K-Akt signaling pathway plays a role in tumor proliferation and growth and is involved in IncRNA regulating the occurrence and progression of various tumors [33-35].

We also integrated the expression pattern of IncRNA with their 5mC modification levels. To investigate whether DNA methylation regulation is involved in regulating differentially expressed IncRNA, we calculated the 5mC acqui-
Cancer-associated methylated lncRNAs

Figure 3. GO (A) and KEGG pathway enrichment (B) analysis of differentially expressed lncRNAs.
Cancer-associated methylated lncRNAs

3796

Am J Transl Res 2019;11(6):3790-3800

sition and loss motif characteristics on the lncRNA gene. The results showed that certain motifs of DNA 5mC were preferentially gained and lost. At the same time, in the vicinity of TSS, the DNA methylation level of the lncRNA gene super-enhancer region decreased. These findings suggest that expression of BC-associated lncRNA is related to the DNA methylation status near the TSS region of the lncRNA gene. We also observed lncRNAs with opposite changes in lncRNA expression and DNA modification, which suggests the potential of DNA methylation in BC to regulate lncRNA. Previous studies have identified epigenetic modifications of lncRNAs that exert either beneficial effects on tumors, such as MINCR, or suppressed effects, such as MEG3 [36]. By contrast, the global profile of differentially methylated lncRNA in BC is poorly understood. The current study has thus provided novel insight into the methylation modifications that control BC-associated lncRNA candidates.

Table 1. Methylated lncRNAs showing differential expression in BC tumor tissue

<table>
<thead>
<tr>
<th>lncRNA</th>
<th>Methylation status</th>
<th>IncRNA expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>LINC00839</td>
<td>down</td>
<td>up</td>
</tr>
<tr>
<td>LINC00301</td>
<td>down</td>
<td>up</td>
</tr>
<tr>
<td>LINC00111</td>
<td>down</td>
<td>up</td>
</tr>
<tr>
<td>VENTXP1</td>
<td>down</td>
<td>up</td>
</tr>
<tr>
<td>LINC00898</td>
<td>down</td>
<td>up</td>
</tr>
<tr>
<td>LINC00887</td>
<td>down</td>
<td>up</td>
</tr>
<tr>
<td>LINC00592</td>
<td>down</td>
<td>up</td>
</tr>
<tr>
<td>LINC00669</td>
<td>down</td>
<td>up</td>
</tr>
<tr>
<td>LINC00643</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC00473</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC00472</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>SNHG9</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>RBAKDN</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC01056</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC01003</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC00626</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC00606</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC00574</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC00486</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC00474</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC00339</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC00303</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC00092</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>LINC0028</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>EPB41L4A-AS1</td>
<td>up</td>
<td>down</td>
</tr>
<tr>
<td>C11orf95</td>
<td>up</td>
<td>down</td>
</tr>
</tbody>
</table>

Figure 4. 5mC is enriched at enhancer regions of lncRNA genes in patients with BC. A. Top five enriched motifs for gain-5mC (UP) and loss-5mC (Down) loci. B. Heat maps showing 5mC patterns in the 3 kb region surrounding the super-enhancers of lncRNA genes.

The functions of most lncRNAs have not been well characterized. However, LINC00643 has been reported to be differentially methylated and expressed in cancers [37] and serves as a tumor enhancer [38-40]. LINC00472 is regulated by DNA methylation in breast cancer and colorectal cancer, and is known to function as a tumor suppressor [41, 42]. Several reports have provided evidence for the roles of SNHG9, LINC00574, and LINC00092 in cancer [43,
We further revealed that differential expression of LINC00574 in BC compared to normal tissue and a gain in methylation of 5mC in BC tissue are associated with reduced survival time. These findings suggest that LINC00574 is a biomarker for BC prognosis. LINC00574 has not been previously reported to have a role in BC, although it has been shown to be upregulated in breast cancer and gliomas [45, 46]. In addition, its 5mC methylation status indicates a link between DNA methylation in regulating lncRNA expression and BC prognosis.

Taken together, the findings of the current study revealed a variety of methylation lncRNAs associated with BC. In particular, LINC00574 was associated with poor outcomes for patients with BC. These findings indicate a role for IncRNA regulation by 5mC methylation in BC prognosis.

Acknowledgements

This study was supported by the Finance Science and Technology Project of Hainan Province (Grant No. ZDYD2019163 and ZDKJ2017007), the Natural Science Foundation of Hainan Province (Grant No. 2017CXTD010), the National Natural Science Foundation of China (Grant No. 81760465, 81460450 and 81760461).
Cancer-associated methylated IncRNAs

Disclosure of conflict of interest

None.

Address correspondence to: Fei Wang, Department of Urology, People’s Hospital of Hainan Province, No. 19 Xiuhua Road, Xinying District, Haikou 570311, Hainan, China. Tel: +86-0898-68642242; Fax: +86-0898-68642242; E-mail: 1149528314@qq.com; Yang Xiang, Department of Urology, Affiliated Haikou Hospital of Xiangya Medical College, Central South University, Haikou 570208, Hainan, China. E-mail: xiangyang200611@126.com

References


Cancer-associated methylated lncRNAs


