Short-lived non-coding transcripts (SLiTs): Clues to regulatory long non-coding RNA

Hidenori Tani*

Environmental Management Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan.

Summary Whole transcriptome analyses have revealed a large number of novel long non-coding RNAs (lncRNAs). Although the importance of lncRNAs has been documented in previous reports, the biological and physiological functions of lncRNAs remain largely unknown. The role of lncRNAs seems an elusive problem. Here, I propose a clue to the identification of regulatory lncRNAs. The key point is RNA half-life. RNAs with a long half-life ($t_{1/2} > 4$ h) contain a significant proportion of ncRNAs, as well as mRNAs involved in housekeeping functions, whereas RNAs with a short half-life ($t_{1/2} < 4$ h) include known regulatory ncRNAs and regulatory mRNAs. This novel class of ncRNAs with a short half-life can be categorized as Short-Lived non-coding Transcripts (SLiTs). I consider that SLiTs are likely to be rich in functionally uncharacterized regulatory RNAs. This review describes recent progress in research into SLiTs.

Keywords: Non-coding RNA, RNA degradation, RNA decay, RNA-Seq, BRIC-Seq

1. Introduction

Recent transcriptome analyses have revealed thousands of intergenic, intronic, and cis-antisense long non-coding RNAs (lncRNAs) that are expressed from mammalian genomes. LncRNAs are defined as RNA molecules greater than 200 nucleotides in length that do not contain any apparent protein-coding potential (1-4). The majority of lncRNAs are transcribed by RNA polymerase II (Pol II), as evidenced by Pol II occupancy, 5' caps, histone modifications associated with Pol II transcriptional elongation, and polyadenylation (5). Although the importance of lncRNAs to processes such as transcriptional regulation, organization of nuclear structure, and post-transcriptional processing has been documented in previous reports (2,6,7), the biological and physiological functions of a great many lncRNAs remain largely unknown. Which lncRNA molecules should we target? Here, I present a clue to the selection of lncRNAs.

In 2012, two independent research groups reported that ncRNA half-lives vary over a wide range, and that they are comparable to those of mRNAs in mice and humans (8,9). Moreover, a genome-wide approach for determining RNA stability, which is called 5'-bromouridine (BrU) immunoprecipitation chase-deep assay (BRIC), or BRIC through deep sequencing (BRIC-Seq) (9-11), revealed that ncRNAs with short half-lives (RNA half-life $t_{1/2} < 4$ h) included known regulatory ncRNAs, such as HOX transcript antisense RNA (HOXAT), antisense noncoding RNA in the inhibitors of CDK4 locus (ANRIL)/CDKN2B antisense RNA 1 (CDKN2B-AS1), and growth arrest specific 5 (GAS5). BRIC-Seq has revealed 785 lncRNAs with short half-lives, termed Short-Lived non-coding Transcripts (SLiTs) (9), and I have selected 26 lncRNAs that are short-lived ($t_{1/2} < 4$ h) in HeLa-Tet-off cells (Table 1), longer than 200 nt, and fulfill the established criteria for lncRNA classification. In this review, we will describe SLiTs identified in studies to date.

2. MIR4435-2HG_v1 and v2

Yang et al. reported that MIR4435-2 host gene (MIR4435-2HG), also known as AK001796, is up-
regulated and acts as an oncogene in lung cancer tissues and cell lines (12). In MIR4435-2HG knockdown experiments, cell proliferation and growth were reduced in lung cancer cells and tumorigenesis was slowed, and cell-cycle arrest was observed with increased numbers of cells in G0/G1. Moreover, they also found that MIR4435-2HG is downregulated in resveratrol-treated lung cancer cells (12).

3. ANRIL/CDKN2B-AS1

Two independent groups reported that ANRIL, also known as CDKN2B-AS1, associates with and recruits polycomb repression complex (PRC)-1 and PRC-2 on the INK4 locus to repress the transcription of p15 and p16 (13,14). Knockdown of ANRIL increases p15 and p16 expression, causing inhibition of cell proliferation and cellular senescence in human fibroblasts (13,14). Higher levels of ANRIL are regarded as a risk factor in several types of human cancers, including hepatocellular carcinoma (15), lung cancer (16), ovarian cancer (17), gastric cancer (18), bladder cancer (19), and colorectal cancer (20). ANRIL also regulates a large number of genes related to gene expression, cell proliferation, cell adhesion, and apoptosis (21), suggesting that ANRIL is involved in various cellular processes.

4. LINC00473_v1 and v2

Reitmair et al. reported that expression of long intergenic non-protein coding RNA 473 (LINC00473_v1, also known as C6orf176), is cyclic adenosine monophosphate (cAMP)-mediated (22). cAMP mediates diverse cellular signals, including prostaglandin E2-mediated intraocular pressure-lowering activity in human ocular ciliary smooth muscle cells. Knockdown of LINC00473_v1 shows modulation of several cAMP-responsive genes. LINC00473_v1 is a potential biomarker and/or therapeutic target in the context of diseases linked to deregulated cAMP signaling.

5. LINC00152 / CYTOR

Recently, reports concerning long intergenic non-protein coding RNA 152 (LINC00152), also known as cytoskeleton regulator RNA (CYTOR), have been increasing. First, LINC00152 expression is increased in gastric cancer tissue (23). The expression level of LINC00152 in gastric carcinoma is significantly increased compared with matched normal tissue and normal mucosa from healthy controls. LINC00152 also acts as a novel biomarker in predicting diagnosis of hepatocellular carcinoma (24). Moreover, LINC00152 acts as an oncogene, because knockdown of LINC00152 inhibits cell proliferation and colony formation, promotes cell cycle arrest at G1 phase, triggers late apoptosis, reduces the epithelial to mesenchymal transition program, and suppresses cell migration and invasion (25). LINC00152 directly binds with epidermal growth factor receptor (EGFR).

---

Table 1. The 26 short-lived long ncRNAs that were investigated in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Another name</th>
<th>Accession No.</th>
<th>Length (nt)</th>
<th>t1/2*</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIR4435-2HG_v2</td>
<td>LINC0541471_v2</td>
<td>NR_024373</td>
<td>557</td>
<td>2.3</td>
</tr>
<tr>
<td>LOC100216545</td>
<td></td>
<td>NR_024586</td>
<td>3,615</td>
<td>2.3</td>
</tr>
<tr>
<td>FLJ34663</td>
<td>LINC-PINT</td>
<td>NR_015431</td>
<td>2,964</td>
<td>2.4</td>
</tr>
<tr>
<td>ANRIL</td>
<td>CDKN2B-AS1</td>
<td>NR_003529</td>
<td>3,857</td>
<td>2.4</td>
</tr>
<tr>
<td>FAM222A-AS1</td>
<td></td>
<td>NR_026661</td>
<td>1,178</td>
<td>2.4</td>
</tr>
<tr>
<td>LINCO0473_v1</td>
<td></td>
<td>NR_026860</td>
<td>1,832</td>
<td>2.4</td>
</tr>
<tr>
<td>MIR22HG</td>
<td></td>
<td>NR_028502</td>
<td>2,699</td>
<td>2.4</td>
</tr>
<tr>
<td>LINCO0473_v2</td>
<td></td>
<td>NR_026861</td>
<td>1,123</td>
<td>2.4</td>
</tr>
<tr>
<td>LINCO00152</td>
<td>CYTOR</td>
<td>NR_024204</td>
<td>828</td>
<td>2.4</td>
</tr>
<tr>
<td>HOTAIR</td>
<td></td>
<td>NR_003716</td>
<td>2,337</td>
<td>2.5</td>
</tr>
<tr>
<td>MIR4435-2HG_v1</td>
<td>LINC0541471_v1</td>
<td>NR_015395</td>
<td>809</td>
<td>2.5</td>
</tr>
<tr>
<td>TTN-AS1</td>
<td></td>
<td>NR_038271</td>
<td>2,033</td>
<td>2.6</td>
</tr>
<tr>
<td>GAS5</td>
<td></td>
<td>NR_002578</td>
<td>651</td>
<td>2.6</td>
</tr>
<tr>
<td>SNHG15</td>
<td></td>
<td>NR_003697</td>
<td>837</td>
<td>2.6</td>
</tr>
<tr>
<td>LINCO0662</td>
<td></td>
<td>NR_027301</td>
<td>2,085</td>
<td>2.7</td>
</tr>
<tr>
<td>HCG18</td>
<td></td>
<td>NR_024052</td>
<td>6,814</td>
<td>2.9</td>
</tr>
<tr>
<td>LOC728431</td>
<td>LINC01137</td>
<td>NR_038842</td>
<td>997</td>
<td>3.0</td>
</tr>
<tr>
<td>TUG1</td>
<td></td>
<td>NR_002323</td>
<td>7,115</td>
<td>3.2</td>
</tr>
<tr>
<td>LOC550112</td>
<td>UBA6-AS1</td>
<td>NR_015439</td>
<td>2,301</td>
<td>3.2</td>
</tr>
<tr>
<td>NEAT1_v1</td>
<td></td>
<td>NR_028272</td>
<td>3,756</td>
<td>3.3</td>
</tr>
<tr>
<td>LINCO0667</td>
<td></td>
<td>NR_015389</td>
<td>3,979</td>
<td>3.4</td>
</tr>
<tr>
<td>GABPB-AS1</td>
<td></td>
<td>NR_024490</td>
<td>4,139</td>
<td>3.4</td>
</tr>
<tr>
<td>LINCO1184</td>
<td>FLJ33630</td>
<td>NR_015360</td>
<td>2,977</td>
<td>3.5</td>
</tr>
<tr>
<td>ZFP91-CNTF</td>
<td></td>
<td>NR_024091</td>
<td>3,544</td>
<td>3.6</td>
</tr>
<tr>
<td>NEAT1_v2</td>
<td></td>
<td>NR_131012</td>
<td>22,743</td>
<td>3.7</td>
</tr>
<tr>
<td>ID2-AS1</td>
<td></td>
<td>NR_024628</td>
<td>1,107</td>
<td>3.7</td>
</tr>
</tbody>
</table>

*These values are taken from a previous report (Tani 2012)
which causes activation of Phosphoinositide 3-kinase (PI3K)/AKT serine/threonine kinase (AKT) signaling (26). LINC00152 is involved in the oncogenesis of hepatocellular carcinoma by activating the mechanistic target of the rapamycin (mTOR) signaling pathway (27).

6. HOTAIR

The most thoroughly studied representative of the modulation chromatin state of lncRNAs is HOTAIR (28). HOTAIR is transcribed within the homeobox C (HOXC) locus, and interacts with the catalytic subunit of PRC2 and enhancer of zeste homolog 2 (EZH2) (29,30). Knockdown of HOTAIR causes a reduction in PRC2 occupancy, a local decrease in H3K27 trimethylation, and the activation of genes within the homeobox D (HOXD) HOXD locus on chromosome 2 (29). HOTAIR acts as a molecular scaffold to connect PRC2 and LSD1-containing repressive complexes (30). While the 5’ end of HOTAIR is required for interaction with PRC2, its 3’ end has been shown to interact with the H3K4-demethylase LSD1 complexes (30). This interaction results in the physical bridging of a small subfraction of PRC2 and LSD1-containing repressive complexes, with HOTAIR knockdown causing loss of either or both complexes at a subset of target genes (30). Recently, RNA-chromatin immunoprecipitation (ChIP) and chromatin Isolation by RNA Purification (ChIRP) reveals that HOTAIR is regulated by estrogens and able to control estrogen receptors (ERs) function by interacting with estrogen receptor (ER)α/ERβ, and HOTAIR is present on pS2, human telomerase reverse transcriptase (hTERT) and HOTAIR promoters at the estrogen response elements (ERE)/endothelial nitric oxide synthase (eNOS) peaks (31).

7. GAS5

GAS5 was originally isolated from a screen for potential tumor suppressor genes expressed at high levels during growth arrest (32). The human GAS5 gene is a multiple small nucleolar RNA (snoRNA) host gene that encodes 10 box C/D snoRNAs within 11 introns, and has been classified as a member of the 5’-terminal oligopyrimidine tract (5’ TOP) gene family, characterized by an upstream oligopyrimidine tract sequence (33). GAS5 transcript abundance is increased during growth arrest induced by either serum starvation or treatment with translation inhibitors (34). GAS5 functions as a starvation- or growth arrest-linked riborepressor for the glucocorticoid receptor (GR) by binding to the DNA-binding domain of the GR, acting as a decoy glucocorticoid response element (GRE), thus competing with DNA GREs for binding to the GR (35). The degradation pathway of GAS5 regulates GAS5 function, which modulates the apoptosis-related genes cellular inhibitor of apoptosis protein 1 (cIAP2) and serum/glucocorticoid regulated kinase 1 (SGK1) (36).

8. TUG1

Taurine up-regulated 1 (TUG1) causes growth-control genes to relocate from the repressive environment of Polycomb bodies, where they interact with co-repressor complexes, to the gene activation milieu of the interchromatin granules, by selectively interacting with methylated and unmethylated polycomb 2 proteins present on growth-control gene promoters (37,38). TUG1 is upregulated by p53 upon DNA damage in p53 wild-type, but not p53 mutant cells. TUG1 also serves as a diagnostic biomarker and therapy target for hepatocellular carcinoma and promotes cell growth and apoptosis by epigenetic silencing of kruppel like factor 2 (KLF2) (39).

9. NEAT1_v1 and v2

Nuclear paraspeckle assembly transcript 1 (NEAT1) has been found to localize specifically to paraspeckles where it forms an essential structural component (40-42). Recently, two independent research groups reported that the NEAT1 - splicing factor proline and glutamine rich (SFPQ) interaction plays roles in both repression and activation of genes, which likely depend on the context of the promoter sequence or interplay with other transcriptional factors (43-44). Hirose et al. reported the role of NEAT1 in transcriptional regulation through sequestering of SFPQ from the RNA-specific adenosine deaminase B2 (ADARB2) gene in response to proteasome inhibition (43). Imamura et al. reported that NEAT1 expression is induced by infection with the influenza virus or herpes simplex virus. This upregulation of NEAT1 results in relocation of SFPQ, a NEAT1- binding paraspeckle protein and repressor of interleukin 8 (IL8) transcription, from the IL8 promoter to the paraspeckles, leading to transcriptional activation of IL8 (44).

10. Other lncRNAs

The functions of other lncRNAs remain to be elucidated; however, we have found that lncRNAs highly and rapidly respond to chemical stresses as follows. In HeLa Tet-off cells, six SLiTs [MIR22 host gene (MIR22HG), GABPB1 antisense RNA 1 (GABPB1-AS1), LINC00152, IDI2 antisense RNA 1 (IDI2-AS1), small nucleolar RNA host gene 15 (SNHG15), and LINC01184] respond to chemical stressors (cisplatin, cycloheximide, or mercury II chloride) (45). In human-induced pluripotent stem cells (hiPSCs), six novel lncRNAs (ANRIL, MIR22HG, GABPB1-AS1, LINC01184, LINC00152, and LINC0541471_v2) respond to chemical stressors (cycloheximide, hydrogen peroxide, cadmium, or arsenic) (46).
11. Concluding remarks

Although the biological and physiological functions of some SLiTs have been documented, those of others remain unknown. We have found unknown SLiTs that respond strongly to chemical stresses; thus, we can use these SLiTs as surrogate indicators of chemical stress responses in human cells. I believe that the relationships of the unknown SLiTs in this review and RNA-binding proteins will be determined in the future.

References


(Received January 11, 2017; Revised February 2, 2017; Accepted February 3, 2017)