Genomics Proteomics Bioinformatics 14 (2016) 73–80

HOSTED BY

ELSEVIER

Genomics Proteomics Bioinformatics

www.elsevier.com/locate/gpb
www.sciencedirect.com

REVIEW

Long Non-coding RNAs in the Cytoplasm

Farooq Rashid [#,a], **Abdullah Shah** [#,b], **Ge Shan** [*,c]

CAS Key Laboratory of Innate Immunity and Chronic Disease, CAS Center for Excellence in Molecular Cell Science, School of Life Sciences, University of Science and Technology of China, Hefei 230027, China

Received 8 January 2016; revised 3 February 2016; accepted 2 March 2016
Available online 6 May 2016

Handled by Er-Wei Song

KEYWORDS

lncRNA;
mRNA stability;
mRNA translation;
ceRNA;
MicroRNA

Abstract An enormous amount of long non-coding RNAs (**lncRNAs**) transcribed from eukaryotic genome are important regulators in different aspects of cellular events. Cytoplasm is the residence and the site of action for many **lncRNAs**. The cytoplasmic **lncRNAs** play indispensable roles with multiple molecular mechanisms in animal and human cells. In this review, we mainly talk about functions and the underlying mechanisms of **lncRNAs** in the cytoplasm. We highlight relatively well-studied examples of cytoplasmic **lncRNAs** for their roles in modulating **mRNA stability**, regulating **mRNA translation**, serving as competing endogenous RNAs, functioning as precursors of **microRNAs**, and mediating protein modifications. We also elaborate the perspectives of cytoplasmic **lncRNA** studies.

Introduction

Mammalian genome is pervasively transcribed into many different complex families of RNA. However, less than 2% of mammalian genome is transcribed into mRNA to encode proteins, whereas a major portion of the genome is transcribed into interweaved and overlapping transcripts that include thousands of non-coding RNA (ncRNA) transcripts [1,2]. ncRNAs more than 200 nucleotides in length are called long ncRNAs (lncRNAs), which are often transcribed by RNA polymerase II [3,4]. These lncRNAs are usually devoid of open reading frames (ORFs), with or without the 3′ polyadenylation [5–8]. Interestingly, expression of lncRNA is more tissue-specific than that of mRNA [9].

In the last several years, a large number of nuclear lncRNAs have been discovered. These lncRNAs play diverse roles in the nucleus through various mechanisms [10]. For example, nuclear lncRNAs control the epigenetic state of particular genes [11], participate in transcriptional regulation [12], get involved in alternative splicing and constitute subnuclear compartments [13,14].

Although for most if not all of the lncRNAs, nucleus is the place of biogenesis and processing, cytoplasm is the final residence and site of action for some lncRNAs. Biogenesis of lncRNAs is quite complicated and share many features of protein-coding RNAs. Within the nucleus, they occupy the chromatin fraction. 17% of lncRNAs *vs.* 15% of mRNAs are enriched in the nucleus, whereas 4% *vs.* 26%, respectively, are enriched in the cytoplasm [6]. Many lncRNA-mediated mechanisms of gene regulation have been identified in the

* Corresponding author.
E-mail: shange@ustc.edu.cn (Shan G).
# Equal contribution.
[a] ORCID: 0000-0003-1017-7769.
[b] ORCID: 0000-0003-0130-9133.
[c] ORCID: 0000-0002-3561-2088.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

http://dx.doi.org/10.1016/j.gpb.2016.03.005

cytoplasm [8,15,16]. In the last decade or so, thousands of cytoplasmic lncRNAs have been discovered, indicating their importance for multiple cellular activities. In this review we highlight the functions and underlying mechanisms of some important cytoplasmic lncRNAs that are responsible for post-transcriptional regulations such as on mRNA stability and translational control.

Modulation of mRNA stability

In the cytoplasm, several lncRNAs target mRNA transcripts and modulate mRNA stability. Some lncRNAs such as half-STAU1-binding site RNAs (*1/2-sbsRNAs*) and growth arrested DNA-damage inducible gene 7 (*gadd7*) decrease the stability of mRNA, while others such as antisense transcript for β-secretase 1 (*BACE1-AS*) and the terminal differentiation-induced ncRNA (*TINCR*) increase mRNA stability.

1/2-sbsRNAs

mRNAs can be degraded via staufen 1 (STAU1)-mediated mRNA decay (SMD), when their 3′ untranslated region (3′ UTR) binds to STAU1 [17]. STAU1 is a double-stranded RNA (dsRNA)-binding protein, which binds within 3′ UTR of translationally-active mRNA [14,18]. STAU1 binds to a complex structure of 19-bp stem with 100 nt apex within the mRNA encoding ADP ribosylation factor 1 (ARF1) [18]. This stem region is conserved in 3′ UTRs of *ARF1* mRNA of mouse and rat [6]. However, such stem structures were not identified in other STAU1 targets [18]. STAU1-binding sites can be formed by imperfect base-pairing between an Alu element in the 3′ UTR of an SMD target and another Alu element in a cytoplasmic, polyadenylated lncRNA [18]. These lncRNAs transactivate the binding of STAU1 to mRNA as only STAU1 could be immunoprecipitated with lncRNAs called *1/2-sbsRNAs*, thus unveiling a pivotal strategy of recruiting proteins to mRNAs and mediating the mRNA decay (**Figure 1A**). However, not all mRNAs containing Alu element in their 3′ UTR are targeted for SMD, despite the presence of complementary 1/2-sbsRNAs that target other mRNA for SMD [17]. One of the 378 identified *1/2-sbsRNAs* in humans, *1/2-sbsRNA1* contains a single Alu element that base pairs with the Alu element in the 3′ UTR of plasminogen activator inhibitor type 1 (*SERPINE1*) and *FLJ21870*. 1/2-sbsRNA1 is present in the cytoplasm but absent in the nucleus of HeLa cells [18] and only STAU1 can be immunoprecipitated with *1/2-sbsRNA1*. Two isoforms of 1/2-sbsRNA1, including *1/2- sbsRNA1*(*S*) (short form) and *1/2-sbsRNA1*(*L*) (long form), have been reported. Both isoforms contain the Alu element and 3' UTR with poly (A) tail, although they differ at the 5′ end. Knocking down *1/2-sbsRNA1*(*S*) increased the level of *SERPINE1* and *FLJ21870* mRNAs by 2–4.5-folds above normal. Other *1/2-sbsRNA* members such as 1/2-sbsRNA2, 1/2-sbsRNA3, and 1/2-sbsRNA4 are largely cytoplasmic and polyadenylated as well, containing a single Alu element. Knocking down these 1/2-sbsRNAs led to upregulation of their mRNA targets [17]. Functional studies showed that *1/2-sbsRNA1* contributed to the reduced cell migration by targeting *SERPINE1* and RAB11-family-interacing protein 1 (*RAB11FIP1*) mRNAs for SMD as confirmed by scarp injury repair assay [17].

gadd7

gadd7 is a 754-nt polyadenylated lncRNA isolated from Chinese hamster ovary (CHO) cells [15,16]. Expression of *gadd7* is induced by several types of DNA damage and growth arrest signals [19,20], and *gadd7* plays a pivotal role in regulating G1/S checkpoint post DNA damage. *gadd7* also regulates lipid-induced oxidative and endoplasmic reticulum (ER) stress [21]. lncRNAs are known to bind to and regulate the functions of proteins. One such example is the binding of *gadd7* with TAR DNA-binding protein (TDP-43), and this interaction is strengthened upon UV exposure [22–24]. TDP-43 is a member of the heterogeneous nuclear ribonucleoprotein (hnRNP) family. HnRNP family members are RNA/DNA binding proteins involved in transcription, splicing, mRNA transport, and mRNA stability [25,26]. TDP-43 is known to repress the expression of cyclin-dependent kinase 6 (*Cdk6*) mRNA in Hela cells, which is important for G1-phase progression [27,28]. Nonetheless, *Cdk6* expression is found to be activated by TDP-43 in CHO cells [24]. UV-induced *gadd7* directly interacts with TDP-43, thus leading to the decreased interaction between TDP-43 and *Cdk6* mRNA. This results in *Cdk6* mRNA degradation, and finally inhibition of cell cycle progression [24]. *gadd7* is not highly conserved at the nucleotide level [29,30]. Since the structure or the functional motif of lncRNAs may be more important, and thus would be more conserved than their nucleotide sequence [9], it remains possible to identify a functional *gadd7* ortholog in humans. This may be important for unveiling the pathogenesis of diseases such as frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS), as dominant mutations in TDP-43 are causative of these two important neurodegenerative diseases [24,31,32].

BACE1-AS

Expression of the conserved non-coding *BACE1-AS* increases *BACE1* mRNA stability when HEK-SW cells are exposed to cellular stressors like amyloid-β1–42 (Aβ1–42) [33]. *BACE1-AS* renders *BACE1* mRNA stability by masking the binding site of miR-485-5p (Figure 1A). *BACE1-AS* and miR-485-5p compete for binding in the sixth exon of *BACE1* mRNA. The sense-antisense RNA duplex between *BACE1* and *BACE1-AS* in the cytoplasm potentially perturb the interaction between miR-485-5p and *BACE1* mRNA, which to some extent, explains the mRNA stabilization by *BACE1-AS* transcript [34].

TINCR

The *TINCR* gene resides on chromosome 19 in humans and encodes a predominantly cytoplasmic 3.7-kb lncRNA. TINCR regulates human epidermal differentiation by post transcriptional mechanism [35]. Previously found as an uncharacterized lncRNA, *TINCR* is now believed to be the most highly-induced lncRNA during epidermal differentiation [35,36]. *TINCR* binds to mRNA through a 25-nt 'TINCR box' motif, which is robustly enriched in the interacted mRNAs. *TINCR* RNA has a strong affinity for STAU1 protein [17,35,37,38]. *TINCR*–STAU1 complex mediates the stabilization of differentiation-related mRNAs, such as *KRT80* encoding