Long noncoding RNA and its role in virus infection and pathogenesis

Kaining Yi¹, Yiling Zhang¹, Yong Wang¹, Zhilong Wang¹, Mengjia Xie¹, Zhigang Jin¹, Tiejun Zhao¹,²

¹College of Chemistry and Life Sciences, Zhejiang Normal University, Jinhua, Zhejiang, China,
²Key Laboratory of Wildlife Biotechnology and Conservation and Utilization of Zhejiang Province,
Zhejiang Normal University, Jinhua, Zhejiang, China

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. LncRNAs characterization, biogenesis, and processing
4. Regulatorly function of IncRNAs
5. Viruses and cellular IncRNAs
6. Virus-encoded IncRNAs
7. LncRNAs and viral diseases
8. Perspective
9. Acknowledgements
10. References

1. ABSTRACT

Long noncoding RNAs (lncRNAs) are RNA molecules longer than 200 nucleotides that regulate gene expression at the transcriptional and post-transcriptional levels. Emerging evidence showed that lncRNAs play important roles in a wide range of biological processes, including cell proliferation, apoptosis, and tumorigenesis. The infection of virus generally can regulate the expression of the cellular lncRNA expression. The lncRNAs which encoded by virus can also modulate the expression of the hosts’ gene which is critical for virus infection. Here, we summarized the recent progress on long noncoding RNA and its relationship with virus, especially the function of long noncoding RNA on virus-induced oncogenesis. Studies on lncRNAs and their relationship with viruses may give new insights into virus-host interactions and therapy of related diseases.

2. INTRODUCTION

Based on the findings from the Human Genome Project (HGP) and the latest high-throughput sequencing technology, it became obvious that a small portion (less than 2%) of the human genome is used to encode about 25,000 protein-encoding genes. The remaining human genome (approximately 98%) is transcribed into a large number of noncoding RNAs (noncoding RNA, ncRNA), among which long noncoding RNAs are the most frequently transcribed. It is suggested that long noncoding RNAs (lncRNAs) may play significant functional roles (1, 2).

LncRNAs regulate gene expression at a number of levels, including epigenetic regulation, transcriptional and post-transcriptional regulation (3). Increasing evidences have showed that lncRNAs play critical roles in a wide range of biological processes, such as stem cell maintenance, cell homoeostasis, differentiation, proliferation and apoptosis. In addition, the differential expression of lncRNAs has recently been linked to cancer and other diseases (4-6). Viral infection is known to induce abnormal lncRNA expression in the host cell. In addition, the lncRNAs encoded by viral genomes can also
modulate the expression of the host genes, which is critical for the infection of virus (7, 8).

The current review focuses on the role of lncRNAs, and their interactions with viruses.

3. LncRNAs CHARACTERIZATION, BIOGENESIS, AND PROCESSING

LncRNAs are noncoding RNA molecules longer than 200 nucleotides that do not code protein or only encode very short polypeptides (3). H19 and XIST were the first discovered lncRNAs (9, 10). Subsequently, in 2002, a large number of lncRNAs were identified by Japanese researchers through large-scale sequencing analysis of 60,770 full-length complementary DNAs (cDNAs) libraries (11).

According to the relative positions of lncRNAs and protein-encoding mRNAs in the genome, lncRNAs can be divided into five types: intergenic lncRNAs, intronic lncRNAs, bidirectional lncRNAs, sense-overlapping lncRNAs and antisense lncRNAs (12) (Figure 1). Similar to endogenous short noncoding RNAs, such as miRNAs, siRNAs, snoRNAs and piRNAs, lncRNAs are usually expressed in a tissue specific manner (13). However, lncRNAs are less conserved among species than endogenous short noncoding RNAs. The secondary structure or tertiary structure of lncRNAs is an important factor that influences lncRNAs normal biological function (3).

The majority of lncRNAs are transcribed by RNA polymerase II, and they are 5’-capped, spliced, and polyadenylated to form a structure similar to mRNA. The structures of lncRNAs are diverse and may or may not contain 5’ and 3’ terminal structures. Because they do not contain reading frame, lncRNAs do not encode proteins (14). At present, the biogenesis of lncRNAs has not been fully elucidated, but based on previous studies, the following five types of lncRNAs biogenesis have been identified: 1) the open reading frame structure of a protein-coding gene is mutated so that the residual coding sequence is merged and transformed into a lncRNA with certain regulatory functions; 2) under the action of chromatin recombination, two distant non-transcribed fragments are arranged side by side, forming a lncRNA with multiple exons; 3) they are produced during reverse transposition, in which noncoding genes produce functional lncRNAs during replication by reversing transcription; 4) local randomized replication leads to the appearance of adjacent repeat sequences in a noncoding RNA gene, resulting in lncRNA formation; and 5) genes are inserted into transcriptional elements to produce functional lncRNAs (15).

4. REGULATORY FUNCTION OF LncRNAs

LncRNAs were initially thought to be spurious transcriptional noise resulting from low fidelity of RNA polymerase. However, later research revealed that lncRNAs have a relatively long nucleotide chain that can be folded to form specific and complicated secondary structures, providing sites for protein binding. Moreover, lncRNAs can associate with DNA and RNA through complementary base
LncRNA and virus

pairing, and dynamically fulfill their biological functions (16). In recent years, the existence of a large and complicated LncRNA-mediated gene expression regulation network has been identified in vivo (17) (Figure 2).

LncRNAs regulate local protein-coding gene expression at the level of chromatin remodelling, transcriptional control and post-transcriptional processing. LncRNAs epigenetically modify gene expression through recruiting chromatin modifying complexes to specific genomic loci and impart their catalytic activity (18) (Figure 2A). In addition, LncRNAs regulate gene transcription via recruiting transcription factors to their target gene promoters, therefore activating gene expression (19) (Figure 2B). However, they can also block binding of general transcription factors, potentially via formation of RNA-DNA-Triplexes (20) (Figure 2C). In addition, LncRNAs antisense transcripts can pair to their specific sense RNA, and interfere with mRNA cleavage at the transcription level (21) (Figure 2D).

5. VIRUSES AND CELLULAR LncRNAs

Since the functions of LncRNAs are highly pleiotropic, it is not surprising that LncRNAs may be involved in virus replication. Viruses with limited coding ability have been found to utilize cellular LncRNA to regulate the expression and function of both host and viral genes.

A variety of animal viruses, such as severe acute respiratory syndrome coronavirus (SARS-CoV) (8), herpes simplex virus (24), Marek’s disease virus (25), human immunodeficiency virus (HIV) (26), avian
leukemia virus (27), and hepatitis B virus (HBV) (28) have been shown to dysregulate the expression of host lncRNAs. Further studies have shown that viral genes regulate the expression of protein-encoding genes by regulating the levels of cellular lncRNAs, thus ultimately promoting viral infection. A number of lncRNAs that are abnormally expressed in host cells during viral infection and their roles in cellular processes are shown in Figure 3.

Hepatitis B virus (HBV) infection is a major cause of hepatocellular carcinoma (HCC). Recent research has shown that some lncRNAs play essential roles in the processes of viral replication and viral infection in hepatocytes. Microarray analysis revealed that 174 lncRNAs were highly expressed in hepatocellular carcinoma cells and that 712 lncRNAs were downregulated (28). Among them, the lncRNA HULC (Highly Upregulated in Liver Cancer) and HEIH (Highly Expressed In HCC) have been reported to play important roles in HBV-related carcinogenesis.

LncRNA HULC is a noncoding RNA of 500 nucleotides and has a structure similar to mRNA. In HCC cells, Hepatitis B virus X protein (HBx) directly interacted with cAMP responsive element binding protein (CREB) and binds to the HULC promoter, thereby induce the expression of HULC. Moreover, high level of HULC inhibits the expression of the tumor suppressor gene p18 to promote the proliferation of hepatoma cells (22, 29, 30) (Figure 3A). The lncRNA HEIH, which is highly expressed in HBV-related HCC, has been reported to play a key role in cell cycle arrest at the G0/G1 phase by binding to the enhancer of Zeste homolog 2 (EZH2), and inhibits the expression of the target gene EZH2 by H3K27 methylation modification (31) (Figure 3B).

After Japanese encephalitis virus (JEV) infection, the murine central nervous system (CNS) has been reported to produce 3.1.8. kb virus inducible noncoding (VINC) transcripts of lncRNA NEAT1 (32). A recent study revealed that NEAT1 can bind to the
RNA-binding proteins P54nrb, PSFand PSP1, and form a subcellular nuclear structure named Paraspeckles (33, 34). Paraspeckles complex was suggested to be involved in nuclear RNA retention, resulting in the inhibition of translation (35). Nakagawa et al (36) found that knockdown of NEAT1 did not influence the expression of the Paraspeckles components, but altered the distribution of these proteins in the nucleus, thus affecting the formation of Paraspeckles (37) (Figure 3C).

Recent studies have shown that HIV-1 infection could also upregulate lncRNA NEAT1 expression (26). Knockdown of NEAT1 enhances virus production through increased nuclear to cytoplasmic export of Rev-dependent INS-containing HIV-1.

A latest study has showed that NEAT1 enhances Hantaan virus (HTNV)-induced IFN responses by acting as positive feedback for retinoic acid-inducible gene I (RIG-I) signaling, and NEAT1 exerts anti-hantaviral effects by modulating host innate immune responses (38). Apart from NEAT1, recent report showed that the lncRNA uc002yug.2 activates HIV-1 latency through regulating the mRNA level of various RUNX1 isoforms and increasing Tat expression (39).

LncRNA BIC was first identified as a lncRNA upregulated by avian leukosis virus (ALV) infection. After ALV infection, the ALV proviral can integrated into the promoter region of BIC RNA, thereby enhancing the expression of lncRNA BIC (27). Moreover, previous studies have shown that lncRNA BIC can be considered as a precursor of oncogenic miR-155, which is induced by several oncogenic viruses, such as Epstein-Barr virus, Hepatitis C virus and reticuloendotheliosis virus strain T (REV-T). Finally, lncRNA BIC inhibits the expression of target genes at the post-transcriptional level through miRNA-155, resulting in the development of leukemia (40) (Figure 3D). Sunantha et al (41) provided evidence that Kaposi’s sarcoma-associated herpesvirus (KSHV) deregulates host lncRNAs in both miRNA-independent fashion through latency-associated proteins, and miRNA-dependent fashion by direct interaction.

Of note, Wang et al (42) have found that lncRNA-ACOD1 promotes viral replication at the metabolic level by enhancing the catalytic activity of metabolic enzyme GOT2 (Glutamic-Oxaloacetic Transaminase 2), significantly promoting viral infection through IFN-I/IRF3-independent pathways. This new mechanism for the lncRNA-mediated regulation of cell metabolism explains the viral metabolic regulation and provides a potential therapeutic target for the development of new antiviral drugs.

Recent studies have shown that some lncRNAs are also involved in viral infection by regulating the host’s innate immune system, and they play an important role in host antiviral responses. For instance, lncRNA#32 positively regulates the expression of ISGs and the antiviral effect of type I IFN through interactions with hnRNPU and ATF2. Therefore, lncRNA#32 plays an important role in antiviral immunity (43). In addition, emerging evidence indicates that lncRNA ITPRIP-1 is involved in MDA5 activation to enhance the innate immune response to hepatitis C virus infection (44).

6. VIRUS-ENCODED LncRNAs

Accumulating data support the fact that lncRNA expression can regulate a large quantity of transcription factors (TFs) and proteins, which is important for both viral replication and activation of viral latency. Numerous studies have shown that cellular lncRNA expression can be regulated by virus infection. To date, several lncRNAs encoded by viruses have also been discovered (Table 1). Moreover, viruses use their own encoded lncRNA to facilitate viral replication or latency.

β2.7., a 2.7. kb transcript, is a highly conserved lncRNA of human cytomegalovirus (HCMV). It is mainly localized in the cytoplasm, and accounts for more than 20% of total viral gene transcription (45). Initial study by McSharry et al (46) showed that β2.7. is not
essential for virus replication in vitro. However, following studies indicated that β2.7. can bind to the mitochondrial enzyme complex I and stabilize mitochondrial membrane potential. This binding protects virus-infected cells from apoptosis and results in continued adenosine triphosphate (ATP) production, which is critical for the successful completion of the viral life cycle (47). Recently, a study further confirmed that lncRNA β2.7. is involved in the host cell apoptosis response, as β2.7. was found to inhibit the apoptosis of rat aortic endothelial cells in ischemia or reperfusion injury (48).

Virus-encoded lncRNAs regulate viral replication not only by RNA-protein interaction, but also by RNA-RNA interaction. The 3′ untranslated region (UTR) of the flavivirus RNA genome contains many secondary stem-loop II structures. These structures protect them from degradation by the nuclease XRN1 and results in the production of the functional lncRNA, sfRNA (49). It was reported that sfRNA played critical role in the regulation of flavivirus genome replication efficiency. During flavivirus infection, the innate immune response of the host could produce miRNA and destroys the viral genomic RNA. However, sfRNA bound to the miRNAs and induced the degradation of these miRNAs to protect viral genomic RNA, thus allowing the replication of virus in the host cell (50).

Viral lncRNA can act as an epigenetic regulator and play a pivotal role in the regulation of gene expression. For example, during HIV infection, the HIV genome encodes antisense RNA transcripts that inhibit the expression of viral genes (51). It has been reported that the 5′ LTR (Long Terminal Repeats) of HIV encodes an antisense lncRNA aspro5, which regulates viral transcription through epigenetic modification. Silencing aspro5 expression in CD4+ cells by using siRNA could activate the expression of HIV genome.
Further studies indicated that lncRNA aspro5 can recruit DNA methyltransferase DNMT3a, histone deacetylase HDAC-1, and histone methyltransferase EZH2 to the 5' UTR of HIV genome, and induce H3K9me2, H3K27me3, and histone deacetylation. Finally, it leads to the inhibition of the viral promoter activity, and suppression of viral genes expression (52).

It is well known that chimeric lncRNAs, which are generated by integration of proviral genome into the host genome, serve as another subclass of viral lncRNA. HBV induces hepatocellular carcinoma mainly through integration of HBV DNA into the host genome. Recently, transcriptome sequencing revealed that HBV DNA insertion site in the short arm of chromosome 8 produced a repeat sequence LINE1. It was transcribed as chimeric lncRNA HBx-LINE1. Further study showed that the expression of HBx-LINE1 promoted the transformation of cells from the epithelial phenotype to the interstitial phenotype, thus increasing the degree of malignancy. Moreover, HBx-LINE1 can activate Wnt/β-catenin signaling and promote β-catenin translocation to the nucleus, resulting in enhanced cell migration (53).

7. LncRNAs AND VIRAL DISEASES

Accumulating evidence shows that the lncRNA expression was dysregulated in virus induced diseases, indicating that lncRNAs may be involved in the development of viral diseases.

Studies have shown that the lncRNA NRAV which is a negative regulator of antiviral innate immune responses, was highly expressed in the cells infected with influenza A virus (IAV). NRAV regulated the expression of MxA and IFITM3 through histone modification of these genes. In addition, NRAV could specifically bind with the transcription factor ZONAB and regulate MxA transcription. In vivo study indicated that transgenic mice which express human NRAV were more susceptible to influenza virus infection. Moreover, the IAV titer in the lungs was much higher than control mice, resulting in higher mortality rates than in wild-type mice. Together, these studies suggested that influenza virus can inhibit interferon-stimulated gene transcription and regulate the antiviral response of host cells by enhancing the expression of the lncRNA NRAV (54). The lncRNA MVIH was also found to be highly expressed in HBV-infected liver tissues. Moreover, MVIH can promote tumor-induced neovascularization by inhibiting the secretion of phosphoglycerate kinase PGK1, thus eventually accelerating tumor growth and intrahepatic metastasis (55).

Some viruses utilize lncRNAs to support virus replication, and ultimately lead to pathogenesis of virus-related diseases. In contrast, host also utilizes lncRNAs as a defense to virus infection. Accumulating evidence shows that viruses regulate the proliferation of cells by its associated lncRNA. Huang et al (56) identified a lncRNA Dreh (down-regulated expression by HBx), which can inhibit HCC growth and metastasis in vitro and in vivo, act as a tumor suppressor in the development of HBV-related HCC. Additionally, lncRNA-Dreh could combine with the intermediate filament protein vimentin and repress its expression, and thus further change the normal cytoskeleton structure to inhibit tumor metastasis.

Another study found an interferon-stimulated lncRNA called Inc-ISG20 acts as a competing endogenous RNA (ceRNA) to bind miR-326 to reduce its inhibition on translation of antiviral gene ISG20, by which Inc-ISG20 inhibits IAV replication (57).

8. PERSPECTIVE

As a class of important molecules involved in gene expression regulation, lncRNAs represent a novel gene expression regulatory network. LncRNAs function as effective molecular regulatory switches in viruses and host cells. Here, we have focused on a few examples of lncRNAs, including cellular and viral lncRNAs.
LncRNA and virus

Our brief survey of advances in lncRNA research shows that viruses regulate the expression of viral genes and host cell genes by interacting with DNA, RNA, and proteins during the process of virus infection, and as a result, viruses are supported by both self-encoded lncRNAs and host lncRNAs. Therefore, further studies on the virus-host lncRNA relationship are needed to provide biomedical researchers with novel information and a broad knowledge base, which will provide new perspectives for the study of viral infections and related diseases.

9. ACKNOWLEDGEMENTS

This work was supported by a grant from the National Natural Science Foundation of China (No. 31470262 to T.Z.), a grant from Zhejiang Province Public Welfare Technology Application Research Project (China) (No. 2015C33149 to T.Z.).

10. REFERENCES


Key Words: Long Noncoding, RNA, Virus, Review

Send correspondence to: Tiejun Zhao, College of Chemistry and Life Sciences, Zhejiang Normal University, 688 Yingbin Road, Jinhua, Zhejiang 321004, China, Tel: 86-579-82291410, Fax: 86-579-82282269, E-mail: tjzhao@zjnu.cn