Critical Review

Research Progress of Long Noncoding RNA in China

Yu-Hang Xing[1,4]*
Zhiqiang Bai[1]*
Chu-Xiao Liu[1,4]
Shi-Bin Hu[1,4]
Meihua Ruan[2]
Ling-Ling Chen[1,3,4]

[1]State Key Laboratory of Molecular Biology, Chinese Academy of Sciences Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai, China
[2]Shanghai Information Center for Life Sciences, Chinese Academy of Sciences, Shanghai, China
[3]School of Life Science, ShanghaiTech University, Shanghai, China
[4]University of Chinese Academy of Sciences, Beijing, China

Abstract

RNA is essential for all kingdoms of life and exerts important functions beyond transferring genetic information from DNA to protein. With the advent of the state-of-the-art deep sequencing technology, a large portion of noncoding transcripts in eukaryotic genomes has been broadly identified. Among them, long noncoding RNAs (lncRNAs) have been emerged as a new class of RNA molecules that have regulatory potential in a variety of physiological and pathological processes. Here we summarize recent research progresses that have been made by scientists in China on lncRNAs, including their biogenesis, functional implication and the underlying mechanism of action at the current stage. © 2016 IUBMB Life, 68(11):887–893, 2016

Keywords: *Long noncoding RNAs; lncRNAs; circular RNAs; biogenesis; China*

Overview of Long Noncoding RNA Studies in China

Long noncoding RNAs (lncRNAs) comprise different types of RNA molecules with >200 nucleotides in length and lack coding potential. Most well characterized lncRNAs are RNA polymerase II (Pol II)-transcribed, capped, polyadenylated and contain exon–exon splice junctions like mRNAs (1,2). The first mammalian lncRNA *H19* was discovered in 1989 (3), followed by the identification of lncRNA *Xist* that plays an essential role in X chromosome inactivation, a process by which one X chromosome in female mammals is transcriptionally silenced to achieve dosage compensation of gene expression in males and females (4,5). It has been accepted that lncRNAs can play important roles in gene regulation in a variety of biological conditions with distinct mechanisms of action, such as recruiting transcription factors or chromatin-modifying complexes to their DNA targets, acting as decoys to sequester RNA binding proteins or miRNAs and directly interacting with DNAs or other RNAs (6).

By searching literatures of lncRNA studies between 1989 and 2015 in the Web of Science database with limited key words such as lncRNAs, lincRNAs and others, a brief summary of the total number of publications in each year is obtained (Table 1). Strikingly, although almost zero work had been reported before 2010 in China, the number of annual Science Citation Index of all publications by Chinese scientists has been rapidly grown since 2010. The number of articles with an impact factor (IF) higher than 4 (IF $\geq$ 4) increases from 2 in 2010 to 207 in 2015, corresponding to a proportion increased from 1.8% to 15.9% among all global publications. The number of articles published in top international journals with an impact factor higher than 10 (IF $\geq$ 10) also increases with a similar tendency. In this respect, it indicates a fact that both the number and the quality of lncRNA research have increased significantly in China.

Volume 68, Number 11, November 2016, Pages 887–893
Address correspondence to: Ling-Ling Chen, 320 Yueyang Road, C Building, Rm 705, Shanghai 200031, China.
E-mail: linglingchen@sibcb.ac.cn
*These authors contributed equally to this work.
Received 14 August 2016; Accepted 8 September 2016
DOI 10.1002/iub.1564
Published online 3 October 2016 in Wiley Online Library (wileyonlinelibrary.com)

TABLE 1 *Statistics of papers on the lncRNA-related research published by Chinese and global scientists (1989–2015).*

	Global			Chinese		
Year	Total	IF≥4	IF≥10	Total	IF≥4	IF≥10
1989–2000	231	33	23	0	0	0
2001–2005	235	124	20	2	0	0
2006	74	39	10	2	0	0
2007	66	28	7	2	1	0
2008	75	41	13	1	0	0
2009	92	44	10	1	0	0
2010	109	48	13	6	2	0
2011	186	77	32	15	4	1
2012	268	150	44	37	15	3
2013	535	254	87	140	49	9
2014	853	311	128	325	102	15
2015	1,302	605	120	656	207	18

The IF is calculated according to Journal Citation Reports (2014). Papers published by Chinese scientists refer to papers of those corresponding authors are from China. These data come from Web of Sciences database. Database search date: June 2, 2016. Database update date: May 31, 2016.

Major Discoveries Of LncRNA Study by Scientists in China

LncRNAs in Stem Cell Pluripotency and Differentiation

Transcriptional and epigenetic networks are regulated by multiple mechanisms in stem cells pluripotency and differentiation. LncRNAs constitute one class of molecules that play roles in this regulatory circuitry. It has been suggested that the cell-type specific expression of lncRNAs may be novel "fine-tuners" in stem cell fate determination (7). By applying single-cell RNA sequencing of polyadenylated RNAs, Tang and colleagues at Peking University analyzed more than 100 individual cells from human preimplantation embryos and human embryonic stem cells (hESCs), of which a total of 8,701 lncRNAs were detected. This work provides a valuable resource to study lncRNAs in early embryo development (8). LncRNA *linc-RoR* was known as a key regulator in stem cell pluripotency. However, function of this lncRNA and mechanism of action during ESCs self-renewal and differentiation have remained unknown (9). A study from Liu's group at Second Military Medical University has shown that *linc-RoR* is an endogenous competing RNA to titrate miRNAs that target Oct4, Nanog and Sox2 in self-renewing hESCs. This observation suggests that *linc-RoR* forms a feedback loop with core transcription factors and miRNAs to regulate ESCs maintenance and differentiation (10). Our group discovered *sno-lncRNA*s by sequencing non-polyadenylated RNAs in hESCs (11). *Sno-lncRNA*s are nuclear-enriched intron-derived lncRNAs that are processed on both ends by the snoRNA machinery. The genomic region encoding one abundant class of *sno-lnc*RNAs is specifically deleted in Prader-Willi syndrome (PWS), a neurogenetic disorder with a worldwide prevalence of 1 in 15,000–30,000 (12). In hESCs the PWS region *sno-lncRNA*s accumulate near their sites of synthesis, associate with Fox family splicing regulators and regulate patterns of splicing. This study reveals a potential role of a new class of lncRNAs in hESCs and a link between lncRNAs and the PWS pathogenesis.

In addition to the lncRNA itself that regulates gene expression in stem cells, recent studies by Shen and colleagues at Tsinghua University began to dissect how the lncRNA and its genomic locus in regulating neighborhood gene expression during stem cell maintenance and differentiation. They have delineated a complex effect of *Haunt* lncRNA and its DNA sequence in regulating *HOXA* gene activation during mouse ESCs (mESCs) differentiation. They found that knockdown of the *Haunt* lncRNA by RNAi, knockin of a transcription stop signal within the *Haunt* DNA locus or deletion of the *Haunt* promoter could all result in an upregulation of *HOXA* during retinoid acids (RA)-induced mESCs differentiation. Interestingly, removal of a large fragment of the *Haunt* DNA locus attenuated *HOXA* expression upon the RA treatment, which could not be rescued by the reinforced expression of *Haunt* RNA. This study suggests that the lncRNA *Haunt*, its genomic DNA and the transcription of this locus all play a role in regulating *HOXA* gene expression (13). Divergent lncRNAs are transcribed on the opposite strand from their neighboring protein-coding genes, representing ~20% of total lncRNAs in mammalian genomes. The same group has shown that divergent lncRNAs can interfere with the transcription of their adjacent genes in pluripotent cells. For instance, the divergent lncRNA *Evx1as* is required for the proper activation of its nearby *EVX1*. *Evx1as* binds to chromatin of its own locus, recruits epigenetic regulators and promotes chromatin looping at the *EVX1* locus (14).

Several studies have also revealed multiple roles of lncRNAs in establishing and maintaining cell-type-specific patterns of gene expression during cell differentiation. Wang and colleagues at Chinese University of Hong Kong reported the lncRNA *Dum* acts as an RNA regulator during myogenesis in mouse myoblast cells. *Dum* is a promyogenic factor during myoblast differentiation. Loss of *Dum* caused the delayed cardiotoxin-induced muscle regeneration in vivo. *Dum* silences the transcription of its neighboring gene developmental pluripotency-associated 2 (DPPA2) by binding and recruiting DNA methyltransferase to the *DPPA2* promoter (15). The same group has also identified *linc-YY1* as a novel lncRNA by analyzing several RNA-seq datasets from proliferating and differentiating mouse muscle myoblast C2C12 cells. They observed that depletion or overexpression of

linc-YY1 in C2C12 cells and muscle satellite cells altered myogenic differentiation and regeneration of injured muscles. Mechanistically, *linc-YY1* was found to interact with transcription factor Yin Yang 1 (YY1) and evict YY1 and Polycomb repressive complex from target promoters (16). Chen and colleagues at Institute of Biophysics, Chinese Academy of Sciences (CAS) characterized the lncRNA *ADINR* as a novel regulator during adipogenesis. Knockdown of *ADINR* impaired expression of adipogenic genes and reduced lipid accumulation in differentiated human bone marrow MSCs. *ADINR* binds to PAXIP1 associated glutamate rich protein 1 (PA1), recruits MLL3/4 histone methyl-transferase complexes to increase H3K4me3 and decrease H3K27me3 histone modifications, which leads to activation of gene at the C/EBPα locus (17). Zhang and colleagues at Kunming Institute of Zoology, CAS reported the DNA methylation signatures of lncRNAs in porcine adipose and muscle tissues by analyzing the methylated DNA immunoprecipitation sequencing datasets, further expanding our knowledge of lncRNAs in domestic animals (18).

LncRNAs in Tumorigenesis and Metastasis

Multiple lines of evidence have shown that lncRNAs are critical regulators in tumorigenesis and metastasis. Hepatocellular carcinoma (HCC) is the most common type of liver cancers and the third most frequent cause of cancer-related deaths (19). Sun's group at Second Military Medical University has identified several key lncRNAs that are involved in HCC progression. *LncRNA-HEIH* is highly expressed in HCC; it regulates cell cycle via binding to the enhancer of Enhancer of Zeste Homolog 2 (EZH2) and inhibits the expression of a substantial groups of genes regulated by EZH2 (20). *LncRNA-Dreh* is a downstream effector of hepatitis B virus X protein (HBx); the expression of this lncRNA is dramatically decreased in HBx-transgenic mice and mouse liver cells that expressing HBx, suggesting its role as tumor suppressor in the development of hepatitis B virus induced HCC (21). *LncRNA-LET* is a regulator of hypoxia-induced cancer cell invasion in hepatocellular carcinomas, colorectal cancers and squamous-cell lung carcinomas; its downregulation in tumors is epigenetically regulated by hypoxia-induced histone deacetylase 3 and required for hypoxia-mediated metastasis (22). LncRNA *PVT1* is a lncRNA found to be highly expressed in HCC; this lncRNA promotes cell proliferation, cell cycle and stem cell-like properties of HCC; *PVT1* binds to and stabilizes NOP2 nucleolar protein, which is required for *PVT1* function during tumorigenesis (23). The *lncRNA-ATB* is highly upregulated during HCC metastasis; this lncRNA promotes the invasion-metastasis cascade in HCC through interacting the miR-200 family to promote zinc finger E-box binding homeobox 1 (ZEB1) and ZEB2 expression and also by activating STAT3 signaling via binding and inducing interleukin 11 (*IL-11*) mRNA expression (24). *PRAL* is yet another lncRNA that has been identified as a regulator in tumor growth and apoptosis; *PRAL* associates with HSP90 and p53; this interaction stabilizes p53 levels by inhibiting MDM2-dependent ubiquitination of p53 (25). Zhou and colleagues identified lncRNA *MVIH* as a risk factor that can predict HCC patients' recurrence-free survival after hepatectomy. *MVIH* binds to phosphoglycerate kinase 1 and inhibits its secretion, a process contributes to the activation of tumor-inducing angiogenesis in HCC (26). A recent work by Zhou's group further uncovered the role of lncRNA *DANCR* in stem cell-like HCC cells. *DANCR* interacts with catenin beta 1 (CTNNB1) mRNA and prevents degradation of *ctnnb1* by miR-214, miR-320a and miR-199a; such an interaction of lncRNA, mRNA and miRNA plays an important role in increasing the stemness feature of HCC (27). Moreover, the lncRNA *lncTCF7* has been reported as a regulator in the self-renewal of liver cancer stem cells by Fan and collaborators at Institute of Biophysics, CAS. *LncTCF7* is upregulated in HCC tumors and liver cancer stem cells; it interacts with Switch/sucrose nonfermentable complex and recruits the complex to the *TCF7* promoter; these activities lead to *TCF7* gene activation and elevate Wnt signaling that can promote self-renewal and tumor propagation (28).

In addition to HCC, multiple lncRNAs that affect gene expression globally or in a gene-specific manner have been reported in other types of cancers. Song and colleagues at Sun Yat-sen University reported the lncRNA *NKILA* was induced by nuclear factor-κB (NF-κB) in breast cancer. *NKILA* prevents IκB kinase (IKK)-induced inhibitor of NF-κB (IκB) phosphorylation by directly masking phosphorylation motifs of IkB, leading to the suppression of breast cancer metastasis (29). Our group identified the human colorectal cancer (CRC)-specific lncRNA *CCAT1-L*. *CCAT1-L* is transcribed from the CRC-specific super-enhancer of *MYC*, accumulates at its sites of transcription and promotes the chromatin looping between the super-enhancer and the *MYC* promoter, leading to enhanced *MYC* transcription in CRC (30). Wu and colleagues at University of Science and Technology of China reported that *lincRNA-p21* was induced by hypoxia and required for glycolysis. *LincRNA-p21* binds to hypoxia-inducible factor-1 (HIF-1α) and von Hippel-Lindau (VHL) and blocks the interaction between these two proteins; such a complex inhibits HIF-1α ubiquitination mediated by VHL, leads to HIF-1α accumulation and promotes glycolysis in tumor (31). *LincRNA-p21* was also found to be a mediator of cell proliferation and apoptosis in atherosclerosis by enhancing p53 activity, shown by Zeng and colleagues at Third Military Medical University (32). Sunitinib has been used for renal cell carcinoma (RCC) treatment for years but some tumors are resistant to it. Wang and colleagues at Second Military Medical University has shown that *lncARSR* plays a crucial role in sunitinib resistance. *LncARSR* acts as a competing endogenous RNA and binds to miR-34/miR-449 that target AXL receptor tyrosine kinase and MET proto-oncogene (c-MET); additionally, hnRNPA2B1 regulates the packaging of *lncARSR* into exosomes, which leads to a transportation of *lncARSR* into sensitive cells and produces sunitinib-resistant RCC (33). Single nucleotide polymorphism (SNP) widely exists in human

genomes. A recent study from Lin's group at Chinese Academy of Medical Sciences and his collaborators at Peking Union Medical College has reported that a SNP can affect the function of lncRNA *LINC00673*. *LINC00673* enhances the interaction between protein tyrosine phosphatase nonreceptor type 11 (PTPN11) and pre-mRNA processing factor 19, a U-box-containing E3 ubiquitin ligase, to promote PTPN11 degradation by ubiquitination. The low level of PTPN11 inhibits proto-oncogene tyrosine-protein kinase Src/extracellular signal regulated kinase (SRC/ERK) oncogenic signaling pathway and activates STAT1-dependent antitumor response in pancreatic cancer. Remarkably, the G > A change at the SNP rs11655237 in the exon 4 of *LINC00673* turns this lncRNA into a target of miR-1231, which leads to *LINC00673* degradation and affects the tumorigenesis in pancreatic cancer (34).

LncRNAs in Immune Cell Differentiation and Function

Several recent studies conducted in China have uncovered lncRNAs as emerging regulators for immune cell differentiation and function. Cao and colleagues at Second Military Medical University has reported a lncRNA called *lnc-DC* that is exclusively expressed in human conventional dendritic cells (DCs). *Lnc-DC* regulates human DCs differentiation by activating transcription factor STAT3 through a direct binding of *lnc-DC* to STAT3. Such an RNA–protein interaction prevents STAT3 from binding to its downstream partner and promotes STAT3 phosphorylation at Tyrosine 705 (35). Zeng's group at Sun Yat-sen University identified *lncRNA-CD244* as a regulator in T-cell immune responses during tuberculosis infection; this lncRNA recruits EZH2 and mediates H3K27 trimethylation at *infg/tnfa* loci toward repressive chromatin states and inhibits interferon gamma/tumor necrosis factor alpha (IFN-γ/TNF-α) expression in CD8(+) T cells (36). Chen and colleagues at Institute of Microbiology, CAS has identified the lncRNA *NRAV* in influenza virus-infected cells. *NRAV* epigenetically inhibits transcription of many interferon-stimulated genes and promotes influenza A virus replication and virulence (37). Furthermore, the lncRNA *NEAT1* was known as the organizer of the nuclear structure called paraspekles. Both *NEAT1* expression and paraspeckle numbers increase upon immune stimuli such as infection and the poly(I:C) treatment (38). Our group recently found that *NEAT1* transcription was repressed by transcription regulator arginine methyltransferase CARM1. Upon poly(I:C) treatment, the paraspeckle-localized CARM1 was significantly reduced that led to both increased transcription of *NEAT1* and number of paraspeckles (39).

LncRNAs in Plant Biology

Dozens of research papers conducted in China have reported the regulatory potential of lncRNAs in different physiological conditions in plants. In 2012, Zhang and colleagues identified lncRNA *LDMAR* as an essential component of photoperiod-sensitive male sterility in hybrid rice (40). Since then over two dozens of lncRNA studies have been reported in plants in China. The majority of these studies largely focused on the profiling of lncRNAs in different plant species under diverse conditions by taking advantage of RNA-seq technology, including *Setaria italic* (41), *Arabidopsis thaliana* (42), *Zea mays L.* (43), *Oryza sativa* (44), *Populus tomentosa* (45), *Fragaria vesca* (46), *Solanum lycopersicum* (47), *Panax ginseng* (48), *Medicago truncatula* (49), *Morus notabilis* (50), *Brassica rapa ssp. Chinensis* (51) and *Gossypium arboreum* (52). For instance, Chen's group at Sun Yat-Sen University has identified a set of lncRNAs that are involved in the sexual reproduction of rice; one such lncRNA *XLOC_057324* has the potential to regulate panicle development and fertility (44). Yu and colleagues at China Agricultural University has identified 584 lncRNAs in millet, among which 19 lncRNAs were found to be responded to the drought stress (41). Lu's group at Tsinghua University identified 245 polyadenylated and 58 non-polyadenylated lncRNAs that were differentially expressed under various stress stimuli in *Arabidopsis thaliana* (42). By profiling lncRNAs in tomato, a study from Ye and colleagues at Huazhong Agricultural University has revealed that the origins of some lncRNAs were associated with transposable elements (53). Meng's group at Hangzhou Normal University has reported that several lncRNAs can generate small RNAs through a Dicer-like 1 dependent pathway in *Arabidopsis thaliana* (54). These studies together expand our knowledge on the widespread expression of lncRNAs in plant species and shed light on the potential regulation of lncRNAs under different stresses of plants.

Circular RNAs

Circular RNAs represent another type of lncRNAs. They are covalently closed single-stranded RNAs that were originally found in more than 20 years ago (55). There are two types of circular RNAs derived from Pol II transcripts: circular intronic RNAs (ciRNAs) that are derived from the inefficient debranching of lariat introns and circular RNAs (circRNAs) that are produced from back-splicing of exons. The RNA-seq of non-polyadenylated transcriptomes and the advent of integrated computational approaches have recently uncovered the widespread expression of circular RNAs (55).

While the majority of annotated lncRNAs are polyadenylated, our group began to explore the non-polyadenylated RNAs in human transcriptomes in 2011 (56). We detected highly expressed RNA-seq signals from both excised introns and excised exons (56), which comprise a number of new non-coding RNA species (11,57,58) that are spliced from RNA precursors (59). In collaboration with Yang lab at CAS-MPG Partner Institute for Computational Biology we uncovered ciRNAs whose formation depend on the consensus sequence containing a 7 nt GU-rich motif near the 5′ splice site and an 11 nt C-rich motif at the branch point site to avoid debranching, resulting in RNA circles carrying the 2′,5′-phosphodiester

bond. Such ciRNAs largely accumulate in the nucleus and promote Pol II transcription in *cis* (57). We also identified thousands of back-spliced circRNAs using an in-house developed computational algorithm called CIRCexplorer (58), which has been proven to be one of the most reliable tools for circRNA prediction (60). Back-spliced circRNA formation is dependent on flanking intronic complementary sequences. Remarkably, alternative formation of inverted repeated *Alu* pairs and the competition between them can lead to alternative circularization, resulting in multiple circRNAs produced from a single gene (58,61). Furthermore, by investigating nascent circRNA processing, we found that circRNA production largely occurs post-transcriptionally and some abundant circRNAs at steady-state level tend to be transcribed quickly and accumulate (55). Most recently, Yang lab reported the existence of pseudogenes retrotransposed from circRNAs in mammalian genomes, revealing a previously unknown impact of circRNAs by altering the host genome through retrotransposition (62).

In addition to the understanding of the biogenesis of circular RNAs, recent progress has also been made to illustrate their functional implications. In an attempt to interrogate ncRNAs associated with Pol II, Shan and colleagues at University of Science and Technology of China found a subset of intron-containing circRNAs produced from exon back-splicing (exon–intron circRNAs, EIciRNAs) that were associated with Pol II and promoted transcription of their parental mRNAs (63). Some EIciRNAs interact with Pol II, U1 small nuclear ribonucleoprotein (snRNP) and their parental gene promoters; depletion of U1 snRNP or blocking the interaction between U1 snRNA and EIciRNA abolished the effect of EIciRNAs on parental gene expression (63). While intron-containing circular RNAs are enriched in the nucleus (57,63), the majority of back-spliced circRNAs are largely localized in the cytoplasm. A recent study from He group at Fudan University revealed the presence of abundant circRNAs in exosomes (64). Exosomes are small membrane vehicles secreted by most of cell types, containing specific protein and RNA molecules that may modulate cell behaviors. Remarkably, over 1,000 circRNAs from serum exosomes of cancer patients were correlated with CRC, indicating that circRNAs may serve as promising biomarkers for cancer diagnosis. The same group also reported another line of evidence that indicates the regulatory potential of some circRNAs. They found that the Exon 2 of *HIPK3*-derived circHIPK3 was involved in cell growth by titrating multiple miRNAs (65).

Identification of circular RNAs depends on sequencing reads that can be mapped to the back-splicing junctions in a reversed order (55), hence specific algorithms need to be developed for their annotation. Several groups in China including Yang lab at PICB, CAS (57,58,62), Zhao lab at Beijing Institutes of Life Science (66) and Qu lab at Sun Yat-sen University (67) have independently developed new algorithms to predict circRNAs. It should be noted that different algorithms should be applied in a combination way in order to accurately mine circular RNAs (60). To overcome the heterogeneity of individual transcriptomes, a collaboration between Tang and Huang's groups at Peking University have developed SUPeR-seq to analyze the non-polyadenylated transcripts in single cells, from which 2,891 cricRNAs were identified in mouse implantation embryos. Finally, a study from Han and colleagues at Shanghai Institute of Plant Physiology and Ecology, CAS identified over two thousands cricRNAs in rice (49), further revealing the widespread expression of circular RNAs.

LncRNA Databases

In addition to the characterization of lncRNA function in normal development and disease states, several research groups in China developed algorithms to predict and annotate lncRNAs. Chen and his collaborators at Institute of Biophysics, CAS established the first version of NONCODE database that provided the comprehensive information of noncoding RNAs in 2005. The database has been updated regularly and the newest version, NONCODEv4, contains 21,0831 lncRNAs (68–71). An earlier study from Zhao's group at Sun Yat-sen University has reported the first computational annotation of lncRNA function using published microarray data (72). Cui and colleagues at Peking University established LncRNADisease, a database for lncRNA-associated disease, which contains 480 entries of lncRNA-disease association from 166 diseases and 1,564 human lncRNAs predicted to be associated with diseases (73). Wu and colleagues at Harbin Medical University has developed a method with an improved accuracy for lncRNA prediction and identified a number of putative lncRNAs with regulatory potential in the vicinity of gene loci during mouse brain development (74).

Conclusions

It is now well accepted that lncRNAs are pervasively transcribed throughout eukaryotic genomes and play important roles in gene regulation. In the past several years, efforts made by Chinese scientists have increased our knowledge on the biogenesis and regulatory potential of lncRNAs in a variety of biological contexts. As the field moves forward, many questions still remain to be addressed, especially the detailed mechanisms of action of lncRNAs that have potentially important functions during normal development and in disease states. In addition, new technologies are also warranted in order to characterize lncRNAs in a more sensitive and quantitative way and to dissect the lncRNAs and their protein complexes in different biological settings.

Acknowledgements

This work was supported by grants from the National Natural Science Foundation of China (91440202 and 31322018). We are grateful to the members of Chen laboratory for suggestions and comments of this manuscript. We apologize to authors whose work could not be cited here owing to space/content limitations.

References

[1] Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., et al. (2005) Antisense transcription in the mammalian transcriptome. Science 309, 1564–1566.

[2] Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., et al. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Gene Dev. 25, 1915–1927.

[3] Groux, H., Huet, S., Aubrit, F., Tran, H. C., Boumsell, L., et al. (1989) A 19-kDa human-erythrocyte molecule H19 is involved in Rosettes, present on nucleated cells, and required for T-cell activation—comparison of the roles of H19 and Lfa-3 molecules in T-cell activation. J. Immunol. 142, 3013–3020.

[4] Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., et al. (1991) A gene from the region of the human X-inactivation center is expressed exclusively from the inactive X-chromosome. Nature 349, 38–44.

[5] Briggs, S. F. and Pera, R. A. R. (2014) X chromosome inactivation: recent advances and a look forward. Curr. Opin. Genet. Dev. 28, 78–82.

[6] Goff, L. A. and Rinn, J. L. (2015) Linking RNA biology to lncRNAs. Genome Res. 25, 1456–1465.

[7] Rinn, J. L. and Chang, H. Y. (2012) Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. 81, 145–166.

[8] Yan, L. Y., Yang, M. Y., Guo, H. S., Yang, L., Wu, J., et al. (2013) Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat. Struct. Mol. Biol. 20, 1131–1139.

[9] Loewer, S., Cabili, M. N., Guttman, M., Loh, Y. H., Thomas, K., et al. (2010) Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. Nat. Genet. 42, 1113–1117.

[10] Wang, Y., Xu, Z. Y., Jiang, J. F., Xu, C., Kang, J. H., et al. (2013) Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. Dev. Cell 25, 69–80.

[11] Yin, Q. F., Yang, L., Zhang, Y., Xiang, J. F., Wu, Y. W., et al. (2012) Long noncoding RNAs with snoRNA ends. Mol. Cell. 48, 219–230.

[12] Cassidy, S. B., Schwartz, S., Miller, J. L., and Driscoll, D. J. (2012) Prader-Willi syndrome. Genet. Med. 14, 10–26.

[13] Yin, Y. F., Yan, P. X., Lu, J. L., Song, G., Zhu, Y. Y., et al. (2015) Opposing roles for the lncRNA Haunt and its genomic locus in regulating HOXA gene activation during embryonic stem cell differentiation. Cell Stem Cell 16, 504–516.

[14] Luo, S., Lu, Y. Y., Liu, L. C., Yin, Y. F., Chen, C. Y., et al. (2016) Divergent lncRNAs regulate gene expression and lineage differentiation in pluripotent cells. Cell Stem Cell 18, 637–652.

[15] Wang, L. J., Zhao, Y., Bao, X. C., Zhu, X. H., Kwok, Y. K. Y., et al. (2015) LncRNA Dum interacts with Dnmts to regulate Dppa2 expression during myogenic differentiation and muscle regeneration. Cell Res. 25, 335–350.

[16] Zhou, L., Sun, K., Zhao, Y., Zhang, S. Y., Wang, X. C., et al. (2015) Linc-YY1 promotes myogenic differentiation and muscle regeneration through an interaction with the transcription factor YY1. Nat. Commun. 6, 10026.

[17] Xiao, T. F., Liu, L. H., Li, H. L., Sun, Y., Luo, H. X., et al. (2015) Long noncoding RNA ADINR regulates adipogenesis by transcriptionally activating C/EBP alpha. Stem Cell Rep. 5, 856–865.

[18] Zhou, Z. Y., Li, A. M., Wang, L. G., Irwin, D. M., Liu, Y. H., et al. (2015) DNA methylation signatures of long intergenic noncoding RNAs in porcine adipose and muscle tissues. Sci. Rep. 5.

[19] Forner, A., Llovet, J. M., and Bruix, J. (2012) Hepatocellular carcinoma. Lancet 379, 1245–1255.

[20] Yang, F., Zhang, L., Huo, X. S., Yuan, J. H., Xu, D., et al. (2011) Long noncoding RNA high expression in hepatocellular carcinoma facilitates tumor growth through enhancer of Zeste Homolog 2 in humans. Hepatology 54, 1679–1689.

[21] Huang, J. F., Guo, Y. J., Zhao, C. X., Yuan, S. X., Wang, Y., et al. (2013) Hepatitis B virus X protein (HBx)-related long noncoding RNA (lncRNA) down-regulated expression by HBx (Dreh) inhibits hepatocellular carcinoma metastasis by targeting the intermediate filament protein vimentin. Hepatology 57, 1882–1892.

[22] Yang, F., Huo, X. S., Yuan, S. X., Zhang, L., Zhou, W. P., et al. (2013) Repression of the long noncoding RNA-LET by histone deacetylase 3 contributes to hypoxia-mediated metastasis. Mol. Cell. 49, 1083–1096.

[23] Wang, F., Yuan, J. H., Wang, S. B., Yang, F., Yuan, S. X., et al. (2014) Oncofetal long noncoding RNA PVT1 promotes proliferation and stem cell-like property of hepatocellular carcinoma cells by stabilizing NOP2. Hepatology 60, 1278–1290.

[24] Yuan, J. H., Yang, F., Wang, F., Ma, J. Z., Guo, Y. J., et al. (2014) A long noncoding RNA activated by TGF-beta promotes the invasion-metastasis cascade in hepatocellular carcinoma. Cancer Cell 25, 666–681.

[25] Zhou, C. C., Yang, F., Yuan, S. X., Ma, J. Z., Liu, F., et al. (2016) Systemic genome screening identifies the outcome associated focal loss of long noncoding RNA PRAL in hepatocellular carcinoma. Hepatology 63, 850–863.

[26] Yuan, S. X., Yang, F., Yang, Y., Tao, Q. F., Zhang, J., et al. (2012) Long noncoding RNA associated with microvascular invasion in hepatocellular carcinoma promotes angiogenesis and serves as a predictor for hepatocellular carcinoma patients' poor recurrence-free survival after hepatectomy. Hepatology 56, 2231–2241.

[27] Yuan, S. X., Wang, J., Yang, F., Tao, Q. F., Zhang, J., et al. (2016) Long noncoding RNA DANCR increases stemness features of hepatocellular carcinoma by derepression of CTNNB1. Hepatology 63, 499–511.

[28] Wang, Y. Y., He, L., Du, Y., Zhu, P. P., Huang, G. L., et al. (2015) The long noncoding RNA lncTCF7 promotes self-renewal of human liver cancer stem cells through activation of Wnt signaling. Cell Stem Cell 16, 413–425.

[29] Liu, B. D., Sun, L. J., Liu, Q., Gong, C., Yao, Y. D., et al. (2015) A cytoplasmic NF-kappa B interacting long noncoding RNA blocks I kappa B phosphorylation and suppresses breast cancer metastasis. Cancer Cell 27, 370–381.

[30] Xiang, J. F., Yin, Q. F., Chen, T., Zhang, Y., Zhang, X. O., et al. (2014) Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. Cell Res. 24, 513–531.

[31] Yang, F., Zhang, H. F., Mei, Y. D., and Wu, M. (2014) Reciprocal regulation of HIF-1 alpha and LincRNA-p21 modulates the Warburg effect. Mol. Cell. 53, 88–100.

[32] Wu, G. Z., Cai, J., Han, Y., Chen, J. H., Huang, Z. P., et al. (2014) LincRNA-p21 regulates neointima formation, vascular smooth muscle cell proliferation, apoptosis, and atherosclerosis by enhancing p53 activity. Circulation 130, 1452–U1464.

[33] Qu, L., Ding, J., Chen, C., Wu, Z. J., Liu, B., et al. (2016) Exosome-transmitted lncARSR promotes sunitinib resistance in renal cancer by acting as a competing endogenous RNA. Cancer Cell 29, 653–668.

[34] Zheng, J., Huang, X., Tan, W., Yu, D., Du, Z., et al. (2016) Pancreatic cancer risk variant in LINC00673 creates a miR-1231 binding site and interferes with PTPN11 degradation. Nat. Genet. 48, 747–757.

[35] Wang, P., Xue, Y. Q., Han, Y. M., Lin, L., Wu, C., et al. (2014) The STAT3-binding long noncoding RNA lnc-DC controls human dendritic cell differentiation. Science 344, 310–313.

[36] Wang, Y., Zhong, H. L., Xie, X. D., Chen, C. Y., Huang, D., et al. (2015) Long noncoding RNA derived from CD244 signaling epigenetically controls CD8(+) T-cell immune responses in tuberculosis infection. Proc. Natl. Acad. Sci. USA 112, E3883–E3892.

[37] Ouyang, J., Zhu, X. M., Chen, Y. H., Wei, H. T., Chen, Q. H., et al. (2014) NRAV, a long noncoding RNA, modulates antiviral responses through suppression of interferon-stimulated gene transcription. Cell Host Microbe 16, 616–626.

[38] Imamura, K., Imamachi, N., Akizuki, G., Kumakura, M., Kawaguchi, A., et al. (2014) Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to Paraspeckle mediates IL8 expression upon immune stimuli. Mol. Cell. 53, 393–406.

[39] Hu, S. B., Xiang, J. F., Li, X., Xu, Y., Xue, W., et al. (2015) Protein arginine methyltransferase CARM1 attenuates the paraspeckle-mediated nuclear retention of mRNAs containing IRAlus. Genes Dev. 29, 630–645.

[40] Ding, J. H., Lu, Q., Ouyang, Y. D., Mao, H. L., Zhang, P. B., et al. (2012) A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. Proc. Natl. Acad. Sci. USA 109, 2654–2659.

[41] Qi, X., Xie, S. J., Liu, Y. W., Yi, F., and Yu, J. J. (2013) Genome-wide annotation of genes and noncoding RNAs of foxtail millet in response to simulated drought stress by deep sequencing. Plant Mol. Biol. 83, 459–473.

[42] Di, C., Yuan, J. P., Wu, Y., Li, J. R., Lin, H. X., et al. (2014) Characterization of stress-responsive lncRNAs in *Arabidopsis thaliana* by integrating expression, epigenetic and structural features. Plant J. 80, 848–861.

[43] Zhang, W., Han, Z. X., Guo, Q. L., Liu, Y., Zheng, Y. X., et al. (2014) Identification of maize long non-coding RNAs responsive to drought stress. Plos One 9
[44] Zhang, Y. C., Liao, J. Y., Li, Z. Y., Yu, Y., Zhang, J. P., et al. (2014) Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice. Genome Biol. 15, 512.
[45] Chen, J. H., Quan, M. Y., and Zhang, D. Q. (2015) Genome-wide identification of novel long non-coding RNAs in *Populus tomentosa* tension wood, opposite wood and normal wood xylem by RNA-seq. Planta 241, 125–143.
[46] Kang, C. Y., and Liu, Z. C. (2015) Global identification and analysis of long non-coding RNAs in diploid strawberry *Fragaria vesca* during flower and fruit development. BMC Genom. 16, 815.
[47] Zhu, B. Z., Yang, Y. F., Li, R., Fu, D. Q., Wen, L. W., et al. (2015) RNA sequencing and functional analysis implicate the regulatory role of long non-coding RNAs in tomato fruit ripening. J. Exp. Bot. 66, 4483–4495.
[48] Wang, K., Liu, C. Y., Zhou, L. Y., Wang, J. X., Wang, M., et al. (2015) APF lncRNA regulates autophagy and myocardial infarction by targeting miR-188-3p. Nat. Commun. 6, 6779.
[49] Lu, T., Cui, L., Zhou, Y., Zhu, C., Fan, D., et al. (2015) Transcriptome-wide investigation of circular RNAs in rice. RNA 21, 2076–2087.
[50] Song, X. B., Sun, L., Luo, H. T., Ma, Q. G., Zhao, Y., et al. (2016) Genome-wide identification and characterization of long non-coding RNAs from mulberry (*Morus notabilis*) RNA-seq Data. Genes 7, E11.
[51] Song, X. M., Liu, G. F., Huang, Z. N., Duan, W. K., Tan, H. W., et al. (2016) Temperature expression patterns of genes and their coexpression with LncRNAs revealed by RNA-Seq in non-heading Chinese cabbage. BMC Genom. 17, 297.
[52] Zou, C. S., Wang, Q. L., Lu, C. R., Yang, W. C., Zhang, Y. P., et al. (2016) Transcriptome analysis reveals long noncoding RNAs involved in fiber development in cotton (*Gossypium arboreum*). Science China-Life Sci. 59, 164–171.
[53] Wang, X., Ai, G., Zhang, C. L., Cui, L., Wang, J. F., et al. (2016) Expression and diversification analysis reveals transposable elements play important roles in the origin of Lycopersicon-specific lncRNAs in tomato. New Phytol. 209, 1442–1455.
[54] Ma, X. X., Shao, C. G., Jin, Y. F., Wang, H. Z., and Meng, Y. J. (2014) Long non-coding RNA A novel endogenous source for the generation of Dicer-like 1-dependent small RNAs in *Arabidopsis thaliana*. RNA Biol. 11, 373–390.
[55] Chen, L. L. (2016) The biogenesis and emerging roles of circular RNAs. Nat. Rev. Mol. Cell. Biol. 17, 205–211.
[56] Yang, L., Duff, M. O., Graveley, B. R., Carmichael, G. G., and Chen, L. L. (2011) Genomewide characterization of non-polyadenylated RNAs. Genome Biol. 12, R16.
[57] Zhang, Y., Zhang, X. O., Chen, T., Xiang, J. F., Yin, Q. F., et al. (2013) Circular intronic long noncoding RNAs. Mol. Cell. 51, 792–806.
[58] Zhang, X. O., Wang, H. B., Zhang, Y., Lu, X., Chen, L. L., et al. (2014) Complementary sequence-mediated exon circularization. Cell 159, 134–147.
[59] Yang, L. (2015) Splicing noncoding RNAs from the inside out. Wiley Interdiscip. Rev. RNA 6, 651–660.
[60] Hansen, T. B., Veno, M. T., Damgaard, C. K., and Kjems, J. (2016) Comparison of circular RNA prediction tools. Nucleic Acids Res. 44, e58.
[61] Chen, L. L., and Yang, L. (2015) Regulation of circRNA biogenesis. RNA Biol. 12, 381–388.
[62] Dong, R., Zhang, X. O., Zhang, Y., Ma, X. K., Chen, L. L., et al. (2016) CircRNA-derived pseudogenes. Cell. Res. 26, 747–750.
[63] Li, Z., Huang, C., Bao, C., Chen, L., Lin, M., et al. (2015) Exon–intron circular RNAs regulate transcription in the nucleus. Nat. Struct. Mol. Biol. 22, 256–264.
[64] Li, Y., Zheng, Q., Bao, C., Li, S., Guo, W., et al. (2015) Circular RNA is enriched and stable in exosomes: a promising biomarker for cancer diagnosis. Cell. Res. 25, 981–984.
[65] Zheng, Q., Bao, C., Guo, W., Li, S., Chen, J., et al. (2016) Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. Nat. Commun. 7, 11215.
[66] Gao, Y., Wang, J., and Zhao, F. (2015) CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. Genome Biol. 16, 4.
[67] Zheng, L. L., Li, J. H., Wu, J., Sun, W. J., Liu, S., et al. (2016) deepBase v2.0: identification, expression, evolution and function of small RNAs, LncRNAs and circular RNAs from deep-sequencing data. Nucleic Acids Res. 44, D196–D202.
[68] Liu, C. N., Bai, B. Y., Skogerbo, G., Cai, L., Deng, W., et al. (2005) NONCODE: an integrated knowledge database of non-coding RNAs. Nucleic Acids Res. 33, D112–D115.
[69] He, S. M., Liu, C. N., Skogerbo, G., Zhao, H. T., Wang, J., et al. (2008) NONCODE v2.0: decoding the non-coding. Nucleic Acids Res. 36, D170–D172.
[70] Bu, D. C., Yu, K. T., Sun, S. L., Xie, C. Y., Skogerbo, G., et al. (2012) NONCODE v3.0: integrative annotation of long noncoding RNAs. Nucleic Acids Res. 40, D210–D215.
[71] Xie, C. Y., Yuan, J., Li, H., Li, M., Zhao, G. G., et al. (2014) NONCODEv4: exploring the world of long non-coding RNA genes. Nucleic Acids Res. 42, D98–D103.
[72] Liao, Q., Liu, C. N., Yuan, X. Y., Kang, S. L., Miao, R. Y., et al. (2011) Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network. Nucleic Acids Res. 39, 3864–3878.
[73] Chen, G., Wang, Z. Y., Wang, D. Q., Qiu, C. X., Liu, M. X., et al. (2013) LncRNA disease: a database for long-non-coding RNA-associated diseases. Nucleic Acids Res. 41, D983–D986.
[74] Lv, J., Liu, H. B., Huang, Z. J., Su, J. Z., He, H. J., et al. (2013) Long non-coding RNA identification over mouse brain development by integrative modeling of chromatin and genomic features. Nucleic Acids Res. 41, 10044–10061.