Long non-coding RNAs (lncRNAs) have been found to perform various functions in a wide variety of important biological processes. To make easier interpretation of lncRNA functionality and conduct deep mining on these transcribed sequences, it is convenient to classify lncRNAs into different groups. Here, we summarize classification methods of lncRNAs according to their four major features, namely, genomic location and context, effect exerted on DNA sequences, mechanism of functioning and their targeting mechanism. In combination with the presently available function annotations, we explore potential relationships between different classification categories, and generalize and compare biological features of different lncRNAs within each category. Finally, we present our view on potential further studies. We believe that the classifications of lncRNAs as indicated above are of fundamental importance for lncRNA studies, helpful for further investigation of specific lncRNAs, for formulation of new hypothesis based on different features of lncRNA and for exploration of the underlying lncRNA functional mechanisms.

Introduction

In contrast to a small proportion of the mammalian genome (e.g., human, mouse) that are transcribed into mRNAs, the vast majority of the genome is transcribed into what was previously regarded as “dark matter”—non-coding RNAs (ncRNAs) that do not encode information about proteins.1,5 Among these ncRNAs, long ncRNAs (lncRNAs) represent the most prevalent and functionally diverse class.6-8 There is no definition of lncRNA that is based on biological argumentation and widely accepted in the community. The most commonly used definition is based on the threshold of 200 nucleotides (nt) of the RNA length.6,7,9 It conventionally divides ncRNAs into lncRNAs that have more than 200 nt in length and the remaining ones that are considered “small” RNAs. Small ncRNAs include many different RNAs, such as microRNAs (miRNAs), small nucleolar RNAs (snoRNAs), piwiRNAs (piRNAs).8 Undoubtedly, the definition of lncRNA merely based on length is arbitrary. One attempt to distinguish lncRNAs from small ncRNAs, based more on the biological argumentation is proposed by Amaral et al. defining lncRNAs as those ncRNAs that function either as primary or spliced transcripts, independent of extant known classes of small ncRNAs.8,9 Therefore, there are some lncRNAs that do not exceed...
the arbitrary threshold in length (such as BCI and snuR, which are less than or close to 200 nt but included in IncRNAdb\(^{10}\)).

LncRNAs are observed in a large diversity of species, including animals,\(^{11,12}\) plants,\(^{13}\) yeast,\(^{14}\) prokaryotes\(^{15}\) and even viruses.\(^{16}\) However, IncRNAs are poorly conserved among different species when compared with the well-studied RNAs (such as mRNAs, miRNAs, snoRNAs),\(^ {4,17,18}\) invoking uncertainty about whether a given lncRNA is functional at all due to poor interspecies conservation, or it conveys functional species-specific characteristics. In addition, IncRNAs are usually low expressed,\(^ {4,19,20}\) making them to look like more as transcriptional noise. Despite this, a lot of evidence has accumulated showing that lncRNAs play a significant role in a wide variety of important biological processes,\(^ {21,22}\) including transcription,\(^ {23,24}\) splicing,\(^ {25,26}\) translation,\(^ {27,28}\) protein localization,\(^ {29,30}\) cellular structure integrity,\(^ {31,32}\) imprinting,\(^ {33-35}\) cell cycle\(^ {36,37}\) and apoptosis,\(^ {38,39}\) stem cell pluripotency\(^ {40}\) and reprogramming\(^ {40}\) and heat shock response.\(^ {51,41}\) It has been suggested that lncRNAs may regulate cancer progression\(^ {43}\) and development of many other human diseases.\(^ {44}\) Moreover, a considerable number of lncRNAs are 3’ polyadenylated, 5’ capped, multi-exonic\(^ {2,4,20}\) and exhibit transcriptional activation activity similar to that of mRNAs.\(^ {1,45,46}\) As a consequence, all these functions and biological features of lncRNAs make them interesting and important research topic.

Recent advances in experimental and computational technologies make it feasible to conduct deep mining on more and more transcribed sequences.\(^ {2,20,47-49}\) At present, there are 73,370 lncRNA entries from 1,239 organisms according to NONCODE v3.0\(^ {50}\) (a database of literature documented lncRNAs). Conversely, among all these lncRNAs, only a small proportion (less than 200 according to LncRNAdb,\(^ {10}\) a database of lncRNA annotation) has been functionally annotated. To better understand their functional significance, it helps to classify lncRNAs into different groups that are useful for exploring their underlying mechanisms of actions, for formulating new hypotheses and for providing insights in differences of such major classes of lncRNAs. Here, we summarize classification methods of lncRNAs according to their different features as discussed in what follows, including their (1) genome location and context, (2) exerted effect on DNA sequences, (3) mechanism of functioning and (4) targeting mechanism. Finally, we provide our perspectives on potential further studies.

**Genomic Location and Context**

**Intergenic lncRNAs and intronic IncRNAs.** There is a large number of non-coding regions (accounting for 98–99% in the human genome) interspersed between coding regions.\(^ {51,52}\) Since lncRNAs are located and transcribed from different genomic locations, those transcribed from intergenic regions are named intergenic lncRNAs (Fig. 1A) and, in contrast, those transcribed entirely from introns of protein-coding genes are named intronic lncRNAs (Fig. 1B).

It is suggested that intergenic lncRNAs and intronic lncRNAs are most likely regulated through different transcription activation mechanisms\(^ {45}\) and may have different poly(A) modifications and manifest activities in different cellular locations.\(^ {48}\) However, only a small portion of intronic lncRNAs has been explored regarding their function. In contrast, there is a large number of long intergenic non-coding RNAs (lincRNAs) that function through different types of mechanisms: cis or trans transcriptional regulation (described below), translational control, splicing regulation, other post-transcriptional regulation, etc. (see Table 1). Also, lincRNAs have been extensively studied about their expression feature and conservation among species.

It is found that lincRNAs are transcriptionally activated similarly to mRNAs,\(^ {45,46,54-55}\) as they are more conserved than introns\(^ {45,46,55}\) and antisense transcripts,\(^ {45}\) more tissue-specificly expressed than protein-coding genes\(^ {40,20}\) and more stable than intronic lncRNAs.\(^ {56}\) “K4-K36” domain (with histone H3K4 trimethylation at their 5’ end and histone H3K36 trimethylation in the body of the gene), an indicator of active transcription in protein-coding genes, is found to prevalently exist in transcriptionally active lncRNAs.\(^ {45,46,54-55}\) Approximately 70% of lincRNAs with “K4-K36” domain show evidence of RNA transcription, which is similar to the proportion (~72%) of protein-coding genes.\(^ {48}\) Most importantly, nearly 70% of the transcription active domains (K4-K36 domain) of lincRNAs in human are conserved in the orthologous region of mouse, which is comparable to the corresponding proportion (80%) of protein-coding genes.\(^ {46}\) In addition, lincRNAs are found to be conserved across multiple vertebrate species.\(^ {20}\) The above evidences (active transcription, a degree of domain conservation, tissue-specific expression, stability) strongly indicate the functional importance of lincRNAs. In reality, lincRNAs are found to perform important functions in many cellular processes, from embryonic stem cell pluripotency to cell proliferation and cancer progression.\(^ {38,46,53,54}\)

**Sense and antisense IncRNAs.** Sense lncRNAs are transcribed from the sense strand of protein-coding genes, containing exons from protein-coding genes. They may overlap with part of protein-coding genes, or cover the entire sequence of a protein-coding gene (Fig. 1C). Antisense IncRNAs, to the contrary, are transcribed from the antisense strand of protein-coding genes. According to GENCODE (a database of manually curated lncRNAs) annotation,\(^ {4}\) antisense lncRNAs may appear in three scenarios: (1) transcripts from the antisense strand of protein-coding genes overlap an exon of a sense gene through lncRNAs’ exons, (2) transcripts from the intron of a sense gene do not have exon-exon overlap with this sense gene and (3) transcripts cover the entire sequence of a sense gene through an intron (Fig. 1D). Sense and antisense lncRNAs are proved to be genuine transcripts by strand-specific assay or sequencing,\(^ {45,47}\) qRT-PCR validation\(^ {45}\) and by sequencing 5’ and 3’ ends of full-length cDNA\(^ {2}\) and, thus, they do not represent truncated CDJs or transcriptional noise. Most of the IncRNAs that come from protein-coding genes or the antisense of the protein-coding genes can be obtained using CAGE (cap-analysis gene expression) and oligo-dT guided reverse transcription, suggesting that they also possess mRNA-like features of 3’ polyadenylation and 5’ capping.\(^ {2}\) Moreover, sense and antisense lncRNAs can also be multi-exonic.\(^ {45,57}\)

Many antisense lncRNAs function through different types of mechanisms (similar to lincRNAs) (Table 1). It has been found that as many as 87% coding transcripts have antisense partners.
in the mouse genome and ~32% of the human IncRNAs are antisense to coding genes, suggesting that antisense regulation is likely to be commonly utilized. However, in comparison with lincRNAs and antisense IncRNAs, sense IncRNAs have been less explored for their functions (Table 1). It is suggested that most IncRNAs tend not to have protein-coding potential. Intriguingly, some sense IncRNAs are special in the sense that they can function as both RNA and protein-coding gene. For instance, SRA (steroid receptor RNA activator) can translate into protein, and the RNA sequence can also act as a scaffold for several co-activator and repressor proteins to form complexes that regulate gene transcription; ENOD40 (early nodulin 40) can translate into proteins and is also needed for correct subcellular localization of RNP particles in legume plants. These findings have challenged our understanding of gene classification and, in the meantime, broadened the known roles of IncRNA. Such special relationship between the sense IncRNAs and protein-coding genes may provide novel insights into the evolution of gene function.

Present studies are mainly focused on lincRNAs and antisense IncRNAs (especially lincRNAs), though less is known about intronic IncRNAs and sense IncRNAs. Some biological features of IncRNA are quite prominent, such as the high conservation of lincRNAs among mammals. Also, another evidence, albeit not extensive, may indicate the special feature of the sense IncRNAs’ coding potential. Additionally, as antisense IncRNAs and sense IncRNAs are correlated differently with coding genes, they are most likely to exert different effects on gene locus or mRNAs. Therefore, genomic location and context can be used for classification of IncRNA, though a classification of IncRNA using exclusively genomic localization and context may not be fully adequate.

**Effects Exerted on DNA Sequences**

It has been found that IncRNAs are predominately localized in nucleus and chromatin, suggesting that IncRNAs may have a significant impact on DNA sequences. Also, a large proportion of IncRNAs are involved in transcriptional regulation (~42% of the 182 assessed entries according to lncRNAdb; Table 1). Therefore, it is meaningful to classify IncRNAs based on their effects exerted on DNA sequences: *cis*-IncRNAs (*cis*-acting IncRNAs) that regulate the expression of genes in close genomic proximity and *trans*-IncRNAs (*trans*-acting IncRNAs) that regulate the expression of distant genes (Fig. 2A).

**Cis-IncRNAs.** It seems plausible that *cis*-IncRNAs function through transcriptional interference or chromatin modification, ranging from yeast [e.g., SRGI (regulatory gene 1) RNAs] to plants [e.g., COLDAIR (cold assisted intronic noncoding RNA)] to mammals [e.g., DHFR (dihydrofolate reductase) upstream transcripts; Xist (X inactive-specific transcript)].

**Antisense IncRNAs.** It is believed that antisense IncRNAs interfere with transcription of genes in close proximity by competing for transcription factors, histone modifications, and chromatin regulators. Additionally, antisense IncRNAs can regulate gene expression by forming RNA hybrids with target mRNAs or by recruiting protein complexes that affect gene expression.

**Trans-Acting IncRNAs.** Trans-acting IncRNAs are not localized in the nucleus and are not transcribed from protein-coding genes. They are associated with the chromatin of protein-coding genes and can function as *cis*-acting IncRNAs in the same cell or as *trans*-acting IncRNAs in different cells. They can regulate gene expression by recruiting protein complexes that affect transcription factors, histone modifications, or chromatin structure. They can also act as miRNA sponges to reduce the function of miRNAs.

Figure 1. Genomic location and context of IncRNAs. Protein-coding genes and their exons are represented by blue color, while IncRNAs and their exons are represented by red color. Panels are mainly based on IncRNA location annotation of GENCODE. (A) Intergenic IncRNA, transcribed intergenically from both strands. (B) Intronic IncRNA, transcribed entirely from introns of protein-coding genes. (C) Sense IncRNA, transcribed from the sense strand of protein-coding genes and contain exons from protein-coding genes, overlapping with part of protein-coding genes or covering the entire sequence of a protein-coding gene through an intron. (D) Antisense IncRNA, transcribed from the antisense strand of protein-coding genes, overlapping with exonic or intronic regions or covering the entire protein-coding sequence through an intron.
Likewise, SER3 region of (Ser3p), are found to cover the promoter of SER3 and, thus, repress the expression of SER3.60,61 Also, the (maternally expressed 3) (a ~1.6 kb lncRNA) and (transcription complex) formation, or by interacting with transcription initiation complex to block PIC (preinitiation complexes). The most studied chromatin modification complexes is PRC and one well-known example is (HOX antisense intergenic RNA), a ~2.2 kb lncRNA from the promoter region of HOX antisense intergenic RNA (Gal10p-noncoding RNA) (a ~4 kb lncRNA) has been reported to recruit the Rpd3S HDAC complex, resulting in a decrease in some histone 3 acetylation to repress the expression of GAL1.4 Although the above-mentioned examples of PRC and Rpd3S HDAC are all implicated in negative regulation of gene expression, interaction of lncRNA with chromatin modification complexes is also involved in positive regulation. A case example is (HOX transcript at the distal tip) (a ~3.8 kb lncRNA in human) that recruits a MLL chromatin modifying complex to maintain a domain of active chromatin over the 5′ end of HOXA (homeobox A cluster) gene cluster.67

**Table 1. Functional mechanisms and genomic locations of lncRNAs**

| Function mechanism         | Gene symbol |
|----------------------------|-------------|
| Transcriptional regulation | Cis, Unclear |
| Splicing regulation        | MIATβ, LUSTβ, Malat1β, SAFβ, VL30 RNAs, Zeb2NATβ |
| Translational regulation   | BCIα, BC200β, Gadd7, SNHG1β, SNHG6β, smaR, Zeb2NATβ |
| Other post-transcriptional regulation | 21Aα, 1/2-sbsRNA1α, At4β, BACE1ASβ, CDR1 ASβ, Dio3osβ, E2F4 antisenseβ, Emx2osβ, Gadd7, H19β, HULCβ, HSUR1 and HSUR2, IFS1β, KRASP1β, Linc-MD1β, psvA antisense RNA, PTENP1β, tie-1asβ, WT1-ASβ |
| Other functional mechanisms | 7SLβ, Beta 2.7 RNA, Centromeric α-satellite RNA, ENOD40β, EBER1 and EBER2 RNAs, G22β, LIP16β, hsr omega transcriptsβ, meiRNA, Maternal RNA templatesβ, Maternal somatic nucleus RNASc, MER11Cc, NRONβ, mhc1-1β, roX1 and roX2β, sfrNA, TERRA, TERCβ, Va1 and Va2l RNAs, VegT RNA, XlSirts, Y RNAsβ |

Note: lncRNAs listed are collected from the database of lncRNAdb and published papers. Cis, lncRNAs that regulate expression of genes in close genomic proximity; Trans, lncRNAs that regulate expression of distant genes; Unclear, lncRNAs that regulate gene expression at transcriptional level, either in cis or trans. LncRNAs. Sense lncRNAs. Antisense lncRNAs. Intronic lncRNAs.

(detailed below). Regarding the mechanism of transcriptional interference, lncRNAs may influence the transcription activity of target genes through promoter binding to block PIC (preinitiation complex) formation, or by interacting with transcription factors (Fig. 2A). Such cis-lncRNAs may be transcribed from genes’ promoter regions. For instance, DHFR upstream transcripts, ~0.8–7.3 kb lncRNAs from the promoter region of DHFR, can form stable triplex structures with the promoter of DHFR and interact with TFIIB to efficiently dissociate PIC. Likewise, SRG1 RNAs, ~0.4–1.9 kb lncRNAs from the promoter region of SER3 (Ser3p), are found to cover the promoter of SER3 coding gene to prevent transcription factor binding to the promoter and, thus, repress the expression of SER3. Also, the transcription of lncRNAs in promoter regions may induce chromatin remodeling and, thus, activate the downstream protein-coding genes’ transcription, such as fbpl1 promoter RNAs. Cis-lncRNAs that function through chromatin modification often recruit chromatin modification complexes, e.g., PRC (polycomb repressive complex) or Rpd3S HDAC (Rpd3 small histone deacetylase complexes). The most studied chromatin modification complex is PRC and one well-known example is Xist (a 19 kb lncRNA in human), which binds to PRC2 to induce H3K27me3 modification and, thus, leads to transcriptional silencing of genes on the X chromosome. Similar examples can be found in MEG3 (maternally expressed 3) (a ~1.6 kb lncRNA) and COLDAIR (a ~1.1 kb lncRNA). Also, there are other chromatin modification complexes that are recruited by lncRNAs. GAL10-ncRNA (Gal10p-noncoding RNA) (a ~4 kb lncRNA) has been reported to recruit the Rpd3S HDAC complex, resulting in a decrease in some histone 3 acetylation to repress the expression of GAL1.4 Although the above-mentioned examples of PRC and Rpd3S HDAC are all implicated in negative regulation of gene expression, interaction of lncRNA with chromatin modification complexes is also involved in positive regulation. A case example is HOTTIP (HOX transcript at the distal tip) (a ~3.8 kb lncRNA in human) that recruits a MLL chromatin modifying complex to maintain a domain of active chromatin over the 5′ end of HOXA (homeobox A cluster) gene cluster.67

**Trans-lncRNAs.** Although it may be easier for lncRNAs to influence genes in their immediate vicinity probably based on sequence complimentary to the locus from which they are transcribed, lncRNAs can also function in trans-acting mode to target distant gene loci. For instance, HOTAIR (HOX antisense intergenic RNA), a ~2.2 kb lncRNA that is transcribed from the HOXC (homeobox C cluster) gene locus in chromosome 12, can be transported by the Suz-Twelve protein to regulate the homologous target sites at HOXD (homeobox D cluster) gene locus in chromosome 2.24 Also, HOTAIR are found to bind to many other genomic loci that tend to possess specific DNA motifs and influence gene expression by recruiting chromatin modification complexes.
Therefore, unlike *cis*-lncRNAs, *trans*-lncRNAs may function independently of sequence complementary to target gene locus. In addition to chromatin modification complexes, they may bind to transcription elongation factors or RNA polymerases to affect transcription. It is reported that 7SK RNA (a ~330 bp lncRNA) functions as a central scaffold to coordinate protein-protein interactions in 7SK snRNP (small nuclear ribonucleoproteins), which comprises transcription elongation factor—P-TEFb (positive transcription elongation factor b). This activity consequently leads to repression of transcription elongation at many gene loci. Another lncRNA, *B2 SINE* RNA, has been found to stably bind to polymerase II complex to block its activity during heat shock response.

**Mechanisms of Functioning**

While only a small number of lncRNAs has been well documented, it is believed that lncRNAs are involved in a wide variety of cellular molecular functions. According to their mechanisms of functioning, lncRNAs roughly fall into three groups that affect transcriptional regulation, post-transcriptional regulation or other functions (Fig. 2).

**Transcriptional regulation.** As mentioned above, there is a large number of lncRNAs that regulate gene transcription through transcriptional interference (e.g., *DHFR* upstream transcripts, *SRG1* RNAs, 7SK snRNA, *B2 SINE* RNA) and chromatin remodeling (e.g., *fbp1* (fructose-1,6-bisphosphatase-1) promoter RNAs, *Xin*, *MEG3*, *GAL10*-ncRNA, *HOTAIR*, *HOTTIP* and *COLDAIR*). Therefore, lncRNAs responsible for
transcription regulation can be sub-divided according to the mechanism of their functioning: (1) transcriptional interference and (2) chromatin remodeling (Fig. 2A). Besides, there are other related functional mechanisms, for example, the regulation effect: a set of lncRNAs transcribed from enhancers are termed eRNAs (enhancer RNAs) as they positively regulate genes' transcription, such as ncRNA-a1 (activating long ncRNA 1), Eif-2 (embryonic ventral forebrain-2) RNA, Alpha-250/Alpha-280.

**Post-transcriptional regulation.** There are two common post-transcriptional regulation mechanisms that IncRNAs get involved in, namely, splicing regulation and translational control (Fig. 2B). IncRNAs that influence mRNA splicing may function through binding to or modulating splicing factors, or directly hybridizing with mRNA sequences to block splicing. MIAF (myocardial infarction associated transcript), a ~9–10 kb IncRNA, contains strong intron branch point sequences (UACUAAC repeats) and is able to bind to SF1 (splicing factor 1) to inhibit splicing and splicesomal complex formation. Malat1 (metastasis-associated lung adenocarcinoma transcript 1), a ~7 kb IncRNA, can bind to SR splicing factor [serine–arginine (SR)-rich splicing factor] and regulate its distribution in nuclear speckle domains. Also, it is suggested that Malat1 may modulate the pools of phosphorylated SR and, thus, influence alternative splicing of pre-mRNAs. Additionally, other splicing regulation mechanisms may exist. LUST (LUCA-15-specific transcript), a ~1.4–2.4 kb IncRNA, is the antisense transcript of RBM5 (RNA binding motif protein 5) and is hypothesized to regulate the expression of RBM5 splice variants through masking a sense-strand regulatory sequence.

IncRNAs that participate in translational control may function through binding to translation factors or ribosome. There are two IncRNAs, BC1 (brain cytoplasmic RNA 1) and BC200 (200 nt brain cytoplasmic RNA), which can bind eIF4A (eukaryotic translation initiation factor 4A), PABP (poly(A)-binding protein) and other factors, to repress translation initiation by blocking assembly of the required complex. snR (small NFI90-associated RNAs), a cytoplasmic IncRNAs, can bind to ribosome, presumably influencing translation of mRNAs. Gadd7 (growth arrested DNA-damage inducible gene 7), which is associated with active translation, is hypothesized to bind to ribosome. It should be noted that in some cases translation and splicing are associated with each other. Zeb2 (zinc finger E-box binding homeobox 2) translation requires retention of an intron. Zeb2NAT (Zeb2 natural antisense transcript) (a lncRNA that is more than 1.2 kb), which overlaps the 5’ splicing site of an intron, can inhibit splicing of the intron to allow translation of Zeb2.

Aside from splicing regulation and translational control, there are other post-transcriptional regulation mechanisms utilized by IncRNAs. The findings of siRNA (small interfering RNA) mechanism and competing endogenous RNAs have opened up new aspects of post-transcriptional regulation and recent studies suggest that IncRNAs are also implicated in these processes, displaying direct siRNA mechanisms or interfering with miRNAs. LncRNAs may function as natural antisense inhibitors to promote degradation of mRNA (Fig. 2B). It has been found that 21A, a ~300 bp lncRNA, which shows high sequence homology to CENP-F (centromere protein F) intronic portions, can reduce CENP-F expression at both mRNA and protein level through antisense inhibitor. 1/2-dsRNA1 (half-STAU1-binding site RNA1), a ~0.7 kb lncRNA, has been found to bind to mRNAs’ 3’ UTR through Alu elements, and reduces mRNA abundance.

Moreover, there are many IncRNAs that interact directly or indirectly with miRNAs to stabilize target mRNAs. These IncRNAs are called ciRNAs (competing endogenous RNAs). For instance, inc-MD1 (long intergenic ncRNA that is associated with muscle differentiation), a ~0.5 kb lncRNA, acts as sponge/target mimic of miR-133 and miR-135 to regulate the expression of two transcription factors: MAML1 (mastermind-like protein 1) and MEF2C (myocyte-specific enhancer factor 2C), which activate the expression of muscle-specific genes. Similar examples can be found in IPSI (induced by phosphate starvation 1) RNA and HULC (highly up-reglated in liver cancer) RNA. Some pseudogenes can function as a sponge/target mimic for miRNAs to stabilize their homologous mRNAs, such as KRASp1 and PTENp1, which are pseudogenes of KRAS (V-K-ras2 Kirsten rat sarcoma viral oncogene homolog) and PTEN (phosphatase and tensin homolog), respectively. In addition, some antisense IncRNAs may bind to mRNA to mask the binding sites of miRNA and thus stabilize mRNA. BACE1AS (BACE1 antisense RNA), a ~2 kb lncRNA, has been reported to be activated in Alzheimer disease to form an RNA duplex with BACE1 (β-secretase 1) mRNA, which may mask the binding site for miR-485-5p and, thus, prevent translational repression of BACE1 mRNA by miRNA.

**Other mechanisms of lncRNA functioning.** In addition to transcriptional regulation and post-transcriptional regulation, lncRNAs may function through other mechanisms (Fig. 2C), such as protein localization, telomere replication, RNA interference, beyond transcription and translation regulation, etc. Considering the limitation of the currently available knowledge, it is difficult to categorize these IncRNAs into some more “stable” groups and, thus, we roughly place them into lncRNAs with “other functional mechanisms.” For instance, meiRNA, a ~0.5 kb lncRNA, is required for nuclear localization of Meiz2. ENOD40 RNA is required for correct subcellular localization of RNPs particles in legume plants. TERC (telomerase RNA component), which is part of telomerase reverse transcriptase, acts as template to extend telomere during DNA replication in eukaryote. In ciliated protozoa, maternal RNA templates have been found to guide reproducible rearrangement during the transition from germline nucleus to somatic nucleus. In addition, IncRNAs may be involved in RNA interference by regulation of Dicer1. It is suggested that rncs-1 (RNA non-coding and starvation upregulated) reduces Dicer-generated siRNA and affects levels of Dicer-regulated genes.

**Targeting Mechanisms of lncRNAs**

According to their mode of action, lncRNAs may also be classified based on their targeting mechanisms, mainly associated with the following categories: (1) signal: show cell type-specific expression and respond to diverse stimuli, such as Xist, COLDAIR; (2) decoy: bind and titrate away a protein target, but does not
exert any additional functions, such as DHFR upstream transcripts,\textsuperscript{23} PANDA;\textsuperscript{93} (3) guide: bind proteins and then direct the localization of ribonucleoprotein complex to specific targets, such as Xist,\textsuperscript{64} HOTAIR,\textsuperscript{24,68} (4) scaffold: serve as central platforms to bring together multiple proteins to form ribonucleoprotein complexes, such as HOTAIR,\textsuperscript{24,68} 7SL.\textsuperscript{24,95} Alternatively, lncRNAs can be grouped based on the types of interactions they make with their targets: RNA-RNA pairings, RNA-DNA hybrids, RNA structure mediated interactions and protein linkers.\textsuperscript{96}

However, a single targeting archetype in referring to mode of action may not be sufficient to fully describe one lncRNA since one lncRNA may contain multiple archetypes. There are many lncRNAs that are induced by endogenous and exogenous signals to express, and they also can possess the binding sites of chromatin modification complexes (such as PRC) to repress or activate expression of a set of genes, e.g., Xist,\textsuperscript{64} Air,\textsuperscript{97} COLDAIR,\textsuperscript{52} HOTTIP,\textsuperscript{67} HOTAIR,\textsuperscript{24,68} lncRNA-p21.\textsuperscript{38} Therefore, these lncRNAs operate in a dual mode as both signal and guide. Some signal lncRNAs may also function as decoys that bind and titrate away a protein target, e.g., PANDA.\textsuperscript{93} More complexly, some lncRNAs have more than two archetypes, such as HOTAIR that functions according to three archetypes: as anatomic signal, guiding the chromatin-modifying complexes to the target gene, and as a scaffold for PRC2 and LSD1.\textsuperscript{24,68}

Albeit, there are multiple groups of lncRNAs based on different classification methods as mentioned above, different groups appear to be linked closely to one another. LncRNAs that function via transcriptional interference mechanism may target gene loci through RNA-DNA hybrids and operate as the archetypes of decoy, e.g., DHFR upstream transcripts;\textsuperscript{23} lncRNAs that affect chromatin modification may target gene loci through RNA structure mediated interaction, which, thus, acts as the archetypes of signal, guide and scaffold, e.g., HOTAIR,\textsuperscript{24,68} ceRNAs that directly or indirectly regulate mRNA level target miRNAs or mRNAs through RNA-RNA pairing, may also operate as the archetype of decoy, for example, Krasp1 and PTENP1;\textsuperscript{86} lncRNAs that regulate splicing may function through RNA-RNA pairing to influence their targets and operate as the archetype of decoy (e.g., Malat1\textsuperscript{26}).

Perspectives

Here, we summarized classification methods of lncRNAs according to their four major characteristics (genomic location and context, effects on DNA sequences, functional mechanisms and targeting mechanisms). Based on function annotations presently available, we explored potential relationships between different classification categories and investigated biological features of different lncRNAs within each category. Although lncRNA could be described in terms of other features, such classifications of lncRNAs as summarized here are of fundamental importance for lncRNA studies, helpful for further investigation of specific groups of lncRNA, for generation of new hypothesis based on different lncRNA groups and for exploration of lncRNA underlying functional mechanisms.

Classification of lncRNAs. When studying lncRNAs, it is straightforward to investigate their functional features by classifying them into different groups. However, classification of lncRNAs is highly dependent on the current existing knowledge, thus requiring frequent validation of the classification system, exploring new classification systems and, when necessary, abandoning old ones. One possible example is the definition of lncRNA that is longer than 200 nt. We investigated the length of lncRNAs and compared their distribution between human and mouse (data from NONCODE V3.0\textsuperscript{50}). Clearly, lncRNAs can be further divided into different groups based on their length distribution (Fig. 3): small-lncRNA (200–950 nt), medium-lncRNA (950–4,800 nt) and large-lncRNA (4,800 nt–). Density distributions of lncRNA length are shown in (A) and percentages of three lncRNA groups are depicted in (B).
lincRNAs that recruit chromatin modification complexes have attracted a lot of attention. It is reported that lincRNAs tend to bind to specific chromatin-modifying complexes, so that lincRNA-associated chromatin modification may represent a specific regulation mechanism. Additionally, lincRNAs recruiting the same chromatin-modifying complex may diverge greatly between different species. Therefore, lincRNAs may affect specific variety of target genes in different biological processes, such as PRC2, which regulates a variety of prostate cancer implicated genes through recruiting PRC2. These findings shed lights on the functional significance of the lincRNA group (that functions through recruiting PRC2) and stimulate further work on other chromatin-modifying complexes. Notably, there is a large number of antisense lincRNAs, which function through recruiting chromatin modification complexes to regulate transcription activity according to available evidences (Table 1). It has been suggested that most of the PRC2 binding transcripts are antisense and sense lincRNAs rather than lincRNAs. Therefore, it would be desirable for future studies to explore more chromatin complex binding features of the antisense and sense lincRNAs.

**Regulation network of lincRNA and small ncRNA.** The discovery of lincRNAs and their regulatory roles challenges the initially miRNA-centered regulatory networks, which may help comprehensively understand gene regulation by both small ncRNA and lincRNA and, accordingly, provide new insights into complex processes of gene regulation. As mentioned above, cell RNAs, a group of lincRNAs interacting with miRNAs, can impose an additional level on post-transcriptional regulation. In addition to those lincRNAs that act as sponge/target mimic of miRNAs, it should be noted that more than 50% of the sense and antisense transcript pairs may be composed of coding and antisense non-coding transcripts, and the formed RNA duplex may also influence the interaction between miRNAs and their target mRNAs. According to our previous integrative analysis of miRNAs and mRNAs in NSCLC (non-small cell lung carcinoma), we found that most of the target mRNAs do not vary significantly, in spite of the dramatic increase or decrease of their miRNAs, indicating that gene regulation network may be more sophisticated.

### References

1. Ozaizi Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, et al; FANTOM Consortium; RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature 2002; 418:597-600; PMID:12068581; http://dx.doi.org/10.1038/nature02669.
2. Carinci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, et al; FANTOM Consortium; RIKEN Genome Exploration Research Group and Genome Science Group (Genome Newton Project Core Group). The transcriptional landscape of the mammalian genome. Science 2005; 309:1595-9; PMID:16141072; http://dx.doi.org/10.1126/science.1112014.
3. Johnson JM, Edwards S, Shoemaker D, Schadt EE. Dark matter in the genome: evidence of widespread transcription directed by microarray tiling experiments. Trends Genet 2005; 21:95-102; PMID:15661355; http://dx.doi.org/10.1016/j.tig.2004.12.009.
4. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 2012; 22:1775-89; PMID:22955988; http://dx.doi.org/10.1101/gr.132159.111.
5. Basia B, Ji J, Khurana J, Wood E, Riekh, B, Guldberg WE Jr., et al. Long noncoding RNAs are rarely translated in two human cell lines. Genome Res 2012; 22:1646-57; PMID:22955977; http://dx.doi.org/10.1101/gr.134707.111.
6. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet 2009; 10:155-9; PMID:19188922; http://dx.doi.org/10.1038/nrg2521.
7. Williez J, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. Genes Dev 2009; 23:1494-504; PMID:19571179; http://dx.doi.org/10.1101/gad.180909.
8. Tafj RJ, Pang KC, Mercer TR, Dinger M, Mattick JS. Non-coding RNA: regulators of disease. J Pathol 2010; 220:126-39; PMID:19882673; http://dx.doi.org/10.1002/path.2638.
9. Kapranov P, Cheng J, Dike S, Nie DA, Darragrupa R, Willingham AT, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science 2007; 316:1484-8; PMID:17510325; http://dx.doi.org/10.1126/science.1138541.
10. Amaral PP, Clark MB, Gascoigne DK, Dinger ME, Mattick JS. IncRNAdb: a reference database for long noncoding RNAs. Nucleic Acids Res 2011; 39(Database issue):D146-51; PMID:2112873; http://dx.doi.org/10.1093/nar/gkq1138.
11. Brown CJ, Hendrich BD, Rupert JL, Lafrenière RG, Xing Y, Lawrence J, et al. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell 1992; 71:527-42; PMID:1423611; http://dx.doi.org/10.1016/0092-8674(92)90528-M.
12. Clemson CM, McNeil JA, Willard HF, Lawrence JB. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nucleate/chromosome structure. J Cell Biol 1996; 132:259-75; PMID:8636206; http://dx.doi.org/10.1083/jcb.132.3.259.

Considering their similar background in gene expression and function overlap between lncRNA and small ncRNA, it is attractive to investigate whether there is a certain evolutionary association between the two components. It is found that some lncRNAs contain miRNAs in their gene locus. H9 (gene comes from colon pH19), a ~2.3 kb lncRNA, contains mir-675 in its exon and also serves as a precursor of this miRNA in addition to its transcription regulation activity. Also, IncRNA LOC554202 contains mir-31 in its intron, and both miR-31 and the host IncRNA are found to be lowly expressed in triple-negative breast cancer. It has been found that five IncRNAs, namely, MEG3, MEG8 (maternally expressed 3), MEG9 (maternally expressed 3), antiPeg11 (antisense transcript to Peg11/Rtl1) and Rian (RNA imprinted and accumulated in nucleus), contain a lot of miRNAs and snoRNAs and function through transcriptional regulation; interestingly, these five IncRNAs and their endogenous small ncRNA may target the same gene, but it remains unclear whether those five IncRNAs can serve as precursors of their corresponding miRNAs. In addition, it is suggested that some lncRNAs in human may be preferentially post-processed into snoRNAs. Taken together, future studies focused on this aspect may bring new unexpected insights into the evolutionary relationships between small ncRNA and lncRNA.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**Acknowledgments**

We thank Hao Wu and Gang Wu for their valuable comments on this manuscript. This work was supported by grants from National Natural Science Foundation of China (Grant No. 31200978 to Z.Z.), the Ministry of Science and Technology of the People’s Republic of China.
22. Ponting CP, Oliver PL, Reik W. Evolution and function of non-coding RNAs. Nature 2004; 431:757; PMID:15483594; http://dx.doi.org/10.1038/nature03166.

19. van Bakel H, Nislow C, Blencowe BJ, Hughes TR. Most "non-coding" complementary DNAs. Nature 2004; 431:757; PMID:15483594; http://dx.doi.org/10.1038/nature03166.

17. Pang KC, Frith MC, Mattick JS. Rapid evolution of non-coding RNAs. Nature 2004; 431:757; PMID:15483594; http://dx.doi.org/10.1038/nature03166.

15. Bernstein HD, Zopf D, Freymann DM, Walter P. Non-coding RNA regulation of the p53 tumor suppressor pathway. Proc Natl Acad Sci U S A 2010; 107:112-7; PMID:20178073; http://dx.doi.org/10.1073/pnas.0907559107.

13. Swiezewski S, Liu Z, Lee S, Schmitt J, Dean C. Cold-shock induction of the Escherichia coli arabinose operon. Genes Dev 2003; 17:1345-8; PMID:12953267; http://dx.doi.org/10.1101/gad.134554.103.

11. Sinicrope FA, Santin AL, Escandon E, Lee J, Yamasaki S, Jelinek J, Mitchell SP, Parry DA, Chiba S, Lipton DS. Potential structural role of non-coding and coding RNAs in the organization of the cytoskeleton at the vegetal cortex of Xenopus oocytes. Development 2005; 132:3445-57; PMID:16000384; http://dx.doi.org/10.1242/dev.01919.

9. Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS. Mouse transcriptome: neutral evolution of 'non-coding' complementary DNAs. Nature 2004; 431:757; PMID:15483594; http://dx.doi.org/10.1038/nature03166.

7. van Bakel H, Nislow C, Blencowe BJ, Hughes TR. Most "non-coding" complementary DNAs. Nature 2004; 431:757; PMID:15483594; http://dx.doi.org/10.1038/nature03166.

5. Li B, Dinger ME, Wilusz JE, Amaral PP, Mattick JS. Novel RNAs identified across a prostate cancer cohort identifies PCAT-1 as a prostate cancer marker. Proc Natl Acad Sci U S A 2010; 107:112-7; PMID:20178073; http://dx.doi.org/10.1073/pnas.0907559107.

3. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dole-Gratien V, FitzHugh W, Goodstein D, Hardenback JR, Hed Wallace J, Holt RA, Egholm M, Internacional Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 2001; 409:860-921; PMID:11237011; http://dx.doi.org/10.1038/35077670.

1. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science 2001; 291:1280-96; PMID:11189959; http://dx.doi.org/10.1126/science.1061510.
65. Hirota K, Miyoshi T, Kugou K, Hoffman CS, Shibata
63. Blume SW, Meng Z, Shrestha K, Snyder RC, Emanuel
59. Girard G, Roussis A, Gultyaev AP, Pleij CW, Spaink
58. Leygue E. Steroid receptor RNA activator (RanRA)
57. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
56. Wang KC, Yang YW, Liu B, Sanyal A, Corces-Vargas
53. Nguyen VT, Kiss T, Michels AA, Bensaude
52. The 5'-untranslated RNA of the human dhfr
51. Parrott AM, Tsi A, Batchu P, Ryan K, Ozer HL, Tian
50. Girard G, Roussis A, Gultyaev AP, Pleij CW, Spaink
49. Leygue E. Steroid receptor RNA activator (RanRA)
48. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
47. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
46. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
45. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
44. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
43. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
42. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
41. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
40. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
39. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
38. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
37. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
36. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
35. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
34. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
33. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
32. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
31. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
30. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
29. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
28. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
27. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
26. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
25. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
24. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
23. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
22. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
21. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
20. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
19. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
18. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
17. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
16. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
15. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
14. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
13. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
12. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
11. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
10. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
9. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
8. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
7. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
6. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
5. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
4. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
3. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
2. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,
1. Zhang X, Rice K, Wang Y, Chen W, Zhong Y,