

Long noncoding RNAs in biology and hematopoiesis

Vikram R. Paralkar¹ and Mitchell J. Weiss²

¹Division of Hematology/Oncology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; and ²Division of Hematology, The Children's Hospital of Philadelphia, Philadelphia, PA

Genome and transcriptome sequencing have revealed a rich assortment of non-coding RNAs in eukaryote cells, including long noncoding RNAs (lncRNAs), which regulate gene expression independent of

protein coding potential. lncRNAs modulate protein coding gene expression in many cell types by regulating multiple processes, including epigenetic control of transcription, mRNA stability, and protein

localization. Although little is known about lncRNAs in hematopoiesis, they are likely to exert widespread roles in this process. (Blood. 2013;121(24):4842-4846)

New layers of genetic regulation through noncoding RNAs

A remarkable discovery of whole-genome sequencing studies initiated at the turn of this century was that protein coding exons account for less than 2% of mammalian DNA.¹ Subsequent large-scale transcriptome studies have established that at least two-thirds of the genome is nonetheless transcribed into RNA, exposing novel and exciting layers of genetic regulation.² Throughout eukaryotic evolution, in yeast, nematodes, flies, and mammals, genomic regions previously believed to be so-called “junk DNA” are now known to be rich with transcripts encoding a large and diverse population of RNAs without protein-coding potential, referred to as noncoding RNAs (ncRNAs). Studies of these RNAs have provided new insights into the development of virtually all mammalian tissues and opened up new fields of biology. It is now known that a class of ncRNAs termed microRNAs exerts diverse roles in normal and pathological hematopoiesis.³ The same is likely to be true for long noncoding RNAs (lncRNAs), which are the topic of this *Blood* Spotlight review.

Classification of ncRNAs

The earliest recognized ncRNAs (ribosomal RNAs and transfer RNAs) make up more than 95% of cellular RNA content. Other ncRNAs are broadly classified on the basis of length. Short ncRNAs are less than 200 nucleotides and include microRNAs, small nucleolar RNAs, piwi-interacting RNAs, and others. By arbitrary definition, lncRNAs are greater than 200 nucleotides and represent a diverse group with many functions. This class of RNAs was appreciated more than 20 years ago but is now rapidly gaining prominence as a pervasive player in gene regulation. Here we provide a succinct summary of lncRNAs, with a focus on potential roles in hematopoiesis. For more in-depth information on lncRNAs, we refer readers to several excellent review articles.⁴⁻⁶

What are lncRNAs?

The concept that long RNAs can regulate gene expression by nonclassical mechanisms has been known for some time. For

example, Xist, a lncRNA that mediates X-chromosome inactivation, was discovered in 1991.⁶ This lncRNA and a handful of others were considered to be relatively infrequent until large-scale transcriptome sequencing identified thousands more.² Now, candidate lncRNAs are recognized by computational algorithms that identify transcripts without protein coding potential, as evidenced by their lack of long open reading frames, conserved codons, or homology with protein databases. Follow-up studies have revealed that some of these RNAs have noncoding functions mediated through direct interaction with effector proteins, DNA, and/or protein-coding RNAs.

lncRNAs range in size from 200 nucleotides to beyond 10 kilobases. They can be polyadenylated or not, can be nuclear or cytoplasmic, and are generally expressed at lower levels compared with protein-coding mRNAs. Similar to protein-coding genes, many lncRNA genes are bound by essential cell-type-specific nuclear factors, transcribed from what appear to be conventional promoters, and are spliced.⁷ Other lncRNAs, both polyadenylated and nonpolyadenylated, arise from enhancers.⁸ It is estimated that many thousands of lncRNAs are encoded in the human genome and are expressed in exquisitely tissue-specific patterns. Thus, it is likely that deep sequencing of specific cell types, including normal and diseased blood lineages, will discover new lncRNAs.

lncRNAs regulate gene expression through diverse mechanisms

lncRNAs are versatile molecules that can interact physically and functionally with DNA, other RNAs, and proteins, either through nucleotide base pairing or via formation of structural domains generated by RNA folding. These properties endow lncRNAs with a versatile range of capabilities that are only beginning to be appreciated. The best-characterized role of lncRNAs is in epigenetic regulation. For example, in female (XX) cells, one X chromosome produces the prototypical lncRNA Xist, which coats that chromosome and recruits repressive chromatin complexes to condense and silence it in a process termed Lyonization.⁹ The lncRNA HOTAIR

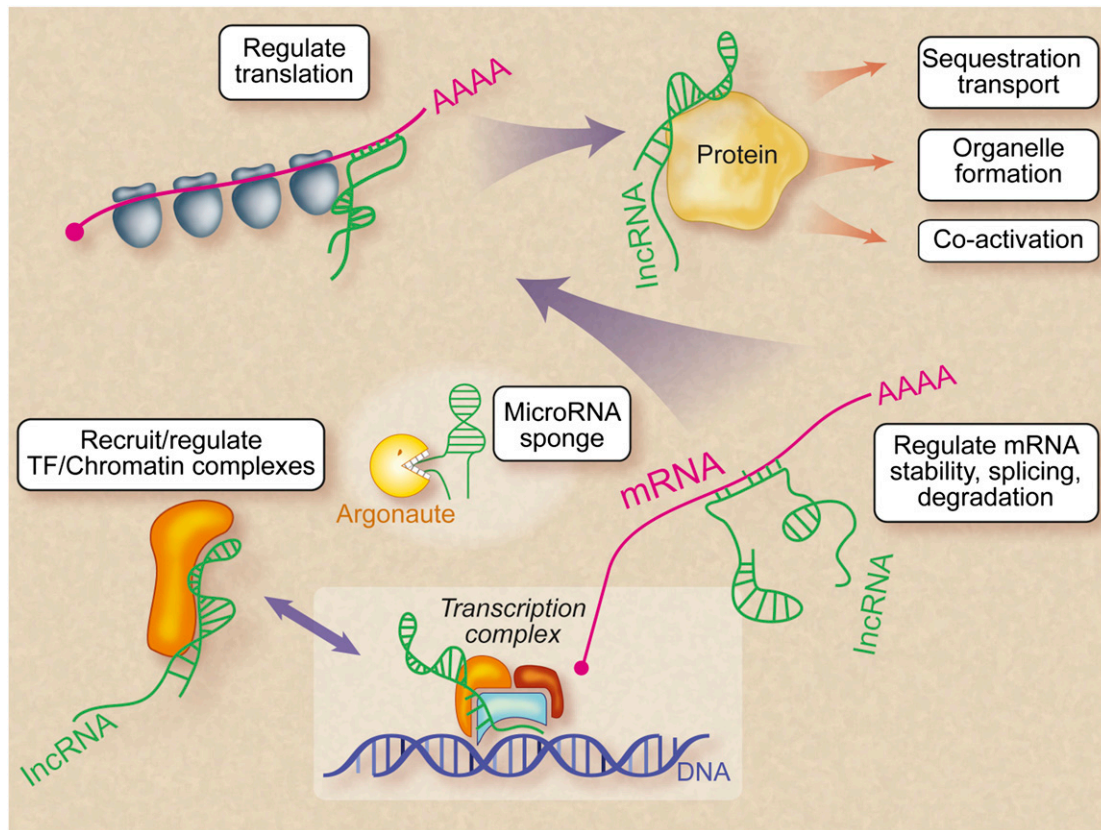


Figure 1. Mechanisms of LncRNA action. In numerous tissues, lncRNAs (indicated in green) have been shown to regulate gene expression at multiple levels: chromatin, transcription, mRNA, translation, and protein. Hematopoietic lncRNAs may act at any of these levels. “MicroRNA sponge” refers to the ability of lncRNAs to sequester cellular microRNAs and prevent them from binding mRNA targets.¹⁴ Professional illustration by Debra T. Dartez.

is transcribed from the HOXC locus and silences HOXD locus genes by recruiting repressive complexes.¹⁰ LncRNAs were thought initially to act only on neighboring genes (in cis), and HOTAIR was the first lncRNA shown to act on another chromosome (in trans). Numerous *cis*- and *trans*-acting lncRNAs have now been described. In addition to their role in recruiting chromatin-modifying proteins, lncRNAs have been shown to interfere with protein–DNA binding,¹¹ organize nuclear architecture,¹² regulate mRNA stability and translation,¹³ modulate mRNA levels by competing for microRNA binding,¹⁴ and directly alter protein function¹⁵ (Figure 1). LncRNAs are therefore capable of acting at multiple levels in the hierarchy of gene expression.

LncRNAs regulate normal biological processes, including hematopoiesis

LncRNAs modulate cell survival, division, and differentiation to facilitate the development of many tissues. For example, RNAi knockdown of 150 mouse embryonic stem cell lncRNAs identified more than 20 lncRNAs that are required for maintenance of pluripotency.¹⁶ Cyrano is a zebrafish lncRNA required for normal embryonic development, and its depletion produces central nervous system and notochord defects. HOTTIP, a lncRNA transcribed adjacent to the HOXA locus, is required for the normal expression of multiple genes therein. Depletion of HOTTIP in the embryonic chick forelimb produces shortened, abnormal limb bones.¹⁷ Altogether, functional studies have identified about 100 lncRNAs that regulate

various aspects of vertebrate cell biology and development, indicating that this class of molecules exerts broad critical functions.

Only a few hematopoietic lncRNAs have been studied closely. LincRNA-EP5 is a mouse nuclear lncRNA that is upregulated during terminal erythropoiesis and represses Pycard, a pro-apoptotic gene. RNAi knockdown of LincRNA-EP5 in erythroblasts derepresses Pycard, causing apoptosis. Conversely, *in vitro* overexpression of LincRNA-EP5 protects erythroblasts from apoptosis caused by erythropoietin deprivation.¹⁸ The lncRNA EGO regulates eosinophil granule protein expression, and HOTAIRM1, a lncRNA in the HOXA cluster, is upregulated during myeloid development and is required for normal induction of certain HOXA and myeloid differentiation genes.⁴ Several abstracts presented at the 2012 American Society of Hematology Annual Meeting identified and categorized lncRNAs in hematopoietic stem cells, myeloid cells, erythroblasts and megakaryocytes. We anticipate that functions for some of these lncRNAs will emerge during the next few years.

LncRNAs and human disease

A handful of lncRNAs have been shown to influence malignant transformation and progression. Elevated expression of HOTAIR in breast cancer cells worsens survival by promoting tissue invasion and metastases.¹⁹ MALAT1 lncRNA is highly expressed in various malignancies, and disruption of its genomic locus impairs the metastatic potential of a human lung cancer cell line in animal models.²⁰ The lncRNA p15AS is antisense to the tumor-suppressor

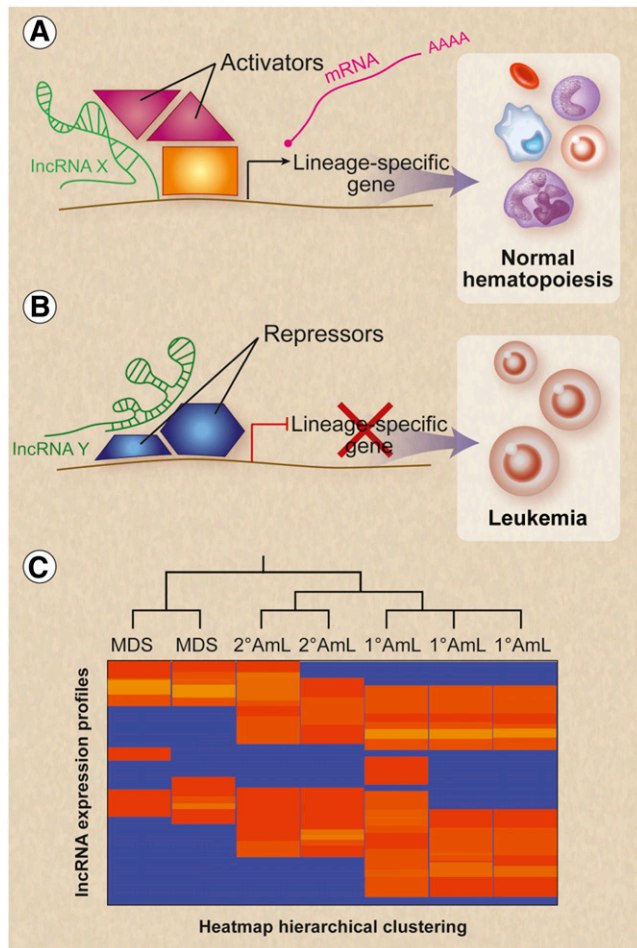


Figure 2. Potential functions for lncRNAs in normal and malignant hematopoiesis. (A) On the basis of their known activities in many cell types (Figure 1), lncRNAs may function during normal hematopoiesis to recruit transcription factors, coactivators, and activating chromatin factors to lineage-specific genes, thereby modulating gene expression to facilitate the differentiation and maturation of blood lineages. lncRNAs are indicated in green. (B) Absence or inappropriate expression of relevant lncRNAs could alter transcription to block normal differentiation and activate leukemic pathways. (C) Hematological malignancy subtypes may express distinct lncRNA expression profiles that reflect cell type of origin, mechanism of transformation, treatment sensitivities, and prognosis. Professional illustration by Debra T. Dartez.

gene p15, which it suppresses by inducing local chromatin condensation. The expression of p15AS is elevated and is associated with p15 silencing in acute myeloid leukemia and acute lymphoblastic leukemia blasts compared with normal hematopoietic cells.²¹ Deletion of Xist in the hematopoietic compartment of female (but not male) mice produces a highly penetrant and aggressive form of myeloproliferative/myelodysplastic syndrome and leukemia through de-repression of numerous genes on the second X-chromosome, a finding that may explain why human malignancies sometimes show supernumerary X chromosomes.²² Most recently, RNA-Seq of Mycosis fungoides and Sézary syndrome cells from affected patients identified 12 unannotated lncRNAs that are differentially expressed in normal and malignant T-lymphocytes.²³

Dysregulation of lncRNAs can also contribute to inherited diseases. Fascioscapulohumeral muscular dystrophy is a rare autosomal dominant disorder associated with deletions on chromosome 4q35. These deletions appear to activate transcription of DBE-T, a chromosome 4 lncRNA that recruits chromatin-activating complexes and abnormally activates neighboring genes that mediate the

disease phenotype.²⁴ Germline mutations that interfere with lncRNA function can also cause diseases. Two families with brachydactyly type E (shortened metacarpals and metatarsals) have translocations of chromosome 12, leading to the displacement of lncRNA DA125942 from its location near *PTH1H* (an important chondrocyte gene) to another chromosome. The normal role of DA125942 is to enhance *PTH1H* expression in cis, but on translocation, the lncRNA is displaced to another chromosome, causing reduced *PTH1H* expression and abnormal cartilage development.²⁵ lncRNAs are also implicated in the pathogenesis of Prader-Willi and HELLP (hemolysis, elevated liver enzymes, low platelet count) syndromes.²⁵ Other than the examples listed earlier, limited lncRNA data are available for human hematopoietic abnormalities, although this is likely to change as noncoding transcripts are extensively mapped and DNA sequence analysis of congenital and acquired hematological disorders interrogates noncoding regions of the genome.

Challenges, controversies, and future directions

Bioinformatic methods have identified thousands of lncRNAs, and an interesting debate has emerged regarding how many of these actually exert important biological functions.²⁶ For example, enhancers frequently produce lncRNA transcripts. Some of these may simply represent nonfunctional byproducts of transcription factor chromatin occupancy,²⁷ whereas others are apparently required in cis for the regulatory activity of the enhancer.¹⁷ To date, only a small fraction of lncRNAs has been studied mechanistically, and publications are probably biased toward lncRNAs with detectable functions. Systematic functional studies on large numbers of lncRNAs will help address their general significance. However, screening studies must be interpreted with caution. Manipulation of the lncRNAs MALAT1 and NEAT1 produce distinct effects in tissue culture cells, yet the corresponding knockout mice exhibit no obvious phenotype.²⁸ Further testing may reveal abnormalities in these mutant mice, analogous to many microRNAs, in which in vivo loss-of-function phenotypes are only apparent under specific stresses.

Understanding the functions of lncRNAs is also complicated by their interesting and unusual evolutionary properties. For example, lncRNA promoters and loci show greater sequence conservation than background DNA,⁷ although it appears that many lncRNAs detected in one species might not even be transcribed in another species.²⁹ Whether this implies lack of function or an ability to gain and lose function more rapidly during evolution than protein-coding genes requires further study. Supporting the latter possibility, a lncRNA required for cardiovascular development in mice is undetectable in rats and humans.³⁰ Thus, failure to exhibit cross-species conservation in expression does not necessarily indicate lack of lncRNA function. In addition, some lncRNAs exhibit little overall sequence homology, yet seem to share common functions. For example, although the human or mouse orthologs of Cyran exhibit low sequence similarity to the zebrafish gene, they can partially substitute for its loss.³¹ In this and other cases, evolutionary constraint of lncRNAs may occur at secondary and tertiary structural levels as well as selectively on small domains, all of which will be challenging to define.

It is likely that future studies will uncover prominent roles for lncRNAs in hematopoiesis and associated disorders. The discovery

of lincRNA-EPS, which promotes the survival of murine erythroblasts,¹⁸ may help to identify novel pathways that govern red blood cell production and may potentially be manipulated for therapeutic purposes. Similarly, multiple other lincRNAs are likely to promote the specification, maintenance, and function of other hematopoietic lineages.

Given their established roles in gene regulation, lincRNAs are also likely to influence hematological malignancies by modulating the expression of oncogenes, tumor suppressors, cell cycle regulators, and proteins that regulate apoptosis via multiple diverse mechanisms outlined in Figure 1. For example, a canonical function of lincRNAs is to modulate epigenetic landscapes. As mutations or altered expression of genes encoding epigenetic modifier proteins are known to promote various leukemias,³² derangements of currently unidentified lincRNA genes may produce similar effects (Figure 2). LincRNAs could also promote malignant transformation through their known abilities to regulate splicing, a process that has been implicated recently in myelodysplastic syndromes and myeloproliferative neoplasms.³³ To investigate these possibilities, it is important for future genome and transcriptome sequencing studies of malignant hematopoietic disorders to examine not only protein-encoding genes but also lincRNAs. In this regard, lincRNAs exhibit highly cell-type-specific expression³⁴ and have already been shown to have differential expression profiles in cutaneous T-cell lymphomas.²³ It will be of interest to determine whether lincRNA expression patterns are useful for substratifying hematopoietic malignancies with respect to underlying mutational profiles, prognosis, and treatment sensitivities (Figure 2C).

LncRNAs may eventually themselves become pharmacologic targets that enhance our ability to modulate pathological gene expression. For example, many (but not all) lincRNAs can be suppressed by RNA interference. In addition, lincRNAs are likely to act by adopting specific folding conformations, which could

present surfaces that are amenable to manipulation by peptides, other RNAs, or small molecules. In many ways, the lincRNA field today is where the microRNA field was a decade ago, and there is considerable enthusiasm for the potential to mine biological insights from its study. We predict that the information unearthed by these studies will advance science and medicine in exciting and unexpected ways.

Acknowledgments

The authors thank Harvey Lodish, Gerd Blobel, Douglas Higgs, Wenqian Hu, Juan Alvarez, and Ana Marques for comments on the manuscript. Noncoding RNA studies in our laboratory are funded by National Institutes of Health grants R01DK092318, P30 DK090969, and DK065806 and by the Roche Foundation for Anemia Research. V.R.P. is a recipient of the American Society of Hematology Research Training Award for Fellows, the ASH Fellow Scholar Award, and the Measey Research Fellowship Award.

Authorship

Contribution: V.R.P. and M.J.W. wrote the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Mitchell J. Weiss, The Children's Hospital of Philadelphia, 316B ARC, 3615 Civic Center Blvd, Philadelphia PA, 19104; e-mail: weissmi@email.chop.edu.

References

- Lander ES, Linton LM, Birren B, et al; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921.
- Carninci P, Kasukawa T, Katayama S, et al; FANTOM Consortium; RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group). The transcriptional landscape of the mammalian genome. *Science*. 2005;309(5740):1559-1563.
- O'Connell RM, Baltimore D. MicroRNAs and hematopoietic cell development. *Curr Top Dev Biol*. 2012;99:145-174.
- Hu W, Alvarez-Dominguez JR, Lodish HF. Regulation of mammalian cell differentiation by long non-coding RNAs. *EMBO Rep*. 2012;13(11):971-983.
- Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem*. 2012;81:145-166.
- Lee JT. Epigenetic regulation by long noncoding RNAs. *Science*. 2012;338(6113):1435-1439.
- Guttman M, Amit I, Garber M, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*. 2009;458(7235):223-227.
- Kim TK, Hemberg M, Gray JM, et al. Widespread transcription at neuronal activity-regulated enhancers. *Nature*. 2010;465(7295):182-187.
- Heard E, Disteche CM. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev*. 2006;20(14):1848-1867.
- Rinn JL, Kertesz M, Wang JK, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*. 2007;129(7):1311-1323.
- Kino T, Hurt DE, Ichijo T, Nader N, Chrousos GP. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci Signal*. 2010;3(107):ra8.
- Mao YS, Sunwoo H, Zhang B, Spector DL. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat Cell Biol*. 2011;13(1):95-101.
- Gong C, Maquat LE. lincRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature*. 2011;470(7333):284-288.
- Cesana M, Cacchiarelli D, Legnini I, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell*. 2011;147(2):358-369.
- Willingham AT, Orth AP, Batalov S, et al. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science*. 2005;309(5740):1570-1573.
- Guttman M, Donaghey J, Carey BW, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature*. 2011;477(7364):295-300.
- Wang KC, Yang YW, Liu B, et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*. 2011;472(7341):120-124.
- Hu W, Yuan B, Flygare J, Lodish HF. Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. *Genes Dev*. 2011;25(24):2573-2578.
- Gupta RA, Shah N, Wang KC, et al. Long noncoding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*. 2010;464(7291):1071-1076.
- Gutschner T, Hämmerle M, Eissmann M, et al. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res*. 2013;73(3):1180-1189.
- Yu W, Gius D, Onyango P, Muldoon-Jacobs K, Karp J, Feinberg AP, Cui H. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature*. 2008;451(7175):202-206.
- Yildirim E, Kirby JE, Brown DE, Mercier FE, Sadreyev RI, Scadden DT, Lee JT. Xist RNA is a potent suppressor of hematologic cancer in mice. *Cell*. 2013;152(4):727-742.
- Lee CS, Ungewickell A, Bhaduri A, et al. Transcriptome sequencing in Sezary syndrome identifies Sezary cell and mycosis fungoides-associated lincRNAs and novel transcripts. *Blood*. 2012;120(16):3288-3297.
- Cabianca DS, Casa V, Bodega B, Xynos A, Ginelli E, Tanaka Y, Gabellini D. A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell*. 2012;149(4):819-831.

25. Troy A, Sharpless NE. Genetic "lnc"-age of noncoding RNAs to human disease. *J Clin Invest*. 2012;122(11):3837-3840.
26. Kowalczyk MS, Higgs DR, Gingeras TR. Molecular biology: RNA discrimination. *Nature*. 2012;482(7385):310-311.
27. Kowalczyk MS, Hughes JR, Garrick D, et al. Intragenic enhancers act as alternative promoters. *Mol Cell*. 2012;45(4):447-458.
28. Zhang B, Arun G, Mao YS et al. The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Rep*. 2012;2(1):111-123.
29. Kutter C, Watt S, Stefflova K, et al. Rapid turnover of long noncoding RNAs and the evolution of gene expression. *PLoS Genet*. 2012;8(7):e1002841.
30. Klattenhoff CA, Scheuermann JC, Surface LE, et al. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell*. 2013;152(3):570-583.
31. Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell*. 2011;147(7):1537-1550.
32. Shih AH, Abdel-Wahab O, Patel JP, Levine RL. The role of mutations in epigenetic regulators in myeloid malignancies. *Nat Rev Cancer*. 2012;12(9):599-612.
33. Yoshida K, Sanada M, Shiraishi Y, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*. 2011;478(7367):64-69.
34. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*. 2011;25(18):1915-1927.