

Short Conceptual Overview

Nadia Vadaie and Kevin V. Morris*

Long antisense non-coding RNAs and the epigenetic regulation of gene expression

Abstract: Shortly after the completion of the human genome project in 2003, the Encode project was launched. The project was set out to identify the functional elements in the human genome, and unexpectedly it was found that >80% of the genome is transcribed. The Encode project identified those transcribed regions of the genome to be encoded by non-coding RNAs (ncRNAs). With only 2% of the genome carrying gene-encoding proteins, the conundrum was then, what is the function, if any, of these non-coding regions of the genome? These ncRNAs included both short and long RNAs. The focus of this review will be on antisense long non-coding RNAs (lncRNAs), as these transcripts have been observed to play a role in gene expression of protein-coding genes. Some lncRNAs have been found to regulate protein-coding gene transcription at the epigenetic level, whereby they suppress transcription through the recruitment of protein complexes to target loci in the genome. Conversely, there are lncRNAs that have a positive role in gene expression with less known about mechanism, and some lncRNAs have been shown to be involved in post-transcriptional processes. Additionally, lncRNAs have been observed to regulate their own expression in a positive feedback loop by functioning as a decoy. The biological significance of lncRNAs is only just now becoming evident, with many lncRNAs found to play a significant role in several human diseases.

Keywords: antisense; epigenetic; long non-coding RNA; ncRNA; transcription.

*Corresponding author: Kevin V. Morris, Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA 92037, USA; and School of Biotechnology and Biomedical Sciences, The University of New South Wales, Sydney 2052, NSW, Australia, e-mail: kmorris@scripps.edu

Nadia Vadaie: Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA 92037, USA

High-throughput studies have contributed to uncovering the complexity of the human transcriptome and the pervasive genome-wide transcription. The Encode project revealed that >80% of the human genome is transcribed, and that the majority of the transcripts do not encode proteins. These newly discovered ncRNAs are functionally different from the canonical ncRNAs, such as transfer RNA and ribosomal RNA, which are involved in general cellular processes. Notably, many of these new ncRNAs are not as ubiquitously expressed as the housekeeping ncRNAs and do not appear to be translated (1). The new ncRNAs are classified into small ncRNAs (<200 nucleotides), such as microRNAs, small interfering RNAs, PIWI-interacting RNAs, and long ncRNAs (lncRNAs), which are at least 200 nucleotides in length (2, 3). This review focuses on antisense lncRNAs and their regulatory roles.

lncRNAs can be categorized into several groups. For example, one type of lncRNA can be characterized with respect to its complementary protein-coding gene, which can be orientated in a sense and/or antisense direction. The lncRNAs and the protein-coding genes can therefore be in a head-head, tail-tail, or head-tail orientation. Other lncRNAs can be localized in introns (intronic) and within two genes (intergenic). Moreover, lncRNAs are expressed in a cell-/tissue-specific manner and under developmental stages, which highlights their functions in regulating gene expression (4, 5). lncRNAs have been found to modulate gene expression by various mechanisms, and a well-studied mechanism by which lncRNAs exert their functions is through their epigenetic regulation (6). This mechanism includes genetic imprinting and dosage compensation. In case of imprinting, the expression of either paternal or maternal allele is silenced. Imprinted lncRNAs, such as *Kcnqot1* and *Air*, which are expressed from paternal alleles, repress a cluster of protein-coding genes on maternal alleles, for instance, *Air* silences *Igf2r*, *Slc22a2*, and *Slc22a3*. The epigenetic mechanisms by which these imprinted lncRNAs suppress gene expression are DNA methylation of promoters and the formation of repressive chromatin state by the histone methyltransferase G9a (7–10).

In dosage compensation, one maternal X-chromosome is silenced by the lncRNA Xist, which interacts and recruits histone-modifying complexes to target sites, and this, in turn, results in the formation of repressive chromatin marks, such as H3K9me3 and H3K27me3 (11, 12). Moreover, lncRNAs can modulate the chromatin state by functioning as a structural scaffold, whereby they provide the platform for the formation of chromatin remodeling complexes. The lincRNA HOTAIR forms such a scaffold to gather the histone methyltransferase polycomb repressive complex 2 (PRC2) and the histone demethylase LSD1. Through these interactions and complex formations, HOTAIR represses the expression of the *HOXD* gene (13).

At the epigenetic level, lncRNAs repress transcription by interacting and recruiting histone and chromatin

remodeling proteins to target sites (Figure 1). These antisense lncRNAs can be expressed in a *cis*- or *trans*-based manner from RNA polymerase II promoters (14) and can be differentially localized within the cell (15). Some of the first such examples of antisense lncRNAs regulating protein-coding genes by targeting epigenetic complexes to the promoters of particular protein-coding genes were observed with antisense lncRNAs for tumor suppressor genes (16, 17). These lncRNAs were observed to epigenetically modulate the promoter of the targeted protein-coding gene, specifically altering the chromatin to a more compacted state (Figure 1). It is interesting to note that most of the tumor suppressor genes reported to date have been found to exhibit some level of antisense lncRNA expression associated with them. These tumor suppressor genes are also often the ones found to be

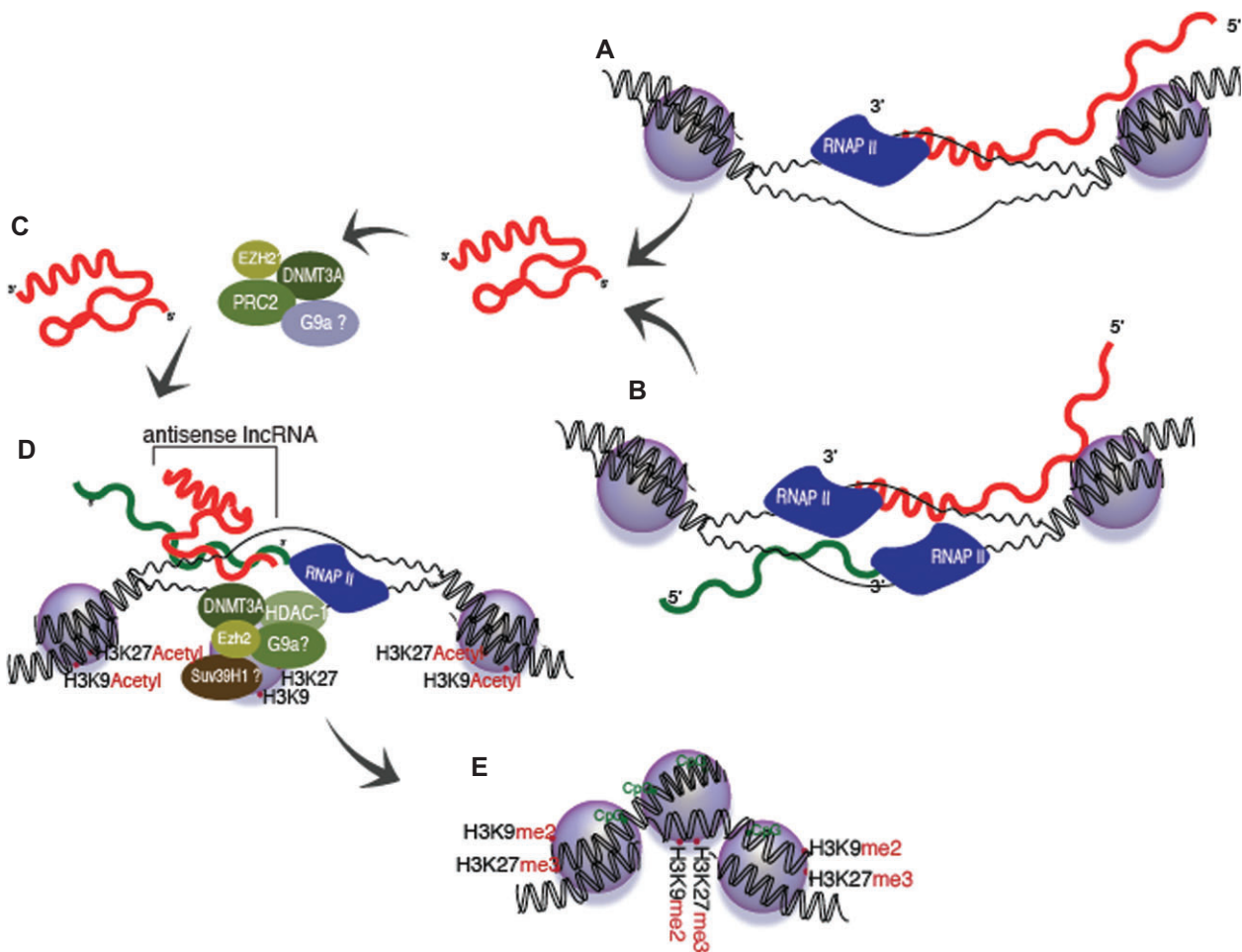


Figure 1 Antisense lncRNA-directed epigenetic regulation in human cells.

Antisense lncRNAs can be expressed in *trans* (A) or *cis* (B). The lncRNAs can then interact with various protein components (C): Ezh2, DNMT3a, PRC2, and G9a are shown schematically, as all have been observed to be involved in lncRNA-mediated function, although not as one complete complex (8, 9, 19, 20). The antisense lncRNA protein complex can then target epigenetic remodeling of the homology containing loci (D), resulting ultimately in compaction of the target loci (E) and epigenetic silencing.

epigenetically silenced in human cancers [reviewed in ref. (18)].

More recent observations have begun to suggest that pseudogenes, which are non-protein-coding transcripts containing sequence homology to known protein-coding genes, are active epigenetic modulators of their respective protein-coding counterpart (19, 20). In one study, the Oct4 transcription factor, which is a dominant player in stem cell control as well as transcriptional fidelity, was found to be under epigenetic regulation of an antisense lncRNA emanating from a pseudogene (19). In a more recent body of work, the PTEN tumor suppressor gene was found to be bimodally regulated by its pseudogene, PTENpg1, whereby both PTEN protein translation and transcription appear to be controlled by one pseudogene network node (20, 21). Such bimodal regulation raises the level of lncRNA regulatory complexity to a state that has not before been observed and suggests that a vast regulatory network of interesting functions remains to be determined with regard to lncRNA-directed epigenetic regulation.

Higher-ordered RNA:RNA complexes forming from lncRNAs have also been observed in the control of gene clusters. In mouse testis, tubulin and the assembly of microtubule structures appear regulated by sense and antisense lncRNAs forming higher-ordered RNA:RNA complexes (22). In another study, the antisense lncRNA AS1DHRS4, which emanates in a head-head manner from the *DHRS4* gene, regulates the gene cluster in both a *cis*- and *trans*-directed manner (23). AS1DHRS4 pairs with the sense counterpart in *cis* to mediate deacetylation of the *DHRS4* gene, whereas the antisense lncRNA in *trans* is involved in epigenetic-mediated regulation of the locus through interactions with G9a and enhancer of zeste 2 (*Ezh2*) (23). These studies illustrate the vast complexity of antisense lncRNAs in both form and function with regard to control of gene expression.

Some lncRNAs can be sense oriented relative to protein-coding genes, emanating across the promoters of genes (22, 24), and function in an allosteric blocking-based manner. While previous studies have found that promoter-associated RNAs are required for RNA-directed transcriptional gene silencing (25), these promoter-associated lncRNAs have been found to obstruct epigenetic activation by binding the translocated in liposarcoma (TLS) protein, which is an RNA-binding protein that is also known to bind CREB (cAMP response element binding protein)-binding protein CBP/adenovirus p300. The ultimate effect of this lncRNA-binding TLS is the inhibition of histone acetyltransferase activity (26).

Other lncRNAs can regulate gene expression independently of changing the chromatin state, and by

directly binding to transcriptional machinery and DNA elements within the protein-coding genes. The human dihydrofolate reductase *DHFR* gene encodes an lncRNA, which binds to the promoter and the transcription factor IIB, forming a complex. Through these interactions, the lncRNA disassembles the complex, which in turn leads to suppression of DHFR expression (27). Interestingly, lncRNAs can have a transcriptional activator role, for instance, the lncRNA EVF-2 forms a complex with the transcription factor *Dlx-2*, which then binds to the homeobox bicluster genes *Dlx-5* and *Dlx-6* and induces the activation and expression of the genes (28). Additionally, lncRNAs can have a regulatory function at the post-transcriptional level. The lncRNA ZEB2NAT modulates alternative splicing of the protein-coding gene *ZEB2*; in one case, the lncRNA splices an mRNA that evades the translational machinery and prevents the translation of ZEB2 protein, and in another instance ZEB2NAT interferes with the spliceosome and generates an alternatively spliced mRNA, which then can be translated (29). The lncRNA MALTA-1 is also involved in regulating the splicing by influencing the phosphorylation state of the splicing factors (30).

In addition to their role as transcriptional regulators of protein-coding genes, lncRNAs can modulate their own expression by interacting with other ncRNAs. For example, the lncRNA HULC binds to miRNA-372 and inhibits the expression of the miRNA target genes. One of these genes is a cAMP-dependent protein kinase catalytic subunit β (*Prkacb*), which, when inhibited, promotes the phosphorylation of the CREB domain of HULC. Then, phosphorylation of HULC leads to upregulation of its own expression (31).

lncRNAs elicit their functions both at the transcriptional and post-transcriptional levels; therefore, it is not surprising that the deregulation of lncRNAs has been implicated in diseases, such as cancers. For instance, aberrant expression of HOTAIR has been observed in breast, colon, and pancreatic cancers (32–34). Additionally, dysregulation of MALTA-1 has been linked to breast, colon, liver, lung, ovary, pancreas, and prostate cancers (35, 36). Moreover, some lncRNAs have been shown to become biomarker tools for cancer diagnosis. For instance, prostate cancer-related lincRNA PCA3 has been detected in urine samples of patients (37), and HULC in blood samples of hepatocellular cancer (38). The overexpression of lncRNAs in many cancers, and the growing evidence that lncRNAs are detected in cell-free environments, can shed more light on our understanding of the regulatory role of the lncRNAs in human diseases, and potentially can provide the therapeutic tools to target those diseases.

Acknowledgments: This publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award numbers R01AI084406 and P01AI099783-01 (KVM), and by the National Institute of Cancer of the National

Institutes of Health under award numbers CA151574 and R01 CA153124.

Received May 23, 2013; accepted June 17, 2013

References

- Banfai B, Jia H, Khatun J, Wood E, Risk B, Gundling WE Jr, Kundaje A, Gunawardena HP, Yu Y, Xie L, Krajewski K, Strahl BD, Chen X, Bickel P, Giddings MC, Brown JB, Lipovich L. Long noncoding RNAs are rarely translated in two human cell lines. *Genome Res* 2012; 22: 1646–57.
- Mattick JS, Makunin IV. Non-coding RNA. *Hum Mol Genet* 2006; 15 Spec No 1: R17–29.
- Chen LL, Carmichael GG. Long noncoding RNAs in mammalian cells: what, where, and why? *Wiley Interdiscip Rev RNA* 2010; 1: 2–21.
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 2007; 129: 1311–23.
- Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 2009; 458: 223–7.
- Hung T, Chang HY. Long noncoding RNA in genome regulation: prospects and mechanisms. *RNA Biol* 2010; 7: 582–5.
- Seidl CI, Stricker SH, Barlow DP. The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *EMBO J* 2006; 25: 3565–75.
- Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, Fraser P. The air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* 2008; 322: 1717–20.
- Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell* 2008; 32: 232–46.
- Peters J, Robson JE. Imprinted noncoding RNAs. *Mamm Genome* 2008; 19: 493–502.
- Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Wang H, de la Cruz CC, Otte AP, Panning B, Zhang Y. Role of histone H3 lysine 27 methylation in X inactivation. *Science* 2003; 300: 131–5.
- Lee JT. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. *Genes Dev* 2009; 23: 1831–42.
- Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 2010; 329: 689–93.
- Conley AB, Jordan IK. Epigenetic regulation of human cis-natural antisense transcripts. *Nucleic Acids Res* 2012; 40: 1438–45.
- Chiba M, Miura T, Kasai K, Monzen S, Kashiwakura I, Yasue H, Nakamura T. Identification of up-regulated and down-regulated cis-natural antisense transcripts in the human B lymphoblastic cell line IM-9 after X-ray irradiation. *Mol Med Rep* 2012; 5: 1151–7.
- Morris KV, Santoso S, Turner AM, Pastori C, Hawkins PG. Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. *PLoS Genet* 2008; 4: e1000258.
- 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: 202–6.
- Morris KV, Vogt PK. Long antisense non-coding RNAs and their role in transcription and oncogenesis. *Cell Cycle* 2010; 9: 2544–7.
- Hawkins PG, Morris KV. Transcriptional regulation of Oct4 by a long non-coding RNA antisense to Oct4-pseudogene 5. *Transcription* 2010; 1: 165–75.
- Johnsson P, Ackley A, Vidarsdottir L, Lui WO, Corcoran M, Grandér D, Morris KV. A pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human cells. *Nat Struct Mol Biol* 2013; 20: 440–6.
- Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 2010; 465: 1033–8.
- Nolasco S, Bellido J, Gonçalves J, Tavares A, Zabala JC, Soares H. The expression of tubulin cofactor A (TBCA) is regulated by a noncoding antisense TbcA RNA during testis maturation. *PLoS One* 2012; 7: e42536.
- Li Q, Su Z, Xu X, Liu G, Song X, Wang R, Sui X, Liu T, Chang X, Huang D. AS1DHRS4, a head-to-head natural antisense transcript, silences the DHRS4 gene cluster in cis and trans. *Proc Natl Acad Sci USA* 2012; 109: 14110–5.
- Song X, Wang X, Arai S, Kurokawa R. Promoter-associated noncoding RNA from the CCND1 promoter. *Methods Mol Biol* 2012; 809: 609–22.
- Han J, Kim D, Morris KV. Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. *Proc Natl Acad Sci USA* 2007; 104: 12422–7.
- Wang X, Arai S, Song X, Reichart D, Du K, Pascual G, Tempst P, Rosenfeld MG, Glass CK, Kurokawa R. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* 2008; 454: 126–30.

27. Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* 2007; 445: 666–70.
28. Feng J, Bi C, Clark BS, Mady R, Shah P, Kohtz JD. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev* 2006; 20: 1470–84.
29. Beltran M, Puig I, Peña C, García JM, Alvarez AB, Peña R, Bonilla F, de Herrerros AG. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev* 2008; 22: 756–69.
30. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya PA, Blencowe BJ, Prasanth SG, Prasanth KV. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell* 2010; 39: 925–38.
31. Wang J, Liu X, Wu H, Ni P, Gu Z, Qiao Y, Chen N, Sun F, Fan Q. CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic Acids Res* 2010; 38: 5366–83.
32. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB, van de Vijver MJ, Sukumar S, Chang HY. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 2010; 464: 1071–6.
33. Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, Tanaka F, Shibata K, Suzuki A, Komune S, Miyano S, Mori M. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res* 2011; 71: 6320–6.
34. Kim K, Jutooru I, Chadalapaka G, Johnson G, Frank J, Burghardt R, Kim S, Safe S. HOTAIR is a negative prognostic factor and exhibits pro-oncogenic activity in pancreatic cancer. *Oncogene* 2013; 32: 1616–25.
35. Lin R, Maeda S, Liu C, Karin M, Edgington TS. A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. *Oncogene* 2007; 26: 851–8.
36. Lai MC, Yang Z, Zhou L, Zhu QQ, Xie HY, Zhang F, Wu LM, Chen LM, Zheng SS. Long non-coding RNA MALAT-1 overexpression predicts tumor recurrence of hepatocellular carcinoma after liver transplantation. *Med Oncol* 2012; 29: 1810–6.
37. Tinzl M, Marberger M, Horvath S, Chypre C. DD3PCA3 RNA analysis in urine – a new perspective for detecting prostate cancer. *Eur Urol* 2004; 46: 182–6; discussion 187.
38. Panzitt K, Tschernatsch MM, Guelly C, Moustafa T, Stradner M, Strohmaier HM, Buck CR, Denk H, Schroeder R, Trauner M, Zatloukal K. Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. *Gastroenterology* 2007; 132: 330–42.