Long non-coding RNA in liver diseases

Kenji Takahashi, **Irene Yan**, **Hiroaki Haga**, and **Tushar Patel**

Mayo Clinic, Jacksonville, Florida

Kenji Takahashi: Takahashi.kenji@mayo.edu; Irene Yan: Yan.irene@mayo.edu; Hiroaki Haga: Haga.hiroaki@mayo.edu; Tushar Patel: patel.tushar@mayo.edu

Abstract

The identification of the presence of large RNA transcripts that do not code proteins but that may have biological functions has provided an important new perspective in gene regulation. These long non-coding RNAs (lncRNAs) are being increasingly recognized to contribute to many biological processes through diverse mechanisms. The roles of these emerging genes are being recognized across kingdoms. These findings are profoundly altering our understanding of disease pathobiology and leading to the emergence of new biological concepts underlying human diseases. Strategies for the discovery and characterization of lncRNAs are highlighted. Several lncRNAs have been described in liver disease, and in liver cancers in particular. Their molecular mechanisms of action, function and contributions to disease pathophysiology are reviewed. LncRNA genes associated with liver diseases have potential roles as biomarkers of disease diagnosis, prognosis, or therapeutic response as well as direct targets for therapeutic intervention. Conclusion: The emerging knowledge in this rapidly advancing field offers promise for new fundamental knowledge and clinical applications that will be relevant for human liver diseases.

Keywords

Gene expression; Hepatocellular cancer; MicroRNA; RNA genes; Genome

Introduction

Although most studies to understand the role of genes have focused on protein coding genes, < 2% of the human genome encodes for protein coding genes. The Encyclopedia of DNA Elements (ENCODE) project revealed that while less than 2% of the genome encodes for protein coding genes, more than 80% of the genome is proposed to have biological activity, and a portion of this is actively transcribed but does not encode for protein-coding genes (Figure 1) (1). These include the long non-coding RNA (lncRNA). Defining their functions and their mechanistic basis has been the focus of recent intense research, and many lncRNA may organize multicomponent complexes that regulate gene expression. lncRNA are

becoming increasingly appreciated in pathogenesis of liver diseases and have potential diagnostic, prognostic or therapeutic importance.

What are long non-coding RNA?

Non-coding (nc) RNA transcripts are divided into two heterogeneous groups based on an arbitrary transcript size of greater or less than 200 nucleotides (Figure 2). The short ncRNAs include micro-RNAs (miRNAs), small nuclear RNAs, and many other classes of small RNAs. miRNA are 21-nucleotide ncRNA that post-transcriptionally regulate gene expression through the RNA-induced silencing complex, and use of Argonaute family proteins to cleave or inhibit translation of target gene mRNA transcripts. In some estimates, upto 60% of protein-coding genes can be targeted by miRNA (2). Their involvement in liver diseases have been recently reviewed (3).

Long ncRNAs are greater than 200 nucleotides. We use the term lncRNA to refer to transcribed RNA genes that are polyadenylated and are transcribed by RNA polymerase II (RNA pol II). They are less well understood than miRNA and other small RNAs. Similar to protein-coding genes, lncRNA show tissue-specific expression, chromatin marks, independent gene promoters, regulation by transcription factors and splicing of multiple exons into a mature transcript. Indeed, the use of chromatin marks that are observed at sites of initiation or active transcription of protein coding genes have been used to identify candidate lncRNA. Some ncRNAs are transcribed off a nearby or associated gene promoter (e.g. promoter/termini associated RNAs, intronic ncRNAs), whereas others may be transcribed from independent promoters (4). The genomic locations of lncRNA can be described in terms of their relationship to protein-coding genes (Figure 3). Intergenic lncRNAs (lincRNAs) have distinct transcriptional units that do not overlap those of protein-coding genes. Other types of lncRNA include enhancer RNAs, transcribed ultraconserved genes, antisense RNA and many others (Figure 2). Circular RNAs (circRNAs) are a distinct group best considered separately from other linear ncRNA, although most identified circRNAs are greater than 200 nucleotides (5). The diversity of genomic locations and types of ncRNA highlights the high complexity of the genetic information encoded within the human genome.

How are long non-coding RNA identified?

Large-scale genome wide sequencing and next generation sequencing identified several lncRNA transcripts, first within the mouse transcriptome, and subsequently also in humans (6). More recently, tiling microarrays and RNA-sequencing have increased the numbers of known lncRNA (7). Quantitatively, the number of postulated lncRNA transcripts may even exceed the number of protein coding mRNAs in mammals. Similar to mice, the identification of 9,640 lncRNA loci in human was reported by the ENCODE project. This effort involved a large number of sequence-based studies to map functional elements across the human genome. The elements mapped include RNA transcribed regions, protein-coding regions, transcriptional factor binding sites, chromatin structure, and DNA methylation sites. ENCODE project data suggest that more than 80% of the genome had biochemical functions, with a small proportion of this representing lncRNA transcripts (1).

Epigenetic marks associated with transcription can identify lncRNAs within the genome. Genes actively transcribed by Pol II are enriched with trimethylation of lysine 4 on histone H3 and trimethylation of lysine 36 on histone H3 within their coding regions. A search for these domains in genome-wide chromatin-state maps in mice identified ~1600 lincRNAs with these epigenetic marks and that were conserved across mammals. These lincRNA undergo 5′-capping and poly-adenylation and are transcribed by Pol II. Transcriptional regulation was demonstrated for selected lincRNAs associated with biological processes such as embryonic stem cell pluripotency to cell proliferation (7). A similar approach identified about 300 lincRNAs in human cells.

Some lincRNAs have been functionally validated and their association with biological processes has been shown. About 20% of lincRNAs are bound to polycomb repressive complex 2 (PRC2). Through this interaction, lincRNA are associated with chromatin remodeling and altered gene expression and with biological functions (7, 8). LincRNAs may be involved in different kinds of biological pathways such as pluripotency, differentiation, proliferation or cell survival. For example, lincRNA-Regulator of Reprogramming (lncRNA-RoR) plays a role in promoting survival in liver cancer cells during hypoxic stress by modulating hypoxia signaling pathways or in pluripotent or embryonic stem cells by preventing the activation of cellular stress pathways such as the p53 response (9, 10).

A search for conserved genomic regions has identified novel lncRNAs. Ultraconserved elements are evolutionary conserved sequences that are 100% identical across the genomes of humans, mice, and rats. A genome wide bioinformatics study identified 481 segments, ranging in size from 200 to approximately 800 base pairs between orthologous regions (11). The expression of some transcribed ultraconserved regions, or ultraconserved RNA (ucRNA), are significantly altered in diseases, such as adult chronic lymphocytic leukemia, colorectal cancer and neuroblastoma (12). We identified and cloned transcribed ultraconserved RNA 338 (TUC338), a ucRNA that is increased in both cirrhosis and HCC (13).

Although the ability to detect lncRNA within the human genome has been facilitated by genomic sequencing and bioinformatics analyses, validation of putative candidate genes is necessary. This can be done using conventional molecular biological approaches such as cloning, and expression analysis using northern blots or quantitative PCR. However, a major challenge remains to identify and validate functional biological roles of these lncRNA candidates. This is difficult in part because of the lack of proteins encoded by these genes, the diverse molecular functions and the wide diversity of potential mechanisms by which these RNA genes can have effects on gene expression. Consequently, the function of most of the lncRNA that have been implicated in liver and other diseases remain poorly described. Understanding these functions will be critical to recognizing the contribution of these genes on biological processes involved in hepatic functioning.

Identification of mRNA targets or specific proteins that bind to a candidate lncRNA may provide clues to its function (14). Potential mRNA targets could be identified by computational searches for mRNA that may be recognized by base-pairing complementarity, by analysis of RNA following cross-linking to a lncRNA or its binding protein or by

transcriptomic analysis after modulation of lncRNA expression. Potential binding proteins could be identified by isolation of proteins bound to tagged RNA or RNA binding proteins, or by using three-hybrid systems using lncRNA as a bait. The subcellular location or patterns of expression within specific cell types, stages of development or disease states can be assessed using in situ hybridization (Figure 4) or Northern blotting. Functional studies can also be performed by knock-down of lncRNA in model organisms such as mice using similar strategies to those used for protein coding mRNAs (15).

Are long non-coding RNA involved in human liver diseases?

Many lncRNA are known to contribute to liver regeneration, neoplasia and other liver diseases. H19 and XIST were the first lncRNAs to be identified and studied in the liver. Although their discoveries predated that of the microRNAs, the latter have been extensively studied. Others include two novel lncRNA involved in liver cancers that we have cloned (13, 16), maternally expressed 3 (Meg3) which we have shown to be a tumor suppressor gene in hepatocellular cancer (HCC) (17), highly up-regulated in liver cancer (HULC) (18), and metastasis-associated lung adenocarcinoma transcript 1 (MALAT-1) which functions in regulation of alternate splicing and has an oncogenic role in HCC development (19, 20). The new insights provided by these lncRNA into disease pathogenesis or behavior merely represent the tip of the iceberg given the large numbers of lncRNA that remain to be characterized.

The identification of disease-associated lncRNAs has significant potential for new clinically relevant biomarkers. Thus focused efforts to characterize disease relevant lncRNA are very likely to identify new RNA genes that will be relevant to liver diseases or biology. Several recent studies have reported expression of selected lncRNA within the liver that may contribute to liver pathobiology or are associated with specific liver diseases, and in particular with liver cancers. Similar to expression of other mRNA, expression of lncRNA can be detected using PCR or by *in situ* hybridization in tissue sections.

lncRNA associated with liver cancers—Several lncRNA have been identified in HCC (Table 1). Several of these lncRNA have been linked to clinical prognostic roles raising hope that lncRNA may provide effective biomarkers that may be helpful in practice.

Viral-hepatitis related lncRNA—lncRNA expression profiling between HBV-HCC and normal liver tissues identified a greater than two-fold change in 4.3% of all lncRNA analyzed (21). lncRNA down-regulated expression by HBx (lncRNA-Dreh) was altered in hepatitis B virus X protein (HBx) transgenic mice compared with wild type mice (22). This murine lncRNA can inhibit HCC growth and metastasis *in vitro* and *in vivo*, and acts through cytoskeletal modulation by repressed expression of vimentin. The human ortholog RNA of Dreh (hDREH) is down-regulated in human HBV-related HCC compared with adjacent noncancerous tissues, with the decrement correlating with poor survival of HCC patients.

lncRNA in genetic liver diseases—An interesting study linked a lincRNA to a mendelian disorder with autosomal-recessive inheritance contributing to liver disease. The

HELLP syndrome is characterized by hemolysis, elevated liver enzymes, and low platelets occurs during pregnancy, This syndrome was linked to the HELLP lncRNA and is functionally implicated in the cell cycle (23). Blocking potential mutation sites identified in HELLP families decreased the invasion capacity of extravillous trophoblasts.

lncRNA involved in hepatic pathophysiology—Differential expression of several lncRNA was noted during liver regeneration after partial hepatectomy in mice (24). Amongst these was long noncoding RNA associated with liver regeneration 1 (lncRNA-LALR1). lncRNA-LALR1 enhances cell cycle progression and hepatocyte proliferation. This lncRNA facilitates cyclin D1 expression through activation of Wnt/β-catenin signaling by a mechanism involving suppression of Axin1 by recruiting CTCF to the AXIN1 promoter region (24). A human ortholog RNA of lncRNA-LALR1 (lncRNA-hLALR1) has been shown to be expressed in human liver tissues.

A novel biological function as a signaling mediator has been reported for TUC339. Although expression of this ucRNA is not increased in malignant hepatocytes compared with normal cells, it is highly enriched within extracellular vesicles released from HCC cells (16). The high extracellular vesicle enrichment of TUC339 within tumor cells, and its detection in serum from patients with HCC raise the potential for its use as a circulating biomarker of disease.

What are the molecular functions of long non-coding RNA, and how do they modulate gene expression?

Amongst the lncRNA that have been functionally characterized, a wide and diverse range of effects on the regulation of gene expression have been recognized. These include effects on gene expression through modulation of chromatin remodeling, control of gene transcription, post-transcriptional mRNA processing, protein function or localization and intercellular signaling (Figure 5). This diversity precludes a function-based classification that would encompass all lncRNA. Mechanisms that have been described for selected lncRNA involved in liver diseases include widely diverse functions such as DNA imprinting, X-inactivation, DNA demethylation, gene transcription and generation of other RNA molecules.

Chromatin remodeling—Within the liver, circadian oscillations of expression of several lncRNA within the liver may contribute to cycling histone modifications and gene transcription that can contribute to homeostasis of metabolic systems (25). Many lncRNA can regulate gene expression through targeted effects on chromatin remodeling, or through X-chromosome inactivation and imprinting, as exemplified by the lncRNAs Xist and H19 respectively. Loss of imprinting at the H19 locus has been implicated in the increased H19 expression noted in HCC. The MEG3 lncRNA is a putative tumor suppressor in HCC (17). MEG3 directly and specifically binds to PRC2 in mouse embryonic stem cells. MEG3/Gtl2 is a single-copy gene and these isoforms are transcribed from the same gene but generated by alternative splicing, using different exons (26). Gtl2 and PRC2 interactions regulate Dlk1 by targeting PRC2 to Dlk1 in cis (27). Moreover, MEG3 can significantly increase p53 protein levels and stimulate p53-dependent transcription from a p53-responsive promoter. All MEG3 isoforms contain three different secondary folding motifs, M1, M2 and M3.

Because M2 and M3 play a critical role in activation of p53, the RNA structural motifs are important to MEG3 function (26, 28) emphasizing the contribution of secondary and tertiary RNA structure in lncRNA function.

Reprogramming of gene expression during cancer development by the lncRNA HOTAIR involves chromatin remodeling. HOTAIR can act as a molecular scaffold capable of binding to at least two distinct histone modification complexes. A 5′ domain of this lncRNA binds to the mammalian PRC2 complex responsible for H3K27 methylation, whereas a 3′ domain of the RNA binds to the histone lysine demethylase LSD1 (29, 30).

Transcriptional control—A role for lncRNA in direct control of transcription is emerging. The HOTTIP gene is located in contiguity with HOXA13 and controls HOXA locus gene expression via interaction with the WDR5/MLL complex (31). Studies have shown a bidirectional regulatory loop between HOTTIP/HOXA13. HEIH associates with EZH2, an important subunit of the PRC2 complex, and this association contributes to the expression of EZH2-regulated target genes. Chromatin isolation by RNA purification revealed that TUC338 interacts with genomic DNA and plasminogen activator inhibitor-1 (PAI-1) mRNA binding protein (PAI-RBP1), an RNA binding protein that is associated with (PAI-1). TUC338 increases PAI-RBP1 and PAI-1 expression, and PAI-RBP1 promotes transformed cell growth. Furthermore, TUC338 may prevent binding of Pax6 and P53 through binding to their DNA motifs. Thus, TUC338 might mediate gene expression through interaction with trans-acting proteins and binding to cis-acting elements for abnormal cell proliferation in HCC. Other lncRNAs such as Evf-2 have been shown to activate transcriptional activity by directly influencing Dlx-2 activity.

Post-transcriptional processing—Several lncRNA such as MALAT-1 can regulate gene expression through post-transcriptional effects. The full length of 7.5 kb MALAT1 transcript is processed by RNaseP and RNaseZ to generate two non-coding RNAs. One of these is mascRNA, a small ncRNA and is exported to the cytoplasm where the other is the lncRNA MALAT-1 and is retained in the nuclear speckles (32). MALAT-1 may regulate alternate splicing by modulating the activity of serine/arginine (SR) splicing factors that regulate alternative splicing (20). MALAT-1 depletion increases the levels of cellular SR proteins and alters the distribution and the ratio of phosphorylated to dephosphorylated pools of SR proteins. HULC co-purifies with ribosomes in the cytoplasm, and has been implicated in the regulation of tumor cell proliferation through down-regulation of p18 and involving ATM/ATR and p53 dependent signaling (33). HULC also down-regulates several miRNA, including miR-372. Reduction of miR-372 decreases translational repression of PRKACB, which in turn induces phosphorylation of cAMP response element binding protein (CREB). Phosphorylated CREB forms part of the RNA pol II transcriptional machinery to activate HULC expression. Thus, HULC may act as an endogenous 'miRNA sponge' to regulate miRNAs activities in HCC cells (34). lincRNA-RoR can associate with epigenetic but also negatively regulates p53 expression and inhibits p53 mediated cell cycle arrest and apoptosis through direct interaction with the heterogeneous nuclear ribonucleoprotein I. The effects of lincRNA-RoR on pluripotency or cell survival may therefore be mediated in part through translation regulation.

Protein function and localization—Some lncRNA have also been shown to be capable of modulating protein localization, protein stability and protein activity. NRON is a noncoding repressor of the nuclear factor of activated T cells (NFAT). This lncRNA interacts with proteins such as members of the importin-beta superfamily and regulates NFAT localization as an RNA component of protein complex that represses NFAT activity (35). Cyclin D1 (CCND1) ncRNAs are transcribed from the promoter region of CCND1 gene. CCND1 ncRNA interacts with RNA-binding protein, translocated in liposarcoma (TLS), and prevents CRBE-binding protein and histone acetyltransferase activities to repress CCND1. This lncRNA negatively regulates CCND1 transcription by recruiting TLS to the CCND1 promoter. Thus, CCND1 ncRNA might be involved in protein activity (36).

These examples illustrate the diversity of mechanisms by which lncRNA can exert functional effects on gene expression (Figure 6).

What is the diagnostic and therapeutic potential of long non-coding RNAs?

lncRNA can be useful for diagnosis or as prognostic markers.

Risk prediction—lncRNA SNPs may identify disease risk. A case-control study examined the association between SNPs rs7763881 in HULC and rs619586 in MALAT1, and susceptibility to HCC or chronic HBV infection was examined in a case-control study of 1300 HBV positive HCC patients, 1344 HBV persistent carriers and 1344 subjects with HBV natural clearance (37). rs7763881 was associated with decreased HCC risk whereas the association with rs619586 was of borderline significance. Thus, the variant genotypes of rs7763881 in HULC may contribute to decreased susceptibility to HCC in HBV carriers. Variants in genes that could affect lncRNA expression or function may also be useful. The deletion allele of a polymorphism (rs10680577) within the distal promoter of Prolyl hydroxylase 1 (EGLN2) was significantly associated with increased risk for HCC in two case-control cohorts, particularly in smokers (38). The deletion allele correlated with expression of both EGLN2 and RERT-lncRNA *in vivo* and *in vitro*, potentially through actions on the RERT-lncRNA structure, and could be a promising biomarker for early diagnosis of HCC.

Diagnostic markers—Detection of circulating lncRNA raises the potential for use as markers of disease. For example, circulating HULC is detected in patients with HCC (18), while HULC expression is increased in liver tumor tissues and observed in the plasma of patients with HBV+ status (39). Other lncRNA that are deregulated in liver diseases such as TUC338 (increased) or MEG3 (decreased) in HCC may also have potential utility as diagnostic markers and warrant further exploration (13, 17).

Predictors of prognosis—Several lncRNA have prognostic potential. lncRNA-HEIH is associated with cancer recurrence in HBV-related HCC, with high lncRNA-HEIH expression predicting a worse prognosis (40). Similarly, plasma HULC is increased with higher HCC Edmonson grades (39). Increased HOTAIR expression is associated with risk of recurrence after hepatectomy for HCC (41, 42) whereas overexpression of MVIH is associated with microvascular invasion and decreased survival after hepatectomy (43).

MALAT-1 expression predicts recurrence in patients with HCC undergoing liver transplantation, particularly in patients beyond Milan criteria (19). As data regarding the role of lncRNA emerges in other liver diseases, the potential for lncRNA based prognostic markers will expand

Therapeutic potential—New therapeutic possibilities are arising from targeting RNA for the treatment of disease (44). Therapeutic approaches based on H19 have entered clinical trials. An example is gene therapy using a vector carrying the diphtheria toxin A-chain gene under the control of H19 regulatory sequences that may potentially be useful for liver tumors with a high expression of H19 (45). Likewise, targeting lncRNA-LALR1 may be therapeutically beneficial in inducing liver regeneration following liver injury because this lncRNA has been shown to accelerate liver cell proliferation (24). Similar to emerging strategies to target microRNA, targeting lncRNA holds considerable promise.

Perspectives

Extensive investigations into the content of the human genome have identified several thousands of lncRNA. Emerging studies are implicating these in a wide range of biological processes involved in development, differentiation, growth tissue injury, repair, regeneration and metabolism. New candidate disease-associated lncRNA genes are being identified and their molecular mechanisms are being elucidated. Epigenetic regulation of the genome by lncRNA is an exciting nascent emerging paradigm that will provide new biological insights. lncRNA offer new possibilities for understanding disease pathogenesis, as diagnostic or prognostic biomarkers, or as direct therapeutic targets. Indeed, several lncRNA associated with liver cancers may have potential as markers of disease progression or prognosis. The involvement of lncRNA in diseases other than cancer remains less well characterized but is likely to be a fruitful area of investigation. Given the large number of lncRNA, and the intense research to identify and evaluate these genes, novel lncRNAs with relevance to liver diseases are likely to be identified and their effects on disease pathophysiology defined. An improved understanding resulting from focused study of these novel genes in liver diseases will undoubtedly provide useful insights and generate new hypothesis regarding disease pathogenesis that will eventually lead to novel clinical applications.

Acknowledgments

Abbreviations

lncRNAs	long non-coding RNAs
ENCODE	Encyclopedia of DNA Elements
ncRNAs	non-coding RNAs
miRNAs	micro-RNAs

RNA pol II	RNA polymerase II
lincRNAs	long intergenic non coding RNAs
circRNAs	circular RNAs
PRC2	polycomb repressive complex 2
MEG3	Maternally expressed 3
HCC	hepatocellular cancer
HULC	highly up-regulated in liver cancer
MALAT-1	metastasis-associated lung adenocarcinoma transcript 1
lncRNA-RoR	lincRNA-Regulator of Reprogramming
ucRNA	ultraconserved RNA
TUC338	Transcribed ultraconserved RNA 338
IGF2	Insulin-Like Growth Factor 2
HBV	hepatitis B virus
MDR1	multi-drug resistance 1
HOTAIR	Hox transcript antisense intergenic RNA
HOTTIP	HoxA distal transcript antisense RNA
lncRNA-LALR1	Long noncoding RNA associated with liver regeneration 1
lncRNA MVI	Long noncoding RNA associated with microvascular invasion in HCC
CUDR	Cancer up-regulated drug resistant
lncRNA-HEIH	Highly expressed in hepatocellular carcinoma long non-coding RNA
TUC339	Transcribed ultraconserved RNA 339
HELLP	hemolysis, elevated liver enzymes, low platelets
lncRNA-Dreh lncRNA	down-regulated expression by HBx
PAI-1	plasminogen activator inhibitor-1
PAI-RBP1	plasminogen activator inhibitor-1 mRNA binding protein
CCND1	Cyclin D1
NFAT	nuclear factor of activated T cells
EGLN2	Prolyl hydroxylase 1

References

1. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, Epstein CB, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. [PubMed: 22955616]

2. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. 2009; 19:92–105. [PubMed: 18955434]

3. Takahashi K, Yan I, Wen HJ, Patel T. microRNAs in liver disease: From diagnostics to therapeutics. Clin Biochem. 2013

4. Venters BJ, Pugh BF. Genomic organization of human transcription initiation complexes. Nature. 2013; 502:53–58. [PubMed: 24048476]

5. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature. 2013; 495:333–338. [PubMed: 23446348]

6. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, et al. The transcriptional landscape of the mammalian genome. Science. 2005; 309:1559–1563. [PubMed: 16141072]

7. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature. 2009; 458:223–227. [PubMed: 19182780]

8. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A. 2009; 106:11667–11672. [PubMed: 19571010]

9. Loewer S, Cabili MN, Guttman M, Loh YH, Thomas K, Park IH, Garber M, et al. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. Nat Genet. 2010; 42:1113–1117. [PubMed: 21057500]

10. Takahashi K, Yan I, Haga H, Patel T. Modulation of hypoxia-signaling pathways by extracellular long non-coding RNA regulator of reprogramming. J Cell Sci. 2014 In press.

11. Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, Haussler D. Ultraconserved elements in the human genome. Science. 2004; 304:1321–1325. [PubMed: 15131266]

12. Calin GA, Liu CG, Ferracin M, Hyslop T, Spizzo R, Sevignani C, Fabbri M, et al. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. Cancer Cell. 2007; 12:215–229. [PubMed: 17785203]

13. Braconi C, Valeri N, Kogure T, Gasparini P, Huang N, Nuovo GJ, Terracciano L, et al. Expression and functional role of a transcribed noncoding RNA with an ultraconserved element in hepatocellular carcinoma. Proc Natl Acad Sci U S A. 2011; 108:786–791. [PubMed: 21187392]

14. Huttenhofer A, Vogel J. Experimental approaches to identify non-coding RNAs. Nucleic Acids Res. 2006; 34:635–646. [PubMed: 16436800]

15. Eissmann M, Gutschner T, Hammerle M, Gunther S, Caudron-Herger M, Gross M, Schirmacher P, et al. Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. RNA Biol. 2012; 9:1076–1087. [PubMed: 22858678]

16. Kogure T, Yan I, Lin W-L, Patel T. Extracellular vesicle mediated transfer of a novel long non-coding RNA TUC339: A mechanism of intercellular signaling in human hepatocellular cancer. Genes and Cancer. 2013; 4:261–272. [PubMed: 24167654]

17. Braconi C, Kogure T, Valeri N, Huang N, Nuovo G, Costinean S, Negrini M, et al. microRNA-29 can regulate expression of the long non-coding RNA gene MEG3 in hepatocellular cancer. Oncogene. 2011; 30:4750–4756. [PubMed: 21625215]

18. Panzitt K, Tschernatsch MM, Guelly C, Moustafa T, Stradner M, Strohmaier HM, Buck CR, et al. Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. Gastroenterology. 2007; 132:330–342. [PubMed: 17241883]

19. Lai MC, Yang Z, Zhou L, Zhu QQ, Xie HY, Zhang F, Wu LM, et al. Long non-coding RNA MALAT-1 overexpression predicts tumor recurrence of hepatocellular carcinoma after liver transplantation. Med Oncol. 2012; 29:1810–1816. [PubMed: 21678027]

20. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell. 2010; 39:925–938. [PubMed: 20797886]

21. Pan YF, Qin T, Feng L, Yu ZJ. Expression profile of altered long non-coding RNAs in patients with HBV-associated hepatocellular carcinoma. J Huazhong Univ Sci Technolog Med Sci. 2013; 33:96–101. [PubMed: 23392715]

22. Huang JF, Guo YJ, Zhao CX, Yuan SX, Wang Y, Tang GN, Zhou WP, et al. 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. 2013; 57:1882–1892. [PubMed: 23239537]

23. van Dijk M, Thulluru HK, Mulders J, Michel OJ, Poutsma A, Windhorst S, Kleiverda G, et al. HELLP babies link a novel lincRNA to the trophoblast cell cycle. J Clin Invest. 2012; 122:4003–4011. [PubMed: 23093777]

24. Xu D, Yang F, Yuan JH, Zhang L, Bi HS, Zhou CC, Liu F, et al. Long noncoding RNAs associated with liver regeneration 1 accelerates hepatocyte proliferation during liver regeneration by activating Wnt/beta-catenin signaling. Hepatology. 2013; 58:739–751. [PubMed: 23483581]

25. Vollmers C, Schmitz RJ, Nathanson J, Yeo G, Ecker JR, Panda S. Circadian oscillations of protein-coding and regulatory RNAs in a highly dynamic mammalian liver epigenome. Cell Metab. 2012; 16:833–845. [PubMed: 23217262]

26. Zhang X, Rice K, Wang Y, Chen W, Zhong Y, Nakayama Y, Zhou Y, et al. Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: isoform structure, expression, and functions. Endocrinology. 2010; 151:939–947. [PubMed: 20032057]

27. Zhao J, Ohsumi TK, Kung JT, Ogawa Y, Grau DJ, Sarma K, Song JJ, et al. Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol Cell. 2010; 40:939–953. [PubMed: 21172659]

28. Zhou Y, Zhong Y, Wang Y, Zhang X, Batista DL, Gejman R, Ansell PJ, et al. Activation of p53 by MEG3 non-coding RNA. J Biol Chem. 2007; 282:24731–24742. [PubMed: 17569660]

29. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, et al. Long noncoding RNA as modular scaffold of histone modification complexes. Science. 2010; 329:689–693. [PubMed: 20616235]

30. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature. 2010; 464:1071–1076. [PubMed: 20393566]

31. Quagliata L, Matter MS, Piscuoglio S, Arabi L, Ruiz C, Procino A, Kovac M, et al. lncRNA HOTTIP/HOXA13 expression is associated with disease progression and predicts outcome in hepatocellular carcinoma patients. Hepatology. 2013

32. Wilusz JE, Freier SM, Spector DL. 3′ end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. Cell. 2008; 135:919–932. [PubMed: 19041754]

33. Du Y, Kong G, You X, Zhang S, Zhang T, Gao Y, Ye L, et al. Elevation of highly up-regulated in liver cancer (HULC) by hepatitis B virus X protein promotes hepatoma cell proliferation via down-regulating p18. J Biol Chem. 2012; 287:26302–26311. [PubMed: 22685290]

34. Wang J, Liu X, Wu H, Ni P, Gu Z, Qiao Y, Chen N, et al. CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. Nucleic Acids Res. 2010; 38:5366–5383. [PubMed: 20423907]

35. Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, Hogenesch JB, et al. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. Science. 2005; 309:1570–1573. [PubMed: 16141075]

36. Wang X, Arai S, Song X, Reichart D, Du K, Pascual G, Tempst P, et al. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature. 2008; 454:126–130. [PubMed: 18509338]

37. Liu Y, Pan S, Liu L, Zhai X, Liu J, Wen J, Zhang Y, et al. A genetic variant in long non-coding RNA HULC contributes to risk of HBV-related hepatocellular carcinoma in a Chinese population. PLoS One. 2012; 7:e35145. [PubMed: 22493738]

38. Zhu Z, Gao X, He Y, Zhao H, Yu Q, Jiang D, Zhang P, et al. An insertion/deletion polymorphism within RERT-lncRNA modulates hepatocellular carcinoma risk. Cancer Res. 2012; 72:6163–6172. [PubMed: 23026137]

39. Xie H, Ma H, Zhou D. Plasma HULC as a promising novel biomarker for the detection of hepatocellular carcinoma. Biomed Res Int. 2013; 2013:136106. [PubMed: 23762823]

40. Yang F, Zhang L, Huo XS, Yuan JH, Xu D, Yuan SX, Zhu N, et al. Long noncoding RNA high expression in hepatocellular carcinoma facilitates tumor growth through enhancer of zeste homolog 2 in humans. Hepatology. 2011; 54:1679–1689. [PubMed: 21769904]

41. Yang Z, Zhou L, Wu LM, Lai MC, Xie HY, Zhang F, Zheng SS. Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. Ann Surg Oncol. 2011; 18:1243–1250. [PubMed: 21327457]

42. Geng YJ, Xie SL, Li Q, Ma J, Wang GY. Large intervening non-coding RNA HOTAIR is associated with hepatocellular carcinoma progression. J Int Med Res. 2011; 39:2119–2128. [PubMed: 22289527]

43. Yuan SX, Yang F, Yang Y, Tao QF, Zhang J, Huang G, Wang RY, et al. 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. 2012; 56:2231–2241. [PubMed: 22706893]

44. Braconi C, Patel T. Non-coding RNAs as therapeutic targets in hepatocellular cancer. Curr Cancer Drug Targets. 2012; 12:1073–1080. [PubMed: 22873215]

45. Amit D, Hochberg A. Development of targeted therapy for bladder cancer mediated by a double promoter plasmid expressing diphtheria toxin under the control of H19 and IGF2-P4 regulatory sequences. J Transl Med. 2010; 8:134. [PubMed: 21162716]

46. Tsang WP, Wong TW, Cheung AH, Co CN, Kwok TT. Induction of drug resistance and transformation in human cancer cells by the noncoding RNA CUDR. RNA. 2007; 13:890–898. [PubMed: 17416635]

47. Anwar SL, Krech T, Hasemeier B, Schipper E, Schweitzer N, Vogel A, Kreipe H, et al. Loss of imprinting and allelic switching at the DLK1-MEG3 locus in human hepatocellular carcinoma. PLoS One. 2012; 7:e49462. [PubMed: 23145177]

48. Zhang L, Yang F, Yuan JH, Yuan SX, Zhou WP, Huo XS, Xu D, et al. Epigenetic activation of the MiR-200 family contributes to H19-mediated metastasis suppression in hepatocellular carcinoma. Carcinogenesis. 2013

49. Iizuka N, Oka M, Yamada-Okabe H, Mori N, Tamesa T, Okada T, Takemoto N, et al. Comparison of gene expression profiles between hepatitis B virus- and hepatitis C virus-infected hepatocellular carcinoma by oligonucleotide microarray data on the basis of a supervised learning method. Cancer Res. 2002; 62:3939–3944. [PubMed: 12124323]

50. Matouk IJ, DeGroot N, Mezan S, Ayesh S, Abu-lail R, Hochberg A, Galun E. The H19 non-coding RNA is essential for human tumor growth. PLoS One. 2007; 2:e845. [PubMed: 17786216]

51. Tsang WP, Kwok TT. Riboregulator H19 induction of MDR1-associated drug resistance in human hepatocellular carcinoma cells. Oncogene. 2007; 26:4877–4881. [PubMed: 17297456]

52. Ishibashi M, Kogo R, Shibata K, Sawada G, Takahashi Y, Kurashige J, Akiyoshi S, et al. Clinical significance of the expression of long non-coding RNA HOTAIR in primary hepatocellular carcinoma. Oncol Rep. 2013; 29:946–950. [PubMed: 23292722]

53. 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; 26:851–858. [PubMed: 16878148]
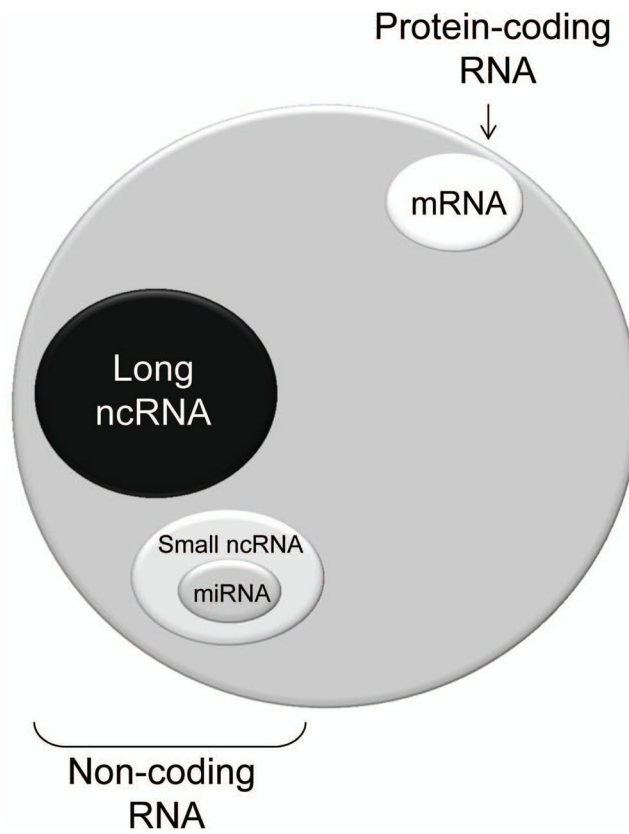
Figure 1. Schematic of mRNA and non-coding RNA encoded within the human genome
Protein coding RNA comprises of less than 2% of the genome. Biological function is ascribed to ~ 80% of the genome. The long non-coding RNA comprise of an unknown number of genes, but are likely to exceed the total number of mRNAs.
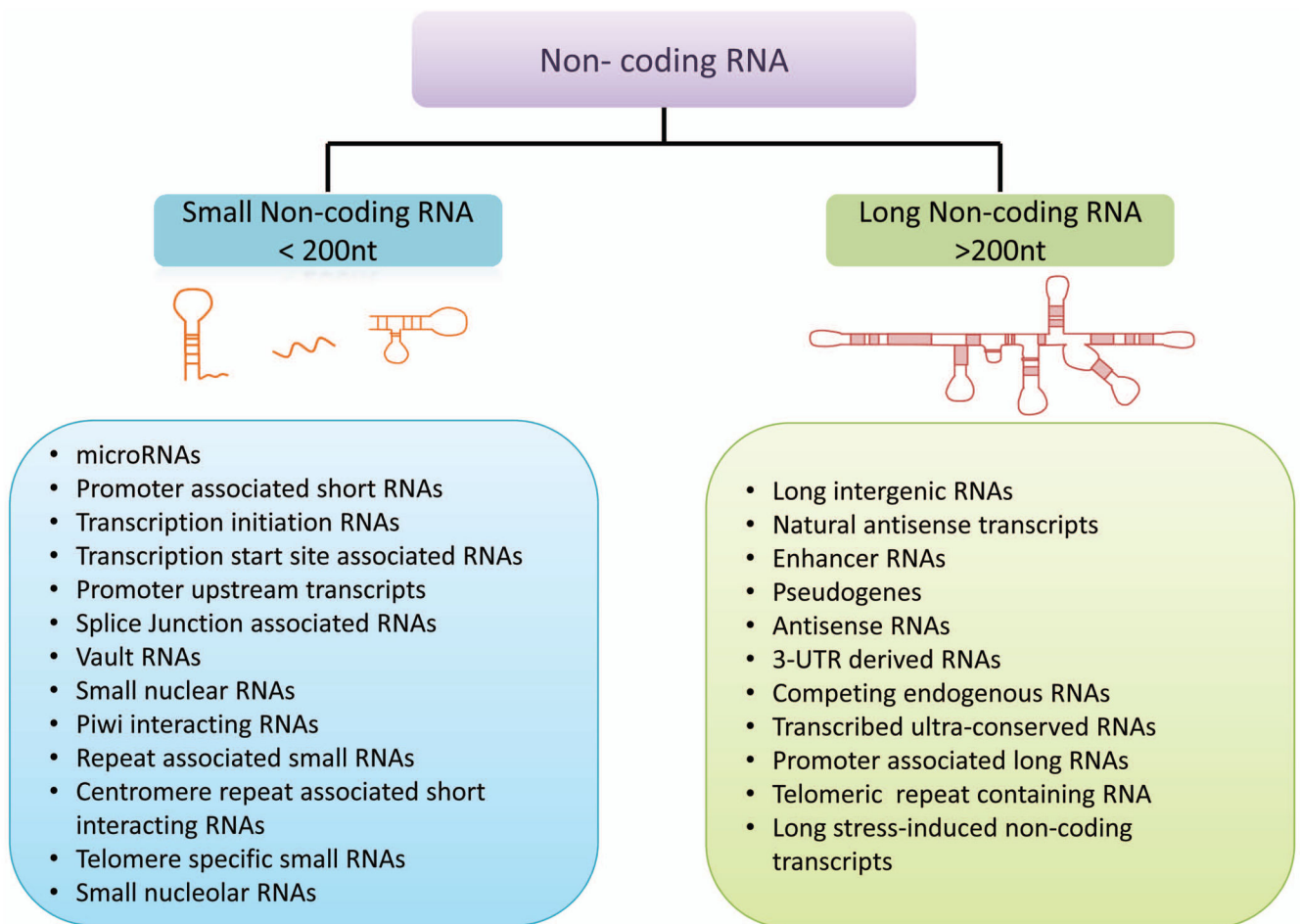
Figure 2. Classification of non-coding RNA (ncRNA)
ncRNAs are generally divided into two groups based on transcript size, small ncRNA (< 200 nucleotides) and long ncRNA (> 200 nucleotides). Examples of each group are provided to illustrate the heterogeneity of these groups.
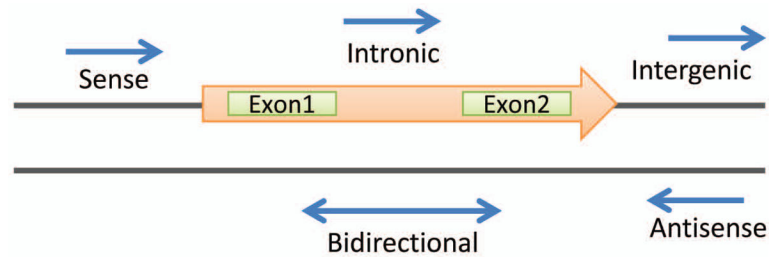
Figure 3. Genomic loci of lncRNA

The genomic locations for lncRNAs can be distinct from or can overlap regions that encode can encode for mRNA and proteins. LncRNAs are described in relationship to their location relative to that of nearby protein-coding genes. LncRNA can be transcribed from sense, antisense or both strands relative to the promoter of a protein-coding gene. Intronic lncRNAs initiate inside of an intron of a protein-coding gene in either direction and terminate without overlapping exons. Intergenic lncRNAs, termed lincRNAs, are lncRNAs with separate transcriptional units from protein-coding genes. Overlap may exist between categories, but is not shown. For example, some genes may be partly exonic and partly intronic. Orange arrow indicates a protein coding region. Blue arrows depict lncRNAs.
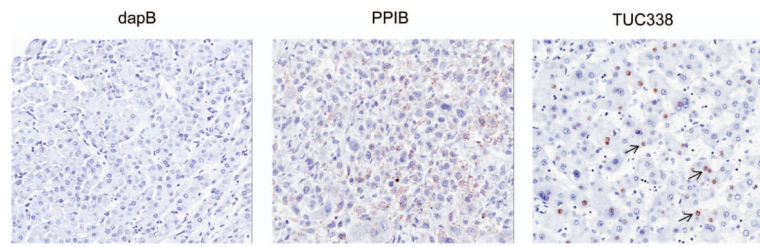
Figure 4. Expression of TUC338 in human HCC

In situ hybridization using RNAscope shows punctate expression of TUC338 in HCC tissues. Expression of dapB and PPIB are shown as negative and positive controls respectively. Original magnification, x 20.
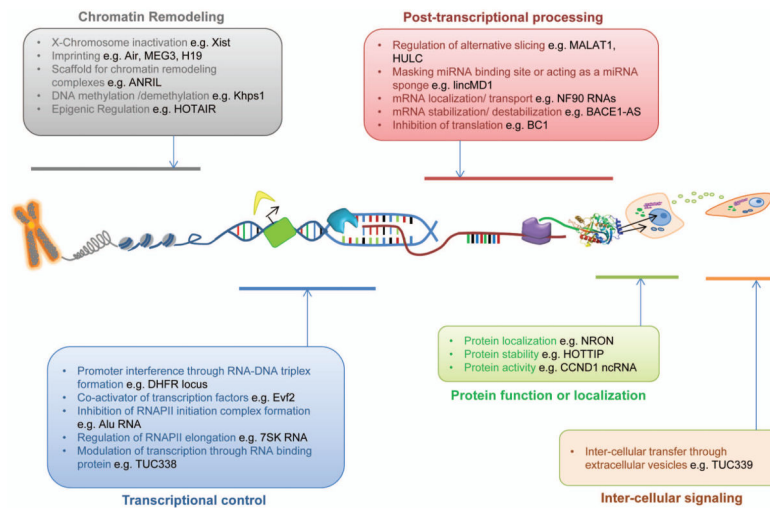
Figure 5. Functions of long non-coding RNAs

Overview and examples of mechanisms by which lncRNAs are involved in diverse gene regulations. lncRNAs can regulate gene expression through diverse mechanisms such as chromatin remodeling, transcriptional control, post-transcriptional processing, protein functioning and localization and through intercellular signaling.
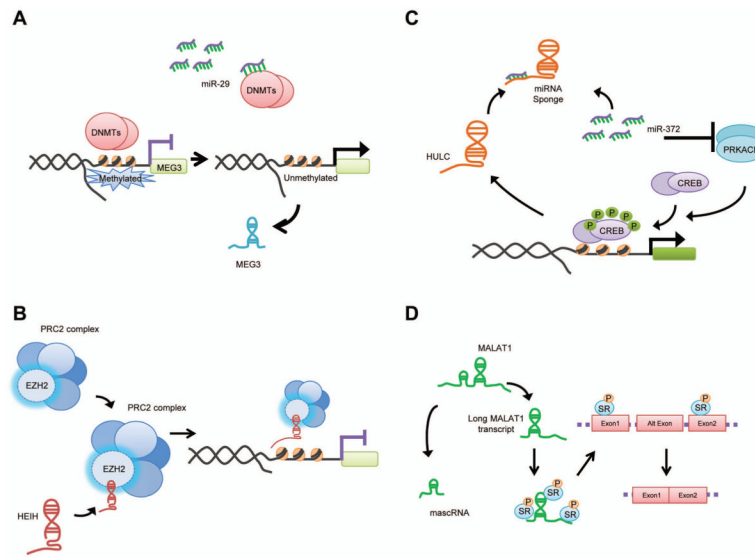
Figure 6. Mechanisms of action of selected lncRNA

lncRNA can be regulated or act through a variety of mechanisms that involve diverse processes such as transcriptional and epigenetic regulation of gene expression, regulation of splicing and interactions with miRNA. Mechanisms that have been described for MEG3, HEIH, HULC and MALAT-1 are depicted to illustrate the diversity of interactions between lncRNA and other proteins and RNA molecules. (A) miRNA dependent regulation of lncRNA expression. Methylation at the MEG3 promoter is regulated by DNMTs. Modulation of DNMTs by miR-29 contributes to regulation of MEG3 expression. (B) Epigenetic regulation. HEIH associates with the EZH2 subunit of the PRC2 complex, and recruits this complex to a specific gene locus and contributes to the expression of EZH2-regulated target genes. (C) Targeting miRNA. HULC acts as a molecular sponge to sequester miR-372 and reduce the repression of its target gene PRKACB thereby enabling phosphorylation and activation of the CREB transcription factor, which in turn stimulates HULC expression. (D) Regulation of alternative splicing. The full length of MALAT-1 transcript is processed to generate two non-coding RNAs. The larger MALAT-1 RNA transcript is retained in the nuclear speckles and can modulate the activity of serine/arginine (SR) splicing factors to potentially regulate alternative splicing.

Table 1

Examples of lncRNA associated with hepatocellular carcinoma

Chromosomal location and transcript size are shown for selected lncRNA involved in hepatocellular carcinoma. lncRNA have been identified to be transcribed from all chromosomes and can be of varying sizes.

lncRNA	Chromosomal location	Size (Kb)	Potential role in liver cancer	Refs
HULC	Chr6	0.5	Up-regulated in HCC. Higher expression associated with histological grade or HBV	(33, 34, 39)
TUC338	Chr12	0.59	Increased in cirrhosis and HCC. Modulates cell growth	(13)
UCA1/CUDR	Chr19	1.4 – 2.7	Chemotherapeutic resistance	(46)
HEIH	Chr5	1.7	Associated with HBV-HCC	(40)
MEG3	Chr14	~1.8	De-regulated in HCC, associated with methylation. Predictive biomarker for monitoring epigenetic therapy	(17, 47)
H19	Chr11	2.3	Increased in HCC, or in peritumoral areas Low peritumor/tumor expression correlates with prognosis. Suppression of tumor metastases through miR-220 dependent inhibition of EMT, drug resistance	(48–51)
HOTAIR	Chr12	2.3	HCC recurrence after LT. Inhibition reduces invasion, increases chemosensitivity	(41, 42, 52)
HOTTIP	Chr 7	7.9	Up-regulated in HCC. Predicts disease outcomes and tumor progression	(31)
MALAT-1	Chr11	8.7	Upregulated in HCC. Associated with cancer metastasis and recurrence	(19, 53)
MVIH			Microvascular invasion in HCC. Predicts recurrence-free survival, overall survival	(43)
LINC-ROR	Chr 18	22.8	Tumor cell survival during hypoxia	(10)

Kb: kilobases