Cardiovascular Disease and Long Noncoding RNAs

Tools for Unraveling the Mystery Lnc-ing RNA and Phenotype

Elizabeth J. Hennessy, PhD

The last decade has ushered in a surge of genetic information with budget-friendly and more efficient sequencing technologies adding to our understanding of human development and disease. The sequencing of the human genome was a remarkable feat, yet using this information to understand human health and disease has proven to be challenging. The mammalian genome is comprised of a complex infrastructure of defined nucleotide sequences and dynamic epigenetic modifications that result in shifts in gene expression patterns and subsequent developmental and phenotypic outcomes. Genome wide association studies (GWAS) have demonstrated the link between alterations in nucleotide sequences created by mutations, such as single nucleotide polymorphisms (SNPs) and disease. However, it remains unclear how SNPs alter gene expression patterns and phenotypes. One hypothesis links SNP containing regions to mutational phenotypes via changes to epigenetic processes that control how genes are regulated, such as DNA methylation and histone acetylation.

SNPs are changes in nucleotide sequences that occur in at least 1% of the population. It is estimated that there are 10 to 30 million SNPs in humans that occur every 100 to 300 bases, and this variation is the major source of heterogeneity among people. A nonsynonymous SNP changes the amino acid sequence of a protein-coding gene. Less than 10% of SNPs are nonsynonymous, whereas 90% occur in nonprotein coding regions of the genome.[1] SNPs can be found in regions of deoxyribonuclease I (DNase I) hypersensitivity or promoters affecting transcription factor–binding sites and chromatin state. SNPs can create or delete microRNA-binding sites in 3′ untranslated regions (UTRs) affecting microRNA target mRNA expression.[2] SNPs can also be found in regions expressing noncoding RNAs, such as long noncoding RNAs (lncRNAs) leading to alterations in their expression patterns. SNPs can affect alternative splicing and the secondary structure of an RNA transcript leading to altered function (Figure 1). It has been estimated that 7% of SNPs that associate with autoimmune diseases are found in lncRNAs located in intergenic regions, signifying the increasing importance of these noncoding regions.[3] LncRNAs represent undiscovered disease-associated loci that could be identified through GWAS analysis.

The National Institutes of Health sponsored ENCODE (Encyclopedia of DNA Elements) project set out to uncover all of the functional elements of the human genome. It revealed that 85% of the human genome is transcribed into several classes of noncoding RNAs, whereas only 3% is translated into protein.[4,5] RNA-sequencing (RNA-seq) is currently the most widespread method for detecting the expression of RNA transcripts and for identifying novel noncoding transcripts.[6] The importance of noncoding RNA is evident by the fact that as the complexity of an organism increases, the abundance of noncoding RNA sequences found in its genome also grows illustrating the requirement for more sophisticated gene regulation in evolved eukaryotic species.[7] LncRNAs are a class of noncoding RNA that is broadly defined as nonprotein coding transcripts >200 nucleotides. They share some features of mRNAs, including polyadenylation and splicing, so they can be difficult to identify and distinguish from neighboring genes.[8,9] They exhibit specific and regulated patterns of expression in cells and tissues, and this can help when identifying them and trying to determine their function.[10–16] They have no strict sequence conservation restraints like protein-coding genes, so lncRNA sequences are often poorly conserved between species.

LncRNAs do not share a common mode of action, but they can be categorized into 4 general methods used to execute their functions, as (1) signals, (2) decoys, (3) guides, or (4) scaffolds where lncRNAs interact with DNA, RNA, and proteins (Figure 2). These interactions can result in cellular epigenetic modifications, such as changes to DNA methylation status,[17] modifications to histones, and the remodeling of chromatin, leading to changes in the expression of target genes. Although several thousand lncRNAs have been identified in the genome,[9,18,19] the function of only a limited number has, thus, far been described. Functions for lncRNAs have been described in various cellular processes, such as development and disease, including X-chromosome inactivation (Xist/Tsix), cancer metastasis (HOTAIR), nuclear import (Nron), type 2 diabetes mellitus (HI-LNC25 and KCNQ1OT1), and inflammation (LincRNA-COX-2 and LETHE).[20–32] Interestingly, 2 lncRNAs have been identified that were once thought to be noncoding for proteins but were re-examined for open reading frames and are now recognized to form micropeptides that play essential roles in muscle function.[33,34] It is becoming more apparent that the criteria for identifying lncRNAs and their functional implications are constantly evolving; so at the present time, there is no fixed set of rules.

From the Perelman School of Medicine, University of Pennsylvania, Philadelphia.
Correspondence to Elizabeth J. Hennessy, PhD, Perelman School of Medicine, University of Pennsylvania, 3400 Civic Center Blvd, Philadelphia, PA 19104. E-mail ehenn@upenn.edu

(*Circ Cardiovasc Genet*. 2017;10:e001556. DOI: 10.1161/CIRCGENETICS.117.001556.)

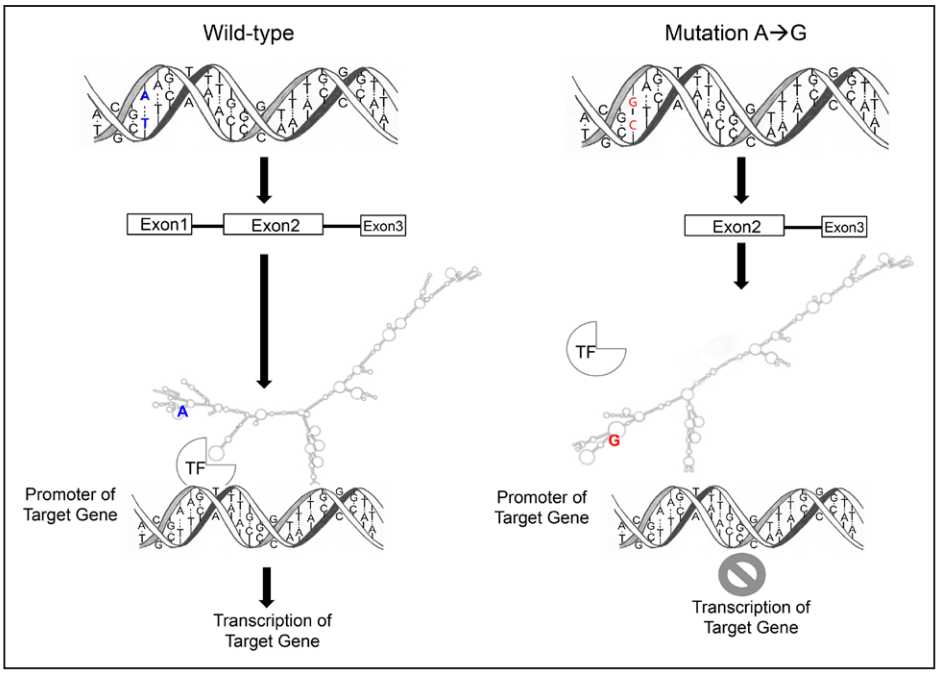

Figure 1. The Effect of SNPs on RNA Secondary Structure. SNPs found in long noncoding genes can alter the splicing pattern of a RNA transcript and the subsequent secondary structure leading to effects on the functional interactions of the lncRNA.

Recent studies have indicated that noncoding RNAs play important roles in the regulation of genes involved in the development and progression of cardiovascular disease (CVD) with several being classified as epi-lncRNAs or lncRNAs involved in epigenetic regulation through their interactions with epigenetic modifiers.[35] FENDRR (Foxf1 adjacent noncoding developmental regulatory RNA) is an essential regulator of heart development and is required for the proper development of the mouse, and this was shown using 2 different knockout models, 1 disrupting FENDRR transcription through the insertion of transcription termination signals in the first exon and the other inserting a lacZ reporter cassette in place of the FENDRR gene.[15,16,36] Both mouse models demonstrated the lethality of FENDRR deficiency but that it occurs at different stages of development: 1 prenatally[16] and the other perinatally,[36] and the authors attribute this to the differences in targeting strategies. FENDRR modifies the chromatin signature of genes involved in the formation and differentiation of the lateral mesoderm lineage. It interacts with the polycomb repressor complex 2 (PRC2) bringing it to target gene promoters (including its neighbor gene Foxf1) and increasing H3K27me3 (Histone3 lysine27 trimethylation), a repressive mark at these sites suppressing target gene transcription. Like FENDRR, Braveheart

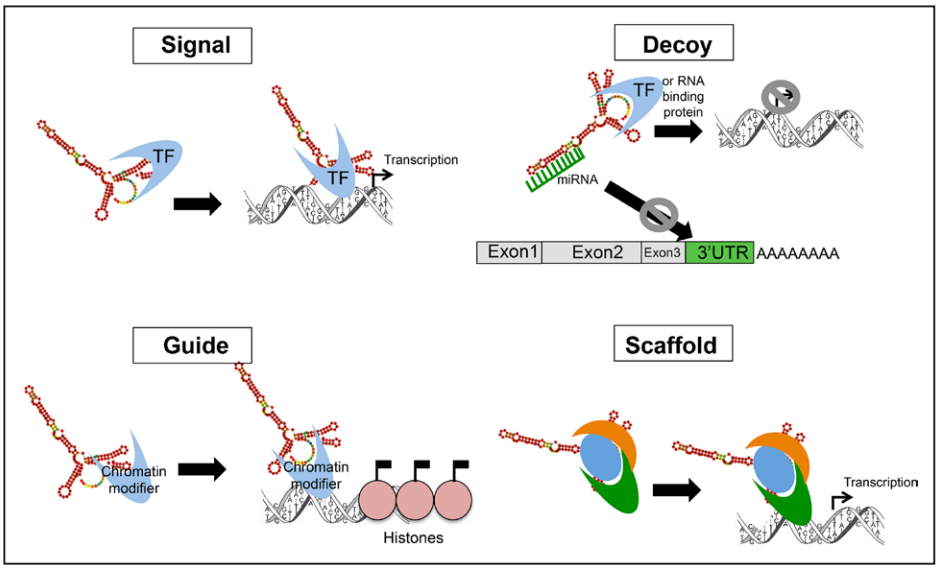

Figure 2. Four General Mechanisms of Action for Long Noncoding RNAs. LncRNAs can act as signaling molecules bringing transcription factors to promoter regions; decoys sequestering away TFs, RNA-binding proteins or miRNAs; guides for chromatin modifying enzymes or scaffolds for bringing together complexes of proteins.

is an epi-lncRNA that is also required for cardiovascular lineage commitment by interacting with a component of the PRC2 complex, SUZ12, to control gene regulatory networks, including key transcription factors (MesP1, Gata4, Hand2, Nkx2.5, and Tbx5).[37] In Braveheart deficient cells, SUZ12 and the repressive mark H3K27me3 were enriched at the promoters of these transcription factors decreasing their activity and the ability of the cells to differentiate into cardiomyocytes suggesting that Braveheart sequesters away the SUZ12 component of PRC2 from these transcription factor genes. However, it remains unknown whether Braveheart plays a direct role in the epigenetic regulation of these transcription factors. A structural study of Braveheart uncovered a key motif in the lncRNA that interacts with a nucleic acid–binding protein and affects cardiomyocyte differentiation.[38] Currently, there is no knockout model for Braveheart function and no human homologue, so translational studies are limited. Chaer (cardiac-hypertrophy-associated epigenetic regulator) is another epi-lncRNA that is heart specific and acts as an epigenetic checkpoint. Like FENDRR and Braveheart, it interacts with the repressor protein PRC2 and prevents it from targeting the promoters of genes involved in cardiac hypertrophy and repressing them via H3K27 trimethylation at the sites.[39] Despite displaying required roles in development, cardiac lineage commitment and hypertrophy through epigenetic regulation, the precise mechanisms for these epi-lncRNAs action remains unclear.

SENCR (smooth muscle and endothelial cell–enriched migration/differentiation-associated long noncoding RNA) is a lncRNA that was identified in a study aimed at distinguishing the lncRNA profile of vascular cells, including smooth muscle cells and endothelial cells. Its expression affects smooth muscle cell migration. Silencing SENCR led to increased expression of promigratory and contractility genes including myocardin, a key transcriptional switch for smooth muscle cell contractile gene expression, and 2 migratory genes MDK and PTN.[40] MALAT1 (Metastasis-associated lung adenocarcinoma transcript 1) is expressed at high levels in endothelial cells. It regulates migration and vascular sprouting, and its expression is increased in response to hypoxia. Silencing of MALAT1 in vitro promotes a migratory endothelial cell phenotype, and pharmacological inhibition of MALAT1 in vivo reduces vascular growth.[41] Despite showing a vascular phenotype in vitro and in vivo, the mechanism underlying these changes is unknown. MALAT1 has been shown to interact with a protein in the polycomb complex CBX4 to regulate histone modifications and control cell proliferation and lung metastasis, but this epigenetic function for MALAT1 has not been shown in endothelial cells. In contrast to these studies, 2 different MALAT1 knockout mouse models have demonstrated that this lncRNA is dispensable for functions that could be linked to the in vitro phenotypes, including pressure overload–induced heart failure in mice, mouse pre and postnatal development, and global gene expression.[42,43] However, in a MALAT1 conditional knockout mouse model and in mice treated with a MALAT1 antisense inhibitor, there was a decrease in lung tumor metastasis, and the expression of *cis*-genes was affected. These conflicting findings in the in vivo modulation of MALAT1 illustrate that MALAT1 like other lncRNAs may play redundant roles in the whole organism for some functions but are still relevant when studying distinct cellular processes. The varied observations between cellular and whole organism studies illustrate the complexity of lncRNA research. We do not know enough about their mechanisms to disregard them. Each of the lncRNA studies described demonstrates their cell type specificity and the potential for targeting specific cardiovascular phenotypic processes through the modulation of their expression. It is evident that the identification of lncRNAs and their effect on the cellular epigenetic signature has created new possibilities for researchers to develop novel therapeutics for CVD.

SNPs in Noncoding Genes and Their Link to CVD

Several lncRNAs that contain SNPs implicated in the development and severity of CVD have been identified through GWAS analysis. The lncRNA MIAT (myocardial infarction associated transcript) was identified in a 2000-person case-control association study of a myocardial infarction (MI) susceptibility locus located on chromosome 22.[44] A follow-up study confirmed this finding and that altered expression of MIAT resulting from 6 independent SNPs in the MIAT gene at location chromosome 22q12.1 conferred genetic susceptibility to MI.[45] One of these SNPs located in exon 5 increased the expression of MIAT and altered its ability to bind to an as yet uncharacterized nuclear protein. MIAT has been implicated in microvascular dysfunction with increased expression in diabetic retinas and in response to high glucose linking it to diabetic pathology.[46] Another study showed that MIAT expression was significantly decreased in peripheral blood cells and platelets of patients with acute MI, whereas showing a positive association with lymphocytes and a negative association with neutrophils and platelets.[47] MIAT has also been described as a competing endogenous RNA (ceRNA) targeting miR-150-5p and affecting levels of miR-150-5p target mRNA vascular endothelial growth factor (VEGF).[46] Lastly, in a study of 414 patients with acute MI, there was an association between levels of several lncRNAs, including MIAT and ANRIL, with inflammatory markers, such as matrix metalloproteinase 9 (MMP9), illustrating their potential role in the regulation of the inflammatory response.[48]

ANRIL (antisense noncoding RNA in the INK4 locus) was identified through GWAS analysis that linked its locus on human chromosome 9p21 to coronary artery disease (CAD) and myocardial infarction susceptibility.[22,23,48–50] ANRIL is expressed in endothelial, smooth muscle, and inflammatory cells, and its expression is associated with risk for atherosclerosis, peripheral artery disease, and other vascular diseases.[51] A primate specific ALU element–containing motif was identified in ANRIL that acts to regulate genes involved in proliferation, cell adhesion, and apoptosis.[50] SNPs in the ANRIL gene result in the alternative splicing of the transcript into many isoforms, including a circular form (circANRIL), and each variant correlates with CAD to a different degree. CircANRIL regulates the maturation of precursor ribosomal RNA and ribosome biogenesis leading to reduced proliferation and protection against atherosclerosis.[52] Carriers of the ANRIL risk

alleles exhibit increased whole blood RNA expression levels of 2 of the shortest ANRIL variants and decreased expression of the longest variant demonstrating the differential impact of the different isoforms of ANRIL in response to SNPs in the gene. However, despite these correlation studies, the precise biological function of ANRIL remains unknown.

Linc-VWF is a lncRNA that is highly expressed in endothelial cells and is induced by lipopolysaccharide. Linc-VWF is located 105 kb from the VWF gene—an endothelial and platelet-derived circulating plasma glycoprotein that plays a central role in hemostasis and thrombosis. A study by Liu et al[53] showed that linc-VWF contains an SNP (rs1558324) associated with mean platelet volume—a predictor of cardiovascular disease; however, there is no linkage disequilibrium between the linc-VWF SNP and SNPs in the VWF gene that have been associated with circulating VWF levels. Additional work needs to be done to elucidate the mode of action for linc-VWF after stimulation with LPS, including potential protein-binding partners that could explain the association of the SNP to platelet volume.

SNPs have also been identified in the lncRNA H19 conferring either susceptibility or protection from CVD (rs217727 and rs2067051). The expression of H19 is induced during embryogenesis and downregulated after birth except in the heart. It is induced in response to homocysteine and is increased in the aortae of mice with hyperhomocysteinemia—a risk factor for CAD. H19 has been linked to different types of cancers, and this is thought to occur via the interaction between H19 and the microRNA let-7. H19 binds to and sequesters let-7 preventing it from inhibiting the expression of its target genes. Impaired let-7 has also been implicated in CVD, so it is thought that H19 could be a regulator of let-7 expression, and, therefore, CVD.[54] The few examples presented here (Summarized in Table 1) describe the potential for lncRNAs to affect cardiovascular phenotypes via SNPs found within their nucleotide sequences. How these noncoding RNA-associated SNPs interact with other factors, such as via their secondary structures, is currently being explored and is proving to be essential to understanding their function.

The Potential for SNPs to Alter RNA Secondary and Tertiary Structures

The genes encoding lncRNAs are thought to evolve rapidly, so their sequences are often poorly conserved between species. Because of the rapid evolution of lncRNAs, SNPs are likely to be created. Unlike protein coding genes, lncRNAs do not need to maintain sequence conservation to maintain their functionality because it is thought that the secondary structure of the RNA molecule is important for function, and this can be conserved. One of the first lncRNAs to be described, Xist regulates X-chromosome inactivation, and it has undergone rapid sequence evolution, whereas preserving its function. Despite moderate sequence conservation, Xist displays conserved RNA secondary structure between various species.[55] Another lncRNA, MEG3 was used as an example to demonstrate the importance of secondary structure over nucleotide conservation. One hundred and fifty-seven nucleotides from the wild-type MEG3 sequence that formed a stem-loop structure were replaced with a different nucleotide sequence that could also form a similar stem loop forming a hybrid transcript. According to RNAfold, the artificial hybrid RNA molecule displayed a similar secondary folding structure to the wild-type MEG3 and had an equivalent effect on p53 activation demonstrating that for some lncRNAs, it is not the sequence that is important for function but the resulting secondary structure that forms.[56]

Single base-pair changes in the genetic code can alter sequence elements that are essential for correct splicing or translation of an mRNA, as well as change the structure and folding of an mRNA molecule and, therefore, its function (Figure 1).[57] Computational and experimental methods can be used to define the secondary structure for a given RNA. Online algorithms include Vienna RNAFold,[58] RNAStructure,[59] Sfold,[60] and RNAMutants.[61] These algorithms use thermodynamics to calculate the minimum free energy presented by the RNA structure. Negative values signify that less energy is required for base pairs to interact, so it is more likely to be a valid and stable predicted structure. Stem-loop structures are thought to be the functional elements of the transcript that will interact with DNA, RNA, or proteins; however, current bioinformatics tools are not able to accurately detect the short sequences that may form important stem-loop structures. The database lncRNASNP provides information about SNPs encoded in lncRNA genes, such as the potential effects of SNPs on structure, lncRNA:miRNA-binding relationships, and their possible functions.[62] Databases like lncRNASNP are adding details to the mystery between SNPs encoded in lncRNA genes and structure/function relationships. However, it is unknown how in vitro examples of RNA structures compare with their in vivo counterparts. There could be alternatively folded secondary structures depending on the scenario, and more studies are needed to conclusively determine the effect of SNPs on disease pathology.

Methods are rapidly being sought to experimentally validate in silico-predicted secondary structures and to identify key motifs in RNA molecules that will help predict interactions between RNA, DNA, and protein.[38] Earlier studies have probed genome-wide RNA secondary structures using dimethyl sulfate sequencing (DMS-seq).[63] Specific bases in the RNA molecule are methylated with DMS preventing natural hydrogen bonds from forming between modified bases. The resulting altered RNA is then sequenced, and the interruptions in the signal indicate modifications to the structure like single-stranded stem-loops. Advancing from the DMS protocol, selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) can predict RNA secondary structure at single-nucleotide resolution and uses chemical probing agents like N-methylisatoic anhydride (NMIA) or 1-methyl-7-nitroisatoic anhydride (1M7) to target hydroxyl groups and acylate single-stranded or flexible regions of RNA (Figure 3A).[64,65] The sites of chemical modification are detected by reverse transcription. The reverse transcriptase enzyme pauses at RNA nucleotides that are modified by 1M7 indicating single strandedness. Nucleotides that are constrained by base-pairing show less product formation than nucleotides that are unpaired. The resulting cDNA library maps ribonucleotides that are single stranded in the context of the folded RNA. Products of this reaction are

Table 1. List of Cardiovascular-Associated Long Noncoding RNAs

LncRNA	Species	Cell/Tissue Type	Function	In Vivo Phenotype	References
FENDRR	Mouse/ Human	Embryonic stem cells Caudal end of the lateral plate mesoderm of midgestation embryos	Modifies chromatin signature of genes involved in the formation and differentiation of lateral mesoderm lineage	Embryonic lethal	15,16, 35, 36
Braveheart	Mouse	Embryonic stem cells Cardiomyocytes	Required for cardiovascular lineage commitment	None	37,38
Chaer	Mouse/ Human	Heart Epicardium, where progenitor cells for endothelial cells and fibroblasts reside	Required in development, cardiac lineage commitment, and hypertrophy	Regulates hypertrophy	39
SENCR	Human	Smooth muscle cells Endothelial cells	Affects smooth muscle cell migration	None	40
MALAT1	Mouse/ Human	Endothelial cells	Regulates migration and vascular sprouting, and its expression is increased in response to hypoxia	Pharmacological inhibition in vivo reduces vascular growth In vivo ko showed it is dispensable for pressure overload–induced heart failure Antisense inhibitor in vivo showed a decrease in lung tumor metastasis, and expression of cis-genes were affected	41–43
MIAT	Mouse/ Human	Positive association with lymphocytes Negative association with neutrophils and platelets	SNPs in exon 5 increased the expression of MIAT and altered its ability to bind to an as yet uncharacterized nuclear protein	Implicated in microvascular dysfunction Decreased in peripheral blood cells and platelets of patients with acute MI Associated with MMP9	44–48
ANRIL	Human	Endothelial cells Smooth muscle cells Macrophages	Unknown	Associated with risk for CAD, MI, atherosclerosis, and peripheral artery disease	22, 23, 48–52
Linc-VWF	Human	Endothelial cells	LPS inducible	SNP associated with mean platelet volume	53
H19	Human	Heart Aorta	Induced in response to homocysteine	Increased in aorta of mice with hyperhomocysteinemia	54

CVD has several lncRNAs implicated in its pathology. LncRNAs are listed with the cell or tissue type where it is primarily expressed, known functions and in vivo evidence.

then fractionated by capillary electrophoresis (CE) generating electropherograms that are converted into nucleotide reactivity tables that are then converted into pseudoenergy constraints by the prediction algorithms like RNAStructure. The 2D RNA structures obtained by combining SHAPE with computational RNA secondary structure prediction programs are more accurate than structures obtained using either method alone.

A key motif in the lncRNA Braveheart that is required for the development of cardiovascular cells was identified using SHAPE.[38] This study set out to investigate the molecular mechanism of Braveheart action by determining the secondary structure of in vitro transcribed full-length Braveheart using SHAPE and DMS probing. The Braveheart transcript is organized into a highly modular structure, including an asymmetrical G-rich internal loop (AGIL) at the 5′ end. The authors then used CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats) technology to delete this AGIL motif and observed an inability of embryonic stem cells to differentiate into cardiomyocytes. The AGIL motif interacts with key factors expressed in the heart, such as the zinc-finger protein CNBP/ZNF9, which is known to bind to single-stranded G-rich sequences and repress cardiomyocyte differentiation. Through its AGIL motif, Braveheart antagonizes CNBP to promote differentiation. The authors have yet to identify a human homolog to the mouse Braveheart transcript, but the structure of the mouse Braveheart lncRNA can be used to analyze human lncRNA molecules and identify similar structures/motifs that may have similar functions. Motifs found in RNA transcripts using secondary structure analysis may contain key regions of interaction for SNPs, and they can be used as fingerprints to make a catalog of key motifs.

Currently, there are no examples of cardiovascular disease–associated SNPs in lncRNAs that affect the secondary structure of the lncRNA, but an example of a polymorphism associated with celiac disease was found in the lncRNA Lnc13 that affects its secondary structure.[66] Lnc13 is downregulated in small intestine biopsies from celiac patients, and LPS downregulates Lnc13 via degradation by the mRNA decapping enzyme Dcp2. In healthy people, Lnc13 acts as a brake to keep a subset of inflammatory genes downregulated through interaction with the nuclear proteins heterogeneous nuclear ribonucleoprotein D (hnRNPD) and histone deacetylase 1 (HDAC1). However, in patients with the disease-associated SNP, Lnc13 binds hnRNPD less efficiently than its wild-type counterpart leading to increased expression of the inflammatory genes. The predicted secondary structures for wild-type and disease-associated Lnc13 were significantly different. In vitro transcribed

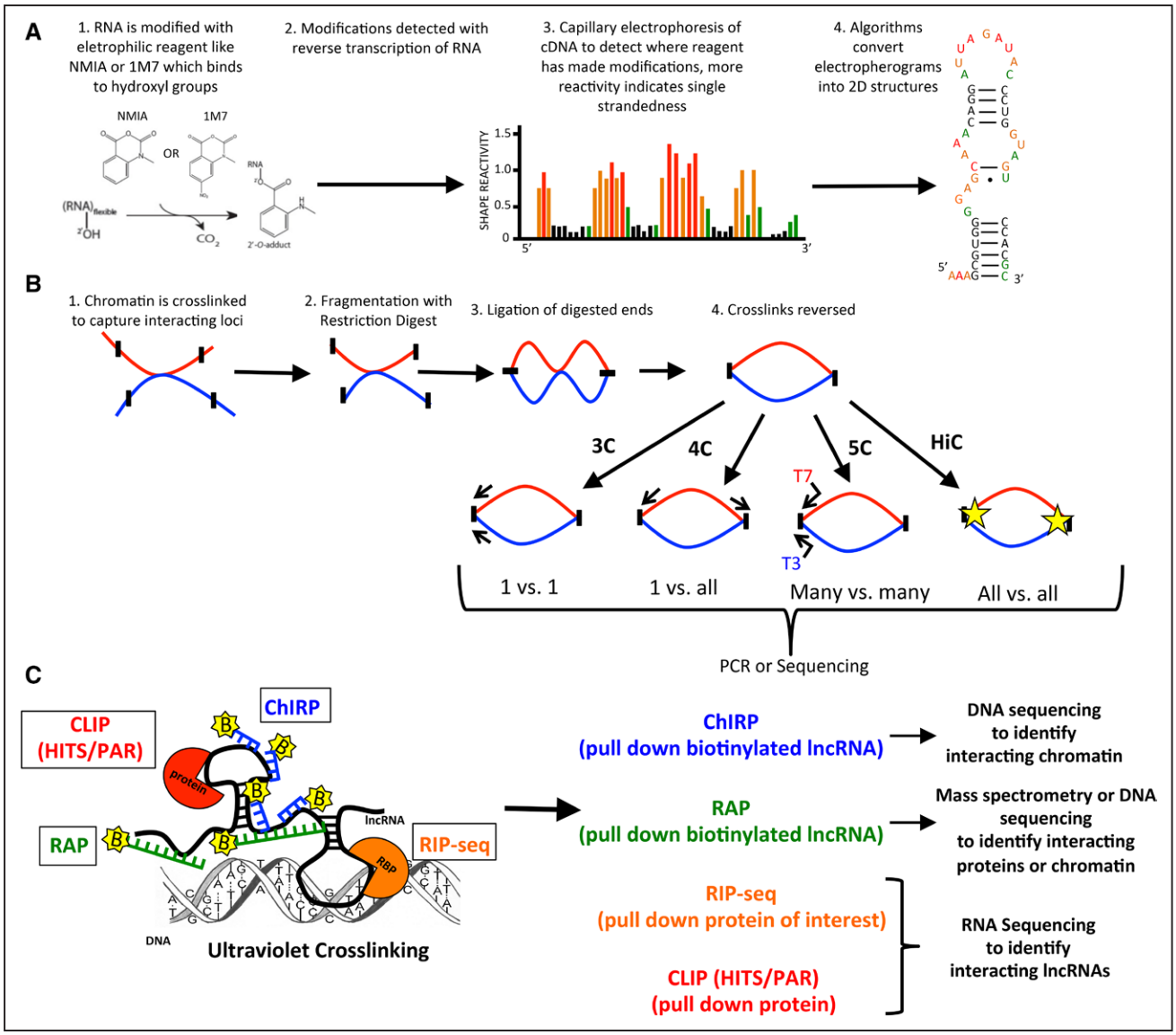

Figure 3. Techniques for Determining the Functions of Long Noncoding RNAs. **A**) SHAPE uses reagents to detect nucleotides that are constrained by base-pairing, they will show less product formation (see nucleotides in black) when fractionated by capillary electrophoresis than nucleotides that are unpaired. (**B**) Chromosome capture experiments allow for the identification of interacting loci in the genome, a typical scenario for lncRNAs. DNA is crosslinked, cut into fragments with restriction enzymes and then randomly ligated back together. 3C, 4C, 5C and Hi-C differ only in the methods used to identify interacting loci. (C) Various pull-down techniques can be used to detect interactions between a single target lncRNA or a protein of interest.

RNA for wild-type and disease-associated Lnc13 showed dramatically different mobilities on a native agarose gel signifying that the structures for these RNA transcripts indeed differ, and this may play a role in their functional differences.

It is becoming evident that in addition to secondary structures, tertiary structures of RNA molecules could be key to understanding their functional consequences. A recent study described both the secondary and tertiary structures for the well-characterized lncRNA Xist and the lncRNA RepA whose sequence is identical to the 5′ region of Xist. RepA is thought to function in the initiation and spreading of X-chromosome inactivation by Xist.[67] This study was able to show that 3 dimensional structures were able to form in the absence of protein partners signifying that Xist and RepA do not require the presence of their protein partners to form their structures.

The laboratory of Rhiju Das is making advances toward understanding the relationship between secondary and tertiary RNA structures of noncoding RNAs by studying benchmarking RNAs like ribozymes, riboswitches, and ribosomal RNA domains in cells and viruses with high-resolution computational approaches and multidimensional chemical mapping to uncover 3 dimensional structures.[68,69] The low abundance of many lncRNAs and the complexity of proteins binding to RNAs can make in vivo interpretation of in vitro signals more complicated. An attempt to examine RNA structures in vivo was made with in vivo click selective 2′ hydroxyl acylation and profiling experiment (icSHAPE-seq) in living cells. It revealed active unfolding of mRNA structures suggesting that RNA structures are contributing to global RNA processing.[70] Studying the secondary and tertiary structures of lncRNAs is

contributing to a better understanding of how noncoding transcripts and alterations in sequences because of SNPs affect disease progression. These studies will aid in the creation of guidelines for streamlining lncRNA studies and ultimately the development of new therapies targeting disease-causing SNPs.

Noncoding RNA Interactions and Their Effect on the Epigenetic Signature of the Cell

It is well established that the epigenetic landscape of the cell is responsible for alterations in the accessibility to chromatin by molecules like chromatin-modifying enzymes, transcription factors, and more recently, lncRNAs resulting in changes in gene expression, cellular differentiation, development, and disease pathology. Specific combinations of modifications constitute a unique epigenetic signature, and variations in the genomic sequence caused by SNPs can affect this signature. A prominent role for lncRNAs has recently been established in epigenetic regulation through chromatin remodeling via the targeting of specific loci. Many lncRNAs have been shown to complex with chromatin-modifying enzymes and guide them to specific sites in the genome altering histone acetylation and affecting the methylation status of DNA through interactions with histone lysine methyltransferases.[17] Acetylation is generally thought to open up chromatin allowing for proteins to access it, whereas an increase in methylation at a gene's promoter can open or close the chromatin depending on the type of methylation. These epigenetic changes alter the ability of proteins like transcription factors to bind to the region leading to changes in the expression of the gene.

Understanding the interactions between lncRNAs with DNA or proteins is proving to be the key to their function, so techniques are rapidly being developed to elucidate these interactions and functional consequences. Chromosome conformation capture (3C) is used to identify chromatin loops. Cellular DNA is cross-linked to freeze interactions between genomic loci. The DNA is then cut into fragments and randomly ligated, and the proximity of fragments is quantified because fragments are likely to ligate to neighboring fragments. The ligated fragments are detected using primers to specific regions. This technique allows for the study of interactions between single genomic loci with other single genomic loci. Updates to the 3C technique are quickly expanding for more detailed and specific analysis of chromatin interactions (Figure 3B).[71,72] Chromosome conformation capture-on-chip (4C) captures the interactions between a single specific gene locus with other loci from across the genome.[73] Chromosome conformation capture carbon copy (5C) examines interactions within a specific region not greater than a megabase using primers that have a universal sequence at the end, such as T7 or T3 allowing, for the amplification of all the ligation products.[74,75] Lastly, Hi-C is a genome-wide high-throughout method for detecting all potential interactions across the entire genome. Biotinylated nucleotides are annealed to the ends of the DNA before the ligation step.[76] The DNA is then pulled down to isolate only fragments containing biotin, and these products are then sequenced for interactions. These methods can be used for determining the *cis* and *trans* effect of a particular lncRNA containing loci and to see how the genomic architecture can be affected by disease-associated SNPs, particularly those encoded in a noncoding locus.[77] Chromosome capture techniques are beginning to reveal the complex organization of the genome and the potential transcriptional regulatory mechanisms governing it.

To examine direct interactions of a particular lncRNA with chromatin, methods like chromatin isolation by RNA purification (ChIRP) can be used on a genome-wide scale and without bias to identify regions of DNA that interact with a lncRNA. Cultured cells are UV cross-linked, and chromatin is extracted. Short biotinylated labeled oligonucleotides that are complementary to the lncRNA of interest are hybridized to the samples and pulled down for analysis of bound chromatin that can be identified with deep sequencing.[78] RNA antisense purification (RAP) uses a similar method to ChIRP to identify interacting regions of DNA or proteins for a given lncRNA with the exception that it uses much longer biotinylated labeled primers for lncRNA detection supposedly leading to greater specificity.[79] Interacting proteins are identified with mass spectrometry, and deep sequencing is used to identify regions of chromatin that are interacting. A recent study from Atianand et al[80] used RAP to demonstrate the interaction of the lncRNA lincRNA-EPS with chromatin and the heterogeneous ribonucleoprotein hnRNPL to repress immune response genes. Unbiased mass spectrometry protocols are being used to identify protein-binding partners for a specific lncRNA of interest. In a protocol by Xing et al,[81] the lncRNA is in vitro transcribed from DNA, and biotins are incorporated, the biotinylated RNA is applied to cellular lysates to allow for proteins to bind, the sample is then pulled-down using streptavidin beads, and bound proteins are identified using mass spectrometry. There are advantages to each of these approaches, ChIRP and RAP use endogenous RNA as the bait for identifying interacting chromatin or proteins, whereas the unbiased mass spectrometry approach uses artificially generated RNA that might not have the same secondary structure as endogenous RNA. However, this approach can be beneficial in that the in vitro-transcribed RNA that is used as a template for the pull-down will likely result in many more protein hits than either ChIRP or RAP, and these hits can than be validated using an approach like RNA-immunoprecipitation sequencing (RIP-seq) to confirm associated proteins.

Specific proteins, such as those in the chromatin-modifying complex PRC2 whose components have been shown to interact with many lncRNAs, can be pulled down using RIP-seq. RNA transcripts that are associating with the complex can then be identified using RNA-seq.[82] Several cross-linking immunoprecipitation (CLIP) techniques can be used to isolate RNA and the associated proteins. High throughput sequencing-CLIP (HITS-CLIP) identifies transcriptome-wide RNAs bound to a particular protein.[83] Photoactivatable ribonucleoside enhanced–CLIP (PAR-CLIP) looks for interacting RNAs and RNA-binding proteins through the incorporation of photoreactive ribonucleoside analogs into nascent RNA transcripts increasing the efficiency of cross-linking and specificity. The RNA-binding protein of interest is pulled down, and the RNA attached is identified with next generation sequencing (Figure 3C).[84] Each of these techniques is contributing to the picture of the transcriptional landscape of the cell through the elucidation and characterization of lncRNA interactions, and adding GWAS studies and the locations of SNPs

to this picture will enable a greater understanding of CVD progression.

The Power of Gene Editing for Correcting Disease-Associated SNPs

The manipulation of the genome with nucleotide-based therapeutics, such as antisense oligonucleotides and CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats), holds tremendous potential as new treatments for disease by potentially correcting disease-associated SNPs. RNA therapies are thought to be particularly well suited for nucleic acid targeting because as little as 8 nucleotides is needed for targeting and disrupting an RNA transcript. RNA is malleable and, therefore, more tolerant of mutations even within regions responsible for target recognition. However, depending on the localization of a particular lncRNA, different approaches may be used to disrupt its expression, so caution needs to be taken when selecting a method for modulation. Antisense oligonucleotides can be modified to avoid degradation by ribonucleases and target nuclear localized transcripts, whereas RNA interference (RNAi) uses unmodified short interfering RNAs (siRNAs) or short hairpin RNAs (shRNAs) that will only target transcripts found in the cytoplasm of the cell.

CRISPR-Cas9 is a method that can be used to target the DNA of the cell and is used to create permanent changes to the genome with the potential to correct genetic mutations. Using a short single-strand guide RNA (sgRNA), the Cas9 enzyme can home in on SNPs within the 3 billion nucleotides of the human genome and cut out or alter the problematic sequence. The Cas9 enzyme creates a site-specific double-stranded break that can be repaired using nonhomologous end joining (NHEJ) or homology-directed repair (HDR). NHEJ nonspecifically repairs the DNA resulting in insertions and deletions (INDELs) at the target locus. HDR can be used to alter the DNA specifically by creating mutations, deletions, or inserting DNA elements into the double-strand break. These changes can result in the correction of an SNP or the silencing or upregulation of a target gene or genes.[85,86] However, once inside the cell, Cas9 can dock and cut at undesired sites potentially producing off-target edits, so researchers are rapidly working to find better tools for predicting edit sites and more efficient ways for getting CRISPR into cells.

A few key examples have emerged demonstrating the therapeutic power of CRISPR-Cas9 in fixing specific mutations in a particular gene. In a recent study by Yin et al,[87] CRISPR-Cas9 was used in vivo to correct a mouse model of hereditary tyrosinemia type 1—a rare condition resulting from the lack of a metabolic enzyme required for tyrosine metabolism. A point mutation in the gene encoding this metabolic enzyme results in a truncated and unstable transcript, and without the enzyme, toxic metabolites accumulate in hepatocytes resulting in severe liver damage. CRISPR-Cas9 has also been used to correct the dystrophin gene (Dmd) mutation in a mouse model of Duchenne muscular dystrophy[88–90] and repair a point mutation responsible for sickle cell disease, so the cells produce enough hemoglobin to have a potential clinical benefit.[91] Beyond the scope of just repairing disease-causing SNPs, CRISPR-Cas9 has excised HIV-inserted genes from immune cells and blocked HIV from entering blood stem cells but with complications, including the virus eventually overcoming the CRISPR edits and using the edits made to its genome to its advantage.[87,92]

These initial studies can be extended to the study of CVD-associated SNPs in lncRNAs, such as MIAT and ANRIL, but it must be kept in mind that some lncRNAs are considered untargetable because of their location in the genome. A recent study measured the CRISPR ability of human lncRNAs and determined that the position of a lncRNA, whether it is intragenic, bidirectional, intergenic, or proximal to another gene, determined the ability of CRISPR-Cas9 to effectively target it.[93] The authors conclude that for many lncRNAs, it will be necessary to use multiple approaches to modulate the expression of the lncRNA and to confirm findings seen with CRISPR-Cas9. However, it is a possibility that different effects on lncRNA-targeted genes or pathways will be observed because of mRNA targeting by antisense oligonucleotides versus genomic targeting by CRISPR-Cas9. Attention must be taken when using any nucleotide-based approach to alter the expression of lncRNAs.

A recently described CRISPR-Cas9 approach is allowing for the high throughput screening of lncRNAs and the examination of their role in particular functions. CRISPRi(nterference) uses a deactivated Cas9 (dCas) enzyme fused with the KRAB domain of the transcriptional repressor protein ZNF10 along with a specific sgRNA for targeting. Instead of cutting the DNA, these proteins block RNA polymerase from gaining access to the locus and silencing transcription of the target gene without genetically altering the target sequence. A 2016 study by Liu et al[94] designed a CRISPRi noncoding library targeting 16 401 lncRNA genes to identify lncRNA genes that modify cell growth in 7 different cell types. Their initial hypothesis was that there would be a core set of lncRNAs that would be essential to the survival of all cell types, but no such core set of lncRNAs was found. Eighty-nine percent of the lncRNAs screened were crucial for only 1 cell type and had no effect on the others. High throughput screens like this study will help to prioritize the lncRNAs that should be pursued for an in vivo phenotype because as demonstrated for MALAT1, the in vivo models did not recapitulate the in vitro phenotypes observed, and this could be because of the cell types examined and the functions pursued. The few examples provided of CRISPR-Cas9 gene-modifying technology demonstrate that this technology has massive translational relevance to human health in its potential ability to treat diseases like cancer and correct genetic mutations responsible for inherited disorders.

Conclusions and the Future of Noncoding RNA Studies

Elucidating the functions of lncRNAs has proven to be a challenge because unlike protein-coding genes, lncRNAs exhibit little sequence conservation, and it is thought that their function is ultimately determined by the RNA secondary structure and the DNA or protein that interacts with it. To gain a deeper understanding of the role of lncRNAs in the cell, a streamlined pipeline of experiments can be used to tackle the functional questions surrounding a newly discovered lncRNA of interest (Table 2). RNA-sequencing experiments are first used to identify and quantify lncRNAs in a particular cell type or disease scenario. Subcellular localization of the

Table 2. List of Techniques for Determining the Functions of Long Noncoding RNAs

Method	Purpose	Drawbacks/ Limitations	References
RNA-sequencing	Identify and quantify RNA	Different alignment algorithms will give varying results, especially with identifying novel ncRNAs transcripts	4–9
RNAFold RNAStructure Sfold RNAMutants	Predict secondary structures of single stranded RNA and accessibility using mean free energy	Does not look at tertiary structures or interactions with proteins that could act to stabilize or reshape the RNA	58–61
DMS sequencing SHAPE icSHAPE sequencing	Chemically validate RNA secondary structure, in vivo or in vitro	Determines the specific nucleotides that are paired but does not reveal binding partners, inherently noisy and reproducibility still remains an issue	63-65,70
3C 4C 5C Hi-C	Examine chromatin interactions between specific or genome-wide loci	Limited by the length of sites of interest, 2 sites close on a linear chromosome will form ligation junctions regardless whether involved in a loop, so close *cis* interactions will be difficult to observe	71–77
ChIRP RAP Mass spectrometry RIP sequencing HITS-CLIP PAR-CLIP	Identify interactions between lncRNAs with DNA or proteins	Native interactions may be disrupted by crosslinking and the use of a foreign bait to pull down the RNA or protein of interest may change their conformation, overexpression can also lead to artifacts	78–84
CRISPR-Cas9 CRISPRi	Gene-editing tool used to create sequence specific breaks in the genome or interfere with transcription of target gene	The gene loci needs to be thoroughly examined for CRISPRability, and alternative approaches like ASO should be used to complement CRISPR	85-87, 91-94

Techniques are rapidly being discovered for unraveling the mystery of lncRNA and function. The pipeline of this discovery begins with RNA sequencing followed by structural studies, interaction mapping, pull-downs for identifying interacting partners and CRISPR to modulate the expression of a lncRNA directly at its genomic locus.

lncRNA is assessed to determine the site of action. Gain and loss of function approaches, such as CRISPR-Cas9, are then used to determine the impact of the lncRNA on a given pathway or function. Lastly, a variety of pull-down experiments can be used to identify interacting DNA, RNA, or proteins. This pipeline of experiments culminates in the implication of a lncRNA in a particular pathology. The discovery that lncRNAs are playing integral roles in cardiovascular development and the progression of disease is leading researchers to search for a concise lncRNA code that can be used to develop novel drug targets and can identify biomarkers for the diagnosis of cardiovascular events.

Disclosures

None.

References

1. Collins FS, Brooks LD, Chakravarti A. A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res*. 1998;8:1229–1231.
2. Hrdlickova B, de Almeida RC, Borek Z, Withoff S. Genetic variation in the non-coding genome: involvement of micro-RNAs and long non-coding RNAs in disease. *Biochim Biophys Acta*. 2014;1842:1910–1922. doi: 10.1016/j.bbadis.2014.03.011.
3. Ricaño-Ponce I, Wijmenga C. Mapping of immune-mediated disease genes. *Annu Rev Genomics Hum Genet*. 2013;14:325–353. doi: 10.1146/annurev-genom-091212-153450.
4. ENCODE Project Consortium. The encode (encyclopedia of DNA elements) project. *Science*. 2004;306:636–640.
5. Hangauer MJ, Vaughn IW, McManus MT. Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLoS Genet*. 2013;9:e1003569. doi: 10.1371/journal.pgen.1003569.
6. Baruzzo G, Hayer KE, Kim EJ, Di Camillo B, FitzGerald GA, Grant GR. Simulation-based comprehensive benchmarking of RNA-seq aligners. *Nat Methods*. 2017;14:135–139. doi: 10.1038/nmeth.4106.
7. Taft RJ, Pheasant M, Mattick JS. The relationship between non-protein-coding DNA and eukaryotic complexity. *Bioessays*. 2007;29:288–299. doi: 10.1002/bies.20544.
8. 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;22:1775–1789. doi: 10.1101/gr.132159.111.
9. Jia H, Osak M, Bogu GK, Stanton LW, Johnson R, Lipovich L. Genome-wide computational identification and manual annotation of human long noncoding RNA genes. *RNA*. 2010;16:1478–1487. doi: 10.1261/rna.1951310.
10. Amaral PP, Mattick JS. Noncoding RNA in development. *Mamm Genome*. 2008;19:454–492. doi: 10.1007/s00335-008-9136-7.
11. Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell*. 2011;43:904–914. doi: 10.1016/j.molcel.2011.08.018.
12. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell*. 2011;147:358–369. doi: 10.1016/j.cell.2011.09.028.
13. Tay Y, Kats L, Salmena L, Weiss D, Tan SM, Ala U, et al. Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell*. 2011;147:344–357. doi: 10.1016/j.cell.2011.09.029.
14. Wang Y, Xu Z, Jiang J, Xu C, Kang J, Xiao L, et al. Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. *Dev Cell*. 2013;25:69–80. doi: 10.1016/j.devcel.2013.03.002.
15. Grote P, Herrmann BG. The long non-coding RNA Fendrr links epigenetic control mechanisms to gene regulatory networks in mammalian embryogenesis. *RNA Biol*. 2013;10:1579–1585. doi: 10.4161/rna.26165.
16. Grote P, Wittler L, Hendrix D, Koch F, Währisch S, Beisaw A, et al. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev Cell*. 2013;24:206–214. doi: 10.1016/j.devcel.2012.12.012.

17. Zhou J, Yang L, Zhong T, Mueller M, Men Y, Zhang N, et al. H19 lncRNA alters DNA methylation genome wide by regulating S-adenosylhomocysteine hydrolase. *Nat Commun*. 2015;6:10221. doi: 10.1038/ncomms10221.
18. Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res*. 2008;18:1433–1445. doi: 10.1101/gr.078378.108.
19. Ponjavic J, Oliver PL, Lunter G, Ponting CP. Genomic and transcriptional co-localization of protein-coding and long non-coding RNA pairs in the developing brain. *PLoS Genet*. 2009;5:e1000617. doi: 10.1371/journal.pgen.1000617.
20. Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, et al. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science*. 2008;322:1717–1720. doi: 10.1126/science.1163802.
21. Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, et al. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell*. 2008;32:232–246. doi: 10.1016/j.molcel.2008.08.022.
22. Jarinova O, Stewart AF, Roberts R, Wells G, Lau P, Naing T, et al. Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. *Arterioscler Thromb Vasc Biol*. 2009;29:1671–1677. doi: 10.1161/ATVBAHA.109.189522.
23. Burd CE, Jeck WR, Liu Y, Sanoff HK, Wang Z, Sharpless NE. Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet*. 2010;6:e1001233. doi: 10.1371/journal.pgen.1001233.
24. Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, et al. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science*. 2005;309:1570–1573. doi: 10.1126/science.1115901.
25. Carpenter S, Aiello D, Atianand MK, Ricci EP, Gandhi P, Hall LL, et al. A long noncoding RNA mediates both activation and repression of immune response genes. *Science*. 2013;341:789–792. doi: 10.1126/science.1240925.
26. Qureshi IA, Mattick JS, Mehler MF. Long non-coding RNAs in nervous system function and disease. *Brain Res*. 2010;1338:20–35. doi: 10.1016/j.brainres.2010.03.110.
27. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell*. 2010;39:925–938. doi: 10.1016/j.molcel.2010.08.011.
28. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*. 2009;458:223–227. doi: 10.1038/nature07672.
29. Beltran M, Puig I, Peña C, García JM, Alvarez AB, Peña R, et al. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev*. 2008;22:756–769. doi: 10.1101/gad.455708.
30. Whitehead J, Pandey GK, Kanduri C. Regulation of the mammalian epigenome by long noncoding RNAs. *Biochim Biophys Acta*. 2009;1790:936–947. doi: 10.1016/j.bbagen.2008.10.007.
31. Morán I, Akerman I, van de Bunt M, Xie R, Benazra M, Nammo T, et al. Human β cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. *Cell Metab*. 2012;16:435–448. doi: 10.1016/j.cmet.2012.08.010.
32. Rapicavoli NA, Qu K, Zhang J, Mikhail M, Laberge RM, Chang HY. A mammalian pseudogene lncRNA at the interface of inflammation and anti-inflammatory therapeutics. *Elife*. 2013;2:e00762. doi: 10.7554/eLife.00762.
33. Nelson BR, Makarewich CA, Anderson DM, Winders BR, Troupes CD, Wu F, et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science*. 2016;351:271–275. doi: 10.1126/science.aad4076.
34. Anderson DM, Anderson KM, Chang CL, Makarewich CA, Nelson BR, McAnally JR, et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell*. 2015;160:595–606. doi: 10.1016/j.cell.2015.01.009.
35. Wang Z, Wang Y. Dawn of the Epi-LncRNAs: new path from Myheart. *Circ Res*. 2015;116:235–236. doi: 10.1161/CIRCRESAHA.114.305490.
36. Sauvageau M, Goff LA, Lodato S, Bonev B, Groff AF, Gerhardinger C, et al. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *Elife*. 2013;2:e01749. doi: 10.7554/eLife.01749.
37. Klattenhoff CA, Scheuermann JC, Surface LE, Bradley RK, Fields PA, Steinhauser ML, et al. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell*. 2013;152:570–583. doi: 10.1016/j.cell.2013.01.003.
38. Xue Z, Hennelly S, Doyle B, Gulati AA, Novikova IV, Sanbonmatsu KY, et al. A G-rich motif in the lncRNA braveheart interacts with a zinc-finger transcription factor to specify the cardiovascular lineage. *Mol Cell*. 2016;64:37–50. doi: 10.1016/j.molcel.2016.08.010.
39. Wang Z, Zhang XJ, Ji YX, Zhang P, Deng KQ, Gong J, et al. The long noncoding RNA Chaer defines an epigenetic checkpoint in cardiac hypertrophy. *Nat Med*. 2016;22:1131–1139. doi: 10.1038/nm.4179.
40. Bell RD, Long X, Lin M, Bergmann JH, Nanda V, Cowan SL, et al. Identification and initial functional characterization of a human vascular cell-enriched long noncoding RNA. *Arterioscler Thromb Vasc Biol*. 2014;34:1249–1259. doi: 10.1161/ATVBAHA.114.303240.
41. Michalik KM, You X, Manavski Y, Doddaballapur A, Zörnig M, Braun T, et al. Long noncoding RNA MALAT1 regulates endothelial cell function and vessel growth. *Circ Res*. 2014;114:1389–1397. doi: 10.1161/CIRCRESAHA.114.303265.
42. Peters T, Hermans-Beijnsberger S, Beqqali A, Bitsch N, Nakagawa S, Prasanth KV, et al. Long non-coding RNA Malat-1 is dispensable during pressure overload-induced cardiac remodeling and failure in mice. *PLoS One*. 2016;11:e0150236. doi: 10.1371/journal.pone.0150236.
43. Zhang B, Arun G, Mao YS, Lazar Z, Hung G, Bhattacharjee G, et al. The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Rep*. 2012;2:111–123. doi: 10.1016/j.celrep.2012.06.003.
44. Ohnishi Y, Tanaka T, Yamada R, Suematsu K, Minami M, Fujii K, et al. Identification of 187 single nucleotide polymorphisms (SNPs) among 41 candidate genes for ischemic heart disease in the Japanese population. *Hum Genet*. 2000;106:288–292.
45. Ishii N, Ozaki K, Sato H, Mizuno H, Saito S, Takahashi A, et al. Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction. *J Hum Genet*. 2006;51:1087–1099. doi: 10.1007/s10038-006-0070-9.
46. Yan B, Yao J, Liu JY, Li XM, Wang XQ, Li YJ, et al. lncRNA-MIAT regulates microvascular dysfunction by functioning as a competing endogenous RNA. *Circ Res*. 2015;116:1143–1156. doi: 10.1161/CIRCRESAHA.116.305510.
47. Eicher JD, Wakabayashi Y, Vitseva O, Esa N, Yang Y, Zhu J, et al. Characterization of the platelet transcriptome by RNA sequencing in patients with acute myocardial infarction. *Platelets*. 2016;27:230–239. doi: 10.3109/09537104.2015.1083543.
48. Vausort M, Wagner DR, Devaux Y. Long noncoding RNAs in patients with acute myocardial infarction. *Circ Res*. 2014;115:668–677. doi: 10.1161/CIRCRESAHA.115.303836.
49. Holdt LM, Beutner F, Scholz M, Gielen S, Gäbel G, Bergert H, et al. ANRIL expression is associated with atherosclerosis risk at chromosome 9p21. *Arterioscler Thromb Vasc Biol*. 2010;30:620–627. doi: 10.1161/ATVBAHA.109.196832.
50. Holdt LM, Hoffmann S, Sass K, Langenberger D, Scholz M, Krohn K, et al. Alu elements in ANRIL non-coding RNA at chromosome 9p21 modulate atherogenic cell functions through trans-regulation of gene networks. *PLoS Genet*. 2013;9:e1003588. doi: 10.1371/journal.pgen.1003588.
51. Tsai PC, Liao YC, Lin TH, Hsi E, Yang YH, Juo SH. Additive effect of ANRIL and BRAP polymorphisms on ankle-brachial index in a Taiwanese population. *Circ J*. 2012;76:446–452.
52. Holdt LM, Stahringer A, Sass K, Pichler G, Kulak NA, Wilfert W, et al. Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. *Nat Commun*. 2016;7:12429. doi: 10.1038/ncomms12429.
53. Liu Y, Ferguson JF, Xue C, Ballantyne RL, Silverman IM, Gosai SJ, et al. Tissue-specific RNA-Seq in human evoked inflammation identifies blood and adipose LincRNA signatures of cardiometabolic diseases. *Arterioscler Thromb Vasc Biol*. 2014;34:902–912. doi: 10.1161/ATVBAHA.113.303123.
54. Gao W, Zhu M, Wang H, Zhao S, Zhao D, Yang Y, et al. Association of polymorphisms in long non-coding RNA H19 with coronary artery disease risk in a Chinese population. *Mutat Res*. 2015;772:15–22. doi: 10.1016/j.mrfmmm.2014.12.009.
55. Yen ZC, Meyer IM, Karalic S, Brown CJ. A cross-species comparison of X-chromosome inactivation in Eutheria. *Genomics*. 2007;90:453–463. doi: 10.1016/j.ygeno.2007.07.002.
56. Zhang X, Rice K, Wang Y, Chen W, Zhong Y, Nakayama Y, et al. Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: isoform structure, expression, and functions. *Endocrinology*. 2010;151:939–947. doi: 10.1210/en.2009-0657.
57. Shen LX, Basilion JP, Stanton VP Jr. Single-nucleotide polymorphisms can cause different structural folds of mRNA. *Proc Natl Acad Sci USA*. 1999;96:7871–7876.

58. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA websuite. *Nucleic Acids Res*. 2008;36(web server issue):W70–W74. doi: 10.1093/nar/gkn188.
59. Reuter JS, Mathews DH. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*. 2010;11:129. doi: 10.1186/1471-2105-11-129.
60. Ding Y, Chan CY, Lawrence CE. Sfold web server for statistical folding and rational design of nucleic acids. *Nucleic Acids Res*. 2004;32(web server issue):W135–W141. doi: 10.1093/nar/gkh449.
61. Waldispühl J, Devadas S, Berger B, Clote P. RNAmutants: a web server to explore the mutational landscape of RNA secondary structures. *Nucleic Acids Res*. 2009;37(web server issue):W281–W286. doi: 10.1093/nar/gkp477.
62. Gong J, Liu W, Zhang J, Miao X, Guo AY. lncRNASNP: a database of SNPs in lncRNAs and their potential functions in human and mouse. *Nucleic Acids Res*. 2015;43(database issue):D181–D186. doi: 10.1093/nar/gku1000.
63. Rouskin S, Zubradt M, Washietl S, Kellis M, Weissman JS. Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature*. 2014;505:701–705. doi: 10.1038/nature12894.
64. Deigan KE, Li TW, Mathews DH, Weeks KM. Accurate SHAPE-directed RNA structure determination. *Proc Natl Acad Sci USA*. 2009;106:97–102. doi: 10.1073/pnas.0806929106.
65. Wilkinson KA, Vasa SM, Deigan KE, Mortimer SA, Giddings MC, Weeks KM. Influence of nucleotide identity on ribose 2'-hydroxyl reactivity in RNA. *RNA*. 2009;15:1314–1321. doi: 10.1261/rna.1536209.
66. Castellanos-Rubio A, Fernandez-Jimenez N, Kratchmarov R, Luo X, Bhagat G, Green PH, et al. A long noncoding RNA associated with susceptibility to celiac disease. *Science*. 2016;352:91–95. doi: 10.1126/science.aad0467.
67. Liu F, Somarowthu S, Pyle AM. Visualizing the secondary and tertiary architectural domains of lncRNA RepA. *Nat Chem Biol*. 2017;13:282–289. doi: 10.1038/nchembio.2272.
68. Sripakdeevong P, Kladwang W, Das R. An enumerative stepwise ansatz enables atomic-accuracy RNA loop modeling. *Proc Natl Acad Sci USA*. 2011;108:20573–20578. doi: 10.1073/pnas.1106516108.
69. Kladwang W, VanLang CC, Cordero P, Das R. A two-dimensional mutate-and-map strategy for non-coding RNA structure. *Nat Chem*. 2011;3:954–962. doi: 10.1038/nchem.1176.
70. Flynn RA, Zhang QC, Spitale RC, Lee B, Mumbach MR, Chang HY. Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE. *Nat Protoc*. 2016;11:273–290. doi: 10.1038/nprot.2016.011.
71. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. *Science*. 2002;295:1306–1311. doi: 10.1126/science.1067799.
72. Denker A, de Laat W. The second decade of 3C technologies: detailed insights into nuclear organization. *Genes Dev*. 2016;30:1357–1382. doi: 10.1101/gad.281964.116.
73. Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet*. 2006;38:1348–1354. doi: 10.1038/ng1896.
74. Belton JM, Dekker J. Chromosome conformation capture carbon copy (5C) in budding yeast. *Cold Spring Harb Protoc*. 2015;2015:593–598. doi: 10.1101/pdb.prot085191.
75. Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, et al. Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res*. 2006;16:1299–1309. doi: 10.1101/gr.5571506.
76. Jäger R, Migliorini G, Henrion M, Kandaswamy R, Speedy HE, Heindl A, et al. Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci. *Nat Commun*. 2015;6:6178. doi: 10.1038/ncomms7178.
77. Dekker J, Marti-Renom MA, Mirny LA. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet*. 2013;14:390–403. doi: 10.1038/nrg3454.
78. Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol Cell*. 2011;44:667–678. doi: 10.1016/j.molcel.2011.08.027.
79. Engreitz J, Lander ES, Guttman M. RNA antisense purification (RAP) for mapping RNA interactions with chromatin. *Methods Mol Biol*. 2015;1262:183–197. doi: 10.1007/978-1-4939-2253-6_11.
80. Atianand MK, Hu W, Satpathy AT, Shen Y, Ricci EP, Alvarez-Dominguez JR, et al. A long noncoding RNA lincRNA-EPS acts as a transcriptional brake to restrain inflammation. *Cell*. 2016;165:1672–1685. doi: 10.1016/j.cell.2016.05.075.
81. Xing Z, Lin C, Yang L. LncRNA pulldown combined with mass spectrometry to identify the novel LncRNA-associated proteins. *Methods Mol Biol*. 2016;1402:1–9. doi: 10.1007/978-1-4939-3378-5_1.
82. Zhao J, Ohsumi TK, Kung JT, Ogawa Y, Grau DJ, Sarma K, et al. Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol Cell*. 2010;40:939–953. doi: 10.1016/j.molcel.2010.12.011.
83. Zhang C, Darnell RB. Mapping *in vivo* protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nat Biotechnol*. 2011;29:607–614. doi: 10.1038/nbt.1873.
84. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*. 2010;141:129–141. doi: 10.1016/j.cell.2010.03.009.
85. Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*. 2015;517:583–588. doi: 10.1038/nature14136.
86. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339:819–823. doi: 10.1126/science.1231143.
87. Yin H, Xue W, Chen S, Bogorad RL, Benedetti E, Grompe M, et al. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat Biotechnol*. 2014;32:551–553. doi: 10.1038/nbt.2884.
88. Long C, Amoasii L, Mireault AA, McAnally JR, Li H, Sanchez-Ortiz E, et al. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science*. 2016;351:400–403.
89. Nelson CE, Hakim CH, Ousterout DG, Thakore PI, Moreb EA, Castellanos Rivera RM, et al. *In vivo* genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science*. 2016;351:403–407. doi: 10.1126/science.aad5143.
90. Tabebordbar M, Zhu K, Cheng JK, Chew WL, Widrick JJ, Yan WX, et al. *In vivo* gene editing in dystrophic mouse muscle and muscle stem cells. *Science*. 2016;351:407–411. doi: 10.1126/science.aad5177.
91. DeWitt MA, Magis W, Bray NL, Wang T, Berman JR, Urbinati F, et al. Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. *Sci Transl Med*. 2016;8:360ra134. doi: 10.1126/scitranslmed.aaf9336.
92. Hultquist JF, Schumann K, Woo JM, Manganaro L, McGregor MJ, Doudna J, et al. A Cas9 ribonucleoprotein platform for functional genetic studies of HIV-host interactions in primary human T cells. *Cell Rep*. 2016;17:1438–1452. doi: 10.1016/j.celrep.2016.09.080.
93. Goyal A, Myacheva K, Groß M, Klingenberg M, Duran Arqué B, Diederichs S. Challenges of CRISPR/Cas9 applications for long noncoding RNA genes. *Nucleic Acids Res*. 2017;45:e12. doi: 10.1093/nar/gkw883.
94. Liu SJ, Horlbeck MA, Cho SW, Birk HS, Malatesta M, He D, et al. Crispri-based genome-scale identification of functional long noncoding rna loci in human cells. *Science*. 2017;355. pii: aah7111. doi: Y

KEY WORDS: base sequence ■ chromatin ■ genome, human ■ mutation ■ RNA, long noncoding

Circulation
Genomic and Precision Medicine

Cardiovascular Disease and Long Noncoding RNAs: Tools for Unraveling the Mystery Lnc-ing RNA and Phenotype

Elizabeth J. Hennessy

Circ Cardiovasc Genet. 2017;10:e001556
doi: 10.1161/CIRCGENETICS.117.001556
Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 1942-325X. Online ISSN: 1942-3268