De-repressing LncRNA-Targeted Genes to Upregulate Gene Expression: Focus on Small Molecule Therapeutics

Roya Pedram Fatemi1,2, Dmitry Velmeshev3,4 and Mohammad Ali Faghihi2,3

Non-protein coding RNAs (ncRNAs) make up the overwhelming majority of transcripts in the genome and have recently gained attention for their complex regulatory role in cells, including the regulation of protein-coding genes. Furthermore, ncRNAs play an important role in normal development and their expression levels are dysregulated in several diseases. Recently, several long noncoding RNAs (lncRNAs) have been shown to alter the epigenetic status of genomic loci and suppress the expression of target genes. This review will present examples of such a mechanism and focus on the potential to target lncRNAs for achieving therapeutic gene upregulation by de-repressing genes that are epigenetically silenced in various diseases. Finally, the potential to target lncRNAs, through their interactions with epigenetic enzymes, using various tools, such as small molecules, viral vectors and antisense oligonucleotides, will be discussed. We suggest that small molecule modulators of a novel class of drug targets, lncRNA-protein interactions, have great potential to treat some cancers, cardiovascular disease, and neurological disorders.

Molecular Therapy—Nucleic Acids (2014) 3, e196; doi:10.1038/mtna.2014.45; published online 18 November 2014

Insights Into The Complexity of The Human Genome

The Human Genome Project and subsequent large-scale transcriptomics efforts, notably: Functional Annotation of the Mouse and Encyclopedia of DNA Elements have overturned the central dogma of molecular biology and helped to make sense of the “transcriptional noise” that was thought to comprise the majority of the human genome.1–5 These studies have shown that the genome is pervasively transcribed and that many genes are transcribed from both sense and antisense strands of DNA, resulting in protein products, non-protein coding transcripts, or combinations of the two depending on the gene locus.1,2,5,6 Recently, we have come to appreciate that these non-coding RNAs (ncRNAs) are functionally important3,7–9 and that unlike the size of the genome coding for proteins, the size of non-protein coding genome scales with organismal complexity.3,10–13 Non-coding RNAs are divided into two broad classes based on length: small ncRNAs (<200 nucleotides) and long non-coding RNAs (>200 nucleotides) that can span up to hundreds of kilobases in length.7,14 Although much is known about the mechanisms and functions of small noncoding RNAs (e.g., microRNAs and small interfering RNAs (siRNAs)),15–20 the biological functions of long noncoding RNAs (lncRNAs) are still being elucidated.

Although the field of IncRNA research has gained increasing momentum in the past decade, many questions about the in vivo function of IncRNAs remain unanswered. This is in part because of the challenge posed by the primate- or human-specific nature of a majority of these transcripts. Some IncRNAs are conserved in mice and have been studied in vivo in knockout models, with variable results ranging from changes in the expression of a limited repertoire of genes lacking phenotypic defects21 to severe developmental abnormalities22,23 and embryonic lethality.24 For example, the IncRNA MALAT1 has been shown in vitro to regulate splicing,24,25 but interestingly, Malat1 knockout mice are viable and do not have any reported defects in vivo.21,26,27 However, targeted deletion of the IncRNA Hotair results in developmental defects and mimics findings from in vitro studies. Furthermore, a recent study of 18 IncRNA knockout strains revealed embryonic lethality in 3 mutant strains (Fendrr−/−, Peril−/−, and Magf−/−) and distinct growth defects for two other strains (linc-Brm1b−/− and linc-Pint−/−).28 This evidence suggests that IncRNAs are a heterogeneous class of transcripts, many of which have important biological functions ranging from fine-tuning the expression of a limited number of genes to acting as master regulators of important cellular processes.

Emergence of Long Non-Coding RNAs as Important Regulatory Units Within the Cell

lncRNAs have been separated into several broad classes in terms of their mechanisms of regulation of mRNA transcription and translation: decoys, regulators of translation, enhancers and modular scaffolds that guide chromatin modifying enzymes to specific genomic loci (Figure 1). lncRNAs can regulate apoptosis and cell cycle; the IncRNA Growth arrest-specific 5 (GAS5) is induced during growth factor starvation and binds the glucocorticoid receptor through its hairpin domain, acting as a decoy to physically block the transcription of metabolic genes.26,27 The IncRNA metastasis-associated lung adenocarcinoma transcript 1 (MALAT-1) localizes to nuclear speckles to associate with the serine/arginine-rich (SR) proteins involved in splicing.24,25 Another important function of IncRNAs is to regulate translation, e.g.,
beta-site APP cleaving enzyme 1-antisense (BACE1-AS) masks miRNA-binding sites on its cognate mRNA BACE1 to form a stable duplex and prevent mRNA degradation. LncRNAs can also enhance the expression of neighboring genes, acting as positive regulators of gene expression as is the case with ncRNAs arising from the SCL/TAL1, SNAI1, and SNAI2 loci. Finally, IncRNAs play an important role in epigenetic regulation, functioning as modular scaffolds to bring protein complexes into proximity to position epigenetic enzymes to specific genomic loci, resulting in changes in gene expression in cis (at the same gene locus) or in trans (at different loci). The most extensively studied IncRNA is Xist, a roughly 17 kb mammalian-specific IncRNA that regulates X-chromosome inactivation (through an epigenetic mechanism that involves interactions with polycomb proteins), by silencing one of the two X-chromosomes in females. LncRNAs are also involved in imprinting; Kcnq1ot1 and Air act as scaffolds to recruit chromatin-modifying enzymes. The 91-kb IncRNA Kcnq1t1 interacts with two chromatin modifying enzymes: G9a and Polycomb repressive complex-2 (PRC2) during imprinting to produce lineage-specific transcriptional silencing patterns. The Air IncRNA silences gene expression at the Igf2r/Slc22a2/Slc22a3 locus of paternal alleles by targeting G9a to chromatin.

Long Non-Coding RNAs as Epigenetic Regulators

IncRNAs are expressed at lower levels than protein-coding genes and are mostly localized in the chromatin fraction of the nucleus, where they interact with epigenetic enzymes. IncRNAs exhibit tissue-specific expression patterns (with a large proportion expressed in the testes and brain) because of the tight regulation of promoters that regulate their transcription. The tight transcriptional regulation and tissue-specificity of IncRNAs, their presence in low-copy number and localization to chromatin, support the notion that IncRNAs are potent regulators of gene accessibility by modulating chromatin structure at developmental-specific timepoints in the cell. Chromatin-associated IncRNAs exert their regulatory effects by acting as modular
scaffolds, conferring specificity to ubiquitously expressed epigenetic enzymes, by guiding these enzymes to specific genomic loci. Therefore, IncRNAs play a pivotal role in fine-tuning gene expression at the chromatin level via their interactions with chromatin-modifying enzymes. It is becoming increasingly evident that these transcripts play an important role in development as well as disease through their interactions with epigenetic effectors (Table 1).

### Long ncRNAs and Cancer: De-Repressing Tumor Suppressor Genes

Hox antisense intergenic RNA, *HOTAIR*, is a 2.2 kb IncRNA transcribed from the *HOXD* locus that negatively regulates the expression of *HOXD* genes in *trans*. *HOTAIR* interacts with the PRC2 complex to silence genes by depositing repressive histone 3 lysine 27 trimethyl marks (H3K27me3) on the PRC2 complex to silence genes by depositing repressive enzymes, by guiding these enzymes to specific genomic loci. Therefore, IncRNAs play a pivotal role in fine-tuning gene expression at the chromatin level via their interactions with chromatin-modifying enzymes. It is becoming increasingly evident that these transcripts play an important role in development as well as disease through their interactions with epigenetic effectors (Table 1).

### Table 1: The various mechanisms through which IncRNAs function during disease

| Transcript | Target | Mechanism | Disease Relevance | Reference |
|------------|--------|-----------|-------------------|-----------|
| GAS5       | Glucocorticoid receptor (GR) | Decoy | Cancer, Autoimmune disease | 28,29 |
| APOA1-AS   | APOA1  | Scaffold, Guide | Cardiovascular disease | 83 |
| BDNF-AS    | BDNF   | Scaffold, Guide | Neurodegenerative disorders | 38 |
| DLEU1, DLEU2 | 13q14.3 gene locus | Scaffold, Guide | Leukemia | 73 |
| H19        | E-cadherin | Scaffold, Guide | Cancers | 75–77 |
| HOTAIR     | HOXD 8-11, 13 | Scaffold, Guide | Cancers | 56–68 |
| KCNO1/OT1  | KCNO1  | Scaffold, Guide | Beckwith-Wiedemann Syndrome, adrenal neuroblastoma | 43,44 |
| ANRIL      | Ink4b/ARF/INK4a | Scaffold, Guide | Cancers | 69,70 |
| PCAT-1     | BRCA2, CENPF, CENPE | Scaffold, Guide | Prostate cancer | 71 |
| PINT       | p53 pathway genes | Scaffold, Guide | Cancers | 78 |
| PTENpg1-α isoform | PTEN | Scaffold, Guide | Cancers | 74 |
| TUG1       | HOXB7  | Scaffold, Guide | Cholangiocarcinoma | 79 |

LncRNAs are involved in the epigenetic regulation of genes and have been implicated in a number of diseases.
IncRNAs that bring chromatin modifying enzymes into contact with specific loci, could be key in developing anticancer therapies. More specifically, targeting the IncRNA or disrupting the IncRNA-protein (HOTAIR-PRC2, ANRIL-CBX7, PCAT-1-PRC2, PTENpg-1c isoform-PRC2, H19-EZH2) interactions, could de-repress cancer-suppressing genes, in a highly specific fashion by altering chromatin structure to promote senescence and slow cancer progression. Furthermore, upregulating TUG1 and PINT IncRNAs that interact with PRC2 to silence oncogenic genes can also be therapeutically valuable.

Although the examples of IncRNAs presented have been shown to exert their effects primarily by interacting with histone modifying complexes, it is important to note that IncRNAs can also interact with DNA methyltransferases directly to regulate DNA methylation. In doing so, IncRNAs have another, perhaps more significant role, that of regulating long-term epigenetic inheritance via their interactions with de novo (DNMT3a and DNMT3b) and maintenance (DNMT1) DNA methyltransferases.

**LncRNAs and Cardiovascular Disease: Upregulating Therapeutic Genes**

Apolipoprotein A-1 (APOA-1) is the major protein component of high-density lipoprotein (HDL) in plasma where it plays an important role in cholesterol efflux. Elevated HDL levels have been shown to protect against coronary heart disease, making APOA1 a strong candidate for pharmacological intervention for patients at risk for heart disease. Analysis of the APOA1 gene locus resulted in the identification of a noncoding antisense transcript, APOA1-AS. The APOA1-AS transcript regulates APOA1 in cis, purportedly through an epigenetic mechanism. Knockdown of APOA1-AS upregulates APOA1, increasing activating histone 3 lysine 4 trimethylation (H3K4me3) marks and reduces repressive H3K27 trimethylation spanning the promoter and enhancer regions of the APOA1 gene. The reduction in H3K27me3 suppressive marks correlates with a reduction in the SUZ12 protein, a core component of the PRC2 complex. Importantly, knockdown of APOA1-AS in vivo results in a significant increase in endogenous APOA1 protein. Therefore, APOA1-AS can be used to directly modulate APOA1 protein, making it a potential therapeutic target for cardiovascular disease.

The emerging use of stem cell therapy for the treatment of patients with cardiovascular disease seems to hold great promise. However, in order for this regenerative therapy to be successful, it is critical to understand and re-create the processes that result in the differentiation of mature cardiomyocytes from pluripotent stem cells. A novel IncRNA, Braveheart (Bvht), was recently identified to be important in specifying lineage commitment in mice via a PRC2-dependent mechanism. Bvht interacts with SUZ12 during multiple stages of cardiomyocyte differentiation. This work suggests that mediating the interaction between Bvht and PRC2 might be a means to promote differentiation of stem cells into cardiomyocytes. Understanding the function and mechanisms of IncRNAs in stem cell differentiation might aid in the production of more efficacious stem cell therapies. Furthermore, strengthening the interaction of Bvht with PRC2 might be a means by which to promote cardiomyocyte differentiation in stem cells.

LncRNAs APOA1-AS and human mimics of Bvht are both potential therapeutic targets for the specific upregulation of therapeutic genes in the heart. Blocking interactions between APOA1-AS and epigenetic enzymes might de-repress the APOA1 gene to upregulate apolipoprotein A1. In the case of Bvht, strengthening the interaction of this lncRNA with PRC2 could increase cardiomyocyte formation by de-repressing differentiation in pluripotent stem cells.

**LncRNAs and Neurological Disorders: Upregulating Neuroprotective Proteins**

Neurotrophins have been shown to play an important role in the maintenance of neuronal survival, outgrowth, and differentiation. Brain-derived neurotrophic factor (BDNF), one of the most abundant growth factors in the brain, is a protein that is critical for the development, survival and maintenance of neurons in the nervous system. Furthermore, BDNF is a therapeutic target for several neurodegenerative and neurodevelopmental disorders in which BDNF is severely downregulated. The BDNF locus gives rise to a 2.2 kb antisense transcript, called BDNF-AS, that discordantly regulates BDNF mRNA and protein levels in cis. Blocking or knocking down BDNF-AS correlates with a reduction in H3K27me3 at the BDNF promoter region and a reduction of enhancer of zeste 2 (EZH2, catalytic subunit of the PRC2 complex) occupancy, suggesting that BDNF-AS exerts its repressive effect on the BDNF locus by interacting with PRC2 to silence BDNF expression in cis. Therefore, BDNF-AS or its site of interaction with PRC2 could be a viable therapeutic target to upregulate BDNF protein expression. This mechanism of specific gene upregulation might be extended to other neuroprotective genes, e.g., glial-derived neurotrophic factor (GDNF), and ephrin B2 receptor (EPHB2), which were also shown to be discordantly regulated by their cognate antisense transcripts, opening the door for further study of the epigenetic regulation of these potentially therapeutic CNS genes.

An analysis of genes both enriched in the brain and implicated in Autism revealed that 40% of “disease” genes had overlapping antisense transcription that exhibited region-specificity in the brain of Autism spectrum disorder (ASD) patients. An antisense transcript to SYNGAP1, called SYN-GAP1-AS, was differentially expressed in the brain regions of patients with ASD compared to non-ASD control subjects. SYNGAP1 is a protein that is important during the maturation of learning, memory and cognition and its dysregulation has been implicated in ASD. Furthermore, SYNGAP1-AS is found primarily in the chromatin compartment of the cell and is discordantly expressed with respect to SYNGAP1 in autism patients. Although a mechanism for SYNGAP1-AS has not been explored, the data suggest that SYNGAP1-AS, due to its localization to chromatin, may exert its effects on SYNGAP1 through an epigenetic mechanism that could be of therapeutic value to ASD patients.
Gene repression by IncRNAs is prevalent in the brain and occurs at several loci that encode for proteins involved in nervous system development and maintenance as well as synapse formation and maturation. Since these IncRNAs are potent regulators of transcription, inhibition of these repressive transcripts could result in locus specific upregulation of therapeutic genes. Importantly, one of the first examples of in vivo endogenous gene upregulation was described by blocking mouse Bdnf-AS. The potential to target IncRNAs and produce changes in neurotrophin and synaptic proteins has major implications in treating neurodegenerative and neurodevelopmental diseases where BDNF protein is dysregulated, e.g., Alzheimer’s Disease, Rett’s Syndrome, and Parkinson’s Disease. These disorders all exhibit marked reductions in BDNF and can potentially be protected against or in part rescued by the de-repression and upregulation of endogenous BDNF protein.

Designing LncRNA-Directed Therapeutics

Despite the multiple ways that genes can be silenced, methods for upregulating the expression of therapeutic genes remains a challenge. LncRNAs represent attractive therapeutic targets due to their presence in a development and cell-type specific manner as well as their ability to target a specific subset of genes. However, it must be noted that although cis or trans regulatory mechanisms have been identified for certain lncRNAs, effects on global gene expression changes are in most cases unknown and should also be carefully considered. LncRNAs have mainly been targeted using antisense oligonucleotides (ASOs), siRNAs and viral vectors that contain shRNAs (Figure 2). Although IncRNAs represent appealing pharmacological and therapeutic targets, inhibiting IncRNAs in vivo remains a challenge.

Figure 2
Inhibition of IncRNAs to upregulate gene expression by blocking interactions with epigenetic complexes. (a) LncRNAs bind to the PRC2 complex and guide the complex to target gene promoters (red) to induce H3 lysine 27 trimethylation (H3K27me3) and repress transcription. (b) Potential approaches to inhibit the repressive action of IncRNAs. Top: Small molecules can inhibit interaction of the PRC2 complex at the IncRNA binding site to de-repress transcription. Middle: Naked or liposome-encapsulated DNA oligonucleotides can target IncRNAs to form DNA-IncRNA duplexes that recruit RNase H. Bottom: Viral particles containing IncRNA-targeting shRNAs can deliver shRNAs that generate siRNAs (targeting the RISC complex to the IncRNA) in the cell. (c) Top: Small molecule drugs prevent the formation of a functional PRC2-IncRNA complex. Middle: RNase H results in IncRNA cleavage and decay. Bottom: The RISC complex cleaves and degrades IncRNAs. These IncRNA-targeting approaches result in the dissociation of the PRC2 complex from the target gene promoter, de-repressing genomic loci to upregulate the expression of therapeutic genes.
Knockdown of disease genes, in vitro, usually involves the use of siRNAs that are unstable and non-penetrating to target cells in vivo. Modified antisense oligonucleotides are stable, nuclease-resistant, short (13–25 nucleotides), single-stranded DNA oligonucleotides that are complementary to RNAs of interest, have been used to induce RNase H-dependent cleavage of IncRNA targets. Various modified nucleotides have been tested both in vitro and in vivo and these modifications are continually undergoing enhancements in chemistry. Modified oligonucleotides targeting antisense IncRNAs, also called AntagoNATs, usually contain a mixture of 2′-O-methyl RNA and Locked Nucleic Acid (LNA) modifications at their 5′ and 3′ ends to protect against nuclease cleavage and phosphothioate backbone to enhance their cellular uptake.

Knockdown of IncRNAs in vivo has led to successful upregulation of therapeutic target genes and proteins resulting in measurable phenotypic responses. Continuous in vivo infusion of AntagoNATs targeting mouse Bdnf-AS produced a locus-specific upregulation in Bdnf mRNA and protein. Mice treated with BDNF AntagoNATs display a marked increase in neurotrophins in vivo, even though small molecule modulators of neurotrophin receptors are available. Repeated administration of AntagoNATs targeting monkey ApoA1-AS resulted in the upregulation of HDL levels. Nuclear-localized lncRNA knockdown of lncRNAs targeting MALAT1 was successfully targeted for downregulation in skeletal muscle with the systemic administration of a gapmer, that might have been deleterious to RNAs of interest, have been used multiple times in vivo. Modified antisense oligonucleotides and RNAi therapies targeting repressive IncRNAs are promising but require additional developments in oligonucleotide chemistry and use of vehicles for safe and effective delivery.

Small Molecule Modulators of lncRNA-Protein Interactions

Many chromatin-modifying enzymes have a flexible backbone and it is believed that they fold and form distinct structures only when bound to other proteins, RNA or DNA. This flexibility is believed to allow proteins to interact with several partners and form stronger, more stable interactions. Chromatin-modifying enzymes are mostly ubiquitously expressed proteins that lack rigidity in protein structure, which makes them undesirable or difficult therapeutic targets using a small molecule approach.

Chromatin-modifying enzymes that interact with IncRNAs are being studied for their ability to specifically bind RNA. Although some reports suggest that in the case of PRC2, the histone methyltransferase EZH2 is a promiscuous RNA binding protein, other studies have shown that the EZH2 component of the PRC2 complex does exhibit RNA-binding specificity when studying well-known PRC2-interacting IncRNAs Xist and HOTAIR as compared to control transcripts. It is important to mention that novel methods need to be developed in order to reliably identify RNA-protein interactions. Current methodologies have serious limitations, for example, in immunoprecipitation experiments, a protein of interest might interact with many RNAs, however due to methods of crosslinking, might grossly over or underestimate RNA protein interactions. Furthermore, RNA-centered approaches, using a labeled RNA of interest to pull down interacting proteins followed by mass spectrometry can detect abundant RNA binding proteins which might be experimental artifacts as only a few of these RNAs have been purified. Additionally, the reported non-specific RNA binding to EZH2 might be an artifact due to technical limitations in studying RNA-protein interactions (particularly using EMSA alone) or could be a general feature of RNA-binding proteins. We suggest that better in vitro assays are needed to test IncRNA-protein interactions and that these studies should incorporate valid RNA and protein controls, while taking the RNA-binding affinity of proteins of interest in the assay design.

Chromatin-modifying enzymes usually need IncRNAs to form complexes with other effector proteins and to be targeted to specific genomic loci. Therefore, one viable approach would be to target interactions between IncRNAs...
and proteins with small molecules, in order to increase specificity and to overcome problems with tissue and cell penetration. We suggest the use of small molecules as an alternative therapeutic approach and are actively pursuing this by designing measurable pharmacological assays to quantitate lncRNA-protein interactions (Fatemi RP, Salah-Uddin H, Faghhi MA, manuscript in preparation). Targeting lncRNA-protein interactions, instead of ubiquitously expressed effector proteins, is a superior approach due to cell, tissue and developmental time point specificity of lncRNA expression. To design small molecules targeting RNA, it is necessary to understand higher order RNA structures. Although lncRNAs lack sequence conservation, they are structurally conserved, making them more stable and predictable physical targets, with their promoters conserved in a manner similar to that of protein-coding genes. Although the “druggability” of RNAs with small molecules has been studied with some success, a more detailed understanding of RNA-ligand interactions, additional RNA targets and active compounds that target RNAs are needed. The most widely studied RNA targets for small molecule therapy include the bacterial 16S rRNA A site, HIV-1 Rev Response element (RRE) and HIV-1 trans-activating region (TAR) RNAs. The most successful RNA-targeting approach to date has been the use of aminoglycosides to target the bacterial 16S rRNA, however these compounds are limited to in vitro use. Netilmicin, an aminoglycoside derivative, has been used to target TAR and has been effective in vitro; however, due to poor absorption, can only be injected for use as an antibiotic in extreme cases.

Therefore, targeting lncRNAs at their site of interaction with chromatin modifying enzymes is a well-defined alternative approach as compared to targeting either proteins or lncRNAs alone. Targeting lncRNA-protein interactions would increase the specificity of compounds, reduce off-target effects, and provide reversible inhibition of chromatin modifying enzymes at non-catalytic domains of the protein.

Considering that lncRNAs and their epigenetic effectors are implicated in gene regulation in a wide variety of diseases, it would stand to reason that targeting lncRNAs at their site of interaction with epigenetic modifying enzymes would alter gene expression. In targeting the interaction between lncRNAs and enzymes of interest, it is possible to modify gene expression without altering the protein’s other effects in the cell, e.g., the epigenetic enzymes JARID2 and EZH2 have distinct sites at which they bind lncRNAs. A great degree of specificity can be achieved when targeting lncRNAs that act in cis compared to more widespread effects in trans-acting lncRNAs, however one important caveat to this is that global changes in gene expression upon lncRNA modulation must first be characterized, since lncRNAs might act both in cis and in trans and might even have cytoplasmic effects. Furthermore, tissue and cell-type as well as developmental-specific lncRNA expression patterns add another level of specificity to this approach. However, in order to target this specific interaction, it is crucial to understand the biochemistry of how RNAs and proteins interact.

Small molecules are used to treat various diseases as these compounds have greater cellular uptake and less administration challenges than antisense oligonucleotides or viral vectors for the delivery of RNAs. In order to use small molecule drugs to target lncRNA-protein interactions, pharmacological rules for rational drug discovery must be applied. Applying the rules for rational drug design necessitates two main criteria: (i) to have a clearly defined target whose modulation is of therapeutic value and (ii) that the target is in fact druggable. We have described above evidence to support the first criteria and find that the modulation of these lncRNAs is indeed of therapeutic value in several disease contexts. It is also critical to understand the drug targets themselves: RNA higher order structures, protein tertiary structure and the unique conformational changes that occur when an RNA and protein interact, i.e., an lncRNA interacts with a chromatin modifying enzyme.

The next task is to define whether these targets are receptive to small molecule modulation. To answer this question, there is a need for the development of target-based pharmacological assays that can quantify lncRNA-protein interactions for the purpose of high throughput compound screening. There are several widely used non-quantitative methods for measurement or RNA-protein interactions such as, RNA footprinting, various RNA pulldown methods (Crosslinking and immunoprecipitation (CLIP), RNA Immunoprecipitation (RIP), Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP)), co-staining by fluorescence in situ hybridization and immunofluorescence (FISH-IF), and RNA electrophoretic mobility shift assays (RNA EMSA). However, these methods are cumbersome and are not amenable to compound screening. Recently, two new methods for the quantitative analysis of RNA-protein interactions have been introduced, RNA on a massively parallel array (RNA-Map) and high throughput sequencing-RNA affinity profiling (HiTS-RAP). These two methods are currently being used to characterize the RNA-protein interaction, and when targets are identified, have the potential to be modified for compound screening. By and large, a pharmacological assay that can quantitate the degree and affinity of lncRNA-protein interactions and is amenable to screening will be a tremendous asset to the field.

Because of the malleability of proteins and RNAs, the most efficient site to target lncRNAs with small molecules should ideally be at the binding cleft where the lncRNA and chromatin-modifying enzyme interact. In order to achieve this, modeling the conformation of proteins and RNAs and understanding the unique conformations these molecules undertake when they interact is key. Several studies have focused on high throughput approaches to understanding RNA structure and RNA-protein interactions on a large-scale. However, methods to assay for small molecules must also be implemented in order to make use of this structural information. Furthermore, X-ray crystallography and NMR spectroscopy should be incorporated to produce models to better understand how small molecules would both fit and bind in the lncRNA-protein binding cleft. This approach has been successfully used in modeling a single protein or protein-protein interactions, but should also be extended large-scale to lncRNA-protein interactions. It should be noted that these approaches have important limitations, e.g., the length of time and difficulty in producing crystal structures and the limitations in nucleic acid length and protein size in NMR, thereby making it difficult to produce a structure for most full-length
long noncoding RNAs.\(^{136}\) Biochemical studies are necessary to accurately model the minimum regions required to allow IncRNAs to interact with specific domains of chromatin modifying enzymes. Furthermore, accurate methods to measure the strength of these IncRNA-protein interactions are lacking in the field but are necessary to develop assays to screen for small molecule modulators.\(^ {123}\) Therefore, quantitative assay methods should be employed in conjunction with biochemical approaches and structural modeling so that high throughput compound screening can be performed and hit compounds can be optimized using accurate structure models. The use of small molecules would overcome many of the challenges observed in the above-mentioned therapies and holds tremendous potential therapeutically.

Long noncoding RNA-protein interactions represent an important class of molecular effectors within the cell. As such, they also represent a new and unexplored set of drug targets with the potential for developing highly specific targeted therapeutics that can surpass the efficacy of existing drugs used to treat diseases in various contexts. The use of small molecules that target specific IncRNA-protein interactions to finely tune gene expression and modify chromatin state can overcome many of the complications introduced by using other approaches. Furthermore, directing therapies to these specific interactions can help reduce off-target side effects that plague drugs currently on the market.

**Conclusions**

Although a wealth of transcriptomics data exists indicating the potential presence of thousands of IncRNAs, validation of the function of these IncRNAs still remains a challenge. To date, the function and mechanisms of only a few of these IncRNAs have been studied, yet their mechanisms and global effects on gene expression require further research.\(^ {30,41,56}\) However, in order to design effective therapeutics, we must further study the function of IncRNAs, elucidate their mechanisms, understand their structures, and develop approaches to quantitatively assay for drugs that modulate their interactions with proteins to ultimately de-repress genomic loci and to effectively upregulate therapeutic genes.

**Acknowledgments.** The authors thank Marco Magistri for critical reading of the manuscript. This work was done during the term of an Award from the American Heart Association Greater Southeast Affiliate to R.P.F. The US NIH NINDS R01NS081208-01A1 awarded to M.A.F supported this work. The authors declare no conflict of interest.

1. Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, TR, Margulies, EH et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447: 799–816.
2. Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakaniishi, M., Nakamura, M. et al.; RIKEN Genome Exploration Research Group; Genome Science Group (Genome Network Project Core Group); FANTOM Consortium. (2005). Antisense transcription in the mammalian transcriptome. Science 309: 1564–1566.
3. Carmi, P., Kasukawa, T., Katayama, S., Gough, J., Frith, MC, Maeda, N. et al.; FANTOM Consortium; RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group). (2005). The transcriptional landscape of the mammalian genome. Science 309: 1559–1563.
4. Bernstein, BE, Birney, E, Dunham, I, Green, ED, Gunter, C, Snyder, M et al. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature 489: 57–74.
5. Djebali, S., Davis, CA, Merkel, A, Cobin, A, Lassmann, T., Mortazavi, A et al. (2012). Landscape of transcription in human cells. Nature 489: 101–108.
6. Derrien, T., Johnson, R., Bussetti, G., Tanzer, A, Djebali, S., Tilgner, H et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 22: 1775–1789.
7. Kapranov, P., Cheng, J., Dike, S., Nix, DA, Duttagupta, R., Willingham, AT et al. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science 316: 1484–1488.
8. Derouet, F., Kapranov, P., Ucis, C., Frankish, A., Castello, R., Drekeow, J. et al. (2007). Prominent use of distal 5′ transcription start sites and discovery of a large number of additional exons in ENCODE regions. Genome Res 17: 746–759.
9. Findeiss, S., Schmitdke, C., Stadler, PF and Bonas, U. (2010). A novel family of plasmid-transferred anti-sense ncRNAs. RNA Biol 7: 120–124.
10. Lander, ES, Linton, LM, Birren, B., Nusbaum, C., Zody, MC, Baldwin, J et al.; International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. Nature 409: 860–921.
11. Waterston, RH; The International Human Genome Sequencing Consortium. (2004). Finishing the euchromatic sequence of the human genome. Nature 431: 931–945.
12. Mattick, JS (2004). RNA regulation: a new genetics? Nat Rev Genet 5: 316–323.
13. Kapranov, P. St Laurent, G., Paz, T., Ozsolak, F., Reynolds, CP, Sorensen, PH et al. (2010). The majority of total nuclear-encoded non-ribosomal RNA in a human cell is ‘dark matter’ un-annotated RNA. BMC Biol 8: 149.
14. Ambros, V. (2001). microRNAs: tiny regulators with great potential. Cell 107: 823–826.
15. Elbashir, SM, Harborth, J, Lendeckel, W, Yalcin, A, Weber, K and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 411: 494–498.
16. Kaplen, P., Parish, S., Imani, F., Fire A, and Morgan, RA (2001). Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. Proc Natl Acad Sci USA 98: 9742–9747.
17. Lee, RC and Ambros, V (2001). An extensive class of small RNAs in Caenorhabditis elegans. Science 294: 852–854.
18. Lau, NC, Lim, LP, Weinstein, EG and Bartel, DP (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science 294: 858–862.
19. Lagos-Quintana, M., Rauhut, R., Lendeckel, W and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science 294: 853–856.
20. Zhang, B., Arun, G., Mao, YS, Lazar, Z., Hugn, G, Bhattacharjee, G et al. (2012). The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. Cell Rep 2: 111–123.
21. Li, L., Liu, B., Wapinski, OL, Tsai, MC, Ov, K., Zhang, J et al. (2013). Targeted disruption of Hotair leads to homotypic transformation and gene derepression. Cell Rep 5: 3–12.
22. Sauvageau, M., Golf, LA, Ladoto, S., Bonev, B, Groff, AF, Gerhardinger, C et al. (2013). Multiple knockout mouse models reveal lncRNAs are required for life and brain development. Elife 2: e01749.
23. Tripathi, V., Ellis, JD, Shen, Z., Song, DF, Pan, Q., Watt, AT et al. (2010). The nuclear-repressed non-coding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell 39: 925–938.
24. Bernard, D., Prasanth, KV, Tripathi, V., Colaiste, S., Nakamura, T., Xuan, Z. et al. (2010). A long nuclear-repressed non-coding RNA regulates synaptogenesis by modulating gene expression. EMBO J 29: 3092–3093.
25. Nakagawa, S., Ip, JY, Shirai, G, Tripathi, V., Zong, X., Hrosse, T et al. (2012). Malat1 is not an essential component of nuclear speckles in mice. RNA 18: 1467–1499.
26. Eissmann, M., Gutschner, T., Hammelme, M., Gunther, S, Caudron-Herger, M., Groff, M et al. (2012). Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. RNA Biol 9: 1076–1087.
27. Kino, T., Hurt, DE, Ichh, T., Nader, N and Chrousos, GP (2010). Noncoding RNA-gas1 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. Sci Signal 3: ra1.
28. Pickard, MR, Mourad-Masabouni, M and Williams, GT (2013). Long non-coding RNA GASS regulates apoptosis in prostate cancer cell lines. Biochem Biophysics Acta 1832: 1613–1623.
29. Faghfhi, MA, Zhang, M., Huang, J., Modarresi, F., Van der Brug, MP, Nalis, M. et al. (2010). Evidence for natural antisense transcript-mediated inhibition of microRNA function. GenoME Biol 11: R56.
30. Faghfhi, MA, Modarresi, F., Khaili, AM, Wood, DE, Sahagan, BG, Morgan, TE et al. (2008). Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. Nat Med 14: 723–730.
31. Andersson, P., Gebhardt, C., Miguel-Escalada, I., Hool, I., Bornholdt, J., Boyd, M et al.; FANTOM Consortium. (2014). An atlas of active enhancers across human cell types and tissues. Nature 507: 455–462.
32. Örom, UA, Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Busotti, G et al. (2010). Long noncoding RNAs with enhancer-like function in human cells. Cell 143: 46–58.
33. Gutmann, M., Donahghey, J., Carey, BW, Barber, M., Greiner, JK, Munson, G et al. (2011). IncRNAs act in the circuitry controlling pluripotency and differentiation. Nature 477: 295–300.
Enhanced coding RNA HOT AIR promotes esophageal squamous cell carcinoma metastasis and poor prognosis. J Clin Oncol 2013; 31:3056–3063.

Hajeri, M., Behnemann, M., Sadeghizadeh, M. and Zenzodini, M. (2013). Up-regulation of HOT AIR long non-coding RNA in human gastric adenocarcinoma tissues. Mol Oncol 2013; 7:607–618.

Endo, H., Hirakawa, J., Nagakawa, T., Yokoyma, Y., Tamai, K., Yamamini, H. et al. (2013). Enhanced expression of long non-coding RNA HOTAIR is associated with the development of gastric cancer. PLoS ONE 8: e77070.

Fumagalli, F., Racagni, G. and Riva, M. (2010). The expanding role of BDNF: a therapeutic target for Alzheimer’s disease? Pharmacogenomics 11: 67–78.

Huang, S., Muñoz-Cabello, S., Muñoz-Cabello, A. and del Cerro, L.F. (2012). The e77070 pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human normal and non-small cell lung cancer cells. Cell Mol Life Sci 2013; 70: 463–475.

Kane, C.J., Hendrich, B., Bagni, R., He, X., Ren, Y., Yang, X. et al. (2012). The human XIST gene: structure, function, and evolution of lncRNA repertoires and expression patterns in tetrapods. J Cell Sci 2012; 125: 4249–4259.

Kobayashi, S., Doi, Y., Nakagawa, H. and Wang, Y. (2010). Long non-coding RNA HOTAIR is a prognostic marker for esophageal squamous cell carcinoma progression and survival. PLoS ONE 5: e13516.

Kuwabara, T., Nakagawa, H., Li, H., Zhao, Z., Wu, Q. and Sone, Y. (2009). The long non-coding RNA HOTAIR is overexpressed and regulates PTEN methylation in laryngeal squamous cell carcinoma. Am J Pathol 2013; 182: 64–70.

Kwon, J., Yang, J., Wu, T. and Chang, Y. (2008). Polycystic proteins targeted by a short repeat RNA to the mouse X chromosome. Science 2008; 322: 750–756.

Lombardo, L., Pogliani, A., Giaccone, G., Delgado, D., Delgado, D. and Delgado, D. (2013). The mechanism of action of BDNF secreted by human glioblastoma cells. PLoS ONE 2013; 8: e77070.
of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to RNA-bound proteins from an atlas of mammalian mRNA-binding proteins. A, Montini, A, Fischer, K, Horos, V, Wang, S, Rao, J, Kim, A, Pagano, D, Schroth, H, Saldaña-Meyer, LM, Layton, CJ, Chang, GP and Lis, JD (2014). Methods to enable the design of a new class of small molecules targeting RNA. Proc Natl Acad Sci USA 111: 36ra43.

Xiang, Z, Xu, D, Chen, L, Li, X, Yang, J, Fan, D et al. (2014). RAID: a comprehensive resource for human RNA-associated (RNA-RNA:protein) interaction. RNA 20: 899–903.

Ouyang, Z, Snyder, MP and Chang, HY (2013). SeqFold: genome-scale reconstruction of RNA secondary structure integrating high-throughput sequencing data. Genome Res 23: 377–387.

Underwood, JG, Uzilov, AV, Katzman, S, Onodera, CS, Mainzer, JE, Mathews, DH et al. (2010). FragSeq: transcriptome-wide RNA structure probing using high-throughput sequencing. Nat Methods 7: 995–1001.

Wan, Y, Xu, K, Zhang, Q, Flynn, RA, Manor, O, Ouyang, Z et al. (2014). Landscape and variation of RNA secondary structure across the human transcriptome. Nature 505: 706–709.

Novikova, IV, Henmelly, SP and Sanbonmatsu, KY (2012). Sizing up long non-coding RNAs: do ncRNAs have secondary and tertiary structure? Biochimie 2: 186–199.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/