The bright side of dark matter: lncRNAs in cancer

Joseph R. Evans,1 Felix Y. Feng,1,2,3,4 and Arul M. Chinnaiyan2,5,6,7

1Department of Radiation Oncology and 3Michigan Center for Translation Pathology, University of Michigan, Ann Arbor, Michigan, USA. 2Departments of Radiation Oncology, Urology, and Medicine, and 4Helen Diller Family Comprehensive Cancer Center, UCSF, San Francisco, California, USA. 5Comprehensive Cancer Center, 6Department of Pathology, and 7Howard Hughes Medical Institute, University of Michigan, Ann Arbor, Michigan, USA.

The traditional view of genome organization has been upended in the last decade with the discovery of vast amounts of non–protein-coding transcription. After initial concerns that this “dark matter” of the genome was transcriptional noise, it is apparent that a subset of these noncoding RNAs are functional. Long noncoding RNA (lncRNA) genes resemble protein-coding genes in several key aspects, and they have myriad molecular functions across many cellular pathways and processes, including oncogenic signaling. The number of lncRNA genes has recently been greatly expanded by our group to triple the number of protein-coding genes; therefore, lncRNAs are likely to play a role in many biological processes. Based on their large number and expression specificity in a variety of cancers, lncRNAs are likely to serve as the basis for many clinical applications in oncology.

Introduction

Completion of sequencing of the human genome led to a surprising downward revision of the number of “genes” to approximately 25,000 (1), as prior estimates based on expressed sequence tag (EST) data ranged from 45,000 to 140,000 (2). However, disappointment was soon replaced by hopeful intrigue when the comprehensive expression analysis of the ENCYclopedia of DNA Elements (ENCODE) project found that 60%–70% of the bases in the human genome could be found in transcripts, while only approximately 1.5% of the genome codes for a protein. This seeming discrepancy posed a large looming question of the nature of this noncoding “dark matter” of the genome (3).

Initial concern that this noncoding RNA was simply leaky transcription noise, as had been shown in yeast (4, 5), gave way relatively quickly, as numerous noncoding RNAs were shown to have specific functions. While a number of classes of small noncoding RNA were already known, including microRNA and small nucleolar RNA (snRNA), a new class of longer RNAs emerged that seemed to hold particular functional promise. Long noncoding RNAs (lncRNAs) were defined by length greater than 200 nt and similarities to protein-coding genes, including transcription mediated by RNA polymerase II, a 5′ cap, multiple exons, poly-adenylation, histone 3, lysine 4 tri-methylation (H3K4me3) around the transcription start site, and histone 3, lysine 36 tri-methylation (H3K36me3) throughout the transcribed gene body (6). Unbiased genome-wide searches were quickly able to identify thousands of lncRNAs (7–9), and more directed experiments are identifying lncRNAs involved in specific contexts, such as the effects of cigarette smoking (10).

In this Review, we will highlight the role of lncRNAs in cancer. As the number of lncRNAs has exploded, so too has the number of lncRNAs involved in cancer biology grown such that an exhaustive review is impossible. Therefore, we discuss lncRNA cancer biology themes as well as the translation of lncRNAs from bench research to clinical use as biomarkers and therapeutic targets. For several reasons that we discuss throughout, lncRNAs — the dark matter — now appear as a very promising class of genes for exploitation in the battle against cancer.

MiTranscriptome. Early lncRNA search efforts, described above, were in the genomics era, but they also quickly fell victim to the extraordinary rapidity of technology development and their own success. The noncoding insights from ENCODE and the studies mentioned above were performed with microarray technology (3, 7–9). It quickly became apparent from these studies that transcription was much more complex and much less discrete than traditionally thought and therefore not assayed very accurately by microarray. Next-generation sequencing (NGS) technology, on the other hand, allows for global unbiased single-transcript interrogation. Based on this, our group undertook a pilot study to identify novel prostate cancer lncRNAs using early NGS technology to profile around 102 prostate tissues and cell lines, including 20 benign adjacent prostate, 47 localized prostate tumors, and 14 metastatic tumors. We identified 121 noncoding transcripts that were dysregulated in prostate cancer (termed prostate-associated cancer transcripts [PCATs]) and were not found in any gene annotation databases. We have gone on to show that several of the PCATs play important and disparate roles in prostate cancer biology. Importantly, one of the identified transcripts, which was renamed second chromosome locus–associated prostate-1 (SChLAP1), is a powerful prognostic biomarker of metastatic progression risk after prostatectomy (11–15).

Conflict of interest: F.Y. Feng and A.M. Chinnaiyan have filed a patent on the clinical uses of the PCATs (U.S. Patent Filing 13/299,000). A.M. Chinnaiyan holds patents on the urine PCA3/TMPRSS2-ERG prostate fusion (U.S. Patent 8,743,222 and U.S. Patent 8,759,301). A.M. Chinnaiyan is a cofounder and serves on the scientific advisory boards for OncoFusion Therapeutics, MedSyn Biopharma LLC, Amune BioScience, and Esanik Therapeutics. SChLAP1 has been licensed to GenomeDx Biosciences Inc. to be developed as a biomarker of prostate cancer (A.M. Chinnaiyan is named as an inventor). F.Y. Feng has received funds for travel, accommodations, or expenses from GenomeDx Biosciences Inc. None of these companies played a role in the design and conduct of this study, in the collection, analysis, or interpretation of the data, or in the preparation, review, or approval of the article.

Based on the success of this pilot study, we undertook a much larger effort that took advantage of the vast RNA sequencing (RNA-seq) resource built by The Cancer Genome Atlas (TCGA, cancergenome.nih.gov) (16). We developed a robust bioinformatics pipeline to predict novel transcripts and genes with specific emphasis on lncRNAs that are often expressed at low levels. Rather than analyzing the sequencing data from each sample individually, which has limited sensitivity for lowly expressed transcripts, we used the sequencing information from all samples within a given cancer type coupled with a custom sequencing noise reduction algorithm. This pipeline was applied to over 7,000 sequencing samples across 18 organ systems from 25 independent studies, though 80% of these studies were from TCGA (16). This transcript discovery pipeline identified nearly three times as many lncRNA genes as protein-coding genes, and approximately 80% of these lncRNA genes did not exist in any annotated database (Figure 1). Interestingly, we found that 5.6% of lncRNAs contain conserved sequence elements from 25 independent studies, though 80% of these studies applied to over 7,000 sequencing samples across 18 organ systems with a sufficient number of normal samples for comparison. There were nearly 8,000 lncRNAs that were lineage and/or cancer specific, representing a vast tumor-specific resource for cancer biomarker and therapeutic target research (16).

of the data on these 8,000 lncRNAs is available for public use at mitranscriptome.org, which will be maintained and updated in support of cancer lncRNA research efforts.

**InncRNAs in cancer biology**

A precedent for involvement of lncRNAs in cancer biology had been set by the pregenomics era lncRNA H19 (19), and convincing evidence for the modern lncRNAs came relatively quickly with elucidation of the role of the HOX locus HOTAIR lncRNA in breast cancer. HOTAIR is overexpressed in breast tumor cells and causes genomic redistribution of the master epigenetic regulatory complex polycomb repressive complex 2 (PRC2). Consequently, in the repressive epigenetic mark histone H3 lysine 27 methylation (H3K27me) mediate increased tumor cell metastatic potential (20). This example is illustrative of an early dominant theme in lncRNA function. Numerous lncRNAs have been demonstrated to interact with and modulate epigenetic regulatory complexes, most notably PRC2 (e.g., Xist [ref. 21], H19 [ref. 22], and the INK4/ARF tumor suppressor locus lncRNA ANRIL [refs. 22, 23]), but also PRC1 (24) as well as neuronal gene repressor H3K17 methylase CoREST (25), H3K4me3 demethylase SMCX (8), and histone methyltransferase MLL1 (26, 27). While these studies are quite intriguing, there remains some skepticism about the specificity of these interactions based on unbiased studies that detected thousands of lncRNAs (up to 24% of all lncRNAs known at the time) interacting with PRC2 (8, 28), CoREST, or SMCX (8). Beyond epigenetic markers, our group showed that SChLAP1 interacts with the nucleosome positioning complex SWI/SNF through its component member SNF5 (also known as SMARCB1). This interaction results in genomic redistribution of SWI/SNF and elicits a global gene expression program that enhances metastatic potential in prostate cancer (14). Epigenetic regulation has been a dominant primary mechanism in lncRNA biology that affects many cellular processes, including oncogenic signaling. Through myriad molecular mechanisms, lncRNAs have been implicated in many classic cancer biology pathways. This list has grown so rapidly that it is impossible to describe exhaustively, and we highlight several illustrative examples in Figure 2.

**p53 pathway.** Several roles for lncRNAs have been found in the p53 pathway; interestingly, one of these novel RNA functions is found in p53 itself. TP53 mRNA can enhance p53 activation by directly binding to and inhibiting the E3 ubiquitin ligase MDM2, and this feed-forward loop for p53 activation can be mitigated by the p53 pathway; interestingly, one of these novel RNA functions is found in p53 itself. TP53 mRNA can enhance p53 activation by directly binding to and inhibiting the E3 ubiquitin ligase MDM2, and this feed-forward loop for p53 activation can be mitigated by...
Consistent with this antineoplastic effect, and apoptosis in vitro as well as modulation of autophagy (32, 33). The lncRNA p53-activated enhancers that do not have a p53-binding site (31). Consistent with this antineoplastic effect, MEG3 selectively enhances transcriptional activation dependent kinase 1A (CDKN1A), thus providing a mechanism for CDKN1A target gene that activates strong enhancers including in cyclin-ing network (30). Consistent with this, the lncRNA LED is a p53 target gene that activates strong enhancers including in cyclin-dependent kinase 1A (CDKN1A), thus providing a mechanism for p53-activated enhancers that do not have a p53-binding site (31). The lncRNA MEG3 selectively enhances transcriptional activation by p53 and downregulates MDM2, resulting in cell-cycle arrest and apoptosis in vitro as well as modulation of autophagy (32, 33). Consistent with this antineoplastic effect, MEG3 is downregul-at ed in multiple cancer types (34–37). Several lncRNAs have been identified as p53 target genes. The lncRNA loc85194 is a p53 target gene that inhibits proliferation through binding of the prop-roliferation miR-211 (38). The CDKN1A tumor-suppressor locus lncRNA p21-associated ncRNA DNA damage activated (PANDA) dampens the apoptotic response following p53 activation by inter-acting with the transcription factor NF-YA to decrease induction of proapoptotic genes (39).

Linc-p21 is also a p53 target gene and provides a mechanism for the longstanding mystery of transcriptional repression in the p53 response. Linc-p21 directly interacts with and is required for genomic localization of heterogeneous nuclear ribonucle-oprotein K (hnRNP-K), which mediates p53-associated tran-scriptional repression. RNAi-mediated knockdown of linc-p21 decreases apoptosis after doxorubicin treatment, suggesting a possible mechanism for tumor cells avoiding apoptosis after p53 activation (40). These examples demonstrate the importance of lncRNAs in the p53 response.

Hypoxia signaling and EMT. Tumor cells often utilize hypoxia signaling to maintain a proliferative response in normoxia and escape growth arrest in hypoxia. lncRNA-LET normally represses hypoxia signaling by promoting degradation of nuclear factor 90 (NF90), which is required for hypoxia signaling. Under hypoxic conditions or in cancer cells, hypoxia-inducible histone deacetyl-ase 3 (HDAC3) downregulates lncRNA-LET expression by promot-er deacetylation, thus allowing hypoxia signaling to proceed (41). Moreover, hypoxia signaling often stimulates a cellular process known as the epithelial-mesenchymal transition (EMT), which is a critical mediator of metastasis. Several lncRNAs are known to affect EMT signaling in cancer cells. For example, HOTTIP is involved in EMT by activating Wnt/β-catenin signaling, which leads to E-cadherin (CDH1) downregulation (22). Additionally, the function of key EMT regulator Snail1 has recently been found to involve an antisense lncRNA produced from the region of the first intron of the Zeb2 gene and is upregulated by Snail1 expression. Expression of the lncRNA retains a Zeb2 5’-UTR intron, which con-tains an internal ribosomal entry site (IRES) that facilitates Zeb2 translation. ZEB2 protein then transcriptionally represses E-cadherin to facilitate EMT (42).

Telomere maintenance. The lncRNA TERC has long been known to be involved in telomere maintenance (43). More recent-ly, another ncRNA in this process has been discovered that pro-vides insight into telomere maintenance in cancer cells. Telomeres produce a large heterogeneous ncRNA named TERRA that binds to both telomeres and TERT to inhibit TERT activity (44). Tumor cell lines and immortalized primary human cells that use TERT to maintain telomeres escape TERRA-mediated TERT repression through heavily methylated subtelomeric DNA to repress TERRA expression (45).

Hormone receptor signaling. Signaling by the nuclear hor-mone receptors for androgens (AR) and estrogens (ER) is a fundamen-tal aspect of prostate and breast cancer, and the role for the AR is expanding. A number of lncRNAs have been identified that participate in AR or ER signaling and may provide novel target-ing strategies.

The lncRNA PCGEMI is known to be overexpressed in prosta-tate cancer patient samples, and its overexpression inhibited apoptosis in a prostate cancer cell line in an AR-dependent man-ner, confirming a functional interaction with AR (46). Subsequent-ly, a detailed mechanism for PCGEMI was uncovered that involves another lncRNA, PRNCR1, which was identified in our pilot study as PCAT-8 (11). PRNCR1 binds to the AR C-terminal which, in association with histone H3 methyltransferase DOT1L, facilitates recruitment of PCGEMI to the AR N terminus. This recruitment enhances AR-mediated transcriptional activation by inducing looping of remote AR-bound enhancers to target gene promoters. Thus, overexpression of PRNCR1 and PCGEMI contributes to ca-stration-resistant AR signaling in prostate cancer (47). While this mechanism is intriguing, it was difficult to reproduce in our hands and further studies will be needed to verify this interaction (48). CTBP1-AS is an antisense lncRNA of the AR corepressor CTBP1. CTBP1-AS directly inhibits CTBP1 transcription via recruitment.
of the RNA-binding transcriptional repressor PSF and histone deacetylases. This mechanism was found to inhibit several tumor suppressor genes across the genome in an AR-dependent manner, resulting in enhanced cell-cycle progression (49).

A recent study demonstrated that approximately 25% of the genome is transcriptionally regulated by ER in breast cancer cells, including over 1,500 unannotated intergenic, antisense, and divergent transcripts; a substantial proportion of these transcripts are likely to be IncRNAs (50). More specifically, HOTAIR is transcriptionally upregulated by estrogen in a breast cancer cell line where it supports proliferation and suppresses apoptosis, a potential mechanism underlying its overexpression in primary breast tumors (51). In prostate cancer, transcriptional regulation by ERα upregulates the IncRNA NEAT1, whose expression was associated with poor outcome (52). These data demonstrate that IncRNAs are regulated by estrogen signaling and play a major role in estrogen-related cancers.

**Competitive endogenous RNA.** IncRNA genes also have a functional interaction with another major class of noncoding RNAs, namely miRNA. Poliseno and colleagues hypothesized a regulatory role for long RNAs (such as IncRNAs) in the binding of miRNAs and uncovered a mechanism that has important implications in cancer biology (53). They found that the phosphatase and tensins homolog (PTEN) pseudogene PTENP1 acts as a molecular sponge for miRNAs that target PTEN mRNA for degradation. They named this class of IncRNAs competitive endogenous RNAs (ceRNAs). Interestingly, the PTENP1-sequestered miRNAs also target other tumor-suppressor genes, including E2F transcription factor 1 (E2F1), CDKN1A, and programmed cell death 4 (PDCD4). The authors were able to show that RNAi against PTENP1 resulted in downregulation of CDKN1A and increased proliferation in PTEN-null cells. Conversely, the PTENP1 locus was deleted in a cohort of colon tumor samples, which also exhibited decreased PTEN expression levels. As suspected, this mechanism is not specific to PTENP1. KRAS and its pseudogene KRASIP share miRNA let-7-binding sites; their expression was positively correlated in a breast cancer cohort and the KRASIP locus is narrowly amplified in several cancer types (53). These findings demonstrate that the ceRNA function of IncRNAs has a novel gene dysregulation function in cancer through modulation of miRNA function.

**RNA processing.** Several IncRNA functions that appear to contribute to oncogenic phenotypes do not fall into classical pathways, and MALAT-1 provides one interesting example. MALAT-1 is a well-studied IncRNA that was initially identified by subtractive hybridization in early lung cancer tumor samples from patients who did or did not eventually develop metastasis; thus, MALAT-1 was associated with aggressive disease (54). Further research efforts bolstered this role in lung and prostate cancer by demonstrating that MALAT-1 promoted neoplastic behavior in multiple preclinical cancer models, including lung, colorectal, and prostate cancer (55–57). The molecular mechanisms underlying this role in cancer remain vague. MALAT-1 is involved in modulation of mRNA splicing and is found in nuclear paraspeckles, which are sites of RNA processing and editing (58, 59); however, there are conflicting reports over whether the mRNA-processing function of MALAT-1 is responsible for its neoplastic effects (60).

**Difficulties in modeling IncRNA functions in mouse models.** Historically, transgenic mice have been arguably the most valuable preclinical model system for evaluating gene function and testing experimental therapeutic agents. However, IncRNA transgenic mice present inherent difficulties, few models have been developed, and the resulting data have been mixed. The primary inherent difficulty of IncRNA transgenic mouse models is that IncRNAs are conserved at much lower rates than protein-coding genes, so that many human IncRNAs do not exist as expressed genes in mice. Among the few that have been made, H19, MALAT-1, and NEAT1 knockout mice are grossly normal (61), while deletion of Xist in hematopoietic progenitors led to hematologic neoplasms (62). Two knockout mice for HOTAIR have been generated with differing results, which is instructive. A mouse harboring deletion of most of the HOXC locus, which includes HOTAIR, was reported to have a minimal phenotype at the molecular or developmental level (63). Two more targeted HOTAIR knockout mice showed very similar mild but reproducible homeotic phenotypes that were consistent with the known function of HOTAIR (64, 65). These differences demonstrate the difficulty and importance of designing “clean” IncRNA mouse models. These difficulties may be addressed by advanced transgenic mouse models, wherein large human genome portions, including whole chromosomes, are added to or replace portions of the mouse genome (66).

**IncRNA deregulation in cancer.** There is intense interest in uncovering exactly how tumor cells co-opt IncRNA function to contribute to oncogenic phenotypes. As described above, the primary mode described to date is up- or downregulation of IncRNA expression levels, but the molecular mechanisms underlying these transcriptional changes have yet to be studied in detail. A global study comparing proximal promoter sequences of protein-coding and noncoding genes found statistical differences in the presence of transcription factor–binding sites and chromatin states. This study also used machine learning to build a protein-coding versus noncoding promoter classifier; however, it is unclear whether these findings explain lineage or cancer specificity of IncRNA expression (67). Copy-number alteration (CNA) can alter IncRNA expression levels in a manner similar to that of protein-coding genes. The IncRNA FAL1 was identified through an unbiased global search for IncRNAs overexpressed by CNAs and for IncRNAs within recurrent CNAs from nearly 2,400 tumor samples. FAL1 expression is associated with outcome in ovarian cancer and interacts with PRC1 component BMI1 to repress numerous genes, including CDKN1A (68). Perhaps more strikingly, the IncRNA PVT1 is located on the 8q24.21 amplicon in MYC amplification and upregulation of PVT1 by coamplification is required for the oncogenic effects of MYC amplification (69, 70).

Oncogenic alterations of protein-coding genes, including point mutations, deletion, and gene fusion, are by and large easy to detect and predict based on the thorough understanding of protein-reading frames and domains. However, this level of understanding does not yet exist for IncRNAs, and so the effects of point mutations, deletion, or gene fusion are very difficult to predict. Indeed, methods for confident detection of IncRNA point mutation or deletion have not yet been developed, though several gene fusions involving IncRNAs have been identified. One of the IncRNAs identified
in our PCAT pilot study, PCAT-I4, can be found in a somatic gene fusion with the Ets transcription factor family member ETV1. This fusion retains the PCAT-I4 promoter, which contains an AR-binding site, allowing for androgen regulation of expression (11). In a B cell lymphoma patient, a GA35-BCL6 fusion was found as a result of a t(3;14) translocation. Here again, the lncRNA GA35 essentially supplies only its promoter; the entire coding sequence of BCL6 is retained, which is almost certainly the functional portion, given that other BCL6 fusions are common in this disease (71). It seems likely that oncogenic alterations of lncRNAs are occurring in cancer and that they will be uncovered with further refinement of bioinformatics and the sheer explosion in sequencing data.

**IncRNAs in cancer risk and SNP studies**

There have been numerous GWAS for germline SNPs that associate with a cancer predisposition. A troublesome aspect of these studies has been that most of the SNP loci discovered so far do not have a clear relationship with a known protein-coding gene. A recent survey of the GWAS catalogue identified 301 SNPs associated with increased cancer risk, of which only 12 (3.3%) had an effect on the amino acid sequence of a protein-coding gene (72). Using a catalogue of 11,194 disease-associated SNPs (not restricted to cancer) in our global MiTranscriptome study, we observed that our newly defined transcripts overlapped 2,181 intergenic SNPs (16). This finding provides strong support, though not proof, for the hypothesis that unassociated SNPs are actually associated with unidentiﬁed noncoding transcripts (73).

Several studies on individual cancer risk SNPs provide proof of principle for the mechanistic aspect of this hypothesis. The deletion allele of SNP rs10680577 correlates with increased hepatocellular carcinoma (HCC) risk and with upregulation of egl-9 family hypoxia-inducible factor 2 (EGLN2) and RERT-lncRNA, while the insertion allele alters the function of RERT-lncRNA (74). In papillary thyroid cancer, the 14q13.3 lncRNA PTCS3C is repressed by the SNP rs944289 risk allele through reduction of C/EBP binding to the lncRNA locus and resultant derepression (75). Further analysis of other unexplained cancer risk regions may help identify particularly important lncRNAs for further study.

**IncRNAs as diagnostic and prognostic biomarkers**

Cancer molecular biomarkers have improved dramatically in the last 2 decades. Diagnostic, prognostic, and predictive biomarkers allow for confident identification of lung adenocarcinoma with transcription termination factor, RNA polymerase I (TTF-1) (76), poor prognosis of neuroblastoma with MYCN amplification (77), and anti-estrogen or -HER2 therapy in breast cancer with ER expression or HER2 amplification, respectively. TCGA results show marked molecular heterogeneity among most cancer types, suggesting that further tumor subclassification is needed; such analysis will almost by definition require molecular biomarkers. The primacy of specificity in diagnostic biomarkers points directly to IncRNAs, as there is mounting evidence that IncRNAs are expressed in a more tissue-specific manner than protein-coding genes (9). Consistent with this, the nearly 8,000 cancer- and/or lineage-speciﬁc IncRNAs from our MiTranscriptome study represent a rich resource for biomarker studies (16).

Several examples from our group and others highlight the potential of lncRNA biomarkers, including use in noninvasive body fluid tests. SChLAP1 was originally identiﬁed as an unannotated, noncoding transcript with outlier expression in metastases in the relatively small cohort for our PCAT pilot study (14). In nearly 1,100 patients undergoing prostatectomy followed by high-density microarray proﬁling of the prostatectomy sample and long-term follow-up, we performed an unbiased analysis wherein SChLAP1 was independently renominated and validated as a biomarker of metastatic progression risk (13). Based on these studies, our group has developed and validated an ISH test for SChLAP1 expression that could be used to guide therapy intensiﬁcation (78). PCA3 is a prostate-speciﬁc IncRNA that is highly overexpressed in the majority of prostate cancers. A urine test developed by our group to detect PCA3 and TMPRSS2-ERG fusion transcripts for noninvasive prostate cancer diagnosis outperforms PSA or PCA3 tests alone, has been approved by the US FDA, and is showing promise in clinical settings (79, 80). The commercial prostate cancer prognostic signature Decipher includes several noncoding transcripts (81). Lung cancer IncRNA MALAT-1 as a
plasma biomarker in non–small cell lung cancer had a sensitivity of 56% and a specificity of 96% (82). Additionally, several other lncRNAs have been detected in body fluids and may allow for noninvasive detection and monitoring of different cancers. For example, HULC can be found in HCC patient plasma (83), H19 in gastric cancer patients (84), and several lncRNAs in oral cavity squamous cell carcinoma patient saliva (85). These studies strongly indicate that lncRNAs will be developed as useful noninvasive biomarkers across a range of cancer types.

**lncRNAs as therapeutic targets**

lncRNA biology has already suggested many promising therapeutic targets. Few lncRNAs have been thoroughly validated as targets, but MALAT-1 provides an illustrative and promising example. MALAT-1 knockout mice are resistant to carcinogenesis in breast and lung cancer models and show a minimal phenotype, indicating that toxicity resulting from disruption of MALAT-1 would be unlikely (86–88). The most straightforward therapeutic targeting strategy is to directly target the RNA by sequence, which can be accomplished through several technologies. Antisense oligonucleotides (ASOs) have recently received FDA approval for two nonmalignant diseases, familial hypercholesterolemia (89) and transthyretin amyloidosis (90). Though ASOs have yet to be proven as an anticancer therapy, MALAT-1 ASOs have shown efficacy in a preclinical breast cancer model (91). siRNA-based therapeutics are in development and have entered clinical trials with therapeutics targeting a range of mRNAs in cancer and other diseases (92). Hammerhead ribozymes have self-contained nucleolytic activity and high-specificity sequence recognition, but have not been tested in humans (93). Synthetic RNAs or small RNAs could be used to redirect chromatin-modifying complexes for gene expression modulation, and antagoNATs could be used to target the natural antisense transcript (NAT) class of lncRNAs (94, 95). Sequence-specific therapeutics have many advantages, but require further development.

There is a host of other therapeutic strategies seeking to exploit other aspects of lncRNA biology. Efforts are underway to disrupt RNA-protein binding sites (96, 97). Another strategy takes advantage of the restricted expression of lncRNAs by using lncRNA regulatory elements. BC-819 is a plasmid containing the diphtheria toxin gene under H19 promoter control that has shown promising results as a cytoreduction agent in bladder, ovarian, and pancreatic tumors (98). An alternative strategy in myotonic dystrophy targets RNA structure with small molecules, in this case binding aberrant ncRNA structure caused by repeat expansion, thereby preventing the pathological binding to and inhibition of muscleblind-like splicing regulator 1 (MBNL1) (99). Based on their signal transduction relationship, mTOR inhibition can increase GASS levels in androgen-responsive prostate cancer cell lines and reduce proliferation (100). There does not appear to be a clear winner among these alternative technologies.

**Conclusions**

The role of lncRNAs in basic, translational, and clinical oncology is likely to equal and perhaps even surpass the role of protein-coding genes. We envision that lncRNA-based clinical tools will expand rapidly in the near future, including as diagnostic and prognostic biomarkers, and as therapeutic targets (Figure 3). While the door to this new world has been opened for us to see the vast potential of the dark matter of the genome, including several clinically available biomarker tests, we still await the first big steps that will benefit cancer patients.

**Acknowledgments**

We wish to thank John Prensner for his helpful discussions about the organization of the article, and Karen Giles for her diligence in managing correspondence. We thank RSNA (Research Resident Grant #RR1420 to JRE) and Prostate Cancer Foundation (Challenger Award to FYF and JRE) for funding lncRNA research.

Address correspondence to: Arul M. Chinnaian, Michigan Center for Translational Pathology, Howard Hughes Medical Institute, The University of Michigan Cancer Center, 1500 E. Medical Center Drive, 5309 Comprehensive Cancer Center, SPC5940, Ann Arbor, Michigan 48109, USA. Phone: 734.615.4062; E-mail: arul@med.umich.edu.
Erho N, et al. Discovery and validation of a pros-

Laxman B, et al. A first-generation multiplex bio-

Tomlins SA, et al. Urine TMPRSS2:ERG fusion


Huang M, Weiss WA. Neuroblastoma and

Srodon M, Westra WH. Immunohistochemical

Jendrzejewski J, et al. The polymorphism

Zhu Z, et al. An insertion/deletion polymorphism

Edwards SL, Beesley J, French JD, Dunning AM.

REVIEW

PLoS

tate cancer genomic classifier that predicts early

Lopez-Berestein G. Preclinical and clinical devel-

Tedeschi L, Lande C, Cecchettini A, Gitti L.
Hammerhead ribozymes in therapeutic target

Moazed D. Small RNAs in transcriptional

Wahlestedt C. Targeting long non-coding RNA to

for small-molecule modulators of long noncod-

Falk SP, Weisblum B. Aptamer displacement
screen for flaviviral RNA methyltransferase

Smaldone MC, Davies BJ. BC-819, a plasmid
comprising the H19 gene regulatory sequences
d and diphtheria toxin A, for the potential tar-

Lee MM, et al. Controlling the specificity of mod-
ularly assembled small molecules for RNA via
ligand module spacing: targeting the RNAs that

Yacqub-Usmann K, Pickard MR, Williams GT.
Reciprocal regulation of GAS5 IncRNA levels and