

Shedding Light on the Dark Cancer Genomes: Long Noncoding RNAs as Novel Biomarkers and Potential Therapeutic Targets for Cancer



Lin Zhang^{1,2}, Dan Peng³, Anil K. Sood^{4,5}, Chi V. Dang^{6,7}, and Xiaomin Zhong³

Abstract

Recently there have been explosive discoveries of new long noncoding RNAs (lncRNA) obtained by progress in the technology of second-generation sequencing. Genome scale analysis of transcriptome, in conjunction with studies on chromatin modifications at the epigenetic level, identified lncRNAs as a novel type of noncoding transcripts whose length is longer than 200 nucleotides. These transcripts are later

found as major participants in various physiologic processes and diseases, especially in human cancers. lncRNAs have been found to function as novel types of oncogenes and tumor suppressors during cancer progression through various mechanisms, which endow them with the potential of serving as reliable biomarkers and novel therapeutic targets for cancers. *Mol Cancer Ther*; 17(9); 1816–23. ©2018 AACR.

Introduction

In recent years, there has been an explosive growth in the identification of long noncoding RNAs (lncRNA). Identification of new lncRNAs has been facilitated by progress in the technology of second-generation sequencing. This has allowed genome-wide analyses of the transcriptome and chromatin modifications at the epigenetic level. Previous research has shown the majority of the human genome produces a great number of transcripts without protein coding potential (1, 2). lncRNAs are one such class of noncoding transcripts. They are operationally defined as RNA molecules longer than 200 nucleotides that do not appear to have canonical protein-coding potential, i.e., they contain no classic open reading frame (3–8). Considering that nearly 70% of the human genome is transcribed to RNA products, among which only a minority encodes proteins (2), the number of lncRNA genes appears very large. Following the initial cloning of lncRNAs such as XIST (9) and H19 (10, 11) from cDNA libraries, two independent groups reported that the number of lncRNA genes is no less than that of protein-coding genes by using the tiling array technology (12, 13). Progress in tiling arrays (12–15), analysis of

chromatin epigenetic modification (16), computational analysis of cDNA libraries (17, 18), and RNA sequencing (1, 19–21) have demonstrated that thousands of lncRNAs are widely expressed in human with noticeable specificity in tissue distribution. Recently, the GENCODE consortium (version 27) has reported 27,908 manually annotated and evidence-supported human lncRNAs originating from 15,778 gene loci (1, 2).

Many studies have revealed universal features of lncRNAs to help better understand this new class of transcripts. First, lncRNAs are independent transcriptional units. Many of the first identified lncRNAs were located in genomic regions that did not contain any previously known genes, for example, lncRNAs HOTAIR and H19 (14, 22). Now it is widely recognized that lncRNAs are unlikely to exist as "transcriptional noise" or "gene trash" as perceived before. Second, lncRNA transcripts contain fewer exons than mRNAs and utilize canonical splicing sites. Third, lncRNA transcripts are under weaker selection pressures than protein-coding ones during evolution, and many are primate-specific. It is reasonable to conclude that lncRNAs bear more possibility in nucleic acid variations during evolution because they have no need to preserve sequences critical for coding conserved amino acids of functional protein domains. Under such weaker selection constraints, primates might have flexibly evolved a set of specific lncRNAs to help accommodate unique physiologic functions of the species. Fourth, lncRNA transcription is under the control of epigenetic modifications of chromatin similar to that for protein-coding mRNAs. Many lncRNAs were initially identified by large-scale sequencing of cDNA libraries and subsequently characterized by transcriptional signatures from RNA pol II binding and epigenetic modifications of chromatin (3, 14). The transcription of lncRNAs is regulated by DNA modifications and histone codes similar to that for protein-coding mRNAs. Fifth, lncRNAs and protein-coding mRNAs also utilize the same transcriptional machinery. Transcription of the majority of lncRNAs is RNA pol II-dependent. lncRNA transcripts usually have a 5' terminal methyl-guanosine cap and are polyadenylated. There is also evidence showing certain lncRNAs deficient in 3' poly(A) tails, which are probably produced from RNA pol III promoters (23, 24). In addition, the splicing of nucleolar RNAs generates a subset of

¹Center for Research on Reproduction and Women's Health, University of Pennsylvania, Philadelphia, Pennsylvania. ²Department of Obstetrics and Gynecology, University of Pennsylvania, Philadelphia, Pennsylvania. ³Center for Stem Cell Biology and Tissue Engineering, Key Laboratory for Stem Cells and Tissue Engineering, Ministry of Education, Department of Biology, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou, China. ⁴Center for RNA Interference and Non-coding RNA, University of Texas MD Anderson Cancer Center, Houston, Texas. ⁵Department of Gynecologic Oncology and Reproductive Medicine, University of Texas MD Anderson Cancer Center, Houston, Texas. ⁶Wistar Institute, Philadelphia, Pennsylvania. ⁷Ludwig Institute for Cancer Research, New York City, New York.

Corresponding Authors: Lin Zhang, University of Pennsylvania, 1207 BRB II/III, Philadelphia, PA 19104. Phone: 215-573-4780; Fax: 215-573-7627; E-mail: linzhang@pennmedicine.upenn.edu; and Xiaomin Zhong, 74 Zhongshan Road 2nd, Medical Science and Technology Building, Rm 643, Guangzhou 510080, China. E-mail: zhongxm23@mail.sysu.edu.cn

doi: 10.1158/1535-7163.MCT-18-0124

©2018 American Association for Cancer Research.

lncRNAs lacking in 5' caps and 3' poly(A) tails (25). This suggests that lncRNAs share the same molecular machinery for biogenesis with protein-coding mRNAs. Crosstalk and reciprocal regulation in the transcription of lncRNAs and protein-coding mRNAs may exist. Last, lncRNA expression is comparatively low and highly tissue-specific. Comprehensive analysis of human lncRNAs revealed that lncRNAs are generally expressed at lower levels than protein-coding genes, and they manifest more tissue-specific expression patterns. For example, a large number of lncRNAs are expressed specifically in the brain (1).

Molecular Mechanisms of lncRNAs

Despite the numerous functions lncRNAs possess, they utilize some common working mechanisms. These common mechanisms involve control of epigenetic modifications of chromatin, gene transcription, and mRNA stability, as well as protein translation (16, 19, 26–30). First, lncRNAs usually regulate gene expression by modulating epigenetic modifications of chromatin. lncRNA transcripts can form scaffold structures and mediate the interaction between chromatin and epigenetic modification factors. For example, lncRNA HOTAIR recruits the silencing complexes PRC2 and LSD1/CoREST/REST to targeted DNA regions for H3K27 methylation and H3K4 demethylation (27), and consequently represses the expression of a large amount of HOTAIR target genes, such as HOXD cluster genes.

Second, lncRNAs promote or repress gene expression at the transcriptional level through various mechanisms. A specific group of lncRNAs termed enhancer RNAs (eRNA) have been suggested to upregulate gene expression by acting as transcriptional enhancers (19, 28). On the other hand, lncRNAs can effect on transcription factors to control gene expression. The lncRNA NRON (noncoding repressor of NFAT) suppresses nucleocytoplasmic shuttling of the transcription factor NFAT, leading to repression on downstream gene transcription (30). In some cases, lncRNAs function as decoys for transcription factors or compete for DNA binding sites with transcription factors. For example, the lncRNA PANDA can sequester NF- κ B protein away from its target genes involved in apoptosis (31). Interestingly, lncRNAs also regulate gene expression by directly affecting the organization of the nuclear compartment. Examples are MALAT1 and TUG1, who reside in distinct nuclear structures of interchromatin granules and Polycomb bodies respectively, can cooperatively control relocation of growth-control genes between the two nuclear compartments, and mediate the assembly of coactivators or corepressors to regulate gene expression in response to growth signals (32).

Third, in addition to the functions in epigenetic and transcriptional control on gene expression, lncRNAs also have proven activities in posttranscriptional regulation, affecting the processing, stability, and translational efficiency of mRNAs. MALAT1 was reported to modulate the nuclear localization and levels of phosphorylated Ser/Arg splicing factors, which directly affects the alternative splicing pattern of target gene pre-mRNAs (33). Like protein-coding genes, some lncRNAs and transcripts from pseudogenes act as miRNA sponges to sequester miRNAs away from target transcripts and promote the stabilization of mRNA molecules (29, 34, 35). lncRNAs also exert their functions in translational control. The association of lincRNA-p21 and the general translation repressor Rck was shown to repress the translational

efficiency of β -Catenin and JUNB mRNAs (26). Similarly, a translational regulatory lncRNA (treRNA) was also reported to inhibit the translation of E-Cadherin mRNA (36).

Meanwhile, in addition to the noncoding functions supported by a vast amount of evidence as mentioned, recent studies implied that lncRNAs may also perform their functions by translation into proteins. Using ribosome profiling or ribosome footprinting methods, these studies showed that lncRNAs are physiologically associated with ribosomes, and the pattern of ribosome protection suggested the possibility of lncRNA transcripts in translating into short peptides (37–40). Furthermore, Schier and colleagues identified a 58-amino acid peptide encoded by a previously annotated noncoding RNA Toddler in zebrafish, and established its essential role in embryonic signaling (41). Nevertheless, whether lncRNAs play their biological roles through encoding short peptides still remains a controversial question (42). It will be an attractive field to help fully elucidate the working mechanism of lncRNAs.

lncRNAs in Human Cancers

Although the biological functions of lncRNAs have not been fully explored, more and more studies indicate that a most prominent function of lncRNAs is their regulatory roles in the development of human cancers. lncRNAs are active participants in tumor development by widely modulating cancer cell proliferation, apoptosis, migration, metabolism, and differentiation (6–8, 15, 31, 43–45). The lncRNA HOTAIR promotes metastasis of cancers through PRC2-mediated epigenetic modification on target genes (15). Correspondingly, increased expression of HOTAIR correlates with poor prognosis of many types of cancers (46, 47). ANRIL represents another typical example of oncogenic lncRNA. ANRIL expression is upregulated in prostate cancer. Its association with both CBX7/PRC1 (48) and SUZ12/PRC2 (49) leads to repression of the tumor suppressor *INK4A/INK4B*. The lncRNA *FAL1* with amplification in gene copy number showed increased expression levels in several types of cancers. *FAL1* enhances the growth rate of cancer cells by interacting with BMI1/PRC1 to suppress *CDKN1A* expression (50). Similarly, the lncRNA *DANCR*, which is widely overexpressed in many types of human cancers, was found to promote cancer progression partly by modulating *CDKN1A* expression (51). A TP53-regulated lncRNA, *LINP1*, enhances double-strand DNA break repair through the nonhomologous end joining pathway and desensitizes triple-negative breast cancer cells to radiotherapy (52). The lncRNA ceruloplasmin (NRCP) modulates cancer metabolism and promotes cancer growth by increasing the expression of genes related to glycolysis (45). In addition, a large number of lncRNAs were also found to be associated with tumorigenesis, such as *PCAT-1* (20), *PANDA* (31), *MALAT-1* (53), *lincRNA-p21* (44), *LINC00673* (54), etc.

It is worth noticing that genomic alterations and transcriptional deregulation of lncRNA genes are widely present in cancer cells. The somatic copy-number alterations (SCNA) of lncRNA genes across 5,860 tumor samples from 13 tumor types from the Cancer Genome Atlas project were extensively analyzed. It was found that a high frequency of copy-number gain and loss existed in an average of 13.16% and 13.53% of lncRNA genes, respectively (55). In line with the genome-wide landscape of SCNAs, individual lncRNAs with copy-number variation were identified. For example, the oncogenic lncRNA *FAL1* with amplification in gene

copy number shows increased expression levels in many types of cancers (50). Focal amplifications containing both the lncRNA *SAMMSON* and the melanoma-specific oncogene *MITF* were located to chromosome 3p13 and 3p14. *SAMMSON* was characterized as a critical player in promoting melanoma cell growth and survival (56). In addition to SCNAs, SNP is another common genomic alteration in lncRNA genes associated with cancer risk and clinical outcomes. Genome-wide studies have indicated frequent overlap between lncRNAs and cancer-related SNPs (55, 57). The tumor suppressor lncRNA *LINC00673* antagonizes pancreatic cancer cell proliferation by promoting degradation of the oncogene *PTPN11*. A germline variant of *LINC00673* has impaired activity and sensitizes cells to malignant transformation (54). *NBAT-1* is an lncRNA that controls neuroblastoma cell proliferation and invasion. Interestingly, an SNP in the intron of *NBAT-1* indicates high risk and poor clinical outcome of neuroblastoma (58). Besides genomic alterations, transcriptional deregulation of lncRNA genes is another mechanism responsible for their differential expression patterns in cancers. A study showed that DNA methylation patterns in the promoter regions of lncRNA genes are intrinsically distinct in cancerous and normal tissues (55). Furthermore, lncRNA transcription is under control of important oncogenes or tumor suppressors. A subset of lncRNA genes is either suppressed (such as *CONCR* and *LINP1*) or promoted (such as *lincRNA-p21* and *DINO*) by *TP53* (44, 52, 59, 60). The oncogenes *RAS* and *MYC* are also involved in transcriptional regulation of lncRNAs such as *Orilnc1* and *DANCR* (51, 61).

lncRNAs May Serve as Biomarkers and Therapeutic Targets for Human Cancers

The explosion in lncRNA research has evoked great enthusiasm for the possibility of lncRNAs serving as diagnostic markers and therapeutic targets for human cancer. Hopefully, the unique features of lncRNAs, i.e., their differential expression patterns in cancer, temporal and tissue specificity, and diverse biological functions, will be advantages taken in clinical applications. The first inspiring case of lncRNA as a diagnostic marker is the FDA-approved testing of *PCA3* for detection of prostate cancer (62). Together with the traditional serum prostate-specific antigen (PSA) testing, *PCA3* testing has made diagnosis of this disease more accurate and specific. Potential lncRNA diagnostic markers for other types of cancers have been reported, including hepatocellular carcinoma and gastric cancer (63, 64). In addition, lncRNAs have been widely investigated as prognostic markers for cancer patients. Previous studies have revealed a correlation between lncRNA *HOTAIR* expression and colorectal cancer recurrence, lymph node metastasis, and poor prognosis (46, 65). On the other hand, lncRNAs are considered ideal biomarkers due to their presence in body fluids, such as in the form of exosome-containing RNA (66–68), which enables noninvasive diagnosis of cancers. Although numerous studies have shown differential expression of lncRNAs in tumors compared with normal tissues, extensive research and clinical data are needed to confirm the consistency of lncRNA expression in tumor tissue and body fluids for diagnostic accuracy and specificity. Despite limited progress in the development of lncRNA biomarkers so far, the diversity in biological functions of lncRNAs endows them with great potential in serving as therapeutic targets for human cancers. A supportive case is lncRNA *MALAT1*, an oncogene first discovered in

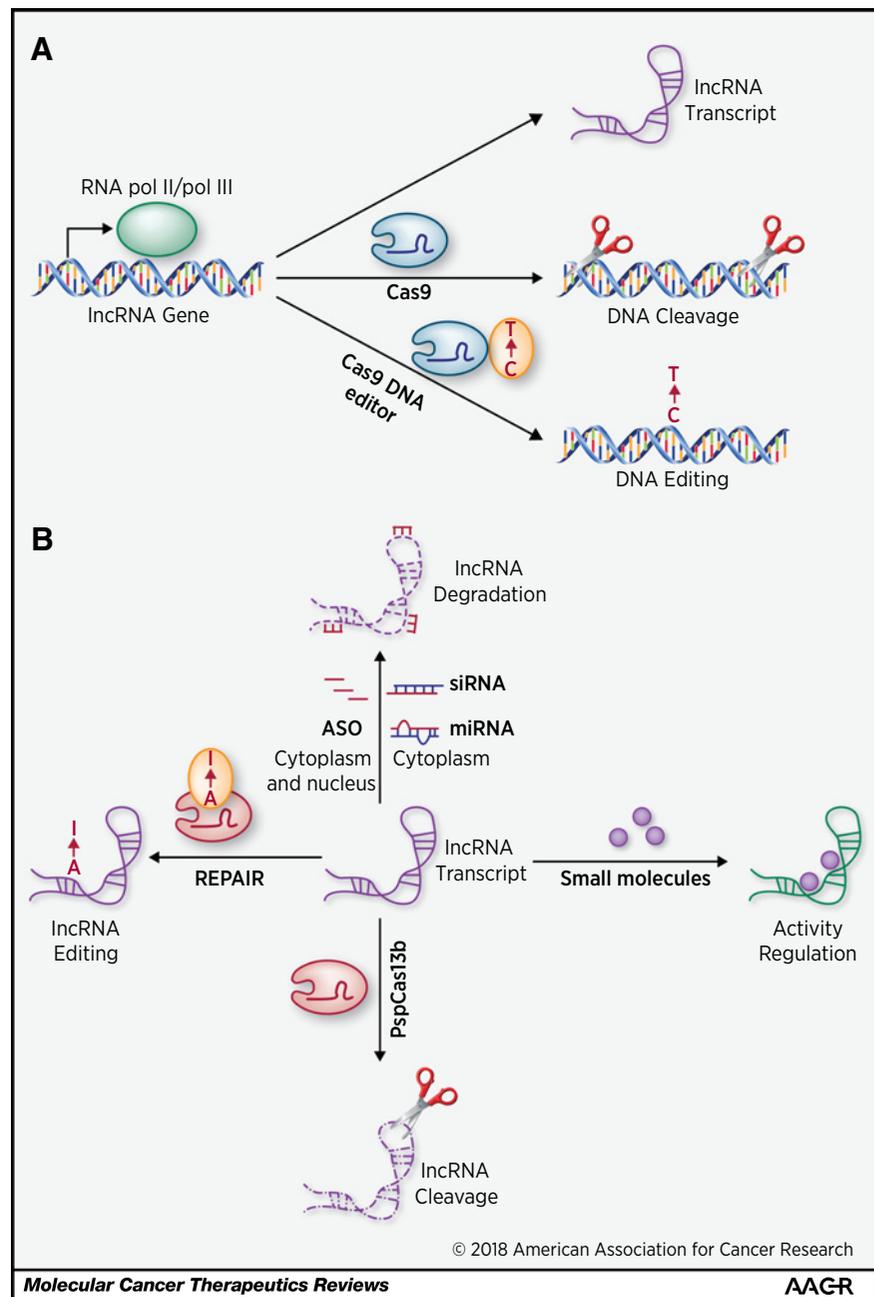
metastatic lung adenocarcinoma patients (53) and later found highly expressed in lung, breast, and prostate cancers (69). It plays critical roles in promoting proliferation and metastasis of cancer cells. Animal studies indicated that silencing *MALAT1* expression by antisense oligonucleotide (ASO) efficiently impaired the tumor growth and metastatic ability of breast and lung cancer cells, respectively (70, 71), implying a potential therapy for cancers by manipulating oncogenic lncRNA expression.

Strategies for Developing Targeted Therapy for Cancers with lncRNAs

As critical players in tumor development, lncRNAs represent promising therapeutic targets for treating cancers. However, targeting lncRNA molecules is different from targeting protein-coding transcripts. First, the lack of protein products encoded by lncRNAs means that targeting methods are restricted to RNA molecules only. However, methods for targeting RNA are currently limited. Second, unlike proteins with conserved domains and specific conformations, which serve as good targeting sites by small-molecule drugs, the three-dimensional structures of lncRNAs have been poorly explored. Furthermore, it has been speculated that lncRNA functions are independent on conserved three-dimensional structures, as they show relatively low evolutionary conservation of sequence compared with protein-coding genes. This makes structure-based drug design and screening difficult for lncRNA targets. Third, the working mechanism and regulation network of most lncRNAs have not been fully understood. This increases difficulties in developing specific targeting strategies for lncRNAs.

Despite the mentioned problems, diverse targeting methods for lncRNAs are under investigation (Figure 1 and Table 1). Using small interfering RNA (siRNA) and ASO to induce loss-of-function effect on lncRNAs represents the most common strategy. Similar strategies for targeting protein-coding genes have proven successful, and the systemic delivery methods have been widely explored (72–74). It has been well proved that both siRNA and ASO function by base pairing with mRNA to form double-stranded RNAs or RNA-DNA hybrids, resulting in degradation of target mRNA by the RNAi mechanism or by RNase H activity. This method has been adapted for knocking down lncRNAs with as potent effect as for mRNA silencing. For example, *MALAT1* expression was suppressed by ASO to impair the tumor growth and metastatic ability of breast and lung cancer cells (70, 71). Furthermore, siRNA and ASO are usually subjected to certain modifications in order to enhance their stability, meanwhile retaining targeting specificity and alleviating the interferon induction effect. Such modifications include adding 2-nt 3' overhangs to siRNA, 2'-O-methyl to siRNA and ASO, and locked nucleic acid to ASO. To overcome poor stability and increase intracellular uptake of siRNA and ASO, delivery methods such as lipid-based carriers, polymersomes, and biocompatible nanoparticles can be employed. Last, the subcellular localization of lncRNAs should be considered when adapting the siRNA/ASO strategy. For lncRNAs located in the cytoplasm, both siRNA and ASO have satisfactory silencing effects. But for those located in the nucleus, ASO would be a better choice than siRNA due to the lack of RNAi machinery in the nuclear compartment. Meanwhile, considering that some lncRNAs function as sponges for miRNAs to alleviate the suppressive effects of miRNAs on endogenous target mRNAs (34, 54), targeting lncRNAs by miRNAs would also be a promising

Figure 1. Strategies for developing lncRNA-targeted therapy for cancer treatment. **A**, lncRNA genes are transcribed by RNA pol II/pol III to give rise to lncRNA transcripts (top). The Cas9 system can be applied to target and cleave DNA fragments from lncRNA genes (middle). And the Cas9 DNA editor system is capable of editing specific DNA bases (e.g., C to T) of lncRNA genes to correct tumorigenic SNP (bottom). **B**, lncRNA transcripts can be targeted by multiple ways. ASO, siRNA, and miRNA represent the most popular methods to knock down the expression level of lncRNA (top). The mechanism of ASO, siRNA, and miRNA involves base pairing with lncRNA which leads to the degradation of lncRNA. Alternatively, lncRNA expression level can also be modulated by the PspCas13b system by its ability in RNA-guided recognition of RNA targets and its RNase activity (bottom). A modified Cas13 system with nucleotide editing activity (e.g., A to I), named REPAIR, can be further used to edit specific nucleotides in lncRNA, hence change the biological functions of lncRNA rather than cleave lncRNA molecules (left). Small molecules that can potentially bind to lncRNA transcripts may also be a prospective way to regulate the structural conformation and biological activities of lncRNAs (right).



strategy. The prominent advantages of using miRNAs include their compatibility with the endogenous regulatory machinery for lncRNAs and the well-established delivery methods (74, 75).

Besides traditional methods using siRNA, ASO, and miRNA, the recently characterized clustered regularly interspaced short palindromic repeats (CRISPR)-associated deoxyribonuclease Cas9 gene editing system may be employed to achieve loss-of-function effects on lncRNAs. In this system, the Cas9 endonuclease is targeted to specific DNA sequences with a protospacer adjacent motif (PAM) by single-guide RNAs (sgRNA) that are artificially designed (76–79). This system has been widely applied for targeted DNA cleavage. For oncogenic lncRNA genes with well-

characterized genomic localization, Cas9 can be directed by a pair of specifically designed sgRNAs targeting the 5' and 3' ends of the genes to generate knockout loci. In this way, successful knockouts of lncRNA-21A, UCA1, and AK023948 in various human cell lines have been accomplished (80). Similarly, utilizing a paired-guide RNA library, CRISPR-Cas9-mediated genome-scale deletion of lncRNAs was achieved (81). The Cas9 system is emerging as a powerful tool for targeting lncRNAs in cancer. Major advantages of the Cas9 system include, but not limited to, programmable targeting effects (e.g., gene knockout, as well as gene mutation and transcriptional activation) and convenience for high-throughput screening of lncRNA knockout phenotypes. Furthermore, a recent

Table 1. Strategies for targeting lncRNA in cancer therapy

Strategies	Effects	Advantage	Limitation
siRNA, ASO, miRNA	RNA degradation	Specific, potent effect	Stability, delivery
Cas9	Gene knockout, DNA editing, gene mutation, etc.	Versatile and long-term effect	Genomic context dependent, PAM requirement, targeting efficiency, <i>in vivo</i> delivery
PspCas13b	RNA cleavage	Programmable targeting, less sequence constraint	Targeting efficiency, <i>in vivo</i> delivery
REPAIR (dCas13-ADAR)	RNA editing	Programmable editing	Targeting efficiency, <i>in vivo</i> delivery
Small molecules	RNA binding	Convenient <i>in vivo</i> delivery	Nonspecificity, nonprogrammable targeting

Abbreviations: ADAR, adenosine deaminase acting on RNA; Cas9, CRISPR-associated 9; dCas13, catalytically inactive PspCas13b; PspCas13b, CRISPR-associated 13 from *Prevotella sp. P5-125*; REPAIR, RNA Editing for Programmable A to I Replacement.

development of the Cas9 system was reported by Liu group, which enables conversion of cytidine to uridine, thereby leading to a C → T (or G → A) substitution, by fusing the Cas9 protein with a cytidine deaminase enzyme. This DNA base editing system can mediate base conversion in a programmable manner, without requiring double-strand DNA cleavage or DNA repairing templates (82). Thus this modified Cas9 system bears great potential especially in correcting tumorigenic SNPs on lncRNA genes with higher efficiency and accuracy. In addition, currently the *in vivo* delivery methods of the Cas9 system are developing very fast. These include viral vectors of lentivirus (LV) and adeno-associated viruses (AAV) for delivering DNAs encoding Cas9 and sgRNA (83–86); nonviral vectors of lipid-based carriers and polymers for delivering the DNA, mRNA, and proteins of the Cas9 system (87–89); and physical approaches, such as microinjection, electroporation, and hydrodynamic injection, for delivering DNA, mRNA, and proteins of the Cas9 system (88, 90). These versatile approaches enable efficient *in vivo* delivery of the Cas9 system for lncRNA targeting in cancer treatment in the future. However, the Cas9 strategy for lncRNA knockout is genomic context dependent. It is well known that numerous lncRNAs are transcribed from bidirectional promoters, or overlap with promoters or bodies of sense or antisense genes. This unique feature of lncRNAs may expose their neighboring genes to the risk of inadvertent deregulation when using the Cas9 system for lncRNA gene knockout purposes. A study showed that only one third of 15,929 lncRNA loci may be safely targeted for gene knockout via the Cas9 system without perturbation of neighboring genes (91). Thus, careful study is needed before adopting the Cas9 system for targeting of any lncRNA gene.

In addition to the Cas9 system, novel CRISPR-Cas systems capable of directly targeting RNA, rather than working through DNA recognition, have been established. The class 2 type VI CRISPR-Cas effectors, Cas13a and Cas13b, were recently identified by Zhang group as single-component programmable RNA-guided RNA-targeting RNases that have both RNA processing and RNA cleavage activities (92, 93). Cas13b from *Prevotella sp. P5-125* (PspCas13b) is an RNA-targeted RNase which does not require a PAMmer for target RNA recognition. PspCas13b displays consistent, robust, and specific knockdown of a reporter mRNA in mammalian cells. Thus, it holds great promise as a powerful tool for knocking down lncRNA transcripts in cancer therapy, especially for oncogenic lncRNAs with higher expression levels in cancer cells than in normal tissues and which are unsuitable for gene knockout by the traditional Cas9 system. Furthermore, the PspCas13b system was modified to an RNA-editing system by fusing the catalytically inactive PspCas13b (dCas13b) with the catalytic domain of adenosine (A) to inosine (I) deaminase 2 (ADAR2), namely RNA Editing for Programmable A to I Replace-

ment (REPAIR; refs. 92, 93). The REPAIR system can convert A to I in mRNA molecules without disturbing genomic sequences, which avoids unexpected effects such as frameshift and nonsense mutations of the genome. Furthermore, this system is theoretically capable of transformation to other types of base editing platforms, e.g., C-to-U editing, if dCas13b is fused with other RNA base editors. Collectively, the REPAIR system offers advantages in the programmable editing of a specific nucleotide without sequence constraints (no PAM requirement), and independence from endogenous DNA repair pathways, which are usually needed for the Cas9 system. Hopefully, the REPAIR system will function as an accurate tool for editing lncRNAs in cancer. For example, the REPAIR system may reduce cancer risk by altering a single nucleotide in tumorigenic SNPs of lncRNAs (54, 58). It might also be used to regulate the association of lncRNAs and onco-protein partners by changing specific nucleotides critical for the interaction.

Previous studies have found that small molecules represent a huge source for drugging proteins. It has always been an interesting question as to whether RNAs can also directly bind to small-molecule compounds. Previously, scientists have been pessimistic about the answer until a recent breakthrough in the field (94). A small molecule LMI070 was found to bind to the pre-mRNA of spinal muscular atrophy-related gene *SMN2* which encodes survival motor neuron. The association boosts the processing of exons and translation of the protein product to antagonize the disease (95). Later, by screening a library of noncoding RNAs against a library of small molecules to find strong interactions, another small molecule, targaprimir-96, was identified as a binding partner of primary miRNA-96 at a key processing site. This interaction blocks the maturation of mir-96 and induces apoptosis in cancer cells (96). Using the same strategy, authors also found a small molecule, targaprimir-210, which binds primary miR-210 (97). These inspiring results demonstrate that both coding and noncoding RNAs are druggable just like proteins. Although no relevant studies on drugging lncRNAs have been reported, these studies provide one more prospective strategy for targeting lncRNAs for cancer therapy.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We apologize to colleagues whose work was not discussed or cited in this review due to the space limitation. This work was supported, in whole or in part, by the US Department of Defense (PC140683 to C.V. Dang), the National Key Research and Development Program of China (2017YFA0105501 to X. Zhong), the Guangdong Province Science and Technology Project (2015A020212019 to X. Zhong), the US National Institutes of Health (R01CA142776 to L. Zhang,

R01CA190415 to L. Zhang, P50CA083638 to L. Zhang, P50CA174523 to L. Zhang, P50CA083639 to A.K. Sood, P50CA098258 to A.K. Sood, and R35 CA209904 to A.K. Sood), the Breast Cancer Alliance (L. Zhang and C.V. Dang), the Frank McGraw Memorial Chair in Cancer Research (A.K. Sood), the American Cancer Society Research Professor Award (A.K. Sood), the Marsha Rivkin Center for Ovarian Cancer Research (L. Zhang), the Bassler Center for

BRCA (L. Zhang), the Harry Fields Professorship (L. Zhang), and the Kaleidoscope of Hope Ovarian Cancer Foundation (L. Zhang).

Received February 2, 2018; revised May 9, 2018; accepted July 5, 2018; published first September 4, 2018.

References

- 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;22:1775–89.
- Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. *Nature* 2012;489:101–8.
- 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;458:223–7.
- Lee JT. Epigenetic regulation by long noncoding RNAs. *Science* 2012;338:1435–9.
- Lieberman J, Slack F, Pandolfi PP, Chinnaiyan A, Agami R, Mendell JT. Noncoding RNAs and cancer. *Cell* 2013;153:9–10.
- Orom UA, Shiekhattar R. Long noncoding RNAs usher in a new era in the biology of enhancers. *Cell* 2013;154:1190–3.
- Prensner JR, Chinnaiyan AM. The emergence of lncRNAs in cancer biology. *Cancer Discov* 2011;1:391–407.
- Ulitisky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. *Cell* 2013;154:26–46.
- Brown CJ, Ballabio A, Rupert JL, Lafreniere RC, Grompe M, Tonlorenzi R, et al. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 1991;349:38–44.
- Pachnis V, Brannan CI, Tilghman SM. The structure and expression of a novel gene activated in early mouse embryogenesis. *EMBO J* 1988;7:673–81.
- Bartolomei MS, Zemel S, Tilghman SM. Parental imprinting of the mouse H19 gene. *Nature* 1991;351:153–5.
- Kapranov P, Cawley SE, Drenkow J, Bekiranov S, Strausberg RL, Fodor SP, et al. Large-scale transcriptional activity in chromosomes 21 and 22. *Science* 2002;296:916–9.
- Rinn JL, Euskirchen G, Bertone P, Martone R, Luscombe NM, Hartman S, et al. The transcriptional activity of human chromosome 22. *Genes Develop* 2003;17:529–40.
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 2007;129:1311–23.
- Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 2010;464:1071–6.
- Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A* 2009;106:11667–72.
- Maeda N, Kasukawa T, Oyama R, Gough J, Frith M, Engstrom PG, et al. Transcript annotation in FANTOM3: mouse gene catalog based on physical cDNAs. *PLoS genetics* 2006;2:e62.
- 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;16:1478–87.
- Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, et al. Long noncoding RNAs with enhancer-like function in human cells. *Cell* 2010;143:46–58.
- Prensner JR, Iyer MK, Balbin OA, Dhanasekaran SM, Cao Q, Brenner JC, et al. Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. *Nat Biotechnol* 2011;29:742–9.
- Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Develop* 2011;25:1915–27.
- Brannan CI, Dees EC, Ingram RS, Tilghman SM. The product of the H19 gene may function as an RNA. *Mol Cell Biol* 1990;10:28–36.
- Dieci G, Fiorino G, Castelnuovo M, Teichmann M, Pagano A. The expanding RNA polymerase III transcriptome. *Trends Genet* 2007;23:614–22.
- Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 2007;316:1484–8.
- Yin QF, Yang L, Zhang Y, Xiang JF, Wu YW, Carmichael GC, et al. Long noncoding RNAs with snoRNA ends. *Mol Cell* 2012;48:219–30.
- Yoon JH, Abdelmohsen K, Srikantan S, Yang X, Martindale JLDe S, et al. lincRNA-p21 suppresses target mRNA translation. *Mol Cell* 2012;47:648–55.
- Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, et al. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 2010;329:689–93.
- Wang D, Garcia-Bassets I, Benner C, Li W, Su X, Zhou Y, et al. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature* 2011;474:390–4.
- 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.
- 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;309:1570–3.
- Hung T, Wang Y, Lin MF, Koegel AK, Kotake Y, Grant GD, et al. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat Genet* 2011;43:621–9.
- Yang L, Lin C, Liu W, Zhang J, Ohgi KA, Grinstein JD, et al. ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. *Cell* 2011;147:773–88.
- 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;39:925–38.
- 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;147:358–69.
- Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 2010;465:1033–8.
- Gumireddy K, Li A, Yan J, Setoyama T, Johannes GJ, Orom UA, et al. Identification of a long non-coding RNA-associated RNP complex regulating metastasis at the translational step. *EMBO J* 2013;32:2672–84.
- Bazzini AA, Johnstone TG, Christiano R, Mackowiak SD, Obermayer B, Fleming ES, et al. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *EMBO J* 2014;33:981–93.
- Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 2011;147:789–802.
- Juntawong P, Girke T, Bazin J, Bailey-Serres J. Translational dynamics revealed by genome-wide profiling of ribosome footprints in Arabidopsis. *Proc Natl Acad Sci U S A* 2014;111:E203–12.
- Ruiz-Orera J, Messegue X, Subirana JA, Alba MM. Long non-coding RNAs as a source of new peptides. *Elife* 2014;3:e03523.
- Pauli A, Norris ML, Valen E, Chew GL, Gagnon JA, Zimmerman S, et al. Toddler: an embryonic signal that promotes cell movement via Apelin receptors. *Science* 2014;343:1248636.
- Guttman M, Russell P, Ingolia NT, Weissman JS, Lander ES. Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Cell* 2013;154:240–51.
- Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 2011;477:295–300.

44. Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D, et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* 2010;142:409–19.
45. Rupaimoole R, Lee J, Haemmerle M, Ling H, Previs RA, Pradeep S, et al. Long noncoding RNA ceruloplasmin promotes cancer growth by altering glycolysis. *Cell reports* 2015;13:2395–402.
46. Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, et al. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res* 2011;71:6320–6.
47. Niinuma T, Suzuki H, Nojima M, Noshio K, Yamamoto H, Takamaru H, et al. Upregulation of miR-196a and HOTAIR drive malignant character in gastrointestinal stromal tumors. *Cancer Res* 2012;72:1126–36.
48. Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, et al. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* 2010;38:662–74.
49. Kotake Y, Nakagawa T, Kitagawa K, Suzuki S, Liu N, Kitagawa M, et al. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. *Oncogene* 2011;30:1956–62.
50. Hu X, Feng Y, Zhang D, Zhao SD, Hu Z, Greshock J, et al. A functional genomic approach identifies FAL1 as an oncogenic long noncoding RNA that associates with BMI1 and represses p21 expression in cancer. *Cancer Cell* 2014;26:344–57.
51. Lu Y, Hu Z, Mangala LS, Stine ZE, Hu X, Jiang D, et al. MYC targeted long noncoding RNA DANCR promotes cancer in part by reducing p21 levels. *Cancer Res* 2018;78:64–74.
52. Zhang Y, He Q, Hu Z, Feng Y, Fan L, Tang Z, et al. Long noncoding RNA LINP1 regulates repair of DNA double-strand breaks in triple-negative breast cancer. *Nat Struct Mol Biol* 2016;23:522–30.
53. Ji P, Diederichs S, Wang W, Boing S, Metzger R, Schneider PM, et al. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* 2003;22:8031–41.
54. Zheng J, Huang X, Tan W, Yu D, Du Z, Chang J, et al. Pancreatic cancer risk variant in LINC00673 creates a miR-1231 binding site and interferes with PTPN11 degradation. *Nat Genet* 2016;48:747–57.
55. Yan X, Hu Z, Feng Y, Hu X, Yuan J, Zhao SD, et al. Comprehensive genomic characterization of long non-coding RNAs across human cancers. *Cancer Cell* 2015;28:529–40.
56. Leucci E, Vendramin R, Spinazzi M, Laurette P, Fiers M, Wouters J, et al. Melanoma addiction to the long non-coding RNA SAMMSON. *Nature* 2016;531:518–22.
57. Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, et al. The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet* 2015;47:199–208.
58. Pandey GK, Mitra S, Subhash S, Hertwig F, Kanduri M, Mishra K, et al. The risk-associated long noncoding RNA NBAT-1 controls neuroblastoma progression by regulating cell proliferation and neuronal differentiation. *Cancer Cell* 2014;26:722–37.
59. Marchese FP, Grossi E, Marin-Bejar O, Bharti SK, Raimondi I, Gonzalez J, et al. A long noncoding RNA regulates sister chromatid cohesion. *Mol Cell* 2016;63:397–407.
60. Schmitt AM, Garcia JT, Hung T, Flynn RA, Shen Y, Qu K, et al. An inducible long noncoding RNA amplifies DNA damage signaling. *Nat Genet* 2016;48:1370–6.
61. Zhang D, Zhang G, Hu X, Wu L, Feng Y, He S, et al. Oncogenic RAS regulates long noncoding RNA orilnc1 in human cancer. *Cancer Res* 2017;77:3745–57.
62. Bussemakers MJ, van Bokhoven A, Verhaegh GW, Smit FP, Karthaus HF, Schalken JA, et al. DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer Res* 1999;59:5975–9.
63. Panzitt K, Tschernatsch MM, Guelly C, Moustafa T, Stradner M, Strohmaier HM, et al. Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. *Gastroenterology* 2007;132:330–42.
64. Shao Y, Ye M, Jiang X, Sun W, Ding X, Liu Z, et al. Gastric juice long noncoding RNA used as a tumor marker for screening gastric cancer. *Cancer* 2014;120:3320–8.
65. Cai B, Wu Z, Liao K, Zhang S. Long noncoding RNA HOTAIR can serve as a common molecular marker for lymph node metastasis: a meta-analysis. *Tumour Biol* 2014;35:8445–50.
66. Huang X, Yuan T, Tschannen M, Sun Z, Jacob H, Du M, et al. Characterization of human plasma-derived exosomal RNAs by deep sequencing. *BMC Genomics* 2013;14:319.
67. Li Q, Shao Y, Zhang X, Zheng T, Miao M, Qin L, et al. Plasma long noncoding RNA protected by exosomes as a potential stable biomarker for gastric cancer. *Tumour Biol* 2015;36:2007–12.
68. Qu L, Ding J, Chen C, Wu ZJ, Liu B, Gao Y, et al. Exosome-Transmitted lncARSR promotes sunitinib resistance in renal cancer by acting as a competing endogenous RNA. *Cancer Cell* 2016;29:653–68.
69. Ren S, Peng Z, Mao JH, Yu Y, Yin C, Gao X, et al. RNA-seq analysis of prostate cancer in the Chinese population identifies recurrent gene fusions, cancer-associated long noncoding RNAs and aberrant alternative splicings. *Cell Res* 2012;22:806–21.
70. Arun G, Diermeier S, Akerman M, Chang KC, Wilkinson JE, Hearn S, et al. Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA loss. *Genes Dev* 2016;30:34–51.
71. Gutschner T, Hammerle M, Eissmann M, Hsu J, Kim Y, Hung G, et al. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res* 2013;73:1180–9.
72. Wu SY, Lopez-Berestein G, Calin GA, Sood AK. RNAi therapies: drugging the undruggable. *Sci Translat Med* 2014;6:240ps247.
73. Wu SY, Yang X, Gharpure KM, Hatakeyama H, Egli M, McGuire MH, et al. 2'-Ome-phosphorodithioate-modified siRNAs show increased loading into the RISC complex and enhanced anti-tumour activity. *Nat Commun* 2014;5:3459.
74. Nishimura M, Jung EJ, Shah MY, Lu C, Spizzo R, Shimizu M, et al. Therapeutic synergy between microRNA and siRNA in ovarian cancer treatment. *Cancer Discov* 2013;3:1302–15.
75. Ma L, Reinhardt F, Pan E, Soutschek J, Bhat B, Marcusson EG, et al. Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model. *Nat Biotechnol* 2010;28:341–7.
76. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013;339:819–23.
77. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 2014;157:1262–78.
78. Komor AC, Badran AH, Liu DR. CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell* 2017;168:20–36.
79. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. *Science* 2013;339:823–6.
80. Ho TT, Zhou N, Huang J, Koirala P, Xu M, Fung R, et al. Targeting non-coding RNAs with the CRISPR/Cas9 system in human cell lines. *Nucleic Acids Res* 2015;43:e17.
81. Zhu S, Li W, Liu J, Chen CH, Liao Q, Xu P, et al. Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. *Nat Biotechnol* 2016;34:1279–86.
82. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 2016;533:420–4.
83. Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera Mdel C, Yusa K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol* 2014;32:267–73.
84. Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* 2014;159:440–55.
85. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen T, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 2014;343:84–7.
86. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* 2014;343:80–4.
87. Liang X, Potter J, Kumar S, Zou Y, Quintanilla R, Sridharan M, et al. Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J Biotechnol* 2015;208:44–53.
88. Zuckerman M, Hovestadt V, Knobbe-Thomsen CB, Zapotka M, Northcott PA, Schramm K, et al. Somatic CRISPR/Cas9-mediated tumour suppressor disruption enables versatile brain tumour modelling. *Nat Commun* 2015;6:7391.

89. Zuris JA, Thompson DB, Shu Y, Guilinger JP, Bessen JL, Hu JH, et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat Biotechnol* 2015;33:73–80.
90. Han X, Liu Z, Jo MC, Zhang K, Li Y, Zeng Z, et al. CRISPR-Cas9 delivery to hard-to-transfect cells via membrane deformation. *Sci Adv* 2015;1:e1500454.
91. Goyal A, Myacheva K, Gross M, Klingenberg M, Duran Arque B, Diederichs S. Challenges of CRISPR/Cas9 applications for long non-coding RNA genes. *Nucleic Acids Res* 2017;45:e12.
92. Cox DBT, Gootenberg JS, Abudayyeh OO, Franklin B, Kellner MJ, Joung J, et al. RNA editing with CRISPR-Cas13. *Science* 2017;358:1019–27.
93. Yang L, Chen LL. Enhancing the RNA engineering toolkit. *Science* 2017;358:996–7.
94. Mullard A. Small molecules against RNA targets attract big backers. *Nat Rev Drug Discov* 2017;16:813–5.
95. Palacino J, Swalley SE, Song C, Cheung AK, Shu L, Zhang X, et al. SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. *Nat Chem Biol* 2015;11:511–7.
96. Velagapudi SP, Cameron MD, Haga CL, Rosenberg LH, Lafitte M, Duckett DR, et al. Design of a small molecule against an oncogenic noncoding RNA. *Proc Natl Acad Sci U S A* 2016;113:5898–903.
97. Costales MG, Haga CL, Velagapudi SP, Childs-Disney JL, Phinney DG, Disney MD. Small molecule inhibition of microRNA-210 reprograms an oncogenic hypoxic circuit. *J Am Chem Soc* 2017;139:3446–55.

Molecular Cancer Therapeutics

Shedding Light on the Dark Cancer Genomes: Long Noncoding RNAs as Novel Biomarkers and Potential Therapeutic Targets for Cancer

Lin Zhang, Dan Peng, Anil K. Sood, et al.

Mol Cancer Ther 2018;17:1816-1823.

Updated version Access the most recent version of this article at:
<http://mct.aacrjournals.org/content/17/9/1816>

Cited articles This article cites 97 articles, 32 of which you can access for free at:
<http://mct.aacrjournals.org/content/17/9/1816.full#ref-list-1>

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link <http://mct.aacrjournals.org/content/17/9/1816>. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.