Beyond classic editing: innovative CRISPR approaches for functional studies of long non-coding RNA

Dahlia A. Awwad*

Center of X-Ray Determination of Structure of Matter (CXDS), Helmi Institute of Biomedical Research, Zewail City of Science and Technology, Giza, Cairo, Egypt

*Correspondence address: Center of X-Ray Determination of Structure of Matter (CXDS), Helmi Institute of Biomedical Research, Zewail City of Science and Technology, Giza, Cairo 11461, Egypt. E-mail: dawwad@zewailcity.edu.eg

Abstract

Long non-coding RNAs (lncRNAs) makeup a considerable part of the non-coding human genome and had been well-established as crucial players in an array of biological processes. In spite of their abundance and versatile roles, their functional characteristics remain largely undiscovered mainly due to the lack of suitable genetic manipulation tools. The emerging CRISPR/Cas9 technology has been widely adapted in several studies that aim to screen and identify novel lncRNAs as well as interrogate the functional properties of specific lncRNAs. However, the complexity of lncRNAs genes and the regulatory mechanisms that govern their transcription, as well as their unique functionality pose several limitations the utilization of classic CRISPR methods in lncRNAs functional studies. Here, we overview the unique characteristics of lncRNAs transcription and function and the suitability of the CRISPR toolbox for applications in functional characterization of lncRNAs. We discuss some of the novel variations to the classic CRISPR/Cas9 system that have been tailored and applied previously to study several aspects of lncRNAs functionality. Finally, we share perspectives on the potential applications of various CRISPR systems, including RNA-targeting, in the direct editing and manipulation of lncRNAs.

Keywords: Long non-coding RNA; CRISPR/Cas9; lncRNA function; Gene editing

Introduction

Although non-coding regions make up around 97% of the human genome, little is known about these regions functionality [1, 2]. Large-scale biochemical studies such as Encyclopaedia of DNA Elements (ENCODE) project and Road-map Epigenomics, indicate that the majority of non-coding DNA is functional [3, 4]. Despite lacking a concrete universal definition, long non-coding RNAs (lncRNAs) are commonly defined as transcripts that are longer than 200 nt with no or limited protein-coding potential and are highly tissue-specific [5]. They share some characteristics with protein-coding RNAs such as the 5' caps, the 3' poly-A tail, as well as similar histone modifications profile and splicing mechanisms [6, 7]. LncRNAs could regulate gene expression by several mechanisms, including acting as scaffolds, decoys, guides and signals [7]. In addition, a number of lncRNAs exert their effects by the mere act of their transcription [8, 9]. The human genome is rich in lncRNAs; GENCODE v26 (a manually curated database of lncRNAs) contains 15 787 annotate lncRNA genes and 2720 lncRNA transcripts [10]. However, it has been difficult to identify the genomic loci of lncRNAs as well as dissect their functionality or their interaction with other molecular pathways until recently, largely due to the challenges associated with manipulating their expression [2, 11].

Several genomic editing tools had recently emerged, such as zinc-finger nucleases (ZFNs) [12] and transcription activation...
Clustered regularly interspaced palindromic repeats (CRISPR) systems were initially identified as means of bacterial adaptive immunity. They generally comprise arrays of DNA repeats interspersed with sequences that had been acquired from invading organisms, such as phages [19, 20]. Among these systems, the type II CRISPR/Cas9 system from Staphylococcus pyogenes is the most widely studied and utilized in genomic editing. The system in essence consists of two major components: the Cas9 nuclease, guided by a crisprRNA (crRNA) and a tracrRNA that together form the guide RNA (gRNA) duplex [21]. Specific recognition and cleavage of the invading DNA by the gRNA–Cas9 complex is facilitated by the presence of a protospacer adjacent motif (PAM), a sequence of 2–6 nucleotides that is present exclusively in the viral DNA but not the bacterial. Ever since its discovery, the CRISPR/Cas9 system has been employed by several groups as a genomic engineering tool, due to the unprecedented ability of the gRNA–Cas9 complex to target and cleave genomic regions in a sequence-specific manner [22–26]. In genomic editing applications, a gRNA sequence that is complementary to a given target genomic sequence is designed, and the CRISPR/Cas9 machinery [34]. CRISPR/Cas9 is also effective in producing homology-directed repair (HDR) pathway [27, 28], although newly emerging variations of these DSBs are repaired through the homology-directed repair thereby causing double-stranded breaks (DSBs). Traditionally, the CRISPRn HR, depends on the HDR when repairing the DSBs and is used in gene corrections, gene knock-in or overexpression, targeting as well as knock-out [37, 38]. Deletion of certain DNA stretches could also be achieved using Cas9 by inducing multiple DSBs, an approach referred to as CRISPRn excision [39, 40] (Fig. 1b and c). Finally, a version of Cas9 that lacks nuclease activity (deactivated or ‘dead’ Cas9; dCas9) while maintaining the RNA-dependent recognition of DNA could be fused with functional domains, thereby producing customized transcription factors. In bacteria, recruiting dCas9 to a promoter region is sufficient to create steric hindrance that might obstruct proper functioning of transcription machinery, hence causing reduced expression [41]. In eukaryotic cells, however, dCas9 should be combined with additional inhibitory domains, such as the KRAB (Kruppel-associated box) domain of ZNF10, in order to form a potent transcription inhibition complex (CRISPR interference or CRISPRi) [42]. Similarly, fusing dCas9 with activator domains such as p65, VP64 or Rta results in activating the targeted genes in cis (CRISPR activation or CRISPRa) [43–45] (Fig. 1d). In the context of lncRNA functional studies, the CRISPRi/a approaches have major advantages: first is the ability to detect in cis effects which is not possible when using plasmid-based overexpression or inhibition by RNAi, both produce in trans effects. A second advantage is the ability to activate endogenous promoters through CRISPRa, producing variant transcripts that are often non-coding.

Classic CRISPR editing and the complex architecture of lncRNA genomic loci

The genomic regions coding for lncRNAs are distributed over the whole genome, including intra- and intergenic regions [5]. Long intergenic non-coding RNAs (lincRNAs) are the subset of lncRNAs that are produced from non-coding regions between two coding genes. They are either produced from intergenic exclusive promoters, or from bidirectional promoters that could be shared with other coding or non-coding genes [46, 47] (Fig. 2).

On the other hand, ‘internal’ lncRNAs lie fully within the body of other ‘host’ genes [48–50] (Fig. 2i–iv). Internal lncRNAs could be transcribed from the coding strand, in which case they are called ‘sense’, and usually share exons with the protein-coding genes, either partially overlapping or covering the entire length of their host genes. Inversely, internal lncRNAs that arise from antisense strands of protein-coding genes are named ‘antisense’ lncRNAs. According to GENCODE [5], antisense lncRNAs could fall under one of three categories, (i) lncRNA exon over-lapping with a part of a sense gene, (ii) non-coding transcripts that span the whole sequence of sense gene, or (iii) the whole transcription unit of the lncRNA is embedded within an intron of a coding host gene, in the latter case they are termed ‘intronic’. Inversely, coding genes that are located within lincRNAs introns are called ‘overlapping’ genes [47, 50]. Both sense and antisense lncRNAs could comprise more than one exon [51, 52]. In addition to intergenic promoters, intragenic lncRNAs might be regulated by promoters that lie within gene bodies and therefore termed ‘internal’ promoters [46, 53, 54]. Expectedly, compared to lincRNAs that were extensively studied [55–58], only a small fraction of intronic lncRNAs are deeply explored due to the risk of disrupting the expression of their host genes [50]. In the mouse genome, around 87% of genes produce antisense transcripts [59], while 23% of human lncRNAs are produced from the antisense strands of coding genes [5], indicating an important functional role of antisense lncRNAs in regulating gene expression.
This intricate architecture of lncRNA genomic loci (outlined in Fig. 2), in addition to the enigmatic mechanisms of lncRNAs functions pose several limitations on the utilization of the CRISPR toolbox in the functional characterization of lncRNAs. For example, methods that rely on NHEJ for single base mutagenesis or knocking out expression by inducing small frame-shift mutations (e.g. CRISPRn, base editing) are generally not applicable to lncRNAs genes, because the exact sequence motifs responsible for exerting their effect remain largely uncharacterized. Besides, if the lncRNA of interest exert their effect by the act of transcription per se, such mutations will not affect its function. In the latter case, however, it might be useful to target regulatory elements such as promoters, although for most lncRNAs those remain to be identified. In addition, due to potential intersections of promoters, this could affect the expression of other coding or non-coding regions which might

**Figure 1:** Classic CRISPR/Cas9 systems and editing of lncRNAs. (a) General scheme of the basic strategy of CRISPR-mediated double-stranded breaks. The guide RNA, which encompasses a scaffold tracrRNA bound to the Cas9 nuclease and a crRNA that recruits the gRNA-Cas9 complex to a specific genomic regions based on base complementarity. The PAM sequence (yellow) functions as a recognition and binding sequence for the Cas9 nuclease, resulting in a DSB in the genomic sequence; (b) Knocking out the expression of lncRNAs by eliminating a large portion of their transcription unit and/or promoter (red), usually using multiple gRNAs. (c) Introducing a donor DNA sequence, usually a transcription inhibition signal, this leads eventually to knocking out lncRNA expression. (d) Deactivated Cas9 (dCas9) fused with either repression or activation domains. When recruited to promoter regions, dCas9 could sterically interfere with the binding of transcriptional activation factors causing transcription inhibition. In mammalian cells, however, dCas9 alone was not as effective in inhibiting transcription as in other cell types (reference). Thus, fusing dCas9 with repressor or activator domains yields a transcription repression/activation complex that could effectively inhibit/activate gene expression in mammalian cells. This strategy was utilized to manipulate the expression of lncRNAs both in a targeted manner and in a wide-scale screening format.
A good solution might be to partially delete selected lncRNA promoters if they are internal or bidirectional. In these cases, regions cannot be applied to lncRNAs whose expression is controlled by internal promoters or promoters proximal to other genes without potently disrupting their sequence and expression. An added layer of complexity comes from the high frequency of sequence repeats within lncRNA loci [62], which could pose certain limitations on designing efficient gRNAs and shRNAs and is probably why RNAi and CRISPR are not so systematically used for the study of lncRNAs.

Functional lncRNA knockouts could be generated using CRISPR excision, by either deleting the target lncRNA promoters [61, 63–66] or deleting the entire lncRNA genes [61, 67]. However, excision of whole-length lncRNA would be inapplicable if the lncRNA loci intersect with other coding or non-coding regions. Similarly, it would be impossible to delete lncRNA promoters if they are internal or bidirectional. In these cases, a good solution might be to partially delete selected lncRNA exons that do not intersect with other genes and that are distant from their promoters [17], although it remains a possibility that the undeleted part would still contain functional domain(s). Moreover, removing only a distal part of the lncRNA but not its TSS might result in generating a new transcript. Deleting large genomic regions in general could lead to inadvertently deleting uncharacterized regulatory DNA elements that influence the expression of other genes, potentially giving rise to phenotypes falsely attributed to lncRNAs [47, 60, 61, 68, 69].

Therefore, applying CRISPR methods to interrogate lncRNAs requires specific knowledge about their genomic locus, their impact on other genes and whether they exert this impact in cis or in trans. In a key genome-wide analysis by Goyal et al. [47], they set out to systematically assess the efficiency of targeting lncRNAs by different CRISPR methods. They found that 62% of the total lncRNAs were deemed ‘non-CRISPRable’, either due to having internal promoters (35%) or bidirectional promoters (20%). Furthermore, targeting 15 929 lncRNA loci by CRISPR applications was specific in only 38% of them, while almost two-thirds were susceptible to inadvertently affecting neighbouring genes. Together, these results demonstrated that the complex organization of lncRNA genomic loci could greatly limit the potential to target them experimentally in a specific manner. The study concluded, however, with three recommendations to ensure accurate attribution of resulting phenotype to the targeted lncRNA in CRISPR-mediated editing experiments: (i) careful examination of the targeted locus when designing the gRNAs to avoid potential perturbations to neighboring genes, (ii) monitoring the expression of the surrounding genes throughout the experiment alongside that of the targeted lncRNA and (iii) a validation step using RNAi or antisense oligonucleotides (ASOs) should follow and the outcome phenotype should match that of the CRISPR editing step. Alternatively, the knockout (KO) phenotype should be rescued by exogenous expression.

In addition to the limitations of Cas9 targeting, applying dCas9-based methods such as CRISPRa/i to lncRNAs could also be problematic. This was highlighted in a recent study that systematically compared the non-specific targeting and the subsequent discrepancies in gene expression profiles of RNAi, ASO-locked nucleic acids (LNA) and CRISPRi using the lncRNA MALAT1 as an illustrative example [70]. Although CRISPRi showed the fewest off-target effects, the three methods resulted in different sets of differentially expressed genes and cellular phenotypes upon depletion.

Despite these limitations, CRISPR approaches have been widely used to investigate lncRNAs functions in biological processes, under both physiological and disease conditions, either through gRNA libraries-based functional screening or through targeting individual lncRNAs for deeper understanding of their contributions to specific phenotypes. Over the past few years, more than 300 lncRNAs have been targets for studies that employ CRISPR methods according to CRISPRInc, a database for manually curated lncRNA-targeting sgRNAs [71]. Beyond the aforementioned ‘classic’ CRISPR methods, several approaches have been tailored over the past years to suit the unique complexity of lncRNAs expression and function, and are utilized in lncRNA functional studies. This review highlights some of these adaptations and briefly discuss their applications in different areas of lncRNA functional studies such as tagging, manipulation of their expression and visualization, to name some examples. Finally, we share perspectives on using some of the newly
emerging RNA-targeting CRISPR systems in the direct editing of lncRNAs.

**Beyond classic editing of lncRNAs: innovative CRISPR approaches**

**Insertion of inhibitory and/or tagging signals**

The complex arrangement of genomic regions of some lncRNAs requires alternative approaches to the direct deletion of either whole length or parts of the lncRNA gene bodies. Using CRISPR/Cas9 system, silencing signals could be inserted in critical key regions along the lncRNA locus, thereby inhibiting its expression. Inserting a premature inhibitory signal, such as a polyadenylation (poly-A) site is an efficient and less invasive strategy to silence lncRNA expression that was applied to a number of lncRNAs in mice with considerable success [72–75]. This approach was recently utilized to silence the human lncRNA MALAT1, in addition to two coding genes [76].

The strategy is based on the biallelic integration of the poly-A signal into key sites of the target genomic regions via CRISPR/Cas9-induced HDR (Fig. 3b). The authors employed a double marker selection to screen for clonal cells with the poly (A) signal successfully integrated. In the case of MALAT1, the poly (A) signal was inserted immediately upstream from its promoter, resulting in efficient silencing and a sharp decrease in the MALAT1 transcript to 0.1% compared with the control cells. Interestingly, a recent study compared different CRISPR strategies of silencing lncRNAs in zebrafish concluded that the poly-A insertion was more efficient and led to complete inactivation of MALAT1 compared to deletion of either TSSs or highly-conserved sequence motifs, both resulted in attenuated expression due to usage of alternative TSSs and production of truncated transcripts, respectively [77].

Conversely, CRISPR-based targeted insertion through the NHEJ pathway in the absence of a homologous donor sequence was recently reported [30, 78], demonstrating the applicability of achieving targeted insertion with one universal vector. Gene-trap mutagenesis has been long established as an insertional mutagenesis strategy where gene inactivation and tagging of the disrupted gene with selection marker is achieved simultaneously [79]. The gene-trap method is based on hijacking the transcriptional regulatory elements of an endogenous gene to express both endogenous gene and the selection marker carried on the gene-trap vector and inserted at the insertion site (reviewed in [79] and [80]). The gene trap vectors usually contain a splice site upstream of a promoter-less reporter. Upon insertion, the cis regulatory elements of the trapped gene causes the expression of both the reporter and the trapped gene, resulting in the simultaneous inactivation of the targeted endogenous gene and the expression of a selection marker.

In a recent development, the CRISPR-mediated tagging and regulation of lncRNAs (CTRL) method was introduced to enable tagging and manipulating the expression of lncRNAs in mammalian cells [81]. The system in essence consists of a modified gene trap vector, which contains puromycin selection cassette and MS2 tagging sequence, and a plasmid that contains Cas9 and 2 sgRNA sequences (referred to as Cas9-2sgRNA), one is genome-targeting while the other is targeting the borders of the selection cassette thereby linearizing the trap vector (Fig. 3a). When introduced simultaneously to the mammalian cells, the selection cassette is inserted near the transcriptional

![Figure 3: Different variations of CRISPR/Cas9 and their applications in lncRNA functional studies.](image-url)

(a) The CRISPR-mediated tagging and regulation of lncRNAs (CTRL) method, where deactivation of lncRNA expression and tagging with a selection marker are achieved simultaneously. (b) Inhibition of lncRNA expression by inserting a poly-A signal immediately upstream from its promoter. (c) Inhibition of lncRNA expression by Cas9-mediated excision of TSSs identified by the overlap between DNase hypersensitivity sites (HSS) and the active histone mark H3K4me3. (d) The CRISPR display (CRISP-Disp) method, where the gRNA is fused to either whole ncRNAs or specific functional domains, and targeted to genomic regions to reveal the function of these ncRNAs/ domains in regulating gene expression. (e) Visualizing lncRNA by targeting with multiple gRNAs and a dCas9-GFP complex, followed by an imaging technique such as immunohistochemistry to enable imaging of the cell (e.g. imaging Xist).
termination sites of lncRNA genes inducing the expression of puromycin selection marker. At the same time, the insertion of an exogenous DNA fragment within the body of lncRNAs allows for examining the phenotype and the resulted functional disruption. For example, targeted insertion of puromycin cassette into transcription termination sites of the following six lncRNAs successfully upregulated their levels: HOX transcript antisense RNA (HOTAIR), taurine up-regulated 1 (TUG1), DICER1 antisense RNA 1 (Dicer1-AS1), ZEB1 antisense RNA1 (Zeb1-AS1), PtenP1 and myocardial infarction-associated transcript (MIAT).

Interestingly, targeted insertion into the TSSs of these lncRNAs with a modified promoter trap vector that carries an exogenous poly-A signal resulted in the inhibition of only two lncRNAs (TUG1 and DICER1-AS1) and stimulation of the expression of the rest, highlighting the complexity of the regulatory circuits that govern the expression of lncRNAs. Nonetheless, the method enables successful tagging of the lncRNAs at either 5’or 3’ as well as monitoring their expression status. Taken together, the CTRL system represents a valuable tool for comprehensive analysis of lncRNA functions and could be utilized in a high-throughput format for future functional screening.

Engineering gRNAs for lncRNA relocation (CRISPR-display)

Strategies of gRNA engineering include fusing gRNAs with additional aptamer (e.g. MS2) or RNA scaffolds, for recruitment of additional effector proteins (e.g. VP64, KRAB and APOBEC) and enhanced manipulation of gene expression [82-85]. The pioneering CRISPR-Display (CRISP-Disp) method is a recent variation of the classic CRISPR system that combines both the modular nature of lncRNA [86] and the high feasibility of sgRNA re-engineering. The method employs the basics of the CRISPR technology to investigate the functional relevance of regulatory domains in a wide range of non-coding RNA molecules, including lncRNAs [87]. The method shows for the first time that, in the dCas9-gRNA complex, the gRNA sequence could be fused to large exogenous RNA domains up to 4.8 kb long, around the length of natural lncRNAs, without affecting the integrity and efficiency of the dCas9 targeting (Fig. 3d). Reporter gene activation was measured using two assays: direct activation, which involves expression of dCas9-VP64 fusion protein, and ‘bridged’ activation, with co-expression of dCas9 and a separate complex of the aptamer-binding protein PP7 coupled with VP64. The direct activation assay demonstrated the efficiency of the targeting mechanism, while the bridged activation showed that the fused RNA accessory domains retained their functional integrity and accessibility to protein binding partners. The method was then applied to lncRNAs, by fusing the RepA domain of the lncRNA Xist to a gRNA that targets a reporter gene. Xist is 17 kb in length and acts in cis to inactivate one of the X chromosomes by recruiting epigenetic ‘writer’ complexes that lay repressive histone marks on the Xist-coated X chromosome [88]. Consistent with the canonical Xist function, the authors reported an observable repression of the reporter gene, indicating the inhibitory function of the RepA domain and that this function is independent from the X chromosome context. As such, further experimentation could follow, such as immunoprecipitation and mass spectrometry, to identify proteins associated with functional lncRNA domains [89]. In addition, the study identified several fusion topologies that allow for engineering gRNA-ncRNA complexes while maintaining efficient targeting of dCas9. The innovative CRISPR-Disp system enables repurposing and functional interrogation of either whole-length or partial domains of ncRNA, in addition to several potential applications in synthetic biology and structural characterization of ncRNAs [90]. It also allows for separately investigating the targeting and effector mechanisms of lncRNAs, and facilitates monitoring the transcription of a reported gene with concomitant imaging of another DNA locus, thus allowing for simultaneous analysis of several targets based on the available RNA motifs [91].

Epigenetic silencing of lncRNA

As discussed previously, precise knockout of lncRNAs through CRISPR-mediated frame-shift mutagenesis within genomic loci could be rather challenging. Instead, inhibiting the expression of lncRNA via targeting active histone marks is a plausible alternative. Modulating the epigenome is a well-established application of CRISPR where dCas9 is traditionally fused with an epigenetic effector protein (such as DNA methyltransferase or histone modifiers) and targeted to specific genomic loci (reviewed in [92]).

Recently, Janga et al. [93] developed for the first time a universal CRISPR-mediated knockout approach guided by epigenetic marks, that enables robust silencing of ncRNA loci through the excision of active histone marks (Fig. 3c). The method builds on previous work showing an overlap between the active mark histone H3 lysine 4 tri-methylation (H3K4me3) and DNase I hypersensitivity around the TSSs of most genes, and thus enables targeting loci that are poorly annotated [94]. Excision of TSS-associated active epigenetic signatures of ncRNAs genes successfully inhibited several microRNAs as well as the lncRNA MALAT1, which is constitutively transcribed in human monocytes. Although the results confirmed a role of miR-146a and miR-155 in regulating inflammatory response, they indicated no significant role of MALAT1. This epigenetically guided CRISPR-based approach represents a ncRNA knockout strategy with minimal alteration of the genomic sequence. By capitalizing on this approach, improved CRISPR libraries that alter the epigenetic code rather than the underlying sequences could be developed for systematic loss-of-function screenings of lncRNAs.

Visualization of lncRNAs

Fusion of DNA-binding proteins and fluorescent proteins (e.g. GFP) enables in principle imaging of specific genomic loci. Recently, several applications of CRISPR technology have enabled chromatin visualization via linking dCas9 to a fluorescent tag either through direct protein fusion or through an RNA scaffold. Using GFP–dCas9 fusion protein combined with an array of gRNAs tiled over a certain genomic locus enables visualization of non-repetitive genomic regions, a strategy that was used with several loci (reviewed in [95]). Similar to DNA imaging, RNA imaging is seemingly feasible using conventional methods such as fluorescent in situ hybridization (FISH) [96]. Nonetheless, efficient visualization of lncRNAs could be rather challenging with complementary FISH probes. The IncRNA XIST, for example, appeared as a low-resolution ‘cloud’, due to its considerably large size and complex secondary structure as well as the condensed nature of the surrounding repressed chromatin that limit the access of FISH probes [97]. Following the strategy of using dCas9–GFP fusion protein, combined with multiple sgRNAs and immunofluorescence, Wasko et al. were able to visualize the prototypical IncRNA Xist in female fibroblast cell line [98] (Fig. 3e). Expectedly, the images showed nuclear co-localization of both Xist and the repressive histone marks H3K27me3.

A different research direction is to exploit the specificity of gRNA targeting of genomic regions to label RNA. A recent study
reported using such strategy to conduct an imaging-based pooled library screening using sgRNAs co-delivered to cells along with barcodes that are linked to reporter genes [99]. High-throughput identification of sgRNA followed by multiplexed error-robust fluorescence in situ hybridization (MERFISH) [100, 101] allowing for screening of RNA-binding protein. Notably, the screen revealed several modulators of the IncRNA MALAT1 localization in nuclear speckles, reflecting the potential application of this method to studies of the dynamic localization of IncRNAs to subcellular compartments as well as their interactions with other proteins.

Methods developed to visualize DNA and associated proteins in vivo are important to monitor the dynamic organization of chromatin within the cell nucleus. Similarly, live-cell imaging is crucial for RNA studies, albeit challenging because washing of unbound labeled probes could not be performed and hence it is necessary to increase the signal-to-background ratio. For that purpose, turn-on probes that fluoresce only upon intercalating with nucleic acids are more efficient, such as molecular beacons [102–104]. Alternatively, live-cell tracking and imaging of transgenic RNA is possible using MS2-MCP fluorescence system, which harnesses the specific binding between the phage MS2 coat protein (MCP) and the MS2 stem–loop RNA aptamer. The system in essence consists of GFP-tagged MCP protein in addition to multiple copies of the MCP-binding RNA aptamer, usually linked to the RNA of interest [105, 106]. Although the MS2-MCP fluorescence system was previously used in an attempt to visualize a transgenic Xist [107], it is obviously not efficient when visualizing endogenous IncRNAs. Another attempt took advantage of the sequence-specific RNA binding of the Pumilio homology domain (the PUF protein family) [108]. A Pumilio protein typically consists of eight domains, each specifically recognizes one nucleotide within an eight-nucleotide RNA sequence [109]. However, this approach is less common than MS2-MCP probably because engineering an efficient construct is rather laborious and time-consuming, limiting its potential applications in analyzing large transcripts or scaling up the techniques to transcriptome-wide studies.

Cas9 has been reportedly effective in targeting ssRNA with high specificity when combined with PAM-presenting PAM/RNA hybrid oligonucleotides (PAMmers). This RNA-targeting Cas9 (RCas9) system is a recent modification of the classic CRISPR/Cas9 system, where the PAM sequence is introduced in vitro as part of the PAMmRNA oligonucleotide [110]. A mismatch in the PAMmRNA allows for exclusive recognition of the target RNA and not the genomic DNA [111]. Capitalizing on this strategy, Nelles et al. [112] successfully used RCas9 to visualize and track RNA in live cells, by fusing dCas9 to a fluorescent protein (mCherry or GFP), a targeting sgRNA and a nuclear localization signal in addition to an exogenous PAMmRNA oligonucleotide. Although only mRNAs were tracked in that study, it would be interesting to explore potential applications of RCas9 in live-cell tracking and visualizing of IncRNAs. Nonetheless, proper visualization of IncRNAs might require using multiple copies of the RCas9–sgRNA complex in order to cover the whole length of molecule.

**Perspectives on post-transcriptional editing of IncRNAs using CRISPR systems**

Due to the aforementioned limitations of Cas9-based genomic editing, a plausible alternative might be to edit RNA molecules post-transcriptionally. Functional studies of IncRNAs often depend on loss-of-function or gain-of-function methods, whether in cell cultures or in animal models. RNA interference (RNAi) has traditionally been the method of choice to manipulate gene expression post-transcriptionally [113, 114]. Although RNAi has been efficiently utilized to deplete coding transcripts [115], it showed limited efficiency when used to manipulate IncRNA expression [116]. RNAi, which depends on short double-stranded RNA sequence to target and silence expression of target RNAs [117], had been utilized for functional studies of IncRNAs [47]. Nonetheless, there are several shortcomings for using RNAi to study IncRNAs; firstly, unlike mRNA, IncRNAs are mostly nuclear [118], and although the RNAi machinery was found to function in the nucleus [119], short interfering RNAs (siRNAs) designed to target nuclear IncRNAs showed limited efficiency [120]. Secondly, the function of some IncRNAs can be transcription-coupled, meaning that the mere act of transcription is responsible for exerting their function rather than the transcript itself [8, 121]. Lastly, some IncRNAs (e.g. MALAT1) are expressed at high levels rendering RNAi targeting and inhibition insufficient to produce a complete loss of function [122]. Alternative to RNAi, other posttranscriptional silencing approaches, such as RNA targeting and cleaving via antisense oligonucleotides (ASOs)-directed RNase H activity, can target nuclear IncRNAs with rather high success [120] and even delete nascent transcripts [53, 123], although they could efficiently be used for only a fraction of IncRNAs. A major drawback for RNA-targeting methods such as RNAi and ASOs is that their effects remain transient. Thus, to achieve long-term silencing, functional studies of IncRNAs still required alternative approaches that act on the genomic level. Although short RNAs were successfully used to achieve stable transcriptional gene silencing (TGS) via targeting proximal regions of genes [124, 125], only a few examples of RNA-induced TGS of IncRNA were reported [126, 127].

The RCas9 system was effective in targeting and cleaving single-stranded RNA (ssRNA) in vitro as shown by recent reports [110]. As in the DNA targeting CRISPRi/i system, fusing dCas9 to protein effectors could allow applications of the RCas9 system in RNA studies (reviewed in [128]). In the context of IncRNAs, fusing dCas9 with protein effectors known to interfere with a certain IncRNA could be a method of fine-tuning their level of expression. For example, fusing dCas9 with IncRNA-stabilizing signals may increase their half-life inside cells. It could be fused with epigenetic effectors to investigate the function of IncRNAs in the epigenetic modulation of gene expression. Engineering gRNA by fusing it with either whole IncRNAs or specific structural domains could represent another layer of customizing the RCas9 system to IncRNA interrogation. The aforementioned are few examples, although one could only imagine a myriad of applications for RNA-targeting CRISPR technology. Nonetheless, the off-target effects of RCas9 are yet to be evaluated, and extensive validation is needed before it could be used in vivo applications.

The newly emerging type VI CRISPR-Cas systems that recognize and, upon activation, degrade ssRNAs were recently discovered in bacteria as an adaptive immunity means [129–131]. The unique effector protein of the type VI-A CRISPR systems, Cas13 (formerly C2c2), had been recently adapted into multiple applications including nucleic acid detection [122], plant and mammalian RNA knockdown and tracking [133, 134]. In addition, a catalytically inactive form of Prevotella sp. Cas13b (PpCas13b) was fused with the deaminase domain (ADAR2Δc) of the adenosine deaminase acting on RNA (ADAR) cDNA, and efficiently used in RNA base editing, a system that was referred to as RNA Editing for Programmable A to I Replacement.
(REPAIR) [135]. Although the aforementioned RNA-targeting CRISPR systems were only used to target coding transcripts, it would be interesting to explore their applicability to edit lncRNAs which, if successful, could represent an attractive tool for dissecting the functions of lncRNAs with rather high specificity, as well as enable structural probing at a single nucleotide resolution. Alternatively, the type VI effectors could be adapted for high-throughput functional screens for lncRNAs (2). On the other hand, the recent development of CRISPR/Cas9 splicing manipulation tools, such as CRISPR-SKIP [136] and TAM [137], that are based on single base editing via cytidine deaminase, could enable the functional identification of lncRNA splice variants [138]. It will probably require combining multiple tools in order to fully elucidate lncRNA function.

**Conclusion**

Although classic methods of CRISPR editing could be efficient when studying some lncRNAs, the intricate architecture of lncRNA genes and their transcriptional regulatory circuits require innovative approaches for their functional studies. Several innovative adaptations to the classic CRISPR system are emerging and are being widely applied in various aspects of lncRNA functional studies such as tagging and expression manipulation and visualization of lncRNAs. In addition to targeting lncRNAs on the genomic level, newly emerging CRISPR methods that directly target RNA could facilitate presumptive direct editing and/or manipulation of lncRNAs, opening the door to a myriad of applications in the study of both the functional and structural aspects of lncRNA biology.

**Acknowledgments**

The author would like to thank Fareed Abouelela, CXDS, Egypt and Rachid Rahmoni, Center of Molecular Biophysics (CBM), CNRS, France for their valuable remarks, and Sara Almeida Barreira for kindly helping with the figures.

**References**

1. Ohno S. An argument for the genetic simplicity of man and other mammals. *J Hum Evol* 1972;1:651–62.
2. Montalbano A, Canver MC, Sanjana NE. High-throughput approaches to pinpoint function within the noncoding genome. *Molecular Cell* 2017;68:44–59.
3. ENCODE Project Consortium TEP, Consortium TEP. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;489:57–74.
4. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J et al. Integrative analysis of 111 reference human epigenomes. *Nature* 2015;518:317–29.
5. Derrien T, Johnson R, Bussotti G 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.
6. Carninci P, Kasukawa T, Katayama S et al. The transcriptional landscape of the mammalian genome. *Science* 2005;309:1559–63.
7. Wang KC, Chang HY. Molecular mechanisms of long non-coding RNAs. *Molecular Cell* 2011;43:904–14.
8. Kornienko AE, Guenzl PM, Barlow DP, Pauler FM. Gene regulation by the act of long non-coding RNA transcription. *BMC Biol* 2013;11:59.
9. Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet* 2016;17:47–62.
10. Harrow J, Frankish A, Gonzalez JM et al. GENCODE: the reference human genome annotation for the ENCODE project. *Genome Res* 2012;22:1760–74.
11. Joung J, Engeit JM, Konermann S et al. Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature* 2017;548:343–6.
12. Le Provost F, Lillico S, Passet B et al. Zinc finger nuclease technology heralds a new era in mammalian transgenesis. *Trends Biotechnol* 2010;28:134–41.
13. Oost JVD. New tool for genome surgery. *Science* 2013;339:768–70.
14. Kim YG, Cha J, ChandraSagar S. Hybrid restriction enzymes: zinc finger fusions to FokI cleavage domain. *Proc Natl Acad Sci* 1996;93:1156–60.
15. Wright DA, Li T, Yang B, Spalding MH. TALEN-mediated genome editing: prospects and perspectives. *Biochem J* 2014;462:15–24.
16. Gutschner T. Silencing long noncoding RNAs with genome-editing tools. In: Pruett-Miller S (ed.), *Chromosomal Mutagenesis*. Methods in Molecular Biology (Methods and Protocols). 1239. New York: Humana Press, 2015, 241–50.
17. Sauvageau M, Goff LA, Lodato S et al. Multiple knockout mouse models reveal lncRNAs are required for life and brain development. *Elife* 2013;2:e01749.
18. Balios S, Adamaki M, Kyriakopoulos AM et al. Role of the CRISPR system in controlling gene transcription and monitoring cell fate. *Mol Med Rep* 2018;17:1421–7.
19. Bhaya D, Davison M, Barrangou R. CRISPR-Cas systems in Bacteria and Archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet* 2011;45:273–97.
20. Horvath P, Barrangou R. CRISPR/Cas, the immune system of Bacteria and Archaea. *Science* 2010;327:167–70.
21. Wang H, La Russa M, Qi LS. CRISPR/Cas9 in genome editing and beyond. *Annu Rev Biochem* 2016;85:227–64.
22. Jinek M, Chylinski K, Fonfara I et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012;337:816–21.
23. Mali P, Yang L, Esvelt KM et al. RNA-guided human genome engineering via Cas9. *Science* 2013;339:823–6.
24. Shalem O, Sanjana NE, Hartenian E et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 2014;343:84–7.
25. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* 2014;343:80–4.
26. Korkmaz G, Lopes R, Ugalde AP et al. Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat Biotechnol* 2016;34:192–8.
27. Iyama T, Wilson DM. DNA repair mechanisms in dividing and non-dividing cells. *DNA Repair* 2013;12:620–36.
28. Ortwein A, Noordermeer SM, Wilson MD et al. A mechanism for the suppression of homologous recombination in G1 cells. *Nature* 2015;528:422–6.
29. Suzuki K, Izpisua Belmonte JC. In vivo genome editing via the HITI method as a tool for gene therapy. *J Hum Genet* 2018;63:157–64.
30. Suzuki K, Tsunekawa Y, Hernandez-Benitez R et al. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* 2016;540:144–9.
31. Auer TO, Durourre K, De Cian A et al. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res* 2014;24:142–53.
Hand2 expression and heart development. Nature 2016;539:433–6.
75. Ballarino M, Cipriano A, Tita R et al. Deficiency in the nuclear long noncoding RNA Charme causes myogenic defects and heart remodeling in mice. EMBO J 2018;37:e99697.
76. Liu Y, Han X, Yuan J et al. Biallelic insertion of a transcriptional terminator via the CRISPR/Cas9 system efficiently silences expression of protein-coding and non-coding RNA genes. J Biol Chem 2017;292:5624–33.
77. Lavalue P, Eckert H, Danny L et al. Strategies for genetic inactivation of long noncoding RNAs in zebrafish. RNA 2019;25:897–9. 069484.118.
78. Schmid-Burgk JL, Höning K, Ebert TS, Hornung V. CRISPaint allows modular base-specific gene tagging using a ligase-4-dependent mechanism. Nat Commun 2016;7:12338.
79. Stanford WL, Cohn JB, Cordes SP. Gene trap mutagenesis: past, present and beyond. Nat Rev Genet 2001;2:756–68.
80. Friedel RH, Soriano P. Gene trap mutagenesis in the mouse. Meth Enzymol 2010;77:243–69.
81. Cheng TL, Qiu Z. Long non-coding RNA tagging and non-coding RNA genes. J Biol Chem 2017;292:5624–33.
82. Dahlman JE, Abudayyeh OO, Joung J et al. Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. Nat Biotechnol 2015;33:1159–61.
83. Zalatan JG, Lee ME, Almeida R et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell 2015;160:339–50.
84. Hess GT, Frésard L, Han K et al. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. Nat Methods 2016;13:1036–42.
85. Kweon J, Jang AH, Kim DE et al. Fusion guide RNAs for orthogonal gene manipulation withCas9 and Cpf1. Nat Commun 2018;9:303.
86. Guttmann M, Rinn J. Modular regulatory principles of large non-coding RNAs. Nature 2012;482:339–46.
87. Shechner DM, Hacisuleyman E, Younger ST, Rinn J. Multiplexable, locus-specific targeting of long RNAs with CRISPR Display. Nat Methods 2015;12:664–70.
88. Zhao J, Sun BK, Erwin JA et al. Polycistrons proteins targeted by a short repeat RNA to the mouse X chromosome. Science 2008;322:750–6.
89. Perez-Pinera P, Jones MF, Lal A, Lu TK. Putting non-coding RNA on display with CRISPR. Mol Cell 2015;59:146–8.
90. Lin PC, Corn JE. Co-opting CRISPR to deliver functional RNAs. Nat Methods 2015;12:613–4.
91. Plummer RJ, Guo Y, Peng Y. A CRISPR reimagining: new twists and turns of CRISPR beyond the genome-engineering revolution. J Cell Biochem 2018;119:1299–308.
92. Pulecio J, Verma N, Mejía-Ramírez E et al. CRISPR/Cas9-based engineering of the epigene. Cell Stem Cell 2017;21:431–47.
93. Janga H, Aznaurova M, Boldt F et al. Cas9-mediated excision of proximal DNaseI(HFk4em3 signatures confers robust silencing of microRNA and long non-coding RNA genes. PLoS One 2018;13:e0193066.
94. Thurman RE, Rynes E, Humbert R et al. The accessible chromatin landscape of the human genome. Nature 2012;489:75–82.
95. Wu X, Mao S, Ying Y et al. Progress and challenges for live-cell imaging of genomic loci using CRISPR-based platforms. Genomics, Proteomics and Bioinformatics 2019;17:119–28.
96. Larsson C, Grundberg I, Söderberg O, Nilsson M. In situ detection and genotyping of individual mRNA molecules. Nat Methods 2010;7:395–7.
97. Clemson CM, McNeil JA, Willard HF, Lawrence JR. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/structure chromosome. J Cell Biol 1996;132:259–75.
98. Waiko U, Zheng Z, Bhatnagar S. Visualization of Xist long noncoding RNA with a fluorescent CRISPR/Cas9 system. Methods Mol Biol 2019;1870:41–50.
99. Wang C, Lu T, Emanuel G et al. Imaging-based pooled CRISPR screening reveals regulators of lncRNA localization. Proc Natl Acad Sci USA 2019;116:10842–51.
100. Moffitt JR, Hao J, Wang G et al. High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. Proc Natl Acad Sci USA 2016;113:11046–51.
101. Moffitt JR, Zhuang X. RNA imaging with multiplexed error-robust fluorescence in situ hybridization (MERFISH). Methods Enzymol 2016;572:1–49.
102. Sokol DL, Zhang X, Lu P, Gewirtz AM. Real time detection of DNA RNA hybridization in living cells. Proc Natl Acad Sci 2019;95:11538–43.
103. Köhler O, Jarijote DV, Seitz O. Forced intercalation probes (FIT Probes): Thiazole orange as a fluorescent base in peptide nucleic acids for homogeneous single-nucleotide polymorphism detection. Chem Bio Chem 2005;6:69–77.
104. Hövelmann F, Gaspar I, Ephrussi A, Seitz O. Brightness enhanced DNA FIT-probes for wash-free RNA imaging in tissue. J Am Chem Soc 2013;135:19025–32.
105. Fouts D. Functional recognition of fragmented operator sites by R17/MS2 coat protein, a translational repressor. Nucleic Acids Res 1997;25:4464–73.
106. Bertrand E, Chartrand P, Schaefer M et al. Localization of ASH1 mRNA particles in living yeast. Mol Cell 1998;2:437–45.
107. Ng K, Daigle N, Bancaud A et al. A system for imaging the regulatory noncoding Xist RNA in living mouse embryonic stem cells. MBoC 2011;22:2634–45.
108. Ha N, Lai LT, Chelliah R et al. Live-cell imaging and functional dissection of Xist RNA reveal mechanisms of X chromosome inactivation and reactivation. Science 2018;361:8–14.
109. Ozawa T, Natori Y, Sato M, Umezawa Y. Imaging dynamics of endogenous mitochondrial RNA in single living cells. Nat Methods 2007;4:413–9.
110. O’Connell MR, Oakes BL, Sternberg SH et al. Programmable RNA recognition and cleavage by CRISPR/Cas9. Nature 2014;516:263–6.
111. Sternberg SH, Redding S, Jinek M et al. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature 2014;507:62–7.
112. Nelles DA, Feng MY, O’Connell MR et al. Programmable RNA tracking in live cells with CRISPR/Cas9. Cell 2016;165:488–96.
113. Tuschl T, Zamore PD, Lehmann R et al. Targeted mRNA degradation by double-stranded RNA in vitro. Genes Dev 1999;13:3191–7.
114. Zamore PD, Tuschl T, Sharp PA, Bartel DP. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell 2000;101:25–33.
115. Barragou R, Birmingham A, Wiemann S et al. Advances in CRISPR-Cas9 genome engineering: lessons learned from RNA interference. Nucleic Acids Res 2015;43:3407–19.
116. Nhn K, Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet 2014;15:7–21.
117. Elbashir SM, Harborth J, Lendeckel W et al. Duplexes of 21+ nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 2001;411:494–8.
118. Cabili MN, Dunagin MC, McClanahan PD et al. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. Genome Biol 2015;16:20.

119. Gagnon KT, Li L, Chu Y et al. RNAi factors are present and active in human cell nuclei. Cell Rep 2014;6:211–21.

120. Lennox KA, Behlke MA. Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides. Nucleic Acids Res 2016;44:863–77.

121. Engreitz JM, Haines JE, Perez EM et al. Local regulation of gene expression by lncRNA promoters, transcription and splicing. Nature 2016;44:863–77.

122. Gutschner T, Baas M, Diederichs S. Noncoding RNA gene silencing through genomic integration of RNA destabilizing elements using zinc finger nucleases. Genome Res 2011;21:1944–54.

123. Vickers TA, Koo S, Bennett CF et al. Efficient reduction of target RNAs by small interfering RNA and RNase H-dependent antisense agents. A comparative analysis. J Biol Chem 2003;278:7108–18.

124. Morris KV, Chan SW, Jacobsen SE, Looney DJ. Small interfering RNA-induced transcriptional gene silencing in human cells. Science 2004;305:1289–92.

125. Weinberg MS, Morris KV. Transcriptional gene silencing in humans. Nucleic Acids Res 2016;44:6505–17.

126. Stojic L, Niemczyk M, Orjalo A et al. Transcriptional silencing of long noncoding RNA GNG12-AS1 uncouples its transcriptional and product-related functions. Nat Commun 2016;7:10406.

127. Golding MC, Magri LS, Zhang L et al. Depletion of Kcnq1ot1 non-coding RNA does not affect imprinting maintenance in stem cells. Development 2011;138:3667–78.

128. Nelles DA, Fang MY, Aigner S, Yeo GW. Applications of Cas9 as an RNA-programmed RNA-binding protein. BioEssays 2015;37:732–9.

129. Abudayyeh OO, Gootenberg JS, Konermann S et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science 2016;353:aa5573.

130. East-Seletsky A, O’Connell MR, Knight SC et al. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. Nature 2016;538:270–3.

131. Smargon AA, Cox DBT, Pyzocha NK et al. Cas13b is a type VI-B CRISPR-associated RNA-guided RNase differentially regulated by accessory proteins Csx27 and Csx28. Mol Cell 2017;65:618–30.

132. Gootenberg JS, Abudayyeh OO, Lee JW et al. Nucleic acid detection with CRISPR-Cas13a/C2c2. Science 2017;356:438–42.

133. Abudayyeh OO, Gootenberg JS, Essletzbichler P et al. RNA targeting with CRISPR-Cas13. Nature 2017;550:280–4.

134. East-Seletsky A, O’Connell MR, Burstein D et al. RNA targeting by functionally orthogonal type VI-A CRISPR-Cas enzymes. Mol Cell 2017;66:373–83.

135. Cox DBT, Gootenberg JS, Abudayyeh OO et al. RNA editing with CRISPR-Cas13. Science 2017;358:1019–27.

136. Gapinske M, Luu A, Winter J et al. CRISPR-SKIP: programmable gene splicing with single base editors. Genome Biol 2018;19:107.

137. Yuan J, Ma Y, Huang T et al. Genetic modulation of RNA splicing with a CRISPR-guided cytidine deaminase. Mol Cell 2018;72:380–94.

138. Sen R, Doose G, Stadler P. Rare splice variants in long non-coding RNAs. ncRNA 2017;3:23.