PORTLAND PRESS

Review Article

Functional annotation of lncRNA in high-throughput screening

Chi Wai Yip[1], Divya M. Sivaraman[1,2], Anika V. Prabhu[1] and ⓘ Jay W. Shin[1]

[1]RIKEN Center for Integrative Medical Sciences, Yokohama, Kanagawa 230-0045, Japan; [2]Department of Pathology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Kerala 695011, India

Correspondence: Jay W. Shin (jay.shin@riken.jp)

Recent efforts on the characterization of long non-coding RNAs (lncRNAs) revealed their functional roles in modulating diverse cellular processes. These include pluripotency maintenance, lineage commitment, carcinogenesis, and pathogenesis of various diseases. By interacting with DNA, RNA and protein, lncRNAs mediate multifaceted mechanisms to regulate transcription, RNA processing, RNA interference and translation. Of more than 173000 discovered lncRNAs, the majority remain functionally unknown. The cell type-specific expression and localization of the lncRNA also suggest potential distinct functions of lncRNAs across different cell types. This highlights the niche of identifying functional lncRNAs in different biological processes and diseases through high-throughput (HTP) screening. This review summarizes the current work performed and perspectives on HTP screening of functional lncRNAs where different technologies, platforms, cellular responses and the downstream analyses are discussed. We hope to provide a better picture in applying different technologies to facilitate functional annotation of lncRNA efficiently.

Introduction

The characterization of the mammalian transcriptome by the Encyclopedia of DNA Elements (ENCODE) and the Functional Annotation of the Mammalian Genome (FANTOM) projects have revealed a large collection of long non-coding RNAs (lncRNAs), which are defined as greater than 200 nucleotides and account for the majority of the transcriptome [1,2]. A study summarizing multiple transcript collections has further revealed approx. 20 000 human lncRNAs with functional insight [3]. Indeed, the single-nucleotide polymorphisms (SNPs) identified by genome-wide association studies (GWASs) that are associated with diseases or traits are mainly located in the non-coding region [4], suggesting that both DNA regulatory elements and lncRNAs play some functional roles in different diseases. Except for the canonical lncRNAs that were initially characterized such as *XIST*, *NEAT1* and *MALAT1* [5,6], the vast majority of lncRNAs were characterized only in the last decade. These lncRNAs have been shown to regulate pluripotency maintenance [7], cellular reprogramming [8], lineage commitment [9], carcinogenesis [10], and pathogenesis of various diseases [11]. While the pervasive transcription of lncRNAs implies they should acquire function over evolutionary time [12], functional role of the majority of lncRNAs remains to be characterized. Notably, lncRNAs express, localize, and function specifically across different cell types [13–15], suggesting that an lncRNA may be functional in one cell type or a particular cellular phenotype but not in the others.

The diverse roles of lncRNAs in organismal development, physiological processes, and disease pathology has been revealed by the advent of high-throughput (HTP) screenings [15–25]. Various technologies have allowed loss-/gain-of-function HTP screening to become popular and affordable. This has included the application of short hairpin RNA (shRNA) pooled libraries, CRISPR-based single guide RNA (sgRNA) pooled libraries, antisense oligonucleotide (ASO), and next-generation sequencing (NGS). While these HTP screening platforms have been frequently applied to mRNA functional screening, they have not yet
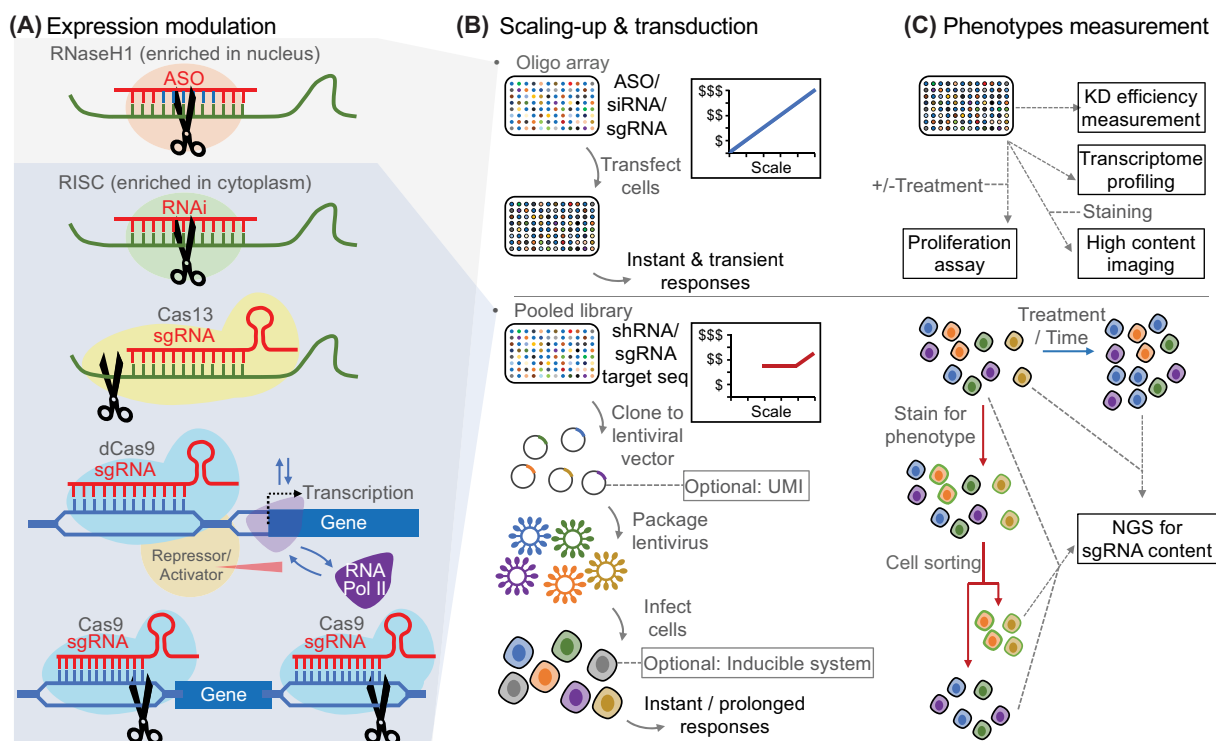
Figure 1. Functional screening for lncRNAs

(**A**) Methods for modulating lncRNA expression levels, by cleavage of the RNA directly, modulating the transcription, and knockout of the gene. (**B**) Expression modulation can be scaled to achieve HTP screening by oligo arrays and pooled libraries and (**C**) the resulting phenotypes for each of them.

become common for lncRNAs. Unlike protein-coding genes, the novelty of lncRNAs makes it challenging to hypothesize the connection between the lncRNA hits and the tested phenotypes. Thus, this review will summarize the use of HTP platforms for lncRNA screening, and the downstream analyses that help to narrow the gap between lncRNA hits and the phenotype.

Selection and prioritization of lncRNA targets

Currently, the number of known human lncRNA transcripts are over 173 000 according to NONCODE [26] and over 268 000 according to LncBook [27]. However, this could be far lower in a particular cell type since lncRNAs are highly cell type-specific [14]. Therefore, expression level is one of the key factors to select and prioritize lncRNAs in a screen to improve the success rate. Indeed, loss-of-function screening on lncRNAs with a broad range of expression levels showed positive correlations between expression level and the cell growth phenotype [15,16].

Furthermore, acquiring differential expression data by microarray or NGS-based methods between the desired phenotype and control can improve target selection. For example, lncRNAs associated with cancer subtypes and clinical prognosis were identified from microarray data of four cancer types [28], and the list was adopted in a CRISPR-based screening study [17]. Similarly, Xu and colleagues [18] selected 25 lncRNAs to screen for drug resistance by choosing the up-regulated lncRNAs after drug treatments using RNA-seq analysis. Although computationally it is possible to further prioritize to predict the most relevant lncRNAs [29–32], loss-/gain-of-function HTP screening is necessary to identify lncRNAs with a specific and relevant cellular phenotype.

Modulation of lncRNA expression

Many loss-/gain-of-function methods are available for modulating the expression of lncRNAs [33]. However, only several of them are scalable for use in HTP screening. We will only cover ASO, RNA-interference (RNAi) and CRISPR/Cas13 systems, which target the RNA directly, as well as CRISPR/Cas9 (CRISPRn) and use of the catalytically dead Cas9 (dCas9) (CRISPRi/a) which target genomic DNA (Figure 1A). For a clearer goal of the review, we

will start by comparing compatible modulation methods directly.

ASO versus RNAi

Both ASO and RNAi (including shRNA exogenous expression and siRNA transfection) directly target the RNA molecules for degradation. The action of RNAi depends on the RNA-induced silencing complex (RISC) to cleave the RNA target, while ASOs form a DNA–RNA duplex with the RNA target for RNase H recognition and cleavage. Typical ASOs for degrading RNA utilize LNA gapmer technology [34] and often include a $2'$MOE modification for stability and nuclease resistance. Both methodologies have been used in HTP screening for functional lncRNAs and shown to knockdown lncRNAs efficiently [16,22–24].

In comparing ASO and RNAi, the key difference is the localization of the endogenous enzymes used to target the RNA molecule—RNase H is enriched in the nucleus, while RISC is enriched in the cytoplasm. Consequently, nuclear lncRNAs were more efficiently suppressed by ASOs, compared with cytoplasmic lncRNAs which responded better to RNAi [35]. However, since most lncRNAs are enriched in the nucleus where chromatin is regulated by lncRNA [36,37], targeting the nuclear lncRNAs is often more beneficial. Additionally, ASOs were shown to be equally effective in targeting introns and exons of lncRNAs [16], indicating ASOs could knockdown cytoplasmic RNA during their transcription in the nucleus. Indeed, lncRNA localization negatively affects RNAi more than ASO [35]. Therefore, ASO is recommended for modulating the expression of lncRNA for array-based screens. On the contrary, shRNA pooled library allows the RNAi modality to be scalable. Besides expression modulation, RNase H-inactive ASO has been used to block splicing of nascent RNA [38,39]. The same method has been applied to block the splicing of an lncRNA, resulting in chromatin retention and malfunction of the lncRNA [40]. This highlights the added benefits of ASO in studying lncRNA biology.

CRISPRn versus CRISPRi/a

CRISPR/Cas9-mediated genome editing has been commonly used in mRNA functional screening and also for lncRNA functionality. Zhu and colleagues [17] adopted paired-guide RNA to mediate deletion of various lengths from 700 lncRNA genes and identified 51 lncRNAs that affected cell growth.

In the CRISPRi system, dCas9 is fused to transcriptional repressor domains KRAB or methyl-CpG binding protein 2 (MeCP2) to achieve targeted suppression of gene expression [41–43]. Besides those for mRNA, several large-scale CRISPRi screenings have been performed to elucidate lncRNA functionality [15,21,25]. More recently, dCas9 fused with transcriptional activator VP64 and other synergistic activators have been described [44]. Unlike conventional plasmid-based overexpression, CRISPRa activates lncRNA expression from the endogenous genomic locus, which has the advantage of capturing *cis*-acting and nuclear lncRNA functions. Thus, CRISPRa has been utilized for the functional screening of lncRNAs [19,20].

When comparing CRISPRn knockout and CRISPRi knockdown in pooled library studies targeting mRNAs, CRISPRn is more effective than CRISPRi [45,46]. However, there are several limitations when applying CRISPRn to modulate lncRNA [47]. Firstly, unlike mRNA, partial deletion of the lncRNA genes may not ablate their functions, with lncRNA gene size often too long for complete deletion [48]. While other options such as deletion of the promoter could be considered, the varied efficiency of suppression for different loci is challenging in large-scale screening. Secondly, CRISPRn could affect other proximal functional elements and their topological interactions thus confounding the mechanistic activity of lncRNAs. CRISPRn cleavage also hinders the detection of *cis*-regulation of lncRNAs. Finally, as many of the lncRNAs overlap with other genes, deletion by CRISPRn is less applicable. Additionally, ds-DNA damage mediated by CRISPRn is known to trigger non-specific false positives [49], which does not occur with CRISPRi [50]. Therefore, CRISPRi/a is more applicable than CRISPRn for lncRNA expression modulation.

CRISPRi versus shRNA versus CRISPR/Cas13

The performance of pooled libraries for RNAi, CRISPRi, and CRISPRn were compared using 46 essential and 47 non-essential mRNAs in a negative selection screen [46]. This study found that (1) CRISPRn screening performs the best in both sgRNA- and gene-based analyses, (2) shRNA screening results reflected off-target effects of individual shRNAs, which can be reduced by using multiple shRNAs, and (3) CRISPRi screening shows virtually no off-target effects, but only 50% of the sgRNA is effective, leading to a lower hit rate. As discussed earlier, CRISPRn is not the most suitable screening method for lncRNAs. For both the shRNA and CRISPRi methods, their shortcomings can be compensated by including more shRNA or sgRNA constructs. Additionally, sgRNA design can be improved with accurate transcription start sites (TSSs) positioning using FANTOM CAGE annotation [51]. Similarly, tools to improve the selection of lncRNA targets according to expression level and TSS annotation for sgRNA design are available

[52]. Therefore, both CRISPRi and shRNA pooled screening are applicable for lncRNA functional screening, with the choice of targeting the DNA or RNA, respectively.

However, a major advantage of CRISPR-based screening is the ongoing development of relevant technologies and supportive analytical algorithms. Recently, an RNA-guided RNA-targeting CRISPR effector Cas13 was characterized and shown to function in both the nucleus and cytoplasm [53–55]. Notably, CRISPR/Cas13 is reported to have a high specificity, allowing the possibility of using closely related mismatch controls in knockdown studies [54]. However, Cas13 exhibits collateral activity after target recognition and cleaves any RNA in close proximity regardless of complementarity [54,56], which may be a hindrance in its utility for RNA knockdown. While Cas13 has already been used in a functional lncRNA screen [18], a comprehensive comparison with other knockdown methods is still needed.

Targeting the lncRNA transcripts versus lncRNA loci

In summary, targeting the transcripts of lncRNA, and not their DNA loci, is advantageous for (1) not interfering with the function of the lncRNA promoter, which may act as an enhancer for other genes, (2) targeting only individual isoforms since different isoforms could have opposing effects [57], and (3) allowing additional lncRNAs to be considered for screening, since the DNA-targeting approach must avoid affecting other genes on the same loci of intragenic, divergent, and antisense lncRNA genes. On the contrary, targeting the DNA loci could also investigate enhancer-like activities, broadening the coverage of functional non-coding regulatory elements in the genome. Furthermore, both ASO and RNAi methods exhibit independent off-target effects [16,58] whereas CRISPRi/a modulation of the DNA is less prone to off-target effects because of the narrow targeting window.

Scaling to HTP screening

HTP screening has shown promise in identifying individual lncRNAs which regulate a cellular phenotype. Moreover, as lncRNAs constitute the majority of the transcriptome and harbor many GWAS SNPs, using HTP screenings to identify the proportion of lncRNAs with the information of their genomic characteristics that participate in cellular mechanisms has become a key tool to understand the genome and its role in disease. Additionally, HTP screening can identify associations between drug compounds and their gene targets, which include both protein-coding genes [59–62] and lncRNAs [18–20,63]. For example, Bester and colleagues [19] utilized a CRISPRa-system to identify genes contributing to cytarabine resistance in an acute myeloid leukemia cell line. They revealed a group of lncRNAs driving cytarabine resistance via *cis*-regulation. Such studies further highlight the interconnected role lncRNAs play in various pathological mechanisms.

Screening in an array format has the advantage of separating the individual perturbations. This allows measurement by qPCR to rule out the unsuccessful perturbation [16,22], where the degree of knockdown significantly affects the hit rate [16]. This also provides the flexibility to directly record cellular phenotype, such as through high-content imaging [24]. Oligonucleotide arrays are available for chemically synthesized siRNA, ASO, and sgRNA (Figure 1B). Previously the relatively high cost of LNA gapmer ASO limited its use in HTP screening. However, due to the lifting of the LNA patent protection, the cost is now comparable with siRNA and sgRNA (~$200 USD per oligo). Alternatively, both shRNA and sgRNA can be generated by cloning into viral vectors in an array format, however, this task can be laborious to scale-up. When the throughput of an arrayed screen is limited by cost and experimental intensiveness, pooled library screening provides another option where the throughput can be near unlimited (~$20 000 USD per library with 30 000 constructs).

Screening by oligonucleotide array
Design of the array platform

When designing the arrayed screens, at least two constructs with independent sequences showing effective knockdown and the same cellular phenotype are necessary to call a hit. From our previous study, which used 2021 ASOs to target 285 lncRNAs, 43.5% of ASOs were effective and there were 68.1% lncRNA targets with at least two effective ASOs [16]. Our results showed that higher expressing lncRNAs are more susceptible to ASO knockdown, suggesting inclusion of such targets could improve the chances of an effective ASO. For RNAi, Guttman and colleagues [22] designed five shRNAs per target for 214 intergenic lncRNAs, in which 65% had at least one effective knockdown. Designing a functional ASO sequence without off-targeting for lncRNA is challenging, since lncRNA genes harbor many repetitive elements [64] and ASOs have the potential to target the intronic sequences of nascent transcripts. For RNAi, off-target effects were partly due to the dependence on a relatively short complementary seed sequence in the 3′ end of RNAi [65]. Therefore, in order to reach enough effective oligos and deal with the off-target effects, the starting number of oligos should be sufficiently high (e.g., five or more). The phenotypic response of transfecting

Table 1 Screenings applied to lncRNAs

Technologies	Cell types	LncRNAs	Constructs	Phenotype (% hit)	References
Array screening					
siRNA	HeLa	2231	4 (pooled)	Mitotic progression (0.1%) Chromosome segregation (0.1%) Cytokinesis (0.1%)	[24]
ASO	Human dermal fibroblast	285 (194), 119	5–15 (2–10), ≥2	Proliferation (7.7%), CAGE sequencing (10.9%)	[16]
shRNA	Mouse ES	214 (147)	5 (1–2)	Microarray (93%)	[22]
Pooled library screening					
CRISPRi	Epidermal keratinocyte	2263	5	Proliferation (0.4%)	[25]
Cas13	K562	25	10	Proliferation with three anti-cancer drug treatments (64%)	[18]
CRISPRi	Human glioblastoma	5689	10	Proliferation with fractionated radiation (8.2%)	[21]
CRISPRa	MOLM14 AML	14701	≥4	Proliferation with Cytarabine treatment (19.5%)	[19]
CRISPRa	Human melanoma A375	10504	~10	Proliferation with Vemurafenib treatment (0.2%)	[20]
CRISPRi	iPSC, MCF7, U87, K562, MDA-MB-231, HeLa, HEK293T	5543, 5725, 5689, 16401, 5725, 6158, 5785	10	Proliferation (5.9%), Proliferation (1%), Proliferation (1.1%), Proliferation (0.4%), Proliferation (0.5%), Proliferation (0.4%), Proliferation (0.3%)	[15]
CRISPR	Huh7.5$_{OC}$	671	~20	Proliferation (7.6%)	[17]
shRNA	Mouse ES	1280	≥3	OCT4 expression (1.6%)	[23]

oligos to cells is transient. For a longer phenotypic assay, such as differentiation, lentivirus-mediated shRNA array can be considered, as done by Guttman and colleagues [22] to identify lncRNAs that are important in mouse ES cells.

Phenotypes tested in array platform

When adopting an array method of screening, the cellular phenotypes of interest are abundant as compared with pooled screening. Numerous quantitative cellular assays have been utilized, such as those measuring growth, differentiation [66], infection [67], and endocytosis [68]. Among them, the inclusion of imaging is a major advantage for array screening. For instance, 50 lncRNAs were shown to affect cell morphology in human dermal fibroblasts by real-time imaging after ASO knockdown [16]. By applying high-content imaging to RNAi array screening, Stojic and colleagues [24] identified six lncRNAs for regulating mitotic progression, chromosome segregation, and cytokinesis.

Screening by pooled library
Design of the pooled library platform

When designing a pooled library screening, several factors determine the scale of the study. These include the total number of lncRNA targets, how many sgRNA/shRNA constructs per lncRNA, the additional non-target scramble sequences, and the size of the coverage. A higher number of constructs could compensate for the off-target effects of shRNA and increase the number of effective sgRNAs. Typically, the number of constructs is in a range of 5–20 shRNA/sgRNA per target (Table 1). From an analytical point of view, the number of constructs should be at least 4, while a higher number allows for statistical analysis that can incorporate technical and biological variability to improve power [69]. The number of non-target controls reflects the variation of the screen, which is complicated with the randomness of construct distribution and cell-to-cell variation. Therefore, it is necessary to include a large pool of non-target controls, which usually ranges from ~100 to ~1000 [19,21,25]. Assuming a library of 10 000 constructs, 10 constructs per lncRNA with 1000 non-target controls, the throughput of the library will be 900 lncRNAs. The infection coverage represents the number of cells uptaking the same construct, where each cell contains only one construct by

restricting the multiplicity of infection (MOI) (usually ≤ 0.3). As the genomic integration event is random, a sufficient size of infection coverage can normalize the variability. The common infection coverage for each construct is approx. 300–500× [15,19,25] while sequencing coverage is approx. 1000×. Additional independent experimental replications of the same library are also necessary. Therefore, for a single replicate, the number of infected cells needed for 300× coverage is 3 million, and the starting number of cells is 10 million at an MOI of 0.3.

Unique molecular identifiers

Since cell-to-cell variability has posed challenges in interpreting phenotypes, strategies such as incorporating unique molecular identifiers (UMIs) in sgRNA libraries have been established [70]. The UMIs have allowed for the screening of clonally expanded and individually tagged cells, resulting in an increased sensitivity and robustness compared with conventional analyses. The statistical methods, including using the UMIs as internal replicates and in lineage dropout analyses, increase both the precision and the accuracy of the screen, as well as reducing the infection coverage needed to reach the same statistical power [71].

Phenotypes tested in pooled library platform

The cellular phenotypes assessed after genetic perturbation are diverse (Figure 1C), including survival advantage for robust cell growth [15,17,25], after drug treatments [18–20] or with fractionated radiation [21]. By combining with cell sorting, a wide variety of phenotypes can be measured, such as pluripotency maintenance [15,23,72], differentiation [73], protein transport [74], oxidative stress [75], and many more.

Almost all the pooled library screenings for lncRNAs thus far have relied on survival advantage as the phenotype of interest. For example, Liu and colleagues [21] identified 434 and 33 lncRNAs, that respectively support and reduce cell growth, in human glioblastoma cells in the presence of clinically relevant doses of fractionated radiation. Additionally, Cai and colleagues [25] identified 9 lncRNAs that support robust cell proliferation of epidermal keratinocyte cells. The lower hit rate of this study may be due to a lower number of sgRNAs designed per target and thus lower statistical power. Another reason is that including treatments that place the cells under stress, as in the Liu and colleagues' study [21], can provoke the expression or function of lncRNAs, which constitute a significant fraction of the genes differentially expressed in response to cell stress [76]. As summarized for all lncRNA screenings in Table 1, the hit rate is generally higher if drug treatment is included.

When combined with cell sorting, cell loss from the staining and washing steps of fluorescence-activated cell sorting (FACS) is a factor to consider, although CRISPRi screen combined with FACS has been reported [15]. A larger population of transduced cells are needed to compensate for this cell loss and reach the final coverage. Formaldehyde fixation can reduce the degree of cell loss while de-crosslinking is required to rescue the genomic DNA for PCR library construction [15]. Indeed, many screens rely on expression of exogenous genes carrying fluorescent signals [73,74] or fluorescent probe live trackers [75]. For example, Liu and colleagues [73] performed a CRISPRa screen in mouse ES cells, where the cell surface marker hCD8 was inserted downstream of Tubb3. Differentiated neurons were then separated by magnetic-activated cell sorting (MACS), which combined with use of cell surface markers, can minimize cell loss.

Analytical efforts

Pooled library screening requires rigorous bioinformatics analyses to interpret the results, with detection of false positives remaining a critical issue. Nevertheless, advanced analysis strategies for CRISPR applications are currently available [77], with several distinct algorithms established for evaluating the results of pooled library screening. Briefly, redundant siRNA activity [78] and HiTSelect [79] are designed for RNAi screening. Redundant siRNA activity ranks the targeted genes by log fold change and generates *P*-values from the ranking against a uniform distribution, while HiTSelect ranks target genes by considering both the effect on the phenotype and the number of active constructs using a random-effects model. The MAGeCK robust ranking algorithm [80] is commonly used in CRISPR-based screens. It uses the raw sgRNA read counts and adopts a negative binomial model to generate sgRNA *P*-values, which are combined to gene level by a modified robust ranking algorithm. CRISPhieRmix [81] uses the log fold change value of each sgRNA generated from standard count software such as DESeq2 [82], and provides empirical FDR for the target genes using a hierarchical mixture distribution. BAGEL [83] uses data from prior screens to build null distribution and positive effects to rank target genes. Besides, some algorithms such as MAGeCK maximum likelihood estimation [84] and JACKS [85] are designed to compare and pool multiple screens. CERES [86] is specifically designed to correct the side effects mediated by DNA damage from CRISPR cuts for cancer cells which exhibit large copy number variation. Bodapati and colleagues [69] compared these algorithms with CRISPR pooled screening data and suggested

using MAGeCK robust ranking algorithm in most cases for its robustness and performance, while CRISPhieRmix was touted as the only algorithm taking the various sgRNA efficiencies into account for CRISPRi/a screenings. If UMIs are included, there is an advantage to incorporating additional statistical methods, such as internal replicate analysis and lineage dropout analysis [71]. Because the hits identified by pooled library screening are statistical outcomes from a population of perturbations with considerable false positives [87], validation by reproducing the phenotypic result with individual perturbation is sometimes necessary.

Annotate the roles of the lncRNA hits
Transcriptome profiling

Except for several of the lncRNAs with structural implications [88] or those regulating translation [89], the majority of lncRNAs regulate other genes at the transcriptional [36,37] or RNA [90] levels. This signature allows us to unveil the functions of most lncRNAs by studying the transcriptomic changes. Therefore, follow-up studies restricted to lncRNA hits using Perturb-seq or CROP-seq [91–93] will be the most compatible with pooled library screening. Perturb-seq and CROP-seq are sequencing platforms designed to combine single cell RNA-seq and CRISPR-based genetic screens. To facilitate the detection of the non-polyadenylated gRNA in single cell transcriptome, Perturb-seq lentiviral vector harbors a gRNA-matched barcode upstream of the poly-A tail of the puromycin gene [91], while CROP-seq introduces an additional gRNA copy upstream of the poly-A of the puromycin gene [93]. Besides, a single cell RNA-seq platform has been applied to the shRNA screen by using a pol II-dependent shRNA backbone [94]. Single-cell transcriptomes containing the sgRNA/shRNA identity can unveil the mechanism mediated by specific lncRNAs, while a single-cell readout is advantageous if the cell population is heterogeneous (e.g., study design with differentiation). More recently, targeted Perturb-seq [95] was developed, allowing profiling of a subset of the transcriptome (e.g., genes near the loci of lncRNA hits). Alternatively, molecular phenotyping with bulk RNA-seq or CAGE-seq after lncRNA perturbation can reveal genes or pathways modulated by the lncRNA [16]. Such transcriptomic profiling will identify the global transcriptomic changes, which can be captured by differential expression profile followed by Gene Set Enrichment Analysis (GSEA) [96] and Gene Ontology (GO) [97] analyses. Ramilowski and colleagues [16] characterized *ZNF213-AS1* by GSEA in controlling migration in dermal fibroblast and validated this in a wound closure migration assay. Besides global transcriptomic changes, transcriptomic profiling combined with computational analyses will also identify direct effector genes of the lncRNA.

Prediction of effector genes with known lncRNA mechanisms

Predicting direct effector genes is often necessary to connect lncRNA hits to the tested phenotype with cellular mechanisms. This will also reflect the functional mechanisms of the lncRNAs so that validation experiments can be designed. Since *cis*-regulation is one of the major mechanisms mediated by lncRNA [36,37], defining proximal genes either by 2D distance or 3D chromatin structure from Hi-C data will yield functional interactions between the lncRNA locus and the effector genes (Figure 2A). From certain disease phenotypes, effector genes can be identified by connecting lncRNAs to GWAS [98] and expression quantitative trait loci (eQTL) [99–101]. Additionally, co-expression analysis among the tissue-wide or cell type-wide data from various consortia [1,102] between the lncRNAs and effector genes should improve confidence in these networks.

Moreover, lncRNAs are known to function by interacting with protein, DNA and RNA [103,104], while their subcellular localization can suggest their functional mechanisms [105]. Fractionation data will be useful to estimate mechanisms of the lncRNAs, but data with matched cell type are necessary as subcellular localization of lncRNAs is cell type specific [13]. For *trans*-regulation, experimental genome-wide RNA–DNA interaction analysis, such as GRID-seq [106] and RADICL-seq [107] and in silico RNA–DNA interaction such as triplex formation prediction [108] can be used as references (Figure 2B). Both *cis* and *trans* RNA–DNA interactions are likely to involve proteins as the executors while linking a protein (chromatin modifier or transcription factor) with RNA–DNA interaction will help characterizing the downstream effects. Many DNA-binding proteins were found to be capable of binding RNA [109,110]. RNA-binding protein (RBP) is a class of protein that cooperates with lncRNAs for post-transcriptional processes, such as splicing, cleavage and polyadenylation. ENCODE Phase III has generated RNA–RBP interaction for 356 RBPs in K562 and HepG2 [109] while ChIP-seq data are also available for 58 of these RBPs [111]. Combining these two datasets could reveal RNA-DNA interaction with RBP content (Figure 2C). Last but not least, sequestering miRNA is one of the known mechanisms of lncRNA while the differential expression of the mRNA competitor can be revealed from transcriptome profiling. A number of databases supporting the prediction of miRNA sponge interaction have been described [112] (Figure 2D).
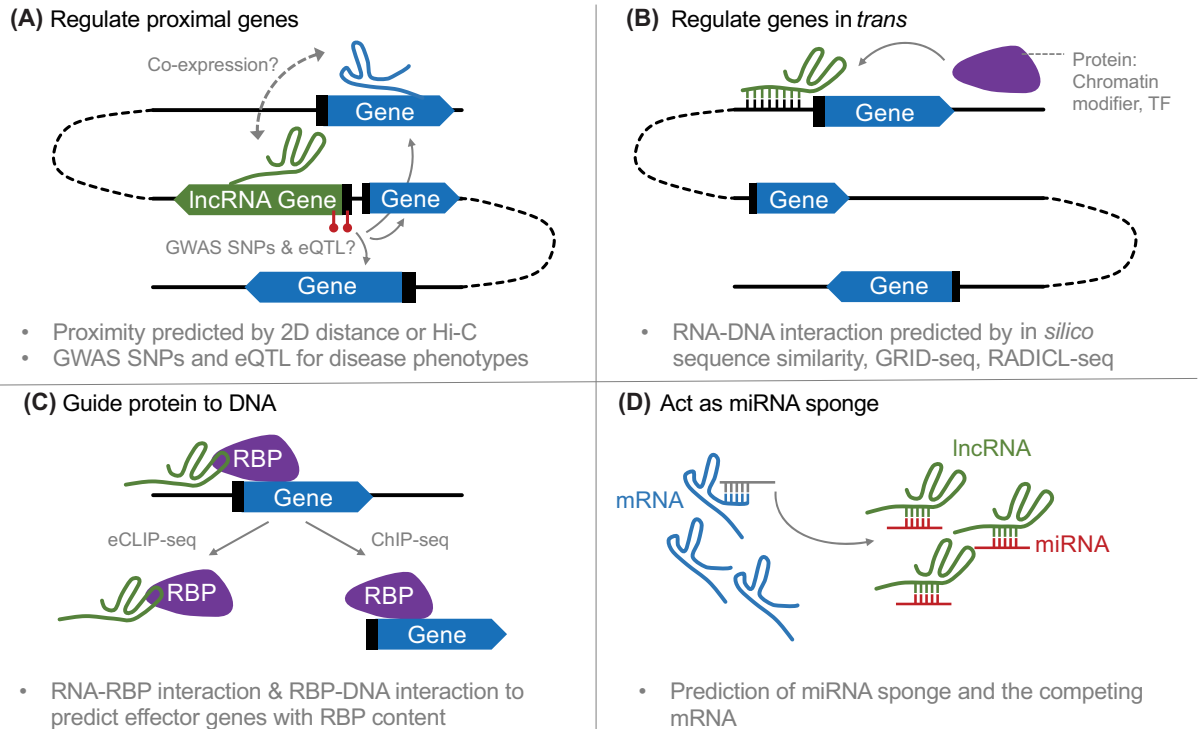
Figure 2. Mechanisms of lncRNA in gene regulation

(**A**) Regulation of proximal genes. (**B**) Regulation of genes in *trans*. (**C**) Mediation of protein-DNA interaction to regulate gene expression. (**D**) Sequestering miRNA from mRNA by acting as miRNA sponge.

Conclusion

Array and pooled library screenings have been established in many studies to identify functional protein coding genes in different cellular phenotypes, but the work on lncRNAs is lagging. However, the majority of lncRNAs have yet to be characterized. Advances in HTP screening platforms present an opportunity to explore the functionality of lncRNAs. Following the genome-wide screening of lncRNAs, it will also be imperative to investigate molecular mechanisms of individual lncRNAs for determining their roles across different cell types, including in disease, to identify functional conservation and redundancy.

Summary

- Direct RNA-targeting perturbation methods present the advantage of (1) distinguishing between isoforms, (2) avoiding interference with the enhancer function of lncRNA promoters, and (3) targeting lncRNAs even when their loci overlap with other genes.

- DNA-targeting perturbation methods provide more consistent results by having fewer off-target effects and allowing functional screening of *cis*-regulatory elements.

- Array screening can be easily combined with different phenotypic readouts and allow the quantification of perturbation efficiency.

- Pooled library screening benefits from high throughput, but the results are preliminary and often require secondary screening and validations.

PORTLAND
PRESS

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

Author Contribution

C.W.Y. and J.W.S. conceived the content of the paper. C.W.Y. wrote the majority of the paper. D.M.S. and A.V.P. contributed to the writing. All authors reviewed, edited and approved the final version of the paper.

Abbreviations

ASO, antisense oligonucleotide; dCas9, dead Cas9; ENCODE, Encyclopedia of DNA Elements; ES, embryonic stem; FANTOM, Functional Annotation of the Mammalian Genome; FDR, false discovery rate; GSEA, gene set enrichment analysis; GWAS, genome-wide association study; HTP, high-throughput; lncRNA, long non-coding RNA; MOI, multiplicity of infection; NGS, next-generation sequencing; pol II, RNA polymerase II; qPCR, quantitative PCR; RBP, RNA-binding protein; RISC, RNA-induced silencing complex; RNAi, RNA-interference; sgRNA, single guide RNA; shRNA, short hairpin RNA; SNP, single-nucleotide polymorphism; TSS, transcription start site; UMI, unique molecular identifier.

References

1   Forrest, A.R.R., Kawaji, H., Rehli, M., Baillie, J.K., de Hoon, M.J.L., FANTOM Consortium and the RIKEN PMI and CLST (DGT) et al. (2014) A promoter-level mammalian expression atlas. *Nature* **507**, 462–470, https://doi.org/10.1038/nature13182

2   ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature.* **489**, 57–74, https://doi.org/10.1038/nature11247

3   Hon, C.-C., Ramilowski, J.A., Harshbarger, J., Bertin, N., Rackham, O.J.L., Gough, J. et al. (2017) An atlas of human long non-coding RNAs with accurate 5′ ends. *Nature* **543**, 199–204, https://doi.org/10.1038/nature21374

4   Tak, Y.G. and Farnham, P.J. (2015) Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics Chromatin* **8**, 57, https://doi.org/10.1186/s13072-015-0050-4

5   Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R. et al. (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349**, 38–44, https://doi.org/10.1038/349038a0

6   Hutchinson, J.N., Ensminger, A.W., Clemson, C.M., Lynch, C.R., Lawrence, J.B. and Chess, A. (2007) A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **8**, 39, https://doi.org/10.1186/1471-2164-8-39

7   Fico, A., Fiorenzano, A., Pascale, E., Patriarca, E.J. and Minchiotti, G. (2019) Long non-coding RNA in stem cell pluripotency and lineage commitment: functions and evolutionary conservation. *Cell Mol. Life Sci.* **76**, 1459–1471, https://doi.org/10.1007/s00018-018-3000-z

8   Luginbühl, J., Sivaraman, D.M. and Shin, J.W. (2017) The essentiality of non-coding RNAs in cell reprogramming. *Non Coding RNA Res.* **2**, 74–82, https://doi.org/10.1016/j.ncrna.2017.04.002

9   Chen, J., Wang, Y., Wang, C., Hu, J.-F. and Li, W. (2020) LncRNA functions as a new emerging epigenetic factor in determining the fate of stem cells. *Front. Genet.* **11**, 277, https://doi.org/10.3389/fgene.2020.00277

10   Bhan, A., Soleimani, M. and Mandal, S.S. (2017) Long noncoding RNA and cancer: a new paradigm. *Cancer Res.* **77**, 3965–3981, https://doi.org/10.1158/0008-5472.CAN-16-2634

11   Yang, S., Lim, K.-H., Kim, S.-H. and Joo, J.-Y. (2020) Molecular landscape of long noncoding RNAs in brain disorders. *Mol. Psychiatry* **26**, 1060–1074, https://doi.org/10.1038/s41380-020-00947-5

12   Palazzo, A.F. and Koonin, E.V. (2020) Functional long non-coding RNAs evolve from junk transcripts. *Cell* **183**, 1151–1161, https://doi.org/10.1016/j.cell.2020.09.047

13   Cabili, M.N., Dunagin, M.C., McClanahan, P.D., Biaesch, A., Padovan-Merhar, O., Regev, A. et al. (2015) Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol.* **16**, 20, https://doi.org/10.1186/s13059-015-0586-4

14   Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A. et al. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **25**, 1915–1927, https://doi.org/10.1101/gad.17446611

15   Liu, S.J., Horlbeck, M.A., Cho, S.W., Birk, H.S., Malatesta, M., He, D. et al. (2017) CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science* **355**, eaah7111, https://doi.org/10.1126/science.aah7111

16   Ramilowski, J.A., Yip, C.W., Agrawal, S., Chang, J.-C., Ciani, Y., Kulakovskiy, I.V. et al. (2020) Functional annotation of human long noncoding RNAs via molecular phenotyping. *Genome Res.* **30**, 1060–1072, https://doi.org/10.1101/gr.254219.119

17   Zhu, S., Li, W., Liu, J., Chen, C.-H., Liao, Q., Xu, P. et al. (2016) Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. *Nat. Biotechnol.* **34**, 1279–1286, https://doi.org/10.1038/nbt.3715

18   Xu, D., Cai, Y., Tang, L., Han, X., Gao, F., Cao, H. et al. (2020) A CRISPR/Cas13-based approach demonstrates biological relevance of vlinc class of long non-coding RNAs in anticancer drug response. *Sci. Rep.* **10**, 1794, https://doi.org/10.1038/s41598-020-58104-5

19   Bester, A.C., Lee, J.D., Chavez, A., Lee, Y.-R., Nachmani, D., Vora, S. et al. (2018) An integrated genome-wide CRISPRa approach to functionalize lncRNAs in drug resistance. *Cell* **173**, 649.e20–664.e20, https://doi.org/10.1016/j.cell.2018.03.052

20 Joung, J., Engreitz, J.M., Konermann, S., Abudayyeh, O.O., Verdine, V.K., Aguet, F. et al. (2017) Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature* **548**, 343–346, https://doi.org/10.1038/nature23451

21 Liu, S.J., Malatesta, M., Lien, B.V., Saha, P., Thombare, S.S., Hong, S.J. et al. (2020) CRISPRi-based radiation modifier screen identifies long non-coding RNA therapeutic targets in glioma. *Genome Biol.* **21**, 83, https://doi.org/10.1186/s13059-020-01995-4

22 Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G. et al. (2011) lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **477**, 295–300, https://doi.org/10.1038/nature10398

23 Lin, N., Chang, K.-Y., Li, Z., Gates, K., Rana, Z.A., Dang, J. et al. (2014) An evolutionarily conserved long noncoding RNA TUNA controls pluripotency and neural lineage commitment. *Mol. Cell* **53**, 1005–1019, https://doi.org/10.1016/j.molcel.2014.01.021

24 Stojic, L., Lun, A.T.L., Mascalchi, P., Ernst, C., Redmond, A.M., Mangei, J. et al. (2020) A high-content RNAi screen reveals multiple roles for long noncoding RNAs in cell division. *Nat. Commun.* **11**, 1851, https://doi.org/10.1038/s41467-020-14978-7

25 Cai, P., Otten, A.B.C., Cheng, B., Ishii, M.A., Zhang, W., Huang, B. et al. (2020) A genome-wide long noncoding RNA CRISPRi screen identifies PRANCR as a novel regulator of epidermal homeostasis. *Genome Res.* **30**, 22–34, https://doi.org/10.1101/gr.251561.119

26 Fang, S., Zhang, L., Guo, J., Niu, Y., Wu, Y., Li, H. et al. (2018) NONCODEV5: a comprehensive annotation database for long non-coding RNAs. *Nucleic Acids Res.* **46**, D308–D314, https://doi.org/10.1093/nar/gkx1107

27 Ma, L., Cao, J., Liu, L., Du, Q., Li, Z., Zou, D. et al. (2019) LncBook: a curated knowledgebase of human long non-coding RNAs. *Nucleic Acids Res.* **47**, D128–D134, https://doi.org/10.1093/nar/gky960

28 Du, Z., Fei, T., Verhaak, R.G.W., Su, Z., Zhang, Y., Brown, M. et al. (2013) Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. *Nat. Struct. Mol. Biol.* **20**, 908–913, https://doi.org/10.1038/nsmb.2591

29 Elkouris, M., Kouroupi, G., Vourvoukelis, A., Papagiannakis, N., Kaltezioti, V., Matsas, R. et al. (2019) Long non-coding RNAs associated with neurodegeneration-linked genes are reduced in Parkinson's disease patients. *Front. Cell Neurosci.* **13**, 58, https://doi.org/10.3389/fncel.2019.00058

30 D'haene, E., Jacobs, E.Z., Volders, P.-J., De Meyer, T., Menten, B. and Vergult, S. (2016) Identification of long non-coding RNAs involved in neuronal development and intellectual disability. *Sci. Rep.* **6**, 28396, https://doi.org/10.1038/srep28396

31 Wang, M., Wang, L., Pu, L., Li, K., Feng, T., Zheng, P. et al. (2020) LncRNAs related key pathways and genes in ischemic stroke by weighted gene co-expression network analysis (WGCNA). *Genomics* **112**, 2302–2308, https://doi.org/10.1016/j.ygeno.2020.01.001

32 Shi, Y., Liu, H., Yang, C., Xu, K., Cai, Y., Wang, Z. et al. (2020) Transcriptomic analyses for identification and prioritization of genes associated with Alzheimer's disease in humans. *Front. Bioeng. Biotechnol.* **8**, 31, https://doi.org/10.3389/fbioe.2020.00031

33 Liu, S.J. and Lim, D.A. (2018) Modulating the expression of long non-coding RNAs for functional studies. *EMBO Rep.* **19**, e46955, https://doi.org/10.15252/embr.201846955

34 Hagedorn, P.H., Persson, R., Funder, E.D., Albæk, N., Diemer, S.L., Hansen, D.J. et al. (2018) Locked nucleic acid: modality, diversity, and drug discovery. *Drug Discov. Today* **23**, 101–114, https://doi.org/10.1016/j.drudis.2017.09.018

35 Lennox, K.A. and Behlke, M.A. (2016) Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides. *Nucleic Acids Res.* **44**, 863–877, https://doi.org/10.1093/nar/gkv1206

36 Lee, J.T. (2012) Epigenetic regulation by long noncoding RNAs. *Science* **338**, 1435–1439, https://doi.org/10.1126/science.1231776

37 Mercer, T.R. and Mattick, J.S. (2013) Structure and function of long noncoding RNAs in epigenetic regulation. *Nat. Struct. Mol. Biol.* **20**, 300–307, https://doi.org/10.1038/nsmb.2480

38 Morcos, P.A. (2007) Achieving targeted and quantifiable alteration of mRNA splicing with Morpholino oligos. *Biochem. Biophys. Res. Commun.* **358**, 521–527, https://doi.org/10.1016/j.bbrc.2007.04.172

39 Lim, K.H., Han, Z., Jeon, H.Y., Kach, J., Jing, E., Weyn-Vanhentenryck, S. et al. (2020) Antisense oligonucleotide modulation of non-productive alternative splicing upregulates gene expression. *Nat. Commun.* **11**, 3501, https://doi.org/10.1038/s41467-020-17093-9

40 Ntini, E., Louloupi, A., Liz, J., Muino, J.M., Marsico, A. and Ørom, UAV. (2018) Long ncRNA A-ROD activates its target gene DKK1 at its release from chromatin. *Nat. Commun.* **9**, 1636, https://doi.org/10.1038/s41467-018-04100-3

41 Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P. et al. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183, https://doi.org/10.1016/j.cell.2013.02.022

42 Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E. et al. (2013) RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826, https://doi.org/10.1126/science.1232033

43 Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E. et al. (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451, https://doi.org/10.1016/j.cell.2013.06.044

44 Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., Iyer, E.P.R. et al. (2015) Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* **12**, 326–328, https://doi.org/10.1038/nmeth.3312

45 Sanson, K.R., Hanna, R.E., Hegde, M., Donovan, K.F., Strand, C., Sullender, M.E. et al. (2018) Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 5416, https://doi.org/10.1038/s41467-018-07901-8

46 Evers, B., Jastrzebski, K., Heijmans, J.P.M., Grernrum, W., Beijersbergen, R.L. and Bernards, R. (2016) CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat. Biotechnol.* **34**, 631–633, https://doi.org/10.1038/nbt.3536

47 Awwad, D.A. (2019) Beyond classic editing: innovative CRISPR approaches for functional studies of long non-coding RNA. *Biol. Methods Protoc.* **4**, bpz017, https://doi.org/10.1093/biomethods/bpz017

48 Liu, Z., Hui, Y., Shi, L., Chen, Z., Xu, X., Chi, L. et al. (2016) Efficient CRISPR/Cas9-mediated versatile, predictable, and donor-free gene knockout in human pluripotent stem cells. *Stem Cell Rep.* **7**, 496–507, https://doi.org/10.1016/j.stemcr.2016.07.021

49 Aguirre, A.J., Meyers, R.M., Weir, B.A., Vazquez, F., Zhang, C.-Z., Ben-David, U. et al. (2016) Genomic copy number dictates a gene-independent cell response to CRISPR/Cas9 targeting. *Cancer Discov.* **6**, 914–929, https://doi.org/10.1158/2159-8290.CD-16-0154

50 Horlbeck, M.A., Gilbert, L.A., Villalta, J.E., Adamson, B., Pak, R.A., Chen, Y. et al. (2016) Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *eLife* **5**, e19760, https://doi.org/10.7554/eLife.19760

51 Radzisheuskaya, A., Shlyueva, D., Müller, I. and Helin, K. (2016) Optimizing sgRNA position markedly improves the efficiency of CRISPR/dCas9-mediated transcriptional repression. *Nucleic Acids Res.* **44**, e141, https://doi.org/10.1093/nar/gkw583

52 Bergadà-Pijuan, J., Pulido-Quetglas, C., Vancura, A. and Johnson, R. (2020) CASPR, an analysis pipeline for single and paired guide RNA CRISPR screens, reveals optimal target selection for long non-coding RNAs. *Bioinformatics* **36**, 1673–1680, https://doi.org/10.1093/bioinformatics/btz811

53 Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B.T. et al. (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573, https://doi.org/10.1126/science.aaf5573

54 Abudayyeh, O.O., Gootenberg, J.S., Essletzbichler, P., Han, S., Joung, J., Belanto, J.J. et al. (2017) RNA targeting with CRISPR-Cas13. *Nature* **550**, 280–284, https://doi.org/10.1038/nature24049

55 Yang, L.-Z., Wang, Y., Li, S.-Q., Yao, R.-W., Luan, P.-F., Wu, H. et al. (2019) Dynamic Imaging of RNA in Living Cells by CRISPR-Cas13 Systems. *Mol. Cell* **76**, 981.e7–997.e7, https://doi.org/10.1016/j.molcel.2019.10.024

56 Burmistrz, M., Krakowski, K. and Krawczyk-Balska, A. (2020) RNA-targeting CRISPR-Cas systems and their applications. *Int. J. Mol. Sci.* **21**, 1122, https://doi.org/10.3390/ijms21031122

57 Johnsson, P., Ackley, A., Vidarsdottir, L., Lui, W.-O., Corcoran, M., Grandér, D. et al. (2013) A pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human cells. *Nat. Struct. Mol. Biol.* **20**, 440–446, https://doi.org/10.1038/nsmb.2516

58 Stojic, L., Lun, A.T.L., Mangei, J., Mascalchi, P., Quarantotti, V., Barr, A.R. et al. (2018) Specificity of RNAi, LNA and CRISPRi as loss-of-function methods in transcriptional analysis. *Nucleic Acids Res.* **46**, 5950–5966, https://doi.org/10.1093/nar/gky437

59 Matheny, C.J., Wei, M.C., Bassik, M.C., Donnelly, A.J., Kampmann, M., Iwasaki, M. et al. (2013) Next-generation NAMPT inhibitors identified by sequential high-throughput phenotypic chemical and functional genomic screens. *Chem. Biol.* **20**, 1352–1363, https://doi.org/10.1016/j.chembiol.2013.09.014

60 Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H. et al. (2014) Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* **159**, 647–661, https://doi.org/10.1016/j.cell.2014.09.029

61 Acosta-Alvear, D., Cho, M.Y., Wild, T., Buchholz, T.J., Lerner, A.G., Simakova, O. et al. (2015) Paradoxical resistance of multiple myeloma to proteasome inhibitors by decreased levels of 19S proteasomal subunits. *eLife* **4**, e08153, https://doi.org/10.7554/eLife.08153

62 Sidrauski, C., Tsai, J.C., Kampmann, M., Hearn, B.R., Vedantham, P., Jaishankar, P. et al. (2015) Pharmacological dimerization and activation of the exchange factor eIF2B antagonizes the integrated stress response. *eLife* **4**, e07314, https://doi.org/10.7554/eLife.07314

63 Smallegan, M.J. and Rinn, J.L. (2019) Linking long noncoding RNA to drug resistance. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 21963–21965, https://doi.org/10.1073/pnas.1915690116

64 Kapusta, A., Kronenberg, Z., Lynch, V.J., Zhuo, X., Ramsay, L., Bourque, G. et al. (2013) Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet.* **9**, e1003470, https://doi.org/10.1371/journal.pgen.1003470

65 Birmingham, A., Anderson, E.M., Reynolds, A., Ilsley-Tyree, D., Leake, D., Fedorov, Y. et al. (2006) 3′ UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nat. Methods* **3**, 199–204, https://doi.org/10.1038/nmeth854

66 Wu, N., Castel, D., Debily, M.-A., Vigano, M.A., Alibert, O., Mantovani, R. et al. (2011) Large scale RNAi screen reveals that the inhibitor of DNA binding 2 (ID2) protein is repressed by p53 family member p63 and functions in human keratinocyte differentiation. *J. Biol. Chem.* **286**, 20870–20879, https://doi.org/10.1074/jbc.M110.169433

67 Genovesio, A., Giardini, M.A., Kwon, Y.-J., de Macedo Dossin, F., Choi, S.Y., Kim, N.Y. et al. (2011) Visual genome-wide RNAi screening to identify human host factors required for Trypanosoma cruzi infection. *PLoS ONE* **6**, e19733, https://doi.org/10.1371/journal.pone.0019733

68 Collinet, C., Stöter, M., Bradshaw, C.R., Samusik, N., Rink, J.C., Kenski, D. et al. (2010) Systems survey of endocytosis by multiparametric image analysis. *Nature* **464**, 243–249, https://doi.org/10.1038/nature08779

69 Bodapati, S., Daley, T.P., Lin, X., Zou, J. and Qi, L.S. (2020) A benchmark of algorithms for the analysis of pooled CRISPR screens. *Genome Biol.* **21**, 62, https://doi.org/10.1186/s13059-020-01972-x

70 Michlits, G., Hubmann, M., Wu, S.-H., Vainorius, G., Budusan, E., Zhuk, S. et al. (2017) CRISPR-UMI: single-cell lineage tracing of pooled CRISPR-Cas9 screens. *Nat. Methods* **14**, 1191–1197, https://doi.org/10.1038/nmeth.4466

71 Schmierer, B., Botla, S.K., Zhang, J., Turunen, M., Kivioja, T. and Taipale, J. (2017) CRISPR/Cas9 screening using unique molecular identifiers. *Mol. Syst. Biol.* **13**, 945, https://doi.org/10.15252/msb.20177834

72 Ihry, R.J., Salick, M.R., Ho, D.J., Sondey, M., Kommineni, S., Paula, S. et al. (2019) Genome-scale CRISPR screens identify human pluripotency-specific genes. *Cell Rep.* **27**, 616.e6–630.e6, https://doi.org/10.1016/j.celrep.2019.03.043

73 Liu, Y., Yu, C., Daley, T.P., Wang, F., Cao, W.S., Bhate, S. et al. (2018) CRISPR activation screens systematically identify factors that drive neuronal fate and reprogramming. *Cell Stem Cell* **23**, 758.e8–771.e8, https://doi.org/10.1016/j.stem.2018.09.003

74 Bassaganyas, L., Popa, S.J., Horlbeck, M., Puri, C., Stewart, S.E., Campelo, F. et al. (2019) New factors for protein transport identified by a genome-wide CRISPRi screen in mammalian cells. *J. Cell Biol.* **218**, 3861–3879, https://doi.org/10.1083/jcb.201902028

75 Tian, R., Abarientos, A., Hong, J., Hashemi, S.H., Yan, R., Nalls, M.A. et al. (2021) Genome-wide CRISPRi/a screens in human neurons link lysosomal failure to ferroptosis. *bioRxiv*, https://doi.org/10.1101/2020.06.27.175679

76 Connerty, P., Lock, R.B. and de Bock, C.E. (2020) Long non-coding RNAs: major regulators of cell stress in cancer. *Front. Oncol.* **10**, 285, https://doi.org/10.3389/fonc.2020.00285

77 Clement, K., Hsu, J.Y., Canver, M.C., Joung, J.K. and Pinello, L. (2020) Technologies and computational analysis strategies for CRISPR applications. *Mol. Cell* **79**, 11–29, https://doi.org/10.1016/j.molcel.2020.06.012

78 König, R., Chiang, C., Tu, B.P., Yan, S.F., DeJesus, P.D., Romero, A. et al. (2007) A probability-based approach for the analysis of large-scale RNAi screens. *Nat. Methods* **4**, 847–849, https://doi.org/10.1038/nmeth1089

79 Diaz, A.A., Qin, H., Ramalho-Santos, M. and Song, J.S. (2015) HiTSelect: a comprehensive tool for high-complexity-pooled screen analysis. *Nucleic Acids Res.* **43**, e16, https://doi.org/10.1093/nar/gku1197

80 Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F. et al. (2014) MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554, https://doi.org/10.1186/s13059-014-0554-4

81 Daley, T.P., Lin, Z., Lin, X., Liu, Y., Wong, W.H. and Qi, L.S. (2018) CRISPhieRmix: a hierarchical mixture model for CRISPR pooled screens. *Genome Biol.* **19**, 159, https://doi.org/10.1186/s13059-018-1538-6

82 Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550, https://doi.org/10.1186/s13059-014-0550-8

83 Hart, T. and Moffat, J. (2016) BAGEL: a computational framework for identifying essential genes from pooled library screens. *BMC Bioinformatics* **17**, 164, https://doi.org/10.1186/s12859-016-1015-8

84 Li, W., Köster, J., Xu, H., Chen, C.-H., Xiao, T., Liu, J.S. et al. (2015) Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biol.* **16**, 281, https://doi.org/10.1186/s13059-015-0843-6

85 Allen, F., Behan, F., Khodak, A., Iorio, F., Yusa, K., Garnett, M. et al. (2019) JACKS: joint analysis of CRISPR/Cas9 knockout screens. *Genome Res.* **29**, 464–471, https://doi.org/10.1101/gr.238923.118

86 Meyers, R.M., Bryan, J.G., McFarland, J.M., Weir, B.A., Sizemore, A.E., Xu, H. et al. (2017) Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784, https://doi.org/10.1038/ng.3984

87 Sharma, S. and Petsalaki, E. (2018) Application of CRISPR-Cas9 based genome-wide screening approaches to study cellular signalling mechanisms. *Int. J. Mol. Sci.* **19**, 933, https://doi.org/10.3390/ijms19040933

88 Sasaki, Y.T.F., Ideue, T., Sano, M., Mituyama, T. and Hirose, T. (2009) MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2525–2530, https://doi.org/10.1073/pnas.0807899106

89 Zucchelli, S., Fasolo, F., Russo, R., Cimatti, L., Patrucco, L., Takahashi, H. et al. (2015) SINEUPs are modular antisense long non-coding RNAs that increase synthesis of target proteins in cells. *Front. Cell Neurosci.* **9**, 174, https://doi.org/10.3389/fncel.2015.00174

90 Thomson, D.W. and Dinger, M.E. (2016) Endogenous microRNA sponges: evidence and controversy. *Nat. Rev. Genet.* **17**, 272–283, https://doi.org/10.1038/nrg.2016.20

91 Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C.P., Jerby-Arnon, L. et al. (2016) Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *cell* **167**, 1853.e17–1866.e17, https://doi.org/10.1016/j.cell.2016.11.038

92 Adamson, B., Norman, T.M., Jost, M., Cho, M.Y., Nuñez, J.K., Chen, Y. et al. (2016) A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell* **167**, 1867.e21–1882.e21, https://doi.org/10.1016/j.cell.2016.11.048

93 Datlinger, P., Rendeiro, A.F., Schmidl, C., Krausgruber, T., Traxler, P., Klughammer, J. et al. (2017) Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods* **14**, 297–301, https://doi.org/10.1038/nmeth.4177

94 Aarts, M., Georgilis, A., Beniazza, M., Beolchi, P., Banito, A., Carroll, T. et al. (2017) Coupling shRNA screens with single-cell RNA-seq identifies a dual role for mTOR in reprogramming-induced senescence. *Genes Dev.* **31**, 2085–2098, https://doi.org/10.1101/gad.297796.117

95 Schraivogel, D., Gschwind, A.R., Milbank, J.H., Leonce, D.R., Jakob, P., Mathur, L. et al. (2020) Targeted Perturb-seq enables genome-scale genetic screens in single cells. *Nat. Methods* **17**, 629–635, https://doi.org/10.1038/s41592-020-0837-5

96 Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A. et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550, https://doi.org/10.1073/pnas.0506580102

97 Gene Ontology Consortium (2008) The Gene Ontology project in 2008. *Nucleic Acids Res.* **36**, D440–D444, https://doi.org/10.1093/nar/gkm883

98 Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C. et al. (2019) The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012, https://doi.org/10.1093/nar/gky1120

99 Kerimov, N., Hayhurst, J.D., Peikova, K., Manning, J.R., Walter, P., Kolberg, L. et al. (2021) eQTL Catalogue: a compendium of uniformly processed human gene expression and splicing QTLs. *bioRxiv*, https://doi.org/10.1101/2020.01.29.924266

100 Wainberg, M., Sinnott-Armstrong, N., Mancuso, N., Barbeira, A.N., Knowles, D.A., Golan, D. et al. (2019) Opportunities and challenges for transcriptome-wide association studies. *Nat. Genet.* **51**, 592–599, https://doi.org/10.1038/s41588-019-0385-z

101 Barbeira, A.N., Dickinson, S.P., Bonazzola, R., Zheng, J., Wheeler, H.E., Torres, J.M. et al. (2018) Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat. Commun.* **9**, 1825, https://doi.org/10.1038/s41467-018-03621-1

102 Saha, A., Kim, Y., Gewirtz, A.D.H., Jo, B., Gao, C., McDowell, I.C. et al. (2017) Co-expression networks reveal the tissue-specific regulation of transcription and splicing. *Genome Res.* **27**, 1843–1858, https://doi.org/10.1101/gr.216721.116

103 Guttman, M. and Rinn, J.L. (2012) Modular regulatory principles of large non-coding RNAs. *Nature* **482**, 339–346, https://doi.org/10.1038/nature10887

104 Marchese, F.P., Raimondi, I. and Huarte, M. (2017) The multidimensional mechanisms of long noncoding RNA function. *Genome Biol.* **18**, 206, https://doi.org/10.1186/s13059-017-1348-2

105 Chen, L.-L. (2016) Linking long noncoding RNA localization and function. *Trends Biochem. Sci.* **41**, 761–772, https://doi.org/10.1016/j.tibs.2016.07.003

106 Li, X., Zhou, B., Chen, L., Gou, L.-T., Li, H. and Fu, X.-D. (2017) GRID-seq reveals the global RNA-chromatin interactome. *Nat. Biotechnol.* **35**, 940–950, https://doi.org/10.1038/nbt.3968

107 Bonetti, A., Agostini, F., Suzuki, A.M., Hashimoto, K., Pascarella, G., Gimenez, J. et al. (2020) RADICL-seq identifies general and cell type-specific principles of genome-wide RNA-chromatin interactions. *Nat. Commun.* **11**, 1018, https://doi.org/10.1038/s41467-020-14337-6

108 Buske, F.A., Bauer, D.C., Mattick, J.S. and Bailey, T.L. (2012) Triplexator: detecting nucleic acid triple helices in genomic and transcriptomic data. *Genome Res.* **22**, 1372–1381, https://doi.org/10.1101/gr.130237.111

109 Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R. et al. (2020) A large-scale binding and functional map of human RNA-binding proteins. *Nature* **583**, 711–719, https://doi.org/10.1038/s41586-020-2077-3

110 Hendrickson, D.G., Kelley, D.R., Tenen, D., Bernstein, B. and Rinn, J.L. (2016) Widespread RNA binding by chromatin-associated proteins. *Genome Biol.* **17**, 28, https://doi.org/10.1186/s13059-016-0878-3

111 Xiao, R., Chen, J.-Y., Liang, Z., Luo, D., Chen, G., Lu, Z.J. et al. (2019) Pervasive chromatin-RNA Binding protein interactions enable RNA-based regulation of transcription. *Cell* **178**, 107.e18–121.e18, https://doi.org/10.1016/j.cell.2019.06.001

112 Le, T.D., Zhang, J., Liu, L. and Li, J. (2017) Computational methods for identifying miRNA sponge interactions. *Brief. Bioinform.* **18**, 577–590, https://doi.org/10.1093/bib/bbw042