

Specificity of RNAi, LNA and CRISPRi as loss-of-function methods in transcriptional analysis

Lovorka Stojic^{1*}, Aaron TL Lun^{1*}, Jasmin Mangei¹, Patrice Mascalchi¹, Valentina Quarantotti¹, Alexis R Barr², Chris Bakal², John C Marioni^{1,3,4}, Fanni Gergely¹ and Duncan T Odom^{1**}

5 ¹ Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK.

² Institute of Cancer Research, 237 Fulham Road London SW3 6JB, UK.

³ European Bioinformatics Institute, European Molecular Biology Laboratory (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK.

10 ⁴ Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK

* Equal contributions

**To whom correspondence should be addressed

Tel: +44 (0) 1223 769 655; Fax: +44 (0) 1223 769 881

15 Email: duncan.odom@cruk.cam.ac.uk

Present Address: *Patrice Mascalchi*: Bordeaux Imaging Center, UMS 3420 CNRS, US4 INSERM, University of Bordeaux, 33000 Bordeaux, France. *Jasmin Mangei*: German Cancer Research Center (DKFZ), Division: Molecular Genetics, Im Neuenheimer Feld 580, D-69120 Heidelberg

20 **Issue Section: RNA, Genomics, Computational biology**

ABSTRACT

Loss-of-function (LOF) methods, such as RNA interference (RNAi), antisense oligonucleotides or CRISPR-based genome editing, provide unparalleled power for studying the biological function of genes of interest. When coupled with transcriptomic analyses, LOF methods allow researchers to dissect networks of transcriptional regulation. However, a major concern is non-specific targeting, which involves depletion of transcripts other than those intended. The off-target effects of each of these common LOF methods have yet to be compared at the whole-transcriptome level. Here, we systematically and experimentally compared non-specific activity of RNAi, antisense oligonucleotides and CRISPR interference (CRISPRi). All three methods yielded non-negligible off-target effects in gene expression, with CRISPRi exhibiting clonal variation in the transcriptional profile. As an illustrative example, we evaluated the performance of each method for deciphering the role of a long noncoding RNA (lncRNA) with unknown function. Although all LOF methods reduced expression of the candidate lncRNA, each method yielded different sets of differentially expressed genes upon knockdown as well as a different cellular phenotype. Therefore, to definitively confirm the functional role of a transcriptional regulator, we recommend the simultaneous use of at least two different LOF methods and the inclusion of multiple, specifically designed negative controls.

INTRODUCTION

The ability to specifically reduce the expression of a particular coding or noncoding gene is fundamental for establishing its loss-of-function phenotype in cells and organisms, and is frequently the only way to infer its function. Three popular strategies for doing this are RNA interference (RNAi),
5 locked nucleic acid (LNA) hybridization and the CRISPR/Cas9 system (1). RNAi uses small interfering RNA oligonucleotides to deplete target transcripts by triggering their degradation (2). Efficient depletion of genes can also be achieved with antisense oligonucleotides (3); the most widely used antisense approach is locked nucleic acids (4), which can trigger RNase-H-mediated degradation of the target in the nucleus (5).

10 Most recently, the CRISPR/Cas9 system has been adapted to inhibit the expression of single genetic locus. Deactivated Cas9 (dCas9) fused to a Krüppel-associated box (KRAB) repression domain can be directed to a specific genomic locus to prevent transcription, an approach known as CRISPR interference (CRISPRi) (6-9). This allows repression of a targeted locus without editing the genome, thus avoiding unintentional deletion of active regulatory elements (10). CRISPRi can either
15 be applied in homogenous populations derived from single cells expressing dCas9-KRAB or in poly- or non-clonal populations (11,12). This approach can successfully deplete both protein-coding genes (13) and noncoding RNAs such as long noncoding RNAs (lncRNAs) (8,14-16) and miRNAs (17). The apparent high specificity of CRISPRi has made it a preferred alternative to conventional gene silencing methods, particularly for lncRNAs, which are frequently located in complex genomic
20 landscapes (10)(14).

Although all of these strategies have been successfully applied against a wide variety of targets in a range of biological systems, a critical consideration is whether the expression levels of non-target genes are inadvertently perturbed. These off-target effects arise from non-specific activity of the knockdown technology when exposed to the endogenous pool of total RNAs. For example,
25 RNAi and antisense oligonucleotides with sufficient complementarity to non-target transcripts may cause unintended repression of those genes (18-21). Widespread genome binding and modest off-target effects have also been reported for the dCas9-KRAB system (22,23).

In this study, we comprehensively quantify the off-target effects associated with each loss-of-function (LOF) strategy, with a particular focus on the transcriptome. Exploiting HeLa cell line as a
30 powerful and easily modifiable model, we performed RNA sequencing (RNA-seq) with a range of negative controls for each method and used differential expression analyses to identify the off-target effects. RNAi and LNA approaches were notably dependent on the precise sequences of in the siRNA/antisense oligonucleotides used in the negative controls. The impact of introducing dCas9-KRAB to generate a polyclonal population of HeLa cells caused few transcriptional perturbations; in
35 contrast, the process of single cell cloning from this population resulted in unique transcriptional signatures. To illustrate the impact that differences between methods can have on understanding gene function, we describe how the different LOF methods used to deplete a nuclear lncRNA can lead to different biological conclusions.

MATERIAL AND METHODS

Cell culture

HeLa and HEK293T cells were maintained in Dulbecco's modified Eagle's medium (Sigma Aldrich, D6429) supplemented with 10% Fetal bovine serum (FBS, Thermo Fisher Scientific). Both cell lines were obtained from American Type Culture Collection (ATCC) and were cultured at 37°C with 5% CO₂. HeLa Kyoto (EGFP- α -tubulin/ H2B-mCherry) cells were obtained from Jan Ellenberg (EMBL, Heidelberg) (24) and cultured in DMEM with 10% FBS (25). All cell lines were verified by short tandem repeat (STR) profiling and tested negative for mycoplasma contamination.

Single-molecule RNA fluorescent *in situ* hybridization (FISH)

Cells were grown on coverslips, briefly washed with PBS and fixed with PBS/3.7% formaldehyde at room temperature for 10 min. Following fixation, cells were washed twice with PBS. The cells were then permeabilised in 70% ethanol for at least 1 hour at 4°C. Stored cells were briefly rehydrated with Wash Buffer (2X SSC, 10% formamide, Biosearch) before FISH. The Stellaris FISH Probes (*Inc289* exonic probes Q570; sequences in Supplementary Table 4) were added to the hybridization buffer (2X SSC, 10% formamide, 10% dextran sulfate, Biosearch) at a final concentration of 250 nM. Hybridization was carried out in a humidified chamber at 37°C overnight. The following day, the cells were washed twice with Wash Buffer (Biosearch) at 37°C for 30 min each. The second wash contained DAPI for nuclear staining (5 ng/ml). The cells were then briefly washed with 2X SSC and then mounted in Vectashield (Vector Laboratories, H-1000). Images were captured using a Nikon TE-2000 inverted microscope with NIS-elements software, a Plan Apochromat 100x objective and an Andor Neo 5.5 sCMOS camera. We acquired 25 optical slices at 0.3 μ m intervals. Images were deconvolved with Huygens Professional and projected in two dimensions using ImageJ.

Plasmids and antibodies

Plasmids used in this study were pHR-SFFV-dCAS9-BFP-KRAB (Addgene, #46911), pU6-sgRNA EF1Alpha-puro-T2A-BFP (Addgene, #60955), second-generation packaging plasmid psPAX2 (Addgene, #12260) and the envelope plasmid pMD2.G (Addgene, #12259). pHIV-ZsGreen (Addgene #18121) and LincExpress-mCherry (modified version of pLenti6.3/TO/V5-DEST, kindly provided by John Rinn, Harvard University) were used as positive controls for transduction efficiency. Cas9 antibody was obtained from Cell Signaling (#14697, dilution 1:1000) and β -tubulin was purchased from Sigma (#T019, dilution 1:2000).

RNAi- and LNA-mediated gene depletion

HeLa cells were transfected with Lipofectamine RNAiMax reagent (Thermo Fischer Scientific) following the manufacturer's instructions. All experiments were done 48 hours after transfection. The siRNAs (Thermo Fischer Scientific) and LNA Gappers (Exiqon) were used at a final concentration of 50 nM and 25nM, respectively. siRNA and LNA sequences are listed in Supplementary Table 5 and 6, respectively.

Western blotting

Cells were grown in a 6 well plate, trypsinized, pelleted and washed twice with PBS. The pellet was lysed in lysis buffer (50 mM Tris-HCl, pH 8, 125 mM NaCl, 1% NP-40, 2 mM EDTA, 1 mM PMSF, and protease inhibitor cocktail [Roche]) and incubated on ice for 25 min. The samples were centrifuged for 3 min at 12 000 x g and 4°C. Supernatant was collected and protein concentration was determined using the Direct Detect® Spectrometer (Merck Millipore). The proteins (25 μ g) were denatured, reduced, and separated with Bolt® 4-12% Bis-Tris Plus Gel (Thermo Fisher Scientific) in MOPS buffer (Thermo Fisher Scientific, B0001-02). The proteins were then transferred to nitrocellulose membrane and blocked with 5% nonfat milk in TBS-T (50mM Tris, 150mM NaCl, 0.1% Tween-20) for 1 hour at

room temperature. The membranes were incubated with primary antibodies in 5% milk in TBS-T. After overnight incubation at 4°C, the membranes were washed with TBS-T and incubated with HRP secondary antibodies (GE Healthcare Life Sciences, 1:5000), and immunobands were detected with a Supersignal West Dura HRP Detection Kit (Thermo-Scientific). An uncropped scan of the immunoblot (Figure 2) is shown in Supplementary figure 12.

Time-lapse microscopy

HeLa cells (10 000 cells) were cultured in eight-well chamber slides (Ibidi) with 200 µl/well of normal HeLa medium (DMEM, 10% FBS). 30 min before live-cell imaging, the medium was replaced with imaging medium (DMEM fluorobrite, A1896701, Thermo Fisher Scientific, supplemented with 10% FBS and 4mM Glutamax) containing 300nM SiR–Hoechst (Spirochrome). SiR–Hoechst was present in the medium throughout imaging. HeLa Kyoto cells were plated in the same way but imaging was performed in DMEM medium with 10% FBS. Time-lapse microscopy was performed for both cell lines 48 hours after transfection with LNA or CRISPRi transduction. Mitotic duration was measured as the time from nuclear envelope breakdown (NEBD) until anaphase onset, based on visual inspection of the images. Live-cell imaging was performed using a Zeiss Axio Observer Z1 microscope equipped with a PL APO 0.95NA 40X dry objective (Carl Zeiss Microscopy) fitted with a LED light source (Lumencor) and an Orca Flash 4.0 camera (Hamamatsu). Four positions were placed per well and a z-stack was acquired at each position every 10 minutes for a total duration of 12 hours. Voxel size was 0.325 µm x 0.325 µm x 2.5 µm. Zen software (Zeiss) was used for data collection and analysis. Throughout the experiment, the cells were maintained in a microscope stage incubator at 37 °C in a humidified atmosphere of 5% CO₂.

CRISPR interference (CRISPRi) and gRNA design

For CRISPRi, we used two negative control sgRNAs and one positive control sgRNA (against the *H19* lncRNA) (8). For *Inc289*, *H19* and *ch-TOG*, we designed gRNA sequences (20 nt) targeting a genomic window of -50 to +200 bp relative to the transcription start site (TSS) (Supplementary Table 7). The location of the TSS was determined using the NCBI RefSeq database. The MIT CRISPR (<http://crispr.mit.edu>) and the gUIDEbook™ gRNA design (Desktop Genetics Ltd) web tools were used to design the gRNA sequence. Potential off-target effects were analysed with the MIT CRISPR and the CRISPR RGEN Cas-OFFinder web tools, for which the position of off-target alignments with equal or less than 4 mismatches were separately checked using the Basic Local Alignment Search Tool. Additional sequences were added to the sgRNA sequences to obtain compatible sticky ends for cloning the DNA insert into the 5'BstXI-BIPI3' digested backbone of a pU6-sgRNA EF1Alpha-puro-T2A-BFP expression plasmid. gRNA oligos were phosphorylated, annealed and cloned into pU6-sgRNA EF1Alpha-puro-T2A-BFP expression plasmid. All inserts were verified with Sanger sequencing.

Lentiviral transduction

To produce lentivirus, 4 x 10⁶ of HEK293T cells were plated in a 10 cm dish one day prior to transduction. Next day, cells were transfected with 15 µg of DNA, composed of 9 µg of the lentiviral vector DNA containing the transgene, 4 µg of psPAX.2 and 2 µg of pMD2.G in the final transfection volume of 1.5 ml (including 45 ul of Trans-Lt1 transfection reagent, Mirus) using OptiMEM medium (Thermo Fisher Scientific). As a positive control for viral infection and to control for any possible effects of lentiviral delivery, we transduced cells with polybrene (5 µg/ml, Sigma) or with pHIV-ZsGreen and LincExpress-mCherry vectors. The transfection mixture was incubated for 25 min at room temperature. Prior to transfection, old medium was replaced by 14 ml of fresh medium and transfection mix was added dropwise to the cells and incubated for 24 hours at 37°C. The following day, old medium was replaced by 7 ml of fresh medium and incubated for another 24 hours at 37°C. Viral supernatant was collected 48 and 72 hours post transfection, spun down at 1800 x g for 5 min at +4°C, and filtered through a 45 µm filter. Ready-to-use virus was stored at +4°C. For long-term storage, viral supernatant was frozen at -80°C.

FACS analysis and cell sorting

HeLa cells were transduced with lentivirus containing the pHR-SFFV-dCAS9-BFP-KRAB vector together with polybrene (5 µg/ml, Sigma). 24 hours after lentivirus transduction, the medium was replaced and the cells were incubated for another 48 hours. HeLa cells were then sorted for the BFP-expressing cells using the BD FACSAria III cell sorter (CRUK Flow Cytometry Core Facility). The expression of BFP fluorescent proteins was detected using MACSQuant VYB (Miltenyi Biotec) and the data were analysed using the FlowJo v7.1 software. BFP-sorted HeLa cells were used for single cell cloning in 96-well plate (clonal cells) or to create a stable non-clonal cell population.

CRISPRi-mediated depletion

Three to four days after FACS sorting, dCas9-KRAB transduced cells were plated on 12-well plates and infected with lentivirus containing gRNAs targeting *Inc289*, *ch-TOG* or *H19*, or with lentivirus containing two negative guide RNAs. The lentivirus was diluted with HeLa medium (1:1 dilution) and cationic polymer polybrene was added to facilitate viral transduction (5 µg/ml, Sigma). After a 24 hour incubation, supernatant was removed and fresh medium was added for another 48 hour incubation before RNA collection to evaluate knockdown of the target gene. Non-transduced cells did not receive virus and were used as a negative control.

RNA extraction, cDNA and Quantitative Real-Time PCR (qPCR)

RNA (1 µg) was extracted with the RNeasy Kit (QIAGEN, 74106) and treated with DNase I following the manufacturer's instructions (QIAGEN, 79254). The QuantiTect Reverse Transcription Kit (QIAGEN, 205313) was used for cDNA synthesis including an additional step to eliminate genomic DNA contamination. Quantitative real-time PCR (qPCR) was performed on a 7900HT Fast Real-Time PCR System (Applied Biosystems) with Fast SYBR Green Master Mix (Life Technologies). Thermocycling parameters were defined as 95°C for 20 sec followed by 40 cycles of 95°C for 1 sec and 60°C for 20 sec. Two reference genes (*GAPDH* and *RPS18*) were used to normalise expression levels using the $2^{-\Delta\Delta CT}$ method. Sequences of qPCR primers are provided in Supplementary Table 8.

Subcellular fractionation

RNA was fractionated as described previously (26). Briefly, cells from a 150 mm dish were used to isolate RNA from cytoplasmic, nucleoplasmic and chromatin fractions by TRIZOL extraction (Life Technologies). Expression of target genes in each fraction was analysed by qPCR. Data were normalised to the geometric mean of *GAPDH* and *β-actin* levels in each cellular compartment. *MALAT1* and *RPS18* were used as positive controls for chromatin and cytoplasmic fractions, respectively.

RNA library preparation, sequencing and analysis

RNA-seq libraries were prepared from HeLa cells using TruSeq Stranded Total RNA Kit with Ribo-Zero Gold (Illumina, RS-122-2303). We performed four biological replicates for RNAi, LNAs and CRISPRi-mediated depletion of *Inc289* and *H19*. Indexed libraries were PCR-amplified and sequenced for 125bp paired-end reads on an Illumina HiSeq 2500 instrument (CRUK Genomics Core Facility). Each library was sequenced to a depth of 20-30 million read pairs. Paired-end reads were aligned to the human genome hg38 (27) and the number of read pairs mapped to the exonic regions of each gene was counted for each library (28). Approximately 80% of read pairs contained one read that was successfully mapped to the human reference genome. On average, 74% of all read pairs in each library were assigned into exonic regions and counted. Any outlier samples with very low depth (resulting from failed library preparation or sequencing) were removed prior to further analysis. Differential gene expression analyses were performed using the voom-limma framework (29), where we tested for differential expression above a \log_2 -fold change threshold of 0.5 in pairwise contrasts

between groups of samples. For each contrast, genes with significant differences in expression between groups were detected at a false discovery rate (FDR) of 5%.

Determination of noncoding potential of lncRNAs

5 The Coding-Potential Calculator (CPC) ((30), <http://cpc.cbi.pku.edu.cn>) and Coding Potential Assessment Tool (CPAT) ((31), <http://lilab.research.bcm.edu/cpat/index.php>) were used to determine noncoding potential. LncRNAs with CPC score >1 and CPAT score >0.364 were predicted to have protein-coding capacity. The PhyloCSF score was taken from UCSC (<https://github.com/mlin/PhyloCSF/wiki>, (32)).

Statistical analysis

10 The statistical significance of data was determined by two-tailed Student's t-test in all experiments using GraphPad Prism unless indicated otherwise. P-values > 0.05 were considered statistically not significant. The differential expression analysis is described in detail in the Supplementary Methods.

RESULTS

LNA and RNAi technologies are associated with non-negligible off-target effects

We first verified that RNAi (pool of four different siRNAs) and LNA oligonucleotides were able to successfully deplete both protein-coding genes and lncRNAs by knocking down *Ch-TOG/CKAP5* and *MALAT1*, respectively (Supplementary Figure 1). Both technologies routinely achieved high knockdown efficiency, indicating that gene depletion was effective.

We then performed RNA-seq on untreated cells and cells treated with one of two negative control siRNAs (Ambion and GE Dharmacon) (Figure 1A). In principle, treatment with the negative control siRNAs should have no effect on gene expression, as no gene is targeted for depletion. Thus, our experimental design allows us to quantify the transcriptional off-target effects of the two negative control siRNAs, based on the number of differentially expressed genes (DEGs) when compared to untreated cells. In this analysis, genes are only considered to be subject to off-target effects if the log-fold change between conditions is significantly greater than 0.5 (33). This ensures that only modest-to-large changes in gene expression are detected, which focuses on DEGs that are more likely to have some biological effect.

We identified 45 (GE Dharmacon) and 169 genes (Ambion) affected by the introduction of each negative control siRNA at a FDR of 5%, some of which we validated by qPCR (Supplementary Figure 2A). Of these, 30 genes were affected by both negative control siRNAs, likely representing a general effect of siRNA transfection (Supplementary Figure 2B). The off-target genes affected by either of the negative control siRNAs were not obviously associated with a functional pathway. No single KEGG term contained more than 20% of the DEGs (Supplementary Figure 3A and B, Supplemental Table 1), with most terms containing less than 10%. These results indicate that off-target genes associated with addition of negative control siRNAs do not fall into pathways that would be easy to computationally predict or remove.

We further identified DEGs between two cell cultures, each of which had been treated separately with one of the two negative control siRNAs. In this comparison, only the siRNA sequence differs between the two controls – thus, any DEGs between the controls must represent sequence-dependent off-target effects, rather than a general effect of siRNA transfection. Comparison between the negative control treatments yielded 56 DEGs, indicating that the perturbations due to nonspecific targeting is dependent on the exact sequence of the siRNA used to treat the cells. Again, no common function for this set of DEGs was detectable by KEGG pathway analysis (Supplementary Figure 3C, Supplemental Table 1), with fewer than 10% of DEGs associated with any KEGG term.

We also generated transcriptomic profiles for negative controls at each step of the depletion protocol using locked nucleic acid (LNA) Gapmers (Exiqon). Treating cells only with transfection reagent led to perturbation of 31 genes (Figure 1B) compared to untreated cells. Treatment using either of two different negative control LNAs yielded zero (control LNA A, Exiqon part number 300611-00) and 113 (control LNA B, Exiqon part number 300615-00) DEGs compared to the transfection control, which we validated by qPCR (Supplementary Figure 4A). Comparison between the two negative control LNAs identified 89 DEGs. Applying the same reasoning as described above for RNAi, these 89 genes represent the typical scale of sequence-dependent off-target effects of the LNA approach. Similar to RNAi, KEGG pathway analysis of off-target effects using negative LNA controls did not reveal any common function for the DEGs (Supplementary Figure 4B and C; Supplemental Table 1).

In summary, a non-negligible number of genes display off-target activity with both RNAi and LNA technologies. The sequence-dependent nature of the off-target effects has important implications for how these methods can be used to study transcriptional regulation. In particular, the high sequence specificity of off-target effects in both methods strongly suggests that generic negative

controls cannot accurately recapitulate nonspecific changes in expression that arise when targeting a particular gene.

CRISPRi with single cell cloning introduces transcriptional variation

5 CRISPRi-mediated transcriptional inhibition can target gene expression using both non-clonal (8,9,15) and single cell derived clonal populations (11,15). To directly compare CRISPRi with other LOF methods, we generated HeLa cells expressing dCas9-KRAB using lentiviral transduction, and confirmed the expression of dCas9-KRAB in both clonally isolated populations of cells and non-clonally isolated populations (Figure 2A). We verified that our CRISPRi system was effective at knocking down protein-coding genes by targeting *Ch-TOG/CKAP5*, a microtubule associated protein required for mitotic spindle assembly (Supplementary Figure 5A) (34). Moreover, depletion of *Ch-TOG* with CRISPRi recapitulated known cellular phenotype such as mitotic delay (Supplementary Figure 5B) (35).

15 We then transduced one CRISPRi clone (CRISPRi clone 2) and a non-clonal cell population with two negative control guide RNAs (negative control guide 1 or 2). We performed RNA-seq to quantify the off-target effects by comparing the transcriptional profile before and after treatment with each of the negative control guides. The addition of the negative guide RNAs had modest effects in the clonal CRISPRi cells (8 and 106 genes, Figure 2B) and minor effects in the non-clonal CRISPRi cells (1 and 8 genes, Figure 2C). This is noticeably lower than the extent of off-target effects associated with RNAi or LNAs, consistent with previous studies suggesting that CRISPRi-mediated gene repression is highly specific (7). Moreover, only 2 genes were differentially expressed between the negative guide RNAs in each CRISPRi strategy. These results indicate that the off-target effects of CRISPRi are minor, not specific to the guide sequence, and can be expected to cancel out when different guide RNA treatments (negative control or targeting a specific gene) are compared.

25 Single cell cloning after introduction of dCas9-KRAB provides another potential source of off-target effects. Indeed, we found 203 DE genes between the parental HeLa cells and clonally-derived cells expressing dCas9-KRAB in the absence of a guide RNA (Figure 2B); in contrast, no such difference (only three genes) was observed between the parental HeLa cells and the non-clonal cells (Figure 2C). To investigate whether the single cell dilution protocol for obtaining CRISPRi targeted cell clones could be responsible for the increase in the number of DEGs, we profiled additional CRISPRi clones expressing different levels of dCas9-KRAB (clones 1 and 3; Figure 2D) along with additional independent replicates of CRISPRi clone 2. The DEG analysis yielded a core set of 37 DE genes that were consistently and strongly downregulated in all three clones compared to non-transduced control cells (Figure 2D and E). We did not observe any common pathway for these 37 genes using KEGG analysis (Supplementary Figure 6, Supplementary Table 1) and found no relationship between these genes and their chromosomal location (Supplementary Figure 7).

40 Most of the genes (33 out of 37) were downregulated in all clones, suggesting that they could potentially be direct targets of the repressive KRAB domain. We validated the repression of two of these 37 genes (*SCIN* and *CDH2*) by qPCR, as well as in HeLa cells that were transduced with dCas9-KRAB but without single cell clonal selection (Figure 2F). Importantly, these changes in gene expression are not caused by lentiviral transduction (Supplementary Figure 8), but are a consequence of single cell cloning. Indeed, no equivalent effect was observed with non-clonal cells (Supplementary Figure 9). Thus, stable expression of dCas9-KRAB causes marked changes in the transcriptomic background in addition to the depletion of gene of interest.

45 Widespread genome binding and modest off-target effects have been reported for the dCas9-KRAB system (22,23,36). Therefore, we compared the published binding sites for dCas9 in HEK293T cells (37) with the gene bodies for our core set of 37 differentially expressed genes (Figure 2E), and found no overlap. We performed a similar analysis using the genome-wide mapping of dCas9-KRAB in K562 cells (23) and found only two dCas9-KRAB binding events in these 37 genes. We also found no overlap between our core set of genes with those reported as being differentially expressed upon

dCas9-KRAB transduction (23). These poor overlaps may reflect the potential dependency of CRISPRi off-target effects on the cell type, the epigenetic landscape, or other factors (38). More generally, a common blacklist for *in silico* removal of likely affected genes is unlikely to be effective. The absence of a consistent set of genes across three independent studies that have attempted to identify common CRISPRi off-targets suggests that purely computational approaches may not be able to account for off-target effects *a priori*.

A case study in using LOF methods to deplete a nuclear long noncoding RNA

We applied the three most commonly used LOF methods – RNAi, LNA oligonucleotides and CRISPRi – to study the regulatory function of a previously uncharacterised lncRNA in HeLa cells. This represents a common use of LOF strategies, given that tens of thousands of lncRNAs exist in the mammalian genome (39,40). Many of these are involved in regulation of diverse cellular processes acting at the transcriptional, post-transcriptional and post-translational level (41-43), but the function of the majority of lncRNAs is unknown. LNAs have been shown to be particularly efficient for depletion of nuclear lncRNAs (44-46). The CRISPRi system has also been used successfully to deplete lncRNAs in a large-scale screen (12). In specific cases, RNAi was also shown to be effective in depleting lncRNAs (26,47), including nuclear lncRNAs (most likely due to the presence of active RNAi machinery in the nucleus (48)).

We selected a prototypical lncRNA following published guidelines (10) with specific characteristics, including: (i) previously uncharacterized, (ii) consistently expressed at more than one molecule per cell, (iii) low coding potential, (iv) chromatin hallmarks of active transcription, and (v) nuclear localization. Using these criteria, we chose *loc100289019* (also known as *SLC25A25-AS1*) as a spliced and functionally uncharacterized intragenic lncRNA with three promoters and a single 3' polyadenylation site (Figure 3A); hereafter called *lnc289*. Among all ENCODE cell lines, *lnc289* was most highly expressed in the nucleus of HeLa cells (Figure 3B). We confirmed that *lnc289* has low protein coding potential using PhyloCSF, Coding-Potential Calculator (CPC) and Coding-Potential Assessment Tool (CPAT) (Figure 3C). Furthermore, computational analysis of previously published ribosomal occupancy data indicated no translation of the *lnc289* transcript in HeLa cells (49) (Figure 3A, Riboseq track). Nevertheless, the genomic locus was actively transcribed based on the presence of both histone H3 lysine 4 trimethylated (H3K4me3) and histone H3 lysine 27 acetylated (H3K27ac) histones at the predicted *lnc289* locus (Figure 3A). We experimentally confirmed nuclear localization of *lnc289* by single-molecule RNA FISH (Figure 3D), and demonstrated by cellular fractionation that *lnc289* is enriched in chromatin (Figure 3E). We note that the vast majority of the lncRNAs currently under active investigation satisfy our selection criteria above. This suggests that our experimental evaluation of the three most widely used LOF methods is likely to be relevant to most studies of lncRNA function.

We then used the different LOF methods to identify the transcript(s) robustly regulated by *lnc289*, regardless of the method used for depletion. Only modest reduction of *lnc289* levels was observed with RNAi, while the other methods were able to successfully reduce *lnc289* levels by at least 50% (Figure 4 A and B). As such, we discarded the RNAi results and attempted to identify a common set of DEGs that were detected with LNA and CRISPRi methods (clonal and non-clonal cells). Remarkably, the only transcript common to these methods was *lnc289* itself (Figure 4C, Supplemental Table 2). However, the depletion of *lnc289* with LNA resulted in hundreds of DEGs, whereas many fewer DEGs were observed with CRISPRi-based methods.

The fact that *lnc289* is the only gene transcriptionally impacted by depletion with LNA and CRISPRi suggests two possible explanations for our results: i) *lnc289* has no function in transcriptional regulation, the lack of detection with the CRISPRi-based methods is correct, and the DE genes identified with LNA knockdown are off-target effects; or ii) *lnc289* does regulate the transcription of other genes, the LNA method correctly identifies its downstream targets, and the CRISPRi-based methods are somehow failing to recapitulate the effect.

5 These two explanations are mutually exclusive and cannot be easily distinguished, as the underlying problem stems from the deficiencies of our experimental tools for perturbing the biological system. We could, perhaps, infer that option (i) might be more likely if the number of DE genes after LNA depletion of *Inc289* were comparable to the number of off-target genes. However, the former (370, Figure 4B) is four times larger than the latter (89, Figure 1B), which would require large variation in the extent of the off-target effects between oligonucleotide sequences. Option (ii) is equally unappealing as it requires complex mechanisms e.g., compensatory effects in CRISPRi-depleted cells to mask the cellular phenotype (1) (50), or differences in the aspects of lncRNA biology that are disrupted by LNA and CRISPRi (1) (51).

10 We also examined whether the changes to the transcriptome might impact the *Inc289* knockdown phenotype. We noticed that two of the top DEGs upon LNA-mediated knockdown of *Inc289*, namely *SEPTIN2* and *GM130* (Figure 4D), had known roles in mitosis (52,53). Depletion of *SEPTIN2* leads to mitotic delay and incomplete cytokinesis whereas inhibition of *GM130* function by antibody injection leads to mitotic delay and multipolar division. Therefore, we assayed if mitotic delay also occurred as a result of the off-target effects upon *Inc289* depletion. To investigate this, we quantified the time required for HeLa cells to transition through mitosis before and after knockdown of *Inc289*. We observed a significant mitotic delay after LNA-mediated knockdown of *Inc289* (Figure 4E), whereas no such effect was observed with the CRISPRi-based methods. We further confirmed the mitotic delay with additional LNA oligonucleotide (Figure 4F) in HeLa Kyoto cells, a cell line stably expressing EGFP-alpha tubulin and mCherry-histone H2B (24). This demonstrates that the differences in the genes disrupted by each method have real consequences on the inferred biological function. Using LNA oligonucleotides, one might conclude that *Inc289* regulates mitosis, whereas the same conclusion cannot be made with CRISPRi.

25 We further tested CRISPRi to deplete *H19*, a multifunctional and well-characterized lncRNA with activity in the nucleus and in the cytoplasm (54)(55). Knockdown efficiency was similar in clonal and non-clonal cells (Supplementary Figure 10A), though there were modest differences in the number of DEGs by RNA-seq. Specifically, we observed 5 and 29 DEGs in clonal and non-clonal populations, respectively, compared to cells treated with negative guides (Supplementary Figure 10B; Supplementary Table 3). This difference hints at the presence of compensatory mechanisms in clonal cells that may be countering dCas9-KRAB activity, possibly as a result of the altered transcriptional background (Supplementary Figure 9B). We also examined the expression of genes previously reported to be regulated by *H19* from experiments using RNAi in different human cell lines (56) (57) (58) or genetic deletion in mice (59), but we observed no evidence of differential expression for these genes in HeLa cells after *H19* depletion (Supplementary Figure 10C). This highlights one challenge in using CRISPRi-mediated depletion to infer lncRNA function, as they often operate in a cell type-specific manner (12) (45).

DISCUSSION

Here, we systematically compared three widely used LOF methods and evaluated the transcriptome-wide changes attributable to each individual method. We describe off-target effects associated with each LOF method that need to be considered when investigating gene function, consistent with previous studies (18-20,22,23,36,60). In particular, we identified large off-target effects in the RNAi and LNA methods, which were highly dependent on the siRNA or LNA oligonucleotide sequence. While CRISPRi was less sensitive to the guide sequence, the introduction of dCas9-KRAB provides another source of off-target effects that can significantly change the transcriptional context in the depleted cells. Single cell cloning of dCas9-KRAB-expressing cells results in strong transcriptional changes even in the absence of guide RNAs, indicating that polyclonal populations should be used for CRISPRi experiments.

Differences between the three LOF methods can also lead to significant differences in the molecular or cellular phenotype after depletion of a gene of interest, as observed in our case study with a nuclear lncRNA *lnc289*. Our results are consistent with previous studies of lncRNA depletion in mammalian cells (61,62) and zebrafish (63,64). Different strategies to deplete lncRNAs in mice have also yielded different phenotypes (65,66). Such discrepancies are not limited to lncRNAs, but have been observed when depleting protein-coding genes using RNAi and CRISPR-based methods (8,51,67,68) as well as in high-throughput screens (51) (69).

Our results suggest that CRISPRi in non-clonal populations of dCas9-KRAB-expressing cells provides the cleanest depletion of the target gene, with the fewest off-target effects (sequence-dependent or otherwise). This is consistent with previous studies demonstrating the superiority of CRISPRi compared to RNAi (7,12,60,69). However, CRISPRi has a number of limitations, especially when investigating the function of lncRNAs. Currently, CRISPRi can not differentiate *cis*- and *trans*-acting functions of RNA transcripts (41), *cis*-mediated regulation related to lncRNA transcription (70-72) and/or enhancer-like functions of some lncRNA loci (73-75). In addition, CRISPRi cannot be used to target bidirectional promoters (76). Similarly, CRISPRi is not ideal for targeting lncRNAs near other transcriptional units (14), as neighboring genes may be unintentionally repressed.

For studies that use LOF methods to characterize the regulatory roles of targeted genes, our recommendations are to generate libraries (1) from controls obtained at each step of the method and (2) from multiple negative control sequences. This allows accurate quantification of the extent of transcriptional off-target effects introduced upon sequential manipulations in the LOF protocol. Affected genes can then be excluded from the DE analysis in samples where the gene of interest has been depleted, thus reducing the impact of off-target effects on the biological conclusions. It may also be necessary to discard experiments where the number or log-fold changes of the off-target genes are comparable to or greater than the number or log-fold changes of DEGs detected upon knockdown of the target gene. In such cases, there is no meaningful way to distinguish between off-target effects and genuine knockdown effects.

Furthermore, we recommend performing differential expression analyses with a minimum log-fold change threshold, in order to avoid detecting genes with small changes in expression. Indeed, when we repeated the analyses using a test for any differential gene expression (i.e., without a minimum log₂-fold change threshold), we obtained a far greater number of affected genes in all comparisons for all depletion methods (Supplementary Figure 11). In particular, we observed a large increase in the number of genes detected in comparisons between control groups for all methods, corresponding to a disproportionate increase in the scale of the off-target effects. The use of a log-fold change threshold mitigates the effect of non-specific activity on the results for each LOF method, by focusing on larger and arguably more biologically relevant effects of depletion.

Finally, the prevalence of off-target effects in most of the LOF methods emphasizes the need for rigorous validation of putative downstream targets of the target gene, such as independent rescue experiments. In the case of RNAi, off-target effects could be reduced by using alternative approaches

such as siPOOL (77) or C911 mismatch siRNAs (78). It remains to be seen to what extent other CRISPR genome editing strategies, including whole locus or promoter deletion, insertion of transcriptional termination sites into the gene body as well as CRISPR-Cas13 system (79), may have off-target effects that impact functional characterization of gene of interest.

- 5 Modern LOF technologies allow researchers to deplete any transcript of interest in a variety of biological systems, and provide an essential experimental toolkit for studying the biological function of transcripts and dissecting networks of transcriptional regulation. Here, we have empirically formulated recommendations to minimize technical artefacts and avoid - or at least prudently manage – the off-target effects of these commonly used LOF methods.

10

DATA AVAILABILITY

Sequencing data are available in the ArrayExpress database (<http://www.ebi.ac.uk/arrayexpress>) under the accession number E-MTAB-5308.

SUPPLEMENTARY DATA

- 5 Supplementary Data are available at NAR online.

FUNDING

10 This work was made possible by funding from Cancer Research UK (C14303/A17197 to F. G and A20412 to D.T.O). We also acknowledge the support of the University of Cambridge, the Wellcome Trust Investigator Award (D.T.O), European Research Council (615584; D.T.O) and Hutchison Whampoa Limited.

Conflict of interest statement

None declared.

Authors' contributions

15 Conception and design of study: LS, ATLL, FG and DTO. Data acquisition: LS, ATLL, JM, PM, ARB, VQ. Data analyses and interpretation: LS, ATLL, FG, JCM, CB, DTO. Writing the paper: LS, ATLL and DTO with input from all of the authors wrote the manuscript.

ACKNOWLEDGEMENTS

20 We thank all the members of Odom, Gergely and Marioni groups for helpful discussions. We also thank the Genomics, FACS and Microscopy Core Facilities at the CRUK Cambridge Institute and John Rinn for providing us with lincExpress vector.

REFERENCES

1. Housden, B.E., Muhar, M., Gemberling, M., Gersbach, C.A., Stainier, D.Y., Seydoux, G., Mohr, S.E., Zuber, J. and Perrimon, N. (2016) Loss-of-function genetic tools for animal models: cross-species and cross-platform differences. *Nat Rev Genet*.
- 5 2. Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. and Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **411**, 494-498.
3. Sahu, N.K., Shilakari, G., Nayak, A. and Kohli, D.V. (2007) Antisense technology: a selective tool for gene expression regulation and gene targeting. *Curr Pharm Biotechnol*, **8**, 291-304.
- 10 4. Sarma, K., Levasseur, P., Aristarkhov, A. and Lee, J.T. (2010) Locked nucleic acids (LNAs) reveal sequence requirements and kinetics of Xist RNA localization to the X chromosome. *Proc Natl Acad Sci U S A*, **107**, 22196-22201.
5. Tse, M.T. (2012) Antisense therapeutics: Nuclear RNA more susceptible to knockdown. *Nat Rev Drug Discov*, **11**, 674.
- 15 6. Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P. and Lim, W.A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, **152**, 1173-1183.
7. Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A. *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, **154**, 442-451.
- 20 8. Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C. *et al.* (2014) Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, **159**, 647-661.
9. Kearns, N.A., Genga, R.M., Enuameh, M.S., Garber, M., Wolfe, S.A. and Maehr, R. (2014) Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. *Development*, **141**, 219-223.
- 25 10. Bassett, A.R., Akhtar, A., Barlow, D.P., Bird, A.P., Brockdorff, N., Duboule, D., Ephrussi, A., Ferguson-Smith, A.C., Gingeras, T.R., Haerty, W. *et al.* (2014) Considerations when investigating lncRNA function in vivo. *Elife*, **3**, e03058.
- 30 11. Mandegar, M.A., Huebsch, N., Frolov, E.B., Shin, E., Truong, A., Olvera, M.P., Chan, A.H., Miyaoka, Y., Holmes, K., Spencer, C.I. *et al.* (2016) CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. *Cell Stem Cell*, **18**, 541-553.
12. Liu, S.J., Horlbeck, M.A., Cho, S.W., Birk, H.S., Malatesta, M., He, D., Attenello, F.J., Villalta, J.E., Cho, M.Y., Chen, Y. *et al.* (2017) CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science*, **355**.
- 35 13. Horlbeck, M.A., Gilbert, L.A., Villalta, J.E., Adamson, B., Pak, R.A., Chen, Y., Fields, A.P., Park, C.Y., Corn, J.E., Kampmann, M. *et al.* (2016) Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *Elife*, **5**.
14. Goyal, A., Myacheva, K., Gross, M., Klingenberg, M., Duran Arque, B. and Diederichs, S. (2016) Challenges of CRISPR/Cas9 applications for long non-coding RNA genes. *Nucleic Acids Res*.
- 40 15. Liu, X., Gally, C., Kjos, M., Domenech, A., Slager, J., van Kessel, S.P., Knoops, K., Sorg, R.A., Zhang, J.R. and Veening, J.W. (2017) High-throughput CRISPRi phenotyping identifies new essential genes in *Streptococcus pneumoniae*. *Mol Syst Biol*, **13**, 931.
- 45 16. Ghosh, S., Tibbit, C. and Liu, J.L. (2016) Effective knockdown of *Drosophila* long non-coding RNAs by CRISPR interference. *Nucleic Acids Res*, **44**, e84.
17. Zhao, Y., Dai, Z., Liang, Y., Yin, M., Ma, K., He, M., Ouyang, H. and Teng, C.B. (2014) Sequence-specific inhibition of microRNA via CRISPR/CRISPRi system. *Sci Rep*, **4**, 3943.
18. Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S.V., Burchard, J., Mao, M., Li, B., Cavet, G. and Linsley, P.S. (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol*, **21**, 635-637.
- 50 19. Krueger, U., Bergauer, T., Kaufmann, B., Wolter, I., Pilk, S., Heider-Fabian, M., Kirch, S., Artz-Oppitz, C., Isselhorst, M. and Konrad, J. (2007) Insights into effective RNAi gained from large-scale siRNA validation screening. *Oligonucleotides*, **17**, 237-250.
- 55 20. Sigoillot, F.D., Lyman, S., Huckins, J.F., Adamson, B., Chung, E., Quattrochi, B. and King, R.W. (2012) A bioinformatics method identifies prominent off-targeted transcripts in RNAi screens. *Nat Methods*, **9**, 363-366.

21. Adamson, B., Smogorzewska, A., Sigoillot, F.D., King, R.W. and Elledge, S.J. (2012) A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. *Nat Cell Biol*, **14**, 318-328.
- 5 22. Kucsu, C., Arslan, S., Singh, R., Thorpe, J. and Adli, M. (2014) Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol*, **32**, 677-683.
- 10 23. Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E. and Gersbach, C.A. (2015) Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods*, **12**, 1143-1149.
24. Neumann, B., Walter, T., Heriche, J.K., Bulkescher, J., Erfle, H., Conrad, C., Rogers, P., Poser, I., Held, M., Liebel, U. *et al.* (2010) Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature*, **464**, 721-727.
- 15 25. Neumann, B., Held, M., Liebel, U., Erfle, H., Rogers, P., Pepperkok, R. and Ellenberg, J. (2006) High-throughput RNAi screening by time-lapse imaging of live human cells. *Nat Methods*, **3**, 385-390.
- 20 26. Stojic, L., Niemczyk, M., Orjalo, A., Ito, Y., Ruijter, A.E., Uribe-Lewis, S., Joseph, N., Weston, S., Menon, S., Odom, D.T. *et al.* (2016) Transcriptional silencing of long noncoding RNA GNG12-AS1 uncouples its transcriptional and product-related functions. *Nat Commun*, **7**, 10406.
27. Liao, Y., Smyth, G.K. and Shi, W. (2013) The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res*, **41**, e108.
28. Liao, Y., Smyth, G.K. and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923-930.
- 25 29. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W. and Smyth, G.K. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*, **43**, e47.
- 30 30. Kong, L., Zhang, Y., Ye, Z.Q., Liu, X.Q., Zhao, S.Q., Wei, L. and Gao, G. (2007) CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res*, **35**, W345-349.
31. Wang, L., Park, H.J., Dasari, S., Wang, S., Kocher, J.P. and Li, W. (2013) CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Res*, **41**, e74.
- 35 32. Lin, M.F., Jungreis, I. and Kellis, M. (2011) PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics*, **27**, i275-282.
33. McCarthy, D.J. and Smyth, G.K. (2009) Testing significance relative to a fold-change threshold is a TREAT. *Bioinformatics*, **25**, 765-771.
34. Gergely, F., Draviam, V.M. and Raff, J.W. (2003) The ch-TOG/XMAP215 protein is essential for spindle pole organization in human somatic cells. *Genes Dev*, **17**, 336-341.
- 40 35. Barr, A.R. and Bakal, C. (2015) A sensitised RNAi screen reveals a ch-TOG genetic interaction network required for spindle assembly. *Sci Rep*, **5**, 10564.
36. Wu, X., Scott, D.A., Kriz, A.J., Chiu, A.C., Hsu, P.D., Dadon, D.B., Cheng, A.W., Trevino, A.E., Konermann, S., Chen, S. *et al.* (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol*, **32**, 670-676.
- 45 37. O'Geen, H., Henry, I.M., Bhakta, M.S., Meckler, J.F. and Segal, D.J. (2015) A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. *Nucleic Acids Res*, **43**, 3389-3404.
38. Dominguez, A.A., Lim, W.A. and Qi, L.S. (2016) Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol*, **17**, 5-15.
- 50 39. Zhao, Y., Li, H., Fang, S., Kang, Y., Wu, W., Hao, Y., Li, Z., Bu, D., Sun, N., Zhang, M.Q. *et al.* (2016) NONCODE 2016: an informative and valuable data source of long non-coding RNAs. *Nucleic Acids Res*, **44**, D203-208.
40. Lagarde, J., Uszczyńska-Ratajczak, B., Carbonell, S., Perez-Lluch, S., Abad, A., Davis, C., Gingeras, T.R., Frankish, A., Harrow, J., Guigo, R. *et al.* (2017) High-throughput annotation of full-length long noncoding RNAs with capture long-read sequencing. *Nat Genet*.
- 55 41. Rinn, J.L. and Chang, H.Y. (2012) Genome regulation by long noncoding RNAs. *Annu Rev Biochem*, **81**, 145-166.
42. Quinn, J.J. and Chang, H.Y. (2016) Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet*, **17**, 47-62.

43. Geisler, S. and Coller, J. (2013) RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat Rev Mol Cell Biol*, **14**, 699-712.
44. Ideue, T., Hino, K., Kitao, S., Yokoi, T. and Hirose, T. (2009) Efficient oligonucleotide-mediated degradation of nuclear noncoding RNAs in mammalian cultured cells. *RNA*, **15**, 1578-1587.
- 5 45. Tripathi, V., Shen, Z., Chakraborty, A., Giri, S., Freier, S.M., Wu, X., Zhang, Y., Gorospe, M., Prasanth, S.G., Lal, A. *et al.* (2013) Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB. *PLoS Genet*, **9**, e1003368.
- 10 46. 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.
47. Golding, M.C., Magri, L.S., Zhang, L., Lalone, S.A., Higgins, M.J. and Mann, M.R. (2011) Depletion of Kcnq1ot1 non-coding RNA does not affect imprinting maintenance in stem cells. *Development*, **138**, 3667-3678.
- 15 48. Gagnon, K.T., Li, L., Chu, Y., Janowski, B.A. and Corey, D.R. (2014) RNAi factors are present and active in human cell nuclei. *Cell Rep*, **6**, 211-221.
49. Ingolia, N.T., Ghaemmaghami, S., Newman, J.R. and Weissman, J.S. (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, **324**, 218-223.
- 20 50. Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Holper, S., Kruger, M. and Stainier, D.Y. (2015) Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature*, **524**, 230-233.
51. Morgens, D.W., Deans, R.M., Li, A. and Bassik, M.C. (2016) Systematic comparison of CRISPR/Cas9 and RNAi screens for essential genes. *Nat Biotechnol*, **34**, 634-636.
- 25 52. Spiliotis, E.T., Kinoshita, M. and Nelson, W.J. (2005) A mitotic septin scaffold required for Mammalian chromosome congression and segregation. *Science*, **307**, 1781-1785.
53. Wei, J.H., Zhang, Z.C., Wynn, R.M. and Seemann, J. (2015) GM130 Regulates Golgi-Derived Spindle Assembly by Activating TPX2 and Capturing Microtubules. *Cell*, **162**, 287-299.
- 30 54. Gabory, A., Jammes, H. and Dandolo, L. (2010) The H19 locus: role of an imprinted non-coding RNA in growth and development. *Bioessays*, **32**, 473-480.
55. Monnier, P., Martinet, C., Pontis, J., Stancheva, I., Ait-Si-Ali, S. and Dandolo, L. (2013) H19 lncRNA controls gene expression of the Imprinted Gene Network by recruiting MBD1. *Proc Natl Acad Sci U S A*, **110**, 20693-20698.
- 35 56. Kallen, A.N., Zhou, X.B., Xu, J., Qiao, C., Ma, J., Yan, L., Lu, L., Liu, C., Yi, J.S., Zhang, H. *et al.* (2013) The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Mol Cell*, **52**, 101-112.
57. Imig, J., Brunschweiler, A., Brummer, A., Guennewig, B., Mittal, N., Kishore, S., Tsikrika, P., Gerber, A.P., Zavolan, M. and Hall, J. (2015) miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. *Nat Chem Biol*, **11**, 107-114.
- 40 58. Yan, L., Zhou, J., Gao, Y., Ghazal, S., Lu, L., Bellone, S., Yang, Y., Liu, N., Zhao, X., Santin, A.D. *et al.* (2015) Regulation of tumor cell migration and invasion by the H19/let-7 axis is antagonized by metformin-induced DNA methylation. *Oncogene*, **34**, 3076-3084.
59. Gabory, A., Ripoché, M.A., Le Digarcher, A., Watrin, F., Ziyat, A., Forne, T., Jammes, H., Ainscough, J.F., Surani, M.A., Journot, L. *et al.* (2009) H19 acts as a trans regulator of the imprinted gene network controlling growth in mice. *Development*, **136**, 3413-3421.
- 45 60. Smith, I., Greenside, P.G., Natoli, T., Lahr, D.L., Wadden, D., Tirosh, I., Narayan, R., Root, D.E., Golub, T.R., Subramanian, A. *et al.* (2017) Evaluation of RNAi and CRISPR technologies by large-scale gene expression profiling in the Connectivity Map. *PLoS Biol*, **15**, e2003213.
- 50 61. Dimitrova, N., Zamudio, J.R., Jong, R.M., Soukup, D., Resnick, R., Sarma, K., Ward, A.J., Raj, A., Lee, J.T., Sharp, P.A. *et al.* (2014) LincRNA-p21 activates p21 in cis to promote Polycomb target gene expression and to enforce the G1/S checkpoint. *Mol Cell*, **54**, 777-790.
62. Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-Broz, D., Khalil, A.M., Zuk, O., Amit, I., Rabani, M. *et al.* (2010) A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*, **142**, 409-419.
- 55 63. Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H. and Bartel, D.P. (2011) Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell*, **147**, 1537-1550.
64. Kok, F.O., Shin, M., Ni, C.W., Gupta, A., Grosse, A.S., van Impel, A., Kirchmaier, B.C., Peterson-Maduro, J., Kourkoulis, G., Male, I. *et al.* (2015) Reverse genetic screening reveals

- poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Dev Cell*, **32**, 97-108.
65. Grote, P., Wittler, L., Hendrix, D., Koch, F., Wahrisch, S., Beisaw, A., Macura, K., Blass, G., Kellis, M., Werber, M. *et al.* (2013) The tissue-specific lincRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev Cell*, **24**, 206-214.
- 5 66. Sauvageau, M., Goff, L.A., Lodato, S., Bonev, B., Groff, A.F., Gerhardinger, C., Sanchez-Gomez, D.B., Hacisuleyman, E., Li, E., Spence, M. *et al.* (2013) Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *Elife*, **2**, e01749.
- 10 67. Bassik, M.C., Kampmann, M., Lebbink, R.J., Wang, S., Hein, M.Y., Poser, I., Weibezahn, J., Horlbeck, M.A., Chen, S., Mann, M. *et al.* (2013) A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. *Cell*, **152**, 909-922.
68. Goyal, A., Fiskin, E., Gutschner, T., Polycarpou-Schwarz, M., Gross, M., Neugebauer, J., Gandhi, M., Caudron-Herger, M., Benes, V. and Diederichs, S. (2017) A cautionary tale of sense-antisense gene pairs: independent regulation despite inverse correlation of expression. *Nucleic Acids Res.*
- 15 69. Evers, B., Jastrzebski, K., Heijmans, J.P., Grenrum, W., Beijersbergen, R.L. and Bernards, R. (2016) CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat Biotechnol*, **34**, 631-633.
70. Kornienko, A.E., Guenzl, P.M., Barlow, D.P. and Pauler, F.M. (2013) Gene regulation by the act of long non-coding RNA transcription. *BMC Biol*, **11**, 59.
- 20 71. Orom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q. *et al.* (2010) Long noncoding RNAs with enhancer-like function in human cells. *Cell*, **143**, 46-58.
72. Engreitz, J.M., Haines, J.E., Perez, E.M., Munson, G., Chen, J., Kane, M., McDonel, P.E., Guttman, M. and Lander, E.S. (2016) Local regulation of gene expression by lincRNA promoters, transcription and splicing. *Nature*.
- 25 73. Yin, Y., Yan, P., Lu, J., Song, G., Zhu, Y., Li, Z., Zhao, Y., Shen, B., Huang, X., Zhu, H. *et al.* (2015) Opposing Roles for the lincRNA Haunt and Its Genomic Locus in Regulating HOXA Gene Activation during Embryonic Stem Cell Differentiation. *Cell Stem Cell*, **16**, 504-516.
- 30 74. Paralkar, V.R., Tabora, C.C., Huang, P., Yao, Y., Kossenkov, A.V., Prasad, R., Luan, J., Davies, J.O., Hughes, J.R., Hardison, R.C. *et al.* (2016) Unlinking an lincRNA from Its Associated cis Element. *Mol Cell*, **62**, 104-110.
75. Fulco, C.P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S.R., Perez, E.M., Kane, M., Cleary, B., Lander, E.S. and Engreitz, J.M. (2016) Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science*.
- 35 76. Rosenbluh, J., Xu, H., Harrington, W., Gill, S., Wang, X., Vazquez, F., Root, D.E., Tsherniak, A. and Hahn, W.C. (2017) Complementary information derived from CRISPR Cas9 mediated gene deletion and suppression. *Nat Commun*, **8**, 15403.
77. Hannus, M., Beitzinger, M., Engelmann, J.C., Weickert, M.T., Spang, R., Hannus, S. and Meister, G. (2014) siPools: highly complex but accurately defined siRNA pools eliminate off-target effects. *Nucleic Acids Res*, **42**, 8049-8061.
- 40 78. Buehler, E., Chen, Y.C. and Martin, S. (2012) C911: A bench-level control for sequence specific siRNA off-target effects. *PLoS One*, **7**, e51942.
79. Abudayyeh, O.O., Gootenberg, J.S., Essletzbichler, P., Han, S., Joung, J., Belanto, J.J., Verdine, V., Cox, D.B.T., Kellner, M.J., Regev, A. *et al.* (2017) RNA targeting with CRISPR-Cas13. *Nature*, **550**, 280-284.
- 45 80. Consortium, F., the, R.P., Clst, Forrest, A.R., Kawaji, H., Rehli, M., Baillie, J.K., de Hoon, M.J., Haberle, V., Lassmann, T. *et al.* (2014) A promoter-level mammalian expression atlas. *Nature*, **507**, 462-470.
- 50 81. Derti, A., Garrett-Engle, P., Macisaac, K.D., Stevens, R.C., Sriram, S., Chen, R., Rohl, C.A., Johnson, J.M. and Babak, T. (2012) A quantitative atlas of polyadenylation in five mammals. *Genome Res*, **22**, 1173-1183.

FIGURE LEGENDS

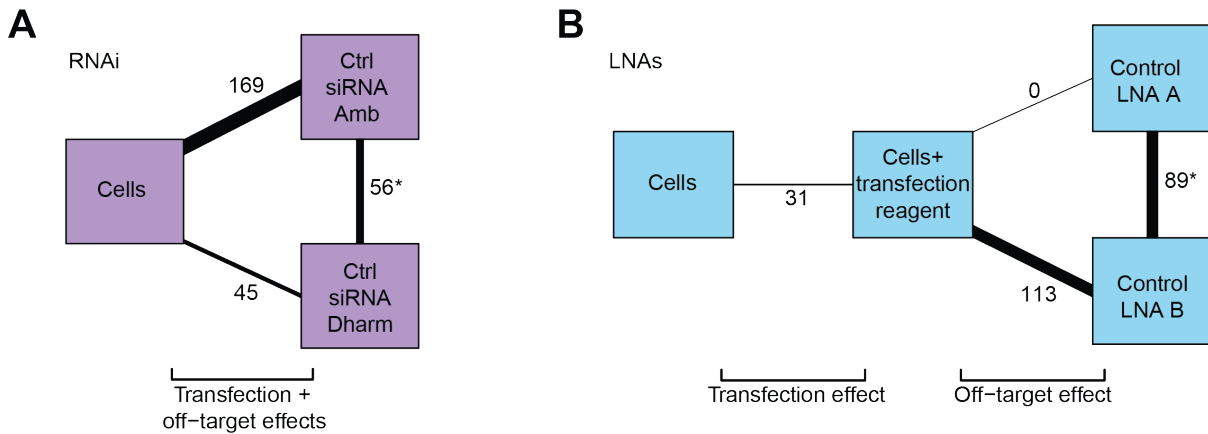


Figure 1. Off-target effects associated with RNAi and LNA oligonucleotides

- 5 **A.** Comparison of the transcriptional differences between untreated cells and cells treated with two negative control siRNAs (Ambion and GE Dharmacon). The number of DEGs between each pair of treatments is labelled and shown as connecting lines of proportional thickness.
- 10 **B.** Comparison of the transcriptional differences between untreated cells, cells treated with transfection reagent (RNAiMax) and two negative control LNAs (A and B). The number of DEGs differing between each pair of treatments is labelled as described in **A**.

DEGs for each pairwise comparison were defined at an FDR of 5% after for a \log_2 -fold change significantly greater than 0.5. Lists of DEGs for each comparison are shown in Supplementary Table 1.

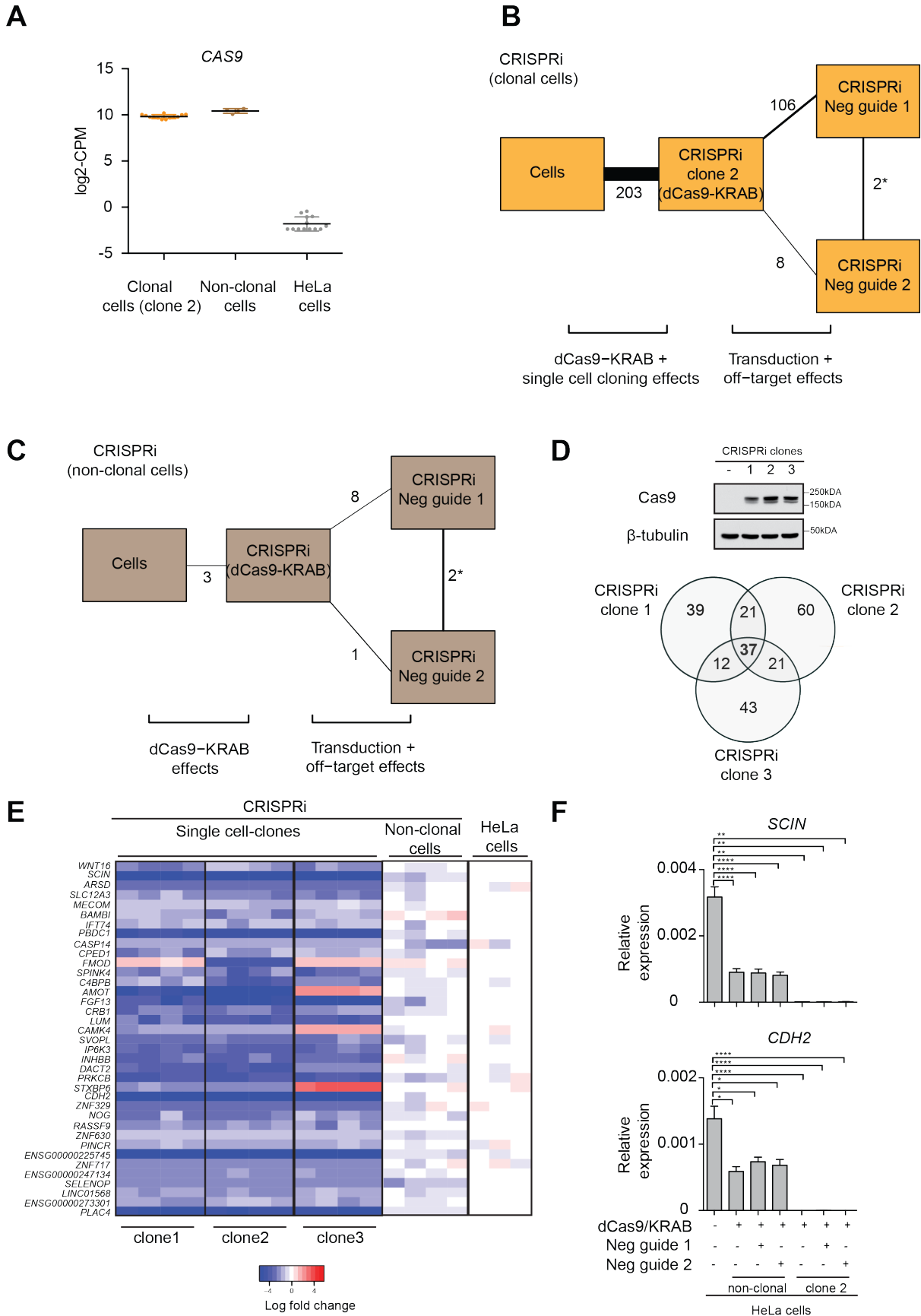


Figure 2. Clonal variations in CRISPRi and their associated off-target transcriptional effects

- A.** Expression in counts-per-million (CPM) of *Cas9* in CRISPRi clonal, CRISPRi non-clonal and untransduced HeLa cells. Clone 2 was used for showing *Cas9* expression in CRISPRi clonal cells.
- B.** Comparison of the transcriptional differences between parental HeLa untreated cells, CRISPRi clones expressing only dCas9-KRAB (clone 2) and clones treated with two negative guide RNAs (negative guide 1 and 2). The number of genes differing between each pair of treatments is labelled and shown as connecting lines of proportional thickness.
- C.** Comparison of the transcriptional differences between parental HeLa untreated cells, non-clonal CRISPRi cells expressing dCas9-KRAB and non-clonal cells treated with two negative guide RNAs (negative guide 1 and 2). The number of genes differing between each pair of treatments is labelled as described in **B**.
- D.** Expression of dCas9-KRAB in three different CRISPRi clones derived from single cell cloning, confirmed by immunoblot using a *Cas9* antibody. β -tubulin was used as a loading control. A Venn diagram of DEGs detected in the three different clones against untransduced cells in the absence of any guide RNAs identified 37 genes as a common transcriptional signature of cloning. The total number of genes in this analysis was 17991 and DEGs were detected at a FDR of 5%.
- E.** Heat map of DEGs from three different CRISPRi clones compared to non-clonal cells and parental untransduced HeLa cells, in the absence of any guide RNAs. 33 out of 37 genes were downregulated in clonal cells compared to the parental population.
- F.** Downregulation of two randomly selected DEGs from **E** (*SCIN* and *CDH2*) was validated by qPCR in clonal cells (clone 2) and in non-clonal populations. Expression levels were normalized to the geometric mean of *GAPDH* and *RPS18*. Error bars, s.e.m. ($n=4$ biological replicates). Statistical significance by two-tailed Student's *t*-test: * $P<0.05$, ** $P<0.01$, *** $P<0.001$ and **** $P<0.0001$.

Lists of DEGs for each pairwise comparison in **B** and **C** are provided in Supplementary Table 1.

25

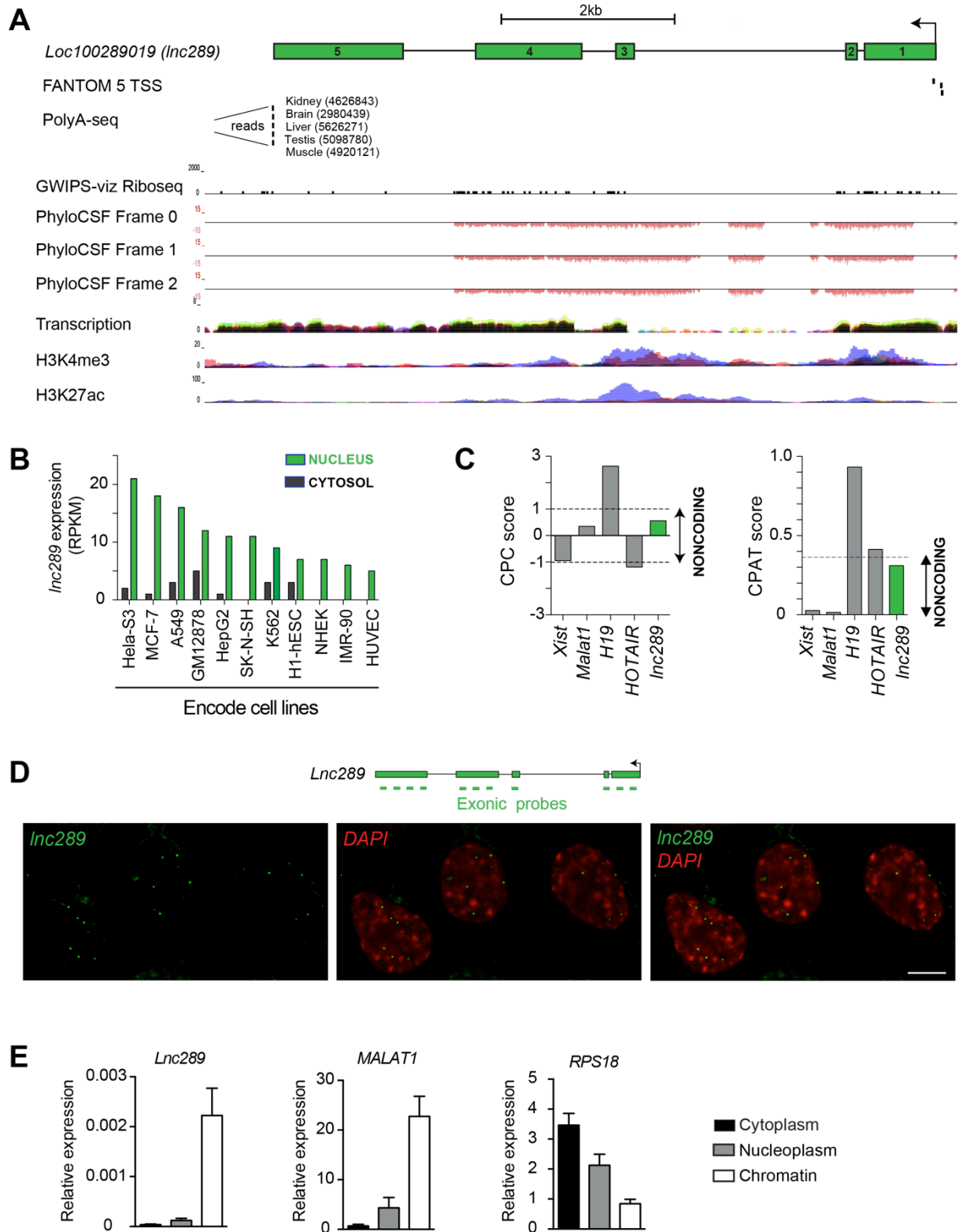


Figure 3. *Inc289* is an archetypical lncRNA expressed in the nucleus

A. Schematic representation of the genomic landscape surrounding *Inc289* (annotated in RefSeq as *loc100289019* or *SLC25A25-AS1*; chr9:128108581-128118693, hg38), including three transcriptional start sites (80) and a polyadenylation site (81). *Inc289* is not occupied by ribosomes (49), shows no protein coding potential (PhyloCSF, (32)), and has clear hallmarks of active transcription in HeLa cells (H3K4me3 and H3K27ac data sets obtained from ENCODE via

the UCSC browser). The arrows denote the direction of transcription, and green boxes represent the five exons. Note that all PhyloCSF scores at this locus are negative.

- B. Expression of *Inc289* in cytosol and nuclei of ENCODE cell lines (www.ebi.ac.uk/gxa/home), shown as reads per kilobase of exon per million reads mapped (RPKM).
- 5 C. Computational analysis of the mature *Inc289* transcript using the CPC and CPAT tools reveals *Inc289* has low coding potential.
- D. Nuclear localization of *Inc289* in HeLa cells was determined using single-molecule RNA FISH with exonic probes (green). The nucleus was stained with 4,6-diamidino-2-phenylindole (DAPI). Scale bar represents 5 μ m.
- 10 E. *Inc289* is enriched in chromatin of HeLa cells. RNA distribution in the cytoplasm, nucleoplasm and chromatin was quantified by qPCR, and *RPS18* and *MALAT1* were used as positive controls for the cytoplasmic and chromatin fraction, respectively. Error bars represent the standard error of the mean (s.e.m) values of four independent experiments.

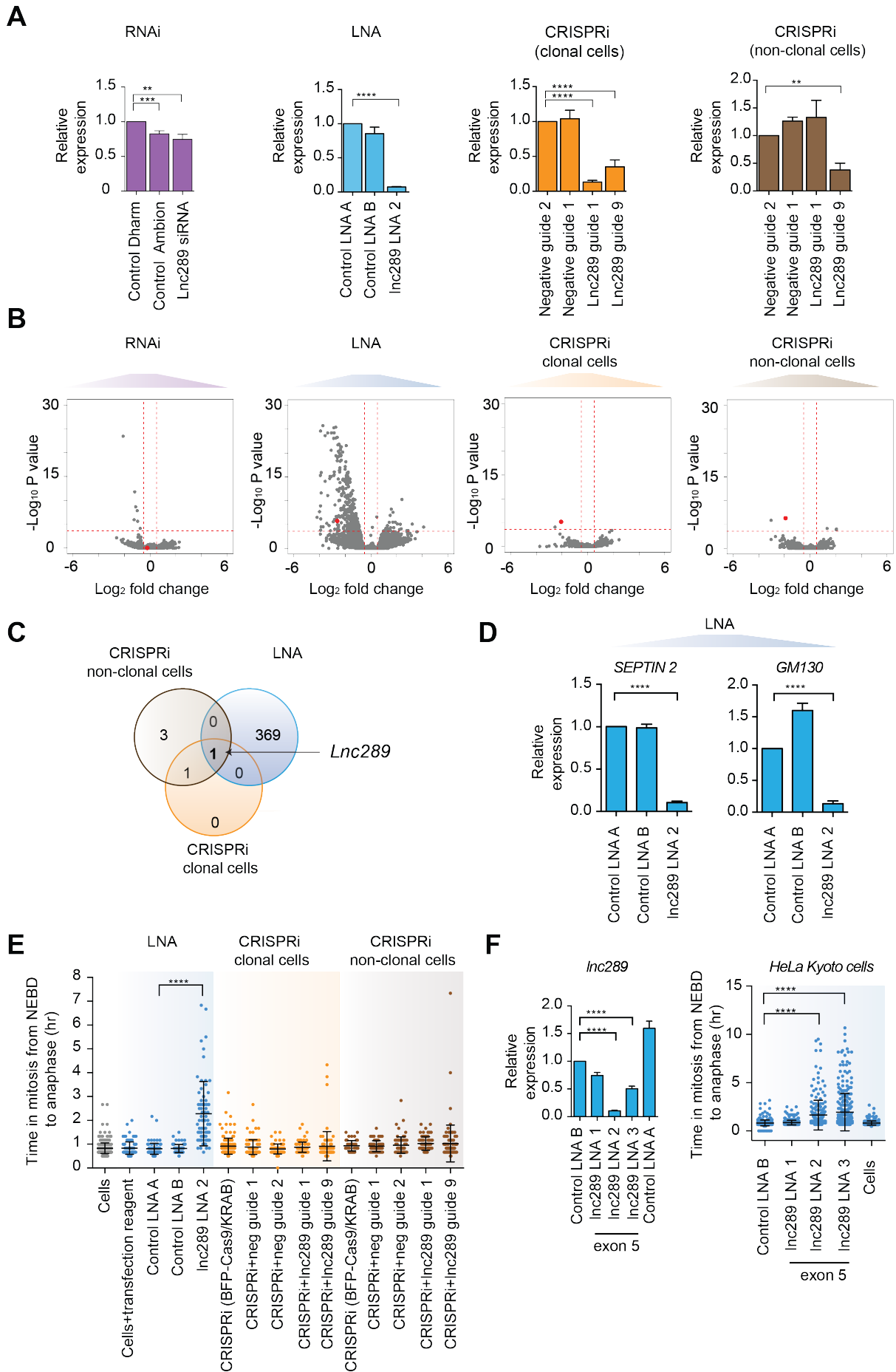


Figure 4. No overlap in DEGs between the different LOF methods upon depletion of nuclear *Inc289*

- A.** Expression levels of *Inc289* after RNAi, LNA and CRISPRi-mediated depletion. qPCR revealed only a 25% reduction in *Inc289* transcription after siRNA-mediated knockdown relative to negative control siRNA from Dharmacon (Control Dharm), and no significant difference relative to negative control siRNA from Ambion (Control Ambion). LNA-mediated knockdown of *Inc289* was performed using LNA oligonucleotide sequence 2 (LNA 2), and showed 90% reduction. CRISPRi-mediated repression of *Inc289* using two guide RNAs targeting TSS of *Inc289* relative to the negative (non-targeting) guide RNA 2 yielded 70-90% knockdown in clonal cells. Only one guide RNA (guide 9) was efficient in depleting *Inc289* in non-clonal cells. Statistical significance by two-tailed Student's *t*-test: ** $P < 0.01$, *** $P < 0.001$ and **** $P < 0.0001$. For all graphs, expression levels of *Inc289* were measured by qPCR using primers spanning exons 1–4, and normalized to the geometric mean of *GAPDH* and *RPS18*. Error bars, s.e.m. ($n=4$ biological replicates).
- B.** Volcano plots of transcriptional differences induced by RNAi, LNA and CRISPRi-mediated depletion of *Inc289*. After siRNA-mediated depletion of *Inc289*, seven genes were differentially expressed compared to the negative control siRNAs. LNA-mediated depletion with LNA 2 identified 370 DEGs compared to negative control oligonucleotides. CRISPRi-mediated depletion of *Inc289* using guide RNA 1 and 9 revealed only two DEGs compared to negative guide RNA 2. In non-clonal cells, only four DEGs were identified using guide RNA 9 compared to the negative guides. The red horizontal line represents the significance threshold corresponding to an FDR of 5%. Red vertical lines are \log_2 -fold change thresholds of ± 0.5 . The red dot corresponds to the *Inc289* itself.
- C.** Venn diagram showing no overlap between the sets of DEGs identified after using LNA and CRISPRi to deplete *Inc289*. The only gene in common between LNA and CRISPRi-mediated depletion is *Inc289* itself. The total number of genes used for this analysis was 17837.
- D.** qPCR confirmation of the downregulation of two DEGs (*SEPTIN2* and *GM130*) identified in **B** after LNA-mediated depletion of *Inc289* with LNA 2. Expression levels were normalized to the geometric mean of *GAPDH* and *RPS18*. Error bars, s.e.m. ($n=3$ biological replicates). Statistical significance by two-tailed Student's *t*-test: **** $P < 0.0001$.
- E.** Quantification results from time-lapse microscopy of mitotic progression of HeLa cells incubated with Sir-Hoechst after LNA and CRISPRi-mediated (clonal and non-clonal) depletion of *Inc289*. Mitotic duration was measured from nuclear envelope breakdown (NEBD) to anaphase onset. Bars show mean \pm s.d. ($n=2$ independent biological replicates). Statistical significance by Mann-Whitney test: **** $P < 0.0001$.
- F.** *Left panel:* Expression analysis of *Inc289* by qPCR using two additional LNAs targeting *Inc289* (LNA 1 and LNA 3). Expression levels were normalized to the geometric mean of *GAPDH* and *RPS18*. Error bars, s.e.m. ($n=3$ biological replicates). Statistical significance by two-tailed Student's *t*-test: **** $P < 0.0001$. *Right panel:* Quantification results from time-lapse microscopy using HeLa Kyoto cells after depletion of *Inc289* using three different LNAs. Bars show mean \pm s.d. ($n=2$ biological replicates). Statistical significance by Mann-Whitney test: **** $P < 0.0001$.

The list of DEGs identified after RNAi, LNA and CRISPRi-mediated depletion of *Inc289* are shown in Supplementary Table 2.