RESEARCH ARTICLE SUMMARY

NONCODING RNA

CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells

S. John Liu,* Max A. Horlbeck,* Seung Woo Cho, Harjus S. Birk, Martina Malatesta, Daniel He, Frank J. Attenello, Jacqueline E. Villalta, Min Y. Cho, Yuwen Chen, Mohammad A. Mandegar, Michael P. Olvera, Luke A. Gilbert, Bruce R. Conklin, Howard Y. Chang, Jonathan S. Weissman,†‡ Daniel A. Lim†‡

INTRODUCTION: The human genome contains tens of thousands of loci that produce long noncoding RNAs (lncRNAs), transcripts that have no apparent protein-coding potential. A subset of lncRNAs have been found to play critical roles in cellular processes, organismal development, and disease. Although these examples are suggestive of the importance and diversity of lncRNAs, the vast majority of lncRNA genes have not been functionally tested.

RATIONALE: Because it is currently not possible to predict which lncRNA loci are functional or what function they perform, there is a need for large-scale, systematic approaches to interrogating the functional contribution of lncRNA loci. We therefore developed a genome-scale screening platform based on CRISPR-mediated interference (CRISPRi), which uses a catalytically inactive CRISPR effector protein, (d)Cas9, fused to a repressive KRAB domain and targeted by a single guide RNA (sgRNA), to inhibit gene expression. By catalyzing repressive chromatin modifications around the transcription start site (TSS) and serving as a transcriptional roadblock, CRISPRi tests a broad range of lncRNA gene functions, including the production of cis- and trans-acting RNA transcripts, cis-mediated regulation related to lncRNA transcription itself, and enhancer-like function of some lncRNA loci.

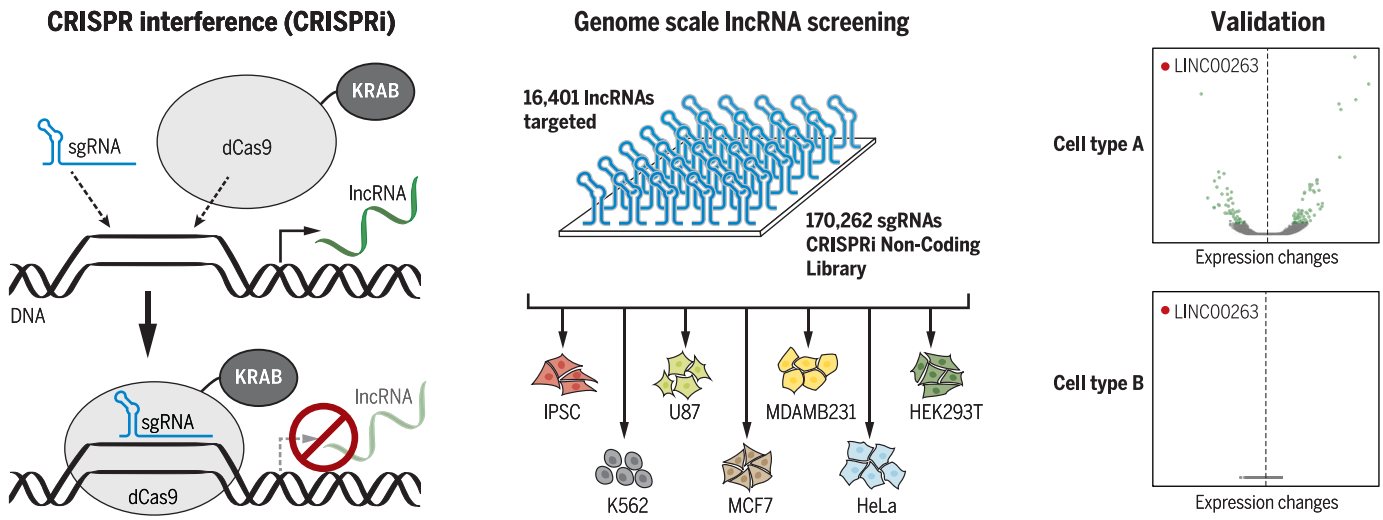
RESULTS: We designed a CRISPRi Non-Coding Library (CRiNCL), which targets 16,401 lncRNA genes each with 10 sgRNAs per TSS, and applied this pooled screening approach to identify lncRNA genes that modify robust cell growth. We screened seven human cell lines, including six transformed cell lines and induced pluripotent stem cells (iPSCs), and identified 499 lncRNA loci that modified cell growth upon CRISPRi targeting; 372 and 299 of these loci were distal from any protein coding gene or mapped enhancer, respectively. Extensive validation confirmed the screen results and demonstrated the robust and specific performance of CRISPRi for repressing lncRNA transcription. Remarkably, 89% of the lncRNA gene hits modified growth in just one of the cell lines tested, and no hits were common to all seven cell lines. Although nearly all of the hit genes were expressed in the cell line in which they exhibited a growth phenotype, expression alone was insufficient to explain the cell type specificity of their function. Transcriptional profiling revealed extensive gene expression changes upon CRISPRi targeting of lncRNA loci in the cells in which they modified growth, whereas targeting the same lncRNA locus in other cell lines resulted in minimal changes to the transcriptome beyond depletion of the targeted lncRNA transcript itself.

CONCLUSION: Our study considerably increases the number of known functional lncRNA loci. More broadly, our CRISPRi approach enables mechanistic studies of specific lncRNA functions and, when applied systematically, supports the global exploration of the complex biology contained in the lncRNA-expressing genome. Finally, in contrast to recent studies that found that essential protein-coding genes typically are required across a broad range of cell types, we show that lncRNA function is highly cell type–specific, a finding that has important implications for their involvement in both normal biology and disease. ■

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aah7111

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†These authors contributed equally to this work.
‡Corresponding author. Email: daniel.lim@ucsf.edu (D.A.L.); jonathan.weissman@ucsf.edu (J.S.W.)
Cite this article as S. J. Liu *et al.*, *Science* **355**, aah7111 (2017).
DOI: 10.1126/science.aah7111



CRISPRi screening of lncRNAs in human cells. CRISPRi can precisely repress transcription of lncRNAs. The CRISPRi Non-Coding Library (CRiNCL) was generated to interrogate the function of thousands of long noncoding RNAs in seven different cell lines. Validation studies confirmed the exquisite cell type–specific function of lncRNAs.

Downloaded from http://science.sciencemag.org/ on January 13, 2020

RESEARCH ARTICLE

NONCODING RNA

CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells

S. John Liu,[1,2]* Max A. Horlbeck,[3,4,5,6]* Seung Woo Cho,[9] Harjus S. Birk,[1,2] Martina Malatesta,[1,2] Daniel He,[1,2] Frank J. Attenello,[1,2] Jacqueline E. Villalta,[3,4,5,6] Min Y. Cho,[3,4,5,6] Yuwen Chen,[3,4,5,6] Mohammad A. Mandegar,[3] Michael P. Olvera,[3] Luke A. Gilbert,[3,4,5,6] Bruce R. Conklin,[3,7,8] Howard Y. Chang,[9] Jonathan S. Weissman,[3,4,5,6]†‡ Daniel A. Lim[1,2,10]†‡

The human genome produces thousands of long noncoding RNAs (lncRNAs)—transcripts >200 nucleotides long that do not encode proteins. Although critical roles in normal biology and disease have been revealed for a subset of lncRNAs, the function of the vast majority remains untested. We developed a CRISPR interference (CRISPRi) platform targeting 16,401 lncRNA loci in seven diverse cell lines, including six transformed cell lines and human induced pluripotent stem cells (iPSCs). Large-scale screening identified 499 lncRNA loci required for robust cellular growth, of which 89% showed growth-modifying function exclusively in one cell type. We further found that lncRNA knockdown can perturb complex transcriptional networks in a cell type–specific manner. These data underscore the functional importance and cell type specificity of many lncRNAs.

Sequencing efforts have revealed that the human genome produces tens of thousands of long noncoding RNAs (lncRNAs), transcripts greater than 200 nucleotides in length that are often spliced and polyadenylated but have no apparent protein-coding potential (*1*–*3*). Certain lncRNAs play critical roles in cellular function, development, and disease (*4*, *5*). However, of the very large set of lncRNAs—many of which are differentially expressed in tissues and disease states—only a very small fraction have established biological functions, and even fewer are known to function in fundamental aspects of cell biology such as cell proliferation. Currently, it is not possible to predict which lncRNAs are functional, let alone what function they perform. Thus, a large-scale, systematic approach to evaluating the function of the vast population of lncRNAs is critical to understanding the roles that these noncoding transcripts play in cell biology.

A central limitation to systematic efforts to evaluate lncRNA function has been the lack of highly specific, scalable tools for inhibiting lncRNA gene activity (*6*). Gene deletion studies conducted in mice, flies, and human cells have yielded important biological insights about lncRNAs, but this approach is difficult to scale up (*7*–*10*). CRISPR/Cas9 nuclease approaches based on introduction of indels are both scalable and useful for targeted loss-of-function studies of protein-coding genes by altering the coding frame, but they are not well suited for the study of lncRNA gene function, as small deletions do not generally disrupt their biological activity (*11*–*13*). Nonetheless, larger Cas9-mediated genetic deletions can be effective at eliminating lncRNA genes (*6*, *14*–*17*). Screens based on RNA interference (RNAi) have been valuable (*18*, *19*) despite challenges with off-target effects (*20*). However, many lncRNAs localize to the nucleus, where RNAi exhibits variable knockdown efficiency (*21*).

We previously developed CRISPRi, a technology that can repress transcription of any gene via the targeted recruitment of the nuclease-dead dCas9-KRAB repressor fusion protein to the transcription start site (TSS) by a single guide RNA (sgRNA) (*22*–*24*). Because CRISPRi acts only within a small window (1 kb) around the targeted TSS (*23*), and because dCas9 occludes only 23 base pairs of the targeted DNA strand (*25*), CRISPRi allows for precise perturbation of any lncRNA gene. By catalyzing repressive chromatin modifications around the TSS and serving as a transcriptional roadblock, CRISPRi tests a broad range of lncRNA gene functions including the production of cis- and trans-acting RNA transcripts (*4*), cis-mediated regulation related to lncRNA transcription itself (*26*–*29*), and enhancer-like function of some lncRNA loci (*14*, *15*, *30*). The repressive chromatin modification H3K9me3 (trimethylation of histone 3 Lys[9]) catalyzed by CRISPRi is highly specific, with little to no off-target effects due to either spurious dCas9 binding or unintended silencing of distal regulatory elements, as measured by chromatin immunoprecipitation sequencing (ChIP-seq) or RNA sequencing (RNA-seq) (*22*, *31*–*34*) (see below). To enhance CRISPRi for large-scale screening, we have improved on the design of CRISPRi sgRNA libraries to optimize on-target activity while further minimizing off-target effects, enabling highly sensitive detection of essential coding genes (*35*).

Here, we developed CRISPRi libraries targeting 16,401 lncRNA loci (with 10 sgRNAs per TSS) and conducted screens for genes that are required for robust growth in seven human cell types—six transformed cell lines and induced pluripotent stem cells (iPSCs) (*36*). These large-scale screens, coupled with extensive validation studies, greatly increased the number of lncRNA genes known to have biological function and revealed lncRNA function to be highly cell type–specific. Our studies thus help to elucidate the biology contained within the lncRNA genome and provide a tool for both large-scale and targeted investigations of lncRNA function.

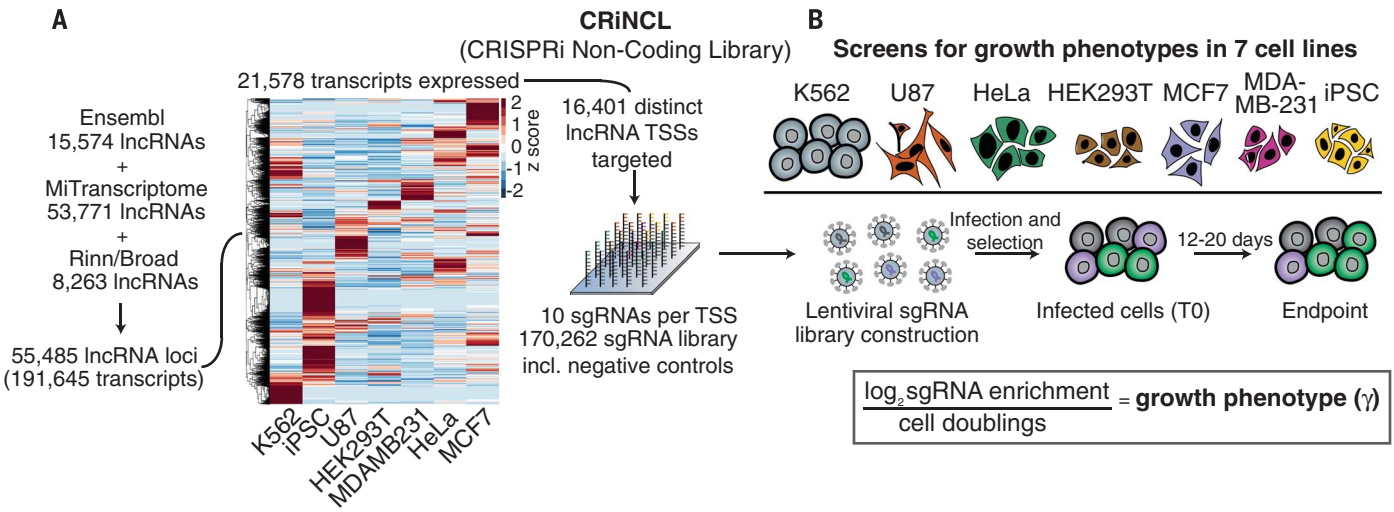
CRISPRi screens identify lncRNA loci that modify cell growth

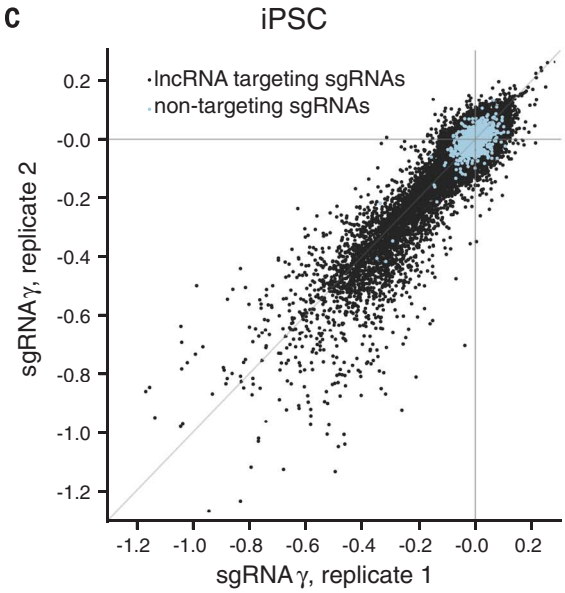
We first designed an sgRNA library to enable genome-scale CRISPRi screening of lncRNA gene function. We generated a comprehensive lncRNA gene set by merging three major noncoding transcriptome annotations (*37*–*39*), prioritized about one-third of these genes based on expression in any of a panel of cancer and nontransformed cell lines (table S1), and used the hCRISPRi-v2.1 algorithm to design 10 sgRNAs targeting each lncRNA TSS (*35*) (Fig. 1A and fig. S1). The cell lines represent a broad range of cell types studied by the ENCODE project (*40*), including a chronic myeloid leukemia cell line (K562), the cervical cancer line HeLa, a glioblastoma line (U87), and two mammary adenocarcinoma lines (MCF7 and MDA-MB-231). We also chose an iPSC line that inducibly expresses CRISPRi components (*33*, *41*). The library, termed CRiNCL (CRISPRi Non-Coding Library), is available as pooled lentiviral plasmid libraries on Addgene and in silico as table S2.

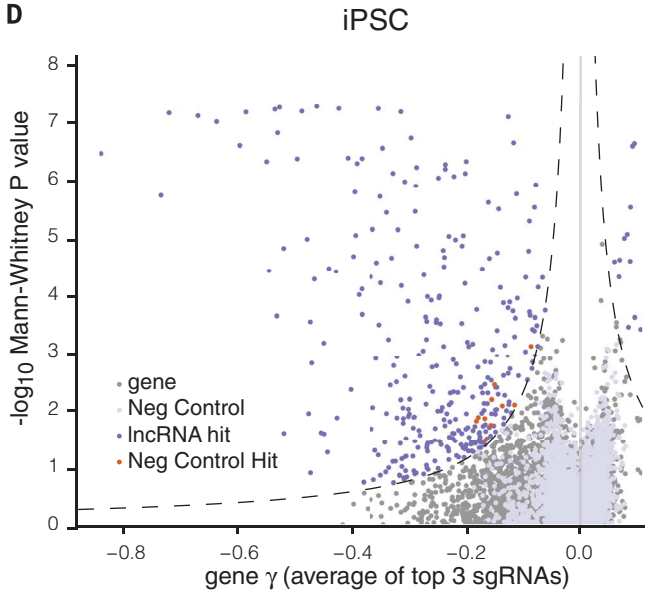
We used this library to conduct screens for lncRNA loci that increase or decrease cell growth in each of seven cell lines. We infected the full lentiviral library or targeted sublibraries (fig. S2A) into each cell line engineered to express dCas9-KRAB (*22*, *23*, *33*, *42*), selected for infected cells by puromycin selection, and cultured for 12 to 20 days, measuring sgRNA enrichment by Illumina sequencing (Fig. 1B and table S3). The fraction of cells infected with the sgRNA library remained

[1]Department of Neurological Surgery, University of California, San Francisco, CA 94143, USA. [2]Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California, San Francisco, CA 94143, USA. [3]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94143, USA. [4]Howard Hughes Medical Institute, University of California, San Francisco, CA 94143, USA. [5]California Institute for Quantitative Biomedical Research, University of California, San Francisco, CA 94143, USA. [6]Center for RNA Systems Biology, University of California, San Francisco, CA 94143, USA. [7]Deparment of Medicine, University of California, San Francisco, CA 94143, USA. [8]Gladstone Institute of Cardiovascular Disease, San Francisco, CA, USA. [9]Center for Personal Dynamic Regulomes, Stanford University, Stanford, CA, USA. [10]San Francisco Veterans Affairs Medical Center, San Francisco, CA, USA.
*These authors contributed equally to this work. †These authors contributed equally to this work. **‡Corresponding author. Email: daniel.lim@ucsf.edu (D.A.L.); jonathan.weissman@ucsf.edu (J.S.W.)**

Downloaded from http://science.sciencemag.org/ on January 13, 2020







E

	Genes targeted in screen	Total hits	lncRNA genes neighboring essential coding genes	**lncRNA hits**
iPSC	5,534	438	112	**326**
MCF7	5,725	117	60	**57**
U87	5,689	88	23	**65**
K562	16,401	144	79	**65**
MDA-MB-231	5,725	44	14	**30**
HeLa	6,158	52	28	**24**
HEK293T	5,785	28	11	**17**

Fig. 1. CRISPRi screens identify lncRNA genes that modify cell growth. (**A**) Schematic of CRISPRi library design strategy. Three lncRNA annotation sets were merged, prioritized by expression in the indicated cell lines, and targeted by 10 sgRNAs per TSS using the hCRISPRi-v2.1 algorithm. Heat map represents expression as *z*-score of FPKM within each cell line (see fig. S1 for TPM values). (**B**) Schematic of growth screens performed in seven different cell lines, and formula for calculation of the growth phenotype (γ). (**C**) Scatterplot of sgRNA phenotypes from two independent replicates of a CRISPRi screen performed in iPSCs. (**D**) Volcano plot of gene γ and *P* value. Screen replicates were averaged, and sgRNAs targeting the same gene were collapsed into a growth phenotype for each gene by the average of the three top-scoring sgRNAs by absolute value and assigned a *P* value by the Mann-Whitney test of all 10 sgRNAs compared to the nontargeting controls. Negative control genes were randomly generated from the set of nontargeting sgRNAs; dashed lines represent a threshold for calling hits by screen score (see supplementary materials). Neighbor hits are not displayed for clarity (see fig. S3, A and B). (**E**) Summary table of all CRISPRi growth screens performed.

Downloaded from http://science.sciencemag.org/ on January 13, 2020

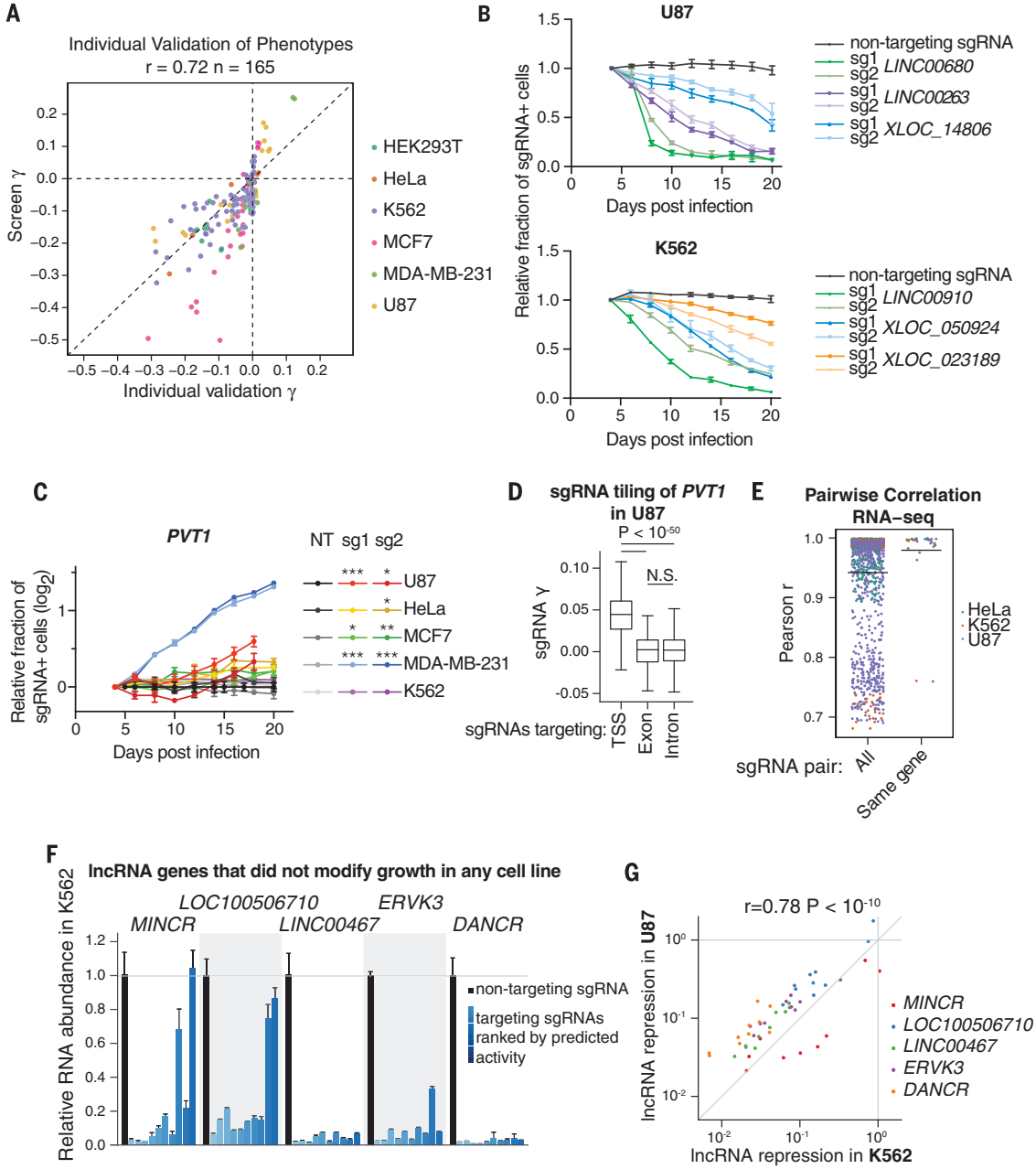


Fig. 2. Validation of screen results shows reproducible phenotypes, correlated transcriptome responses, and robust knockdown of target transcripts. (**A**) Individual sgRNA phenotypes from internally controlled growth assays [(B) and (C)] compared to sgRNA phenotypes from screens. Individual growth phenotypes were calculated from the relative fraction of sgRNA-containing cells at the endpoint divided by the number of doublings from 4 days after infection. Screen growth phenotypes represent the replicate average phenotype from the indicated cell line. (**B**) Internally controlled growth assays performed with sgRNAs targeting lncRNA hit genes in U87 and K562 cells. Cells were infected with lentivirus of the sgRNA expression vector [including a blue fluorescent protein (BFP) marker gene] and passaged for 20 days. The fraction of sgRNA-containing cells was measured as the fraction of high–BFP-expressing cells by flow cytometry and expressed relative to the fraction at 4 days after infection. Points represent the mean and standard deviation of three biological replicates. (**C**) Internally controlled growth assays of *PVT1*-targeting sgRNAs in five cell lines. Assays were performed as in (B). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ [*t* test values compared to the nontargeting (NT) sgRNA at the assay endpoint]. (**D**) Boxplot of sgRNA growth phenotypes from tiling screen of *PVT1* in U87 cells. TSS represents all sgRNAs within 1 kb of the *PVT1* "p1" and "p2" TSSs as annotated by FANTOM; exon represents sgRNAs targeting any *PVT1* exon annotated by Ensembl; intron represents all other sgRNAs (see fig. S7B). sgRNA γs are the average of two replicates. (**E**) Pairwise correlation of gene expression profiles for independent sgRNAs. Expression profiles were measured by RNA-seq and correlations were calculated from transcripts per million (TPM) of genes with significant variation of expression (see supplementary materials). "All" represents every sgRNA pair from the same cell line with the same phenotype direction, except same-sgRNA and same-gene pairs. (**F**) Relative RNA abundance in K562 cells of lncRNA genes that were not hits in any cell line. RNA abundance for all 10 sgRNAs targeting the indicated genes in the CRiNCL library was measured by qPCR. Each bar represents the mean and standard deviation of three biological replicates, and is ordered by decreasing activity as predicted by the hCRISPRi-v2.1 algorithm. (**G**) Correlation of lncRNA repression in K562 and U87 cells. Points represent mean values from (F) and fig. S7C.

Downloaded from http://science.sciencemag.org/ on January 13, 2020

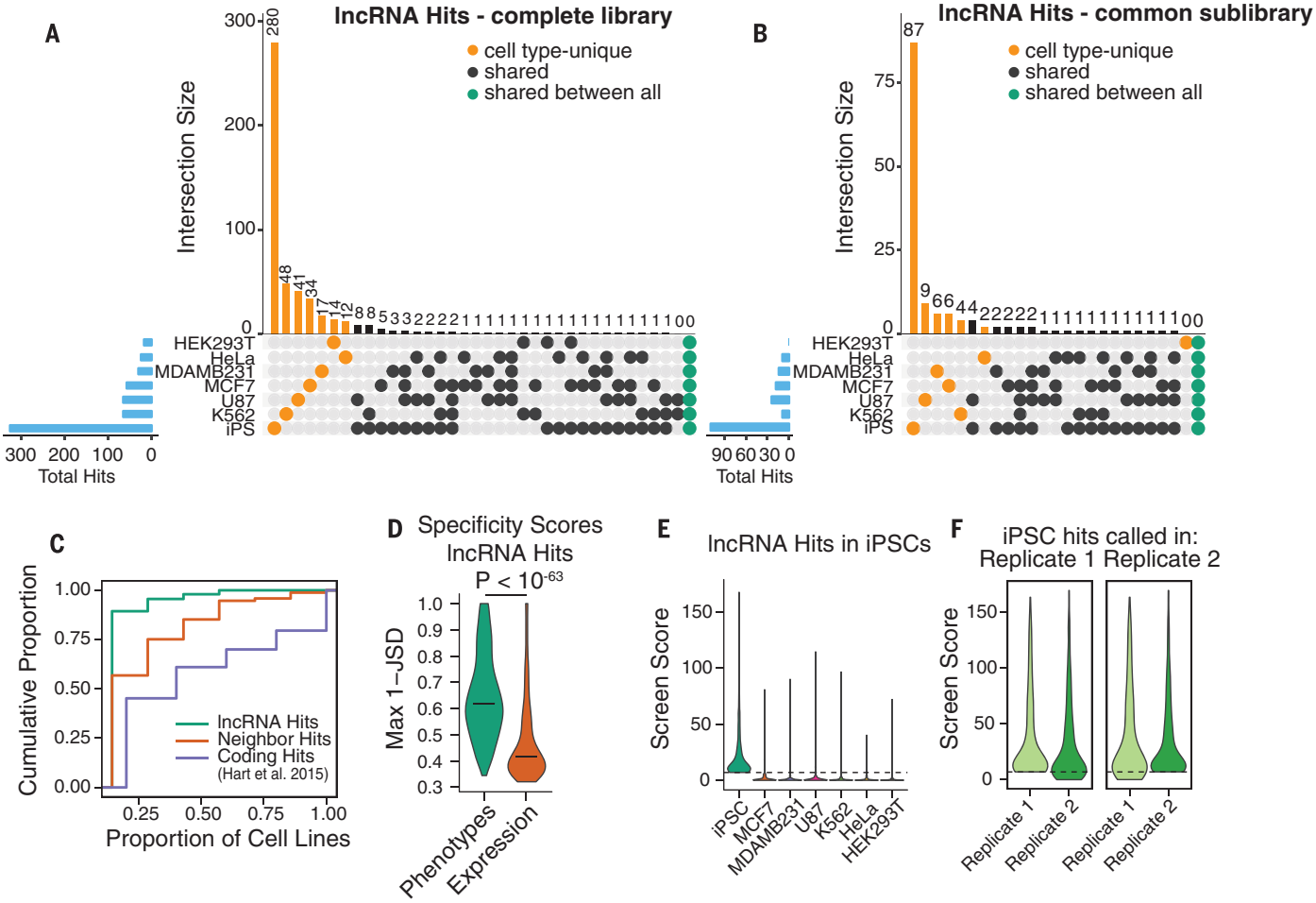


Fig. 3. Growth modifier lncRNA function is highly cell type–specific. (**A**) Numbers of lncRNA hits for each set of cell types in the complete library and (**B**) common sublibrary (lncRNAs that were expressed and screened in all cell types). Blue bars indicate total number of lncRNA hits in each cell type. (**C**) Cumulative distribution function for the proportion of cell types in which each gene is a hit. Protein-coding hits were obtained from Hart *et al.* (*47*) using their 5% FDR Bayes factor threshold. (**D**) Distributions of the maximum 1 – Jenson-Shannon distance (JSD) metric of cell type specificity for lncRNA hit screen scores and expression values. Horizontal lines denote medians. (**E**) Distributions of screen scores across all cell types for lncRNAs that were hits in iPSCs. Dashed line represents screen score threshold for calling hit genes. (**F**) Distributions of screen scores across both replicates of iPS cells, for lncRNAs that would be called as hits in replicate 1 (left) and in replicate 2 (right).

stable over the course of the screen (*23*), indicating that CRISPRi targeting of lncRNA loci does not exhibit nonspecific toxicity (fig. S2B). To facilitate comparisons between screens conducted for different durations and in cell lines with different growth rates, we normalized sgRNA enrichment by total cell doublings to obtain the quantitative growth phenotype γ, which reflects the positive or negative impact on cell growth caused by knockdown of a given gene (*43*) (Fig. 1B).

Analysis of biological replicates revealed that the γ for targeting sgRNAs showed strong and reproducible phenotypes (Pearson r = 0.34 to 0.90), whereas nontargeting control sgRNAs were tightly distributed around 0 (Fig. 1C, fig. S2C, and table S3). We averaged replicate sgRNA phenotypes and used these to score lncRNA genes (*23*, *35*), calculating gene phenotypes from the mean of the top three sgRNAs targeting the gene and Mann-Whitney P values from all 10 sgRNAs compared to nontargeting control sgRNAs (Fig. 1D, fig. S3A, and table S4). Within each screen, we also randomly sampled nontargeting sgRNA phenotypes to generate "negative control genes" and analyzed them as with lncRNA genes (see supplementary materials), enabling us to estimate an empirical false discovery rate (FDR) for each screen as well as the combined screen data set (fig. S2D). We classified lncRNA genes as hits if their combined phenotype effect size and P value (referred to here as "screen score") exceeded a consistent threshold applied to each screen corresponding to an empirical FDR of 5% (fig. S3C). Overall, we found between 28 and 438 lncRNA loci hits in each cell line (Fig. 1E, fig. S3A, and table S4).

We observed that for 169 of these lncRNA hits, the TSS of the noncoding gene was within 1 kb of the TSS of a coding gene previously found to be essential in a CRISPRi screen (*23*), making it difficult to determine whether the observed phenotypes were due to knockdown of the target lncRNA or direct inhibition of the neighboring coding gene (fig. S3B). We thus removed these hits from the total set of hit genes for downstream analyses (Fig. 1E and fig. S3, A and D), resulting in 169 "neighbor hits" and 499 "lncRNA hits," 299 of which are distal from any protein-coding gene (~90% of which would not measurably affect growth upon knockdown). The 1-kb threshold was chosen on the basis of the maximum distance at which CRISPRi is effective, as revealed by analysis of dense sgRNA tiling and genome-scale screens (fig. S4) (*23*); increasing this threshold to 10 kb classified only an additional 19 genes as neighbor hits (fig. S3D).

A larger fraction of lncRNAs hits were observed in the iPSC screen, which suggests either that this cell line is more susceptible to growth perturbations or that iPSCs were differentiating to other cell types with lower growth rates. We therefore investigated iPSC differentiation in a secondary fluorescence-activated cell sorting (FACS)–based screen by assessing loss of pluripotency as indicated by decreased *POU5F1/OCT4* expression. CRISPRi targeting of only nine lncRNA loci

Downloaded from http://science.sciencemag.org/ on January 13, 2020

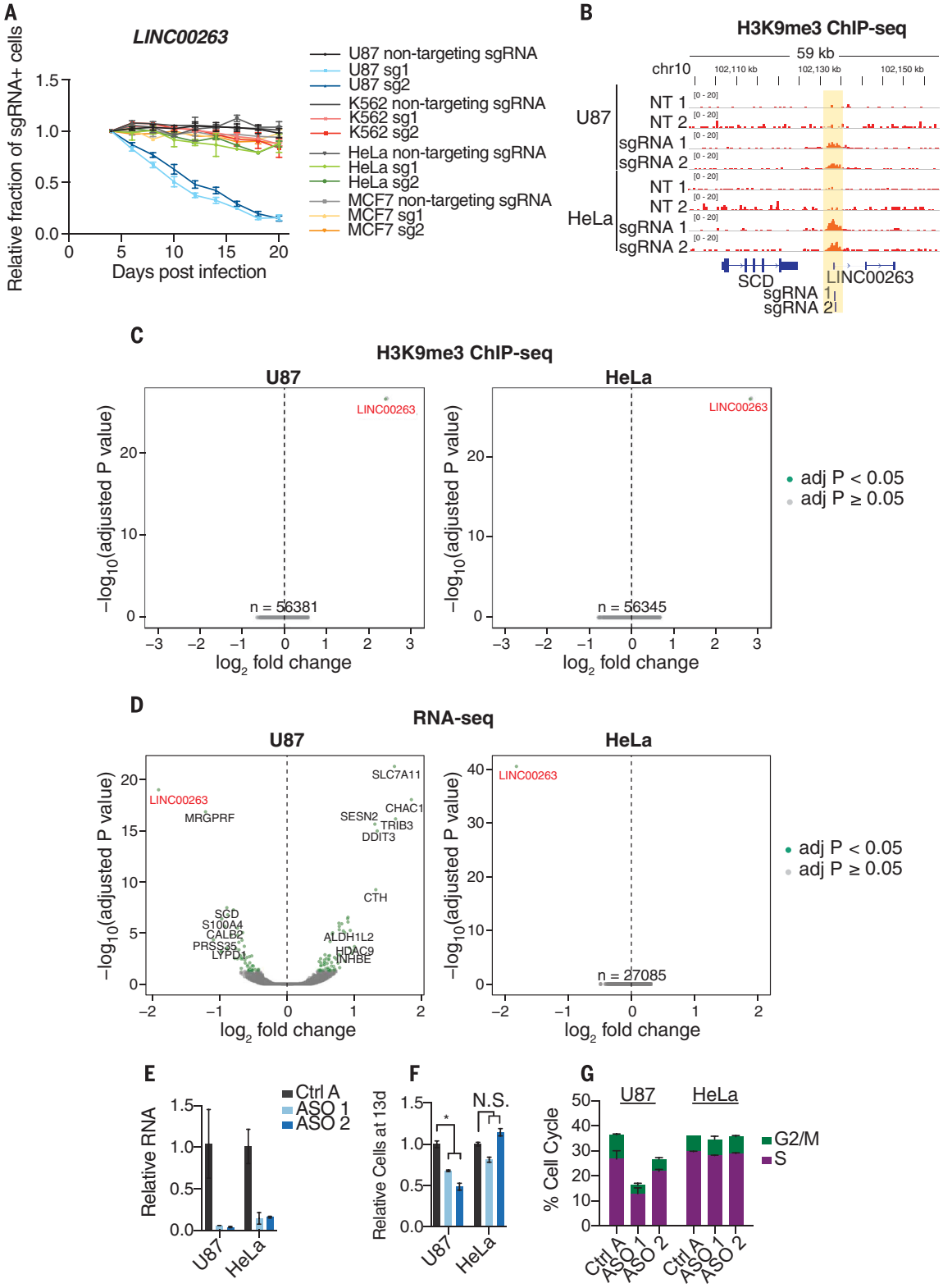


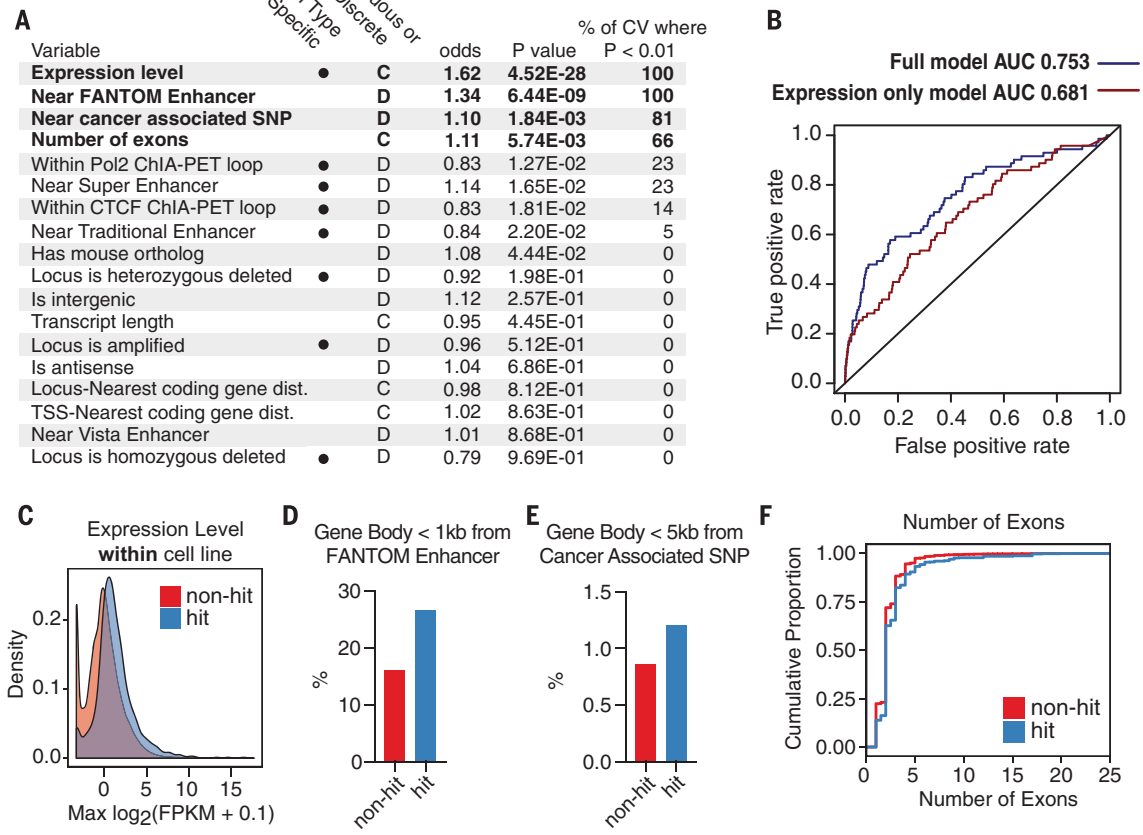
Fig. 4. Dissection of cell type–specific growth modifier lncRNA *LINC00263*. (**A**) Internally controlled growth assays for two independent sgRNAs targeting the TSS of *LINC00263* and nontargeting sgRNA in U87, K562, HeLa, and MCF7 cells. (**B**) ChIP-seq against H3K9me3 in replicates of U87 and HeLa cells infected with nontargeting sgRNAs or *LINC00263* sgRNAs. Values represent normalized reads. (**C**) Volcano plots for ChIP-seq samples in (B), representing genome-wide differential enrichment of H3K9me3 at promoter regions. Relative changes are those of *LINC00263* sgRNAs versus nontargeting sgRNAs. (**D**) Volcano plots for RNA-seq differential expression after infection of *LINC00263* sgRNAs compared to infection of nontargeting sgRNAs. (**E**) qPCR of ASO knockdown of *LINC00263* in U87 and HeLa cells. (**F**) Proportion of cells at 13 days after ASO transfection, relative to control ASO. (**G**) Percentage of cells in S or G_2/M phases after ASO knockdown of *LINC00263*. *P = 0.0029.

Downloaded from http://science.sciencemag.org/ on January 13, 2020

Fig. 5. Machine learning identifies genomic features of growth modifier lncRNAs. (**A**) Results from logistic regression model using 18 classes of genomic data as possible predictors of growth modifier lncRNAs. Cell type dependent variables are marked. Odds ratios represent relative impact of 1 standard deviation increase of given variable. Significant variables ($P < 0.01$) are bolded. Results of 10-fold cross-validation are represented as the percentage of cross-validation iterations where the given variable is significant. (**B**) Receiver operating characteristic (ROC) curves for full model compared to model using only expression data. Model was trained on 75% of screen data, and ROC curves show predictive value on remaining 25%. AUC, area under the curve. (**C**) Density plot of expression levels for lncRNAs that scored as hits and nonhits, aggregated across all cell types. (**D**) Percentage of nonhit (red) and hit (blue) lncRNAs whose gene bodies resided <1 kb from an annotated FANTOM enhancer. (**E**) Percentage of nonhit (red) and hit (blue) lncRNAs whose gene bodies resided <5 kb from a cancer-associated SNP. (**F**) Cumulative distribution function of number of exons for nonhit (red) and hit (blue) lncRNA transcripts.

A

Variable	Cell Type Specific	Continuous or Discrete	odds	P value	% of CV where P < 0.01
Expression level	●	**C**	**1.62**	**4.52E-28**	**100**
Near FANTOM Enhancer		**D**	**1.34**	**6.44E-09**	**100**
Near cancer associated SNP		**D**	**1.10**	**1.84E-03**	**81**
Number of exons		**C**	**1.11**	**5.74E-03**	**66**
Within Pol2 ChIA-PET loop	●	D	0.83	1.27E-02	23
Near Super Enhancer	●	D	1.14	1.65E-02	23
Within CTCF ChIA-PET loop	●	D	0.83	1.81E-02	14
Near Traditional Enhancer	●	D	0.84	2.20E-02	5
Has mouse ortholog		D	1.08	4.44E-02	0
Locus is heterozygous deleted	●	D	0.92	1.98E-01	0
Is intergenic		D	1.12	2.57E-01	0
Transcript length		C	0.95	4.45E-01	0
Locus is amplified	●	D	0.96	5.12E-01	0
Is antisense		D	1.04	6.86E-01	0
Locus-Nearest coding gene dist.		C	0.98	8.12E-01	0
TSS-Nearest coding gene dist.		C	1.02	8.63E-01	0
Near Vista Enhancer		D	1.01	8.68E-01	0
Locus is homozygous deleted	●	D	0.79	9.69E-01	0



reduced *POU5F1/OCT4* expression (fig. S5 and tables S5 and S6), which suggests that the majority of lncRNA hits identified in iPSCs primarily affect cell growth. To confirm that the increased fraction of lncRNA hits in iPSCs was not due to technical differences in CRISPRi function between cell lines, we performed a CRISPRi screen for protein-coding genes required for cell growth in iPSCs (fig. S6A and table S7). These results corresponded well with our previously published K562 growth screen (*35*), both in the number of genes found to have function and in the ability to specifically identify known essential genes (fig. S6, B and C) (*44*). Taken together, our screens identified 499 lncRNA genes that modify cell growth and have no essential coding gene neighbors, representing a large set of unstudied non–protein-coding genes that serve important functions in cell biology.

lncRNA CRISPRi phenotypes are reproducible with robust knockdown

Extensive validation studies support the low false-positive and false-negative rates of our studies. First, we individually cloned the top two sgRNAs targeting 65 representative lncRNA hit loci, 41 of which were hits in only one cell line. We used internally controlled growth assays, in which the fraction of cells infected with an sgRNA were measured over time by flow cytometry, to test whether the observed phenotypes from the screens were reproducible. We monitored the growth effects of sgRNAs in the cell lines in which they exhibited a phenotype in the screen, as well as several sgRNAs in cell lines where they showed no effect, and found that the individual sgRNA growth phenotypes (γ) correlated well with the screen γ (Pearson $r = 0.72$; Fig. 2A). This confirmed both that lncRNA knockdown phenotypes were reproducible and that the difference in lncRNA phenotype between cell lines was not due to technical differences between genome-scale screens. Analyzing these phenotypes over time further revealed distinct kinetics of cell depletion mediated by lncRNA knockdown (Fig. 2B). For 12 lncRNA hits, we measured the levels of knockdown by quantitative polymerase chain reaction (qPCR) and found 70 to 95% knockdown for most of the targeted transcripts (14/14 sgRNAs in U87 cells; 10/16 sgRNAs in MCF7 cells) despite the effect of cellular depletion (fig. S7A).

In four cell lines, knockdown of lncRNA *PVT1* had a pro-growth phenotype. Because *PVT1* had previously been characterized as a proto-oncogene (*45*) and pro-growth phenotypes in cancer cell lines are uncommon (*23*, *46*), we validated the pro-growth phenotype (Fig. 2C and fig. S7A) and investigated this complex locus further by conducting a CRISPRi screen in U87 cells with an sgRNA library tiling every possible site along the locus (17,469 sgRNAs). We found that only sgRNAs within 1 kb of the most upstream TSSs, which are distal to any mapped enhancers, caused a consistent pro-growth phenotype (Fig. 2D, fig. S7B, and table S8). Within this TSS region, the majority of sgRNAs promoted cell growth, and knockdown of the major isoform was confirmed by qPCR (fig. S7A). sgRNAs outside of this 1-kb window around the TSS, which would not be expected to affect transcription of the major *PVT1* isoform (*23*), showed no consistent impact on growth; this finding implies that the observed pro-growth phenotype is mediated by transcriptional interference.

Repression of lncRNA loci elicits lncRNA-specific transcriptome responses

To better understand the consequences of lncRNA CRISPRi, we performed RNA-seq after CRISPRi knockdown of 42 lncRNA hits in three cell types; 32 of these lncRNA loci were hits in only one cell type. Selected lncRNA loci did not have essential

Downloaded from http://science.sciencemag.org/ on January 13, 2020

coding gene neighbors, and two or more sgRNAs per gene were tested individually. Distinct sgRNAs targeting the same lncRNA TSS resulted in highly correlated transcriptome responses (mean Pearson r = 0.980; Fig. 2E) that were generally proximal to each other in hierarchical clustering analysis (fig. S8, A to D). By contrast, pairs of sgRNAs targeting different hit lncRNA loci with the same phenotype direction had transcriptome responses that were more dissimilar (mean Pearson r = 0.942, Mann-Whitney P value compared to same-gene pairs = 6.4×10^{-8}), suggesting distinct molecular mechanisms of the lncRNAs despite having similar phenotypes (Fig. 2E).

RNA-seq analysis of differential gene expression also revealed several clusters of coexpressed genes, suggesting that growth modifier lncRNA loci regulate critical pathways (fig. S8, A to D, and table S9). For instance, two lncRNA knockdowns that caused increased growth in U87 cells clustered by up-regulation of translation genes ($P = 3.2 \times 10^{-37}$), whereas other pro-growth sgRNAs showed correlated changes in expression of DNA replication ($P = 2.0 \times 10^{-10}$) and posttranscriptional regulation ($P = 3.0 \times 10^{-8}$). Clusters enriched for genes in the p53 pathway (e.g., *ATF3*) were up-regulated by many anti-growth sgRNAs in both U87 and HeLa cells. Interestingly, K562 cells showed clusters of genes enriched for platelet degranulation ($P = 1.6 \times 10^{-5}$) and response to decreasing oxygen levels ($P = 5.0 \times 10^{-5}$). The median magnitudes of $\log_2$ fold changes for differentially expressed genes in U87, HeLa, and K562 cells were 0.67, 0.86, and 1.17, respectively (fig. S8E), with several genes exhibiting up- or down-regulation by a factor of >2 consistently across many samples (fig. S8F). These results indicate that different lncRNAs can regulate distinct biological pathways that affect cell growth and proliferation.

Analysis of the chromosomal location of differentially expressed genes did not reveal a global trend toward transcriptional changes on the targeted chromosome (fig. S9). We did, however, find that knockdown of 14 lncRNA loci resulted in local transcriptional changes within a 20-gene window (fig. S10), suggesting that certain lncRNAs may preferentially act locally.

CRISPRi robustly inhibits lncRNA transcription

The fraction of growth modifier lncRNA loci identified in our screens (1 to 8% per cell line) was less than the fraction of essential protein-coding genes in previous reports (10 to 11%) (*35*, *46*). We therefore wanted to assess whether lncRNA genes that did not appear as a hit in any screen were true negatives or were simply a result of ineffective repression by CRISPRi. To this end, using all 10 sgRNAs per gene, we measured the knockdown of five arbitrarily selected lncRNA genes that had no observed phenotype in any cells and were expressed in both K562 and U87 cells (Fig. 2F and fig. S7C). Of these 100 knockdown measurements, 61 showed >90% repression of the targeted lncRNA. Furthermore, with the exception of *LOC100506710* in U87 cells, all lncRNAs were repressed by at least 90% by at least three different sgRNAs. For all sgRNAs, lncRNA knockdown efficiency correlated with their predicted CRISPRi activity, and the efficiency of knockdown was highly correlated between K562 and U87 cells (Pearson r = 0.78; Fig. 2G). On the basis of these findings, with the exception of cases where a small amount of residual transcript is sufficient for lncRNA function, we infer that the majority of lncRNA loci that did not appear as a screen hit produce transcripts that are not essential for robust growth of the cell line screened.

Growth modifier lncRNA function is highly cell type–specific

We next determined the number of lncRNA hits that were unique to a specific cell type or common to any combination of two or more of the cell types screened. The vast majority (89.4%) of lncRNA hits were unique to only one cell type, with none being a hit in five or more cell types (Fig. 3, A to C). Even when we restricted this analysis to the 1329 lncRNAs expressed in all seven cell types, 82.6% of the lncRNA hits modified growth in only one cell type (Fig. 3B). Analysis of cell type specificity scores based on the Jensen-Shannon distance, which quantifies how closely a given distribution resembles "perfect" specificity (*37*), revealed that the specificity of lncRNA screen scores was far greater than the specificity of lncRNA expression for lncRNA hits (Fig. 3D). Therefore, differential expression patterns alone are not sufficient to predict functional lncRNAs. Cross-comparison of screen score distributions for lncRNAs that scored as hits in each cell type revealed that the threshold used for calling hits did not account for the cell type specificity (Fig. 3E and fig. S11, D and E). Furthermore, cross-comparison of screen scores between replicates did not support technical variation as the source of the apparent cell type–specific function (Fig. 3F and fig. S11F).

In contrast to the sparse cell type overlap of lncRNA hits, analysis of published protein-coding screens across similar numbers of cell types (*46*, *47*) revealed that the majority [54.8% in (*47*), 67.3% in (*46*)] of essential protein-coding genes are hits in two or more cell types, with 20.4% and 30.8% being essential to all cell types screened in (*47*) and (*46*), respectively (Fig. 3C and fig. S11, A and B). In addition, "neighbor hits" (lncRNA loci that are within 1 kb of an essential protein-coding gene) were more likely to modify growth in multiple cell types, which suggests that CRISPRi targeted to these loci represses the adjacent essential coding gene, at least in some cases (Fig. 3C and fig. S11, C and E).

Cell type–specific lncRNAs elicit highly divergent phenotypes

We sought to better understand the cell type–specific function of specific lncRNAs. We focused on *LINC00263*, which, despite being expressed in all seven cell lines screened, had a much stronger negative growth phenotype in U87 than in any other cell line (fig. S12A). The abundance of *LINC00263* transcript in a given cell line was also poorly correlated with the corresponding screen phenotype (Pearson i = 0.266). Validation of these screen results in internally controlled growth assays showed that two distinct sgRNAs to the TSS of *LINC00263* reduced the propagation of only U87 cells and not K562, MCF7, or HeLa cells (Fig. 4A). H3K9me3 is a chromatin modification that is a result of local dCas9-KRAB activity (*31*), and in both U87 and HeLa cells with *LINC00263* CRISPRi targeting, ChIP-seq analysis demonstrated equal enrichment of H3K9me3 specifically at the *LINC00263* promoter for two independent sgRNAs (Fig. 4, B and C, and fig. S12, B and C). However, despite such evidence of equivalent and specific CRISPRi targeting, U87 and HeLa cells had substantially different transcriptome changes after *LINC00263* knockdown. Although U87 cells up-regulated genes related to ER stress (e.g., *ATF4*, *CHAC1*; GO term $P = 4.51 \times 10^{-9}$) and apoptosis (e.g., *DDIT3*, *SOD2*; GO term $P = 3.39 \times 10^{-8}$), only *LINC00263* itself was differentially expressed in HeLa cells (adjusted P < 0.05; Fig. 4D). In K562 cells, these same two sgRNAs also produced very little transcriptional change (fig. S12D). Of note, in all three cell lines, the knockdown efficiency of *LINC00263* was equivalent (Fig. 4D and fig. S12D). Consistent with our observations for *LINC00263*, knockdown of *PVT1* and *LINC00909*, which were hits in U87 cells but not in HeLa cells, produced many more differentially expressed genes in U87 cells (fig. S12E). By contrast, depletion of *LINC00680,* which was a hit in both U87 and HeLa cells, resulted in comparable numbers of differentially expressed genes in U87 and HeLa cells (fig. S12E). Our results suggest that the specificity of lncRNA function is not due to differences in CRISPRi activity but is related to differences in transcriptional networks across cell types.

We then targeted the *LINC00263* lncRNA transcript with antisense oligonucleotides (ASOs) that degrade RNA via a ribonuclease H–based mechanism. In both U87 and HeLa cells, ASOs reduced *LINC00263* transcript levels by 85 to 95% (Fig. 4E). However, *LINC00263* ASOs decreased proliferation in U87 cells but not in HeLa cells (Fig. 4, F and G). The magnitude of proliferation decrease was also comparable to CRISPRi (fig. S12, F and G), further supporting the cell type–specific function of this lncRNA. ASO knockdown of three other U87 lncRNA hits also reduced cell proliferation (fig. S12, H and I), providing additional evidence for the functional contribution of the lncRNA molecule in these examples.

Machine learning identifies features predictive of growth modifier lncRNAs

Using data from our genome-scale screens, we sought to identify properties of the lncRNA hits that can distinguish them from nonhit lncRNAs. We compared 18 classes of genomic data such as enhancer maps, expression levels, chromosomal looping data, conservation, and copy number variation from ENCODE (*40*), FANTOM (*48*), Vista (*49*), and other sources (*50*–*52*) with all lncRNA loci screened in this study. Several of

Downloaded from http://science.sciencemag.org/ on January 13, 2020

these properties—expression, Pol2/CTCF looping by chromatin interaction analysis with paired-end tag sequencing (ChIA-PET), enhancers and superenhancers from (*51*), and copy number variation—were cell type–dependent. Generalized linear models were constructed to assess which genomic properties are predictive of lncRNA function (see supplementary materials). Expression levels within each cell line, in each lncRNA gene body within 1 kb of a mapped FANTOM Enhancer, and in each lncRNA gene body within 5 kb of a cancer-associated single-nucleotide polymorphism (SNP) (*50*), as well as the number of exons, were all significant predictors of lncRNA hits ($P < 0.01$) in repeated 10-fold cross-validation (Fig. 5 and table S10); 99.6% of lncRNA genes that were screened but not apparently expressed were not called as hits (Fig. 5C). Whether the 11 growth modifier hits of such "non-expressed" lncRNA loci represent non–lncRNA-mediated effects, inaccurate quantitation of the transcript levels, or effects mediated by lncRNAs acting at low expression remains to be determined. In support of the latter possibility, *HOTTIP* has been reported to function despite being expressed at ~0.3 copies per cell (*53*). Nonetheless, many highly expressed lncRNAs were not hits [e.g., 154 nonhit lncRNAs were detected at >100 fragments per kilobase million (FPKM)], and the accuracy for predicting lncRNA hits was greater for a model using all variables than for a model that relied only on expression levels (Fig. 5B).

Relative to nonhit lncRNAs, hit lncRNA gene bodies were 1.66 times as likely to be within 1 kb of a mapped enhancer (Fig. 5D). This represented 127 of the lncRNA hit loci identified in our screens. However, the FANTOM enhancer annotations used for our analyses were derived from hundreds of different cell types, and thus only a fraction of these enhancers are active in any given cell type in our screen (*48*, *49*). Hit loci were also 1.4 times as likely to be within 5 kb of a cancer-associated SNP (Fig. 5E). That our hits were enriched for multiexonic lncRNAs is consistent with the concept that lncRNA splicing can be an aspect of lncRNA function (*26*) (Fig. 5F). However, the explanatory power of exon number was relatively low, and our screen did identify several single-exon hits such as *NEAT1*. However, no genomic property analyzed, alone or in aggregate, fully predicted growth modifier lncRNAs in a given cell type, underscoring the importance of performing loss-of-function screens for defining sets of functional genes.

Discussion

By using CRISPRi for systematic, large-scale screens for lncRNA function in multiple cell lines, we identified 499 lncRNA loci that are required for robust cell growth. This work increases considerably the number of known functional lncRNAs and reveals that the large majority (89%) of identified lncRNA genes modified growth in just one cell type. Studies of the protein-coding genome with similar large-scale screening efforts showed that an essential gene in one cell type is highly likely to be essential in the other cell types tested (*46*, *47*). In contrast to protein-coding genes, of the 1329 lncRNA genes expressed in all seven cell lines tested, not one lncRNA gene was required for robust cell growth in all cell types, with the large majority of lncRNA gene hits being specific to just one cell line. Our results thus reveal a critical role of cellular context in determining lncRNA function.

Several clues to this specificity of lncRNA function emerge from our analyses. First, although cell type–specific expression of lncRNAs was the strongest predictor of lncRNA hits in our machine learning model (Fig. 5, A and C), it did not fully explain this functional specificity (Figs. 3 and 5B). For example, RNA-seq analysis points to *LINC00263* playing a role in a complex transcriptional network required for U87 cells, but despite being expressed in other cell types, *LINC00263* appears dispensable for the normal expression of nearly all genes in these other cells (Fig. 4D and fig. S12, D and E). Taking advantage of the scale of our data set, we have also begun to discover genomic features that predict growth-modifying function. Our finding that enhancer proximity and chromosome contacts correlate with lncRNA function suggests that higher-order chromatin structure can play a role in such specificity of lncRNA function (*28*–*30*). The extent to which cell type–specific function of enhancer-templated lncRNAs results from repression of the transcript itself or its genomic locus remains an important open question. In any case, the association of lncRNA function with higher-order chromatin structure is consistent with the emerging view that chromosomal looping between lncRNA promoters and target genes differs between cell types (*54*) and is critical to lncRNA function (*55*). Finally, our finding that genomic regions containing growth modifier lncRNAs are enriched for cancer risk SNPs suggests that these lncRNAs may contribute to the pathogenesis of cancer.

Regardless of the mechanism(s) of the observed cell type specificity of lncRNAs, this finding has implications for understanding the biological roles of lncRNAs. lncRNAs appear to have originated much later than protein-coding genes, consistent with their not playing generic housekeeping roles (*3*, *56*). Our study, which focused on lncRNAs required for robust cell growth, underestimates the true number of functional lncRNAs in these cell types, as lncRNAs have been shown to regulate more evolutionarily complex cellular decisions such as cell fate (*7*, *19*, *57*, *58*), cancer metastasis (*59*, *60*), and perhaps neuronal function (*61*). The CRISPRi tools developed here can now be applied to the study of such higher-order cellular processes, where lncRNAs might exhibit even greater richness of function. Finally, the exquisite cell type specificity of lncRNA gene function has clear implications for targeted therapy.

REFERENCES AND NOTES

1. S. Djebali *et al*., Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012). doi: 10.1038/nature11233
2. A. R. R. Forrest *et al*., A promoter-level mammalian expression atlas. *Nature* **507**, 462–470 (2014). doi: 10.1038/nature13182
3. I. Ulitsky, D. P. Bartel, lincRNAs: Genomics, evolution, and mechanisms. *Cell* **154**, 26–46 (2013). doi: 10.1016/j.cell.2013.06.020
4. J. L. Rinn, H. Y. Chang, Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **81**, 145–166 (2012). doi: 10.1146/annurev-biochem-051410-092902
5. C. P. Ponting, P. L. Oliver, W. Reik, Evolution and functions of long noncoding RNAs. *Cell* **136**, 629–641 (2009). doi: 10.1016/j.cell.2009.02.006
6. A. R. Bassett *et al*., Considerations when investigating lncRNA function in vivo. *eLife* **3**, e03058 (2014). doi: 10.7554/eLife.03058
7. M. Sauvageau *et al*., Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *eLife* **2**, e01749 (2013). doi: 10.7554/eLife.01749
8. V. H. Meller, B. P. Rattner, The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. *EMBO J.* **21**, 1084–1091 (2002). doi: 10.1093/emboj/21.5.1084
9. E. Aparicio-Prat *et al*., DECKO: Single-oligo, dual-CRISPR deletion of genomic elements including long non-coding RNAs. *BMC Genomics* **16**, 846 (2015). doi: 10.1186/s12864-015-2086-z
10. T.-T. Ho *et al*., Targeting non-coding RNAs with the CRISPR/Cas9 system in human cell lines. *Nucleic Acids Res.* **43**, 10.1093/nar/gku1198 (2014). doi: 10.1093/nar/gku1198
11. T. Wang, J. J. Wei, D. M. Sabatini, E. S. Lander, Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014). doi: 10.1126/science.1246981
12. O. Shalem *et al*., Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014). doi: 10.1126/science.1247005
13. J. Shi *et al*., Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* **33**, 661–667 (2015). doi: 10.1038/nbt.3235
14. Y. Yin *et al*., Opposing roles for the lncRNA haunt and its genomic locus in regulating HOXA gene activation during embryonic stem cell differentiation. *Cell Stem Cell* **16**, 504–516 (2015). doi: 10.1016/j.stem.2015.03.007
15. V. R. Paralkar *et al*., Unlinking an lncRNA from its associated cis element. *Mol. Cell* **62**, 104–110 (2016). doi: 10.1016/j.molcel.2016.02.029
16. A. F. Groff *et al*., In vivo characterization of Linc-p21 reveals functional cis-regulatory DNA elements. *Cell Rep.* **16**, 2178–2186 (2016). doi: 10.1016/j.celrep.2016.07.050
17. S. Zhu *et al*., Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. *Nat. Biotechnol.* 10.1038/nbt.3715 (2016). doi: 10.1038/nbt.3715
18. M. Guttman *et al*., lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **477**, 295–300 (2011). doi: 10.1038/nature10398
19. N. Lin *et al*., An evolutionarily conserved long noncoding RNA TUNA controls pluripotency and neural lineage commitment. *Mol. Cell* **53**, 1005–1019 (2014). doi: 10.1016/j.molcel.2014.01.021
20. B. Adamson, A. Smogorzewska, F. D. Sigoillot, R. W. King, S. J. Elledge, 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 (2012). doi: 10.1038/ncb2426
21. Y. Zeng, B. R. Cullen, RNA interference in human cells is restricted to the cytoplasm. *RNA* **8**, 855–860 (2002). doi: 10.1017/S1355838202020071
22. L. A. Gilbert *et al*., CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013). doi: 10.1016/j.cell.2013.06.044
23. L. A. Gilbert *et al*., Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* **159**, 647–661 (2014). doi: 10.1016/j.cell.2014.09.029
24. L. S. Qi *et al*., Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013). doi: 10.1016/j.cell.2013.02.022
25. H. Nishimasu *et al*., Crystal structure of *Staphylococcus aureus* Cas9. *Cell* **162**, 1113–1126 (2015). doi: 10.1016/j.cell.2015.08.007
26. J. M. Engreitz *et al*., Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452–455 (2016).
27. A. E. Kornienko, P. M. Guenzl, D. P. Barlow, F. M. Pauler, Gene regulation by the act of long non-coding RNA transcription. *BMC Biol.* **11**, 59 (2013). doi: 10.1186/1741-7007-11-59
28. U. A. Ørom *et al*., Long noncoding RNAs with enhancer-like function in human cells. *Cell* **143**, 46–58 (2010). doi: 10.1016/j.cell.2010.09.001
29. W. Li *et al*., Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* **498**, 516–520 (2013). doi: 10.1038/nature12210

Downloaded from http://science.sciencemag.org/ on January 13, 2020

30. C. P. Fulco *et al*., Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* **354**, 769–773 (2016). doi: 10.1126/science.aag2445
31. P. I. Thakore *et al*., Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* **12**, 1143–1149 (2015). doi: 10.1038/nmeth.3630
32. A. Amabile *et al*., Inheritable silencing of endogenous genes by hit-and-run targeted epigenetic editing. *Cell* **167**, 219–232.e14 (2016). doi: 10.1016/j.cell.2016.09.006
33. M. A. Mandegar *et al*., CRISPR interference efficiently induces specific and reversible gene silencing in human iPSCs. *Cell Stem Cell* **18**, 541–553 (2016). doi: 10.1016/j.stem.2016.01.022
34. C. J. Braun *et al*., Versatile in vivo regulation of tumor phenotypes by dCas9-mediated transcriptional perturbation. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E3892–E3900 (2016). doi: 10.1073/pnas.1600582113
35. M. A. Horlbeck *et al*., Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *eLife* **5**, e19760 (2016). doi: 10.7554/eLife.19760
36. K. Takahashi *et al*., Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007). doi: 10.1016/j.cell.2007.11.019
37. M. N. Cabili *et al*., Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **25**, 1915–1927 (2011). doi: 10.1101/gad.17446611
38. M. K. Iyer *et al*., The landscape of long noncoding RNAs in the human transcriptome. *Nat. Genet.* **47**, 199–208 (2015). doi: 10.1038/ng.3192
39. A. Yates *et al*., Ensembl 2016. *Nucleic Acids Res.* **44**, D710–D716 (2015). doi: 10.1093/nar/gkv1157
40. ENCODE Project Consortium, The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* **306**, 636–640 (2004). doi: 10.1126/science.1105136
41. F. R. Kreitzer *et al*., A robust method to derive functional neural crest cells from human pluripotent stem cells. *Am. J. Stem Cells* **2**, 119–131 (2013).
42. S. J. Liu *et al*., Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biol.* **17**, 67 (2016). doi: 10.1186/s13059-016-0932-1
43. M. Kampmann, M. C. Bassik, J. S. Weissman, Integrated platform for genome-wide screening and construction of high-density genetic interaction maps in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E2317–E2326 (2013). doi: 10.1073/pnas.1307002110
44. T. Hart, K. R. Brown, F. Sircoulomb, R. Rottapel, J. Moffat, Measuring error rates in genomic perturbation screens: Gold standards for human functional genomics. *Mol. Syst. Biol.* **10**, 733 (2014). doi: 10.15252/msb.20145216
45. Y.-Y. Tseng *et al*., PVT1 dependence in cancer with MYC copy-number increase. *Nature* **512**, 82–86 (2014).
46. T. Wang *et al*., Identification and characterization of essential genes in the human genome. *Science* **350**, 1096–1101 (2015). doi: 10.1126/science.aac7041
47. T. Hart *et al*., High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell* **163**, 1515–1526 (2015). doi: 10.1016/j.cell.2015.11.015
48. R. Andersson *et al*., An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455–461 (2014). doi: 10.1038/nature12787
49. A. Visel, S. Minovitsky, I. Dubchak, L. A. Pennacchio, VISTA Enhancer Browser–a database of tissue-specific human enhancers. *Nucleic Acids Res.* **35**, D88–D92 (2007). doi: 10.1093/nar/gkl822
50. X. Yan *et al*., Comprehensive genomic characterization of long non-coding RNAs across human cancers. *Cancer Cell* **28**, 529–540 (2015). doi: 10.1016/j.ccell.2015.09.006
51. D. Hnisz *et al*., Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013). doi: 10.1016/j.cell.2013.09.053
52. J. Chen *et al*., Evolutionary analysis across mammals reveals distinct classes of long non-coding RNAs. *Genome Biol.* **17**, 19 (2016). doi: 10.1186/s13059-016-0880-9
53. K. C. Wang *et al*., A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**, 120–124 (2011). doi: 10.1038/nature09819
54. W. Ma *et al*., Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of human lincRNA genes. *Nat. Methods* **12**, 71–78 (2015). doi: 10.1038/nmeth.3205
55. J. M. Engreitz *et al*., The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341**, 1237973 (2013). doi: 10.1126/science.1237973
56. A. Necsulea *et al*., The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* **505**, 635–640 (2014). doi: 10.1038/nature12943
57. A. D. Ramos *et al*., The long noncoding RNA Pnky regulates neuronal differentiation of embryonic and postnatal neural stem cells. *Cell Stem Cell* **16**, 439–447 (2015). doi: 10.1016/j.stem.2015.02.007
58. M. Kretz *et al*., Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **493**, 231–235 (2013). doi: 10.1038/nature11661
59. T. Gutschner *et al*., The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res.* **73**, 1180–1189 (2013). doi: 10.1158/0008-5472.CAN-12-2850
60. R. A. Gupta *et al*., Long non-coding RNA HOTAIR reprograms chromatinstate to promote cancer metastasis. *Nature* **464**, 1071–1076 (2011). doi: 10.1038/nature08975
61. J. A. Briggs, E. J. Wolvetang, J. S. Mattick, J. L. Rinn, G. Barry, Mechanisms of long non-coding RNAs in mammalian nervous system development, plasticity, disease, and evolution. *Neuron* **88**, 861–877 (2015). doi: 10.1016/j.neuron.2015.09.045

ACKNOWLEDGMENTS

We thank the members of the Lim and Weissman labs, particularly A. Fields, J. Dunn, M. DeVera, M. Cui, and D. Wu, for helpful discussions and assistance; A. Truong for assistance with iPS cell culturing; N. Salomonis for iPSC RNA-seq data; E. Chow and D. Bogdanoff of the UCSF Center for Advanced Technology for sequencing assistance; and L. Bruhn, D. Ryan, L. Fairbairn, and P. Tsang of Agilent Technologies for their assistance on the design and synthesis of oligonucleotide pools. Supported by NIH grant 1R01NS091544-01A1, VA grant 5I01 BX000252-07, NIH Specialized Programs of Research Excellence Developmental Research Program subaward, the Shurl and Kay Curci Foundation, the LoGlio Foundation, and the Hana Jabsheh Initiative (D.A.L.); NIH grant F30 NS092319-01 (S.J.L.); the Howard Hughes Medical Institutes and NIH grants P50 GM102706, U01 CA168370, and R01 DA036858 (M.A.H., J.E.V., M.Y.C., Y.C., L.A.G., and J.S.W.); NIH grants R35-CA209919 and P50-HG007735 (S.W.C. and H.Y.C.); the Gladstone Institutes and NIH grants U01HL100406, P01HL089707, and R01HL130533 (B.R.C. and M.A.M.); and NIH/NCI Pathway to Independence Award K99CA204602 (L.A.G.). Oligonucleotide pools were provided courtesy of the Innovative Genomics Initiative. M.A.H., L.A.G., and J.S.W. are inventors on patent application PCT/US15/40449 submitted by UCSF that covers CRISPRi library design. The human iPSC line WTC expressing the CRISPRi system (WTC-CRISPRi Gen IC) is available from B.R.C. under material transfer agreement from the Gladstone Institutes. The parental iPSC line (WTC) is available from the Coriell Biorepository #GM25256. The CRiNCL libraries are available from J.S.W. (via Addgene for academic users) under a material transfer agreement from UCSF and Agilent Technologies. M.A.H., L.A.G., and J.S.W. are inventors on patent application PCT/US15/40449 submitted by UCSF that covers CRISPRi library design. J.S.W. is a founder of KSQ Therapeutics, a company that uses CRISPR-based screening to identify therapeutic targets. H.Y.C. is a co-founder of Epinomics Inc. and served on the scientific advisory board of RaNA Therapeutics. M.A.H. and L.A.G. are consultants for KSQ Therapeutics.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/cgi/content/355/6320/aah7111/suppl/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 to S11
References (*62*–*70*)

3 August 2016; accepted 30 November 2016
Published online 15 December 2016;
10.1126/science.aah7111

Downloaded from http://science.sciencemag.org/ on January 13, 2020

Science

CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells

S. John Liu, Max A. Horlbeck, Seung Woo Cho, Harjus S. Birk, Martina Malatesta, Daniel He, Frank J. Attenello, Jacqueline E. Villalta, Min Y. Cho, Yuwen Chen, Mohammad A. Mandegar, Michael P. Olvera, Luke A. Gilbert, Bruce R. Conklin, Howard Y. Chang, Jonathan S. Weissman and Daniel A. Lim

Science **355** (6320), eaah7111.
DOI: 10.1126/science.aah7111originally published online December 15, 2016

A very focused function for lncRNAs

The human genome generates many thousands of long noncoding RNAs (lncRNAs). A very small number of lncRNAs have been shown to be functional. Liu *et al.* carried out a large-scale CRISPR-based screen to assess the function of ~17,000 lncRNAs in seven different human cell lines. A considerable number (~500) of the tested lncRNAs influenced cell growth, suggesting biological function. In almost all cases, though, the function was highly cell type—specific, often limited to just one cell type.

Science, this issue p. 10.1126/science.aah7111

ARTICLE TOOLS http://science.sciencemag.org/content/355/6320/eaah7111

SUPPLEMENTARY MATERIALS http://science.sciencemag.org/content/suppl/2016/12/14/science.aah7111.DC1

REFERENCES This article cites 70 articles, 13 of which you can access for free
http://science.sciencemag.org/content/355/6320/eaah7111#BIBL

PERMISSIONS http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2017, American Association for the Advancement of Science

Downloaded from http://science.sciencemag.org/ on January 13, 2020