immune responses in control of viral infection in the liver, and provide a paradigm for HAV pathogenesis that is likely relevant to other hepatotropic human viruses.

REFERENCES AND NOTES

1. C. Dembek, U. Protzer, *Med. Microbiol. Immunol.* **204**, 95–102 (2015).
2. B. Y. Winer, Q. Ding, J. M. Gaska, A. Ploss, *FEBS Lett.* **590**, 1987–1999 (2016).
3. R. E. Lanford, S. M. Lemon, C. Walker, in *Hepatitis C Antiviral Drug Discovery & Development*, Y. He, T. Tan, Eds. (Horizon Scientific, Norwich, UK, 2011), pp. 99–132.
4. R. E. Lanford *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11223–11228 (2011).
5. U. Protzer, M. K. Maini, P. A. Knolle, *Nat. Rev. Immunol.* **12**, 201–213 (2012).
6. I. N. Crispe, *Annu. Rev. Immunol.* **27**, 147–163 (2009).
7. J. Kaiser, *Science* **350**, 1013 (2015).
8. Z. Feng *et al.*, *Nature* **496**, 367–371 (2013).
9. F. Deinhardt, J. B. Deinhardt, in *Hepatitis A*, R. J. Gerety, Ed. (Academic Press, 1984), pp. 185–204.
10. B. Hornei *et al.*, *J. Med. Virol.* **64**, 402–409 (2001).
11. D. A. Feigelstock, P. Thompson, G. G. Kaplan, *J. Virol.* **79**, 2950–2955 (2005).
12. J. F. Drexler *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15190–15195 (2015).
13. Y. Yang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7253–7258 (2007).
14. L. Qu *et al.*, *PLOS Pathog.* **7**, e1002169 (2011).
15. See supplementary materials on *Science* Online.
16. S. M. Lemon, *N. Engl. J. Med.* **313**, 1059–1067 (1985).
17. J. Pott *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7944–7949 (2011).
18. T. J. Nice *et al.*, *Science* **347**, 269–273 (2015).
19. Y. Lei *et al.*, *PLOS ONE* **4**, e5466 (2009).
20. K. Guan *et al.*, *Mol. Cell. Biol.* **33**, 3137–3149 (2013).
21. A. N. Schulman *et al.*, *J. Infect. Dis.* **134**, 80–84 (1976).
22. C. M. Walker, Z. Feng, S. M. Lemon, *Curr. Opin. Virol.* **11**, 7–13 (2015).
23. H. Ikushima, H. Negishi, T. Taniguchi, *Cold Spring Harb. Symp. Quant. Biol.* **78**, 105–116 (2013).
24. H. W. Chen *et al.*, *J. Immunol.* **191**, 4194–4201 (2013).
25. H. M. Lazear *et al.*, *PLOS Pathog.* **9**, e1003118 (2013).
26. S. Faouzi *et al.*, *J. Biol. Chem.* **276**, 49077–49082 (2001).
27. M. G. Wathelet *et al.*, *Mol. Cell* **1**, 507–518 (1998).
28. N. Grandvaux *et al.*, *J. Virol.* **76**, 5532–5539 (2002).
29. R. Lin, C. Heylbroeck, P. Genin, P. M. Pitha, J. Hiscott, *Mol. Cell. Biol.* **19**, 959–966 (1999).
30. B. Rehermann, *Cell. Mol. Gastroenterol. Hepatol.* **1**, 578–588 (2015).
31. F. Marra, F. Tacke, *Gastroenterology* **147**, 577–594.e1 (2014).
32. N. C. Reich, *J. Interferon Cytokine Res.* **33**, 199–205 (2013).
33. R. Besch *et al.*, *J. Clin. Invest.* **119**, 2399–2411 (2009).
34. S. Chattopadhyay, T. Kuzmanovic, Y. Zhang, J. L. Wetzel, G. C. Sen, *Immunity* **44**, 1151–1161 (2016).
35. J. Petrasek *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 16544–16549 (2013).

ACKNOWLEDGMENTS

We thank D. Yamane, K. McKnight, T. Benzine, L. Wai, D. Hilliard, and M. Chua for helpful discussions and technical assistance; R. Lanford (Texas Biomedical Research Institute) for chimpanzee-passaged HAV; and C. Walker (Research Institute of Nationwide Children's Hospital) for the generous gift of HAV peptides. The data presented in this manuscript are tabulated in the main paper and in the supplementary materials. DNA sequences are deposited in GenBank with accession numbers KX343014, KX343015, KX343016, KX343017, and KX343018. *Irf3*$^{-/-}$, *Irf7*$^{-/-}$, and *Irf3*$^{-/-}$*Irf7*$^{-/-}$ mice were provided under a materials transfer agreement from Tokyo University; *Mavs*$^{-/-}$ mice were provided under a materials transfer agreement from the University of Washington. Supported by NIH grants R01-AI103083 and U19-AI109965 (S.M.L.), NIH grants R01-AI074862, R21-AI117575, and R56-AI110682 (J.K.W.), and NCI Center Core Support Grant P30-CA016086 to the Lineberger Comprehensive Cancer Center.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6307/1541/suppl/DC1
Materials and Methods
Figs. S1 to S13
Tables S1 and S2
References (*36*–*52*)

6 April 2016; accepted 2 September 2016
Published online 15 September 2016
10.1126/science.aaf8325

ENHANCER FUNCTION

High-resolution interrogation of functional elements in the noncoding genome

Neville E. Sanjana,[1,2]*†‡ Jason Wright,[1,2]† Kaijie Zheng,[1,2] Ophir Shalem,[1,2] Pierre Fontanillas,[1] Julia Joung,[1,2] Christine Cheng,[1,3] Aviv Regev,[1,3] Feng Zhang[1,2]*

The noncoding genome affects gene regulation and disease, yet we lack tools for rapid identification and manipulation of noncoding elements. We developed a CRISPR screen using ~18,000 single guide RNAs targeting >700 kilobases surrounding the genes *NF1*, *NF2*, and *CUL3*, which are involved in BRAF inhibitor resistance in melanoma. We find that noncoding locations that modulate drug resistance also harbor predictive hallmarks of noncoding function. With a subset of regions at the *CUL3* locus, we demonstrate that engineered mutations alter transcription factor occupancy and long-range and local epigenetic environments, implicating these sites in gene regulation and chemotherapeutic resistance. Through our expansion of the potential of pooled CRISPR screens, we provide tools for genomic discovery and for elucidating biologically relevant mechanisms of gene regulation.

More than 98% of the human genome does not code for proteins; however, unlike for the coding genome, there exists no overarching framework to translate the noncoding genomic sequence into functional elements (*1*, *2*). Evidence from genome-wide association studies suggests that many noncoding regions are critical for human health (*3*, *4*). The implications of these associations, however, have been difficult to assess, in part because we lack the tools to determine which variants alter functional elements. Molecular hallmarks, such as epigenetic state, chromatin accessibility, transcription factor binding, and evolutionary conservation, correlate with putative functional elements in the noncoding genome and can predict regulatory function (*2*, *5*). However, these predictions largely bypass regions lacking hallmarks, and it is difficult to ascertain which hallmarks play a correlative or truly causal role in function or phenotype (*6*, *7*). Efforts to determine causality have used preselected DNA fragments, with expression serving as a proxy for function (*8*), but these methods lack the local chromatin context and broader regulatory interactions. Thus, there is a need for systematic approaches to sift through noncoding variants and determine whether and how they affect phenotypes in a native biological context.

For this purpose, we designed a high-throughput method that uses pooled CRISPR (clustered regularly interspaced short palindromic repeat)–Cas9 single guide RNA (sgRNA) libraries to screen noncoding genomic loci in order to identify functional regions related to phenotype and gene regulation. Previous applications of CRISPR screens to the noncoding genome have focused on specific functional elements (e.g., microRNAs and transcription factor binding sites) or required fluorescent reporters (*9*–*12*). In this work, we comprehensively assayed a total of 715 kilobases (kb) of sequence surrounding three different genes by performing unbiased mutagenesis to identify functional elements relevant to cancer drug resistance.

Vemurafenib inhibits BRAF proteins carrying the V600E mutation, which are found in 50 to 70% of melanomas (*13*). Resistance to vemurafenib arises within months in almost all melanoma patients (*14*), and surviving tumor cells display increased malignancy that rapidly leads to lethality (*15*). A genome-scale CRISPR screen found that loss-of-function mutations in *NF1*, *NF2*, and *CUL3* result in vemurafenib resistance (*16*). To explore whether mutations in the noncoding regions around these three genes could similarly affect drug resistance, we designed three sgRNA libraries tiling across 100-kb regions 5′ and 3′ of each gene's major isoforms (Fig. 1A). For each library, we synthesized the sgRNAs as a pool (6682 for *NF1*, 6934 for *NF2*, and 4699 for *CUL3*; 18,315 sgRNAs in total) and cloned them into a lentiviral vector (fig. S1). We transduced A375 human melanoma cells, which carry the BRAF mutation, with the sgRNA libraries at a low multiplicity of infection and cultured them in 2 uM vemurafenib or control (dimethyl sulfoxide, DMSO) for 14 days. Using deep sequencing, we counted the representation of sgRNAs in both conditions (Fig. 1, B to D) and identified vemurafenib-enriched sgRNAs as those enriched by >4 standard deviations from the control distribution (fig. S2).

[1]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA. [2]McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Howard Hughes Medical Institute, David H. Koch Institute of Integrative Cancer Biology, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA.
***Corresponding author. Email: nsanjana@nygenome.org (N.E.S.); zhang@broadinstitute.org (F.Z.)** †These authors contributed equally to this work. ‡Present address: New York Genome Center and Department of Biology, New York University, New York, NY 10012, USA.

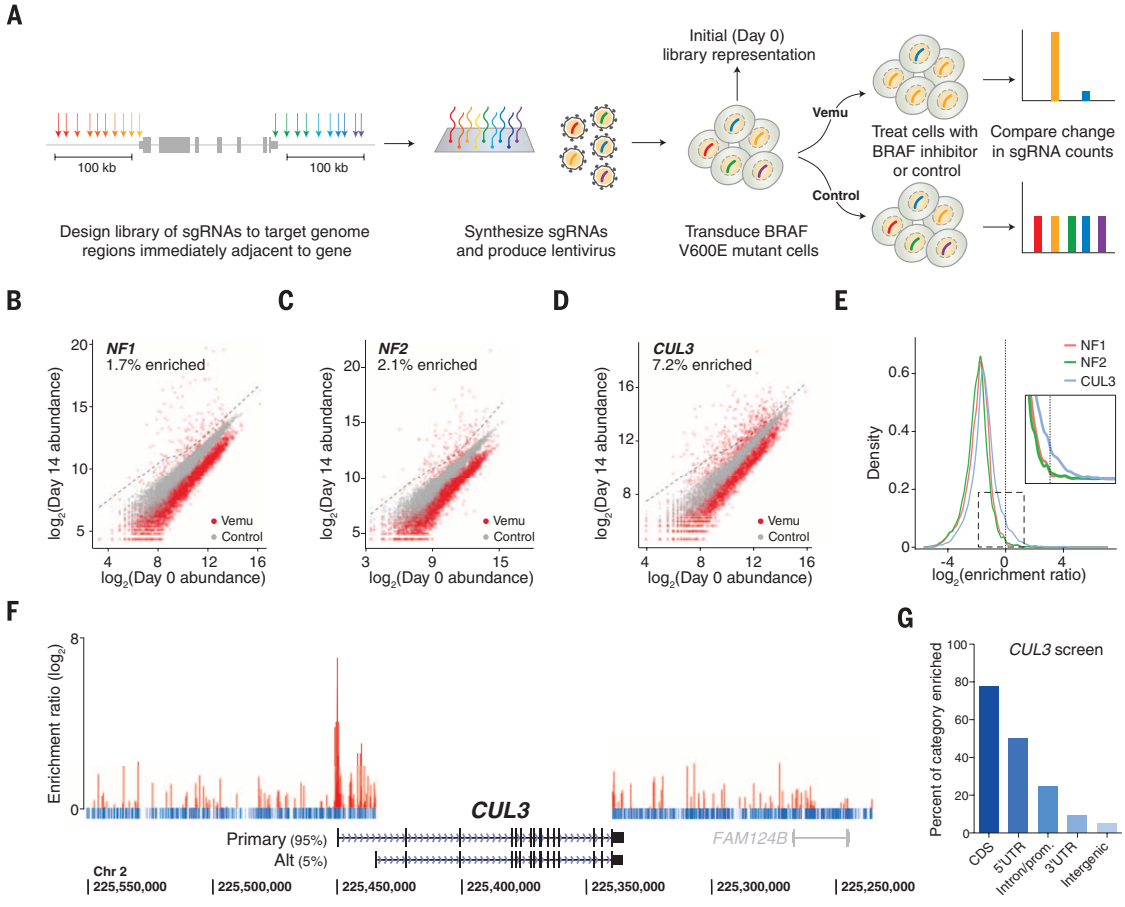

Fig. 1. CRISPR mutagenesis of noncoding regions flanking three genes involved in BRAF inhibitor resistance. (**A**) Design of sgRNA libraries targeting 100 kb 5′ and 100 kb 3′ of a gene. The sgRNAs are array-synthesized and cloned into a lentiviral vector. BRAF mutant cells are transduced with the pooled lentivirus and treated with vemurafenib (vemu) or DMSO (control). A deep sequencing readout identifies sgRNAs enriched after treatment with vemurafenib. (**B**) Scatterplot of normalized read counts (average of the two infection replicates) for *NF1*, (**C**) *NF2*, and (**D**) *CUL3* sgRNAs at days 0 (*x* axes) and 14 (*y* axes). Read counts from control (gray) and vemurafenib-treated cells (red) are shown relative to 4 standard deviations of the control cell distribution (dotted line). The percentage of sgRNAs enriched after vemurafenib treatment (>4 standard deviations) is indicated. (**E**) Distribution of the $\log_2$ ratio of the normalized read count for each sgRNA in vemurafenib to its normalized read count in control (minimum of the two infection replicates). (**F**) All *CUL3* sgRNAs, plotted by human reference genome hg19 coordinates, and the percent expression of the two most highly expressed *CUL3* isoforms [primary and alternate (alt)]. For vemurafenib-enriched sgRNAs, the $\log_2$ enrichment relative to control sgRNAs (minimum value of two replicate screens) is plotted (red); nonenriched sgRNAs are indicated in blue. (**G**) Percent of enriched sgRNAs in each genomic category. CDS, coding sequence; UTR, untranslated region; prom., promoter.

Overall, most sgRNAs were depleted after treatment with vemurafenib, which is expected because vemurafenib targets the oncogene addiction that drives A375 growth (Fig. 1E). However, in all three libraries, we found a small group of sgRNAs that were enriched after vemurafenib treatment ($\log_2$ ratio of vemurafenib to control > 0), with the *CUL3* library having the largest percentage of enriched sgRNAs. We also included a small number of sgRNAs targeting the coding region of each gene, and most of these (70 to 80%) were enriched, as expected (fig. S3A). However, among the sgRNAs targeting noncoding regions, about fourfold more were enriched in the *CUL3* library than in the *NF1* or *NF2* libraries (7.2% in *CUL3*, 1.7% in *NF1*, and 2.1% in *NF2*), suggesting the presence of more gene regulatory elements in the noncoding regions flanking the gene (fig. S3A). To determine whether this increase in putative gene regulatory elements in the 200-kb region surrounding *CUL3* is also reflected in human gene expression and genotyping data, we queried the Genotype-Tissue Expression (GTEx) database (7051 tissue samples from 449 donors). We found that *CUL3* had the largest number of cis-expression quantitative trait loci (eQTL) (n = 161 eQTLs, mean effect size = −0.21), and the region targeted by the sgRNA library overlapped with a large number of these eQTLs (fig. S3B) (*17*). We thus chose to focus our downstream analysis and validation efforts on *CUL3*.

We visualized the enriched sgRNAs in a genome browser–style view (Fig. 1F and fig. S4, A and B). We found that a higher percentage of sgRNAs targeting gene-proximal elements were enriched compared with other noncoding regions (Fig. 1G), and we found greater enrichment for sgRNAs targeting noncoding elements on the 5′ side of the gene than for those on the 3′ side (fig. S4C).

To test whether regions targeted by enriched sgRNAs from the screen physically interact with the *CUL3* promoter through chromatin looping (*18*), we created three independent chromosome conformation capture (3C) libraries (Fig. 2A) (*19*). We quantified the interaction frequency for each site across the ~200-kb region (supplementary methods) and found that regions on the 5′ side of *CUL3* tend to interact more strongly with the promoter. Regions with higher 3C interaction contained, on average, more vemurafenib-enriched sgRNAs (Fig. 2B).

Because chromatin accessibility can identify regulatory regions (*20*, *21*), we performed ATAC-seq (assay for transposase-accessible chromatin with sequencing) in A375 melanoma, MCF7 breast adenocarcinoma, and U87 glioblastoma cells (Fig. 2C). Overall, we found higher sgRNA enrichment near A375-specific ATAC peaks than near those from other cell types, and this finding was replicated with deoxyribonuclease (DNase) I hypersensitivity data (Fig. 2, D and E, and fig. S5). Enrichment in these regions suggests the presence of cell type–specific enhancers (*22*, *23*). Even though the accessible peaks overlapped with enriched sgRNA sites, the chromatin accessibility data by themselves only predict a small fraction of the total number of enriched sgRNA sites (table S1).

Given that evolutionary conservation varies widely across the noncoding genome, we sought to test whether regions exhibiting higher levels of conservation harbor more enriched sgRNAs. We examined phastCons conservation scores over the

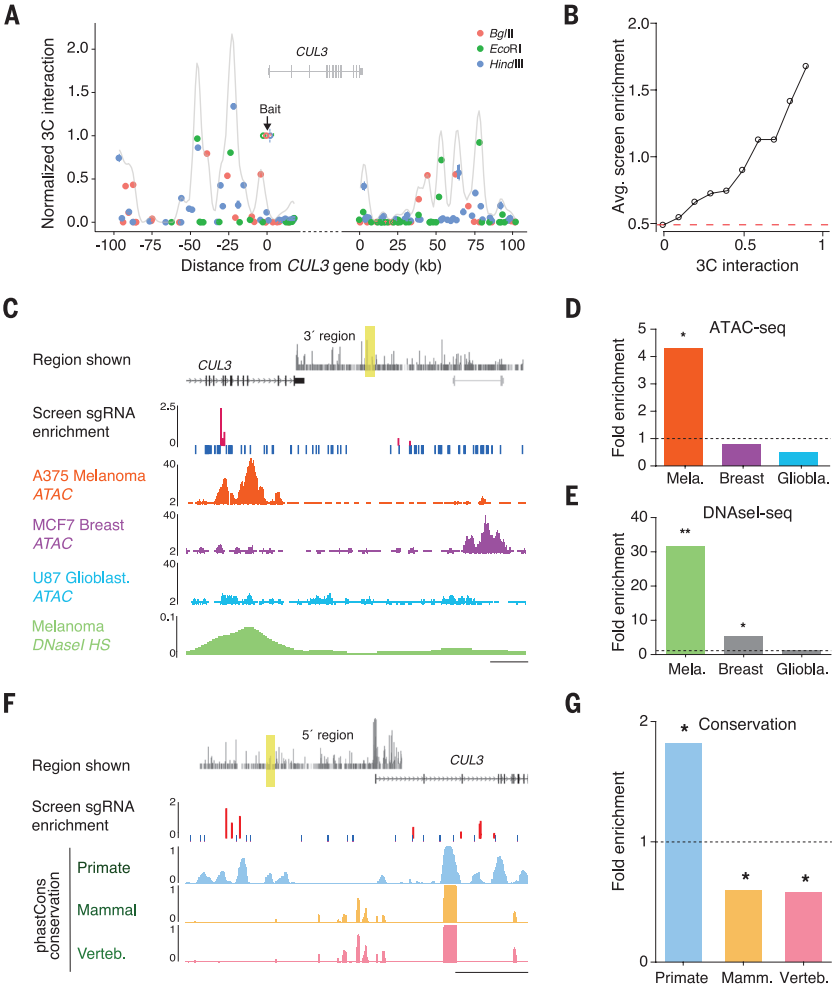

Fig. 2. Characteristics of functional noncoding elements at the *CUL3* locus. (**A**) Plot of the normalized 3C interactions with the *CUL3* promoter in A375 cells. Data points represent independent libraries generated with BglII, EcoRI, and HindIII restriction enzymes. The gray curve shows a smoothed estimate of interaction frequency. Table S3 provides the genomic coordinates of baits and probes. (**B**) Average window enrichment of sgRNAs (log_2 ratio of vemurafenib to DMSO reads) near 3C sites with the specified minimum normalized 3C interaction with the *CUL3* promoter (supplementary methods). The red dashed line indicates average window enrichment for all 3C sites in (A). (**C**) An example of enriched sgRNAs (red) that overlap with a melanoma-specific region of open chromatin. Read counts are shown from ATAC-seq in A375 melanoma (orange), MCF7 breast cancer (purple), and U87 glioblastoma (blue) cells and from melanoma DNase I hypersensitivity (HS) sequencing (green; ENCODE/Colo-829 cell line). Loci investigated with respect to *CUL3* are shown at the top (yellow). Scale bar, 500 bases. (**D**) Fold enrichment of enriched sgRNAs near ATAC-seq open chromatin peaks in melanoma, breast cancer, and glioblastoma cell lines. (**E**) Same as (D), but for DNAse I hypersensitivity sequencing. (**F**) An example of enriched sgRNAs (red) that coincide with regions that show greater primate-specific conservation than placental mammal and vertebrate conservation. Loci investigated with respect to *CUL3* are shown at the top (yellow). Scale bar, 200 bases. (**G**) Fold enrichment of enriched sgRNAs near phastCons peaks in primates, placental mammals, and vertebrates. $^{**}P < 0.01$; $^{*}P < 0.05$.

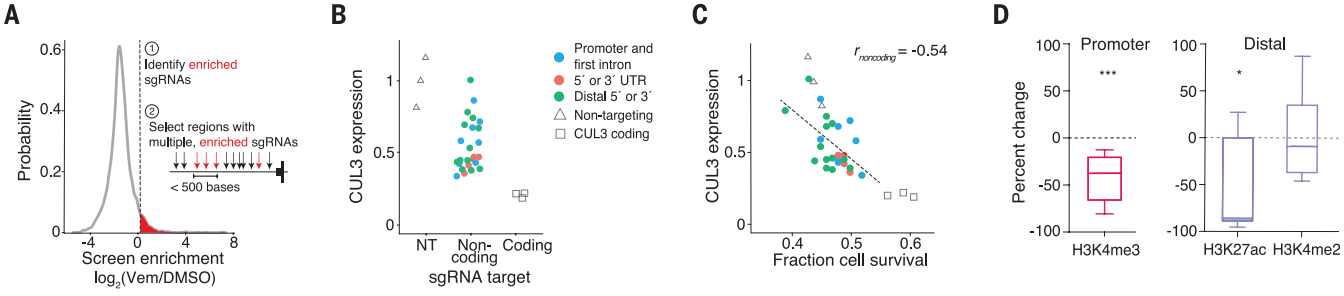

Fig. 3. Noncoding mutations affect *CUL3* expression and histone modifications. (**A**) Criteria for selecting 25 sgRNAs targeting noncoding regions for validation. (**B**) *CUL3* RNA expression (normalized to nontargeting sgRNAs) after transduction with lentivirus expressing nontargeting (triangles), noncoding region–targeting (colored circles), and coding region–targeting (squares) sgRNAs. (**C**) Relationship between *CUL3* expression and cell survival after vemurafenib treatment. The linear fit is to noncoding sgRNAs only ($r_{noncoding}$ = −0.54, P = 0.005) and does not include coding region–targeting or nontargeting sgRNAs. (**D**) Percent change in (left) average H3K4me3 ChIP for all validation sgRNAs within 1 kb of the transcription start site of *CUL3* and (right) average H3K27ac and average H3K4me2 ChIP for all validation sgRNAs >1 kb from the transcriptions start site of *CUL3*. $^{***}P < 0.001$; $^{*}P < 0.05$.

CUL3 locus among primates, placental mammals, and vertebrates (Fig. 2F) (*24*). Overall, enriched sgRNAs were ~1.8-fold more likely to be found near peaks of primate conservation and ~1.7-fold less likely to be found near conservation peaks among mammals and vertebrates (Fig. 2G and fig. S5). In contrast, the genomic sites of sgRNAs targeting coding regions of *CUL3* did not demonstrate differential conservation (phastCons probability of ~0.95 in primates, mammals, and vertebrates). This observation supports recent findings that enhancers evolve rapidly in a lineage- or species-specific manner and that conserved enhancers between mammals tend to be rare (*25*).

To confirm that mutations in these specific noncoding regions were mediated by *CUL3* and lead to altered drug resistance, we transduced cells with individual sgRNAs that have at least one other enriched sgRNA within 500 bases (Fig. 3A). We validated that the sgRNAs created mutations at the intended target sites (fig. S6) and found that 24 of the 25 sgRNAs resulted in decreased *CUL3* expression relative to nontargeting sgRNAs (Fig. 3B). As expected, there was a negative correlation between *CUL3* gene expression and vemurafenib resistance (correlation coefficient r = −0.54, P = 0.005) (Fig. 3C), and increased vemurafenib

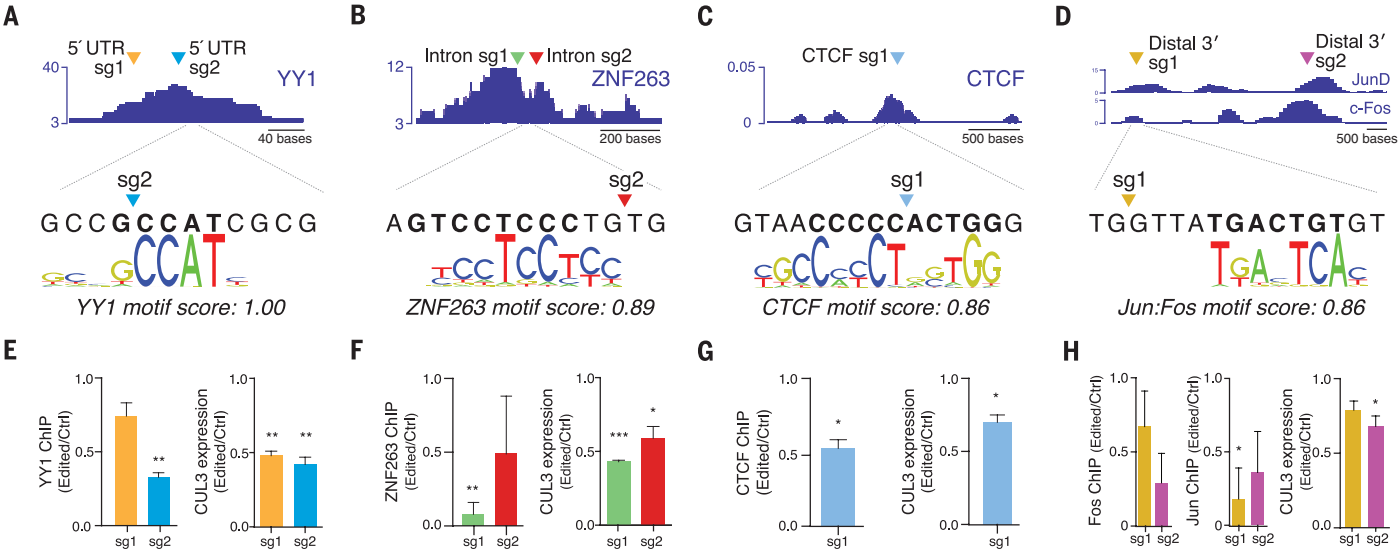

Fig. 4. Cas9 mutagenesis disrupts predicted transcription factors and DNA binding proteins at target sites of vemurafenib-enriched sgRNAs. (**A** to **D**) sgRNA target locations in relation to predicted binding sites. Table S4 gives sgRNA sequences and target locations. (**E** to **H**) Change in transcription factor or DNA binding protein occupancy around cleavage sites and change in *CUL3* expression. Both measurements are normalized to cells transduced with non-targeting sgRNAs. ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$.

resistance could be reversed by restoring *CUL3* expression (fig. S7).

Next, we surveyed changes in posttranslational histone modifications at each target site (fig. S8A). For target sites near the promoter, we found a 56% average decrease in H3K4me3 after editing ($P = 7 \times 10^{-4}$, $n = 9$ sites) (Fig. 3D), which is consistent with the reduced gene expression. At distal sites, we found a 41% average decrease in H3K27ac ($P = 0.02$, $n = 7$ sites) after editing and no significant change in H3K4me2 ($P = 0.82$, $n = 7$ sites) (Fig. 3D), although a subset of H3K4me2 levels decreased after editing (fig. S8B). We also found that mutagenesis of a ~22-kb distal histone acetyltransferase (p300) binding site that has a strong 3C promoter interaction resulted in a 75% decline of promoter H3K27ac and a 50% decrease in *CUL3* expression (fig. S9 and supplementary text).

By examining regions targeted by enriched sgRNAs, we found individual loci containing the canonical transcription factor binding motifs for yin yang 1 (YY1), zinc finger protein 263 (ZNF263), CCCTC-binding factor (CTCF), and activation protein 1 (AP-1) complex, which were disrupted after editing (Fig. 4, A to D, and fig. S10). We found that mutations within these binding sites abrogated transcription factor recruitment, leading to loss of *CUL3* expression (Fig. 4, E to H). For example, specific sgRNAs that target loci near a YY1 chromatin immunoprecipitation (ChIP) peak (Fig. 4A) disrupted the YY1 motif (fig. S11), and vemurafenib treatment selected for mutations that were more deleterious to the binding site (fig. S12). Although both of the sgRNAs targeting loci near the site decreased YY1 binding, the sgRNA whose cut site overlaps the motif disrupted YY1 binding more efficiently (67 versus 26%) (Fig. 4E). In addition, mutagenesis by either sgRNA significantly decreased *CUL3* expression. Similarly, two sgRNAs in the first intron of *CUL3*, spaced 30 bases apart, overlap a ZNF263 ChIP sequencing peak (Fig. 4B). Both resulted in a significant decrease in ZNF263 occupancy and in *CUL3* expression (Fig. 4F).

Although we observed a bias in the presence of regulatory elements 5′ of the transcription start site, we did find several enriched sgRNAs downstream of *CUL3* (Fig. 4, C and D, and supplementary text). One sgRNA cuts inside the core motif of CTCF (Fig. 4C). After editing, CTCF occupancy was decreased by 45%, with a concurrent 30% decrease in *CUL3* expression (Fig. 4G). For AP-1, a heterodimer of FOS and JUN, editing at either of two nearby sites decreased FOS and JUN binding compared with binding in control cells and decreased *CUL3* expression by ~25% (Fig. 4H). Overall, as in the pooled screen, mutagenesis at transcription factor binding sites located on the 3′ side exhibited weaker effects on gene expression than those located on the 5′ side.

Thus, we show how a Cas9-mediated systematic dissection of noncoding loci can identify functional elements involved in gene regulation and cancer drug resistance. In combination with other genome-wide assays, we demonstrate high-throughput identification of regions where changes in chromatin context and transcription factor binding are causally linked to loss of gene expression and a disease-relevant phenotype. This approach is generalizable, and we anticipate that the extension of pooled CRISPR screens into the noncoding genome will provide further insights and methods for unbiased interrogation of the genome.

REFERENCES AND NOTES

1. A. Visel, E. M. Rubin, L. A. Pennacchio, *Nature* **461**, 199–205 (2009).
2. ENCODE Project Consortium, *Nature* **489**, 57–74 (2012).
3. L. A. Hindorff *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9362–9367 (2009).
4. M. T. Maurano *et al.*, *Science* **337**, 1190–1195 (2012).
5. Roadmap Epigenomics Consortium *et al.*, *Nature* **518**, 317–330 (2015).
6. J. C. Kwasnieski, C. Fiore, H. G. Chaudhari, B. A. Cohen, *Genome Res.* **24**, 1595–1602 (2014).
7. R. Mundade, H. G. Ozer, H. Wei, L. Prabhu, T. Lu, *Cell Cycle* **13**, 2847–2852 (2014).
8. A. Melnikov *et al.*, *Nat. Biotechnol.* **30**, 271–277 (2012).
9. S. Chen *et al.*, *Cell* **160**, 1246–1260 (2015).
10. M. C. Canver *et al.*, *Nature* **527**, 192–197 (2015).
11. G. Korkmaz *et al.*, *Nat. Biotechnol.* **34**, 192–198 (2016).
12. N. Rajagopal *et al.*, *Nat. Biotechnol.* **34**, 167–174 (2016).
13. Cancer Genome Atlas Network, *Cell* **161**, 1681–1696 (2015).
14. J. A. Sosman *et al.*, *N. Engl. J. Med.* **366**, 707–714 (2012).
15. I. Zubrilov *et al.*, *Cancer Lett.* **361**, 86–96 (2015).
16. O. Shalem *et al.*, *Science* **343**, 84–87 (2014).
17. GTEx Consortium, *Science* **348**, 648–660 (2015).
18. E. Lieberman-Aiden *et al.*, *Science* **326**, 289–293 (2009).
19. J. Dekker, K. Rippe, M. Dekker, N. Kleckner, *Science* **295**, 1306–1311 (2002).
20. G. E. Crawford *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 992–997 (2004).
21. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, W. J. Greenleaf, *Nat. Methods* **10**, 1213–1218 (2013).
22. N. D. Heintzman *et al.*, *Nat. Genet.* **39**, 311–318 (2007).
23. N. C. Sheffield *et al.*, *Genome Res.* **23**, 777–788 (2013).
24. J. Felsenstein, G. A. Churchill, *Mol. Biol. Evol.* **13**, 93–104 (1996).
25. D. Villar *et al.*, *Cell* **160**, 554–566 (2015).

ACKNOWLEDGMENTS

We thank R. Macrae, D. Scott, and the entire Zhang laboratory. N.E.S. is supported by the NIH through a Pathway to Independence Award (R00-HG008171) from the National Human Genome Research Institute and a postdoctoral fellowship from the Simons Center for the Social Brain at the Massachusetts Institute of Technology. J.W. is supported by the NIH through a Ruth L. Kirschstein National Research Service Award (F32-DK096822). O.S. is a Klarman Fellow of the Broad-Israel Partnership. F.Z. is a New York Stem Cell Foundation–Robertson Investigator and is supported by the NIH through the National Institute of Mental Health (grants 5DP1-MH100706 and 1R01-MH110049); the NSF; the New York Stem Cell, Simons,

Paul G. Allen Family, and Vallee Foundations; and J. and P. Poitras, R. Metcalfe, and D. Cheng. A.R. is on the scientific advisory board for Syros Pharmaceuticals and Thermo Fisher and is a consultant for Driver Group. N.E.S., J.W., O.S., and F.Z. are listed as inventors on a related patent application. Plasmids are available from Addgene, subject to a material transfer agreement. Deep sequencing data are available at Sequence Read Archive under BioProject accession number PRJNA324504, and software tools are available on GitHub (https://github.com/nsanjana/BashRegion).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6307/1545/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S13
Tables S1 to S7
References (*26*–*42*)

25 March 2016; accepted 16 August 2016
10.1126/science.aaf7613

GENE REGULATION

Cyclin A2 is an RNA binding protein that controls *Mre11* mRNA translation

Arun Kanakkanthara,[1] Karthik B. Jeganathan,[1] Jazeel F. Limzerwala,[2] Darren J. Baker,[1] Masakazu Hamada,[1] Hyun-Ja Nam,[1] Willemijn H. van Deursen,[1] Naomi Hamada,[1] Ryan M. Naylor,[2] Nicole A. Becker,[2] Brian A. Davies,[2] Janine H. van Ree,[1] Georges Mer,[2] Virginia S. Shapiro,[3] L. James Maher III,[2] David J. Katzmann,[2] Jan M. van Deursen[1,2]*

Cyclin A2 activates the cyclin-dependent kinases Cdk1 and Cdk2 and is expressed at elevated levels from S phase until early mitosis. We found that mutant mice that cannot elevate cyclin A2 are chromosomally unstable and tumor-prone. Underlying the chromosomal instability is a failure to up-regulate the meiotic recombination 11 (Mre11) nuclease in S phase, which leads to impaired resolution of stalled replication forks, insufficient repair of double-stranded DNA breaks, and improper segregation of sister chromosomes. Unexpectedly, cyclin A2 controlled Mre11 abundance through a C-terminal RNA binding domain that selectively and directly binds *Mre11* transcripts to mediate polysome loading and translation. These data reveal cyclin A2 as a mechanistically diverse regulator of DNA replication combining multifaceted kinase-dependent functions with a kinase-independent, RNA binding–dependent role that ensures adequate repair of common replication errors.

Cyclin A2 is a core cell cycle regulator that activates Cdk1 and Cdk2. Cyclin A2 levels increase upon S phase entry and remain high until its proteasome-dependent destruction in prometaphase (*1*). Cyclin A2–Cdk complexes that assemble at the onset of S phase drive chromosome duplication through phosphorylation of key DNA replication factors. Subsequently, cyclin A2–Cdk activity initiates mitosis by phosphorylating and inactivating the protein kinase Wee1, resulting in activation and nuclear localization of cyclin B1–Cdk1. During early mitosis, cyclin A2 is implicated in mitotic spindle anchoring (*2*) and correction of aberrant kinetochore-microtubule attachments (*3*). Furthermore, in a Cdk-independent manner, cyclin A2 regulates RhoA and RhoC, two guanosine triphosphatases implicated in cell morphogenesis, adhesion, and migration (*4*).

In mice, cyclin A2 is essential for early embryogenesis, limiting investigations of its biological functions (*5*). Studies of conditional knockout mice revealed that cyclin A2 is essential for cell cycle progression of certain cell types, including pluripotent and hematopoietic stem cells and select neuronal progenitors, yet is dispensable in fibroblasts because of redundancy with cyclin E in this cell type (*6*). To uncover novel biological processes that critically depend on a full complement of cyclin A2, we used a combination of knockout (*Ccna2*$^{-}$) and hypomorphic (*Ccna2*^{H}) alleles to markedly down-regulate cyclin A2 expression in mice without overtly affecting embryogenesis or postnatal development. *Ccna2*^{H} was created by targeted insertion of a neomycin resistance cassette into *Ccna2* intron 2; *Ccna2*$^{-}$ was created by gene trap mutagenesis (fig. S1, A to D). *Ccna2*$^{-/H}$ mice showed markedly reduced levels of cyclin A2 protein in tissues with a high mitotic index, including small intestine, bone marrow, and spleen (Fig. 1A). In contrast, cyclin A2 levels appeared normal in tissues with few cycling cells, such as brain, liver, and lung. However, actively cycling cultured lung epithelial cells from *Ccna2*$^{-/H}$ mice had lower amounts of cyclin A2 relative to the wild type (Fig. 1A). Analysis of mouse embryonic fibroblasts (MEFs) confirmed this, with *Ccna2*$^{-/H}$ MEFs expressing only ~25% of normal cyclin A2 levels (Fig. 1A and fig. S1E). In *Ccna2*$^{+/+}$ MEFs, cyclin A2 expression typically started to increase during G_1 and peaked from G_2 until prophase (fig. S2). Cyclin A2 levels remained low in *Ccna2*$^{-/H}$ MEFs throughout the cell cycle, resulting in markedly reduced Cdk activity during S and G_2 phase (fig. S3). Despite these abnormalities, *Ccna2*$^{-/H}$ MEFs showed a normal cell cycle profile (fig. S4).

Both cyclin A2 deficiency and overabundance have been observed in human tumors and predict poor clinical outcome (*7*), but whether and how cyclin A2 deregulation drives malignant growth is unknown. To determine whether reduced cyclin A2 expression contributes to tumorigenesis, we treated *Ccna2*$^{+/+}$ and *Ccna2*$^{-/H}$ mice with 7,12-dimethylbenz(a)anthracene (DMBA), a carcinogen that predisposes mice to lung adenomas and skin papillomas (*8*). *Ccna2*$^{-/H}$ mice showed a marked increase in tumor incidence and multiplicity in both lung and skin (Fig. 1B). Furthermore, lung adenomas of *Ccna2*$^{-/H}$ mice were larger in size. *Ccna2*$^{-/H}$ mice were also more susceptible to spontaneous tumors, particularly lung adenomas (Fig. 1C). Collectively, these data establish that cyclin A2 insufficiency promotes neoplastic transformation.

To identify the underlying defects, we screened for chromosomal instability, a hallmark of human malignancies. Splenocytes and MEFs of *Ccna2*$^{-/H}$ mice had increased aneuploidy (fig. S5A). *Ccna2*$^{-/H}$ MEFs showed predisposition to two types of chromosome segregation errors: chromatin bridges and lagging chromosomes (Fig. 2A). The latter are the result of merotelic attachment, a microtubule-kinetochore malattachment caused by spindle defects, including defects in attachment error correction, microtubule dynamics, mitotic timing, centrosome disjunction, and centrosome movement (*9*). We systematically screened *Ccna2*$^{-/H}$ MEFs for lagging chromosomes, and by measuring the separation between centrosomes in G_2 and prophase, we found that the movement of sister centrosomes to opposite poles was impaired (Fig. 2B and fig. S5, B to F). Cells with delayed centrosome separation form asymmetrical spindles and lagging chromosomes at increased rates (*10*). Consistent with this, spindle geometry defects occurred at high frequency in *Ccna2*$^{-/H}$ MEFs (Fig. 2C). Furthermore, lagging chromosomes were more prevalent in *Ccna2*$^{-/H}$ MEFs with asymmetrical as opposed to symmetrical spindles (fig. S5, G and H).

The motor protein Eg5 accumulates at centrosomes in prophase to drive centrosome movement (*11*). Loading of Eg5 was markedly reduced in *Ccna2*$^{-/H}$ MEFs, as was centrosomal accumulation of cyclin A2 (Fig. 2D and fig. S5I). Centrosome targeting of Eg5 is dependent on Cdk-mediated phosphorylation of Thr^{926} (T926) (*11*). T926 phosphorylation was reduced in *Ccna2*$^{-/H}$ MEFs despite normal expression of Eg5 and alternative Cdk1 and Cdk2 partners, including cyclins B and E (Fig. 2E and fig. S5J). Restoration of cyclin A2 expression in *Ccna2*$^{-/H}$ MEFs normalized T926 phosphorylation and corrected centrosomal loading of Eg5, centrosome movement, and spindle geometry (fig. S5, K to N). Collectively, these results uncover a nonredundant catalytic role of cyclin A2–Cdk in targeting Eg5 to centrosomal

[1]Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, MN, USA. [2]Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN, USA. [3]Department of Immunology, Mayo Clinic, Rochester, MN, USA.
***Corresponding author. Email: vandeursen.jan@mayo.edu**

High-resolution interrogation of functional elements in the noncoding genome

Neville E. Sanjana, Jason Wright, Kaijie Zheng, Ophir Shalem, Pierre Fontanillas, Julia Joung, Christine Cheng, Aviv Regev and Feng Zhang (September 29, 2016)

Science **353** (6307), 1545-1549. [doi: 10.1126/science.aaf7613]

Editor's Summary

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.