

genes. The top two thirds of our predictions exhibited threefold higher efficacy than that of the remaining fraction, confirming the accuracy of the algorithm.

Using this algorithm, we designed a whole-genome sgRNA library consisting of sequences predicted to have higher efficacy (table S8). As with the sgRNA pool used in our screens, this new collection was also filtered for potential off-target matches. This reference set of sgRNAs may be useful both for targeting single genes as well as large-scale sgRNA screening.

Taken together, these results demonstrate the utility of CRISPR-Cas9 for conducting large-scale genetic screens in mammalian cells. On the basis of our initial experiments, this system appears to offer several powerful features that together provide substantial advantages over current functional screening methods.

First, CRISPR-Cas9 inactivates genes at the DNA level, making it possible to study phenotypes that require a complete loss of gene function to be elicited. In addition, the system should also enable functional interrogation of nontranscribed elements, which are inaccessible by means of RNAi.

Second, a large proportion of sgRNAs successfully generate mutations at their target sites. Although this parameter is difficult to directly assess in pooled screens, we can obtain an estimate by examining the “hit rate” at known genes. Applying a *z* score analysis of our positive selection screens, we found that over 75% (46 of 60) of sgRNAs score at a significance threshold that perfectly separates true and false positives on a gene level (fig. S5, A to D). Together, these results show that the effective coverage of our library is very high and that the rate of false negatives should be low, even in a large-scale screen.

Third, off-target effects do not appear to seriously hamper our screens, according to several lines of evidence. Direct sequencing of potential off-target loci detected minimal cleavage at secondary sites, which typically reside in noncoding regions and do not affect gene function. Moreover, in the 6-TG screens the 20 most abundant sgRNAs all targeted one of the four members of the MMR pathway. In total, they represented over 30% of the final pool, which is a fraction greater than the next 400 sgRNAs combined. In the etoposide screen, the two top genes scored far above background levels (*P* values 100-fold smaller than that of the next best gene), enabling clear discrimination between true and false-positive hits. Last, new versions of the CRISPR-Cas9 system have recently been developed that substantially decrease off-target activity (30, 31).

Although we limited our investigation to proliferation-based phenotypes, our approach can be applied to a much wider range of biological phenomena. With appropriate sgRNA libraries, the method should enable genetic analyses of mammalian cells to be conducted with a degree of rigor and completeness currently possible only in the study of microorganisms.

References and Notes

- G. Giaever *et al.*, *Nature* **418**, 387–391 (2002).
- M. Costanzo *et al.*, *Science* **327**, 425–431 (2010).
- J. E. Carette *et al.*, *Science* **326**, 1231–1235 (2009).
- G. Guo, W. Wang, A. Bradley, *Nature* **429**, 891–895 (2004).
- A. Fire *et al.*, *Nature* **391**, 806–811 (1998).
- T. R. Brummelkamp, R. Bernards, R. Agami, *Science* **296**, 550–553 (2002).
- J. Moffat *et al.*, *Cell* **124**, 1283–1298 (2006).
- V. N. Ngo *et al.*, *Nature* **441**, 106–110 (2006).
- H. W. Cheung *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 12372–12377 (2011).
- C. J. Echeverri *et al.*, *Nat. Methods* **3**, 777–779 (2006).
- M. Booker *et al.*, *BMC Genomics* **12**, 50 (2011).
- W. G. Kaelin Jr., *Science* **337**, 421–422 (2012).
- R. Barrangou *et al.*, *Science* **315**, 1709–1712 (2007).
- L. Cong *et al.*, *Science* **339**, 819–823 (2013).
- P. Mali *et al.*, *Science* **339**, 823–826 (2013).
- W. Y. Hwang *et al.*, *Nat. Biotechnol.* **31**, 227–229 (2013).
- M. Jinek *et al.*, *Science* **337**, 816–821 (2012).
- T. Horii, D. Tamura, S. Morita, M. Kimura, I. Hatada, *Int. J. Mol. Sci.* **14**, 19774–19781 (2013).
- H. Wang *et al.*, *Cell* **153**, 910–918 (2013).
- P. D. Hsu *et al.*, *Nat. Biotechnol.* **31**, 827–832 (2013).
- Y. Fu *et al.*, *Nat. Biotechnol.* **31**, 822–826 (2013).
- V. Pattanayak *et al.*, *Nat. Biotechnol.* **31**, 839–843 (2013).
- T. Yan, S. E. Berry, A. B. Desai, T. J. Kinsella, *Clin. Cancer Res.* **9**, 2327–2334 (2003).
- R. D. Kolodner, G. T. Marsischky, *Curr. Opin. Genet. Dev.* **9**, 89–96 (1999).
- D. J. Burgess *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9053–9058 (2008).
- B. Scappini *et al.*, *Cancer* **100**, 1459–1471 (2004).
- S. Xue, M. Barna, *Nat. Rev. Mol. Cell Biol.* **13**, 355–369 (2012).
- C. M. Johnston *et al.*, *PLOS Genet.* **4**, e9 (2008).
- A. Subramanian *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
- F. A. Ran *et al.*, *Cell* **154**, 1380–1389 (2013).
- P. Mali *et al.*, *Nat. Biotechnol.* **31**, 833–838 (2013).
- K. Yoshimoto *et al.*, *Front. Oncol.* **2**, 186 (2012).

Acknowledgments: We thank all members of the Sabatini and Lander labs, especially J. Engreitz, S. Schwartz, A. Shishkin, and Z. Tsun for protocols, reagents, and advice; T. Mikkelsen for assistance with oligonucleotide synthesis; and L. Gaffney for assistance with figures. This work was supported by the U.S. National Institutes of Health (CA103866) (D.M.S.), National Human Genome Research Institute (2U54HG003067-10) (E.S.L.), the Broad Institute of MIT and Harvard (E.S.L.), and an award from the U.S. National Science Foundation (T.W.). The composition of the sgRNA pools and screening data can be found in the supplementary materials. A patent application has been filed by the Broad Institute relating to aspects of the work described in this manuscript. Inducible Cas9 and sgRNA backbone lentiviral vectors and the genome-scale sgRNA plasmid pool are deposited in Addgene.

Supplementary Materials

www.sciencemag.org/content/343/6166/80/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S5
Tables S1 to S8
References (33–43)

8 October 2013; accepted 2 December 2013

Published online 12 December 2013;
10.1126/science.1246981

Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells

Ophir Shalem,^{1,2*} Neville E. Sanjana,^{1,2*} Ella Hartenian,¹ Xi Shi,^{1,3} David A. Scott,^{1,2} Tarjei S. Mikkelsen,¹ Dirk Heckl,⁴ Benjamin L. Ebert,⁴ David E. Root,¹ John G. Doench,¹ Feng Zhang^{1,2†}

The simplicity of programming the CRISPR (clustered regularly interspaced short palindromic repeats)–associated nuclease Cas9 to modify specific genomic loci suggests a new way to interrogate gene function on a genome-wide scale. We show that lentiviral delivery of a genome-scale CRISPR-Cas9 knockout (GeCKO) library targeting 18,080 genes with 64,751 unique guide sequences enables both negative and positive selection screening in human cells. First, we used the GeCKO library to identify genes essential for cell viability in cancer and pluripotent stem cells. Next, in a melanoma model, we screened for genes whose loss is involved in resistance to vemurafenib, a therapeutic RAF inhibitor. Our highest-ranking candidates include previously validated genes *NF1* and *MED12*, as well as novel hits *NF2*, *CUL3*, *TADA2B*, and *TADA1*. We observe a high level of consistency between independent guide RNAs targeting the same gene and a high rate of hit confirmation, demonstrating the promise of genome-scale screening with Cas9.

A major goal since the completion of the Human Genome Project is the functional characterization of all annotated genetic elements in normal biological processes and dis-

¹Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA. ²McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. ³Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, MA 02142, USA. ⁴Division of Hematology, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA.

*These authors contributed equally to this work.

†Corresponding author. E-mail: zhang@broadinstitute.org

ease (1). Genome-scale loss-of-function screens have provided a wealth of information in diverse model systems (2–5). In mammalian cells, RNA interference (RNAi) is the predominant method for genome-wide loss-of-function screening (2, 3), but its utility is limited by the inherent incompleteness of protein depletion by RNAi and confounding off-target effects (6, 7).

The RNA-guided CRISPR (clustered regularly interspaced short palindromic repeats)–associated nuclease Cas9 provides an effective means of introducing targeted loss-of-function mutations at specific sites in the genome (8, 9). Cas9 can be programmed to induce DNA double-strand breaks

(DSBs) at specific genomic loci (8, 9) through a synthetic single-guide RNA (sgRNA) (10), which when targeted to coding regions of genes can create frame shift insertion/deletion (indel) mutations that result in a loss-of-function allele. Because the targeting specificity of Cas9 is conferred by short guide sequences, which can be easily generated at large scale by array-based oligonucleotide library synthesis (11), we sought to explore the potential of Cas9 for pooled genome-scale functional screening.

Lentiviral vectors are commonly used for delivery of pooled short-hairpin RNAs (shRNAs) in RNAi because they can be easily titrated to control transgene copy number and are stably maintained as genomic integrants during subsequent cell replication (2, 12, 13). Therefore, we designed a single lentiviral vector to deliver Cas9, a sgRNA, and a puromycin selection marker into target cells (lentiCRISPR) (Fig. 1A). The ability to simultaneously deliver Cas9 and sgRNA through

a single vector enables application to any cell type of interest, without the need to first generate cell lines that express Cas9.

To determine the efficacy of gene knockout by lentiCRISPR transduction, we tested six sgRNAs targeting enhanced green fluorescent protein (EGFP) in a human embryonic kidney (HEK) 293T cell line containing a single copy of EGFP (fig. S1). After transduction at a low multiplicity of infection (MOI = 0.3) followed by selection with puromycin, lentiCRISPRs abolished EGFP fluorescence in $93 \pm 8\%$ (mean \pm SD) of cells after 11 days (Fig. 1B). Deep sequencing of the EGFP locus revealed a $92 \pm 9\%$ indel frequency ($n \geq 10^4$ sequencing reads per condition) (fig. S2). In contrast, transduction of cells with lentiviral vectors expressing EGFP-targeting shRNA led to incomplete knockdown of EGFP fluorescence (Fig. 1C).

Given the high efficacy of gene knockout by lentiCRISPR, we tested the feasibility of conducting genome-scale CRISPR-Cas9 knockout (GeCKO) screening with a pooled lentiCRISPR library. We designed a library of sgRNAs targeting 5' constitutive exons (Fig. 2A) of 18,080 genes in the human genome with an average coverage of 3 to 4 sgRNAs per gene (table S1), and each target site was selected to minimize off-target modification (14) (see supplementary text).

To test the efficacy of the full GeCKO library at achieving knockout of endogenous gene targets, we conducted a negative selection screen by profiling the depletion of sgRNAs targeting essential survival genes (Fig. 2A). We transduced the human melanoma cell line A375 and the human

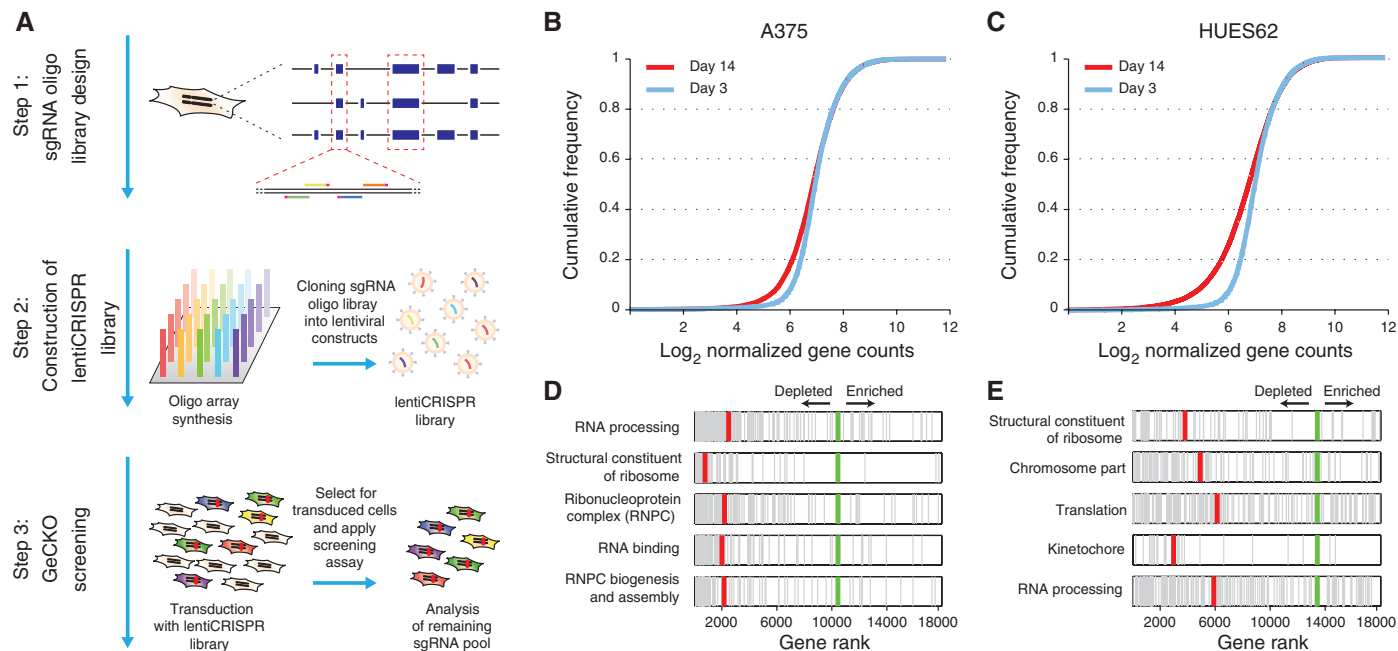
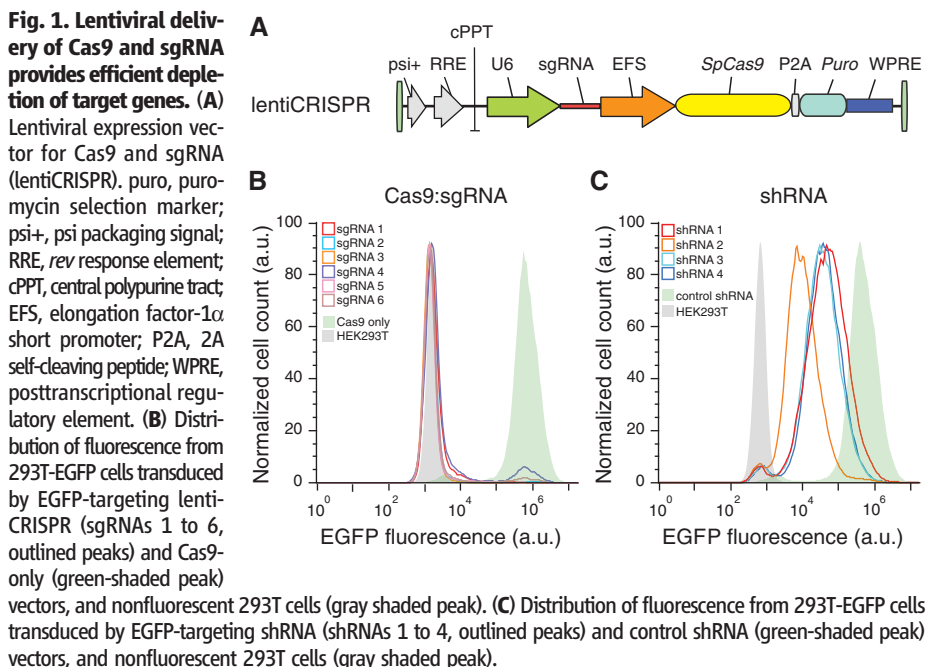


Fig. 2. GeCKO library design and application for genome-scale negative selection screening. (A) Design of sgRNA library for genome-scale knockout of coding sequences in human cells (see supplementary text). (B and C) Cumulative frequency of sgRNAs 3 and 14 days after transduction in A375 and human em-

brionic stem cells, respectively. Shift in the 14-day curve represents the depletion in a subset of sgRNAs. (D and E) The five most significantly depleted gene sets in A375 cells [nominal $P < 10^{-5}$, false discovery rate (FDR)-corrected $q < 10^{-5}$] and HUES62 cells (nominal $P < 10^{-5}$, FDR-corrected $q < 10^{-3}$) identified by GSEA (15).

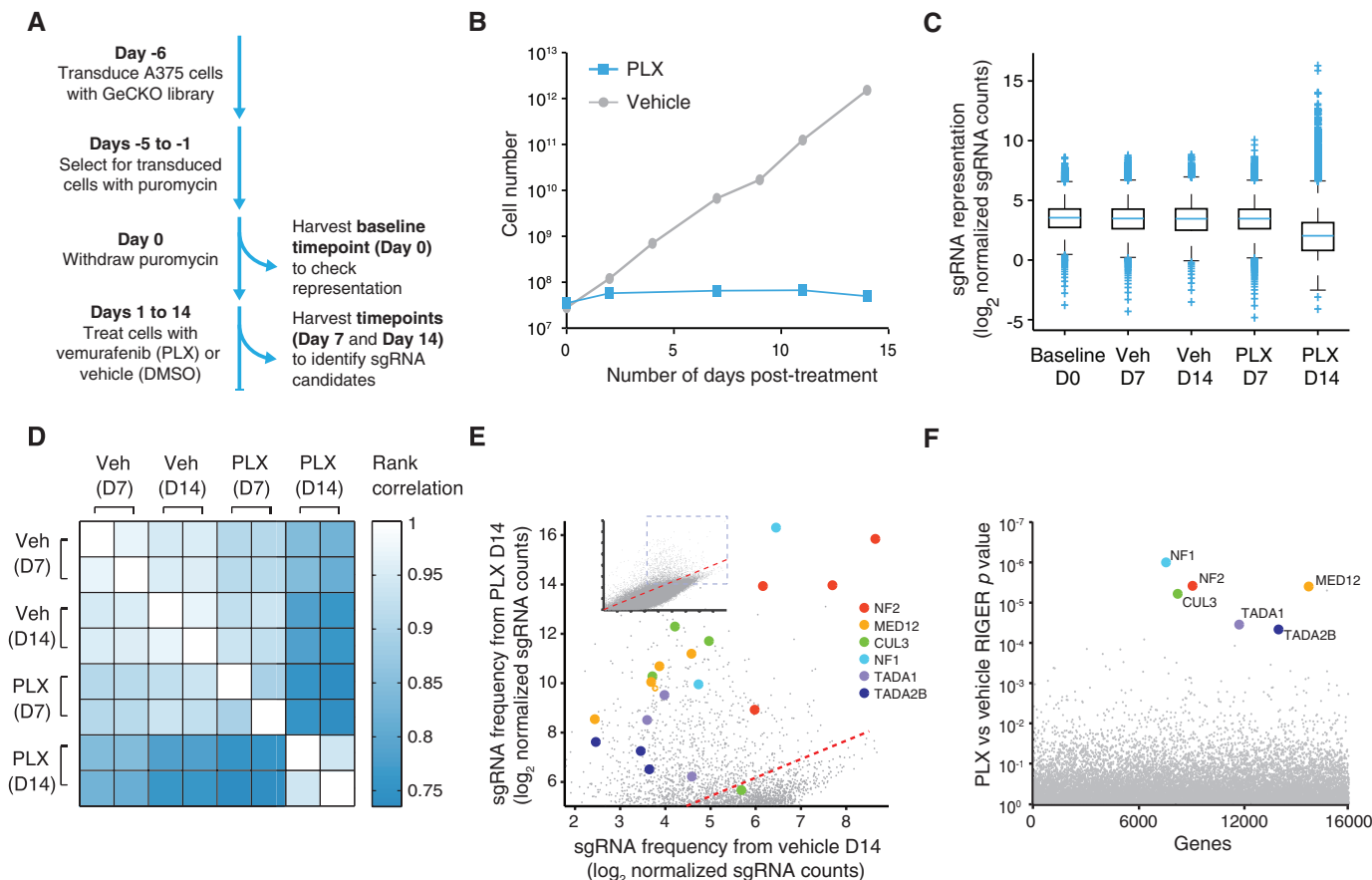


Fig. 3. GeCKO screen in A375 melanoma cells reveals genes whose loss confers PLX resistance. (A) Timeline of PLX resistance screen in A375 melanoma cells. (B) Growth of A375 cells when treated with dimethyl sulfoxide (DMSO) or PLX over 14 days. (C) Box plot showing the distribution of sgRNA frequencies at different time points, with and without PLX treatment (vehicle is DMSO). The box extends from the first to the third quartile with the whiskers denoting 1.5 times the interquartile range. Enrichment of

specific sgRNAs: 7 days of PLX treatment, 1 sgRNA greater than 10-fold enrichment; 14 days of PLX treatment, 379 and 49 sgRNAs greater than 10-fold and 100-fold enrichment, respectively. (D) Rank correlation of normalized sgRNA read count between biological replicates and treatment conditions. (E) Scatterplot showing enrichment of specific sgRNAs after PLX treatment. (F) Identification of top candidate genes using the RIGER P value analysis.

stem cell line HUES62 with the GeCKO library at a MOI of 0.3. As expected, deep sequencing (figs. S3 and S4) 14 days after transduction revealed a significant reduction in the diversity of sgRNAs in the surviving A375 and HUES62 cells (Fig. 2, B and C) (Wilcoxon rank sum test, $P < 10^{-10}$ for both cell types). Gene set enrichment analysis (GSEA) (15) indicated that most of the depleted sgRNAs targeted essential genes such as ribosomal structural constituents (Fig. 2, D and E, and tables S2 and S3). The overlap in highly depleted genes and functional gene categories between the two cell lines (fig. S5) indicates that GeCKO can identify essential genes and that enrichment analysis of depleted sgRNAs can pinpoint key functional gene categories in negative selection screens.

To test the efficacy of GeCKO for positive selection, we sought to identify gene knockouts that result in resistance to the BRAF protein kinase inhibitor vemurafenib (PLX) in melanoma (16) (Fig. 3A). Exposure to PLX resulted in growth arrest of transduced A375 cells, which harbor the V600E gain-of-function BRAF mutation (17) (Fig. 3B), therefore enabling the enrichment of a

small group of cells that were rendered drug-resistant by Cas9:sgRNA-mediated modification. After 14 days of PLX treatment, the sgRNA distribution was significantly different when compared with vehicle-treated cells (Fig. 3C) (Wilcoxon rank-sum test, $P < 10^{-10}$) and clustered separately from all other conditions (Fig. 3D and fig. S6).

For a subset of genes, we found enrichment of multiple sgRNAs that target each gene after 14 days of PLX treatment (Fig. 3E), suggesting that loss of these particular genes contributes to PLX resistance. We used the RNAi Gene Enrichment Ranking (RIGER) algorithm to rank screening hits by the consistent enrichment among multiple sgRNAs targeting the same gene (Fig. 3F and table S4) (12). Our highest-ranking genes included previously reported candidates *NF1* and *MED12* (18, 19) and also several genes not previously implicated in PLX resistance, including neurofibromin 2 (*NF2*), Cullin 3 E3 ligase (*CUL3*), and members of the STAGA histone acetyltransferase complex (*TADA1* and *TADA2B*). These candidates yield new testable hypotheses regarding PLX resistance mechanisms (see supplementary

text). For example, *NF1* and *NF2*, although unrelated in sequence, are each mutated in similar but distinct forms of neurofibromatosis (20). In addition, epigenetic dysregulation resulting from mutations in the mechanistically related STAGA and Mediator complexes (21) may have a role in acquired drug resistance. All of these hits were also identified through a second independent library transduction (figs. S7 and S8 and tables S5 and S6).

A similar screen to identify PLX drug resistance in A375 cells was previously conducted using a pooled library of 90,000 shRNAs (19). To compare the efficacy and reliability of genome-scale shRNA screening with GeCKO, we used several methods to evaluate the degree of consistency among the sgRNAs or shRNAs targeting the top candidate genes. First, we plotted the P values for the top 100 hits using either RIGER (Fig. 4A) or redundant siRNA activity (RSA) (fig. S9) scoring. Lower P values for the GeCKO versus shRNA screen indicate better scoring consistency among sgRNAs. Second, for the top 10 RIGER hit genes, $78 \pm 27\%$ of sgRNAs targeting each gene ranked among the top 5% of enriched sgRNAs, whereas

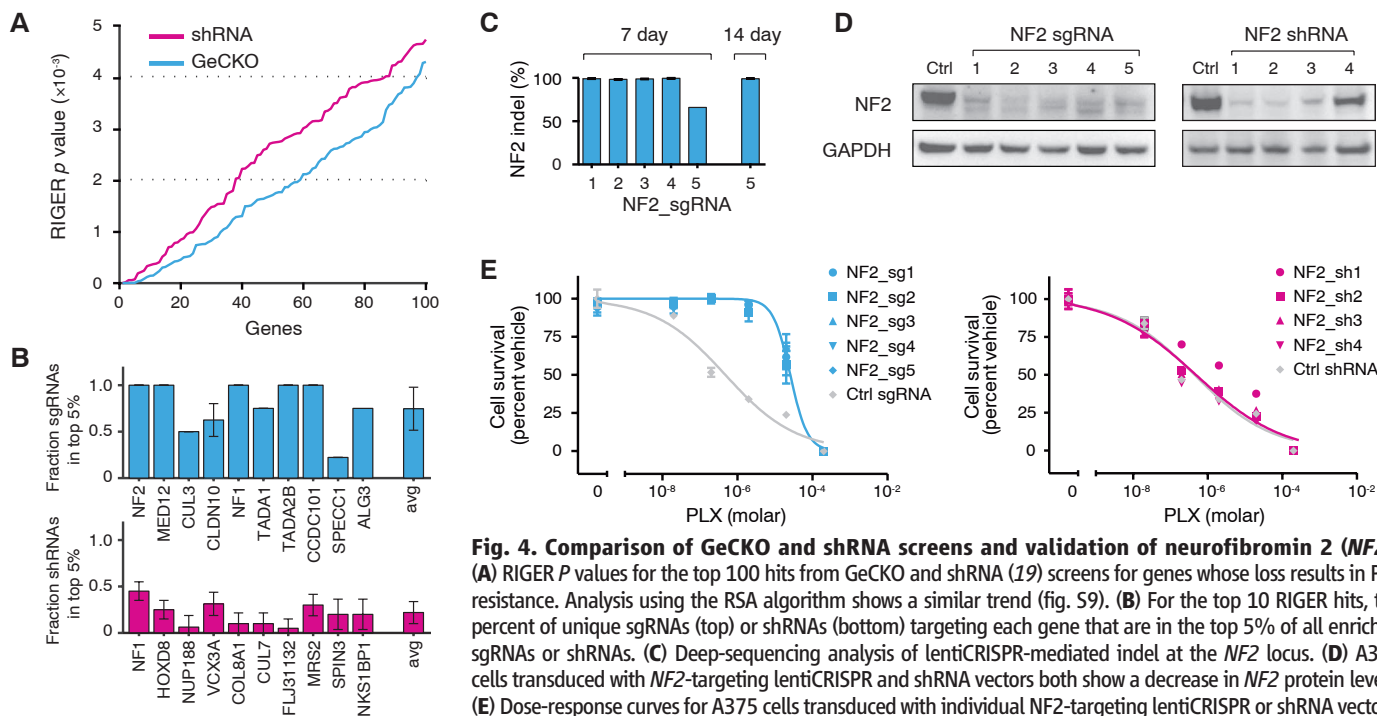


Fig. 4. Comparison of GeCKO and shRNA screens and validation of neurofibromin 2 (NF2). (A) RIGER *P* values for the top 100 hits from GeCKO and shRNA (19) screens for genes whose loss results in PLX resistance. Analysis using the RSA algorithm shows a similar trend (fig. S9). (B) For the top 10 RIGER hits, the percent of unique sgRNAs (top) or shRNAs (bottom) targeting each gene that are in the top 5% of all enriched sgRNAs or shRNAs. (C) Deep-sequencing analysis of lentiCRISPR-mediated indel at the *NF2* locus. (D) A375 cells transduced with *NF2*-targeting lentiCRISPR and shRNA vectors both show a decrease in *NF2* protein levels. (E) Dose-response curves for A375 cells transduced with individual *NF2*-targeting lentiCRISPR or shRNA vectors. Controls were EGFP-targeting lentiCRISPR or null-hairpin shRNA vectors.

lentiCRISPRs show a significant increase ($F_{1,8} = 30.3$, $P < 0.001$, $n = 4$ replicates) in the half-maximal effective concentration (EC_{50}), whereas cells transduced with *NF2*-targeting shRNA vectors do not ($F_{1,8} = 0.47$, $P = 0.51$, $n = 4$ replicates).

20 ± 12% of shRNAs targeting each gene ranked among the top 5% of enriched shRNAs (Fig. 4B).

We validated top-ranking genes from the GeCKO screen individually using 3 to 5 sgRNAs (Fig. 4, C to E, and figs. S10 and S11). For *NF2*, we found that 4 out of 5 sgRNAs resulted in >98% allele modification 7 days after transduction, and all 5 sgRNAs showed >99% allele modification 14 days after transduction (Fig. 4C). We compared sgRNA and shRNA-mediated protein depletion and PLX resistance using Western blot (Fig. 4D) and cell growth assays (Fig. 4E). Interestingly, although all five sgRNAs conferred resistance to PLX, only the best shRNA achieved sufficient knockdown to increase PLX resistance (Fig. 4E), suggesting that even low levels of *NF2* are sufficient to retain sensitivity to PLX. Additionally, sgRNAs targeting *NF1*, *MED12*, *CUL3*, *TADA1*, and *TADA2B* led to a decrease in protein expression and increased resistance to PLX (figs. S10 and S11). Deep sequencing confirmed a high rate of mutagenesis at targeted loci (figs. S12 and S13), with a small subset of off-target sites exhibiting indels (figs. S14 to S16), which may be alleviated using an offset nicking approach (22, 23) that was recently shown to reduce off-target modifications (22).

GeCKO screening provides a mechanistically distinct method from RNAi for systematic perturbation of gene function. Whereas RNAi reduces protein expression by targeting RNA, GeCKO introduces loss-of-function mutations into genomic DNA. Although some indel mutations are expected to maintain the reading frame, homozygous knockout yields high screening sensitiv-

ity, which is especially important in cases where incomplete knockout retains gene function. In addition, RNAi is limited to transcripts, whereas Cas9:sgRNAs can target elements across the entire genome, including promoters, enhancers, introns, and intergenic regions. Furthermore, catalytically inactive mutants of Cas9 can be tethered to different functional domains (23–27) to broaden the repertoire of perturbation modalities, including genome-scale gain-of-function screening using Cas9 activators and epigenetic modifiers. In the GeCKO screens presented here, the efficiency of complete knockout, the consistency of distinct sgRNAs, and the high validation rate for top screen hits demonstrate the potential of Cas9:sgRNA-based technology to transform functional genomics.

References and Notes

1. E. S. Lander, *Nature* **470**, 187–197 (2011).
2. K. Berns *et al.*, *Nature* **428**, 431–437 (2004).
3. M. Boutros *et al.*, Heidelberg Fly Array Consortium, *Science* **303**, 832–835 (2004).
4. R. Rad *et al.*, *Science* **330**, 1104–1107 (2010).
5. J. E. Carette *et al.*, *Science* **326**, 1231–1235 (2009).
6. A. L. Jackson *et al.*, *RNA* **12**, 1179–1187 (2006).
7. C. J. Echeverri *et al.*, *Nat. Methods* **3**, 777–779 (2006).
8. L. Cong *et al.*, *Science* **339**, 819–823 (2013).
9. P. Mali *et al.*, *Science* **339**, 823–826 (2013).
10. M. Jinek *et al.*, *Science* **337**, 816–821 (2012).
11. A. P. Blanchard, L. Hood, *Nat. Biotechnol.* **14**, 1649 (1996).
12. B. Luo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20380–20385 (2008).
13. P. J. Paddison *et al.*, *Nature* **428**, 427–431 (2004).
14. P. D. Hsu *et al.*, *Nat. Biotechnol.* **31**, 827–832 (2013).
15. A. Subramanian *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
16. K. T. Flaherty *et al.*, *N. Engl. J. Med.* **363**, 809–819 (2010).

17. H. Davies *et al.*, *Nature* **417**, 949–954 (2002).
18. S. Huang *et al.*, *Cell* **151**, 937–950 (2012).
19. S. R. Whittaker *et al.*, *Cancer Discovery* **3**, 350–362 (2013).
20. A. L. Lin, D. H. Gutmann, *Nat. Rev. Clin. Oncol.* **10**, 616–624 (2013).
21. X. Liu, M. Vorontchikhina, Y. L. Wang, F. Faiola E. Martinez, *Mol. Cell. Biol.* **28**, 108–121 (2008).
22. F. A. Ran *et al.*, *Cell* **154**, 1380–1389 (2013).
23. P. Mali *et al.*, *Nat. Biotechnol.* **31**, 833–838 (2013).
24. L. A. Gilbert *et al.*, *Cell* **154**, 442–451 (2013).
25. S. Konermann *et al.*, *Nature* **500**, 472–476 (2013).
26. P. Perez-Pinera *et al.*, *Nat. Methods* **10**, 973–976 (2013).
27. M. L. Maeder *et al.*, *Nat. Methods* **10**, 977–979 (2013).

Acknowledgments: We thank G. Cowley, W. Harrington, J. Wright, E. Hodis, S. Whittaker, J. Merkin, C. Burge, D. Peters, C. Cowan, L. P. Club, and the entire Zhang laboratory for technical support and critical discussions. O.S. is a Klarman Family Foundation Fellow, N.S. is a Simons Center for the Social Brain Postdoctoral Fellow, D.A.S. is an NSF Fellow, and J.D. is a Merkin Institute Fellow. D. H. was funded by the German Cancer Foundation. F.Z. is supported by an NIH Director’s Pioneer Award (1DP1-MH100706); a NIH Transformative R01 grant (1R01-DK097768); the Keck, McKnight, Merkin, Vallee, Damon Runyon, Searle Scholars, Klingenstein, and Simons Foundations; and Bob Metcalfe and Jane Pauley. The authors have no conflicting financial interests. A patent application has been filed relating to this work, and the authors plan to make the reagents widely available to the academic community through Addgene and to provide software tools at the Zhang laboratory Web site (www.genome-engineering.org).

Supplementary Materials

www.sciencemag.org/content/343/6166/84/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S16
Tables S1 to S10
References

9 October 2013; accepted 2 December 2013
Published online 12 December 2013;
10.1126/science.1247005