Systematic comparison of CRISPR/Cas9 and RNAi screens for essential genes

David W Morgens¹, Richard M Deans¹,², Amy Li¹ & Michael C Bassik¹,³

We compared the ability of short hairpin RNA (shRNA) and CRISPR/Cas9 screens to identify essential genes in the human chronic myelogenous leukemia cell line K562. We found that the precision of the two libraries in detecting essential genes was similar and that combining data from both screens improved performance. Notably, results from the two screens showed little correlation, which can be partially explained by the identification of distinct essential biological processes with each technology.

Efficient gene knockdown and knockout using RNA interference (RNAi)¹–³ and CRISPR/Cas9 systems⁴–⁸ allow for systematic evaluation of gene function, but it is unclear how the choice of technology may affect results. For example, heterogeneity of reagents has historically been associated with poor performance in RNAi-based screens⁹,¹⁰, and it may also influence CRISPR/Cas9 deletion screens⁴–¹⁰ (Supplementary Figs. 1–3). Whereas the variability of shRNAs in RNAi screens stems from differences in knockdown efficiency¹⁰, variability of single guide RNAs (sgRNAs) in CRISPR/Cas9 screens probably stems from the array of genotypes (true knockouts, heterozygotes, and wild-type cells) created¹⁰. Notably, this depends on the efficiency of guide cutting as well as the relative fitness between these subpopulations. Other possible differences include interference by nonspecific effects such as mRNA deregulation during RNAi¹¹,¹² and intrinsic differences between knockouts and knockdowns. These concerns, as well as others, necessitate a careful comparison between these techniques.

To directly compare the phenotypes obtained using CRISPR/Cas9 and shRNA-based screening technologies, we performed parallel screens in duplicate for genes affecting growth rates in K562 cells using both a 25-hairpin-per-gene shRNA library¹³ and a four-sgRNA-per-gene CRISPR/Cas9 library¹⁴. Briefly, sgRNA and shRNA libraries were lentivirally infected into cells, replicate populations were split at time zero (T0), and the composition of those populations was investigated after 2 weeks of unperturbed growth by comparison to the starting plasmid library (Fig. 1a). The screens were conducted in parallel to minimize technical variation, allowing a quantitative assessment of performance. A previously established gold standard of 217 genes expected to have growth phenotypes in all cell types (essential) and 947 genes (Supplementary Data 1) expected to have growth phenotypes in no cell type (nonessential)¹⁵ was used to estimate true positive and false positive rates.

Using the median enrichment averaged over two replicates, we found that both shRNA and CRISPR/Cas9 screens had very high performance in the detection of essential genes (area under the curve of the receiver operating characteristic (ROC) curve > 0.90) (Fig. 1b and Supplementary Data 2–4). At an ~1% false positive rate, both screens recovered >60% of gold-standard essential genes. However, at a 10% false positive rate there were ~4,500 genes identified in the Cas9 screen versus ~3,100 in the shRNA screen, with ~1,200 genes identified in both (Fig. 1c). This indicates that although our shRNA and Cas9 screens achieved similar levels of precision on the gold standard, the Cas9 and shRNA screens both identified numerous genes that were not in the gold standard or identified by the other screen.

To leverage data from both screening technologies, we developed a statistical framework, Cas9 high-Throughput maximum Likelihood Estimator (casTLE). For each gene, casTLE combines measurements from multiple targeting reagents to estimate a maximum effect size as well as a P value associated with that effect (Supplementary Fig. 4; also see Supplementary Methods). We validated casTLE by analyzing previous RNAi¹, CRISPR deletion¹⁶, and CRISPRi/a¹⁷ screens and found consistent results (Supplementary Fig. 5 and Supplementary Data 5–7). casTLE’s performance in the identification of essential genes was favorable compared to that of previous methods¹¹,¹⁸–²¹, including the median effect used here (Supplementary Fig. 6; also see Supplementary Discussion). Although casTLE performs well on single replicates from many screen types, it can also combine results from diverse data types by separately considering (a) experimental noise and (b) variability caused by heterogeneous reagents.

Using casTLE to combine data from a single replicate of the shRNA and Cas9 screens led to a noticeable improvement in performance, with an area under the curve of 0.98, >85% of gold-standard essential genes identified at an ~1% false positive rate (FPR) (Fig. 1b, Supplementary Figs. 7a and 8a,c, and Supplementary Data 8), and the identification of ~4,500 genes with negative growth phenotypes with evidence from the combination of both screens (Fig. 1c and Supplementary Fig. 8b). To test whether these results depended on the number of targeting elements used, we compared the Cas9 results to a down-sampled four-hairpin shRNA screen, which yielded similar results (Supplementary Fig. 9a,b). The fact that the combination of the two technologies more successfully separated essential and nonessential genes suggests that the screens may be revealing different aspects of biology.

Consistent with the presence of nonredundant information, results from the Cas9 and shRNA screens show low correlation (Fig. 2a). Nonetheless, both screens effectively separated essential
and nonessential genes (Figs. 1b and 2b) and were highly reproducible (Supplementary Figs. 7b and 10a,b). When we compared the enrichment of Gene Ontology (GO) terms for essential complexes, we found that the screens identified different biological processes (Fig. 2c and Supplementary Data 9). For example, genes involved in the respiratory chain were enriched for essential genes in the Cas9 results, whereas all subunits of the chaperonin-containing T-complex were identified as essential by the shRNA screen (Fig. 2d). By using casTLE to combine information from each screen, we were able to recover each of these biological terms (Fig. 2c), further demonstrating the utility of a parallel screening approach. Again, these results did not change when we used a down-sampled four-hairpin library (Supplementary Fig. 9c,d).

Several technical factors might account for the observed lack of correlation and differential GO term enrichment: (1) the presence or absence of reagents effective toward a particular gene, (2) differences in off-target effects as evidenced by the distribution of phenotypes among nontargeting controls (Supplementary Fig. 1a), or (3) differences in the timing of deletion and knockdown (Supplementary Fig. 10c,d). Another possibility highlighted in a recent analysis of human essential genes is that RNAi is less able to perturb genes expressed at low levels. However, we found no clear signature of this in our data (Supplementary Fig. 11a), perhaps because we used a more effective hairpin design capable of targeting these genes. This increased efficacy in our shRNA library may explain the qualitatively higher performance in the detection of essential genes observed here compared to previous studies, which found that RNAi screens perform poorly in the identification of essential genes.

Although these technical differences may account for the low overall correlation observed, it is difficult to explain differential GO term enrichment with technical reasons alone. This observation suggests that Cas9 and shRNA screens are able to detect distinct aspects of biology, although it remains unclear why particular gene sets have strong signatures with one technology but not the other. One possibility, with respect to complexes such as RNA polymerase and the Mediator complex, is that the dependency of shRNA knockdown on ongoing transcription, whereas sgRNAs no longer need to be efficiently expressed after a gene is knocked out. Another possibility is that for certain genes a small loss in gene product via knockdown leads to a completely different phenotype than a large loss via knockout. This could reflect nonmonotonic gene-dose dependence on fitness. Alternatively, deletion of adjacent genes by Cas9 may occur in the context of tandem duplications, as found in a previous study. A final possibility is that distinct hits in Cas9 and shRNA screens may represent genes that interact with the nonspecific effects of RNAi and Cas9.

Figure 1 Parallel shRNA and CRISPR/Cas9 depletion screens used to identify essential genes in K562 cells. (a) Schematic of screen. shRNA and Cas9 libraries were lentivirally integrated into K562 cells and selected via puromycin treatment (time zero). After this, replicate cell populations were maintained in log-phase growth for 14 d. Library representation at each time point was monitored by deep-sequencing of the inserted locus. (b) ROC curves indicating screen performance in identifying essential genes by comparing the library composition between the plasmid library and cells after 2 weeks of growth. True positive rates and false positive rates were calculated using a previously established gold-standard set of essential and nonessential genes. Shown are ROC curves for Cas9 and shRNA screens based on the median score averaged over two replicates and for data from single replicates of both Cas9 and shRNA screens combined using casTLE. The dashed line represents the 10% FPR threshold used. AUC, area under the curve. (c) The number of essential genes at 10% false positive rate and their overlap based on the average median data from Cas9 and shRNA screens, as well as the combination of a single replicate from both screens using casTLE. The false positive rate was estimated using gold-standard nonessential genes.

Figure 2 Differences between Cas9 and shRNA results. (a) Comparison of casTLE scores for single replicates of Cas9 and shRNA data. A large positive casTLE score indicates a high-confidence increase in growth rate, whereas a large negative casTLE score indicates a high-confidence decrease in growth rate (i.e., essential). Density is in log scale. (b) casTLE scores for gold-standard essential and nonessential genes. (c) Adjusted P-values (calculated using GOilara, which uses a modified hypergeometric test) for select GO terms for shRNA and Cas9 screens as well as for data from both screens combined with casTLE. (d) casTLE scores shown for genes involved in the respiratory chain complex (GO:0098803) and the chaperonin-containing T-complex (GO:0005832), which exhibited differential enrichment in Cas9 and shRNA screens.
certain genes would be expected to exhibit an effect on growth in the presence of persistent DNA damage, or interference with miRNA processing, due to Cas9 nuclease activity or shRNA expression, respectively. The potential existence of both false positives and false negatives in each screen presents an analytical problem that could be addressed by combination analysis using casTLE.

In fact, we found that genes found uniquely in either the shRNA or the Cas9 screen—but not found in the combination analysis—did not have key signatures of essential genes. Essential genes are more likely to be highly expressed, and this was clearly seen in both screens (Supplementary Fig. 11a). However, when we limited our analysis to genes found in either the shRNA or the Cas9 screen but not in the combination analysis, this pattern was no longer clear compared to observations for the genes found in both the shRNA and the Cas9 screens (Supplementary Fig. 11b). Similarly, hits from the shRNA and Cas9 screens showed the expected enrichment for homologs of essential yeast genes, and the Cas9 screens showed high precision (Supplementary Fig. 11d). Although there are other explanations for these observations, this suggests that the combination analysis with casTLE can limit technology- or screen-specific false positives. Given that the use of both technologies also seemed to reduce technology-specific false negatives, as evidenced by the more complete capture of GO terms (Fig. 2c), the use of both (1) multiple reagents per gene to control for sequence-specific off-target effects and (2) multiple technologies for perturbation to control for nonspecific effects should provide a more robust determination of a gene's phenotype. However, determining the precise sources of differences between shRNA and CRISPR/Cas9 screens will require further inquiry.

Heterogeneity is a known feature of shRNA libraries, but we have also found this phenomenon in CRISPR/Cas9-based screens (Supplementary Figs. 1 and 2). One likely source of this variability is the occurrence of in-frame indels, creating wild-type and heterozygous subpopulations, which has been observed previously to interfere with screening results. More efficacious guide designs might help, but in some cases sgRNAs targeting conserved functional elements within genes, as opposed to simply the 5' coding region, may be preferred. Nonetheless, heterogeneity in library elements can be a useful feature for genes with a complex gene–dosage/phenotype relationship. In particular, the spectrum of expression levels generated with an shRNA library should allow the identification of non-growth phenotypes in situations where the complete deletion of a gene causes severe growth effects or lethality. Indeed, in our previous work we found that an shRNA screen for drug toxicity identified a presumably obligate biosynthetic gene, DHODH, as the drug's target, which a parallel Cas9 screen failed to identify. This suggests that differences between CRISPR/Cas9 and shRNA technologies may be useful to fully capture relevant biology in non-growth-based screens.

Here we present an experimental side-by-side comparison of CRISPR/Cas9 and RNAi screens for essential genes. Using median enrichment analysis and our casTLE analysis tool (Supplementary Figs. 2–4), we demonstrate that both a recently developed shRNA library and a Cas9 sgRNA library showed high precision (Fig. 1b and Supplementary Fig. 8a,c), but that the Cas9 library identified many more essential genes (Fig. 1c and Supplementary Fig. 8b). The two screening technologies identified different biological categories of genes (Fig. 2a,c). These differences could be exploited to obtain a more complete picture of genes regulating growth by combining information using casTLE (Figs. 1b and 2c and Supplementary Fig. 8a,c) and may be important when researchers are considering which technology to use to probe a given process.

**METHODS**

Methods and any associated references are available in the online version of the paper.

**Accession codes.** Sequence Read Archive: SRP072806. BioProject: PRJNA317269.

**Note:** Any Supplementary Information and Source Data files are available in the online version of the paper.

**ACKNOWLEDGMENTS**

We thank K. Han, G. Hess, M. Duhreul, M. Haney, K. Tsai, and the Bassik lab for technical expertise and helpful discussions. We thank G. Sherlock, N. Boyle, A. Kundaje, R. Tibshirani, and A. Fire for their critical reading of the manuscript and suggestions. Many thanks to A. Regev, N. Hacohen, O. Parnas, and R.H. Herbst for providing data from their published screen. We acknowledge the ENCODE Consortium and the Gingeras lab for the generation and release of the K562 gene expression data used here. This work was funded by the NIH Director’s New Innovator Award Program (project no. DP2HD08469-01) NIH/NHGRI (training grant T32 HG000404 to D.W.M.), and a seed grant from Stanford ChEM-H. This material is based on work supported by the National Science Foundation Graduate Research Fellowship (grant DGE-114747 to D.W.M. and R.M.D.). R.M.D. was additionally supported by a Burt and DeeDee McClurty Stanford Graduate Fellowship.

**AUTHOR CONTRIBUTIONS**

D.W.M., R.M.D., and M.C.B. conceived and designed the study. R.M.D. and A.L. performed screens. D.W.M., R.M.D., and A.L. processed and sequenced screens. D.W.M. designed and wrote casTLE. D.W.M. performed all analysis. D.W.M., R.M.D., and M.C.B. wrote the manuscript. All authors reviewed and approved the manuscript.

**COMPETING FINANCIAL INTERESTS**

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Bassik, M.C. et al. Cell 152, 909–922 (2013).
2. Silva, J.M. et al. Science 319, 617–620 (2008).
3. Barbie, D.A. et al. Nature 462, 108–112 (2009).
4. Shalem, O., Sanjana, N.E. & Zhang, F. Nat. Rev. Genet. 16, 299–311 (2015).
5. Wang, T., Wei, J.J., Sabatini, D.M. & Lander, E.S. Science 343, 80–84 (2014).
6. Shalem, O. et al. Science 343, 84–87 (2014).
7. Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M.D.C. & Yusa, K. Nat. Biotechnol. 32, 267–273 (2014).
8. Zhou, Y. et al. Nature 509, 487–491 (2014).
9. Kaelin, W.G. Science 337, 421–422 (2012).
10. Barrangou, R. et al. Nucleic Acids Res. 43, 3407–3419 (2015).
11. Jackson, A.L. & Linsley, P.S. Nat. Rev. Drug Discov. 9, 57–67 (2010).
12. Grimm, D. et al. Nature 441, 537–541 (2006).
13. Kampmann, M. et al. Proc. Natl. Acad. Sci. USA 112, E3384–E3391 (2015).
14. Dears, R.M. et al. Nat. Chem. Biol. http://dx.doi.org/10.1038/nchembio.2050 (2016).
15. Hart, T., Brown, K.R., Sircoulomb, F., Rottapel, R. & Moffat, J. Mol. Syst. Biol. 10, 733 (2014).
16. Parnas, O. et al. Cell 162, 675–686 (2015).
17. Gilbert, L.A. et al. Cell 159, 647–661 (2014).
18. Li, W. et al. Genome Biol. 15, 554 (2014).
19. König, R. et al. Nat. Methods 4, 847–849 (2007).
20. Luo, B. et al. Proc. Natl. Acad. Sci. USA 105, 20380–20385 (2008).
21. Díaz, A.A., Qin, H., Ramalho-Santos, M. & Song, J.S. Nucleic Acids Res. 43, e16 (2015).
22. Hart, T. et al. Cell 163, 1515–1526 (2015).
23. Wang, T. et al. Science 350, 1096–1101 (2015).
24. Tsygankov, P. et al. Elife 4, e08467 (2015).
25. Frock, R.L. et al. Nat. Biotechnol. 33, 179–186 (2015).
26. Tsai, S.-Q. et al. Nat. Biotechnol. 33, 187–197 (2015).
27. Prewitt-Miller, S.M., Reading, D.W., Porter, S.N. & Porteus, M.H. PLoS Genet. 5, e1000376 (2009).
28. Shi, J. et al. Nat. Biotechnol. 33, 661–667 (2015).
ONLINE METHODS

Cell culture. Cell culture was carried out as previously described. Briefly, K562 cells (ATCC) were cultured in RPMI 1640 (Gibco) media and supplemented with 10% FBS (Hyclone), penicillin (10,000 I.U./mL), streptomycin (10,000 µg/mL), and 1-glutamine (2 mM). We grew cells in log phase during all biological assays by returning the population to 500,000 cells per mL each day. K562 cells were maintained in a controlled humidified incubator at 37 °C, with 5% CO₂.

Genome-wide shRNA screen. A previously designed 25-shRNA-per-gene RNAi library was used. Library infections and shRNA prep and sequencing were performed as previously described. Briefly, to generate sufficient lentivirus to infect the genome-wide shRNA library into K562 cells, we plated 293T cells on 15-cm tissue culture plates. 293T cells were transfected with third-generation packaging plasmids and shRNA-encoding vectors. After 48 h and 72 h of incubation, lentivirus was harvested. We pooled the filtered lentivirus through a 0.45-µm PVDF filter (Millipore) to remove any cellular debris. Approximately 560 million K562 cells were infected with our next-generation genome-wide lentiviral shRNA library to maintain roughly 1,000-fold coverage of the shRNA library after selection. Infected cells grew for 3 d before the cells were selected with puromycin (0.7 µg/mL, Sigma). After 3 d of selection, infection efficiency was monitored using flow cytometry (BD Accuri C6). Once the cells reached 90–100% mCherry+ cells, they were spun out of selection and allowed to recover in normal RPMI 1640 media. At 70, 500 million cells were pelleted by centrifugation (300g for 5 min). Cells were then split into two populations and maintained at logarithmic growth (500,000 cells per mL) each day for 14 d. After 14 d of growth, cells were spun down (300g for 5 min). Genomic DNA was extracted for all three time points separately according to the protocol included with Qiagen’s Blood Maxi Kit. shRNA-encoding constructs were measured by deep sequencing.

Genome-wide CRISPR/Cas9 screen. A previously designed four-sgRNA-per-gene CRISPR/Cas9 library was used targeting 5’ ends of conserved exons with sgRNAs varying in length between 19 and 25 base pairs. Using K562 cells with stably expressed SFFV-Cas9-BFP, we infected the lentiviral genome-wide sgRNA library into approximately 120 million cells using the same protocol as for the genome-wide shRNA library to maintain at least 1,000-fold representation in cells. Infected cells were selected with puromycin (0.7 µg/mL, Sigma) for 3 d. The percentage of mCherry+ cells was measured by flow cytometry (BD Accuri C6). Selected cells were spun out of selection and into normal RPMI 1640 media. At 70, 120 million cells were spun down (300g for 5 min). Cells were then split into two populations and grown for 14 d, with logarithmic growth (500,000 cells per mL) maintained each day. After 14 d of growth, cells were spun down (300g for 5 min). Genomic DNA was extracted for all three time points separately according to the protocol included with Qiagen’s Blood Maxi Kit. sgRNA-encoding constructs were measured by deep sequencing.

Analysis of previous screens. Data from previous shRNA screens for modifiers of ricin toxicity were obtained as pre-computed hairpin-level enrichments averaged from two replicates; pooled data were analyzed with casTLE and compared to signed, log-transformed published results (Supplementary Data 5). Count data for two replicates from a previous CRISPR/Cas9 cutting screen for lipopolysaccharide-induced TNF expression were obtained from the authors of ref. 16, analyzed with casTLE, and compared to signed, log-transformed published DESeq results (Supplementary Data 6). Data from previous CRISPRi and CRISPRa screens were obtained as pre-computed guide-level enrichment averaged from two replicates, analyzed with casTLE, and compared to published results (Supplementary Data 7). Where available, positive and negative results from published low-throughput validations are also presented.

Analysis of screen results. Deep sequencing on an Illumina NextSeq was used to monitor library composition. Trimmed sequences were aligned to libraries using Bowtie, with zero mismatches tolerated. All alignments from multi-mapped reads were used. Enrichment of individual hairpins was calculated as a median-normalized log-ratio of the fraction of counts, as previously described. Raw count files are available (Supplementary Data 2), and raw FASTQ files have been deposited at the Sequence Read Archive (accession SRP072806) and BioProject (accession PRJNA317269).

Briefly, we built a Cas9 high-Throughput maximum Likelihood Estimator (casTLE) that uses an empirical Bayesian framework to account for multiple sources of variability, including reagent efficacy and off-target effects (a more complete description is provided in the Supplementary Methods). For each gene, we have the phenotypes of multiple targetting reagents. From these, as well as from the phenotypes of negative controls, we obtain an effect-size estimate for each gene and an associated log-likelihood ratio. In the figures we present this as the casTLE score, which is twice the log-likelihood ratio, signed to match the effect size. All screens were analyzed with the same parameters, and no optimization was performed using the gold standard or related sets. casTLE is implemented using custom Python scripts. These, along with a complete screen-analysis pipeline, are available at https://bitbucket.org/dmorgen/casTLE.

For comparison, we implemented RIGER with default settings on precomputed element enrichments with GENE-E (http://www.broadinstitute.org/cancer/software/GENE-E/). RSA was implemented with Python scripts available from http://carrier.gnf.org/publications/RSA/. MAGeCK was implemented with default settings with software available from https://sourceforge.net/projects/hitselect/. Median and highest-effect heuristics and the Mann–Whitney test were implemented with custom Python scripts. GO terms were generated using GOst12, available at http://cbi-gorilla.cs.technion.ac.il/, by ranking genes from highest-confidence negative growth phenotypes to lowest-confidence genes to highest-confidence positive growth phenotypes.

For direct comparison to the four-sgRNA-per-gene CRISPR/Cas9 library, the 25-hairpin-per-gene RNAi library was downsamplde to four shRNAs per gene. Hairpins were ranked according to their original computational design and the top four unique shRNAs were used, as well as negative control shRNAs. Note that this is independent of the data set used here and represents the four-shRNA library that would have been designed. Essential gene predictions were then repeated using this reduced shRNA library as described above (Supplementary Fig. 9).

Analysis of gene expression data and yeast essentials. Gene sets were defined by a 10% FPR cutoff for Cas9, shRNA, and combination screens. Cas9 and shRNA unique gene sets were defined as genes with casTLE scores under the 10% FPR cutoff in either the Cas9 or the shRNA screen but not in both or in the combination screen. An overlap set was defined as genes with casTLE scores under the 10% FPR cutoff for both Cas9 and shRNA screens.

Public RNA-seq data (accession ENCFP934YBO) for the K562 cell line was obtained from ENCODE experiment ENCSR000AEM. Genes were filtered for FPKM (fragments per kilobase of transcript per million mapped reads) > 1 and successful mapping to genes present in our libraries, leaving expression data for ~7,000 genes. These were then ranked from highest to lowest expression and binned in increments of 500. The number of genes in each bin was counted for each gene set and normalized by the total number of genes in each gene set present in the RNA-seq data. This fraction of essential genes for each gene set was then graphed versus the mean log(FPKM) value for each bin (Supplementary Fig. 11a,b and Supplementary Data 10).

Yeast essential genes annotated as viable or inviable, as well as their human homologs, were obtained from the Saccharomyces Genome Database. The number of homologs in each gene set was counted for both inviable and viable annotations, and the fraction inviable is presented (Supplementary Fig. 11c and Supplementary Data 11). P values were calculated using Fisher’s exact test. The total fraction of all homologs mapped from inviable yeast genes is also presented.

29. Efron, B. Stat. Sci. 29, 285–301 (2014).
30. Koss, R.E. & Steffey, D. J. Am. Stat. Assoc. 84, 717–726 (1989).
31. Birmingham, A. et al. Nat. Methods 3, 199–204 (2006).
32. Jackson, A.L. et al. Nat. Biotechnol. 21, 635–637 (2003).
33. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. BMC Bioinformatics 10, 48 (2009).
34. ENCODE Project Consortium. Nature 489, 57–74 (2012).
35. Cherry, J.M. et al. Nucleic Acids Res. 40, D700–D705 (2012).

doi:10.1038/nbt.3567