nature biotechnology

CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes

Bastiaan Evers, Katarzyna Jastrzebski, Jeroen P M Heijmans, Wipawadee Grernrum, Roderick L Beijersbergen & Rene Bernards

High-throughput genetic screens have become essential tools for studying a wide variety of biological processes. Here we experimentally compare systems based on clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein 9 (Cas9) or its transcriptionally repressive variant, CRISPR-interference (CRISPRi), with a traditional short hairpin RNA (shRNA)-based system for performing lethality screens. We find that the CRISPR technology performed best, with low noise, minimal off-target effects and consistent activity across reagents.

Functional genetic screening has provided valuable insights in areas such as oncology, where it is used to identify diagnostic markers and therapeutic targets[1], or in virology to identify host cell factors essential for successful viral entry and/or reproduction[2]. Large-scale gene perturbations are often carried out by RNA interference (RNAi), but this approach is limited by promiscuous off-target activity and variable knockdown efficiencies. These disadvantages can be partly mitigated by using multiple constructs targeting the same gene, but it remains difficult to create effective and reliable genome-wide libraries. The recent adaptation of the prokaryotic immune system CRISPR–Cas9 for use in mammalian cells has allowed easy, efficient and cost-effective genome editing in cultured human cells[3,4]. Soon after its first use in human cells, this system was exploited to perform genome-wide genetic screens similar to those carried out using shRNA, with the fate of cells harboring a particular single guide RNA(sgRNA) sequence monitored by means of deep sequencing[5,6]. A related approach, called CRISPRi, makes use of an enzymatically dead Cas9 (dCas9) fused to a Krüppel-associated box (KRAB) transcriptional repression domain, which does not cleave the target gene but reduces its expression when dCas9 is targeted to a transcriptional start site (TSS), where it inhibits transcription[7]. This technology was also applied to genome-wide interrogation of gene function using sequencing as a readout of relative sgRNA enrichment and/or depletion[8].

Although off-target effects and variable on-target efficiency are well-known limitations of shRNA-based screening, the question remains whether CRISPR and/or CRISPRi approaches avoid these problems and would thus provide superior functional genetic screening. Initial efforts to assess the extent of off-target activity of CRISPR technology have clearly shown the existence of off-target effects[9], whereas other reports suggest that these may be limited[10]. Here we compare the performance of shRNA, CRISPR and CRISPRi in discriminating hits from non-hits in functional genetic screens (**Supplementary Fig. 1**). Incomplete phenotype penetrance limits the dynamic range of hit identification in negative-selection (dropout) screens, where shRNA or sgRNA targeting constructs are depleted from the population. Such negative-selection screens are generally more technically demanding than positive-selection screens. Thus, comparing the three technologies in dropout screens should provide the most reliable information on their relative performance.

Genes essential for cellular life form a useful set to compare the relative performance of the different screening platforms. Efficient suppression of these genes is expected to be lethal, independent of context. Recently, Hart *et al.*[11] set out to create a standard list of essential genes obtained by selecting consistently depleted shRNA vectors across hundreds of genome-wide screens. Similarly, by selecting genes that show consistent low expression and do not affect phenotype upon knockdown, a list of nonessential genes was selected. We selected 46 essential and 47 nonessential genes that are expected to have very consistent phenotypes.

We prepared shRNA, CRISPR and CRISPRi libraries targeting these 93 genes and screened for gene essentiality over a 14-d period (t = 1; **Fig. 1a**). An overview of the relative read-counts mapped to the libraries of t = 1 versus t = 0, normalized for sequencing depth of individual samples and averaged over the three replicates, clearly shows that overall, shRNAs and sgRNAs targeting essential genes are depleted in comparison to those targeting nonessential genes in all three screening systems (**Fig. 1b**). The comparison of the t = 1 sample of the CRISPRi screen maintained without doxycycline to the t = 0 control, shows that this system does not suffer from leaky expression, as depletion of essential sgRNAs was not observed. Scatter plots showing all pairwise read-count correlations further indicate that there was very strong concordance of the data between replicates (**Supplementary Figs. 2–4**). These results show that all three screening technologies performed as intended.

However, the signal variation in the population of control shRNAs is much larger than those of the control sgRNAs in the CRISPR and CRISPRi systems. The high average Pearson's correlation coefficients between the replicates of the shRNA screen at t = 0 (0.98) and t = 1 (0.98) rule out that this variation was due to technical variation between screen replicates but was more likely the result of off-target effects of individual hairpins.

To investigate the extent to which each of the screening technologies is able to recall shRNAs or sgRNAs targeting essential genes, we processed our data using the MAGeCK software package (v0.5.0)[12]. *P*-values, and multiple testing corrected *P*-values (*Q*-values) were calculated

Division of Molecular Carcinogenesis, Cancer Systems Biology Centre and Cancer Genomics Centre Netherlands, The Netherlands Cancer Institute, Amsterdam, the Netherlands. Correspondence should be addressed to R.B. (r.bernards@nki.nl).

Received 12 October 2015; accepted 11 March 2016; published online 25 April 2016; doi:10.1038/nbt.3536

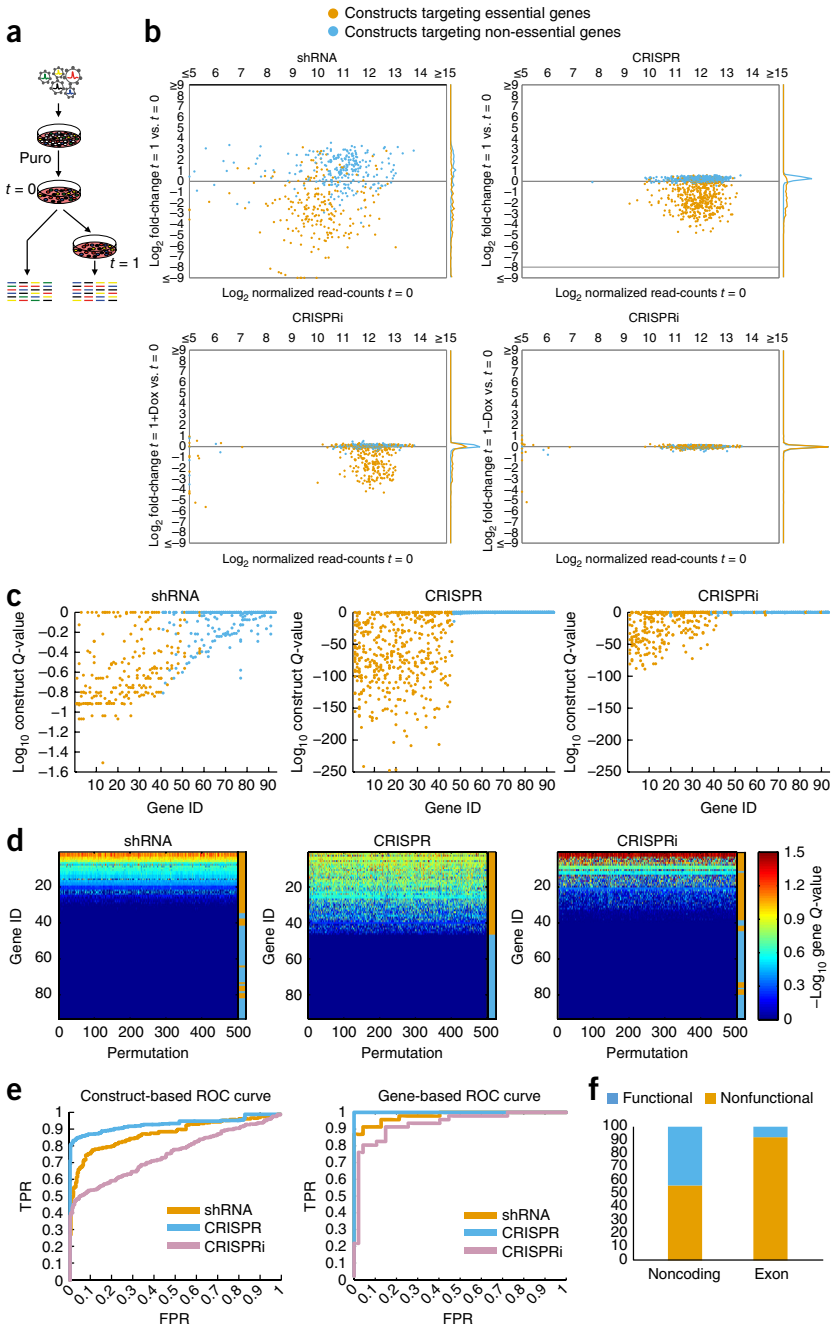

Figure 1 Screen results of RT-112 cells. (**a**) Screen workflow. Low-titer lentiviral supernatant is used to infect cells. Cells are treated with puromycin to select infected cells. Half the cells are harvested for the $t = 0$ sample; the other half is replated and grown for 14 d ($t = 1$). DNA is extracted from all samples and sequenced. (**b**) Constructs for all screening technologies and their relative fold-change of normalized read-counts of $t = 1$ vs. $t = 0$, plotted as a function of their initial averaged normalized read-counts at $t = 0$ and averaged over three biological replicates. To the right of each plot are histograms representing the relative construct density. (**c**) Stacked dot plots showing the construct-based $\log_{10}$ Q-values for depletion at $t = 1$ vs. $t = 0$. Q-values are Benjamini–Hochberg-corrected P-values as calculated by MAGeCK from three independent biological replicate screens. Key as in **b**. (**d**) Heatmaps showing for each gene (row) $-\log_{10}$ Q-values of gene-based P-values as determined by the MAGeCK software of 500 permutations of the data (columns). For the permutations, construct subsets were picked corresponding in size to the minimum number of constructs available for each gene across the three libraries used. Rows were sorted according to average value and the column to the right indicates whether a gene is essential (orange) or nonessential (blue). (**e**) ROC curves for construct-based P-values and gene ranks averaged over the 500 permutations performed. FPR, false-positive rate; TPR, true-positive rate. (**f**) Relative fractions of functional and nonfunctional constructs of sgRNAs targeting exons or noncoding sequences close to the start codon.

for the significance of depletion of every construct from the population based on all three independent biological replicates. The sorted construct-based Q-values clearly show that although most shRNAs targeting essential genes scored better than those targeting nonessential genes, the Q-value differences between the essential and nonessential constructs were much smaller compared to the other technologies, consistent with the presence of more noise in the shRNA screens (**Fig. 1c**).

Apart from these construct-based P-values, MAGeCK also determines whether the rank distribution of P-values targeting the same gene is statistically significantly different from a random rank distribution, leading to a gene-based P-value. To normalize for the different number of constructs targeting each gene across the screening modalities, we performed this analysis 500 times on randomly chosen subsets of constructs that were matched in number to the minimum number available for every gene across the libraries (**Fig. 1d**). We then determined the average rank of gene-based P-values over the 500 analysis runs and used this alongside the construct-based P-values of the complete screens to produce receiver operator characteristic (ROC) curves (**Fig. 1e**). The gene-based ROC curves clearly show that the strategy to use multiple constructs per gene did indeed lead to an overall increase in hit-calling reliability. We further conclude that based on these analyses the CRISPR screen performed best with an absolute separation between essential and nonessential genes and P-values that were overall better than for the other two technologies.

All construct-based ROC curves showed a bipartite nature that suggests all three libraries consist of both functional and nonfunctional constructs. For the CRISPR screen, ~85% of all constructs targeting essential genes were accurately recalled without any false positives, whereas the remaining 15% did not outperform constructs targeting nonessential genes (**Fig. 1e**), suggesting these were not active. For the shRNA screen, while observing that ~75–80% of shRNAs were working well, we noted a ~10% false-positive rate at the point where all active shRNAs were recalled, which is in stark contrast to both the CRISPR and CRISPRi ROC curves that at these transition points showed virtually no off-target effects. When we analyzed the fraction of essential constructs per gene called at these transition points, we observed a quite even spread for shRNA and CRISPR. Fractions of functional CRISPRi constructs, however, varied substantially

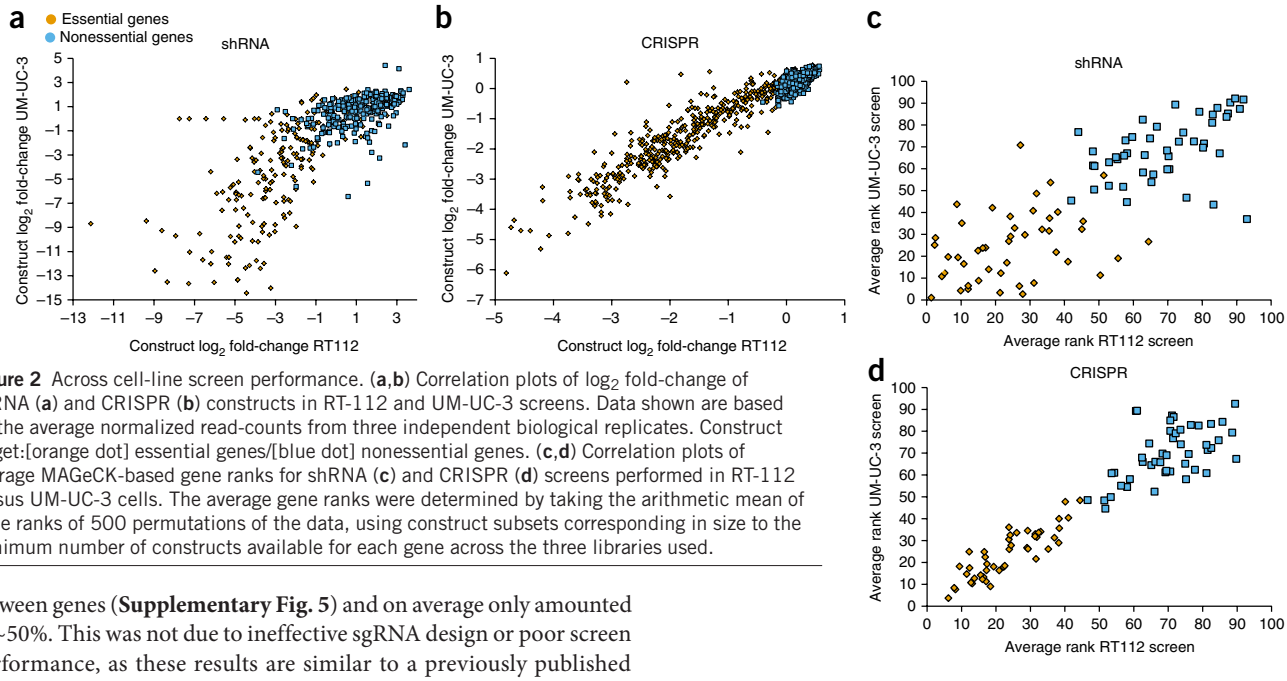

Figure 2 Across cell-line screen performance. (**a**,**b**) Correlation plots of $\log_2$ fold-change of shRNA (**a**) and CRISPR (**b**) constructs in RT-112 and UM-UC-3 screens. Data shown are based on the average normalized read-counts from three independent biological replicates. Construct target:[orange dot] essential genes/[blue dot] nonessential genes. (**c**,**d**) Correlation plots of average MAGeCK-based gene ranks for shRNA (**c**) and CRISPR (**d**) screens performed in RT-112 versus UM-UC-3 cells. The average gene ranks were determined by taking the arithmetic mean of gene ranks of 500 permutations of the data, using construct subsets corresponding in size to the minimum number of constructs available for each gene across the three libraries used.

between genes (**Supplementary Fig. 5**) and on average only amounted to ~50%. This was not due to ineffective sgRNA design or poor screen performance, as these results are similar to a previously published CRISPRi screen (**Supplementary Fig. 6**)[8]. Targeting TSSs of non-transcribed or nonessential transcript variants of essential genes may contribute but does not fully explain the observed variation in sgRNA functionality (**Supplementary Fig. 7**).

We asked whether we could explain the 10–15% of nonfunctional essential CRISPR constructs, to improve future library design. Clearly, even when targeting close to the start codon, sgRNAs that do not cut within the exon itself are much more likely to be nonfunctional. When only those sgRNAs that cut within an exon are considered, the fraction of active sgRNAs increased to >90% (**Fig. 1f**).

The RT-112 cell line that we used in the screens described above has an average ploidy of 2.02 (http://cancer.sanger.ac.uk/cosmic[13]). However, in many screening setups cancer cell lines with higher average copy numbers per gene will be used. Because full CRISPR-mediated functional knockouts are only ensured upon disruptive mutations of all alleles of a gene, we hypothesized that cell lines with higher average ploidy could perhaps perform less well. We therefore repeated both the shRNA and CRISPR screens in UM-UC-3 bladder cancer cells that have an average ploidy of 3.39 (http://cancer.sanger.ac.uk/cosmic[13]). Again, the results of the shRNA screens showed higher variability and more off-target activity compared to the CRISPR screen (**Supplementary Fig. 8**). To further analyze the consistency of the screening methods, we performed a comparison of the screens performed in RT-112 cells to UM-UC-3 cells (**Fig. 2**). Both the depletion of individual constructs targeting essential genes as well as the averaged MAGeCK-based essential gene ranks showed better correlations for the CRISPR screens (correlation coefficients of 0.94 and 0.87, respectively) than for the shRNA screens (correlation coefficients of 0.78 and 0.33, respectively). Thus, our data show that CRISPR screens performed better and more consistently across cell lines, even when applied in cells with a higher average ploidy. These data do not exclude the possibility that highly amplified genes are less efficiently inactivated by CRISPR.

In conclusion, our comparative screens show that the CRISPR system outperformed the shRNA- and CRISPRi-based technologies for identical sets of essential and nonessential genes. This was mostly because of less variation in the data, the presence of more functional constructs, apparent absence of off-target effects resulting in a low false-discovery rate and better consistency across cell lines.

METHODS

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

Accession codes. SRP072947.

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

ACKNOWLEDGMENTS

We thank all members of the Bernards and Beijersbergen groups and G. Bounova for useful discussions, and the Netherlands Cancer Institute genomics core facility for next-generation sequencing. The doxycycline-inducible KRAB-dCas9 expression system, consisting of pHR-TRE3G-KRAB-dCas9-P2A-mCherry and pHR-Tet3G, were kind gifts of Luke Gilbert, the Jonathan Weissman laboratories, UCSF. We thank G. Verhaegh, Nijmegen Institute for Molecular Sciences, the Netherlands, and M. Knowles, Leeds Institute of Cancer Studies and Pathology, UK for the cell lines they kindly provided.

AUTHOR CONTRIBUTIONS

Experiments were designed by B.E., K.J., R.L.B. and R.B., carried out by B.E., K.J., J.P.M.H. and W.G., and analyzed by B.E. The paper was written by B.E., R.L.B. and R.B. This work was supported by the Cancer Genomics Netherlands consortium.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Bernards, R. *Curr. Opin. Genet. Dev.* **24**, 23–29 (2014).
2. Amberkar, S., Kiani, N.A., Bartenschlager, R., Alvisi, G. & Kaderali, L. *World J. Virol.* **2**, 18–31 (2013).
3. Cong, L. *et al. Science* **339**, 819–823 (2013).
4. Mali, P. *et al. Science* **339**, 823–826 (2013).
5. Shalem, O. *et al. Science* **343**, 84–87 (2014).
6. Wang, T., Wei, J.J., Sabatini, D.M. & Lander, E.S. *Science* **343**, 80–84 (2014).
7. Gilbert, L.A. *et al. Cell* **154**, 442–451 (2013).
8. Gilbert, L.A. *et al. Cell* **159**, 647–661 (2014).
9. Fu, Y. *et al. Nat. Biotechnol.* **31**, 822–826 (2013).
10. Kim, D. *et al. Nat. Methods* **12**, 237–243, 1, 243 (2015).
11. Hart, T., Brown, K.R., Sircoulomb, F., Rottapel, R. & Moffat, J. *Mol. Syst. Biol.* **10**, 733 (2014).
12. Li, W. *et al. Genome Biol.* **15**, 554 (2014).
13. Forbes, S.A. *et al. Nucleic Acids Res.* **43**, D805–D811 (2015).

ONLINE METHODS

Cell lines. RT-112 cells were a kind gift of G. Verhaegh, Nijmegen Institute for Molecular Sciences, the Netherlands. The UM-UC-3 cells were a kind gift of M. Knowles, Leeds Institute of Cancer Studies and Pathology, UK. Both cell lines were correctly identified by STR profiling and were tested negative for mycoplasma contamination.

Vector backbones. For the CRISPR system, LentiCRISPRv2.0 (Addgene #52961; ref. 14) plasmid DNA was digested with BsmBI and using GIBSON assembly, a custom-ordered synthetic gene was cloned back in, implementing the A-U flip in the constant domain of the sgRNA as described[15] to create LentiCRISPRv2.1. For the CRISPRi system, we used a doxycycline-inducible KRAB-dCas9 expression system consisting of pHR-TRE3G-KRAB-dCas9-P2A-mCherry and pHR-Tet3G, which were kind gifts of Luke Gilbert, the Jonathan Weissman laboratories, UCSF. For the accompanying sgRNA expression vector, the PacI/EcoRI hUbC-EGFP insert of pFUGW (Addgene #14883) was replaced by an SFFV promoter-driven puromycin-resistance gene, amplified from pHR-SFFV-dCas9-BFP-KRAB (Addgene #46911) and pLenti-Guide-Puro (Addgene #52963), respectively, and cloned in using GIBSON assembly. From this product, again the PacI/EcoRI backbone was used for a GIBSON assembly reaction with a synthetic gene containing a mouse U6 promoter, 2× BfuAI sgRNA cloning sites, an A-U flipped constant sgRNA region and 6×T transcriptional stop signal. For a graphical representation of the three systems used, see **Supplementary Figure 1**.

Libraries. From the core essential genes described by Hart *et al.*[11], we selected a subset of 46 genes belonging to five biological processes that are considered to be essential for all cell types: the COP9 signalosome (7 genes), the proteasome (10 genes), the nuclear pore complex (5 genes), the ribosomal complex (20 genes) and RNA polymerase complex (4 genes). From the nonessential genes identified by Hart *et al.*[11], we randomly selected 47 genes. These two gene sets were used to generate the libraries targeting essential/nonessential gene sets. For the shRNA system, v1.0, v1.5 and v2.0 vectors were picked from glycerol stocks of the publicly available The RNAi Consortium (TRC) libraries (Sigma-Aldrich, MO), grown up on agar plates and combined before maxiprep DNA isolation. For CRISPR, we designed ~10 sgRNAs per gene targeting the first 50% of the ORF or 50 nt surrounding the start codon, minimizing for predicted off-target effects. sgRNAs for the CRISPRi system were designed by first choosing the principal transcripts for every gene using the APPRIS database[16]. Then, sgRNAs were designed to target the region from −50 to +300 relative to the transcription start site (TSS) with a preference for the region between +25 and +100, as this was shown to be the region of highest CRISPRi-based repression[8]. Exact library compositions can be found in **Supplementary Tables 1–3**.

To clone the CRISPR and CRISPRi libraries, oligonucleotide pools were ordered from CustomArray (Bothell, WA) containing sgRNA sequences flanked by 20–30 nt of overlapping vector sequence. A 5′-G was added to all sgRNAs to ensure efficient expression from U6 promoters. Thus, pooled plasmid libraries were prepared and used to produce lentiviral particles using standard procedures.

Screening. Lentiviral supernatants were produced with every library and low titers (multiplicity of infection (MOI) < 0.5) were used to infect RT112 or UM-UC-3 cells, each in three independent biological replicates. The next day, viral supernatant was replaced with medium containing 2 μg/ml puromycin. 48 h later, cells were harvested, keeping half for DNA extraction of the $t = 0$ sample. The other half was reseeded and maintained for another 14 d ($t = 1$), regularly splitting when confluent. In the case of the CRISPRi screen, cells were grown on 250 ng/ml doxycycline from $t = 0$ onwards. At every step of the screening process, sample sizes were at least 1,000-fold the complexity of the library to keep sufficient dynamic range in our measurements.

PCR and next-generation sequencing. For every sample, a maximum of 24 μg genomic DNA was divided over 24 50-μl PCR reactions using barcoded forward primers to be able to deconvolute multiplexed samples after next-generation sequencing (for primers and barcodes used, see **Supplementary Table 4**). PCR mixture per reaction: 10 μl HF Buffer (NEB), 1 μl 10-μM forward primer, 1 μl 10-μM reverse primer, 0.5 μ Phusion Hot Start II polymerase (NEB, cat. # M0530L), 1 μl 10-mM dNTPs, adding mQ and template to 50 μl. Cycling conditions: 2′ @98 °C, 20× (30″ @98 °C, 30″ @60 °C, 30″ @72 °C), 5′ @72 °C. The products of all reactions were pooled and 2 μl of this PCR 1 product was used in a subsequent PCR 2 reaction using primers containing adapters for next-generation sequencing (**Supplementary Table 4**). The same PCR protocol was used, but only for 15 instead of 20 cycles. Next, PCR products were purified using standard PCR purification columns, DNA concentrations were measured and, based on this, samples were equimolarly pooled and subjected to Illumina next-generation sequencing. Mapped read-counts were subsequently used as input for the MAGeCK analysis software package, using version 0.5 (ref. 12).

14. Sanjana, N.E., Shalem, O. & Zhang, F. *Nat. Methods* **11**, 783–784 (2014).
15. Chen, B. *et al. Cell* **155**, 1479–1491 (2013).
16. Rodriguez, J.M. *et al. Nucleic Acids Res.* **41**, D110–D117 (2013).

doi:10.1038/nbt.3536