LETTER

doi:10.1038/nature13166

High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells

Yuexin Zhou[1]*, Shiyou Zhu[1]*, Changzu Cai[1]*, Pengfei Yuan[1], Chunmei Li[2], Yanyi Huang[2] & Wensheng Wei[1]

Targeted genome editing technologies are powerful tools for studying biology and disease, and have a broad range of research applications[1–7]. In contrast to the rapid development of toolkits to manipulate individual genes, large-scale screening methods based on the complete loss of gene expression are only now beginning to be developed[8,9]. Here we report the development of a focused CRISPR/Cas-based (clustered regularly interspaced short palindromic repeats/CRISPR-associated) lentiviral library in human cells and a method of gene identification based on functional screening and high-throughput sequencing analysis. Using knockout library screens, we successfully identified the host genes essential for the intoxication of cells by anthrax and diphtheria toxins, which were confirmed by functional validation. The broad application of this powerful genetic screening strategy will not only facilitate the rapid identification of genes important for bacterial toxicity but will also enable the discovery of genes that participate in other biological processes.

Recent progress in genomic editing, such as ZFN (zinc finger nuclease)[1–3], TALENs (transcription activator-like effector nucleases)[3–5] and the CRISPR/Cas system[6,7], has markedly changed the way that researchers study genes and their functions in mammalian systems. The CRISPR/Cas system was originally found to provide bacteria and archaea with adaptive immunity against viruses and plasmids[10]. By means of the type II CRISPR system, the Cas9 nuclease can be directed by a chimaeric single guide RNA (sgRNA) via Watson–Crick base pairing[11] to the desired genomic loci followed by an NGG protospacer-adjacent motif (PAM) to create double-stranded DNA breaks (DSBs)[6,7]. Taking advantage of this effective gene-editing technology with easy programmability, we developed a focused sgRNA library that enables gene identification from function-based genetic screening.

The CRISPR/Cas system requires both the gene-specific sgRNA and an identical nuclease Cas9 to achieve DSBs at the targeted locus[6,7]. We developed cell lines with a constant expression of Cas9 with OCT1 (Fig. 1a and Supplementary Fig. 1), as it has been reported that overexpression of OCT1 would further boost the sgRNA expression through enhanced U6 promoter activity[12]. We used the Golden Gate cloning method[13] to construct the sgRNA on the lentiviral backbone (Fig. 1b and Supplementary Fig. 1). To verify whether the expression of sgRNA through lentiviral delivery is sufficient to guide Cas9 to the targeted loci to generate DSBs, we arbitrarily selected the *CSPG4* gene for targeting and conducted the T7E1 digestion assay as previously described[14]. *CSPG4*-specific sgRNA was found to induce insertions and deletions (indels) through viral infection in all three lines constantly expressing Cas9 and OCT1 (that is, HEK293T$_{OC}$, HT1080$_{OC}$ and HeLa$_{OC}$, where OC indicates OCT1 and Cas9). However, the pooled HeLa$_{OC}$ cells showed lower efficiency in creating indels than the pooled HEK293T$_{OC}$ and HT1080$_{OC}$ cells at day 5 after infection, whereas one isolated single clone of HeLa$_{OC}$, designated as HeLa$_{OC}$-SC, showed the highest level of efficiency (Fig. 1c). This difference in indel efficiency is unlikely to be due to the variable levels of Cas9 expression (Extended Data Fig. 1), and prolonged culturing enabled HeLa$_{OC}$ at day 8 to reach similar cleavage efficiency as HEK293T$_{OC}$ and HT1080$_{OC}$ at day 5 (Fig. 1c), although still significantly lower than HeLa$_{OC}$-SC. It is therefore critical to isolate a single clone with high DSB mis-rejoining activity, at least for certain cell types. The beneficial role of exogenous expression of OCT1 requires further investigation.

We next examined the off-target activity of virally delivered sgRNA in HeLa$_{OC}$-SC. Three representative sgRNAs targeting three genes (*ANTXR1*, *HBEGF* and *CSPG4*) were chosen for this analysis. The top five off-target sites were selected for each sgRNA, and the T7E1 assays indicated that there were no off-target cleavages except for one

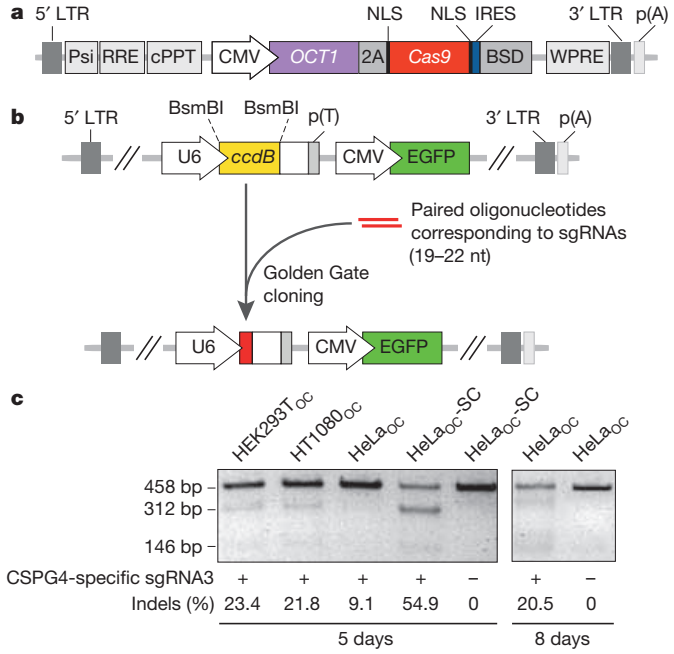


Figure 1 | Lentiviral-delivered sgRNA creates indels with high efficiency in human cells stably expressing Cas9 and OCT1. a, Structure of the lentiviral plasmid expressing 2A-linked OCT1 and Cas9. An IRES-linked blasticidin-resistant gene was fused with the *cas9* gene to facilitate the selection of cell clones with high-level expression of Cas9 and OCT1. The pLenti-CMV-BSD vector was used for plasmid construction. nt, nucleotides; LTR, long terminal repeat. **b**, Schematic diagram of the sgRNA expression construct in a lentiviral backbone. U6 promoter-driven *ccdB* and sgRNA scaffold were cloned into the LL3.7 lentiviral vector. Annealed oligonucleotide pairs encoding customized sgRNA were ligated into the lentiviral backbone using the Golden Gate method. **c**, Indels induced by lentivirus-delivered sgRNA (5′-TTGGCCAGACTTG CATCCG-3′) targeting the *CSPG4* gene (CSPG4-specific sgRNA3) in the indicated cells were assayed by T7E1 digestion. All cells were incubated for 5 or 8 days after infection before assay as indicated. Primers used for the PCR reactions are listed in Supplementary Table 5, and the percentage of cleaved band was measured using the ImageJ program (http://rsbweb.nih.gov/ij/) for this and other figures.

[1]State Key Laboratory of Protein and Plant Gene Research, College of Life Sciences, Peking University, Beijing 100871, China. [2]Biodynamic Optical Imaging Centre (BIOPIC), College of Engineering, Peking University, Beijing 100871, China.
*These authors contributed equally to this work.

©2014 Macmillan Publishers Limited. All rights reserved

case, in which the off-target site of sgRNA targeting *HBEGF* has exactly the same sequence for the last 15 base pairs immediately upstream of PAM (Supplementary Table 1). All of the off-target sites tested were located in the non-coding regions, and therefore unlikely to alter gene function.

To establish our methodology, we created a library targeting 291 human genes that fit our specific research interests. Assisted by a rule[7]-based algorithm, we designed three kinds of sgRNAs targeting each gene for a majority of the 291 genes (Supplementary Table 2). The single clone HeLa$_{OC}$-SC was chosen to produce the library because this particular line generated the highest level of indels (Fig. 1c). The library, designated as HeLa/sgRNA^{library}, was created in such a way that the coverage of sgRNAs is approximately 1,000-fold with a virus MOI (multiplicity of infection) of 0.05. We then applied this library to a functional screening process using two bacterial toxins: diphtheria and chimaeric anthrax (PA/LFnDTA, protective antigen (PA)/N-terminal domain of lethal factor (LF) fused to the catalytic subunit of diphtheria toxin) toxins. Figure 2 displays a schematic of the library construction and functional screening. After three rounds of toxin treatment, the surviving cells were collected, most of which exhibited green fluorescence (Extended Data Fig. 2), indicating the active presence of sgRNA-containing cartridges that carry CMV promoter-driven EGFP (Fig. 1b). The genomic DNA of the pooled surviving cells from the library, as well as the original library of cells before the toxin treatment, was extracted and used for PCR amplification of the sgRNA-coding regions, before being subjected to deep-sequencing analysis (Fig. 2). These sgRNA-coding regions can be amplified rather specifically without interference from the host chromosomal DNA (Extended Data Fig. 3).

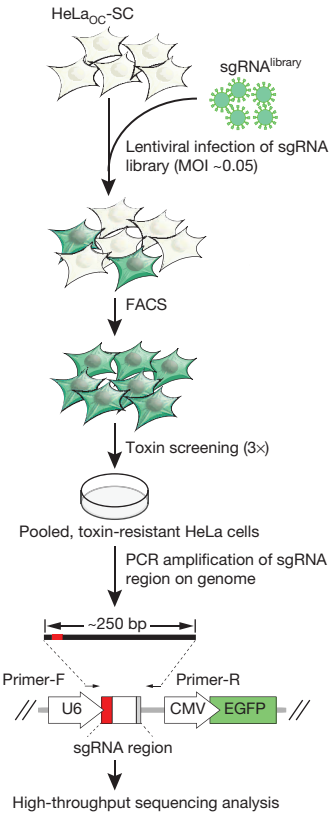


Figure 2 | Schematic of sgRNA library construction and functional screening. sgRNAs were delivered into HeLa$_{OC}$-SC cells by lentiviral infection with a MOI of 0.05. Library cells stably expressing sgRNAs were obtained by FACS for green fluorescence 2 days after infection. Library screening was conducted by three rounds of toxin treatment, followed by PCR amplification of the sgRNA-coding sequence integrated into the chromosomes. Primers used for the PCR reaction are listed in Supplementary Table 5. Purified PCR product was subjected to high-throughput sequencing analysis with Illumina HiSeq 2500.

High-throughput sequencing analysis revealed a total of 863 (99.3%) sgRNA sequences from the control library that was designed to contain 869 kinds of sgRNA targeting a total of 291 genes (Supplementary Table 2). The read counts for most of the sgRNA coding sequences were consistent among three repeats (Supplementary Table 3 and Extended Data Fig. 4a) and after prolonged culturing (Extended Data Fig. 4b). After three rounds of toxin treatment, however, the composition and the read count of the sgRNAs showed distinct differences between the two different screens. We organized all the $\log_2$-fold changes of the normalized counts into two heat maps, ranked according to their enrichment level from either PA/LFnDTA or diphtheria toxin selection. Samples of the top hits in each screening were also shown with their gene names and sgRNA numbers. Most of the sgRNAs were depleted in both cases, as expected (Fig. 3a). Using the analytical tool DESeq2 (ref. 15), we selected 21 sgRNAs targeting 19 genes from PA/LFnDTA screening and 15 sgRNAs targeting 15 genes from diphtheria toxin screening as potential candidates for future studies (Figs 3b, c and Supplementary Table 4). Notably, all three sgRNAs targeting the same known anthrax toxin receptor gene *ANTXR1* (also called *TEM8*) (ref. 16) were enriched and ranked at the top from the screening of PA/LFnDTA (Fig. 3b), and one of the two sgRNAs targeting the diphtheria toxin receptor gene *HBEGF* (ref. 17) ranked at the top from the screening of the diphtheria toxin (Fig. 3c).

Notably, the three *ANTXR1*-targeting sgRNAs varied in their efficacy of inducing gene interruption, as shown by the rates of indels determined by a T7E1 digestion assay. The two most strongly enriched sgRNAs (sgRNA1 and sgRNA3) showed a much higher efficiency in causing indels at the targeted locus of the *ANTXR1* gene (Fig. 3d). Consistently, the two sgRNAs targeting *HBEGF* also had a distinct efficiency in inducing indels (Fig. 3e), with *HBEGF* sgRNA3, with a much higher level of efficiency, being enriched and ranked at the top of the diphtheria toxin screening results (Fig. 3a), whereas the other sgRNA targeting *HBEGF* ranked at number 35 (Supplementary Table 4).

To validate the functions of candidate genes identified from the library screening, we generated individual gene knockouts in HeLa cells for *ANTXR1* and *HBEGF*. The *ANTXR1* knockout in HeLa cells was created by either the TALEN technique[18] or CRISPR/Cas9 system (Extended Data Fig. 5a, b). We tested all three mutant clones (one created by TALENs and the other two created by CRISPR/Cas9) for their susceptibility to the two toxins using both XTT (2,3-bis-(2-methoxy-4-nitro-5-sulfophenyl)-2H-tetrazolium-5-carboxanilide) cell viability and LDH (lactate dehydrogenase) cell death assays. The HeLa *ANTXR1*$^{-/-}$ became totally resistant to PA-mediated toxicity, while remaining susceptible to the diphtheria toxin (Fig. 4a and Extended Data Fig. 5c, d). Similarly, HeLa *HBEGF*$^{-/-}$ was completely resistant to diphtheria toxin, but not to PA/LFnDTA for two independent clones created through the CRISPR/Cas9 system (Fig. 4b and Extended Data Fig. 6). The identification and confirmation of these two known toxin receptors demonstrates the high level of efficiency and reliability of the CRISPR/Cas9 library in functional screening.

To determine whether our screening revealed novel genes involved in the two toxins' intoxication mechanisms, we picked candidate genes from the top of sgRNA enrichment lists (Fig. 3b, c) for further validation. Because only one sgRNA was positively selected for most of the candidates, we designed additional sgRNAs (Supplementary Table 6) and examined their capacity to produce a toxin-resistance phenotype. Four genes (*PLXNA1*, *FZD10*, *PECR* and *CD81*) were confirmed to be potential functional candidates specifically involved in PA-mediated anthrax toxicity because at least two different sgRNAs for each candidate produced a toxin-resistance phenotype (Fig. 4c and Extended Data Fig. 7). As there was only one *CFTR*-targeting sgRNA conferring resistance, further analysis was required to rule out the possibility that this is not due to an off-target effect. Similarly, the involvement of RAB2A in diphtheria toxin toxicity was supported by the effect of

©2014 Macmillan Publishers Limited. All rights reserved

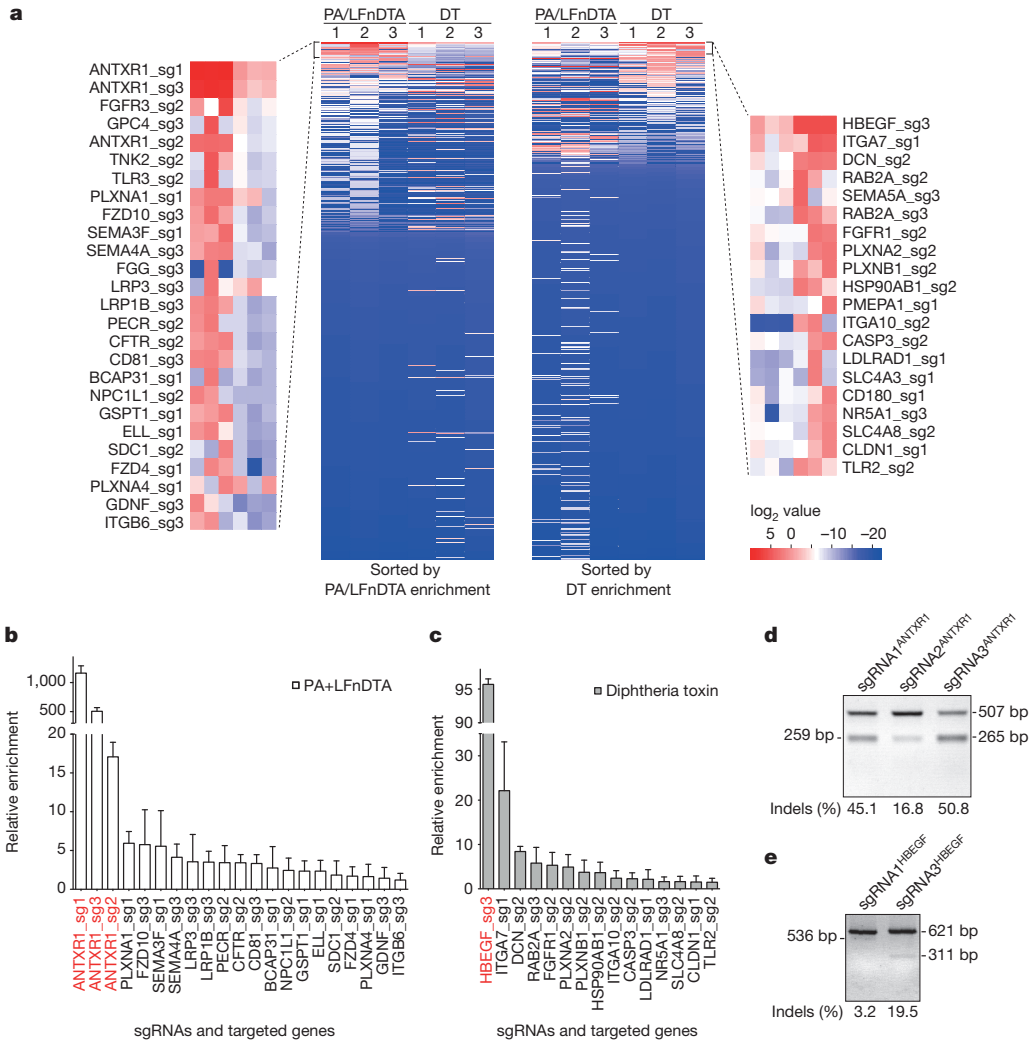


Figure 3 | Screening for essential genes for PA/LFnDTA and diphtheria toxin (DT) toxicity. **a**, Heat maps generated from sgRNA library screening of PA/LFnDTA and diphtheria toxins and high-throughput sequencing analysis. The two heat maps are sorted separately by average enrichment level of sgRNAs obtained from three parallel library screenings with PA/LFnDTA (left) and diphtheria toxin (right). The enrichment level is calculated as $\log_2$(Exp:Ctrl) of the normalized read counts. The partial enlarged images show the top sgRNAs with their targeted genes determined from the screens of two toxins. **b**, **c**, Enrichment of candidate sgRNAs and their targeted genes from PA/LFnDTA (**b**) and diphtheria toxin (**c**) screening. sgRNA enrichment was calculated as the average of normalized experimental read counts divided by the control, and was plotted as the mean of three replicates in descending order. sgRNAs targeting positive genes known to be important for anthrax (**b**) or diphtheria toxin (**c**) toxicity are highlighted in red. Data are presented as the mean ± s.e.m. ($n = 3$). **d**, **e**, T7E1 digestion assays indicating indels induced by indicated sgRNAs targeting *ANTXR1* (**d**) or *HBEGF* (**e**).

multiple *RAB2A*-targeting sgRNAs in $HeLa_{OC}$-SC cells (Extended Data Fig. 8).

To assess further the candidate genes isolated from our library screening and the above validation procedure, we chose the *PECR* gene for further extensive study. *PECR* gene knockouts were isolated through the CRISPR/Cas9 system using two different sgRNAs, and the complete disruption of *PECR* was confirmed by Sanger sequencing (Extended Data Fig. 9a, b). The complete loss of *PECR* in both independently acquired clones conferred on HeLa cells increased levels of resistance to PA/LFnDTA, but not to diphtheria toxin (Fig. 4d and Extended Data Fig. 9c). *PECR*-specific short hairpin RNAs (shRNAs) also generated a resistant phenotype in HeLa cells to PA-mediated killing (Extended Data Fig. 9d), further demonstrating the positive role of PECR in anthrax toxicity.

In addition to *ANTXR1*, we also designed three sgRNAs targeting the other anthrax toxin receptor gene, *ANTXR2* (ref. 19). Only two sgRNAs (sgRNA2 and sgRNA3) were retained in the library. Although $sgRNA2^{ANTXR2}$ induced indels at an efficiency of 21.3% by T7E1 assay, neither sgRNA generated knockout cells that survived the PA/LFnDTA toxin treatment (Supplementary Tables 3 and 4). Using two pairs of *ANTXR1*- and *ANTXR2*-targeting TALENs construct[18], we introduced each pair of plasmids into HeLa and HEK293T cells. From randomly picked clones harbouring either of the two pairs of TALENs constructs, toxin-resistant clones were identified only from the HeLa $TALENs^{ANTXR1}$ (19 of 50) and the HEK293T $TALENs^{ANTXR2}$ (5 of 32), but not from the HeLa $TALENs^{ANTXR2}$ (0 of 50) or the HEK293T $TALENs^{ANTXR1}$ (0 of 32) cells (Extended Data Fig. 10a), suggesting that these two receptors have different roles in these two cell types. Reverse transcription PCR analysis revealed that the basal expression of *ANTXR2* transcripts was barely detectable in HeLa cells (Extended Data Fig. 10b). This explained why ANTXR1, but not ANTXR2, has such a dominant role in HeLa cells. As a more potent receptor[20], ANTXR2 instead of ANTXR1 has a more significant role when the expression of both genes was comparable in HEK293T cells (Extended Data Fig. 10).

Although arrayed and pooled screens using RNA interference (RNAi) libraries have already been developed and are widely used for systematic genetic studies in mammalian cells, they do have certain limitations—in particular, the RNAi-based downregulation of a particular gene is not always sufficient to cause the phenotypic change of interest[21]. Therefore, methods based on gene knockout screening are highly desirable. Here, we established an effective method for the construction of a pooled gene-knockout library and demonstrated that this CRISPR-based strategy can be seamlessly implemented in functional genomics when combined with deep-sequencing analysis. The broad application of this scalable genetic tool may further augment the power of the CRISPR/Cas system in studying gene functions in a high-throughput fashion.

While our manuscript was under revision, two parallel studies[8,9] reported a similar lentiviral-based pooled screening using the CRISPR/Cas9 system, both of which developed large libraries covering over 18,000 (ref. 9) and 7,000 (ref. 8) genes. The genome-scale screening of the CRISPR/Cas9 library is valuable; however, it might be technically challenging to manage the library. Our approach presents an alternative way—that is, via a focused library—to conduct functional genomics,

©2014 Macmillan Publishers Limited. All rights reserved

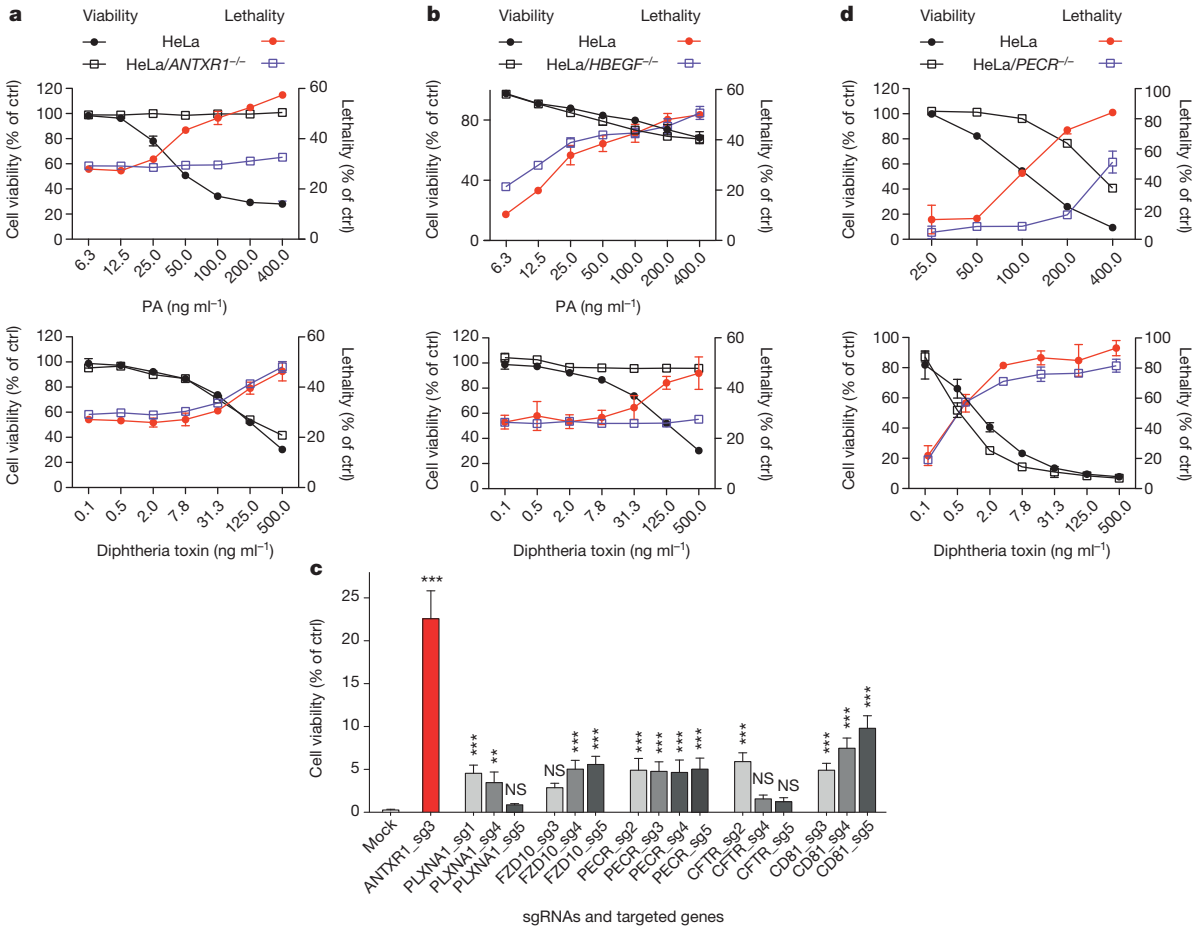


Figure 4 | Genetic validation of candidate genes. **a**, **b**, **d**, Effects of *ANTXR1* (**a**), *HBEGF* (**b**) and *PECR* (**d**) deficiency on PA/LFnDTA- and diphtheria-toxin-triggered cytotoxicity in HeLa cells. The cells were treated with different doses of toxins, and the XTT cell viability assay and LDH lethality assay were performed 48 h after toxin treatment. Data are presented as the mean ± s.d. ($n = 3$). **c**, Relative cell viability for $HeLa_{OC}$-SC cells transfected with indicated sgRNA constructs. Four days after transfection, cells were incubated with PA (70 ng ml^{-1}) plus LFnDTA (50 ng ml^{-1}) for 48 h. Images of toxin-treated and untreated cells were taken (Extended Data Fig. 7) for cell number survey by ImageJ program. The cell viability was calculated as the percentage of surviving cells after toxin treatment. The data are the mean ± s.d. ($n = 6$); $^{**}P < 0.01$; $^{***}P < 0.001$; NS, not significant; Dunnett's multiple comparison test, one-way ANOVA.

and is particularly advantageous for knowledge-based screening. The other difference is that ref. 9 chose to combine Cas9 and sgRNA expression into a single vector, as it is time-consuming to generate the cell lines that express Cas9. However, we found that individual clones harbouring Cas9 varied in their efficiency to generate sgRNA-mediated indels. It is therefore important to select the best single clone for library construction for certain cell types. In addition, the universal background of Cas9 derived from a single clone is beneficial for the accurate assessment of sgRNA contribution manifested by the final counts from the deep-sequencing analysis.

METHODS SUMMARY

The lentiviral plasmid expressing OCT1 and Cas9 was constructed using Gibson's method[22,23] in pLenti-CMV-BSD, and the lentiviral backbone used for the sgRNA library cloning was constructed in the pLL3.7 vector (Addgene, Inc.). The oligonucleotides used for the sgRNA library construction were individually designed (Supplementary Table 2) and synthesized (Ruibiotech, Inc.). $HeLa_{OC}$-SC cells infected by the library viruses were sorted by FACS (MOFLO, Cytomation) based on the presence of EGFP. The library cells were screened by three rounds of toxin treatment: PA (70 ng ml^{-1}) plus LFnDTA (50 ng ml^{-1}); and diphtheria toxin (7.5 ng ml^{-1}). Three replicates were arranged for each toxin screening. The genomic DNA of each replicate as well as the original library was isolated using the DNeasy Blood and Tissue kit (Qiagen). sgRNA-coding regions integrated into the chromosomes were PCR-amplified (TransTaq DNA Polymerase High Fidelity, TransGen) through 26 cycles of reaction with primers annealed to the flanking sequences of the sgRNAs (Supplementary Table 5) before being subjected to high-throughput sequencing analysis (Illumina HiSeq 2500).

PA and LFnDTA, a surrogate of LF consisting of the amino-terminal domain of LF fused to the catalytic subunit of diphtheria toxin, were produced by using plasmid pET-22b-PA and pET-15b-LFnDTA (Addgene plasmid 11079 and 11075), respectively. The diphtheria toxin was purchased from List Biological Laboratories, Inc. Cytotoxicity assays were performed as described[24] with XTT (Roche) according to the product manual. LDH staining and detection were performed as described in the product instruction (CytTox96, Promega).

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 December 2013; accepted 24 February 2014.
Published online 9 April 2014.

1. Porteus, M. H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300,** 763 (2003).
2. Miller, J. C. *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nature Biotechnol.* **25,** 778–785 (2007).
3. Wood, A. J. *et al.* Targeted genome editing across species using ZFNs and TALENs. *Science* **333,** 307 (2011).
4. Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nature Biotechnol.* **29,** 143–148 (2011).
5. Zhang, F. *et al.* Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nature Biotechnol.* **29,** 149–153 (2011).
6. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339,** 819–823 (2013).

©2014 Macmillan Publishers Limited. All rights reserved

7. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339,** 823–826 (2013).
8. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343,** 80–84 (2014).
9. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343,** 84–87 (2014).
10. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315,** 1709–1712 (2007).
11. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337,** 816–821 (2012).
12. Lin, B. R. & Natarajan, V. Negative regulation of human U6 snRNA promoter by p38 kinase through Oct-1. *Gene* **497,** 200–207 (2012).
13. Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS ONE* **4,** e5553 (2009).
14. Mussolino, C. *et al.* A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* **39,** 9283–9293 (2011).
15. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11,** R106 (2010).
16. Bradley, K. A., Mogridge, J., Mourez, M., Collier, R. J. & Young, J. A. Identification of the cellular receptor for anthrax toxin. *Nature* **414,** 225–229 (2001).
17. Naglich, J. G., Metherall, J. E., Russell, D. W. & Eidels, L. Expression cloning of a diphtheria toxin receptor: identity with a heparin-binding EGF-like growth factor precursor. *Cell* **69,** 1051–1061 (1992).
18. Yang, J. *et al.* ULtiMATE system for rapid assembly of customized TAL effectors. *PLoS ONE* **8,** e75649 (2013).
19. Scobie, H. M., Rainey, G. J., Bradley, K. A. & Young, J. A. Human capillary morphogenesis protein 2 functions as an anthrax toxin receptor. *Proc. Natl Acad. Sci. USA* **100,** 5170–5174 (2003).
20. Liu, S. *et al.* Capillary morphogenesis protein-2 is the major receptor mediating lethality of anthrax toxin *in vivo*. *Proc. Natl Acad. Sci. USA* **106,** 12424–12429 (2009).
21. Zuber, J. *et al.* Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nature Biotechnol.* **29,** 79–83 (2011).
22. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6,** 343–345 (2009).
23. Gibson, D. G. Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol.* **498,** 349–361 (2011).
24. Wei, W., Lu, Q., Chaudry, G. J., Leppla, S. H. & Cohen, S. N. The LDL receptor-related protein LRP6 mediates internalization and lethality of anthrax toxin. *Cell* **124,** 1141–1154 (2006).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Xi for providing the Cas9-encoding construct, F. Tang for technical advice regarding the high-throughput sequencing analysis, Z. Jiang for aid with the cell lines, Y. Yang for assistance in primer design, Y. Wang for data analysis, and D. Trono for aid with the lentivirus packaging system. We thank Peking University High-throughput Sequencing Center operated by BIOPIC. This work was supported by funds from the National Basic Research Program of China (2010CB911800), the National Science Foundation of China (NSFC31170126, NSFC31070115), and the Peking-Tsinghua Centre for Life Sciences.

Author Contributions W.W., Y.Z., S.Z. and C.C. conceived the methodology and designed the experiments. Y.Z., S.Z., C.C., P.Y. and C.L. performed the experiments. W.W., Y.Z., S.Z., C.C., P.Y. and Y.H. analysed the data. W.W., C.C., Y.Z. and S.Z. wrote the manuscript with help from the other authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to W.W. (wswei@pku.edu.cn).

©2014 **Macmillan Publishers Limited. All rights reserved**

METHODS

Cells and reagents. HT1080, HeLa and HEK293T cells were maintained in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 10% fetal bovine serum (FBS, Invitrogen) with 5% CO_2 in 37 °C.

Plasmid construction. The lentiviral sgRNA vector (pLenti-sgRNA-Lib) was constructed by substituting its original U6 promoter in pLL3.7 (Addgene, Inc.) for the human U6 promoter, *ccdB* cassette and sgRNA scaffold. The gene *OCT1* was PCR-amplified from a human cDNA library (Abclonal, Inc.) and fused with a 2A-linked[25] humanized *Cas9* gene[26], followed by an IRES (pIRES, Invitrogen)-linked blasticidin-resistance gene. The construction of this plasmid (pLenti-OC-IRES-BSD) was accomplished using Gibson's method[22,23], and was verified by sequencing analysis. Details regarding these two constructs are given in Supplementary Fig. 1.

Construction of the CRISPR/Cas9 sgRNA library. We created a library targeting 291 human genes based on our previous shRNAmir library screening results (our own unpublished data). We have subjectively added additional genes based on previous knowledge and our research interests. We preferentially selected genes encoding cell surface proteins, and those important for endocytosis, trafficking and cell death. The oligonucleotides for each sgRNA-coding sequence for the library were individually designed (Supplementary Table 2) and synthesized (Ruibiotech, Inc.). Paired oligonucleotides were mixed in 96-well plates to the final concentration of 9 μM in 50 μl of TransTaq HiFi Buffer II (1×) for annealing. These annealed oligonucleotide pairs were then mixed, phosphorylated and ligated into the lentiviral sgRNA vector using the Golden Gate method. The ligation mixture was transformed into Trans1-T1 competent cells (Transgen) to obtain the library plasmids.

Lentiviruses of the sgRNA library were obtained by co-transfection of library plasmids with two viral packaging plasmids psPAX2 and pMD2.G (from Didier Trono of EPFL) into HEK293T using the polyethylenimine (PEI) method. HeLa (OCT1-Cas9) cells were infected by the library viruses, followed by FACS (MOFLO, Cytomation) for EGFP-positive cells, 48 h after infection. The viral titre used was 0.05, and the total cell number for viral infection was 4×10^7.

CRISPR/Cas9 sgRNA library screening. A total of 3×10^7 sgRNA library cells were plated onto 100-mm Petri dishes at 1×10^6 cells per plate. After 24 h, the cells were treated by toxins at an appropriate concentration: PA (70 ng ml^{-1}) plus LFnDTA (50 ng ml^{-1}); and diphtheria toxin (7.5 ng ml^{-1}). For each toxin screening, five plates of cells were grouped together as one replicate, and three replicates were arranged for each toxin screen. For each round of screening, the library cells were incubated with the appropriate toxins for 2 days before being changed to fresh DMEM. The surviving cells were re-seeded for growth and subjected to another round of toxin screening. After three rounds of toxin treatment, the resistant clones and the original untreated library cells were collected separately for genomic DNA extraction, followed by PCR amplification of the sgRNA-coding region and deep-sequencing analysis.

Identification of candidate sgRNA sequences. The genomic DNA of six replicates was isolated from 5×10^6 cells using the DNeasy Blood and Tissue kit (Qiagen). sgRNA-coding regions integrated into the chromosomes were PCR-amplified (TransTaq DNA Polymerase High Fidelity, TransGen) with 26 cycles of reaction using primers annealed to the flanking sequences of the sgRNAs (Supplementary Table 4). In this work, 0.5 μg genomic DNA was used in each 50-μl PCR reaction, 16 PCR reactions were performed for each replicate, and different replicates were indexed with different barcodes. The PCR products (16 tubes) of each replicate were pooled and purified with DNA Clean & Concentrator-5 (Zymo Research Corporation), followed by high-throughput sequencing analysis (Illumina HiSeq 2500).

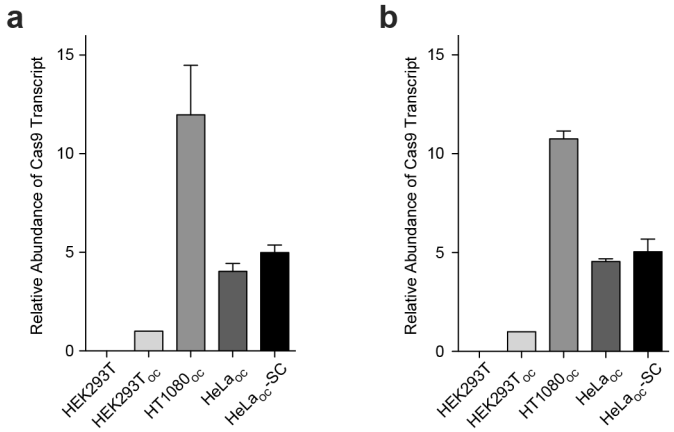
Statistical analysis. The R software package from Bioconductor, DESeq2, was used to perform a statistical analysis of the sequencing data. The enrichment of sgRNA was ranked by the average fold change of normalized counts: $reads^{Exp}/reads^{Ctrl}$. The adjusted *P* value was calculated to evaluate the data quality. The criterion for the selection of candidate sgRNAs and their targeted genes is an adjusted $P \leq 0.05$ for which the fold changes are greater than that of the negative control plus its standard deviation (mean + s.d.). sgRNAs targeting *ANTXR1* and *HBEGF* were chosen to serve as negative controls for the diphtheria toxin and PA/LFnDTA screening, respectively.

Cytotoxicity assay. PA and LFnDTA, a surrogate of LF consisting of the N-terminal domain of LF fused to the catalytic subunit of diphtheria toxin, were produced using plasmid pET-22b-PA and pET-15b-LFnDTA (Addgene plasmid 11079 and 11075), respectively. Diphtheria toxin was purchased from List Biological Laboratories, Inc. Cytotoxicity assays were performed as described[24] with XTT (Roche) according to the product manual. LDH staining and detection were performed as described in the product instruction (CytTox96, Promega). The death signal represented by the amount of LDH release was normalized to the wells based on the maximum LDH activity of the total lysed cells. Each data point and related error bar shown in the figures for the XTT or LDH assays represent the average results from three replicates.

Real-time PCR. RNA of cultured cells was extracted using EasyPure RNA kit (Transgen, ER101-01), and the cDNAs were synthesized using PrimeScript 1st Strand cDNA Synthesis kit (TAKARA, 6110A). Real-time PCR was performed with Platinum SYBR green qPCR SuperMix-UDG (Life Technologies, C11733-038) on Stratagene Mx3005P qPCR system. Two pairs of primers were designed to examine the expression level of each gene. β-actin and GAPDH transcript levels were measured as internal controls.

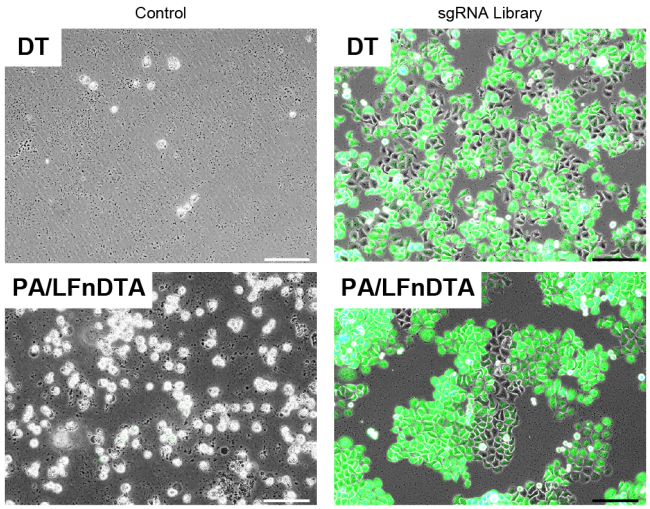
25. Kim, J. H. *et al.* High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS ONE* **6,** e18556 (2011).
26. Chang, N. *et al.* Genome editing with RNA-guided Cas9 nuclease in Zebrafish embryos. *Cell Res.* **23,** 465–472 (2013).

©2014 Macmillan Publishers Limited. All rights reserved



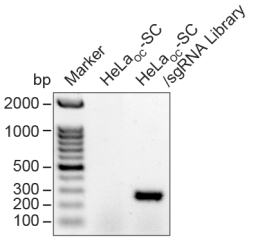
Extended Data Figure 1 | Relative expression level of Cas9 by real-time PCR. Two pairs of primers (**a**, 5′-GGGGGACAGTCTTCACGAGC-3′ and 5′-CACGTACATGTCCCTGCCGT-3′; **b**, 5′-GCACGTGGCCCAAATTCTCG-3′ and 5′-AGCGGTGGCCTTGCCTATTT-3′) were designed to test Cas9 expression level. cDNA of HEK293T serves as the negative control. β-Actin was used as the internal control and the Cas9 expression of $HEK293T_{OC}$ was set as the calibrator. The data are the mean ± s.d. ($n = 3$).

©2014 Macmillan Publishers Limited. All rights reserved



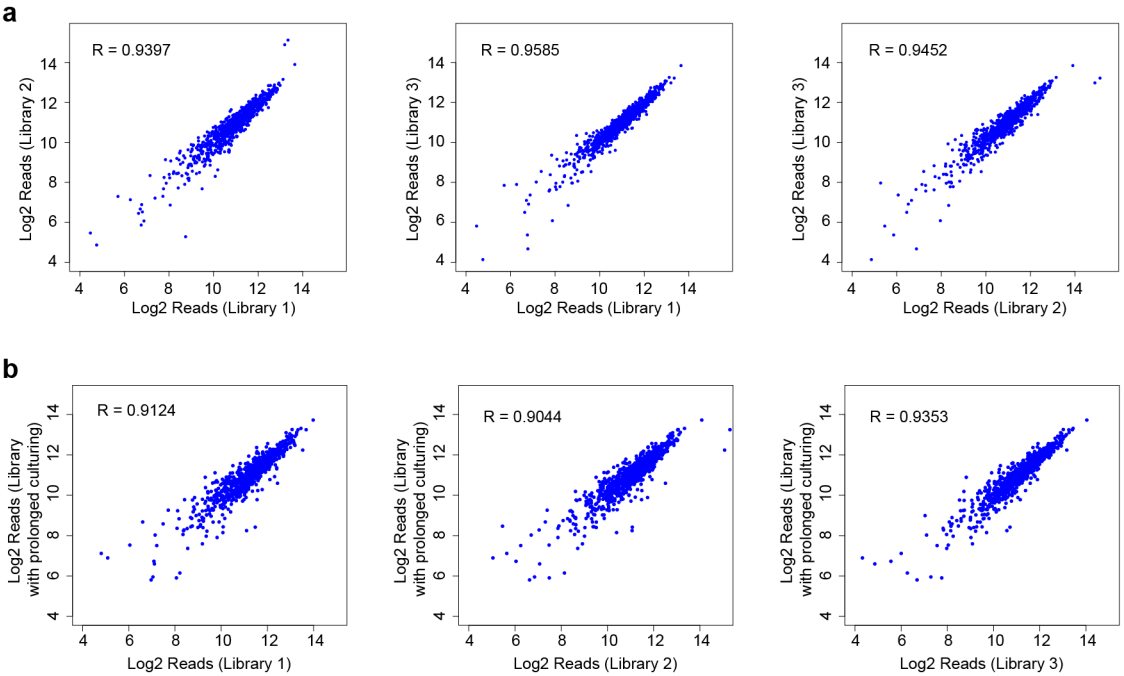
Extended Data Figure 2 | Images of cells after three rounds of toxin treatment: diphtheria toxin (7.5 ng ml^{-1}) for 60 h or PA/LFnDTA (70 ng ml^{-1}/ 50 ng ml^{-1}) for 48 h. The fluorescence signals were superimposed onto white light images. Scale bar: 100 μm.

©2014 Macmillan Publishers Limited. All rights reserved



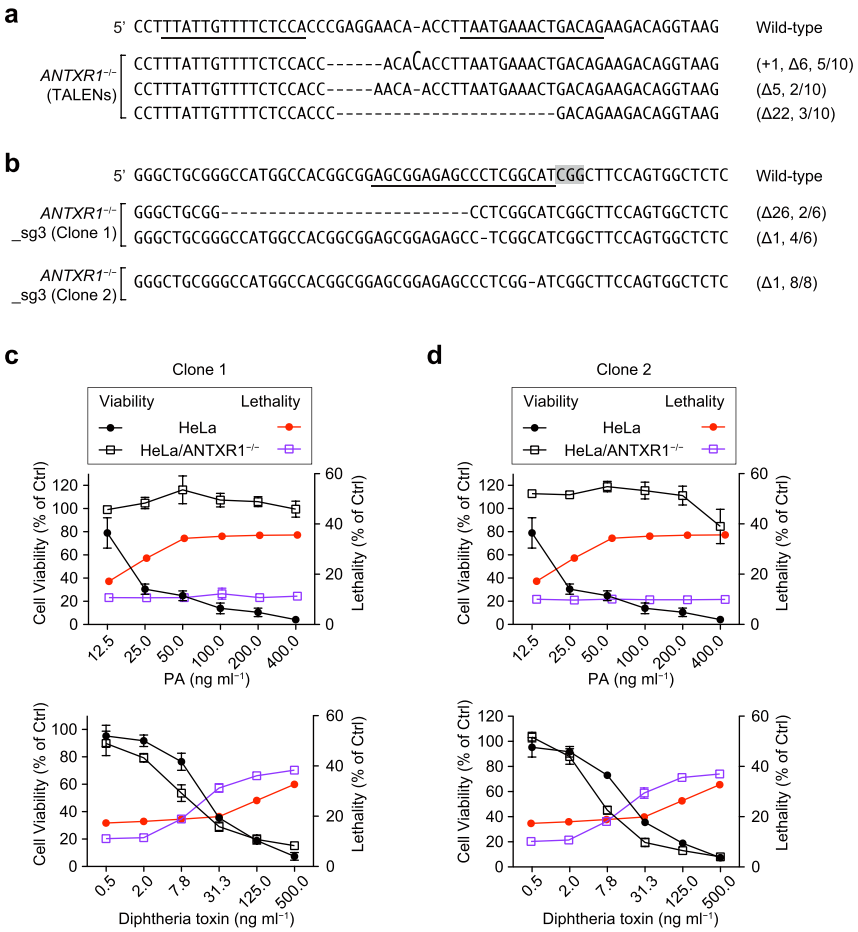
Extended Data Figure 3 | PCR amplification of the sgRNA region from the sgRNA library for deep-sequencing analysis, as indicated by electrophoresis. The genomic DNAs isolated from the indicated cells were used as templates for PCR amplification.

©2014 Macmillan Publishers Limited. All rights reserved



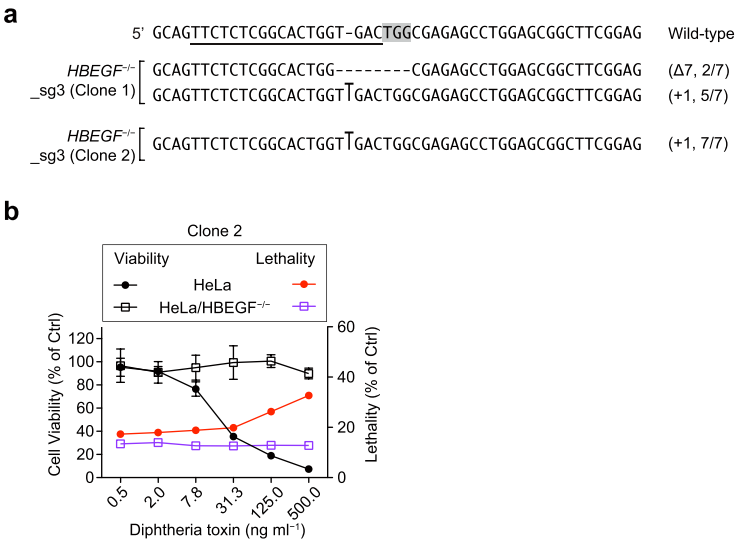
Extended Data Figure 4 | Scatter plots of $\log_2$ normalized read counts between two sgRNA libraries. **a**, Scatter plots of $\log_2$ normalized read counts between two of the three replicated sgRNA libraries. **b**, Scatter plots of $\log_2$ normalized read counts between two sgRNA libraries with or without prolonged culturing (14 days), and three replicates of the original library were used for the comparison.

©2014 Macmillan Publishers Limited. All rights reserved



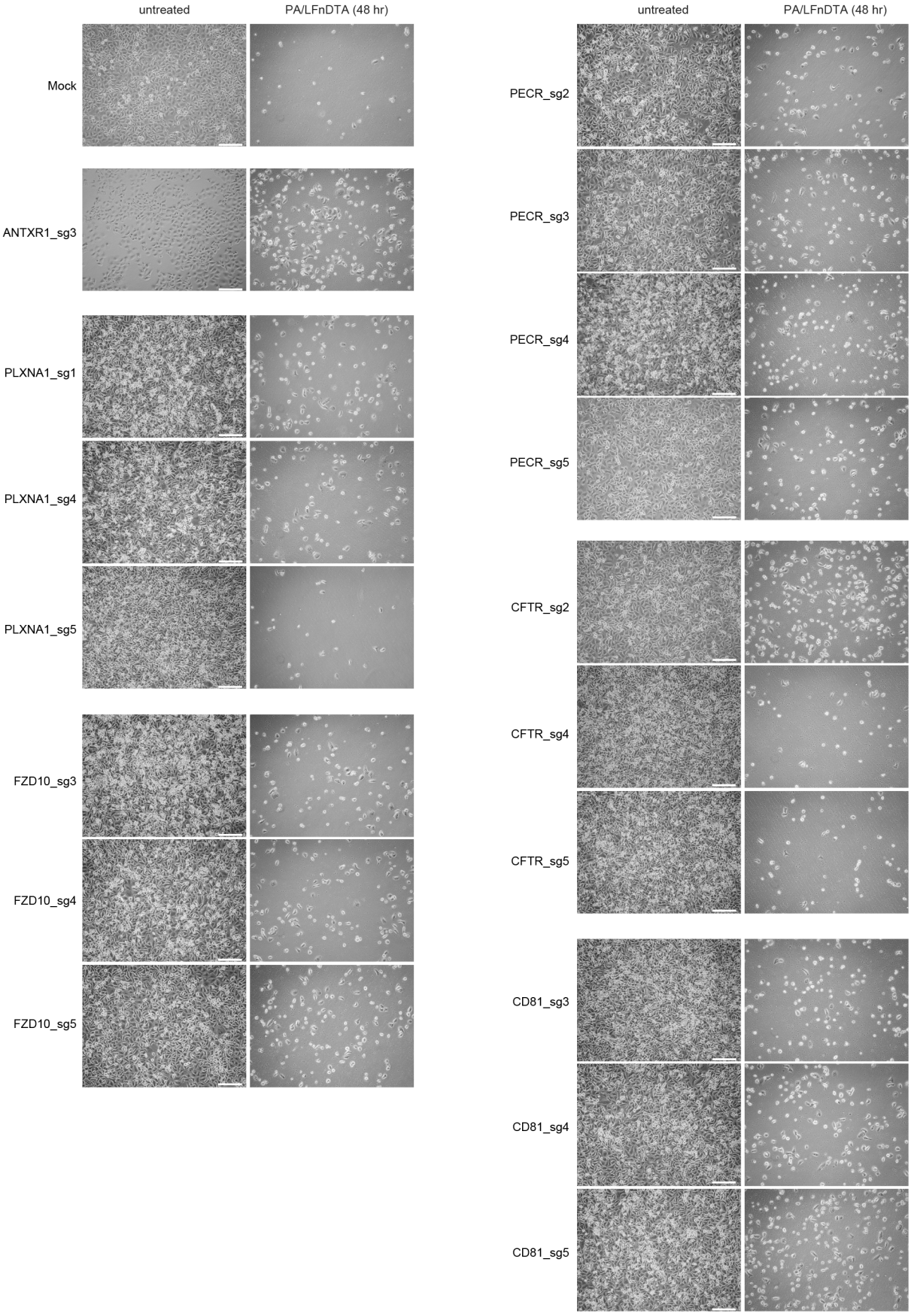
Extended Data Figure 5 | Effect of *ANTXR1* deficiency on PA/LFnDTA- and diphtheria-toxin-triggered cytotoxicity in HeLa cells. **a**, Partial sequences of *ANTXR1* gene in genome containing TALENs binding regions (underlined) and the mutated alleles from 10 randomly selected TALENs clones (in HeLa cells). The tall letter indicates insertion, and the dashes indicate deletions in this and other figures. **b**, DNA sequencing analysis of mutated alleles in the human *ANTXR1* locus. Partial coding sequences of targeted genes in the genome containing the CRISPR/Cas9 sgRNA3^{ANTXR1} binding region (underlined) and sequence analysis of the mutated alleles from two independent clones are shown. Primers used for the PCR reactions are listed in Supplementary Table 5, and the shaded nucleotides represent the PAM sequences that guide the Cas9 for DNA recognition and cleavage in this and other figures. **c**, **d**, Susceptibility assay of two HeLa *ANTXR1*$^{-/-}$ clones on PA/LFnDTA- and diphtheria-toxin-triggered cytotoxicity. The cells were treated with different doses of toxins, and the XTT cell viability assay and LDH lethality assay were performed 48 h after toxin treatment. Data are presented as the mean ± s.d. ($n = 3$).

©2014 Macmillan Publishers Limited. All rights reserved



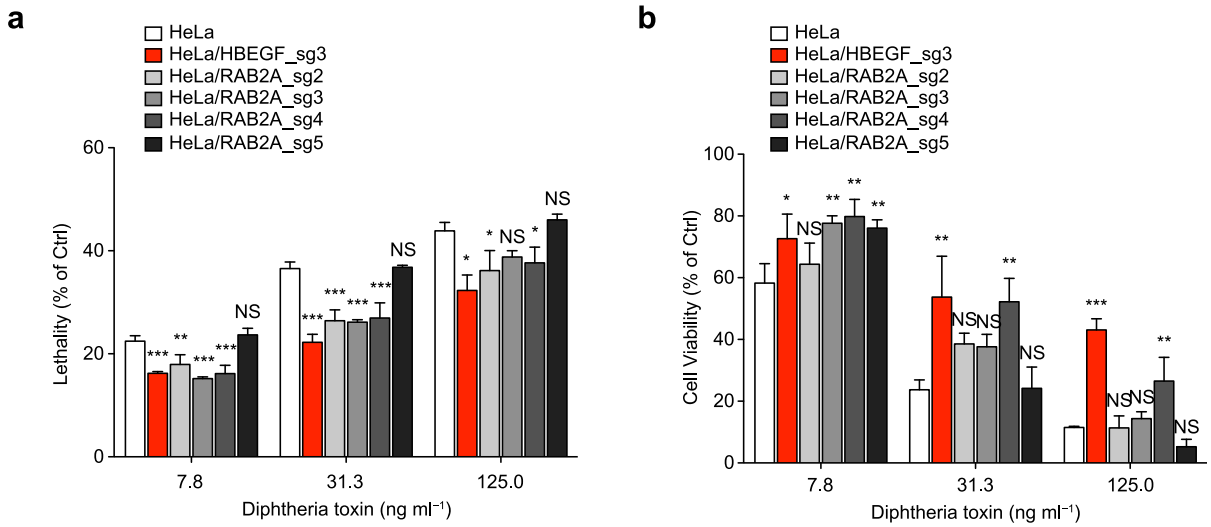
Extended Data Figure 6 | Effect of *HBEGF* deficiency on diphtheria-toxin-triggered cytotoxicity in HeLa cells. **a**, DNA sequencing analysis of mutated alleles in the human *HBEGF* locus. Partial coding sequences of targeted genes in the genome containing the CRISPR/Cas9 sgRNA3^{HBEGF} binding region (underlined) and sequencing analysis of the mutated alleles from two independent clones. **b**, Susceptibility assay of HeLa *HBEGF*$^{-/-}$ clone 2 on diphtheria-toxin-triggered cytotoxicity. The cells were treated with different doses of toxins, and the XTT cell viability assay and LDH lethality assay were performed 48 h after toxin treatment. Data are presented as the mean ± s.d. ($n = 3$).

©2014 Macmillan Publishers Limited. All rights reserved



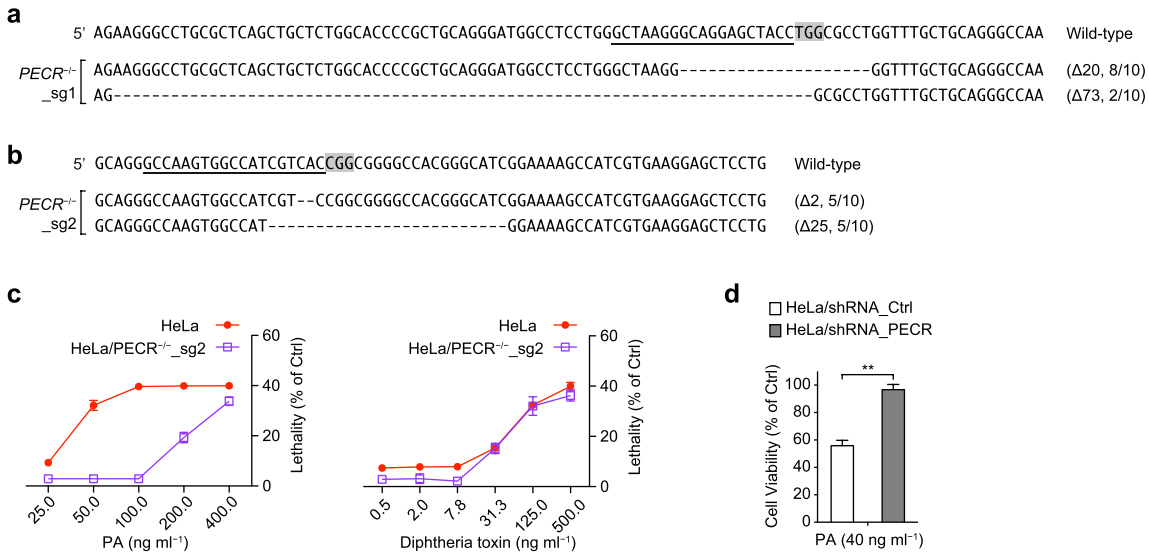
Extended Data Figure 7 | Light microscopic images of HeLa$_{OC}$-SC cells transfected with indicated sgRNA constructs. Four days after transfection, cells were incubated with PA (70 ng ml^{-1}) plus LFnDTA (50 ng ml^{-1}) for 48 h. The representative images for every condition (six images for each) are shown. Scale bar, 200 μm.

©2014 Macmillan Publishers Limited. All rights reserved



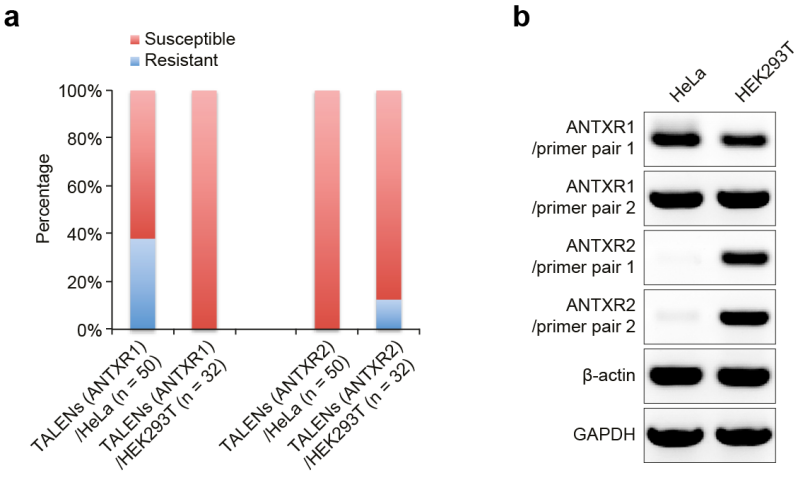
Extended Data Figure 8 | Effects of *RAB2A* sgRNAs on diphtheria-toxin-triggered cytotoxicity in HeLa cells. $HeLa_{OC}$-SC cells were transfected with indicated sgRNA constructs. Four days after transfection, the LDH lethality assay (**a**) and the XTT cell viability assay (**b**) were performed 48 h after toxin treatment. HBEGF_sg3, labelled in red, was used as the positive control. The data are the mean ± s.d. ($n = 3$); $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; NS, not significant; Dunnett's multiple comparison test, one-way ANOVA.

©2014 Macmillan Publishers Limited. All rights reserved



Extended Data Figure 9 | Effect of *PECR* deficiency on PA/LFnDTA- and diphtheria-toxin-triggered cytotoxicity in HeLa cells. **a**, **b**, DNA sequencing analysis of mutated alleles in the human *PECR* locus. Partial coding sequences of targeted genes in the genome containing the CRISPR/Cas9 sgRNA1^{PECR} (**a**) and sgRNA2^{PECR} (**b**) binding region (underlined) and sequencing analysis of the mutated alleles from two independent clones are shown. **c**, Susceptibility assay of HeLa *PECR*$^{-/-}$_sg2 clone on PA/LFnDTA- and diphtheria-toxin-triggered cytotoxicity. The cells were treated with different doses of toxin, and the LDH lethality assay was performed 48 h after toxin treatment. Data are presented as the mean ± s.d. ($n = 3$). **d**, Effect of *PECR*-targeting shRNA smart pool on HeLa cells to PA/LFnDTA toxicity. Three different shRNAs (TRCN0000046538: 5′-CCGGGCAGTGGAATACGGATCAATTCTC GAGAATTGATCCGTATTCCACTGCTTTTTG-3′, TRCN0000046539: 5′-CCGGCCTCACCAAATCTTTAGCTTTCTCGAGAAAGCTAAA GATTTGGTGAGGTTTTTG-3′, and TRCN0000046541: 5′-CCGGCCCGCTAAACGAATTGGTGTTCTCGAGAACACCAATTC GTTTAGCGGGTTTTTG-3′) from MISSION shRNA library (Sigma, Inc.) were chosen, and the MISSION shRNA control vector (SHC002) was used, which contains a sequence that engages with RISC without targeting any known mammalian gene. After transient transfection, 1.5 μg ml^{-1} puromycin was added to obtain the stable shRNA expression cells. The cells were treated with PA (40 ng ml^{-1}) plus LFnDTA (50 ng ml^{-1}), and the XTT cell viability assay was performed 48 h after toxin treatment. Data are presented as the mean ± s.d., $n = 3$, $^{**}P < 0.01$; *t*-test.

©2014 Macmillan Publishers Limited. All rights reserved



Extended Data Figure 10 | Effects of *ANTXR1* and *ANTXR2* gene expression on the susceptibility of HeLa and HEK293T cells to PA/LFnDTA. **a**, Effects of *ANTXR1* and *ANTXR2* gene knockout on the susceptibility of the indicated cells to the treatment of PA and LFnDTA. Details of the TALENs constructs used for gene targeting have been previously described[16]. Cell clones transfected with TALENs were randomly picked and cultured with PA (100 ng ml^{-1}) and LFnDTA (50 ng ml^{-1}) for 3 days to determine their sensitivity to toxins. **b**, Level of *ANTXR1* and *ANTXR2* transcripts in HeLa and HEK293T cells was measured by RT–PCR, and β-actin and GAPDH were included as internal controls. To ensure accuracy, two pairs of primers were used to measure *ANTXR1* and *ANTXR2* transcript levels, respectively. The primers used in PCR reactions were: for *ANTXR1*, primer pair 1 5′-GGAAACGGCTTCCGACATGC-3′and 5′-TGTTGTGGGGCCTGATG CAA-3′, primer pair 2 5′-GCCCGGATTGCGGACAGTAA-3′ and 5′-GCTGACCTGGAGTGCAGCTT-3′; for *ANTXR2*, primer pair 1 5′-TTGGACGGTCTGGTGCCATC-3′ and 5′-TCCTCCCCCACACA GACACT-3′, and primer pair 2 5′-ATGTGGTGGTTTTGGCCCCT-3′ and 5′-GTGTGGGTTTGGGTCGAGGT-3′; for β-actin, 5′-AGAGCT ACGAGCTGCCTGAC-3′ and 5′-AGCACTGTGTTGGCGTACAG-3′; for GAPDH, 5′-ACCACAGTCCATGCCATCAC-3′ and 5′-TCCACCACCC TGTTGCTGTA-3′.

©2014 Macmillan Publishers Limited. All rights reserved