Published in final edited form as:

J Biotechnol. 2014 November 10; 189: 1–8. doi:10.1016/j.jbiotec.2014.08.033.

Improving the specificity and efficacy of CRISPR/CAS9 and gRNA through target specific DNA reporter

Jian-Hua Zhang^{1,§}, Mritunjay Pandey¹, John F. Kahler¹, Anna Loshakov¹, Benjamin Harris¹, Pradeep K. Dagur², Yinyuan Mo³, and William F. Simonds¹

¹Metabolic Diseases Branch, Bldg. 10 Rm 8C-101, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

²Flow Cytometry Core Facility, Bldg. 10 Room 8C-104, National Institute of Heart Lung and Blood Diseases, National Institutes of Health, Bethesda, MD 20892

³Department of Pharmacology and Toxicology, Cancer Institute, The University of Mississippi Medical Center, 2500 N State St., Room G651-3, Jackson, MS 39216

Abstract

Genomic engineering by the guide RNA (gRNA)-directed CRISPR/CAS9 is rapidly becoming a method of choice for various biological systems. However, pressing concerns remain regarding its off-target activities and wide variations in efficacies. While next generation sequencing (NGS) has been primarily used to evaluate the efficacies and off-target activities of gRNAs, it only detects the imperfectly repaired double strand DNA breaks (DSB) by the error-prone non-homologous end joining (NHEJ) mechanism and may not faithfully represent the DSB activities because the efficiency of NHEJ-mediated repair varies depending on the local chromatin environment. Here we describe a reporter system for unbiased detection and comparison of DSB activities that promises to improve the chance of success in genomic engineering and to facilitate large scale screening of CAS9 activities and gRNA libraries. Additionally, we demonstrated that the tolerances to mismatches between a gRNA and the corresponding target DNA can occur at any position of the gRNA, and depend on both specific gRNA sequences and CAS9 constructs used.

Keywords

CRISPR; CAS9; gRNA; EGFP reporter; Efficacy; Specificity

[§]Corresponding author: Jianhua Zhang, PhD, National Institutes of Health, Bldg. 10 Room 8C-205, 10 Center Dr. MSC 1752, Bethesda, MD 20892-1752; Tel: 301-451-1850, Fax: 301-402-0374, jianhuaz@mail.nih.gov.

Conflict of Interest

None of the authors has a conflict of interest that might prejudice the impartiality of the research reported here.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

The target specific genomic engineering through gRNA-directed Type II CRISPR/CAS9 complex has been successfully applied in various biological systems including human, mouse, rat, fly, worm, zebrafish and more recently monkey (1–12). However, the recognition of off-target activities and variations in efficacies indicate that the potentials of the CRISPR/CAS9 system may not be fully unleashed until these problems are adequately addressed (13–15). Although there are several bioinformatic tools available for gRNA design and selection in order to increase the gRNA specificity and enhance the DSB efficacy, the proof of the prediction accuracy is in fact has to be after the completion of an actual experiment by the mismatch nuclease activity assay or NGS of the entire or selected potential off-target sites of a targeted genome. The lack of a simple method to validate the DSB specificity and efficacy of a specific gRNA and CAS9 construct prior an actual experiment has been causing uncertainties, which in many cases might be unacceptable such as in creation of animal models and clinical applications. Additionally, the NGS and mismatch-based nuclease activity assays rely on the indels caused by NHEJ mechanism that varies significantly in efficiency among different chromatin locations, therefore may not faithfully reflect the DSB activities of either CAS9 or gRNA. Recently, GFP was elegantly used as a reporter to evaluate DSB efficiencies of gRNAs against a cloned fragment of target DNA by rescuing a disrupted GFP with a functional copy of the GFP gene through homologous recombination (16). However the DSB results might be biased by the homologous recombination efficiencies. We have developed a target DNA reporter system that can be used for unbiased evaluation of specificities and efficacies of the gRNA guided CAS9 activities. Using this simple reporter system the most favorable combinations of a gRNA and CAS9 construct can be selected *in vivo* before embarking on the actual experiments.

2. Materials and Methods

2.1. Plasmids

pEGFP-C1 plasmid was purchased from Clontech. The following plasmids were purchased from Addgene: hCAS9 (41815), px330-CAS9 (42230), gRNA_Cloning vector (41824), gRNA_AAVS1-T1(41817), gRNA_AAVS1-T2(41818), gRNA_GFP-T1 (41819) and gRNA_GFP-T2 (41820). The pW25 plasmid was a gift from Dr. Kent Golic (University of Utah). The myc-CAS9 was created by replacing 3xFlag with Myc tag. The gRNA_Pfb was created by inserting an exon2 specific protospacer DNA of the Parafibromin gene, *CDC73*, into the gRNA_cloning vector. The psAAVS1-T2 EGFP reporter was created by inserting the target psAAVS1-T2 and a PAM (TGG) sequence flanked by two BsaI restriction sites (CTGTATgagaccGGGGCCACTAGGGACAGGATTGGggtctcTGTTT) after the 568th nt between the CMV promoter and the EGFP reporter through site directed mutagenesis (Bioinnovatise Inc.). The two BsaI sites are on each complementary strand of DNA so that after the BsaI digestion the vector would have non-compatible sticky ends with 4 nt overhangs (Fig. 1c). An internal BsaI site of the pEGFP-C1 vector at positions 3750–3755 nt was removed by changing the C to T at position 3755 in order to facilitate subcloning at the newly inserted BsaI sites. All subsequent and mutant psDNA constructs for specific

psDNAs reporters were created using this psDNA plasmid backbone by directly ligating a double strand oligonucleotides as shown in Fig. 1c into the BsaI linearized psDNA plasmid using the primer pair below:

forward, TGTANNNNNNNNNNNNNNNNNNNNNN and

reverse, AAACNNNNNNNNNNNNNNNNNNNNNN.

Briefly, the forward and reverse primers were mixed (5µl of 10 µM each), heated at 95 °C for 5 min, and incubated at room temperature for 20 min. Transfer 4µl of the cooled primer mixture to a microtube containing 10 ng of BsaI linearized psDNA EGFP reporter plasmid and 5µl of 2x T4 ligase buffer. Mix well by vortexing and then add 1µl of T4 DNA ligase (New England Biolabs) and incubate at room temperature for 1 hr. Mix 5µl of the ligation mix with 50 µl of competent E. coli (Top10 Invitrogen) and incubate on ice for 30min. Heat the bacteria and DNA mixture at 42 °C for 40 sec and transfer it onto ice for 2 min. Add 250 µl of SOC medium (Invitrogen) and incubate on a 37 °C heating block with shaking (>700 RPM) for 1hr. Plate the bacteria out on Kanamycin containing LB agar plate (KD medical).

The gRNA constructs were generated by either gene synthesis or the cloning method using the gRNA_cloning vector and the recommended protocol at Addgene: <http://www.addgene.org/static/data/93/40/adf4a4fe-5e77-11e2-9c30-003048dd6500.pdf> or by PCR using the primer pair:

forward, NNNNNNNNNNNNNNNNNNNNGTTTATAGAGCTAGAAATAGCAAG
and

reverse, NNNNNNNNNNNNNNNNNNNNGGTGTTTCGTCCTTTCCA

followed by the Cold Fusion method (SBI Inc). The full list of the plasmids created in this study is shown in Table 1.

2.2. Mammalian cell culture and transfection

HEK293T cells (ATCC) were grown at 37 °C with 5% CO₂ in DMEM media (Invitrogen) supplemented with 10% heat inactivated fetal bovine serum (Germini). DNA transfection of overnight cell cultures was conducted using the Omni UltraTransfect™ Transfection reagent (Dbio) following the manufacturer's protocol. In order to minimize variations, master mixtures were prepared for common components in any given reactions so that the tested parameters are the only variables. A typical transfection for a 24well plate format is described below: Seeding the cells the day before transfection at ~6 x 10⁴ cells/well in 500 µl of culture medium above. On the day of transfection (the number of cells is ~1.6 x 10⁵ cells/well), for CAS9 dosage experiments, adding the following to a sterilized microcentrifuge tube: 50 µl x N+1 of Opti MEM medium (N=number of wells to be transfected), 5 ng x N+1 of the psDNA EGFP reporter plasmid and 0.2 µg x N+1 of the corresponding gRNA plasmid, mix well by vortexing for 5 second and aliquot 50 µl of the mixture into N microcentrifuge tubes. Add various amount of hCAS9 into each microcentrifuge tube and vortex for 5 sec. The total amount of DNA was kept constant between different reactions by using comparable amount of the irrelevant plasmid pW25. Add 0.7 µl of Omni Ultra-Transfect™ Transfection reagent to each microcentrifuge tube and vortex for 5 sec. Briefly spin to collect the contents to the bottom and incubate at room temperature for 15 min. Add

each transfection mixture drop wise to each well and mix by tilting (not rotating) the plate several times at different orientations. The cells are incubated for 24–48 hrs before their GFP expressing cells were measured. For gRNA dosage experiments, the same process is conducted except premixing the psDNA EGFP reporter plasmid with hCAS9 plasmid and varying the gRNA dosage for each reaction.

2.3. Fluorescence measurement and data analysis

The GFP expressing cells were analyzed 24–48 hr post transfection by one or more of the following methods.

- a. **Flow cytometry analysis:** Flow cytometry was used for the analysis of positivity and intensity of GFP expression in transfected cells as per various experimental conditions. Cells were prepared as follows: For 24 well plate, 100 µl/well of 5% Trypsin-EDTA (Invitrogen) was added to each well and incubated at 37 °C for 5–10 min to ensure complete detachment of the cells from plate. Then 200 µl of complete DMEM medium was added to each well and the cells were dissociated by pipetting up and down 10 times. The completely suspended cells were transferred to a 300 µl tube and were acquired on LSRII (BD Biosciences, USA) equipped with 407, 488, 532, and 633 LASER lines using DIVA 6.1.2 software. Similarly, for 96 well plate, 50 µl/well of 5% Trypsin-EDTA was added to each well and incubated at 37 °C for 5–10 min to completely detach cells. After adding 150µl/well of the complete DMEM culture medium, the cells were dissociated using a multi-Channel pipettor for 10 times. Before analyzing the samples, volume was adjusted to 250 µl with PBS for each well and the plate was acquired on LSRII and High throughput sampler system (HTS) using DIVA 6.1.2 software. From each sample 10000 cells were acquired and % GFP positive and intensity of GFP was calculated using DIVA 6.1.2 software.
- b. **Microscopy.** The fluorescence intensities of transfected cells were examined under fluorescence microscope 24–48hr after transfection. For 24 well plate, 3 images under 10x objective lens at each side and the center of each well along an arbitrary middle line were taken to ensure fair representation of the well. For 96 well plates, one image from the center of the well was taken.
- c. **Cellometer® Vision (Nexcelom Bioscience Inc.).** Cells used for Cellometer analysis were all from 24 well plate. The cells were prepared as described for 24 well plate above and counted following the manufacturer's protocol.
- d. **Data analysis and graphing** were conducted using Prism software version 5.0c (GraphPad Software, Inc.). The percentage numbers from either flow cytometry analysis and/or Cellometer analysis were used separately or in combination after standardization to their corresponding negative controls as indicated for each graph. For statistical analysis, one-way ANOVA were used unless otherwise stated.
- e. **Protein sequence analyses** were performed using the Isoelectric, Helical wheel, Pepplot, and Pltot structure software available at <http://www.helix.nih.gov>

3. Results

3.1. Construction of the target protospacer DNA reporter plasmid

The CRISPR/CAS9 system provides an efficient and powerful tool for genomic engineering of biological systems. However the lack of a simple platform to unbiasedly evaluate the efficacies of a gRNA and/or CAS9 construct creates uncertainties before and impacts the trouble shooting after an actual experiment. Toward this end, we developed an EGFP reporter based on the hypothesis that insertion of a target protospacer DNA (psDNA) and PAM (protospacer adjacent motif) sequence between a promoter and a reporter gene might be used to directly assess the DSB activities of a CRISPR/CAS9 and gRNA complex by measuring the reduction of the reporter expression resulting from CAS9-mediated DNA breaks of the target psDNA (Fig. 1A). To test this hypothesis, we first validated that the CMV-driven EGFP gene expression of the unmodified pEGFP-C1 plasmid was effectively reduced in human 293T cells co-transfected with hCAS9 and the gRNA_GFP-T2 targeting the EGFP gene (Fig. 1B). We then created the psAAVS1-T2 EGFP reporter plasmid by inserting the gRNA_AAVS1-T2 target protospacer DNA (3) between the CMV promoter and the EGFP gene (Fig. 1C, and Table 1). To facilitate other psDNA cloning, two BsaI sites were inserted flanking the psAAVS1-T2 sequence so that it can be easily replaced with any psDNA flanked by the compatible ends.

3.2. Validation of the target protospacer DNA reporter plasmid

In order to know if the psDNA reporter system is suitable for distinguishing DSB activities of different gRNAs and CAS9 constructs, we performed comparisons of DSB efficacies among three CAS9 and three gRNA constructs against the psAAVS1-T2 reporter (Fig. 1C, 2A–B). To minimize the effects of variations of transfection efficiency, we used one of the most transfectable human 293T cell line and the highly efficient transfection reagent, Omni Ultra Transfect™ (Dbio Inc). In addition, for all comparisons, we made master transfection mixes for the common components so that the variations were only from the components in question. Under these conditions, transfection efficiency was broadly uniform over a range of component concentrations (Supplemental Fig. 1). Comparison of three CAS9 constructs revealed that both gRNA_GFP-T2 and gRNA_AAVS1-T2 caused similar reduction of EGFP expressing cells when hCAS9 construct was used, but the former caused more significant reduction of EGFP expressing cells than the latter when either px330-CAS9 or myc-CAS9 were used (Fig. 2A–B). Similar results were obtained using different detection methods including flow cytometry analysis, microscopic imaging, and Cellometer analysis (Supplemental Fig. S2). Dose response studies using this reporter demonstrated that higher doses of either hCAS9 or gRNA_AAVS1-T2 increased the DSB efficacy at the target psAAVS1-T2 site (Fig. 3B and D), but also caused dose-dependent nonspecific reduction of EGFP positive cells in the non-homologous gRNA_AAVS1-T1(3) controls (Fig. 3A and C), consistent with a previous report (14). These results illustrated that the psDNA EGFP reporter system is sensitive to variations of gRNA and CAS9 constructs and dosages, suggesting that the reporter system could be used for direct comparisons of DSB activities among different CAS9 and gRNA constructs.

3.3. DSB efficacy depends on specific combinations of the CAS9 and gRNA constructs

To demonstrate the broader applicability of the reporter system, we systemically compared the efficacies of the three CAS9 constructs in combination with eight different gRNAs and the corresponding psDNA EGFP reporters (Fig. 2A and Table 1). The results revealed greater variations of DSB efficacies among different combinations of gRNA and CAS9 constructs (Fig. 4A). Comparisons between combinations of three CAS9 constructs and the three reported gRNAs (gRNA_AAVS1-T2, gRNA_GFP-T1 and gRNA_GFP-T2) (3) using the psAAVS1-T2 EGFP reporter revealed that gRNA_GFP-T1 was most effective with EGFP expressing cells at 0.5%, 4.3% and 4.6% for hCAS9, px330-CAS9, and myc-CAS9 respectively (Fig. 4A, second row of the heat map) while hCAS9 was most effective with EGFP expressing cells at 1.4%, 0.5%, and 1.6% for gRNA_AAVS1-T2, gRNA_GFP-T1 and gRNA_GFP-T2 respectively (Fig. 4A–D). The px330-CAS9 was moderately more effective than myc-CAS9 only when used with gRNA_AAVS1-T2 (Fig. 4B) and gRNA_GFP-T2 (Fig. 4D). These results confirmed that the psAAVS1-T2 EGFP reporter is capable of distinguishing the DSB activities of different gRNA and CAS9 complexes. To demonstrate the applicability of the EGFP reporter system to other psDNAs as well, we tested five new psDNA EGFP reporters and the corresponding gRNAs for the tumor suppressor gene, *CDC73* and the neuronal RGS7 binding protein gene, *R7bp* (Fig. 4A). The gRNA_Pfb had lower DSB activities for all three CAS9 constructs with 8.3%, 10.0%, and 14.0% EGFP positive cells for hCAS9, px330-CAS9, and myc-CAS9 respectively (Fig. 4A and E). Interestingly, the gRNA_R7bp1 and 2 were most effective when co-transfected with px330-CAS9 (0.1–0.2% of EGFP expressing cells) (Fig. 4A and F–G) while gRNA_R7bp3 and 4 were most effective when co-transfected with hCAS9 (1.3–3.2% EGFP expressing cells) (Fig. 4A and H–I). The myc-CAS9 construct was less effective for all four gRNA_R7bp exon2 constructs tested with EGFP expressing cells from 6.0% to 9.5%. These results are in line with the initial observations that both gRNA sequence and CAS9 N- and/or C-terminal modification impact their DSB efficacies.

3.4. DSB specificity is determined by both CAS9 and gRNA constructs

The tolerance to mismatches between gRNA and the target psDNA has been associated with the off-target activities of the CRISPR/CAS9 system. The PAM and its proximal upstream 11–12 nts are considered invariable in some studies, but in others tolerance to changes in these regions were also reported (13,14,17–19). To address this discrepancy we created several mutant psDNA EGFP reporters for psPfb (psPfb_M1-2, psPfb_M10-11, and psPfb_M19-20) and psR7bp3 (psR7bp3_M1-2, sR7bp3_M10-11, and psR7bp3_M19-20) (Table 1). Comparisons of DSB activities between different combinations of CAS9 and gRNA constructs against these mutant psDNA EGFP reporters revealed that the DSB specificity is primarily determined by gRNA sequences but CAS9 constructs also played a role (Fig. 5A). The psPfb_M10-11 and M19-20 completely abolished the gRNA_Pfb guided DSB activities for all three CAS9 constructs while psPfb_M1-2 only diminished the DSB activity of myc-CAS9 (Fig. 5A top 5 rows, 5B, and Supplemental Fig. S3). In contrast, the psR7bp3_M1-2 and M10-11 completely diminished the DSB activities of all three CAS9 constructs while psR7bp3_M19-20 only reduced the DSB activities for hCAS9 and myc-CAS9, but had no significant effect on px330-CAS9 (Fig. 5A bottom 5 rows, 5C, and

supplemental Fig. S4). These results indicated that the tolerable mismatches between a gRNA and the target psDNA can potentially occur at any positions of a psDNA and are a function of both gRNA sequence and CAS9 N- and/or C-terminal modifications.

4. Discussion

The CRISPR/CAS9 system provides an unparalleled tool for genetic modifications of living organisms. However the potential off-target activities and unpredictable efficacies bioinformatically have been hindering the genome engineering efforts. The described psDNA EGFP reporter system provides a simple, yet powerful platform to objectively evaluate and compare the DSB activities of different CAS9 and gRNA constructs. Using this system the most favorable gRNA and CAS9 combinations for the target DNA of interest can be easily selected before embarking on costly and time consuming experiments such as creation of gene knockout animal models and gene therapy in clinical settings. It is foreseeable that adaptation of this system to a high throughput format would facilitate the large scale screening of gRNA libraries (19). We used the psDNA EGFP reporter system to successfully distinguish the DSB activities among different CAS9 constructs, which might otherwise be undetectable by other methods currently available. We found that even short modifications at the N terminus of CAS9 such as the Flag tag for px330-CAS9 and Myc tag for myc-CAS9 could significantly alter its specificity and efficacy, likely a result of the protein tags having different isoelectric points (pI of 10.1 for N-terminus of myc-CAS9 versus pI of 6.0 for px330-CAS9) and charge distributions (Plot structure) (Supplemental Table 1). Therefore the psDNA reporter system could be a handy tool in the CAS9 optimization efforts through rational design, which is likely to be further sparked by the recent resolution of the crystal structure of SpCAS9 (19).

However the psDNA EGFP reporter system should be used with caution. Firstly the DSB activities tested were against gRNA targets located on the extra-chromosomal plasmid that could be very different against the same target in the context of chromosomal DNA. That being said, if it is assumed that a gRNA-CAS9 complex which is ineffective against a relatively “naked” target on plasmid would most likely also be ineffective against the chromosomal target, then this method should allow the easy screening and weeding out of the ineffective gRNAs and/or CAS9 constructs. This feature is particularly suitable for testing the potential off-target activities of gRNA/CAS9 constructs since the potential off-target activities can be easily detected and compared using the reporter system. The gRNA/CAS9 constructs with high DSB activities against the target sites of choice, but no activities against the potential off-target sites should be selected. However the flip side may not be always true. It is recommended that two or more effective gRNA/CAS9 constructs should be used in real experiments. Secondly, steps should be taken to minimize transfection effects as the DSB activities were calculated based on the reduction of the GFP transfected cells. Because the variations of transfection efficiency (up to 17%, Supplemental Fig. S1) and the lack of normalizer, the highly transfectable cells (such as 293T in this study) and efficient transfection reagents should be used to minimize the transfection effects. Additionally master mixes should be prepared for common components whenever possible so that only the components in question vary.

5. Conclusion

We have developed a target protospacer DNA reporter system for testing the DSB activities of gRNA and CAS9 constructs. Using fluorescent protein as a reporter, the DSB activities of different gRNA and CAS9 constructs can be easily evaluated as compared with the NGS methods mostly used currently. Additionally, adaptation of this system to a high throughput format would greatly facilitate the screening of gRNA libraries and CAS9 constructs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the Intramural Research Programs of the National Institute of Diabetes and Digestive and Kidney Diseases, and the National Heart Lung and Blood Institute.

Abbreviations List

CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CAS9	CRISPR associated protein
gRNA	Guide RNA
NHEJ	None homologous end joining
DSB	Double strand DNA break
NGS	next generation sequencing

References

1. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012; 337:816–821. [PubMed: 22745249]
2. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013; 339:819–823. [PubMed: 23287718]
3. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. *Science*. 2013; 339:823–826. [PubMed: 23287722]
4. Cho SW, Kim S, Kim JM, Kim JS. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol*. 2013; 31:230–232. [PubMed: 23360966]
5. Chang N, Sun C, Gao L, Zhu D, Xu X, Zhu X, Xiong JW, Xi JJ. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res*. 2013; 23:465–472. [PubMed: 23528705]
6. Shen B, Zhang J, Wu H, Wang J, Ma K, Li Z, Zhang X, Zhang P, Huang X. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res*. 2013; 23:720–723. [PubMed: 23545779]
7. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*. 2013; 153:910–918. [PubMed: 23643243]
8. Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, Harrison MM, Wildonger J, O'Connor-Giles KM. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics*. 2013; 194:1029–1035. [PubMed: 23709638]

9. Friedland AE, Tzur YB, Esvelt KM, Colaiacovo MP, Church GM, Calarco JA. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat Methods*. 2013; 10:741–743. [PubMed: 23817069]
10. Bassett AR, Tibbit C, Ponting CP, Liu JL. Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Rep*. 2013; 4:220–228. [PubMed: 23827738]
11. Li JF, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J. Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat Biotechnol*. 2013; 31:688–691. [PubMed: 23929339]
12. Hu X, Chang N, Wang X, Zhou F, Zhou X, Zhu X, Xiong JW. Heritable gene-targeting with gRNA/Cas9 in rats. *Cell Res*. 2013; 23:1322–5. [PubMed: 24145754]
13. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013; 31:827–832. [PubMed: 23873081]
14. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol*. 2013; 31:839–843. [PubMed: 23934178]
15. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*. 2013; 154:1380–1389. [PubMed: 23992846]
16. Malina A, Mills JR, Cencic R, Yan Y, Fraser J, Schippers LM, Paquet M, Dostie J, Pelletier J. Repurposing CRISPR/Cas9 for in situ functional assays. *Genes & development*. 2013; 27:2602–2614. [PubMed: 24298059]
17. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*. 2013; 31:822–826. [PubMed: 23792628]
18. Pennisi E. The CRISPR craze. *Science*. 2013; 341:833–836. [PubMed: 23970676]
19. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell*. 2014; 156:935–49. [PubMed: 24529477]

Highlights

- A target DNA EGFP reporter for fast screening CRISPR/CAS9 and gRNA activities.
- The reporter can be used to pre-screen the most specific CAS9 and gRNA pairs.
- Identification of the N-terminal modification of CAS9 is critical for activities.
- Double strand DNA break specificities are both CAS9 and gRNA dependent.

Figure 1.

Characterization of the psDNA EGFP reporter system. (A) Schematic of the psDNA EGFP reporter design and function mechanism. The template pEGFP-C1 plasmid was linearized at the dual BsaI sites created between the CMV promoter (open black bar) and the EGFP reporter gene (solid black bar when unexpressible and solid green bar when expressible). A target protospacer DNA (psDNA, open red bar) with a protospacer adjacent motif (PAM, solid vertical blue bar at the right side of the psDNA) is inserted at the BsaI sites by ligation to create the psDNA EGFP reporter (solid green bar). Co-transfection of the psDNA EGFP reporter with plasmid(s) expressing a CAS9 (blue crescent) and the corresponding guide RNA (gRNA, green line with a wiggled tail) into a cell (dotted square) results in formation of the gRNA and CAS9 protein complex, which binds to the target psDNA site homologous to the gRNA sequence and generates double strand DNA break (DSB). Consequently the EGFP gene expression is diminished due to the separation of the CMV promoter from the EGFP gene (solid black bar). (B) Optimization of pEGFP-C1 dosage (ng/rx) in transfection. Right panel: percentage of fluorescent cells transfected with various amount of pEGFP plasmid as assessed by flow cytometry. The same amounts of gRNA_GFP-T2 and hCAS9 were used in all transfections. The unrelated gRNA_Pfb was used as negative gRNA control. Horizontal axis: ng pEGFP plasmid used per transfection (rx). Left panel: schematic showing the target psGFP-T2 site of the EGFP gene. (C) Creation of the psAAVS1-T2 EGFP reporter. At the dual BsaI sites created between the CMV promoter and the EGFP reporter gene (shown in solid arrows) the double strand target psDNA of the gRNA_AAVS1-T2, psAAVS1-T2 (20nt in italics) and the PAM sequence (TGG in bold)

flanked by BsaI compatible ends were inserted. The dotted double head arrow indicates the predicated double strand DNA break (DSB) site by gRNA_AAVS1-T2 guided CAS9 complex.

Figure 2.

Validation of the psDNA EGFP reporter system. (A) Schematic sequence alignment of px330-CAS9, hCAS9 and myc-CAS9 plasmid constructs. The black horizontal bars represent the common SpCAS9 sequences (not in scale). The 3xFlag for px330-CAS9 and the myc tag for myc-CAS9 constructs are in red and the nuclear localization signals are underlined. The identical sequences at either N or C termini of px330-CAS9 and myc-cas9 are in bold black. (B) Comparison of EGFP expressing cells transfected with equal amount of psAAVS1-T2 reporter and different combinations of gRNAs and CAS9 constructs as shown. The gRNA_AAVS1-T1 with no homologies to psAAVS1-T2 reporter served as negative control.

Figure 3.

Dosage response profiling of DSB efficacies of hCAS9 and gRNA_AAVS1-T2 against the target psAAVS1-T2 reporter. A and C used the non-homologous gRNA, gRNA_AAVS1-T1, as negative controls. Horizontal axis: μg plasmid DNA or gRNA used per transfection reaction (rx). The total amount of plasmid of DNAs was kept constant by using the comparable amount of the non-relevant plasmid, pW25, to minimize variations in transfection efficiency.

Figure 4.

Profiling of DSB efficacies for different combinations of gRNA and CAS9 constructs. (A) Heatmap of DSB activities displayed by different combinations of gRNA and CAS9 constructs (as shown) against their corresponding psDNA EGFP reporters as labeled. The psAAVS1-T2 reporter was used for gRNA_AAVS1-T2, gRNA_GFP-T1 and T2 constructs. Non-relevant plasmid (pW25) was used as negative control to show the basal level of GFP cells without CAS9 activities. Green, yellow, and red colors represent low, medium and high DSB activities respectively. The numbers in each colored cells represent the average percentages of GFP positive cells counted by flow cytometry analysis from three or more independent biological experiments. (B–D) The psAAVS1-T2 EGFP reporter was used. (E–I) Other psDNA EGFP reporters used: psPfb (E), psR7bp1 (F), psR7bp2 (G), psR7bp3 (H), and psR7bp4 (I). Comparisons with statistically significant differences at $p < 0.05$ level were labeled with (*) and at $p < 0.01$ level were labeled with (***)

Figure 5.

Profiling of DSB specificities for different combinations of gRNA and CAS9 constructs. **(A)** Heatmap of the potential off-target DSB activities displayed by combinations of either gRNA_Pfb (top five rows) or gRNA_R7bp3 (bottom five rows) with different CAS9 constructs and their corresponding target psDNAs as shown. The mutated nucleotides of each psDNA are in red small letters at the positions as labeled. Non-relevant plasmid, pW25 was used in control reactions. Green, yellow, and red colors represent low, medium and high DSB activities respectively. The numbers in each colored cells represent the average percentages of GFP positive cells counted by flow cytometry analysis from three or more independent biological experiments. **(B)** Effects of psDNA mutations on DSB efficacies of gRNA_Pfb and different CAS9 constructs. **(C)** Effects of psDNA mutations on DSB efficacies of gRNA_R7bp3 and different CAS9 constructs. The suffixes M1-2, M10-11, and M19-20 refer to the mutated nt positions of a psDNA respectively. Comparisons with statistically significant differences at $p < 0.05$ level were designed as (*) and at $p < 0.01$ level were designed as (**).

Table 1

gRNAs and primers used for generation of psDNA EGFP reporters¹

Name of EGFP reporter	Forward	Reverse	Corresponding gRNA used
psAAVSI1-T2	CTGTATgagaccGGGGCCCACTAGGGACACAGGATTTGGgggtctctTGTTT	GACATActctggCCCCGGTGATCCCTGTCTCTAAACCccagagA	gRNA_AAVSI1-T2: GGGGGCCCACTAGGGACACAGGA
psPib ³	GGTTTAGTGAAACCGTCAGATCCTGTATGAGACCGCTGCACGTCGGGA	CCAAATCAGCTTGGCAGTCTAG	gRNA_Pib: GCTGCACGTCGGGACATAAAA
psPib_M1-2	TGTAATTTGCACGTCGGGACATAAACCCAGG	AAACCCCTGTTTATGTCCGACGTGCAAT	
psPib_M10-11	TGTAGCTGCACGTTAGACATAAACACAGG	AAACCCCTGTTTATGTCTAAACGTGCAGC	
psPib_M19-20	TGTAGCTGCACGTCGGGACATAAAGTAGG	AAACCCCTACTTATGTCCGACGTGCAGC	
psR7bp1	TGTAGCTCCTTAGCTGGAGATTTGGGG	AAACCCCCCAAAATCTCCAGCTAAGGAGC	gRNA_R7bp1: GCTCCTTAGCTGGGAGATTT
psR7bp3	TGTAGATAAGTTTCTGATAAGTGTTGG	AAACCCCAACACTTATCAGAGAAACTTATC	gRNA_R7bp3: GATAAGTTTCTGATAAGTG
psR7bp3_M1-2	TGTACTTAAAGTTTCTGATAAGTGTTGG	AAACCCCAACACTTATCAGAGAAACTTAAAG	
psR7bp3_M10-11	TGTAGATAAGTTTGTAGATAAAGTGTTGG	AAACCCCAACACTTATCTCAAACTTATC	
psR7bp3_M19-20	TGTAGATAAAGTTTCTGATAAAGTCTATGG	AAACCCCACTGACTTATCAGAGAACTTATC	
psR7bp2	TGTAGAGCTGTGTTATTTGGGATGCTAGG	AAACCCCTAGCATCCCAATAACACAGCTC	gRNA_R7bp2: GAGCTGTGTTATTTGGGATG
psR7bp4	TGTAGATTTGTAGAGAAATGTTCTGAGG	AAACCCCTCAGAACATTCTCTACAAATC	gRNA_R7bp4: AGATTTGTAGAGAAATGTTCT

¹ psDNA sequences are italicized. The BsaI sites are in lower letters. The mutated nucleotides are underlined.

² Sequences used for site-directed insertion by Bioinnovatise Inc.

³ Primers for PCR amplification and cold fusion method.