Characterization of Cas9–Guide RNA Orthologs

Jonathan L. Braff,[1,4] Stephanie J. Yaung,[1,2,3,4] Kevin M. Esvelt,[1] and George M. Church[1,2,5]

[1]*Wyss Institute for Biologically Inspired Engineering, Harvard Medical School, Boston, Massachusetts 02115;*
[2]*Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115;* [3]*Program in Medical*
Engineering and Medical Physics, Harvard-MIT Division of Health Sciences and Technology, Massachusetts
Institute of Technology, Cambridge, Massachusetts 02139

In light of the multitude of new Cas9-mediated functionalities, the ability to carry out multiple Cas9-enabled processes simultaneously and in a single cell is becoming increasingly valuable. Accomplishing this aim requires a set of Cas9–guide RNA (gRNA) pairings that are functionally independent and insulated from one another. For instance, two such protein–gRNA complexes would allow for concurrent activation and editing at independent target sites in the same cell. The problem of establishing orthogonal CRISPR systems can be decomposed into three stages. First, putatively orthogonal systems must be identified with an emphasis on minimizing sequence similarity of the Cas9 protein and its associated RNAs. Second, the systems must be characterized well enough to effectively express and target the systems using gRNAs. Third, the systems should be established as orthogonal to one another by testing for activity and cross talk. Here, we describe the value of these orthogonal CRISPR systems, outline steps for selecting and characterizing potentially orthogonal Cas9–gRNA pairs, and discuss considerations for the desired specificity in Cas9-coupled functions.

INTRODUCTION

As opposed to large multisubunit complexes of Type I and III CRISPR systems, a single Cas9 protein in the Type II system can carry out the prokaryotic adaptive immune response to foreign DNA. In native prokaryotic Type II CRISPR systems, transcribed arrays are processed into CRISPR RNAs (crRNAs) that form a complex with Cas9 and a *trans*-activating RNA (tracrRNA) (Deltcheva et al. 2011). The crRNA guides Cas9 to double-stranded DNA sequences called protospacers that match the sequence of the spacer and are flanked by a protospacer-adjacent motif (PAM) unique to the CRISPR system (Mojica et al. 2009). If spacer–protospacer base-pairing is a close match, Cas9 cuts both strands of DNA.

In June of 2012, Jinek et al. first described a functional chimera of crRNA and tracrRNA called a guide RNA or gRNA (Jinek et al. 2012). The introduction of the gRNA format simplified the Type II CRISPR system from four components (target, tracrRNA, crRNA, and Cas9) to three and sparked rapid technology development in genome engineering. In December of that year, simultaneous publications from the Church and Zhang laboratories showed the chimera as a tool for multiplexable genome engineering in human cells (Cong et al. 2013; Mali et al. 2013c).

To date, numerous Cas9-mediated functions and configurations for implementation have been presented. Besides the initial applications in targeted genome editing in bacteria (Jiang et al. 2013) and eukaryotes (Cong et al. 2013; Mali et al. 2013c), Cas9 has been used for programmable activation of

Cite this introduction as *Cold Spring Harb Protoc*; doi:10.1101/pdb.top086793

endogenous genes in bacteria (Bikard et al. 2012; Qi et al. 2012) and eukaryotes (Gilbert et al. 2013; Maeder et al. 2013; Mali et al. 2013a; Perez-Pinera et al. 2013). Cas9 can be used for epigenome regulation and visualization (Chen et al. 2013), as well as for targeted RNA degradation and localization with PAM-presenting DNA oligonucleotides (O'Connell et al. 2014). Although Cas9 naturally induces double-stranded DNA breaks, Cas9 variants can be engineered to cleave only a single strand (as a nickase) or lack endonucleolytic activity (as nuclease-null); these variants can be functionalized with other domains to recruit a variety of molecular components in a defined manner. In general, gRNAs can serve as a scaffold for assembling complexes of nucleic acids and proteins. One example is transcriptional initiation, for which there are various architectures for enhanced Cas9 activators (Chakraborty et al. 2014; Konermann et al. 2014; Chavez et al. 2015).

ORTHOGONALITY VALUE

Although powerful, a single Cas9 protein is able to mediate only one activity targeted at multiple different sites; it cannot carry out a different activity at other sites. Achieving simultaneous functionalities in a single cell requires multiple Cas9–gRNA orthologs, each engineered with a custom effector domain or modality to independently target its activity to its own array of target sites.

Given the vast potential of different Cas9-mediated functions (Mali et al. 2013b), orthogonal Cas9–gRNA pairs will be critical for independent targeting such that each function exclusively responds to its own set of gRNAs. This exclusivity insulates signal transduction pathways (Podgornaia and Laub 2013); there should be minimal, ideally zero, cross talk between Cas9–gRNA orthologs used in the same cell. Well-insulated, orthogonal biological control elements should vary independently without interfering with other components (Purnick and Weiss 2009). Orthogonal systems are further insulated by PAM specificity resulting in distinct sets of targetable protospacers. The properties of orthogonal Cas9 systems include the Cas9–gRNA interaction specificity as well as target recognition using distinct PAM sequences.

Identifying Putatively Orthogonal Cas9 Proteins

Putatively orthogonal Cas9 proteins can be selected by examining and identifying divergent repeat sequences. Tools like CRISPRfinder (Grissa et al. 2007a) and CRISPRdb (Grissa et al. 2007b) enable identification of CRISPR arrays with their constituent spacer and repeat sequences. There are also methods to experimentally validate expression and coprocessing of tracrRNAs and pre-crRNAs (Chylinski et al. 2013), and the dual RNA format can be engineered into the single gRNA format (Deltcheva et al. 2011). Candidate gRNA sequences can then be assessed to determine whether their sequences are sufficiently divergent to impart specificity to their cognate Cas9 protein. This is especially important because there is evidence for gRNA exchangeability among different Cas9s (Fonfara et al. 2013).

Characterizing Orthogonal Cas9 Proteins

PAMs are required for initial target binding, unwinding for interrogation, and subsequent cleavage of target sequences (Anders et al. 2014; Sternberg et al. 2014). PAMs are generally unique to their cognate Cas9–gRNA pair and can be identified in three ways. First, using bioinformatics tools, a multiple sequence alignment of targeted bacteriophages or plasmids can yield a consensus PAM. However, this method requires searching available sequences of phages and plasmids for matches to CRISPR spacers. It is thereby unable to predict PAMs with confidence if the availability of relevant sequences is insufficient. In addition, the bioinformatics approach may be affected by biases imparted by endogenous spacer acquisition machinery that is not generally used in engineering applications. Experimentally, in vitro cleavage assays with different plasmid substrates can produce a position weight matrix of potential PAMs, but this is relatively low-throughput compared to a library approach (Esvelt et al. 2013), which can interrogate a PAM sequence space of NNNNNNNN (corresponding to $4^8 = 65,536$ sequences). In theory, one could perform a library selection on $4^{10} = 1,048,576$ sequences,

J.L. Braff et al.

TABLE 1. Characterized PAMs for Cas9 orthologs

Cas9 system	PAM	References	Other notes
Streptococcus thermophilus CRISPR1	NNAGAAW	Horvath et al. 2008; Esvelt et al. 2013	NNAAAAW cleaved more efficiently (Fonfara et al. 2013)
Streptococcus thermophilus CRISPR3	NGGNG	Horvath et al. 2008	
Streptococcus pyogenes	NGG	Mojica et al. 2009	
Streptococcus agalactiae	NGG	Mojica et al. 2009	
Listeria monocytogenes	NGG	Mojica et al. 2009	
Streptococcus mutans	NGG	Van der Ploeg 2009	
Neisseria meningitidis	NNNNGATT	Zhang et al. 2013; Esvelt et al. 2013	
Campylobacter jejuni	NNNNACA	Fonfara et al. 2013	
Francisella novicida	NG	Fonfara et al. 2013	
Streptococcus thermophilus LMG18311	NNGYAAA	Chen et al. 2014	NNNGYAAA seems to also work
Treponema denticola	NAAAAN	Esvelt et al. 2013	

given enough sequencing depth, if the PAM is suspected to extend out to 10 bases. For more details on this approach, see Protocol: **Characterizing Cas9 Protospacer-Adjacent Motifs with High-Throughput Sequencing of Library Depletion Experiments** (Braff et al. 2016).

Because the PAM sequence (Table 1) is key to distinguishing "self" versus "nonself" DNA, characterizing the PAM is important not only for Cas9-mediated activity, but also for understanding potential off-targets (Aach et al. 2014; Zhang et al. 2014). Thus, it is an essential first step to determining the orthogonality of CRISPR systems. Many groups have studied ways to improve Cas9 specificity, as Cas9–gRNA complexes are known to tolerate up to three mismatches in their targets. Notably, specificity should be evaluated with the context of the assay in mind; there are observed differences between cleavage assay results and in vivo performance in terms of PAM preference (Fonfara et al. 2013), and off-target binding does not necessarily result in off-target cleavage (Cencic et al. 2014; Wu et al. 2014). Specificity also depends on the concentration and bioavailability of the components. Factors affecting specificity include the amount of Cas9 protein, gRNA expression, and half-life of components, as well as the target and PAM sequence accessibility (Fu et al. 2013; Hsu et al. 2013; Mali et al. 2013a; Pattanayak et al. 2013). Some approaches to increase specificity include using multiple homologs for activity, finding or engineering improved variants, harnessing the dimerization requirement of FokI (Guilinger et al. 2014; Tsai et al. 2014), and using slightly truncated gRNAs (Fu et al. 2014).

CONCLUSION

Expanding the set of characterized Cas9–gRNA orthologs increases the number of engineered functions that can be simultaneously deployed in a single cell. Investigating additional Cas9 proteins can also provide insights into the nuances of varying sensitivity, specificity, and kinetic requirements in the context of the engineering application. A suite of Cas9 homologs from which to choose would allow careful matching and optimization of these parameters for a desired application. Further, the differences between PAMs found in Cas9 homologs increase the range of targetable sequences.

Undoubtedly, Type II CRISPR systems have been widely adopted for their multiplexability and programmability. The focus has been on the performance of a single function at a large number of loci and the ability to rapidly retarget this function. The increasing number of engineered variants capable of performing distinct functions has made the ability to program multiple target action pairs a very promising avenue for technology advancement.

ACKNOWLEDGMENTS

Cold Spring Harbor Protocols
www.cshprotocols.org

REFERENCES

Aach J, Mali P, Church G. 2014. CasFinder: Flexible algorithm for identifying specific Cas9 targets in genomes. *bioRxiv* 1–8.

Anders C, Niewoehner O, Duerst A, Jinek M. 2014. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513:** 569–573.

Bikard D, Hatoum-Aslan A, Mucida D, Marraffini LA. 2012. CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe* **12:** 177–186.

Braff JL, Yaung SJ, Esvelt KM, Church GM. 2016. Characterizing Cas9 protospacer-adjacent motifs with high-throughput sequencing of library depletion experiments. *Cold Spring Harb Protoc* doi:10.1101/pdb.prot090183.

Cencic R, Miura H, Malina A, Robert F, Ethier S, Schmeing TM, Dostie J, Pelletier J. 2014. Protospacer adjacent motif (PAM)-distal sequences engage CRISPR Cas9 DNA target cleavage. *PLoS One* **9:** e109213.

Chakraborty S, Ji H, Kabadi A. 2014. A CRISPR/Cas9-based system for reprogramming cell lineage specification. *Stem Cell Reports* **3:** 940–947.

Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, P R Iyer E, Lin S, Kiani S, Guzman CD, Wiegand DJ, et al. 2015. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* **12:** 1–5.

Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, Park J, Blackburn EH, Weissman JS, Qi LS, et al. 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155:** 1479–1491.

Chen H, Choi J, Bailey S. 2014. Cut site selection by the two nuclease domains of the Cas9 RNA-guided endonuclease. *J Biol Chem* **289:** 13284–13294.

Chylinski K, Le Rhun A, Charpentier E. 2013. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol* **10:** 726–737.

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339:** 819–823.

Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E. 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471:** 602–627.

Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM. 2013. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods* **10:** 1116–1121.

Fonfara I, Le Rhun A, Chylinski K, Makarova KS, Lécrivain A-L, Bzdrenga J, Koonin EV, Charpentier E. 2013. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res* **42:** 2577–2590.

Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD. 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31:** 822–826.

Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol* **32:** 279–284.

Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, et al. 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154:** 442–451.

Grissa I, Vergnaud G, Pourcel C. 2007a. CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* **35:** W52–W57.

Grissa I, Vergnaud G, Pourcel C. 2007b. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* **8:** 172.

Guilinger JP, Thompson DB, Liu DR. 2014. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol.* **32:** 577–582.

Horvath P, Romero DA, Coûté-Monvoisin A-C, Richards M, Deveau H, Moineau S, Boyaval P, Fremaux C, Barrangou R. 2008. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J Bacteriol* **190:** 1401–1412.

Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, et al. 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31:** 827–832.

Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. 2013. RNA-guided editing of bacterial genomes using CRISPR–Cas systems. *Nat Biotechnol* **31:** 233–239.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337:** 816–821.

Konermann S, Brigham M, Trevino A. 2014. Genome-scale transcriptional activation by an engineered CRISPR–Cas9 complex. *Nature* **517:** 583–588.

Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. 2013. CRISPR RNA-guided activation of endogenous human genes. *Nat Methods* **10:** 977–979.

Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM. 2013a. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* **31:** 833–838.

Mali P, Esvelt KM, Church GM. 2013b. Cas9 as a versatile tool for engineering biology. *Nat Methods* **10:** 957–963.

Mali P, Yang L, Esvelt K, Aach J, Guell M. 2013c. RNA-guided human genome engineering via Cas9. *Science* **339:** 823–827.

Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C. 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155:** 733–740.

O'Connell MR, Oakes BL, Sternberg SH, East-Seletsky A, Kaplan M, Doudna JA. 2014. Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* **516:** 263–266.

Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. 2013. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31:** 839–843.

Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, Thakore PI, Glass KA, Ousterout DG, Leong KW, et al. 2013. RNA-guided gene activation by CRISPR–Cas9-based transcription factors. *Nat Methods* **10:** 973–976.

Podgornaia AI, Laub MT. 2013. Determinants of specificity in two-component signal transduction. *Curr Opin Microbiol* **16:** 156–162.

Purnick PEM, Weiss R. 2009. The second wave of synthetic biology: From modules to systems. *Nat Rev Mol Cell Biol* **10:** 410–422.

Qi L, Haurwitz RE, Shao W, Doudna JA, Arkin AP. 2012. RNA processing enables predictable programming of gene expression. *Nat Biotechnol* **30:** 1002–1006.

Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507:** 62–67.

Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, Goodwin MJ, Aryee MJ, Joung JK. 2014. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol* **32:** 569–576.

Van der Ploeg JR. 2009. Analysis of CRISPR in *Streptococcus mutans* suggests frequent occurrence of acquired immunity against infection by M102-like bacteriophages. *Microbiology* **155:** 1966–1976.

Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, Cheng AW, Trevino AE, Konermann S, Chen S, et al. 2014. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol* **32:** 670–676.

Zhang Y, Heidrich N, Ampattu B. 2013. Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol Cell* **50:** 488–503.

Zhang Y, Ge X, Yang F, Zhang L, Zheng J, Tan X, Jin Z-B, Qu J, Gu F. 2014. Comparison of non-canonical PAMs for CRISPR/Cas9-mediated DNA cleavage in human cells. *Sci Rep* **4:** 5405.

Characterization of Cas9−Guide RNA Orthologs

Jonathan L. Braff, Stephanie J. Yaung, Kevin M. Esvelt and George M. Church

Email Alerting Service	Receive free email alerts when new articles cite this article - **click here.**
Subject Categories	Browse articles on similar topics from *Cold Spring Harbor Protocols*. Bioinformatics/Genomics, general (189 articles) Characterization of Protein Complexes (83 articles) Computational Biology (100 articles) Molecular Biology, general (1288 articles) Proteins and Proteomics, general (573 articles) Proteomics (66 articles) RNA (317 articles) RNA, general (269 articles)