Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis

Zhonggang Houa,b,1, Yan Zhangb,1, Nicholas E. Propsona, Sara E. Howdenb, Li-Fang Chua, Erik J. Sontheimerb,2, and James A. Thomsona,c,d,2

aMorgridge Institute for Research, Madison, WI 53715; bDepartment of Molecular Biosciences, Northwestern University, Evanston, IL 60208-3500; cDepartment of Cell and Regenerative Biology, University of Wisconsin–Madison, Madison, WI 53706; and dDepartment of Molecular, Cellular, and Developmental Biology, University of California, Santa Barbara, CA 93106

Contributed by James A. Thomson, July 18, 2013 (sent for review July 1, 2013)

Genome engineering in human pluripotent stem cells (hPSCs) holds great promise for biomedical research and regenerative medicine. Recently, an RNA-guided, DNA-cleaving interference pathway from bacteria [the type II clustered, regularly interspaced, short palindromic repeats (CRISPR)-CRISPR-associated (Cas) pathway] has been adapted for use in eukaryotic cells, greatly facilitating genome editing. Only two CRISPR-Cas systems (from Streplococcus pyogenes and Streplococcus thermophilus), each with their own distinct targeting requirements and limitations, have been developed for genome editing thus far. Furthermore, limited information exists about homology-directed repair (HDR)-mediated gene targeting using long donor DNA templates in hPSCs with these systems. Here, using a distinct CRISPR-Cas system from Neisseria meningitidis, we demonstrate efficient targeting of an endogenous gene in three hPSC lines using HDR. The Cas9 RNA-guided endonuclease from N. meningitidis (NmCas9) recognizes a 5′-NNNNNGATT-3′ protospacer adjacent motif (PAM) different from those recognized by Cas9 proteins from S. pyogenes and S. thermophilus (SpCas9 and StCas9, respectively). Similar to SpCas9, NmCas9 is able to use a single-guide RNA (sgRNA) to direct its activity. Because of its distinct protospacer adjacent motif, the N. meningitidis CRISPR-Cas machinery increases the sequence contexts amenable to RNA-directed genome editing.

H
uman pluripotent stem cells (hPSCs) can proliferate indefinitely while maintaining the potential to give rise to virtually all human cell types (1). These cells are therefore invaluable for regenerative medicine, drug screening, and biomedical research. However, to realize the full potential of hPSCs, it will be necessary to manipulate their genomes in a precise, efficient manner. Historically, gene targeting in hPSCs has been extremely difficult (2). The development of zinc-finger nucleases (ZFNs) and transcription activator-like endonucleases (TALENs) has facilitated genome engineering in human pluripotent stem cells (2–5) and genome editing using Cas9 from Neisseria meningitidis, especially by obviating the need for repeated protein design/optimization. Recently, this vision has become a reality (21–31).

Type II CRISPR-Cas systems are noteworthy in that the essential targeting activities—crRNA binding, target DNA binding, R-loop formation, and double-stranded DNA cleavage—are all executed by a single polypeptide, Cas9 (32–35). In addition to crRNA and Cas9, an additional RNA, transcribing CRISPR RNA (tracrRNA), is essential for interference in bacteria (14, 32, 36) and in vitro (34, 36). The tracrRNA is partially complementary to pre-crRNA repeats, leading to the formation of duplexes that are cleaved by the host factor ribonuclease III (RNase III) (32). The type II crRNA maturation pathway was originally characterized in strains of Streplococcus pyogenes (32) and Streplococcus thermophilus (35, 36), and RNase III-catalyzed pre-crRNA processing is essential for interference in both native systems. Recent studies of a type II CRISPR-Cas locus from Neisseria meningitidis revealed an intrinsically RNase III- and processing-independent system, which nonetheless requires tracrRNA (14). Importantly, crRNA-directed DNA cleavage was reconstituted in vitro with recombinant S. pyogenes Cas9 (SpCas9) (34) or S. thermophilus Cas9 (StCas9) (33, 36). The SpCas9 in vitro system enabled the development of fused crRNA-tracrRNA chimeras called single-guide RNAs (sgRNAs) that bypass processing (34). Subsequent development of eukaryotic genome editing applications has focused on sgRNAs (21–30), although separately encoded pre-crRNAs and tracrRNAs are also effective (21).

Target cleavage by many CRISPR-Cas systems, including those from type II, require proximity to a 2- to 5-nt sequence called protospacers that direct the targeting activity of Cas9. However, not all protospacers are recognized by Cas9 (14). Importantly, crRNA-bound Cas9 can cleave DNA at spacer sequences that are not recognized by crRNA (14). This is consistent with CRISPR-Cas systems providing a genomic memory of past invasions. In addition, crRNA-bound Cas9 cleaves DNA at spacer sequences that are not recognized by crRNA (14). This is consistent with CRISPR-Cas systems providing a genomic memory of past invasions.

Significance

Genome engineering in human pluripotent stem cells holds great promise for biomedical research and regenerative medicine, but it is very challenging. Recently, an RNA-guided nucleosome system called clustered, regularly interspaced, short palindromic repeats (CRISPR)/CRISPR-associated (Cas) has been applied to genome engineering, greatly increasing the efficiency of genome editing. Here, using a CRISPR-Cas system identified in Neisseria meningitidis, which is distinct from the commonly used Streplococcus pyogenes system, we demonstrate efficient genome engineering in human pluripotent stem cells. Our study could have a tremendous impact in regenerative medicine.


Conflict of interest statement: Northwestern University (Y.Z. and E.J.S.) and Wisconsin Alumni Research Foundation (Z.H. and J.A.T.) have filed a related patent: DNA cleavage and genome editing using Cas9 from Neisseria meningitidis. Freely available online through the PNAS open access option.

See Commentary on page 15514.

1Z.H. and Y.Z. contributed equally to this work.

2To whom correspondence may be addressed. E-mail: erik@northwestern.edu or jthomson@morggridgeinstitute.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1313587110/-/DCSupplemental.
a protospacer adjacent motif (PAM) (37–40). Genome-editing applications reported thus far have focused almost exclusively on SpCas9, which has a 5′-NGG-3′ PAM. StCas9 (from the CRISPR1 locus of strain LMD-9) has also been used in eukaryotes (21), and a system has a 5′-NNAGAAA-3′ PAM (W or A or T). Eukaryotic editing capabilities will benefit from the increased frequency of target sites stemming from the development of additional Cas9s with distinct PAMs.

Targeting by sgRNAs usually relies on either of two approaches. First, double-strand break (DSB) repair by nonhomologous end joining (NHEJ) can be used to generate insertions or deletions (indels) that induce frame-shifts. Second, the addition of a homologous repair template can allow Cas9-induced DSBs or nicks to be repaired by homology-directed repair (HDR). The latter strategy is useful for making precise changes such as repairing mutations or introducing transgenes. Most studies thus far have relied on either NHEJ, or on HDR using short DNA fragments or oligos (24–26, 29, 31). Currently there is very limited information available on gene targeting using long DNA donor templates in hPSCs (23).

Here, we report the development of N. meningitidis Cas9 (NmCas9) (14) as a genome-editing platform, and its application to high-efficiency targeting of an endogenous gene in hPSCs. This system uses a 24-nt proto-spacer for targeting and requires a PAM that is different from those of SpCas9 or StCas9. We have achieved ~60% targeting efficiency with two human embryonic stem cell (hESCs) and one human induced pluripotent stem (iPS) cell line. Our work demonstrates the feasibility of using the N. meningitidis CRISPR-Cas system in genome editing in hPSCs using long DNA donor templates. This work also provides an alternative to the S. pyogenes and S. thermophilus CRISPR-Cas system and expands the genomic contexts that are amenable to RNA-directed genome editing in eukaryotes.

Results

Functional Expression of NmCas9 in Mammalian Cells. Our recent work has shown that N. meningitidis strain 8013 has a functional type II-C CRISPR-Cas system (14), and that Cas9 is the only Cas protein required for interference activity. We set out to test whether this system could be used for efficient gene targeting in hPSCs. We cloned the ORF from the 3.25-kb cas9 gene, along with a C-terminal FLAG tag, into a mammalian expression plasmid under the control of an elongation factor-1α (EF1α) promoter (Fig. 1A). This NmCas9-containing vector was transfected into 293FT cells and the expression of NmCas9 protein was analyzed by anti-FLAG Western blot. As shown in Fig. 1A, full-length NmCas9 was efficiently expressed in 293FT cells. We then assayed the nuclelease activity of NmCas9 expressed in mammalian cells by in vitro plasmid cleavage. Cell extract was prepared from 293FT cells 2 d after transfection with the NmCas9-containing vector (the same one as in Fig. 1A). We assembled cleavage reactions using cell extract, various in vitro-synthesized small RNAs, and the plasmid ptdTomato prelinearized by NdeI (Fig. 1B). tdTomato is a fusion of two copies of the dTomato gene, each of which has one consensus PAM sequence (5′-NNNGATT-3′) (blue in Fig. 1B). As shown in Fig. 1C, we achieved efficient plasmid cleavage in the presence of both tracrRNA and a cognate crRNA (Fig. 1C, lane 3). The pattern of the cleavage products was consistent with two predicted cleavage sites in the PAM-proximal regions (Fig. 1C, Right). Importantly, a noncognate crRNA (N), which contains sequences from EGFP, did not direct NmCas9-mediated cleavage (Fig. 1C, lane 4), indicating that the specificity of the NmCas9 nuclease is indeed guided by the spacer-derived sequence in crRNAs. Additionally, plasmid cleavage is deficient when tracrRNA is absent, even in the presence of a cognate crRNA (Fig. 1C, lane 5), suggesting that tracrRNA is necessary for NmCas9 function in vitro. This finding is consistent with the tracrRNA requirement for NmCas9-mediated interference in bacterial cells (14). Two Cas9 orthologs, SpCas9 and StCas9, were previously demonstrated to induce blunt DSBs in their DNA targets, between the third and fourth nucleotide counting from the PAM-proximal end of protospacers (19, 33, 34). We hypothesized that NmCas9 cleaves the DNA target in a similar way, and we tested this by mapping the NmCas9 cleavage site on ptdTomato by Sanger sequencing. Two cleavage products in Fig. 1C (the 1.5- and 2.5-kb fragments) were gel-extracted and sequenced to identify the NmCas9 cleavage sites on the sense strand and the antisense strand, respectively. As expected, NmCas9 induced a blunt-end DSB between the third and fourth nucleotides counting from the PAM-proximal end of the proto-spacer (Fig. 1D).

NmCas9 Functions in RNA-Directed Gene Disruption in hPSCs. Knowing that NmCas9, without any codon optimization, can be efficiently expressed in mammalian cells and is functional in vitro, we next tested its utility in genome editing in hPSCs. We first monitored its localization. We transfected 293FT cells with several NmCas9 constructs with various nuclear localization signal (NLS) arrangements, and analyzed NmCas9 protein localization by either GFP fluorescence or anti-HA immunostaining. NmCas9 with NLSs on both N and C termini localized efficiently to the nucleus (Fig. 2C), but NmCas9 constructs with just one NLS did not (Fig. 2A and B). In addition, the same NmCas9 construct with two NLSs also localized to the nucleus of hESCs (Fig. 2D). We noticed that in hESCs, NmCas9, without any crRNA/tracrRNA, displayed a punctate pattern similar to the organization of the nucleolus in hESCs. It is not yet clear if this phenomenon is related to the organization of the double NLS on the protein.

To test the genome-editing activity of NmCas9, we used an HESC cell reporter line that has a single copy of the dTomato fluorescent protein gene knocked into the highly expressed DNMT3b locus (H9 DNMT3b-tdTomato), leading to tdTomato...
If NmCas9 is able to introduce a DSB in the tdTomato sequence in the genome, repair by NHEJ would likely lead to indels that disrupt tdTomato expression. Accordingly, the appearance of tdTomato-negative cells would be predicted to reflect genome-editing activity.

hESCs are known to have low transfection efficiencies. To achieve maximum genome-editing efficiency in hPSCs, we assembled expression cassettes of all of the necessary components (NmCas9, tracrRNA, and crRNA) onto one single plasmid.

**Fig. 2.** NmCas9 functions in gene disruption in human ES cells. (A–C) The localization of NmCas9 with an NLS at the N terminus (A), C terminus (B), or both termini (C) was analyzed by either EGFP fluorescence (A) or anti-HA immunofluorescence (B and C) in 293FT cells. (Scale bars, 20 μm.) (D) The localization of NmCas9 with the double NLS (see C) was analyzed by anti-HA immunofluorescence in hESCs. (Scale bars, 20 μm.) (E) Design of a single plasmid used for gene editing in hESCs. (F) FACS analysis of tdTomato reporter hESC lines after electroporation of the indicated crRNA/tracrRNA/NmCas9 constructs. The number in the plot indicates the percentage of tdTomato-fluorescence-negative cells 5 d after electroporation. (G) Indels introduced by the targeting CRISPR in the tdTomato negative population (see F) were analyzed by targeted PCR amplification and sequencing. The protospacer sequence is underlined. The numbers in parentheses indicate the number of sequenced clones containing that specific indel.
FACS 4–6 d after electroporation. As shown in Fig. 2F, a subpopulation (5.4%) of tdTomato-negative cells became detectable only when a tdTomato-targeting crRNA was encoded on the plasmid. Importantly, for the control plasmid expressing nontargeting crRNA, only background levels (~0.1%) of tdTomato-negative cells appeared (Fig. 2F), likely because of the low level of spontaneous differentiation in the culture, leading to repression of the DNMT3b promoter. The increased frequency of nonfluorescent cells in the presence of the cognate crRNA suggests successful genome editing by NmCas9.

To confirm that NmCas9 introduced a DSB at the intended genomic site, we FACS-sorted the tdTomato-negative population, PCR-amplified the genomic region flanking the predicted cutting site in the 5′ copy of dTomato, cloned the resulting PCR fragments, and sequenced 22 of the resulting plasmids (selected at random). The sequencing results showed both insertions and deletions in the tdTomato sequence (Fig. 2G; only unique indels are shown) in 95% of the sequenced clones. Most importantly, all of these indels were centered around the NmCas9 cleavage site, indicating that the DSB occurred at the intended position (Fig. 2G).

A Chimeric sgRNA Is Effective for Gene Editing in hPSCs. To simplify the NmCas9 genome-editing system, we explored the possibility of substituting both crRNA and tracrRNA with a chimeric sgRNA. We fused the 5′ end of the 91-nt processed tracrRNA sequence with the 3′ end of the 48-nt mature crRNA using a 6-nt linker (Fig. S1A). This sgRNA was cloned under the control of the U6 promoter and electroporated into the H9 DNMT3b-tdTomato reporter cell line together with a plasmid expressing NmCas9. FACS analysis showed that this sgRNA indeed resulted in tdTomato-negative cells (Fig. S1B) at a level comparable to that achieved by the all-in-one plasmid expressing separate crRNA and tracrRNA (Fig. 2F). These results indicated that an sgRNA could substitute for separate crRNA and tracrRNA in directing NmCas9-mediated gene editing in hESCs.

Specificity of NmCas9 in hPSCs. We next tested the specificity of NmCas9 in mammalian cells by mutational analysis. We introduced single-nucleotide mutations at every odd-numbered position from the 5′ to the 17th nucleotide in the PAM-proximal end [spanning the cleavage site (Figs. 1D and 2G) and the functionally critical “seed sequence] of the spacer in the tdTomato-targeting crRNA construct (Fig. 3A, Upper). We then measured the ability of those constructs to give rise to tdTomato-negative cells in the H9 DNMT3b-tdTomato cell line. As shown in Fig. 3A, mutations at positions 1 through 9 led to background levels of tdTomato-negative cells, indicating that mismatches at these positions in the crRNA/target duplex are not tolerated by NmCas9. As for mutations at positions 11, 13, and 15, some tdTomato-negative cells appeared, but with an efficiency of only 10–25% of that observed with wildtype crRNA (Fig. 3A, Lower). The mismatch at position 17 was ~40% efficient as wild-type. These results imply a crRNA/target specificity comparable to that of the SpCas9 system in mammalian cells (21).

We also investigated PAM sequence requirements for NmCas9 in hESC. We designed five crRNAs that use different sequences as the PAM in the tdTomato coding region (Fig. 3B) and then tested their ability to disrupt tdTomato expression in H9 DNMT3b-tdTomato cells. Four of the sites were associated with a PAM that varied from the 5′-NNNNGATT-3′ consensus by only a single nucleotide. Only a GCTT variant site was efficiently targeted, whereas the other four variants were severely deficient (Fig. 3B, Lower). Our results show that the second nucleotide of the PAM could be tolerated, whereas a G-to-C mutation at the first position, T-to-C at the third, and T-to-G at the fourth likely render the PAM variants nonfunctional. Interestingly, C is the second most-frequent residue at the second nucleotide of the PAM in candidate bacterial protospacers (14), suggesting that GCTT might also be a natural PAM.

NmCas9 Increases Gene-Targeting Efficiency in hPSCs. We next explored whether NmCas9 can increase gene-targeting efficiency in hPSCs compared with the traditional method in which no DSB was intentionally introduced at the target site. We used a donor DNA template previously used to target the endogenous POU5F1 (OCT4) gene (6) (Fig. 4A), creating a fusion of OCT4 with EGFP. We designed the crRNA using the consensus PAM sequence located ~84 bp downstream of the OCT4 stop codon (Fig. 4A). Two hESC lines, H1 and H9, and one human iPSC cell line, iP5005 (42), were used in the experiment. After puromycin selection, we were able to obtain clones for all three cell lines when plasmid expressing the OCT4-targeting crRNA was used. Of these clones, ~60% were correctly targeted with single insertion events (Table 1), comparable to the efficiency obtained using TALENs in a previous report with the same donor DNA (6). Fluorescent images of the targeted clones revealed the expected nuclear localization of EGFP signal because of the fusion with Oct4 protein (Fig. 4B). Southern blots using a probe outside the targeting vector’s homology arm confirmed the correct integration of the donor sequence in the OCT4 locus (Fig. 4C). Most importantly, the EGFP signals respond to differentiation cues as the endogenous Oct4 would (Fig. 4D). In a control experiment with an all-in-one plasmid expressing a nontargeting crRNA, no puromycin-resistant clones were obtained with the H1 ESC line. Only one puromycin-resistant clone each was obtained from H9 ESCs and iP5005 iPSCs, and neither clone was...
correctly targeted (Table 1). All of the above results indicated that the CRISPR-Cas system from *N. meningitidis* was able to generate accurately targeted clones in hPSCs with much increased efficiency compared with the traditional method.

**Discussion**

**Genome Editing by *N. meningitidis* Cas9.** In this report, we have successfully used the type II-C CRISPR-Cas system from *N. meningitidis* to achieve both NHEJ-mediated gene editing and long DNA donor-directed gene targeting of an endogenous locus in hPSCs. The targeting efficiency we obtain with NmCas9 is comparable to that achieved with TALENs. Using the same donor construct, we were able to get ~60% targeting efficiency in all three different hPSC lines tested (Table 1), whereas the targeting efficiency of a TALEN was 48% in the one hESC line tested (6). A previous report using SpCas9 in human iPSCs achieved a targeting efficiency of 48%, close to what we observed with NmCas9 (6). However, that report only identified seven clones and did not perform further analysis to confirm the correct integration of the donor DNA sequence only at the intended site. Therefore, additional work will be needed to compare the efficiency of mammalian gene targeting using these two CRISPR-Cas systems.

**CrRNA/Target Mismatch Tolerance by NmCas9 in Mammalian Cells.** One potential advantage of NmCas9, relative to SpCas9, is that it might offer better targeting specificity by virtue of its longer CrRNA spacer (24 vs. 20 nts) and its longer PAM (14). We chose 24 nt as the crRNA spacer length for NmCas9 because that is the length of the crRNA spacer in *N. meningitidis*. CrRNA-target mismatches distant from the PAM were tolerated to various extents for both NmCas9 (Fig. 3B) and SpCas9 (21) in mammalian cells. However, NmCas9 was more sensitive than SpCas9 to mismatches at the 13th, 15th, and 17th nts (counting from the PAM-proximal end of the proto-spacer). NmCas9 gene-editing efficiencies with mismatches at those positions were no higher than 10–40% of those observed with the perfectly matched crRNA (Fig. 3B), whereas with SpCas9, mismatches at equivalent positions retained 60–90% of the nonmismatched efficiency (21).

**PAM Requirements in Mammalian Cells.** One hallmark of type II CRISPR-Cas systems is the requirement of a nearby PAM on the target sequence. This sequence varies between different Cas9 orthologs. Among Cas9 proteins validated for mammalian genome editing, PAM functional requirements have been defined for three: those from *S. pyogenes* SF370 (21–23, 32, 34), *S. thermophilus* LMD-9 (the CRISPR1 locus) (19, 21, 38), and *N. meningitidis* 8013 (Fig. 3B) (14). On the one hand, the PAM requirement adds a second layer of specificity for gene targeting, beyond that afforded by spacer/protospacer complementarity. For longer PAMs (such as the NmCas9 PAM, 5′-NNNGATT-3′) the frequency of off-target cutting events should potentially drop significantly compared with SpCas9, which requires a 5′-NGG-3′ PAM. On the other hand, longer PAM requirements also constrain the frequency of targetable sites. By developing genome-editing systems using a range of Cas9 proteins with distinct PAM requirements, the genomic regions that can be targeted by CRISPR-Cas editing would expand significantly.

The results in Fig. 3B show that NmCas9 does allow limited deviation from the 5′-NNNGATT-3′ PAM. Having a variable PAM can potentially increase the flexibility during the design of targeting construct. However, it also increases the potential of off-target cleavage. Because of the limited options afforded by the sequence of tdTomato, we only tested one nucleotide substitution in each position of the PAM. It is possible that additional nucleotide substitutions will also be tolerated. A detailed mutational analysis will be needed to fully understand the PAM requirements of NmCas9 in mammalian cells.

**Editing the Genomes of hPSCs.** Compared with two other widely used systems for enhancing gene targeting efficiency (ZFNs and

![Image](https://example.com/image.png)

**Fig. 4.** Gene targeting in hESCs using NmCas9. (A) Donor DNA and crRNA design. The mismatch in the first nucleotide of crRNA is to satisfy the requirement of the U6 promoter for a G residue at the transcription start site. (B) Phase-contrast (Upper) and fluorescent (Lower) images of targeted clones from H1, H9, and iPS005 line. (Scale bars, 50 μM.) (C) Southern blot analysis of H1 WT (lane 1), H9 WT (lane 3), and iPS005 WT (lane 5) and targeted clones of H1 (lane 2), H9 (lane 4), and iPS005 (lane 6) line. Genomic DNA was digested with BamHI. The Southern probe is located outside of donor DNA (see A). The wild-type clone should give one band of 4.2 kb and targeted heterozygous clone should give one additional band of 5.6 kb. (D) Targeted clones (see B) were treated with 10 μM SB431542 and 10 ng/mL BMP4 to initiate differentiation. The EGFP signal was analyzed by FACS 3 d after differentiation. Gray, undifferentiated parental cells before targeting; green, targeted cells before differentiation; blue, targeted cells after differentiation.

| Table 1. Summary of gene targeting efficiency using NmCas9 in hPSCs |
|------------------|-----------------|----------------|-----------------|-----------------|-------------------|------------------|
| H1 (ES) Nontargeting | 0 | 0 | 0 | 0 | 0 | 0 |
| H1 (ES) Targeting | 20 | 5 | 3 | 12 | 60 |
| H9 (ES) Nontargeting | 1 | 1 | 0 | 0 | 0 | 0 |
| H9 (ES) Targeting | 39 | 9 | 7 | 23 | 59 |
| iPS005 (iPS) Nontargeting | 1 | 1 | 0 | 0 | 0 | 0 |
| iPS005 (iPS) Targeting | 10 | 1 | 3 | 6 | 60 |

The gene targeting efficiencies using the donor construct shown in Fig. 4A were calculated as the ratio between the number of correctly targeted clones and the total number of analyzed clones. For H1 cells, no puromycin-resistant clones were obtained using nontargeting crRNA.
Materials and Methods

**Cell Culture.** hESCs and iP5s cells were cultured in E8 medium (42) on Matrigel-coated tissue culture plates with daily medium change at 37 °C with 5% (vol/vol) CO2. Cells were split every 4–5 d with 0.05 mM EDTA in 1× PBS. 293FT cells were cultured similarly in DMEM/F12 media supplemented with 10% (vol/vol) FBS.

9. Cermak T, et al. (2011) Efficient genome editing by using cell lysate was prepared 2 d after transfection. Plasmid digestion was done using Fugene HD (Promega) following the manufacturer

**Cell Culture.** hESCs and iP5s cells were cultured in E8 medium (42) on Matrigel-coated tissue culture plates with daily medium change at 37 °C with 5% (vol/vol) CO2. Cells were split every 4–5 d with 0.05 mM EDTA in 1× PBS. 293FT cells were cultured similarly in DMEM/F12 media supplemented with 10% (vol/vol) FBS.

Supporting Information

Hou et al. 10.1073/pnas.1313587110

SI Materials and Methods

Plasmid Construction. The cas9 gene from Neisseria meningitidis strain 8013 was PCR-amplified and cloned into the pSimpleII plasmid (an OriP containing plasmid) under the control of the elongation factor-1α promoter. Nuclear localization signals and HA tag sequences were incorporated via the PCR primers. An N. meningitidis BsmBI-crRNA cassette and the N. meningitidis transacting clustered, regularly interspaced, short palindromic repeat (CRISPR) RNA (tracrRNA), both under the control of U6 RNA polymerase III promoters, were synthesized as gene blocks (Integrated DNA Technologies) and cloned into pSimpleII-Cas9 via blunt-end cloning, generating the pSimpleII-Cas9-Tracr-BsmBI plasmid that includes all elements needed for targeting. To insert specific spacer sequences into the crRNA cassette, synthetic oligonucleotides containing the desired spacer sequences were annealed to generate a duplex with overhangs compatible with those generated by BsmBI digestion of the pSimpleII-Cas9-Tracr-BsmBI plasmid. The insert was then ligated into the BsmBI-digested plasmid.

NmCas9 DNA Transfection and in Vitro Plasmid Digestion. All transfections with 293FT cells were done using Fugene HD (Promega) following the manufacturer’s instructions. Roughly 2-μg plasmids and 6 μL of Fugene HD were used for one well of a six-well plate. Two days after transfection, 293FT cells expressing N. meningitidis Cas9 (NmCas9) were harvested by TrypLE (Life Technologies), washed once in PBS, and then lysed in PBS by sonication. Cellular debris was cleared by centrifugation and the supernatant was used in plasmid digestion assays. For the digestions, 1 μg tdTomato plasmid (Clontech) linearized by NdeI (New England Biolabs) was mixed with in vitro-transcribed tracrRNA, crRNA, and 293FT cell lysate and incubated at 37 °C for 1–4 h in digestion buffer (1× PBS with 10 mM MgCl2). DNA from the reaction mix was then purified with a PCR clean-up kit (Qiagen) and resolved by agarose gel electrophoresis. To map the cleavage site of NmCas9, the digested plasmid DNA was excised from the agarose gel and purified using Gel Extraction Kit (Qiagen). The purified fragments were then sequenced to map the cleavage site.

In Vitro Transcription. Synthetic oligonucleotides (Integrated DNA Technologies) containing the T7 promoter sequence and N. meningitidis tracrRNA or crRNA sequences were annealed to generate dsDNA templates for run-off transcription. In vitro transcription was done using the MegaScript T7 In Vitro Transcription kit (Ambion) following the manufacturer’s specifications.

Southern Blots. Genomic DNA of targeted clones is purified using PureGene core kit (Qiagen). Five micrograms of genomic DNA was digested with BamHI and then resolved on a 0.8% agarose gel. DIG-labeled DNA probe synthesis, DNA gel transfer, and blot hybridization and visualization were done according to Roche’s DIG application manual.

Genome Editing Using Single-Guide RNA. A single-guide RNA (sgRNA) that targets tdTomato was put under the control of a U6 promoter and cloned into the EcoRV site of pstBlue-1 (NovaGen). For electroporation, 7.5 μg of pstBlue-U6-sgRNA, 7.5 μg of pSimpleII-NLS-NmCas9-HA-NLS(s), and 5 μg of EBNA RNA was mixed with ~1 x 10⁶ cells in a 4-mm cuvette (Bio-Rad) and immediately electroporated with a Bio-Rad Gene Pulser. Cells were then plated into appropriate Matrigel coated culture dishes in E8 supplemented with 10 μM ROCK inhibitor Y-27632.

Fig. S1. sgRNA is able to direct NmCas9-catalyzed genome editing. (A) Design of an sgRNA that targets tdTomato. Red: spacer; green: crRNA repeat; orange: linker; Blue: tracrRNA. (B) FACS analysis of tdTomato reporter human ES cell lines after electroporation of the sgRNA construct and the NmCas9 expressing plasmid. The number in the plot indicates the percentage of tdTomato fluorescence-negative cells 5 d after electroporation.
Table S1. Plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid no.</th>
<th>Plasmid name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pSimpleII-NmCas9-FLAG</td>
<td>Flag tagged NmCas9 without NLS</td>
</tr>
<tr>
<td>2</td>
<td>pSimpleII-NLS-NmCas9-EGFP</td>
<td>NmCas9 EGFP fusion with N-terminal NLS</td>
</tr>
<tr>
<td>3</td>
<td>pSimpleII-NLS-NmCas9-HA-NLS</td>
<td>NmCas9 with dual NLS and HA tags</td>
</tr>
<tr>
<td>4</td>
<td>pSimpleII-U6-tracrRNA-U6-BSmBII-NLS-NmCas9-HA-NLS(s)</td>
<td>All-in-one plasmid containing NmCas9, tracrRNA expression cassette and U6-BSmBII cassette</td>
</tr>
<tr>
<td>5</td>
<td>pSimpleII-U6-tracrRNA-U6-crRNA(tdTomato)-NLS-NmCas9-HA-NLS(s)</td>
<td>All-in-one plasmid containing NmCas9, crRNA expression cassette and tdTomato-targeting crRNA expression cassette</td>
</tr>
<tr>
<td>6</td>
<td>pSimpleII-U6-tracrRNA-U6-crRNA(EGFP)-NLS-NmCas9-HA-NLS(s)</td>
<td>All-in-one plasmid containing NmCas9, crRNA expression cassette and EGFP-targeting crRNA expression cassette</td>
</tr>
<tr>
<td>7</td>
<td>pSimpleII-U6-tracrRNA-U6-crRNA(OCT4)-NLS-NmCas9-HA-NLS(s)</td>
<td>All-in-one plasmid containing NmCas9, crRNA expression cassette and OCT4-targeting crRNA expression cassette</td>
</tr>
<tr>
<td>8</td>
<td>pSTBlue-1-U6-sgRNA (tdTomato)</td>
<td>U6-driven sgRNA targeting tdTomato</td>
</tr>
</tbody>
</table>

Table S2. crRNA-encoding DNA sequences used in this study

<table>
<thead>
<tr>
<th>crRNA</th>
<th>Encoding DNA sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFP targeting crRNA</td>
<td>gttcagcgctcggcaggccggaGTGTAGCTCCCTTTTCATTTTCG</td>
</tr>
<tr>
<td>OCT4 targeting crRNA</td>
<td>GtgctgatgctcggcaggccggaGTGTAGCTCCCTTTTCATTTTCG</td>
</tr>
<tr>
<td>tdTomato targeting crRNA (GATT PAM)</td>
<td>gttacgtgaacccccggccacatGTGTAGCTCCCTTTTCATTTTCG</td>
</tr>
<tr>
<td>tdTomato targeting crRNA (GATG PAM)</td>
<td>GccccgagggcttcaagtgggagcGTGTAGCTCCCTTTTCATTTTCG</td>
</tr>
<tr>
<td>tdTomato targeting crRNA (GACT PAM)</td>
<td>ggacggcggtctggtgaccgtgacGTGTAGCTCCCTTTTCATTTTCG</td>
</tr>
<tr>
<td>tdTomato targeting crRNA (GCTT PAM)</td>
<td>gattacaagaagctgtccttccccGTGTAGCTCCCTTTTCATTTTCG</td>
</tr>
<tr>
<td>tdTomato targeting crRNA (CATT PAM)</td>
<td>GggccgcccctacgagggcacccagGTGTAGCTCCCTTTTCATTTTCG</td>
</tr>
<tr>
<td>tdTomato targeting crRNA (CCAA PAM)</td>
<td>gggccgcccctacgagggcacccagGTGTAGCTCCCTTTTCATTTTCG</td>
</tr>
</tbody>
</table>

All sequences are 5’ to 3’, left to right. Spacer regions are in lowercase and underlined, and CRISPR repeat regions are in uppercase. In some cases, the first nucleotide of the spacer is changed to a G to satisfy the requirement of the U6 promoter.

Table S3. Primers used in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>For cleavage site mapping in tdTomato</td>
<td>ATGGTGAGCAAGGGCGAGGAG</td>
<td>CCGGTGCTCCGGTGCCATGCCCAG</td>
</tr>
<tr>
<td>For tdTomato indel mapping</td>
<td>AACACTGTCCCTCTCATGTCCTGTCTC</td>
<td>CCGGTGCTCCGGTGCCATGCCCAG</td>
</tr>
<tr>
<td>For making a Southern blot probe for OCT4</td>
<td>GTGATGCCAACAAGAACCCT</td>
<td>ACAGCACCGAGCAATAGGT</td>
</tr>
</tbody>
</table>

All sequences are 5’ to 3’, left to right.

Table S4. Sequences of unprocessed and processed tracrRNA and crRNA in *N. meningitidis*

<table>
<thead>
<tr>
<th>tracrRNA and crRNA</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unprocessed tracrRNA</td>
<td>5’TAAUUGGCACUGCGAAUAGGAACCGUUGCUAACAUAAGGCGUCUGAAAAAGUGGCGCCGAGCAAGCGUCUGCGCCUGUUUA 3’</td>
</tr>
<tr>
<td>Processed tracrRNA</td>
<td>5’TAAUUGGAACCGUUGCUAACAUAAGGCGUCUGAAAAAGUGGCGCAACGCUCUGCCCUCUGCUUUA 3’</td>
</tr>
<tr>
<td>Full-length spacer+repeat unit in CRISPR locus</td>
<td>5’NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGTAGCTCCCTTTTCATTTTCGACTGCTA 3’</td>
</tr>
<tr>
<td>Processed crRNA</td>
<td>5’NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGGTAGCTCCCTTTTCATTTTCG 3’</td>
</tr>
</tbody>
</table>

Sequences that are cleaved off during processing are underlined.

Other Supporting Information Files

Dataset S1 (DOCX)
Dataset S2 (DOCX)
Dataset S3 (DOCX)