HHS Public Access

Author manuscript

Nat Chem Biol. Author manuscript; available in PMC 2017 December 19.

Published in final edited form as:

Nat Chem Biol. 2017 August ; 13(8): 839–841. doi:10.1038/nchembio.2410.

Cpf1 proteins excise CRISPR RNAs from mRNA transcripts in mammalian cells

Guocai Zhong[1], **Haimin Wang**[1], **Yujun Li**[1,2], **Mai H. Tran**[1], and **Michael Farzan**[1]

[1]Department of Immunology and Microbiology, The Scripps Research Institute, Jupiter, Florida, USA

[2]Department of Microbiology and Wu Lien-Teh Institute, Harbin Medical University, Harbin, China

Abstract

Cpf1 is a CRISPR-effector protein with greater specificity than *Streptococcus pyogenes* Cas9 (SpCas9) in genome-editing applications. Here we show that *Lachnospiraceae bacterium* (Lb) and *Acidaminococus sp.* (As) Cpf1 orthologs have RNase activities that can excise multiple CRISPR RNAs (crRNAs) from a single Pol II-driven RNA transcript expressed in mammalian cells. This property simplifies modification of multiple genomic targets and can be used to increase the efficiency of Cpf1-mediated editing.

Recently, two effector proteins of type V-A CRISPR systems, the Cpf1 proteins of *Lachnosperiaceae bacterium* ND2006 (LbCpf1) and *Acidaminococus sp.* BV3L6 (AsCpf1) have been shown to efficiently edit mammalian cell genomes with fewer off-target cleavage sites than the widely used *Streptococcus pyogenes* Cas9 (SpCas9) [1-3]. In bacteria, Cas9 relies on RNase III to excise crRNAs from a CRISPR array [4]. In contrast, one Cpf1 protein, that of *Francisella novidica* U112 (FnCpf1) has its own RNase activity that can excise crRNA from a bacterial CRISPR array [5]. However, FnCpf1 does not efficiently edit mammalian genomes [1], and it is not known whether any Cpf1 can excise crRNA from mRNA expressed in mammalian cells. We therefore investigated the ability of LbCpf1 and AsCpf1 to excise crRNA from transcripts expressed in mammalian cells by RNA polymerase II (Pol II), and to use these crRNAs to edit genomes.

To assess Cpf1 RNase activity in mammalian cells, we first inserted the LbCpf1 or AsCpf1 crRNA-scaffold region (Fig. 1a) into the 3′ untranslated region (3′ UTR) of a gene encoding *Gaussia* luciferase (GLuc). Thus when the GLuc mRNA transcript is cleaved in its 3′ UTR, luciferase expression is halted. We cotransfected 293T cells with one of six such constructs and a plasmid expressing either LbCpf1 or AsCpf1 (Fig. 1b). Neither Cpf1 inactivated a GLuc transcript lacking a crRNA scaffold, whereas both LbCpf1 and AsCpf1

Correspondence should be addressed to G.Z. (gzhong@scripps.edu).

Author contributions: G.Z. conceived this study and developed it with assistance from M.F.; G.Z., H.W., Y.L., and M.H.T. performed all experiments; M.F. and G.Z wrote the manuscript.

Competing financial interests: The authors declare they have no competing financial interests.

inactivated a GLuc transcript bearing the Lb scaffold. When that scaffold sequence was modified in its initiating AAUU sequence to UUAA, or completely randomized, Cpf1 activity was abrogated. Interestingly, when the As scaffold replaced the Lb scaffold, only AsCpf1, but not LbCpf1, could efficiently halt luciferase expression. LbCpf1's greater ability to discriminate among scaffolds is consistent with its greater interaction with the loop region of its scaffold [6, 7]. We then tested a panel of LbCpf1 variants, with mutations either in its putative RNase region, or in its RuvC DNase domain [1, 5] (Fig. 1c). As before, wild-type LbCpf1 abrogated GLuc expression, whereas putative RNase LbCpf1 mutants did so less efficiently (K768A, K785A, F789A) or not at all (H759A, HKK/AAA). In contrast, LbCpf1 DNase mutants inactivated GLuc as efficiently as wild-type LbCpf1, indicating that the RNase and DNase activities of LbCpf1 can be segregated. Consistently, LbCpf1 RNase mutants H759A, K768A and K789A retained their DNase activity, as indicated by efficient gene editing mediated by three different U6 (Pol-III) promoter-expressed guide RNAs (gR1, gR3, and gR8;Supplementary Results, Supplementary Figs. 1 and 2). In contrast, F789A and HKK/AAA did not, suggesting that these mutations impair Cpf1 function more globally. Note that H759 and K785 are proximal to the phosphate group of the first adenosine, A(-20), of the Lb scaffold [6] and are thus well positioned to mediate LbCpf1's RNase activity (Supplementary Fig. 2e). We conclude that both LbCpf1 and AsCpf1 have RNase activity that can cleave an mRNA transcript bearing an appropriate Cpf1 scaffold sequence in its 3′ UTR.

To determine if crRNA excised from a Pol II-driven mRNA transcript could be used to edit mammalian cell genomes, we inserted the functional guide, gR1, between two Lb scaffolds (sR) in the 3′ UTR of a GLuc transcript (GLuc-sR-gR1-sR; Fig. 2a). We assessed the efficiency of genome editing with crRNA excised from this GLuc transcript. Wild-type LbCpf1 efficiently introduced indel mutations in an integrated target gene encoding a green-fluorescent protein variant engineered for a short half-life (EGFPd4), as indicated by T7E1 assay (Fig. 2b) and by loss of GFP fluorescence (Supplementary Figs. 3a and b). Nearly identical results were obtained when gR1 was replaced by gR3 (Supplementary Figs. 3c and d). Editing was dependent on both the DNase and RNase activities of Cpf1, as indicated by loss of editing with RNase mutants H759A and with DNase mutants D832A and E925A. This genome editing was also dependent on the presence of the sR-gR1-sR sequence and on the second sR domain, as shown by the induction of a stably integrated and originally out-of-frame firefly luciferase (FLuc) by the GLuc transcript (GLuc-sR-gR1-sR) in a double-strand break (DSB)-induced gain-of-expression reporter assay [8] (Fig. 2c and Supplementary Fig. 4), indicating that two RNase cleavage events were necessary to produce a functional crRNA. Finally, using nine different guide sequences and the DSB-reporter assay depicted in Figure 2c, we compared the efficiency of editing mediated by Pol III (U6)-expressed crRNAs with that by Pol II (CMV)-expressed crRNAs (Fig. 2d). In most cases, editing was more efficient with Pol II expression, and in no case was it less efficient. We conclude that LbCpf1 can efficiently excise and utilize crRNAs embedded in the 3′ UTR of a Pol II-driven RNA transcript.

We then assessed the ability of LbCpf1 to edit multiple genes using crRNAs excised from a single Pol-II transcript. To do so, a second gain-of-expression DSB reporter, expressing *Cypridina* Luciferase (CLuc) (Supplementary Fig. 5a), was stably integrated into reporter

cells employed in Figure 2c. Single gR1 and gR3 mediated activation of the originally out-of-frame FLuc and CLuc genes, respectively, by wild-type LbCpf1, but not its RNase-inactive mutant H759A (Fig. 3a). When both gR1 and gR3 were arrayed in a single Pol-II transcript (gR1-gR3; gR3-gR1), both FLuc and CLuc were activated with efficiencies similar to those mediated by a single crRNA. Nearly identical results were obtained when guide RNAs (gR9 and gR11) and target sequences were derived from previously characterized DNMT1 and EMX1 loci [1] (Supplementary Fig. 5b). Pol II-expressed multiplexed crRNA could also edit the native genome. Thus gR9 and gR11, which target respectively *DNMT1* and *EMX1*, could efficiently edit both genes when expressed from the same Pol-II promoter (gR9-gR11, gR11-gR9; Supplementary Fig. 5c). We then compared the efficiency of multiplex editing mediated by CMV (Pol-II) and U6 (Pol-III) promoters. We observed that the efficiency of a crRNA expressed at the 3′ of a Pol-III transcript decreased as the length of the crRNA array increased. No such decrease was observed with Pol-II transcripts. This observation was confirmed using four guide RNAs, gR1 and gR3 (Fig. 3b), as well as gR9 and gR11 (Supplementary Fig. 5d). These data are consistent with the presence of a low-efficiency Pol-III terminator, TTTCT, in the Cpf1 scaffold regions [1, 9]. To evaluate multiplex editing on a single-cell basis, we used three different crRNA arrays (Array 1-3) to compare Pol II-expressed arrays, Pol III-expressed arrays, and a pool of single Pol II-expressed crRNAs. Gain-of-expression DSB-reporter systems were again applied, except in this case, stably integrated and initially out-of-frame genes expressing the fluorescent proteins, iRFP670 [10] and Superfolder GFP [11], served to report gene editing events (Supplementary Fig. 6a). Pol II expression of Array 1 and Array 2 generated significantly more double-positive cells than their Pol III-expressed or pooled Pol II-expressed counterparts (Fig. 3c, Supplementary Figs. 6b and c). However, when editing was especially efficient, as with Array 3, no differences were observed between Pol II- and Pol III-expressed arrays, suggesting that other factors limit editing efficiency in this case. Collectively, Figure 3 makes clear that LbCpf1 can efficiently generate and utilize multiple crRNAs expressed in tandem by CMV (Pol-II) promoter, and that, in general, it does so more efficiently than when U6 (Pol-III) promoter is used.

Here we describe and assess the utility of the RNase activity of LbCpf1 and AsCpf1 for genome-editing applications. LbCpf1 and AsCpf1 differ in several subtle ways that may be relevant to these applications. It has previously been reported that the AsCpf1 PAM-binding domain is more selective than that of LbCpf1, whereas AsCpf1 DNase exhibits greater off-target activity [1-3]. Our data suggest that the LbCpf1 scaffold-binding region is more selective than that of AsCpf1. For example, AsCpf1 RNase activity works well with both As and Lb scaffold sequences, but LbCpf1 can only function efficiently with its cognate scaffold (Fig. 1b). However, the AsCpf1 RNase activity was more efficient in our hands, while LbCpf1 appears to be more efficient in on-target genome editing, although such differences may be context dependent [2, 3].

Our data show that AsCpf1 and LbCpf1 can excise multiple crRNAs from a single RNA transcript expressed from either a Pol-II or a Pol-III promoter. A recently published paper also reported that these enzymes could excise multiple functional crRNAs, albeit only from a Pol-III transcript [12]. We observed that CMV (Pol II)-expressed crRNAs were consistently as or more efficient at mediating genome editing than those expressed from U6 (Pol-III)

promoter. This observation may in part reflect the fact that only Pol-II transcripts are efficiently exported to the cytoplasm where they might more easily interact with newly translated Cpf1. The ability to excise multiple crRNAs simplifies the targeting of a single gene or exon at multiple sites, or the targeting of multiple genes in a cell. Such targeting can increase the efficiency of a gene knockout, or facilitate editing of many proteins in parallel for combinatorial screens. It also enables *in vivo* genome-editing of multiple targets with the commonly used adeno-associated virus vector, which can only accommodate Cpf1 or Cas9 and a single promoter for crRNA expression [13]. Thus the RNase activity of AsCpf1 and LbCpf1 in mammalian cells expands the utility of these key CRISPR effector molecules.

Online Methods

Plasmids

Wild-type LbCpf1 plasmid, pY016 (pcDNA3.1-hLbCpf1, Addgene plasmid # 69988), and AsCpf1 plasmid, pY010 (pcDNA3.1-hAsCpf1, Addgene plasmid # 69982) were gifts from the laboratory of Dr. Feng Zhang [1]. LbCpf1 mutations (H759A, K768A, K785A, HKK/AAA, F789A, D832A, E925A) were introduced by synthesis of the LbCpf1 fragment on pY016 plasmid between *Asc*I and *Ale*I restriction sites. The synthesized fragments containing the desired mutations were cloned back to the pY016 plasmid to replace the wild-type LbCpf1 sequence between *Asc*I and *Ale*I sites.

To create GLuc expression plasmids with 3′ UTR scaffold RNAs (sR) or CRISPR RNA (crRNA) variants, DNA fragments encoding Cpf1 sRs, single crRNAs, or crRNA arrays were synthesized as gBlocks or custom genes by Integrated DNA Technologies (IDT) and cloned at the 3′ UTR of a CMV-driven *Gaussia* luciferase (GLuc) expression plasmid between *Bam*HI and *Xho*I restriction sites. DNA fragments encoding single crRNAs or crRNA arrays expressed by a U6 promoter were synthesized by IDT and cloned into pcDNA3.1(+) between *Mfe*I and *Nru*I restriction sites. Sequences for all the single crRNAs and crRNA arrays and the targets of the guide RNAs are provided in Supplementary Tables 1 and 3.

Three classes of double-strand break (DSB) reporter plasmids were used in this study. The first encodes a frameshift construct, EGFP(+1)-GLuc(+3), similar to those described in Certo et al. [8] This DSB reporter was constructed by cloning a fusion of the EGFP coding sequence with sequence encoding the foot-and-mouth disease protease (F2A) and GLuc coding sequence into pQCXIP vector plasmid (Clontech) between *Sbf*I and *Eco*RI restriction sites. Both F2A and GLuc were frameshifted 2-bp out of the EGFP (+1) reading frame, and the GLuc start codon was removed. This reporter plasmid was used to produce retroviral particles for generating the DSB-induced gain-of-expression reporter cell line used in Supplementary Figures 1 and 2. The second DSB reporter plasmid encodes an EGFP variant with a 4-hour half-life (EGFPd4). It was constructed by cloning a fusion of EGFP coding sequence with a sequence encoding a mutated version of residues 422-461 of mouse ornithine decarboxylase into pQCXIP vector plasmid between *Not*I and *Bam*HI restriction sites. This plasmid was used to produce retroviral particles for generating the DSB-induced loss-of-GFP-expression reporter cell line used in Figure 2b, and Supplementary Figure 3. The last category of DSB reporter plasmids encodes one of a panel of DSB-induced gain-of-

expression constructs. They were constructed by cloning a fusion of a guide RNA target sequence (gRT) and a reporter gene (FLuc, CLuc, iRFP670 [10], or Superfolder GFP [11]) into pQCXIP vector plasmid between *SbfI* and *BamHI* restriction sites. Guide RNA target sequences (gRTs) were inserted downstream of the +1 frame start codon. A "TCCGGA" sequence, encoding a serine-glycine (S-G) dipeptide linker, and a reporter gene lacking the native start codon were inserted downstream and frame-shifted 2 bp from the +1 reading frame. Twelve reporter plasmids of this category were generated, including gR1T-FLuc, gR3T-CLuc, gR9T-FLuc, gR10T-CLuc, gR11T-CLuc, gR12T-FLuc, gR13T-FLuc, gR14T-FLuc, gR15T-FLuc, gR1T-iRFP670, gR3T-SuperfolderGFP, and gR11T-SuperfolderGFP. These reporter plasmids were used to produce retroviral particles for generating DSB-induced gain-of-expression single- or dual-reporter cell lines used in Figures 2-3, Supplementary Figures 4-6. All of the DSB-induced gain-of-expression reporters and primers used to generate these reporter plasmids are provided in Supplementary Table 2.

Cell culture

Human embryonic kidney 293T (ATCC) and 293T-derived stable cell lines were maintained in Dulbecco′s Modified Eagle Medium (DMEM, Life Technologies) at 37°C in a 5% CO_2-humidified incubator. All growth media were supplemented with 2 mM Glutamax-I™ (Life Technologies), 100 μM non-essential amino acids (Life Technologies), 100 U/mL penicillin and 100 μg/mL streptomycin (Life Technologies), and 10% FBS (Hyclone). 293T cells were confirmed mycoplasma-free by the provider.

Generation of DSB-reporter stable cell lines

293T cells plated at 70% confluence in T75 culture flasks were transfected using CalPhos™ transfection kit (Clontech) with 10 μg of plasmid encoding the Machupo virus (MACV) envelope glycoprotein GP, 15 μg of plasmid encoding murine leukemia virus (MLV) Gag and Pol proteins, and 15 μg of pQCXIP-based DSB-reporter plasmids. Forty-eight hours after transfection, culture supernatants containing the MLV-based retroviral particles were harvested and filtered through 0.45 μm surfactant free cellulose acetate filters (Corning).

To generate the DSB-reporter stable cell lines, 293T cells were plated in 6-well plates, and transduced with the retroviral particles either at low MOI (less than 1% of cells transduced) for EGFP(+1)-GLuc(+3) and EGFPd4 single-reporter cell lines, or at high MOI (more than 80% of cells transduced) for other reporter cell lines. Two days post transduction, the cells were selected with growth medium containing 3 μg/mL puromycin (Life Technologies) to generate a polyclonal stable cell pool. EGFPd4 stable cells were further colonized and a single colony with robust EGFP expression was chosen and expanded for use in the experiments shown Figure 2b, and Supplementary Figure 3.

Cpf1 RNase-mediated loss of GLuc expression

293T cells seeded in 48-well plate with antibiotic-free growth medium were transfected with 100 ng of Cpf1 plasmid and 20 ng of GLuc-sR variant plasmid using 0.3 μL Lipofectamine 2000 (Life Technologies). Four hours later, medium was removed and fresh growth medium containing 2% FBS was added. Culture medium was refreshed on day 1 post transfection,

and GLuc in the supernatant was measured using a luminescence assay on day 2 post transfection.

Cpf1 DNase activity assays

293T or 293T-derived stable cells carrying one or two integrated DSB reporters were seeded in 48-well plate with antibiotic-free growth medium. Cells were transfected with 100 ng of Cpf1 plasmid in combination with 100 ng of single crRNA plasmid, 20 ng of pooled single crRNAs, or 20 ng of crRNA arrays using 0.75 μL Lipofectamine 2000. Six to ten hours later, medium was removed and fresh growth medium containing 2% FBS was added. Culture medium was refreshed every day. On day 3 post transfection, cells were either assayed for reporter expression by luminescence assays (for FLuc and CLuc) or by flow cytometry (for EGFPd4, iRFP670, and Superfolder GFP), or collected for genomic DNA isolation and analyzed by T7E1 assay to quantify Cpf1-mediated indel frequency.

Luminescence assays

To measure GLuc expression, 20 μL of 10-fold diluted cell culture supernatant of each sample and 100 μL of GLuc assay buffer were added to one well of a 96-well black opaque assay plate (Corning), and measured with a Centro LB 960 microplate luminometer (Berthold Technologies) for 0.1 second/well. GLuc assay buffer consists of 7.5 mM sodium acetate, 250 mM sodium sulfate, 250 mM sodium chloride, and 4 μM coelenterazine native (Biosynth Chemistry & Biology) at pH 5.0.

Firefly luciferase (FLuc) expression was measured using Pierce Firefly Luciferase Flash Assay Kit (ThermoFisher Scientific) following the manufacturer's instructions. In brief, cells in 48-well plate were washed once with PBS, then incubated and shaken with 250 μL of Cell Lysis Buffer for 10 min at room temperature. 20 μL cell lysate of each sample and 50 μL of Working Solution (Firefly Flash Assay Buffer + D-luciferin) were added to one well of a 96-well white opaque assay plate (Corning), and measured with the Centro LB 960 microplate luminometer for 0.1 second/well.

Cypridina luciferase (CLuc) expression was measured using Pierce *Cypridina* Luciferase Flash Assay Kit (ThermoFisher Scientific) following the manufacturer's instructions. In brief, 20 μL of cell culture supernatant of each sample and 50 μL of CLuc assay buffer (*Cypridina* Flash Assay Buffer + Vargulin) were added to one well of a 96-well black opaque assay plate (Corning), and measured with the Centro LB 960 microplate luminometer for 0.1 second/well.

T7E1 mismatch cleavage assay

Genomic DNA was isolated from transfected cells using PureLink Genomic DNA Mini Kit (Life Technologies). 100 ng genomic DNA and a 50 μL PCR reaction was used to amplify guide-RNA-targeted loci for 35 cycles. PCR was done with CloneAmp HiFi PCR Premix (Clontech). The sequences of PCR primers are provided in Supplementary Table 4. PCR products were purified with NucleoSpin Gel and PCR Clean-up kit (Machery-Nagel). 300 ng of purified PCR product was denatured and reannealed in NEB buffer 2 (New England Biolabs) using a Matstercycler Pro S thermocycler (Eppendorf) with the following

procedures: 95 °C, 10 min; 95-85 °C, cooling down with 44% ramp rate (-2 °C/s); 85-25 °C, cooling down with 2% ramp rate (-0.1 °C/s); 25 °C, hold for 1 min; 25-4°C, cooling down with 100% ramp rate (-4.5 °C/s); 4 °C, hold. Hybridized PCR products were incubated with 10 U of T7 Endonuclease I (New England Biolabs) at 37 °C for 30min and analyzed by agarose gel electrophoresis. Gels were stained with SYBR Gold DNA stain (Life Technologies) and imaged with ImageQuant LAS4000 Mini (GE Healthcare Life Sciences) gel imaging system. Indel frequency (%) was calculated using the formula, 100×(1-sqrt(1-faction cleaved)), as described in Guschin et al. [14]

Flow cytometry

To quantify fluorescent protein (EGFPd4, iRFP670, and Superfolder GFP) expression level in the DSB-reporter stable cells, cells in 48-well plates were washed once with PBS and trypsinized down. Cells were spun down, washed once, and resuspended in PBS. Then cells were loaded to BD Accuri™ flow cytometer (BD Biosciences), and data from twenty thousand gated cells per sample were collected and analyzed using BD CSampler™ software (BD Biosciences).

Statistical analysis

Data expressed as mean values ± s.d. or s.e.m. Statistical analyses were performed using paired- or two-sample Student's *t*-test using GraphPad Prism 6.0 software when applicable. Differences were considered significant at $P < 0.01$.

Data availability

Our research resources, including methods, cells, and protocols are available upon request to qualified academic investigators for non-commercial research purposes. All reagents developed, such as DSB receptor constructs, as well as detailed methods, will be made available upon written request. The corresponding author adheres to the NIH Grants Policy on Sharing of Unique Research Resources including the “Sharing of Biomedical Research Resources: Principles and Guidelines for Recipients of NIH Grants and Contracts.”

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is supported by NIH R37 AI091476 and P01 AI100263 (M.F).

References

1. Zetsche B, et al. Cell. 2015; 163:759–771. [PubMed: 26422227]
2. Kim D, et al. Nature biotechnology. 2016; 34:863–868.
3. Kleinstiver BP, et al. Nature biotechnology. 2016; 34:869–874.
4. Mohanraju P, et al. Science. 2016; 353:aad5147. [PubMed: 27493190]
5. Fonfara I, Richter H, Bratovic M, Le Rhun A, Charpentier E. Nature. 2016; 532:517–521. [PubMed: 27096362]
6. Dong D, et al. Nature. 2016; 532:522–526. [PubMed: 27096363]

7. Yamano T, et al. Cell. 2016; 165:949–962. [PubMed: 27114038]
8. Certo MT, et al. Nature methods. 2011; 8:671–676. [PubMed: 21743461]
9. Orioli A, et al. Nucleic acids research. 2011; 39:5499–5512. [PubMed: 21421562]
10. Shcherbakova DM, Verkhusha VV. Nature methods. 2013; 10:751–754. [PubMed: 23770755]
11. Pedelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS. Nature biotechnology. 2006; 24:79–88.
12. Zetsche B, et al. Nature biotechnology. 2017; 35:31–34.
13. Samulski RJ, Muzyczka N. Annual review of virology. 2014; 1:427–451.
14. Guschin DY, et al. Methods Mol Biol. 2010; 649:247–256. [PubMed: 20680839]

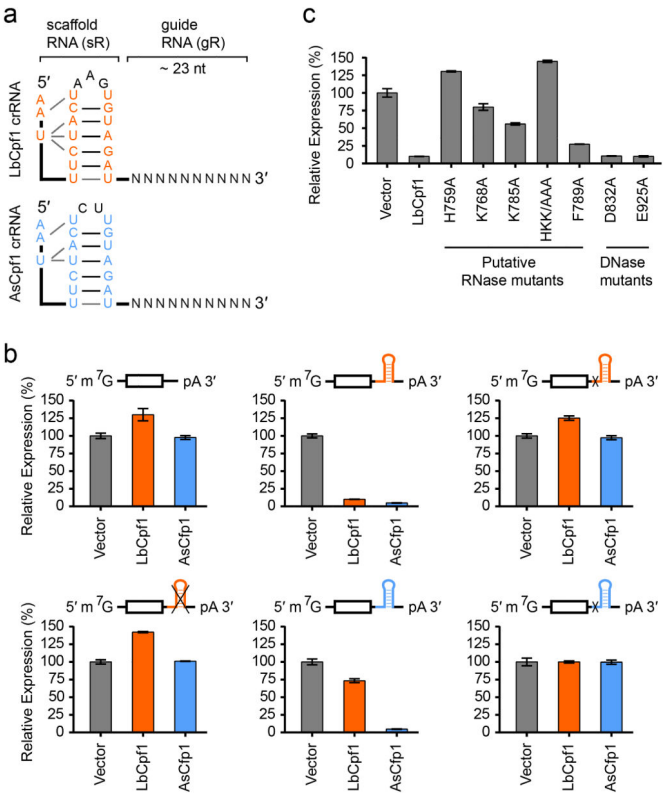

Figure 1.

LbCpf1 and AsCpf1 have RNase activities in mammalian cells. (**a**) The crRNA recognized by LbCpf1 and AsCpf1 are represented. A 19-20 nucleotide scaffold region (sR) is followed by a 23-base guide region (gR) complementary to the target DNA. (**b**) Plasmids encoding LbCpf1, AsCpf1 or vector alone were cotransfected with GLuc-expressor plasmids bearing the indicated Cpf1 scaffold RNA variants, and GLuc activity was measured. When Cpf1 recognizes an appropriate scaffold RNA present in the 3′ UTR of an mRNA encoding *Gaussia* luciferase (GLuc), the message is cleaved and GLuc expression is halted. Orange indicates LbCpf1 or the Lb scaffold, whereas blue indicates AsCpf1 or the As scaffold. A small '×' preceding the scaffold indicates replacement of the initial AAUU sequence with UUAA. A large '×' indicates that the scaffold sequence has been randomized. (**c**) An assay similar to that in (b) except that LbCpf1 variants with mutations in their putative RNase domains or in their RuvC DNase domains were evaluated for their ability to inactivate GLuc expression from the reporter plasmid bearing a wild-type Lb scaffold in the 3′ UTR. Experiments shown are representative of two (panel b), or four (panel c) performed with nearly identical results. Data points in panels b and c represent mean ± s.d. of three biological replicates.

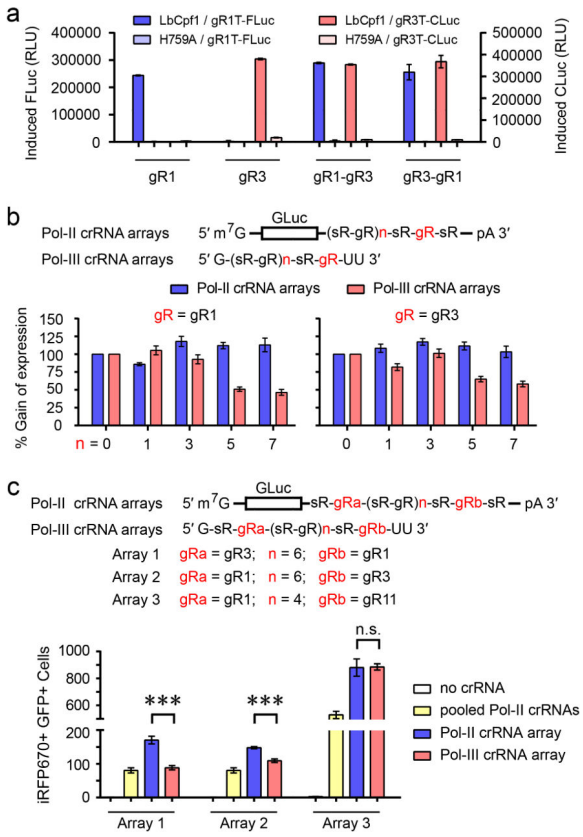

Figure 2.

crRNA excised from an Pol II-expressed RNA transcript can efficiently edit a mammalian genome. (**a**) An mRNA encoding GLuc with two Lb scaffold regions (sR) separated by a 23-base guide (gR1) in its 3′ UTR (GLuc-sR-gR1-sR) can be cleaved by LbCpf1. If both scaffold regions are cleaved, a functional crRNA can be produced and loaded into LbCpf1 to perform genome editing. (**b)** Wild-type LbCpf1 or the indicated LbCpf1 variants were co-expressed with the Pol-II construct, GLuc-sR-gR1-sR, to edit an integrated EGFP gene. gR1 is complementary to the EGFP gene. T7E1 mismatch cleavage assay was performed to measure the editing efficiency. (**c**) Depiction of a double-strand break (DSB)-induced gain-of-expression assay used below. An integrated and initially out-of-frame firefly luciferase (FLuc) or *Cypridina* luciferase (CLuc) gene preceded by a guide RNA target sequence (gRT-FLuc, or gRT-CLuc) is placed in frame with the +1 frame translational start codon, in approximately one third of the editing events by Cpf1-mediated DSB and cellular non-homologous end joining (NHEJ). (**d**) Two crRNA expression systems are depicted and compared using the indicated guide RNAs and the assay system depicted in (c). Pol II-expressed crRNA was significantly more efficient than that expressed by a Pol-III promoter (two-tailed paired-sample *t*-test; P=0.0047). Sequence information of the guide RNAs and related reporter constructs can be found in Supplementary Tables 1 and 2. Experiments shown are representative of at least two performed with nearly identical results. Data points in panel d represents mean ± s.d. of three biological replicates.

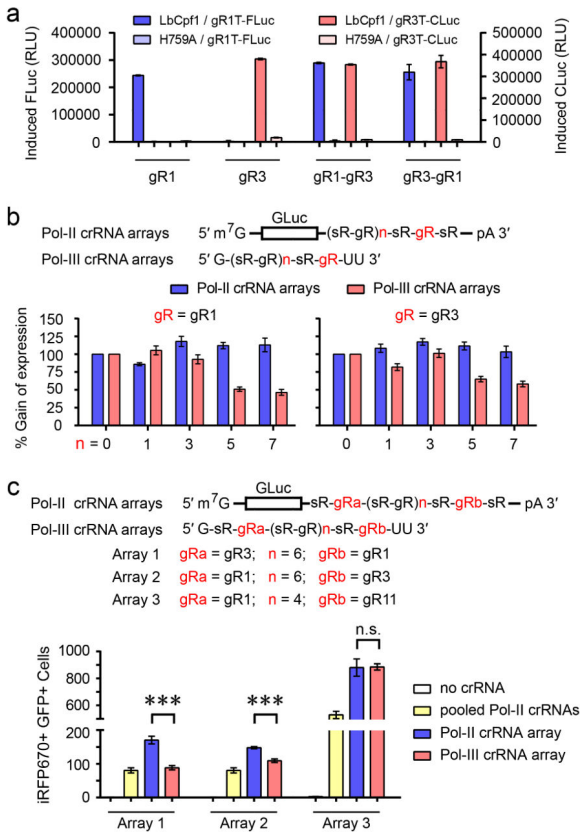

Figure 3.

Cpf1 efficiently excises and uses tandemly arrayed crRNAs expressed from a Pol-II promoter. **(a)** A dual-reporter cell line with integrated and initially out-of-frame FLuc (gR1T-FLuc) and CLuc (gR3T-CLuc) reporter genes was cotransfected with the indicated crRNA expression constructs and wild-type or RNase-inactive (H759A) LbCpf1, and assayed as depicted in Supplementary Figure 5a. Experiments shown are representative of two performed with nearly identical results, and data points represent mean ± s.d. of three biological replicates. (**b**) Tandem arrays with gR1 or gR3 preceded by n (n=0, 1, 3, 4, or 7) additional scaffold and guide RNA pairs (sR-gR) expressed by either Pol-II CMV or Pol-III U6 promoter are depicted. The gR1 and gR3 arrays were compared using cell lines carrying stably integrated DSB-induced gain-of-expression reporter gR1T-FLuc and gR3T-CLuc, respectively. Luciferase expression was normalized to values observed in the single crRNA constructs (n=0). (**c**) Dual-reporter cell lines carrying stably integrated and initially out-of-frame iRFP670 (gR1T-iRFP670) and Superfolder GFP (gR3T- or gR11T-SuperfolderGFP) were used to report editing efficiencies of the indicated crRNA arrays. Numbers of double-edited cells (iRFP670+ GFP+) are indicated for each array and expression system. (Two-tailed two-sample *t*-test; ***P=0.0002 for array 1 and P=0.0008 for array 2; n.s., not significant). Data points in panels b, and c represent mean ± s.e.m. of three or four independent experiments, each with three or four biological replicates. Sequence information of the crRNA arrays and related DSB reporter constructs can be found in Supplementary Tables 1-3.