Published online 13 October 2020

Nucleic Acids Research, 2020, Vol. 48, No. 20 ***e117***
doi: 10.1093/nar/gkaa843

DNA and RNA editing without sequence limitation using the flap endonuclease 1 guided by hairpin DNA probes

Kun Tian[1,†], Yongjian Guo[2,†], Bingjie Zou[3,†], Liang Wang[1], Yun Zhang[1], Zhen Qi[1], Jieying Zhou[2], Xiaotang Wang[2,*], Guohua Zhou[3,*], Libin Wei[1,*] and Shu Xu[1,*]

[1]School of Basic Medical Science and Clinical Pharmacy, China Pharmaceutical University, Nanjing 210000, China, [2]Department of Chemistry and Biochemistry, Florida International University, Miami, FL, USA and [3]Department of Pharmacology, Jinling Hospital, Medical School, Nanjing University, Nanjing 210000, China

Received January 22, 2020; Revised September 11, 2020; Editorial Decision September 17, 2020; Accepted September 19, 2020

ABSTRACT

Here, we characterized a flap endonuclease 1 (FEN1) plus hairpin DNA probe (hpDNA) system, designated the HpSGN system, for both DNA and RNA editing without sequence limitation. The compact size of the HpSGN system make it an ideal candidate for *in vivo* delivery applications. *In vitro* biochemical studies showed that the HpSGN system required less nuclease to cleave ssDNA substrates than the SGN system we reported previously by a factor of ∼40. Also, we proved that the HpSGN system can efficiently cleave different RNA targets *in vitro.* The HpSGN system cleaved genomic DNA at an efficiency of ∼40% and ∼20% in bacterial and human cells, respectively, and knocked down specific mRNAs in human cells at a level of ∼25%. Furthermore, the HpSGN system was sensitive to the single base mismatch at the position next to the hairpin both *in vitro* and *in vivo*. Collectively, this study demonstrated the potential of developing the HpSGN system as a small, effective, and specific editing tool for manipulating both DNA and RNA without sequence limitation.

INTRODUCTION

Homologous recombination and nonhomologous end-joining are necessary but inefficient processes in genomic DNA modification. To improve the efficiency of these processes, precise and efficient genome-targeting cleavage tools are needed to make breaks in genomic DNA. To date, many canonical DNA editing nucleases have been reported, such as ZFNs (1–4), TALENs (4–6) and CRISPR-associated systems (7–10). In particular, CRISPR-associated systems have been applied for genomic and epigenetic modification in both prokaryotic and eukaryotic cells (11–13). CRISPR-associated genetic perturbation is simple and scalable, empowering scientists to establish causal linkages between genetic variations and biological phenotypes. However, almost all CRISPR-associated systems require protospacer-adjacent motif (PAM) sequences to locate targets, which restricts the target space in the human genome. This restriction has been partially overcome by exploiting the family of Cas9 enzymes and their differing PAM requirements (14–16). On the other hand, Cas9 and Cas12 proteins, the key to such CRISPR-associated systems, are relatively hard to pack and deliver along with DNA donors *in vivo* due to their large sizes. Although efforts have been made to develop smaller Cas proteins, such as CasX (17) and Cas14 (18), which contain fewer than one thousand amino acid residues, more studies are warranted to further reduce the sizes of these DNA editing tools.

Recently, we reported a structure-guided nuclease (SGN) DNA editing tool without sequence limitation (19). It was constructed based on fusing flap endonuclease 1 (FEN1) (20) to the cleavage domain of *Fok* I (Fn1) (21). SGN recognizes the target DNA through the FEN1 domain on the basis of the 3′ flap structure formed between DNA targets and guide DNA probes and cuts DNA targets through Fn1 dimerization. Since the formation of the 3′ flap structure is not limited by target sequences, SGN can theoretically cleave any locus in genomic DNA. SGN, composed of ∼500 amino acids, is much smaller than Cas proteins. Furthermore, large fragment deletions, which are more likely to produce a null allele than small indels, were

*To whom correspondence should be addressed. Email: rain_duly@163.com
Correspondence may also be addressed to Xiaotang Wang. Email: wangx@fiu.edu
Correspondence may also be addressed to Guohua Zhou. Email: ghzhou@nju.edu.cn
Correspondence may also be addressed to Libin Wei. Email: wlbiws_1986@aliyun.com
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

observed in the genomic DNA of SGN-treated cells. These features make SGN an alternative tool for cleaving DNA targets.

However, the cleavage efficiency of SGN is slightly low for practical applications. The SGN's loose 'hook'—the guide DNA probe (gDNA), which diffuses separately from nuclease in cells, is considered to be responsible for the low efficiency. The presence of these two individual molecules at the same target locus is a relatively low-probability event. Moreover, the specificity of SGN warrants further improvement. The SGN's 'scissors', i.e. the Fn1 domain, is a non-specific endonuclease (21) that leads to off-target effects in genomic DNA. ZFN, another nuclease that uses Fn1 as 'scissors', is able to avoid off-target effects by optimizing the Fn1 domain to selectively reduce activity at low-affinity off-target sites (22). It remains undefined whether SGN can cleave RNA. Therefore, further study is required to obtain more precise and efficient nucleotide editing tools.

In fact, the FEN1 domain in SGN is an endonuclease with high efficiency and specificity. FEN1 has been widely used in the sensitive and specific detection of single-base mutations (23–26). During DNA replication and repair processes, FEN1 participates in removing either RNA primers or damaged DNA (27), demonstrating the versatile function of FEN1 in cleaving both RNA and DNA. The newly synthesized DNA and the displaced region compete for base pairing with the template strand, resulting in the formation of a double-flap structure (27). FEN1 recognizes this special structure by the 3′ flap and catalyses phosphodiester cleavage of the 5′ flap (28). However, in the SGN system, the gDNA is a linear DNA probe that forms only a 3′ flap structure with the target DNA. In this 3′ flap structure, FEN1 can play a role only in capture but not cleavage.

To maintain the advantage of lacking sequence limitation and improve the efficiency and specificity of both DNA and RNA cleavage, we report a hairpin DNA probe-SGN system (hereafter designated the HpSGN system). The HpSGN system is composed of FEN1 nuclease (gray) and a hairpin DNA probe (designated hpDNA below, green), as shown in Figure 1A. The crystal structure of *A. fulgidus* FEN1 was reported (29). The 3′ part of hpDNA is a stem-loop structure, and FEN1 captures hpDNA by this stem–loop structure to form an FEN1-hpDNA binary complex. The 5′ part of hpDNA (i.e. guide sequence, ≥11 nt in size) is complementary to its target (red) and further forms a FEN1–hpDNA–target ternary complex. In this ternary complex, the target is positioned at the 5′ flap as the substrate of FEN1 cleavage. FEN1 cuts the target between bases 1 and 2 (Figure 1A). With the FEN1-hpDNA binary complex, the HpSGN system is more catalytically competent and offers a higher probability of cleaving the target than the SGN system, in which the gDNA and SGN protein work individually. Furthermore, the low molecular weight (∼35 kDa) of FEN1 is potentially beneficial for *in vivo* delivery along with DNA donors via adeno-associated viruses. This programmable HpSGN system will be tested *in vitro* and *in vivo* to develop an alternative effective and specific editing tool for DNA and RNA without sequence limitation.

MATERIALS AND METHODS

FEN1-expressing construct generation

Different constructs expressing FEN1 were generated. *A. fulgidus* FEN1 ORF was synthesized and inserted into the prokaryotic expression plasmid pET28a(+) to generate pET28a(+)-*A. fulgidus* FEN1. HA tag and NLS were added to the N-terminus of both *A. fulgidus* and *Homo sapiens* FEN1 and then subcloned to pcDNA3.1(+) to generate pcDNA3.1(+)-HA-NLS-*A. fulgidus* (or *Homo sapiens*) FEN1. For mRNA cleaving, NLS was substituted by NES. For MG1655 genomic DNA editing, synthesized *A. fulgidus* FEN1 ORF was subcloned to a vector containing the Tc-inducible promoter, chloramphenicol-selectable marker and p15A replication origin.

Expression and purification of FEN1 recombinant protein

The pET28a(+)-FEN1 was transfected into host bacteria, BL21(DE3), with the $CaCl_2$-heat-shock method. Briefly, cells were cultured at 37°C, and induced with IPTG (0.1 mM) at 25°C for 16 h to express FEN1. The induced cells were collected, lysed with ultrasound and centrifuged. FEN1 was purified from the crude extract by Ni-affinity chromatography (Supplementary Figure S1). FEN1 was then concentrated with ultrafiltration and confirmed with sodium dodecyl sulphate (SDS)-polyacrylamide gel electrophoresis (PAGE; 10%). The molecular weight of FEN1 was ∼35 kDa.

Cleaving ssDNA *in vitro*

A 10-μl reaction mixture consisting of hpDNA or gDNA (5 pmol), single-strand DNA (ssDNA) substrate (5 pmol) (Supplementary Table S1), MOPS (10 mM), 0.05% Tween-20, 0.01% nonidet *P*-40, $MgCl_2$ (7.5 mM). The 5′ end of the target ssDNA was labelled with fluorescent group Cy5. Before enzyme was added, the mixture was incubated at 95°C for 5 min and 55°C for 10 min. Then, FEN1 or SGN was added, and the reaction was incubated at 37°C for 2 h.

Cleaving ssRNA *in vitro*

A 10-μl reaction mixture consisting of hpDNA (10 pmol), single-strand RNA (ssRNA) substrate (10 pmol) (Supplementary Table S1), MOPS (10 mM), 0.05% Tween-20, 0.01% nonidet *P*-40, $MgCl_2$ (7.5 mM), and appropriate amount of FEN1 was prepared and incubated at 37°C for 2 h. The 5′ end of the target ssRNA was labelled with fluorescent FAM. The tips, tubes and reagents used in this experiment were all treated with DEPC (Diethyl pyrocarbonate).

Denatured-polyacrylamide gel electrophoresis (PAGE)

The products derived from the above reactions were analyzed with PAGE under denaturing conditions. The loading buffer contained 90% formamide, 0.5% EDTA, 0.1% xylene cyanol, and 0.1% bromophenol blue. Before loading, the samples (20-μl) were incubated in boiling water for 5 min and then cooled on ice. The samples were then loaded onto a 20% PAGE gel at room temperature, and ran in a buffer

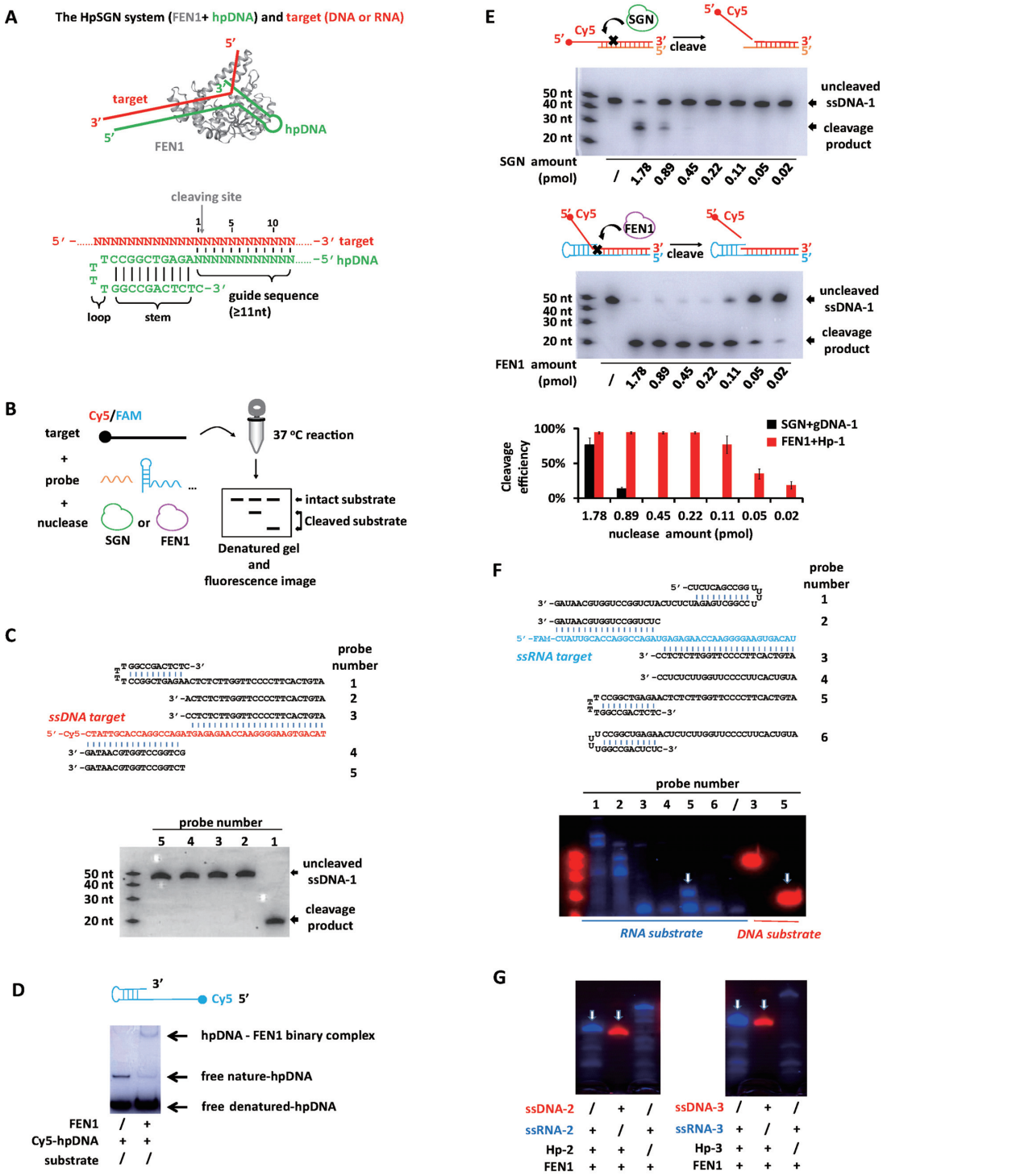

Figure 1. FEN1 cleavage guided by the hpDNA is highly effective *in vitro.* (**A**) Schematic diagram of the HpSGN system including FEN1 protein (gray, PDB ID 1RXV), the hpDNA probe (green), and the target (red). (**B**) Graphic illustration of the *in vitro* biochemical analysis. (**C**) Denaturing gels showing FEN1 catalyzed cleavage of target guided by different types of probes. The probes were hairpin DNA probe (hpDNA, probe number 1), guide DNA probe without mismatch bases at the 3′ end (probe number 2), guide DNA probe with a mismatch base at the 3′ end (gDNA, probe number 3), guide DNA probe with a mismatch base at the 5′ end (probe number 4) and guide DNA probe without mismatch bases at the 5′ end (probe number 5). (**D**) Electrophoretic mobility shift assay showing different migration velocity of hpDNAs with or without FEN1. (**E**) Denaturing gels (top and middle panels) and quantitative graph (bottom panel) showing the cleavage efficiencies of target DNA substrate 1 (ssDNA-1) by different amounts of nucleases using guide DNA probe 1 (gDNA-1, top panel) and hairpin DNA probe 1 (Hp-1, middle manel), $n = 3$. (**F**) Cleavage of ssRNA mediated by FEN1 can only be guided by DNA probes but not RNA probes. Probes used were hairpin guide RNA probe with 5′ flap structure (probe number 1), RNA probe with 5′ flap structure (probe number 2), DNA probe with 3′ flap structure (probe number 3, gDNA), RNA probe with 3′ flap structure (probe number 4), hairpin guide DNA probe with 3′ flap structure (probe number 5, hpDNA) and hairpin guide RNA probe with 3′ flap structure (probe number 6). (**G**) Cleavage of RNA substrate 2 (ssRNA-2, left panel) and RNA substrate 3 (ssRNA-3, right panel) and their ssDNA counterparts by the HpSGN system.

containing urea (8.7 M) and Tris-borate (89 mM). The electrophoresis was run at 9.6 V/cm for 2 h. The gel was imaged by Amersham Imager 600 (GE Healthcare).

Electrophoretic mobility shift assay (EMSA)

A 10-μl reaction mixture consisting of hpDNA (5 pmol) (Supplementary Table S1), MOPS (10 mM), 0.05% Tween-20, 0.01% nonidet *P*-40, $MgCl_2$ (7.5 mM) and the suitable amount of FEN1 was incubated at 37°C for 0.5 h. The 5′ end of the hpDNA was labelled with fluorescent group Cy5. The reacting products were analyzed with 12% native-PAGE. After electrophoresis, the gel was imaged by Amersham Imager 600 (GE Healthcare).

Sequence preference validation

A DNA oligo with a degenerate 7-bp sequence and its hpDNA (Supplementary Table S1) were incubated with FEN1 *in vitro* followed by agarose electrophoresis. The uncleaved (intact) fragment was recovered and amplified by PCR. The PCR program was 95°C 5 min; 35 cycles of (95°C 30 s, 60°C 30 s, 72°C 10 s). The amplicons were cloned into T vector and 50 clones of each groups were sequenced. The frequency of 4 bases at the degenerate 7-bp sequence was analyzed.

DNA editing in bacterial cells

The constructed plasmid expressing FEN1 was transfected into *E. coli* MG1655 or BL21(DE3) bacterial cells. Positive clones were selected and cell cultures were grown at 37°C. The cells were harvested at 4°C and were used for transformation of the hpDNA by electroporation. For electroporation, 50 μl of cells was mixed with hpDNAs targeting *narG* gene for MG1655 (500 pmol, Supplementary Table S1) and hpDNA targeting *GFP* gene for BL21(DE3) (500 pmol, Supplementary Table S1); electroporation was done in a 2-mm Gene Pulser cuvette (Bio-Rad) at 2.5 kV. MG1655 cells were recovered at 37°C for 1 h and induced 1 h by aTc (200 ng/ml) before being diluted 100 000 times and spread onto LB agar containing chloramphenicol (25 mg/ml), then incubated overnight at 37°C. The number of colonies on plates in different groups were counted. BL21(DE3) cells were recovered at 37°C for 1 h and induced 1 h by IPTG (0.1 mM) before fluorescence detection.

Mammalian cell culture and transfection

Cells were maintained in Dulbecco's modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (HyClone), 100 U/ml penicillin, and 100 μg/ml streptomycin at 37°C with 5% CO_2 incubation. Cells were seeded into six-well plates (Corning) 24 h prior to transfection at a density of 200 000 cells per well. Cells were transfected by using Lipofectamine 2000 (Life Technologies) following the manufacturer's recommended protocol. For each well of a six-well plate, 4 μg plasmids were transfected and 500 pmol of hpDNAs (Supplementary Table S1) were transfected 24 h later.

Immunofluorescence (IF)

Cells were fixed with 4% paraformaldehyde in PBS at 1 h intervals, permeabilized in 0.2% (v/v) Triton X-100 for 20 min, and blocked with 3% BSA for 30 min. Incubation with primary antibodies (SANTA, 1:50) against HA tag was done overnight at 4°C. The coverslips were washed three times with PBS and followed by co-incubation with Alexa Fluo 488 Goat anti-Rabbit IgG (H+L) antibody (1:200) for 1 h at 37°C. The nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI, Sigma-Aldrich) for 20 min before imaging. The coverslips were inverted onto slides and immersed in a mounting medium. A laser scanning confocal microscope FV10-ASW [Ver 2.1] (Olympus Corp, MPE FV1000) was used for co-localization analysis.

Western blot (WB)

Protein samples were isolated with lysis buffer, eluted with SDS buffer, separated by SDS-polyacrylamide gels, and electroblotted onto PVDF membranes. The specific protein bands were stained with High-sig ECL Western Blotting Substrate (Tanon) and imaged using the Amersham Imager 600 (GE Healthcare). Primary antibodies included antibodies against β-Actin (Abclonal Technology, AC026, 1:150 000), HA (Santa Cruz Biotech, sc-7392, 1:200), CDK9 (Cell Signaling Technology, C12F7, 1:2000) and AFP (Abclonal Technology, A11865, 1:1000).

PCR and Sanger sequencing analysis for HEK293A cells genome modification

Genomic DNA was extracted using the QIAamp DNA Mini Kit (Qiagen) following the manufacturer's protocol. Genomic region surrounding the FEN1 target site for each gene was PCR amplified. To reduce nonspecific PCR products, nested-PCR and high-fidelity DNA polymerase (PrimeSTAR® HS DNA Polymerase, Takara) were used. Genomic DNA (50 ng) was added as template with the out-primer pairs (EMX1-OutF and EMX1-OutR, DYRK1A-OutF and DYRK1A-OutR, Supplementary Table S1) in 20-μl PCR mixture. The first PCR program was 98°C for 5 min, 35 cycles of (98°C for 10 s, 60°C for 5 s, and 72°C for 360 s), 72°C for 10 min. The amplicons were diluted 100 times as the templates with the in-primer pairs (EMX1-inF and EMX1-inR, DYRK1A-inF and DYRK1A-inR, Supplementary Table S1) in the second PCR amplification. The second PCR program was 98°C for 5 min, 35 cycles of (98°C for 10 s, 58°C for 5 s, and 72°C for 140 s), 72°C for 10 min. The products were purified using QiaQuick Spin Column (Qiagen) following manufacturer's protocol and analyzed on 1% agrose gels. Gels were stained with GelRed stain and imaged. For Sanger sequencing analysis, PCR products corresponding to genomic modifications were purified and cloned into the T-Blunt vector (pMD18-T, Takara). Then, transformants for each group were randomly selected and sequenced to identify mutations using the M13 primer.

Genomic DNA sequencing

Given the mechanism of inducing mutation by our HpSGN system, the positions of these mutation were not fixed and

ranged about kilo bases up/down stream. The target region deep sequencing approach, the library of which is ∼200 bp, can not be used to estimate the HpSGN system. Here whole genome sequencing approach was used. A total amount of 0.5 μg DNA per sample was used as input material for the DNA library preparations. Sequencing library was generated using Truseq Nano DNA HT Sample Prep Kit (Illumina) following manufacturer's recommendations and index codes were added to each sample. Briefly, genomic DNA sample was fragmented by sonication to a size of 350 bp. Then DNA fragments were end-polished, A-tailed and ligated with the full-length adapter for Illumina sequencing, followed by further PCR amplification. After PCR products were purified (AMPure XP system), libraries were analyzed for size distribution by Agilent 2100 Bioanalyzer and quantified by real-time PCR (3 nM). The clustering of the index-coded samples was performed on a cBot Cluster Generation System using Hiseq X PE Cluster Kit V2.5 (Illumina) according to the manufacturer's instructions. After cluster generation, the DNA libraries were sequenced on Illumina Hiseq platform and 150 bp paired-end reads were generated. Valid sequencing data was mapped to the reference human genome by Burrows–Wheeler Aligner (BWA) software (30) to get the original mapping results stored in BAM format. If one or one paired read(s) were mapped to multiple positions, the strategy adopted by BWA was to choose the most likely placement. If two or more most likely placements presented, BWA picked one randomly. Then, SAMtools (31) and Picard (http://broadinstitute.github.io/picard/) were used to sort BAM files and do duplicate marking, local realignment, and base quality recalibration to generate final BAM file for computation of the sequence coverage and depth. The large fragment mutation (including deletion, translocation, insertion, inversion) were expected induced by the HpSGN system with HR or NHEJ. The 'mutation reads divided by total reads on target region' was used to generally estimate the modification efficiency, with a possibility of inaccuracy for the depth of whole genome sequencing approach.

Flow cytometry

Cells were resuspended in PBS, and the fluorescence intensity (488 nm excitation and 525 nm emission) was measured immediately using FACSCalibur (Becton Dickinson). The efficiency of EGFP down-regulation was calculated according to the following equation.

$$\text{Efficiency } (\%) = \frac{(\text{Fc} - \text{Fa}) - (\text{Fc} - \text{Fb})}{\text{Fc}} = \frac{(\text{Fb} - \text{Fa})}{\text{Fc}}$$

Fa, the normalized fluorescence intensity of the group transfected with FEN1 and hpDNA; Fb, the normalized fluorescence intensity of the group transfected with FEN1; Fc, the normalized fluorescence intensity of the control group.

RNA sequencing

Total RNA was prepared using a RNA Isolation Reagent (R701-01, Vazyme). A total amount of 1 μg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext UltraTM RNA Library Prep Kit for Illumina (NEB) following manufacturer's recommendations and index codes were added to attribute sequences to each sample. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumia) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina Novaseq platform and 150 bp paired-end reads were generated. FeatureCounts v1.5.0-p3 was used to count the reads numbers mapped to each gene. And then counts of each gene was calculated based on the length of the gene and reads count mapped to this gene.

Guide:target base-pairing mismatch search

The methods for the identification of potential off-target sites were based on Watson–Crick base pairing mismatch between the guide sequence and targets. Alignment was manually adjusted to allow for insertion and deletion mismatches in the guide:target heteroduplex.

Statistical analysis

All data were expressed as mean ± SD or mean ± SEM from two or three independent experiments performed in a parallel manner. Comparisons between two groups were analyzed using two-tailed Student's *t*-tests. P values <0.05 were considered statistically significant.

RESULTS

FEN1 cleavage guided by hpDNA is highly effective *in vitro*

To verify that FEN1 can be developed as an effective target DNA cleavage tool, the cleavage efficiency of the HpSGN system was tested. The targets were ssDNA modified with the Cy5 group. Theoretically, only the intact target strand (uncleaved) and the cleaved products containing the 5′ end of targets with the Cy5 group can be visibly detected (Figure 1B). As shown in Figure 1C, five kinds of DNA probes were used to guide the FEN1-catalysed cleavage of targets, and only hpDNA (probe number 1) successfully facilitated cleavage. In contrast, the gDNA (probe number 3) did not work (Figure 1C, Supplementary Figure S2).

As demonstrated by EMSA, the hpDNA and FEN1 molecules formed a binary complex before encountering their targets (Figure 1D). It was postulated that the formation of the hpDNA-FEN1 binary complex could increase the probability of cleaving targets compared to the SGN system, in which the gDNA and SGN molecules function individually. To test our hypothesis, we compared the efficiency of SGN plus gDNA and FEN1 plus hpDNA. As shown in Figure 1E, 1.78 and 0.89 pmol of SGN cleaved part of ssDNA-1 with gDNA-1, while 0.45, 0.22, 0.11, 0.05 and 0.02 pmol of SGN did not. Meanwhile, 1.78, 0.89, 0.45 and 0.22 pmol of FEN1 cleaved all ssDNA-1 with hpDNA-1. Lower concentrations (0.11, 0.05 and 0.02 pmol) of FEN1 also cleaved part of ssDNA-1 with hpDNA-1, whereas corresponding concentrations of SGN failed to function. These results demonstrated that ∼40 times more

nuclease was needed to cleave the target with SGN plus gDNA than with FEN1 plus hpDNA. To further confirm this conclusion, we repeated the tests with 0.11 pmol of SGN and FEN1 on two other DNA substrates (i.e. ssDNA-2 and ssDNA-3) with different sequences. The results (Supplementary Figure S3) corroborated our results shown in Figure 1E, demonstrating that FEN1 plus hpDNA performs much better than SGN plus gDNA in substrate cleavage.

Next, we added hpDNA-ccr5 (hpDNA for substrate CCR5) instead of hpDNA-2 into the reaction with ssDNA-2 to determine whether miscleavage would occur. As expected, no cleavage was observed (Supplementary Figure S4). When we mixed hpDNA-ccr5, ssDNA-2 and ssDNA-ccr5 together, only ssDNA-ccr5 but not ssDNA-2 was cleaved (Supplementary Figure S4). These results demonstrated that collateral cleavage did not occur in the HpSGN system.

Given that FEN1 is known to have RNA template-dependent 5′ nuclease activity to handle Okazaki fragments (27), we tested whether FEN1 could be reprogrammed to cut selected synthetic ssRNAs *in vitro*. The ssRNAs were labelled with a fluorescent FAM group on the 5′ end (Figure 1B). We mixed synthetic ssRNA substrates with RNA probes and DNA probes to detect the RNase function of FEN1 in comparison with a control group for ssDNA substrates (Figure 1F). No obvious cleavage products were observed with other probes except for hpDNA (probe number 5), which enabled FEN1 to cut not only the DNA substrate but also the RNA substrate. To confirm this result, we repeated these tests on two other ssRNA substrates with different sequences (i.e. ssRNA-2 and ssRNA-3). Guided by hpDNAs, both ssRNA targets were cleaved by FEN1 (Figure 1G). This result substantiated that FEN1 could recognize and capture hpDNAs and then cleave the corresponding substrates, regardless of whether the substrates were DNA or RNA.

FEN1 cleavage is sensitive to mismatches in the probe-target duplex *in vitro*

Similar to CRISPR-associated systems, most off-targets occur due to misannealing between the guide sequence (the 5′ part of hpDNA) and the wrong locus in genomic DNA. However, it is almost impossible to completely avoid the annealing of guide sequences with wrong loci in genomic DNA. The solution is to determine where and how much misannealing can be tolerated. Therefore, it is necessary to investigate the HpSGN system's sensitivity to mismatches in the probe–target duplex. The higher the sensitivity is, the lower is the possibility of off-target incidents.

First, we compared the effects of hpDNAs of different lengths. As Figure 2A shows, the cleavage efficiency of ssDNA-4 was noticeably reduced when the length of the guide sequence was 10 nt. Then, we tested the sensitivity of the HpSGN system to a single mismatch in the probe-target duplex by mutating single bases into their respective complementary bases (A to T, T to A, G to C and C to G). For hpDNAs with a 20 nt guide sequence, the HpSGN system was tolerant to single mismatches far away from the stem-loop structure but sensitive to single mismatches next to the stem-loop structure (Figure 2B). The test was repeated using hpDNAs with 12 nt (Figure 2C) and 11 nt (Figure 2D) guide sequences. Tests with shorter probes showed that the HpSGN system was very sensitive to single mismatches. We repeated the tests with another DNA substrate (ssDNA-ccr5) and observed similar results (Supplementary Figure S5). Interestingly, the introduction of two consecutive substitutions of the respective complementary bases resulted in a marked reduction in the cleavage of the double mismatches (Supplementary Figure S6). This observation suggested the presence of a mismatch-sensitive 'seed region', specifically the single base next to the hairpin (position 1). The longer the length of the guide sequence was, the narrower was the range of the 'seed region'. This region can decrease the off-target rates by avoiding consecutive annealing between potential off-target sites and bases close to the 'stem–loop' structure in guide sequences.

In summary, the efficiency and specificity of the HpSGN system are sensitive to mismatches in the probe-target duplex. The length of the guide sequence in hpDNA should be at least 11 nt to ensure appropriate and stable annealing with the targets. For applications with strict specificity requirements, such as gene therapy for diseases caused by single-base mutations, the guide sequence should be short. On the other hand, for applications requiring a wide range of coverage, such as cutting the coding gene of a virus with various subtypes, the guide sequence should be long.

Cleavage is sensitive to the structure of hpDNAs but has no requirement for the sequence of the target substrates

In contrast to previously reported gDNAs, hpDNAs contain a single stem-loop structure, suggesting that the secondary structure of the hpDNAs could facilitate the interaction with FEN1. Here, we explored the structural requirements of hpDNAs for mediating DNA cleavage.

We first studied the effect of base number in the stem and loop on DNA cleavage activity. We mutated the bases in the stem to T bases one by one to reduce the length of the stem, ranging from ten bases to five bases (Figure 3A, loop type 1, 2, 3, 4, 5 and wt). The cleavage efficiency of loop type 1 (five paired bases in the stem) and other loop types was reduced by ∼70% and ∼90%, respectively, compared to the wt loop type (Figure 3A), suggesting that too short a stem destroyed the stability of the hpDNA structure and decreased the cleavage activity of the HpSGN system. Next, we reduced the number of bases in the loop from four to one to reduce the size of the loop (Figure 3A, loop type 6, 7, 8 and wt). We observed that a reduction in the number of bases in the loop or the size of the loop failed to disrupt the stem-loop duplex structure or cleavage activity (Figure 3A). We further noted that, to some extent, even one T base in the loop can maintain the stability of cleavage activity. Furthermore, we tested the effects of different flap lengths. As shown in Figure 3B, overlap types 1, 2, 3 and 4 in Figure 3B had two, three, four and five C bases at the 3′ flap. Additionally, we tested hpDNA with a stem-loop structure at the 5′ end and observed that both moving the stem-loop structure to the 5′ end and extending the length of the 3′ flap almost completely disabled the cleavage activity (Figure 3B). Additionally, we determined the base type requirements at the

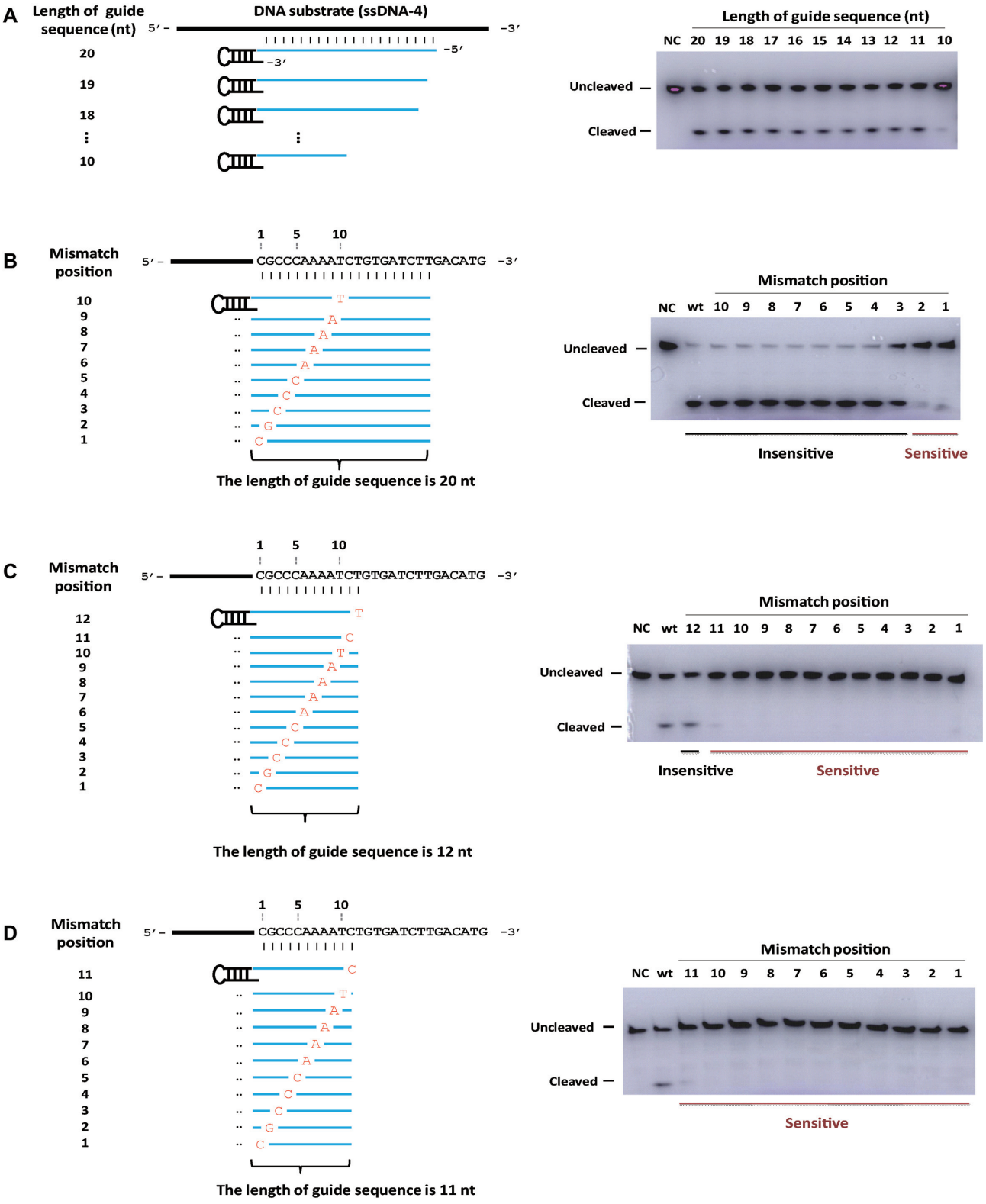

Figure 2. FEN1 cleavage is sensitive to mismatches in probe-target duplex. (**A**) Effect of guide sequences with different lengths ranging from 20 nt to 10 nt at an interval of 1-nt. (**B**) Effect of single mismatches at different positions in hpDNAs with 20 nt guide sequence (mismatch position 1 to 10 with the mismatched based highlighted red). (**C**) Effect of single mismatches at different positions in hpDNAs with 12 nt guide sequence (mismatch position 1 to 12 with the mismatched base highlighted red). (**D**) Effect of single mismatches at different positions in hpDNAs with 11 nt guide sequence (mismatch position 1 to 11 with the mismatched base highlighted red). The wt in electrophoretograms means no mismatch base in the corresponding guide sequence, while NC stands for no hpDNAs in the corresponding groups.

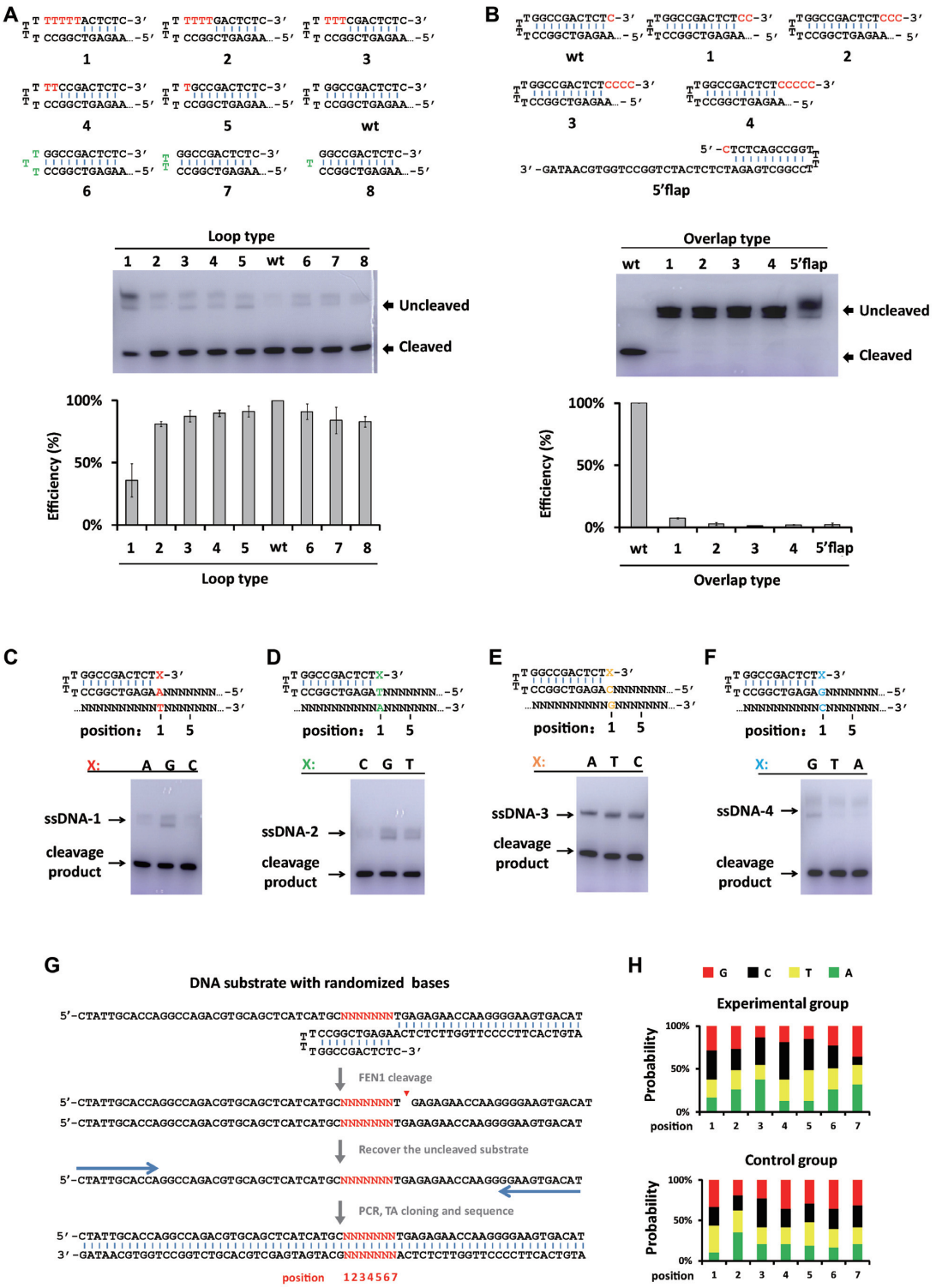

Figure 3. The cleavage is sensitive to the structure of hpDNAs and has no requirement of target sequence. (**A**) FEN1 cleavage activity with hpDNAs having 9, 8, 7, 6, 5, 4, 3, 2 and 1 T base in the loop, termed as loop types 1, 2, 3, 4, 5, wt, 6, 7 and 8, respectively. The intensities of bands marked with 'cleaved' at different loop types were quantified as relative cleavage efficiencies ($n = 3$). Base T was marked red if the number of base T is more than the size of the conventional loop (wt), but green if less than that of wt. (**B**) FEN1 cleavage activity with conventional hpDNAs having 1, 2, 3, 4 and 5 C base (marked red) in the overlap region (the gap between the stem and the guide sequence), termed as the overlap types of wt (regular 1-base overlap), 1, 2, 3 and 4, respectively. Cleavage activity at mutated hpDNAs formed with the loop-stem structure at the 5′ end of the guide sequence was termed as '5′ flap'. The intensities of bands marked with 'cleaved' at different mutation types were quantified as relative cleavage efficiencies ($n = 3$). (**C**)–(**F**) show the requirement of base type for X when the base at position 1 is T, A, G and C. (**G**) The schematic illustration of the *in vitro* cleavage-based method used to identify the first seven positions near the 5′ end of probe-target duplex. (**H**) The detection probability of G, C, T and A base in the recovered uncleaved fragment from the groups of control and FEN1-treated samples.

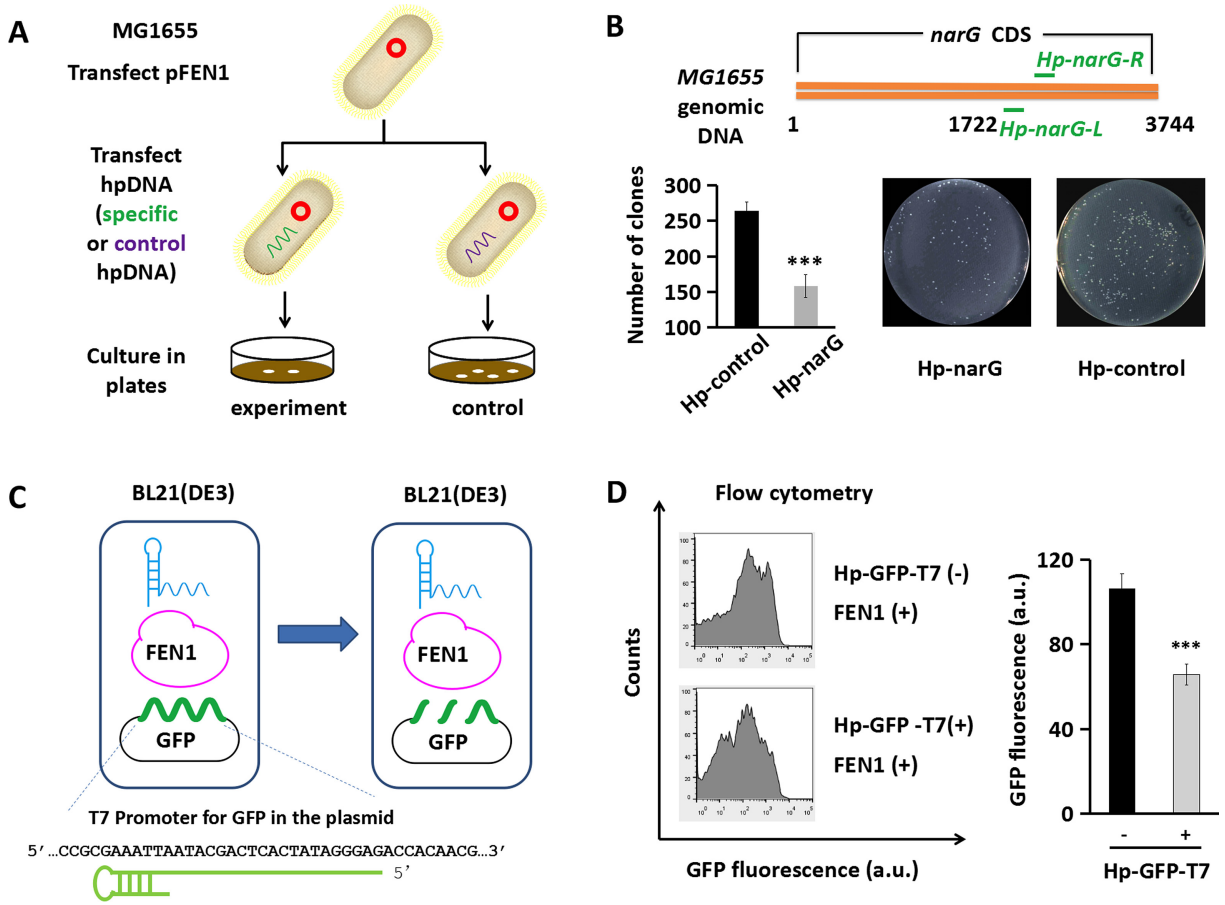

Figure 4. FEN1 can edit genomic DNA in *E. coli.* (**A**) Schematic diagram of hpDNA-guided cleavage of genomic DNA in MG1655 heterologously expressing FEN1. (**B**) The location of hpDNAs for the *narG* gene (Hp-narG-L/R) in MG1655 and sample plates showing different number of clones in control and experimental groups. $n = 3$. *** P value <0.001. (**C**) Schematic diagram of hpDNA-guided break of GFP plasmid in BL21(DE3). (**D**) The fluorescence of groups transfected with FEN1 plus hpDNA probe (Hp-GFP-T7) targeted for *GFP* targeting. $n = 3$. ***P value <0.001.

3′ flap position (the X in Figure 3C–F). For targets with base T at position 1 (such as ssDNA-1 here), similar cleavage efficiency was obtained by changing X from C to G or A (Figure 3C). For the targets with base A, G or C at position 1 (such as ssDNA-2, ssDNA-3 and ssDNA-4 in this test, respectively), this conclusion was still valid (Figure 3D–F).

To identify potential sequence requirements for FEN1 cleavage, we synthesized a DNA fragment with a degenerate 7-bp sequence (5′-NNNNNNN) near the 5′ end of a 25-bp probe-target duplex (Figure 3G). This fragment was incubated with FEN1 plus hpDNA *in vitro* followed by agarose electrophoresis. The uncleaved (intact) fragment was recovered and amplified by PCR. The amplicons were cloned into the T vector and sequenced. We analysed the occurrence frequency of four bases at positions 1 to 7 and observed no obvious sequence preference (Figure 3H). It can be inferred that unlike the CRISPR system, the HpSGN system picked the target substrate without strict requirements for the target sequence.

The HpSGN system can edit plasmids and genomic DNA in *E. coli*

The successful and efficient cleavage of both RNA and DNA substrates by FEN1 in our *in vitro* study described above encouraged us to investigate whether FEN1 can also effectively cleave bacterial genomic DNA. For this purpose, we transfected plasmids expressing FEN1 (tet-induce) and hpDNAs targeting the endogenous nitrate reductase G (*narG*) gene into MG1655 cells, as shown in Figure 4A. Because bacterial cells have very weak self-repair ability, cleavage-induced breaks in bacterial genomic DNA without an assisting repair system are lethal. Therefore, the number of clones on plates will decrease if FEN1 can cause breaks in the genomic DNA of MG1655 cells. Successful expression of FEN1 in MG1655 cells induced by the tetracycline promoter was confirmed (Supplementary Figure S7). To measure the efficiency of editing, we transfected irrelevant hpDNA into MG1655 cells as control groups and counted the clones on plates. As shown in Figure 4B and Supplementary Figure S8, the control group had an average of 263 clones, and the experimental group had an average of only 158 clones, demonstrating that FEN1 can cause breaks in the genomic DNA of MG1655 cells.

Next, we tested the HpSGN system in *E. coli* BL21(DE3) cells targeting the *GFP* gene on a plasmid (Figure 4C). The hpDNA targeting *GFP* was located on the T7 promoter (Hp-GFP-T7). The fluorescence intensity in the group with Hp-GFP-T7 was only 60% of that in the group without hpDNA (Figure 4D). Our results demonstrated that FEN1

could effectively cleave target sites to cause breaks within genomic DNA, suggesting that the HpSGN system has potential applications in editing genomic DNA in *E. coli* through chromosomal integration.

The HpSGN system can be harnessed to facilitate genome editing in human cells

The encouraging results from our *in vitro* and *in vivo* study on *E. coli* with the HpSGN system prompted us to test the possibility of using this system in human cells (Figure 5A). Therefore, we coded the *A. fulgidus*/*Homo sapiens* FEN1 gene and attached an N-terminal nucleus localization signal (NLS) for optimal expression and nucleus targeting in human cells (Figure 5B).

To test whether heterologous expression of FEN1 can achieve targeted cleavage of human chromosomes, a pair of hpDNA targeting sites within the *EMX1* gene in human embryonic kidney 293A cells (HEK293A) was selected (Figure 5C), and the results were subsequently verified by PCR, TA cloning and Sanger sequencing. As shown in Table 1 and Supplementary Figures S9 and S10, an average of 21.4% ($n = 3$, 1/6, 3/9 and 1/7, respectively, for *A. fulgidus* FEN1) and 22.6% ($n = 3$, 2/6, 1/8 and 2/9, respectively, for *Homo sapiens* FEN1) nuclease-induced mutations in the *EMX1* gene were detected. Similarly, an average of 17.8% ($n = 3$, 1/7, 2/8 and 1/7, respectively, for *A. fulgidus* FEN1) and 12% ($n = 3$, 2/8, 0/6 and 1/9, respectively, for *Homo sapiens* FEN1) nuclease-induced mutations in the *DYRK1A* gene were detected (Table 1, Figure 5D, Supplementary Figures S11 and S12).

Small indels are the major types of mutations created by ZFN, TALEN and CRISPR-associated systems; however, larger fragment mutations that are more likely to create a null allele than small indels were detected in our study. As shown in the left panel of Figure 5E, when a pair of 'tail to tail' hpDNAs were located on the target genomic DNA, FEN1 made two nicked breaks by its 5′ flap endonuclease activity. The nicked breaks could in turn be captured and expanded by the 5′ exonuclease activity of FEN1 (20). Therefore, large fragment deletions were made in this way. The nicked breaks were also recognized by DNA polymerase (Figure 5E, right panel). The nicked breaks moved in the 3′ direction by DNA extension, and indels were made by the DNA repair system. However, the high expression of FEN1 in the cells used in our study tended to cause the repair pathway to be overridden by the deletion pathway (Figure 5E, left).

However, the results drawn from the analysis of a limited number of clones alone are not enough to claim that the HpSGN system has the ability to edit genomic DNA. Therefore, a high-throughput sequencing method is necessary to estimate the potential universal application of the HpSGN system in editing genomic DNA and to evaluate the frequency of off-target incidents. Large fragment mutations (including deletion, translocation, insertion, and inversion) were expected to be induced by the HpSGN system with HR or NHEJ. The occurrence frequency of these mutation reads was used to generally estimate the overall modification efficiency. Obvious mutations were observed at the *DYRK1A* locus in the corresponding samples (Figure 5F and Supplementary Figure S13). Compared with that previously reported for SGN, the efficiency of the HpSGN system displayed an improvement of up to ∼20% (Figure 5F). Nonetheless, cleavage by the HpSGN system at 3 of the 5 off-target sites tested did take place at different levels (Figure 5F, Supplementary Figure S14). Therefore, we conducted additional analysis of the *EMX1* locus. Similar results for on- and off-target modifications were observed (Figure 5F, Supplementary Figures S15 and S16). These observations indicated that a degree of off-target modification did occur in HpSGN-treated cells. The levels of off-target effects depended on several factors, such as the number and position of mismatch bases, chromatin secondary structure, and DNA methylation status.

In vitro analysis indicated that the HpSGN system was sensitive to a single-base mismatch next to the hairpin structure. To further understand the sensitivity of the HpSGN system to mismatches *in vivo*, the *EGFP* gene was stably transfected into the genomic DNA of HeLa cells, and a pair of hpDNAs (Hp-EGFP-L and Hp-EGFP-R) was designed (Figure 5G). The *EGFP* guide sequence in hpDNA was systematically mutated to introduce single mismatches at different positions. The efficiency of EGFP downregulation was detected by flow cytometry. Compared with a group of five types of Hp-EGFP-L (Hp-EGFP-L-wt, -m3, -m5, -m7 and -m9), Hp-EGFP-L-m1 showed a noticeably lower level of downregulation. Similar results were obtained with Hp-EGFP-R-m1 compared with Hp-EGFP-R-wt, -m3, -m5, -m7 and -m9. These observations were in good agreement with the results from our *in vitro* studies that the HpSGN system was tolerant of single mismatches far away from the stem-loop structure but sensitive to single mismatches near the stem-loop structure. Furthermore, FEN1 was more sensitive to mismatches in Hp-EGFP-R. In fact, it is inevitable that NLS-FEN1 expressed in the nucleus would encounter *EGFP* mRNA and genomic DNA simultaneously with Hp-EGFP-R in the coding regions. The high sensitivity may result from an additive effect of the combination of DNA cleavage and RNA cleavage.

FEN1 can be reprogrammed to mediate specific mRNA knockdown in human cells

To investigate whether the HpSGN system can be reprogrammed to mediate specific mRNA knockdown *in vivo*, human cells were cotransfected with plasmids encoding FEN1 with NES and hpDNAs targeting specific mRNA (Figure 6A). FEN1 with NES mainly remained outside the nucleus (Figure 6B) and acted on mRNA molecules in the cytoplasm.

First, HeLa-EGFP cells were cotransfected with plasmids encoding FEN1 with NES and hpDNAs targeting the *EGFP* gene. The efficiency of *EGFP* gene downregulation in HeLa-EGFP cells was detected. The normalized fluorescence of the group transfected with both hpDNAs and FEN1 was ∼55% of that of the control group and ∼70% of that of the group transfected with FEN1 alone (Figure 6C), indicating that the HpSGN system can cleave mRNA targets in human cells with an efficiency of ∼23.4% (calculated by the equation described in Materials and Methods).

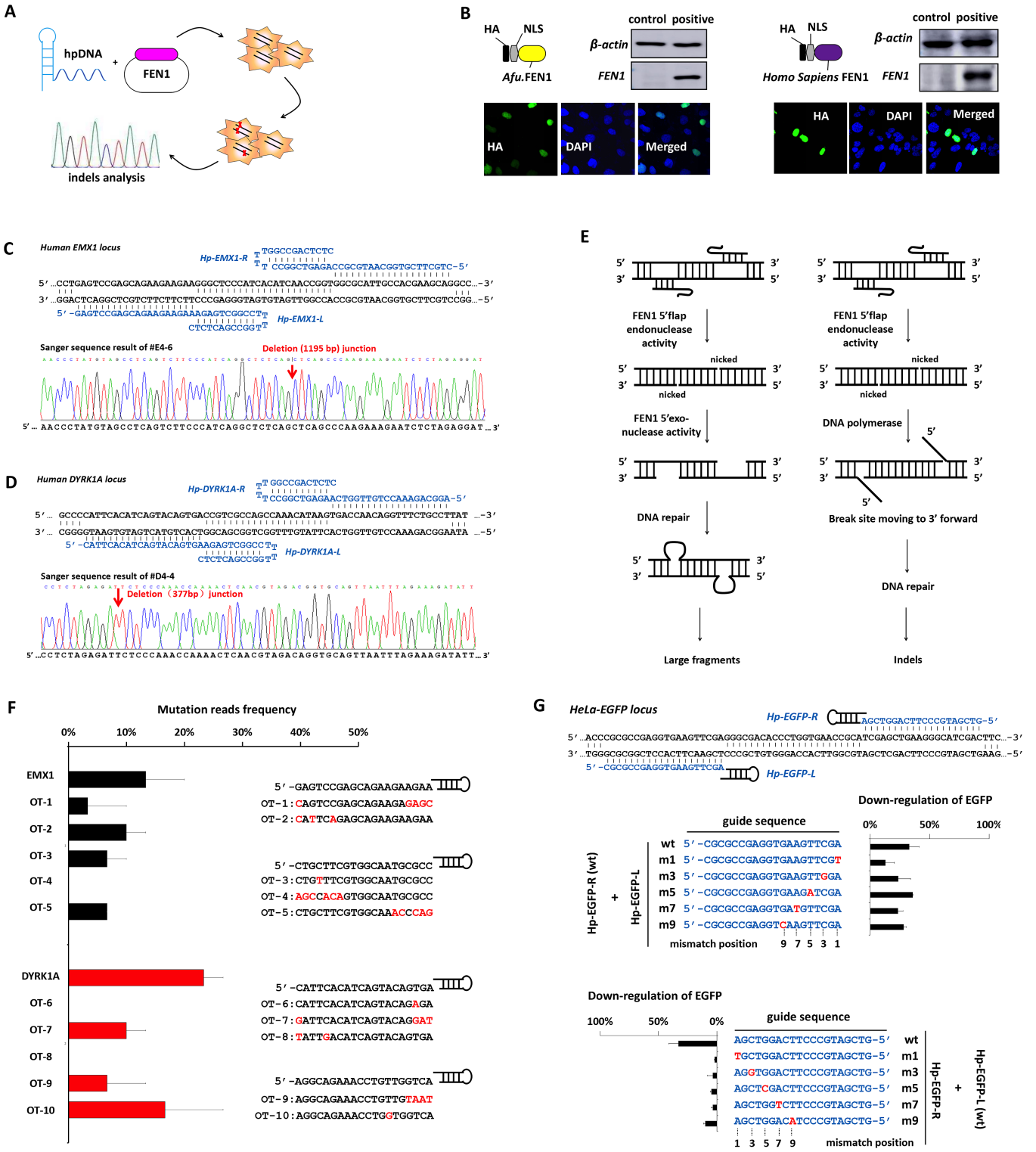

Figure 5. The HpSGN system can be harnessed to facilitate genome editing in human cells. (**A**) Schematic diagram of hpDNA-guided disruption of genomic DNA in HEK293A cells heterologously expressing FEN1. (**B**) Engineering of *A. fulgidus*/*Homo sapiens* FEN1 with NLS enables the import of FEN1 into mammalian nucleus. (**C**) The location of hpDNAs for the *EMX1* gene and a typical Sanger sequencing result of the HpSGN-edited product with large fragment deletions. (**D**) The location of hpDNAs for the *DYRK1A* gene and a typical Sanger sequencing result of the HpSGN-edited product with large fragment deletions. (**E**) The hypothesis of large fragment deletions by the HpSGN system. (**F**) The frequency of mutation reads of the *EMX1* gene and the *DYRK1A* gene edited by the HpSGN system. NGS was employed to analyze HpSGN-edited products. $n = 2$. (**G**) The efficiency of downregulation of EGFP in groups transfected with NLS-FEN1 plus wild/mutated probe of Hp-EGFP-L/R. $n = 2$.

Table 1. Mutations of *EMX1* and *DYRK1A* in the genome of HEK293A induced by the HpSGN system

Orthologue	Gene	Mutation/wild	Clone	Deletion (bp)	Insertion (bp)
A. fulgidus	*EMX1*	1/6	#E7-3	1	1
		3/9	#E8-3	1	/
			#E8-4	1	1
			#E8-7	776	/
		1/7	#E9-4	760	/
Homo sapiens	*EMX1*	2/6	#E4-5	127	/
			#E4-6	1195	/
		1/8	#E5-8	1	1
		2/9	#E6-1	1410	/
			#E6-3	1476	/
A. fulgidus	*DYRK1A*	1/7	#D8-3	448	/
		2/8	#D7-3	330	/
			#D7-5	442	/
		1/7	#D9-1	1	1
Homo sapiens	*DYRK1A*	2/8	#D4-2	96	/
			#D4-4	377	/
		0/6	#D5	/	/
		1/9	#D6-6	320	/

Second, we cotransfected liver hepatocellular cells (HepG2) with plasmids encoding FEN1 with NES and two groups of hpDNAs targeting the mRNA of the Alpha Fetoprotein (*AFP*) gene and the Cyclin-dependent kinase 9 (*CDK9*) gene. Compared with the control groups that were transfected with control plasmids and irrelevant hpDNAs, significant (∼25%) downregulation of AFP expression in the group transfected with both FEN1 and specific Hp-AFP was observed (Figure 6D). The location of Hp-AFP crossed two exons in genomic DNA. Therefore, the knockdown of AFP expression cannot be attributed to DNA cleavage. Knockdown of CDK9 was also observed (∼25%, downregulation) (Figure 6D). No indels were detected by Sanger sequencing of the DNA fragments of *CDK9* genes, suggesting that the knockdown of proteins was mainly due to the destruction of mRNA, not genomic DNA. Thus, it is concluded that FEN1 can be reprogrammed to cleave mRNAs preferentially *in vivo*.

Furthermore, RNA-seq was used to identify off-target effects on mRNA. The expression levels of all genes in RNA-seq libraries of nontargeting control groups and CDK9-targeting condition groups treated with the HpSGN system were analysed. However, the protein encoded by the *CDK9* gene is a member of the cyclin-dependent protein kinase (CDK) family. It is an elongation factor for the directed transcription of RNA polymerase II and plays a role through the C-terminal domain of the largest subunit of phosphorylated RNA polymerase II. As the downregulation of CDK9 affected the expression of many genes, it is not appropriate to evaluate off-target effects here. Given the mismatched hpDNA results *in vitro*, we expected that for any particular hpDNA, FEN1 may cleave mRNA loci that contain a small number of mismatched bases. For the CDK9 target, we computationally selected 13 candidate off-target sites in the human transcriptome that meet the following criteria (Figure 6E): (i) can be misannealed by Hp-CDK9 in positions 1 to 11, which makes them very likely to be cleaved by the HpSGN system and (ii) are not reported to have functional associations with CDK9 (Figure 6E). Then, the expression levels of candidate off-target genes in nontargeting control groups and CDK9-targeting condition groups were compared. As shown in Figure 6E, obvious downregulation was observed only in *CDK9* mRNA, *CDK10* mRNA and *MVD* mRNA but not in the other 11 candidates, indicating that a degree of off-target effects occurs in HpSGN-treated cells.

DISCUSSION

A novel HpSGN system for both DNA and RNA editing, composed of flap endonuclease 1 and a hairpin DNA probe, was designed and characterized in this study. *In vitro* biochemical studies showed that the HpSGN system required less nuclease to cleave ssDNA substrates than the SGN system we reported previously by a factor of ∼40 (19). The HpSGN system cleaved bacterial and human genomic DNA with ∼40% and 20% efficiency, respectively. Compared with the activity of SGN, which was tested in zebrafish cells, the activity of the HpSGN system was obviously improved. However, the efficiency of the HpSGN system was still not at the same level as that of CRISPR-Cas9 (8,9). In CRISPR-associated systems, although single guide RNAs (sgRNAs) are transfected into cells either as PCR amplicons or sgRNA-expressing plasmids (32), large quantities of sgRNA transcripts in the nucleus contribute to the high efficiency of the CRISPR system. Compared with the abundant quantities of sgRNAs in the nucleus, the amount of hpDNAs diffusing freely or carried by FEN1 into the nucleus is limited.

The sensitive dependence of FEN1 cleavage on mismatches in the probe-target duplex provides important information for designing new hpDNAs to further improve the sensitivity and specificity of the HpSGN system described here. We therefore offer some guidelines in this study to help users design HpSGN systems. (i) The guide sequence should consist of a minimum of 11 nt bases to ensure the smooth formation of a stable probe-target duplex with an appropriate Tm via favourable binding of the probe with the target. (ii) For applications that require strict specificity, the guide sequence should be short as long as a stable duplex can be formed, while for applications that require a wide range of coverage, the guide sequence should be long. (iii)

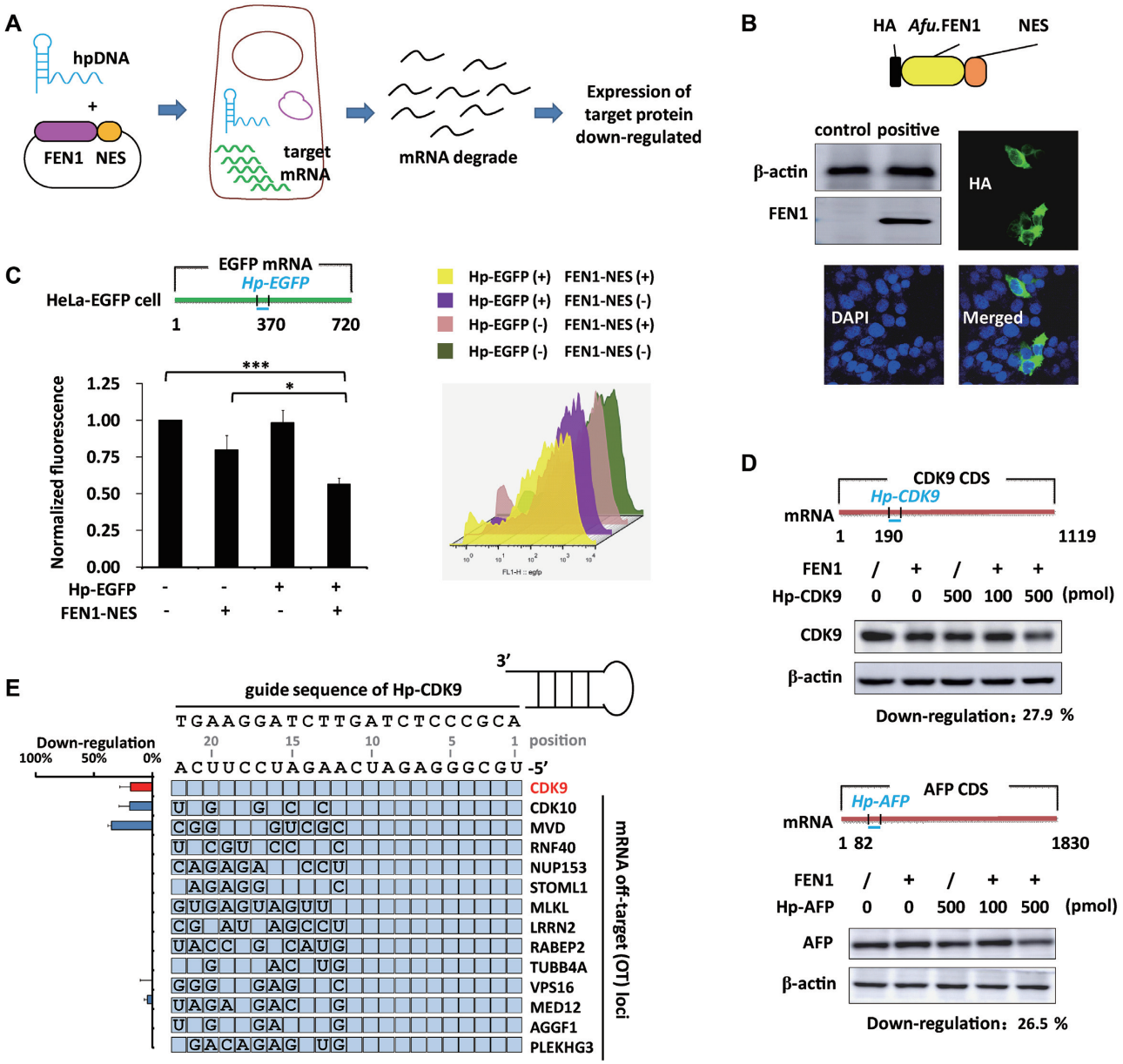

Figure 6. FEN1 mediates hpDNA-guided specific mRNA knockdown in human cells. (**A**) Schematic diagram of hpDNA-guided knockdown in cells heterologously expressing *A. fulgidus* FEN1. (**B**) The construction of *A. fulgidus* FEN1 with NES, and the WB and IF images illustrating the successful expression of the merged protein. (**C**) The normalized fluorescence of groups transfected with FEN1-NES plus hpDNA (Hp-EGFP) targeting for EGFP, $n = 3$. ***P value <0.001. *P value <0.05. (**D**) The location of hpDNAs for the gene *AFP* and *CDK9* with WB results of the control and experimental groups. $n = 3$. (**E**) HpSGN-mediated mRNA knockdown at predicated off-target loci. $n = 2$.

The loop and stem part of the hpDNAs can be fixed as designed here regardless of the nature of the targets. (iv) Misannealing of consecutive sequences close to the stem-loop structure should be avoided when choosing the locations of hpDNAs to decrease the off-target rate.

Distinct from most reported nucleic acid assembly tools, our HpSGN system functions as both a DNase and an RNase. It was also reported in the CRISPR system, with Cas9/12 (7,8) for DNA cleavage and Cas13 (10,33) for RNA cleavage. However, in the HpSGN system, it reaches the goal without no need to change proteins. The dual-functional feature of FEN1 in our HpSGN system makes it possible to develop a broad-spectrum antiviral kit to manage infections caused by either RNA or DNA viruses or both. The versatile functions of our HpSGN system may also find significant applications in regulating protein expression by cleaving coding genomic DNA and mRNA. However, modifications to both RNA and DNA simultaneously may compromise specificity. For example, the HpSGN system cannot specifically target RNA molecules without affecting the integrity of its cohabitating genomic DNA in the nucleus.

In summary, we have designed and characterized a novel and versatile system for editing both DNA and RNA

without sequence limitation. General considerations for the selection of target sites, structural requirements of the probe, evaluation of cleavage efficiency, and analysis of off-target effects have been described. However, further study is needed to improve the functional properties of the HpSGN system. For example, the efficiency of the FEN1-based system needs to be improved, and the level of off-target activity needs to be better controlled by optimizing the reaction conditions, including the concentration of FEN1 and hpDNA as well as the location of hpDNA. In addition, the rate of malignant transformation induced by high FEN1 expression should be inspected. Finally, although we have demonstrated that the HpSGN system could cleave genomic DNA in *E. coli*, its potential application in editing genomic DNA through chromosomal integration with the λ-red system remains to be tested. Compared with the CRISPR-associated system, the FEN1-based system remains to be improved. However, the results reported in this study broadened our understanding of the FEN1 plus hpDNA system and demonstrated its potential as a useful alternative tool for both DNA and RNA editing in biological engineering.

DATA AVAILABILITY

All data generated or analyzed in this study are included in this article and its supplementary information files. The plasmids used in this work can be obtained from the corresponding authors upon request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Professor Zhaoqiu Wu (China Pharmaceutical University) for critical reading and constructive discussion of the manuscript.

FUNDING

National Natural Science Foundation of China [31701161, 81873051, 81830105]; Natural Science Foundation of Jiangsu Province [BK20170729, BK20180005]; Drug Innovation Major Project [2018ZX09711001-003-007, 2017ZX09301014]; 'Double First-Class' University project [CPU 2018GF11]. Funding for open access charge: National Natural Science Foundation of China [31701161].
Conflict of interest statement. None declared.

REFERENCES

1. Miller,J.C., Holmes,M.C., Wang,J., Guschin,D.Y., Lee,Y.L., Rupniewski,I., Beausejour,C.M., Waite,A.J., Wang,N.S., Kim,K.A. *et al.* (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.*, **25**, 778–785.
2. Porteus,M.H. and Baltimore,D. (2003) Chimeric nucleases stimulate gene targeting in human cells. *Science*, **300**, 763.
3. Sander,J.D., Dahlborg,E.J., Goodwin,M.J., Cade,L., Zhang,F., Cifuentes,D., Curtin,S.J., Blackburn,J.S., Thibodeau-Beganny,S., Qi,Y. *et al.* (2011) Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat. Methods*, **8**, 67–69.
4. Wood,A.J., Lo,T.W., Zeitler,B., Pickle,C.S., Ralston,E.J., Lee,A.H., Amora,R., Miller,J.C., Leung,E., Meng,X. *et al.* (2011) Targeted genome editing across species using ZFNs and TALENs. *Science*, **333**, 307.
5. Miller,J.C., Tan,S., Qiao,G., Barlow,K.A., Wang,J., Xia,D.F., Meng,X., Paschon,D.E., Leung,E., Hinkley,S.J. *et al.* (2011) A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.*, **29**, 143–148.
6. Reyon,D., Tsai,S.Q., Khayter,C., Foden,J.A., Sander,J.D. and Joung,J.K. (2012) FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol.*, **30**, 460–465.
7. Zetsche,B., Gootenberg,J.S., Abudayyeh,O.O., Slaymaker,I.M., Makarova,K.S., Essletzbichler,P., Volz,S.E., Joung,J., van der Oost,J., Regev,A. *et al.* (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, **163**, 759–771.
8. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
9. Ran,F.A., Cong,L., Yan,W.X., Scott,D.A., Gootenberg,J.S., Kriz,A.J., Zetsche,B., Shalem,O., Wu,X., Makarova,K.S. *et al.* (2015) In vivo genome editing using Staphylococcus aureus Cas9. *Nature*, **520**, 186–191.
10. Abudayyeh,O.O., Gootenberg,J.S., Konermann,S., Joung,J., Slaymaker,I.M., Cox,D.B., Shmakov,S., Makarova,K.S., Semenova,E., Minakhin,L. *et al.* (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science*, **353**, aaf5573.
11. Hsu,P.D., Lander,E.S. and Zhang,F. (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, **157**, 1262–1278.
12. Wright,A.V., Nunez,J.K. and Doudna,J.A. (2016) Biology and applications of CRISPR Systems: harnessing nature's toolbox for genome engineering. *Cell*, **164**, 29–44.
13. Knott,G.J. and Doudna,J.A. (2018) CRISPR-Cas guides the future of genetic engineering. *Science*, **361**, 866–869.
14. Kleinstiver,B.P., Prew,M.S., Tsai,S.Q., Topkar,V.V., Nguyen,N.T., Zheng,Z., Gonzales,A.P., Li,Z., Peterson,R.T., Yeh,J.R. *et al.* (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*, **523**, 481–485.
15. Wang,J., Li,J., Zhao,H., Sheng,G., Wang,M., Yin,M. and Wang,Y. (2015) Structural and mechanistic basis of PAM-dependent spacer acquisition in CRISPR-Cas systems. *Cell*, **163**, 840–853.
16. Leenay,R.T., Maksimchuk,K.R., Slotkowski,R.A., Agrawal,R.N., Gomaa,A.A., Briner,A.E., Barrangou,R. and Beisel,C.L. (2016) Identifying and visualizing functional PAM diversity across CRISPR-Cas systems. *Mol. Cell*, **62**, 137–147.
17. Liu,J.J., Orlova,N., Oakes,B.L., Ma,E., Spinner,H.B., Baney,K.L.M., Chuck,J., Tan,D., Knott,G.J., Harrington,L.B. *et al.* (2019) CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature*, **566**, 218–223.
18. Harrington,L.B., Burstein,D., Chen,J.S., Paez-Espino,D., Ma,E., Witte,I.P., Cofsky,J.C., Kyrpides,N.C., Banfield,J.F. and Doudna,J.A. (2018) Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science*, **362**, 839–842.
19. Xu,S., Cao,S., Zou,B., Yue,Y., Gu,C., Chen,X., Wang,P., Dong,X., Xiang,Z., Li,K. *et al.* (2016) An alternative novel tool for DNA editing without target sequence limitation: the structure-guided nuclease. *Genome Biol.*, **17**, 186.
20. Hosfield,D.J., Frank,G., Weng,Y., Tainer,J.A. and Shen,B. (1998) Newly discovered archaebacterial flap endonucleases show a structure-specific mechanism for DNA substrate binding and catalysis resembling human flap endonuclease-1. *J. Biol. Chem.*, **273**, 27154–27161.
21. Li,L., Wu,L.P. and Chandrasegaran,S. (1992) Functional domains in Fok I restriction endonuclease. *PNAS*, **89**, 4275–4279.
22. Miller,J.C., Patil,D.P., Xia,D.F., Paine,C.B., Fauser,F., Richards,H.W., Shivak,D.A., Bendana,Y.R., Hinkley,S.J., Scarlott,N.A. *et al.* (2019) Enhancing gene editing specificity by attenuating DNA cleavage kinetics. *Nat. Biotechnol.*, **37**, 945–952.
23. Zou,B., Ma,Y., Wu,H. and Zhou,G. (2011) Ultrasensitive DNA detection by cascade enzymatic signal amplification based on Afu flap endonuclease coupled with nicking endonuclease. *Angew. Chem.*, **50**, 7395–7398.

24. Wang,J., Zou,B., Ma,Y., Ma,X., Sheng,N., Rui,J., Shao,Y. and Zhou,G. (2017) Closed-Tube PCR with nested serial invasion probe visualization using gold nanoparticles. *Clin. Chem.*, **63**, 852–860.
25. Zou,B., Cao,X., Wu,H., Song,Q., Wang,J., Kajiyama,T., Kambara,H. and Zhou,G. (2015) Sensitive and specific colorimetric DNA detection by invasive reaction coupled with nicking endonuclease-assisted nanoparticles amplification. *Biosens. Bioelectron.*, **66**, 50–54.
26. Zou,B., Song,Q., Wang,J., Liu,Y. and Zhou,G. (2014) Invasive reaction assisted strand-displacement signal amplification for sensitive DNA detection. *Chem. Commun.*, **50**, 13722–13724.
27. Kao,H.I., Henricksen,L.A., Liu,Y. and Bambara,R.A. (2002) Cleavage specificity of Saccharomyces cerevisiae flap endonuclease 1 suggests a double-flap structure as the cellular substrate. *J. Biol. Chem.*, **277**, 14379–14389.
28. Harrington,J.J. and Lieber,M.R. (1994) The characterization of a mammalian DNA structure-specific endonuclease. *EMBO J.*, **13**, 1235–1246.
29. Chapados,B.R., Hosfield,D.J., Han,S., Qiu,J., Yelent,B., Shen,B. and Tainer,J.A. (2004) Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell*, **116**, 39–50.
30. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
31. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R. and Genome Project Data Processing, S. (2009) The sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
32. Ran,F.A., Hsu,P.D., Wright,J., Agarwala,V., Scott,D.A. and Zhang,F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.*, **8**, 2281–2308.
33. Cox,D.B.T., Gootenberg,J.S., Abudayyeh,O.O., Franklin,B., Kellner,M.J., Joung,J. and Zhang,F. (2017) RNA editing with CRISPR-Cas13. *Science*, **358**, 1019–1027.