1 # Supplementary Information

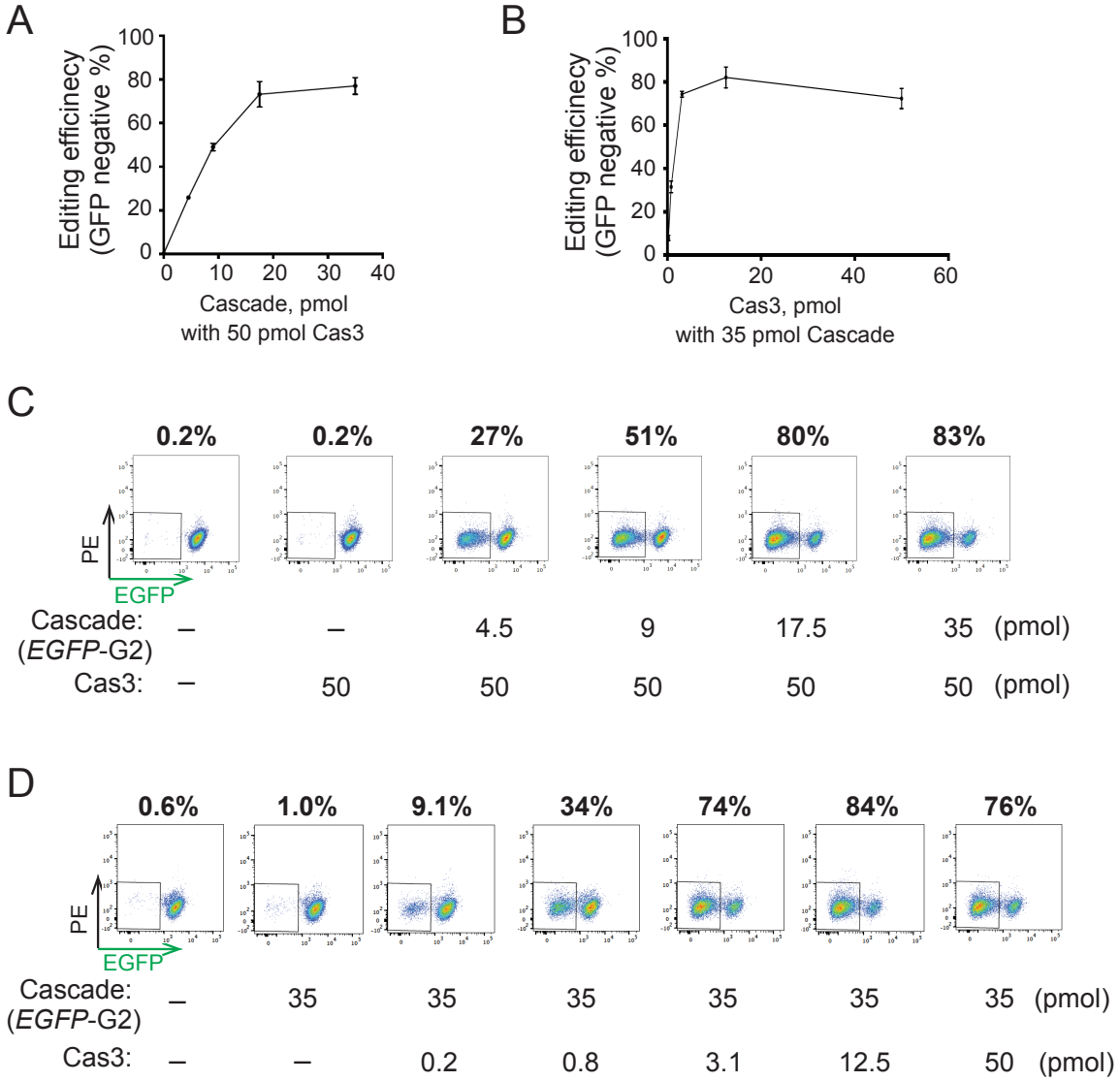
Cas11 enables genome engineering in human cells with compact CRISPRCas3 systems

2 Renke Tan*, Ryan K Krueger*, Max J Gramelspacher, Xufei Zhou, Yibei Xiao, Ailong Ke,

3 Zhonggang Hou[#] and Yan Zhang[#]

4 **Supplemental Figures**

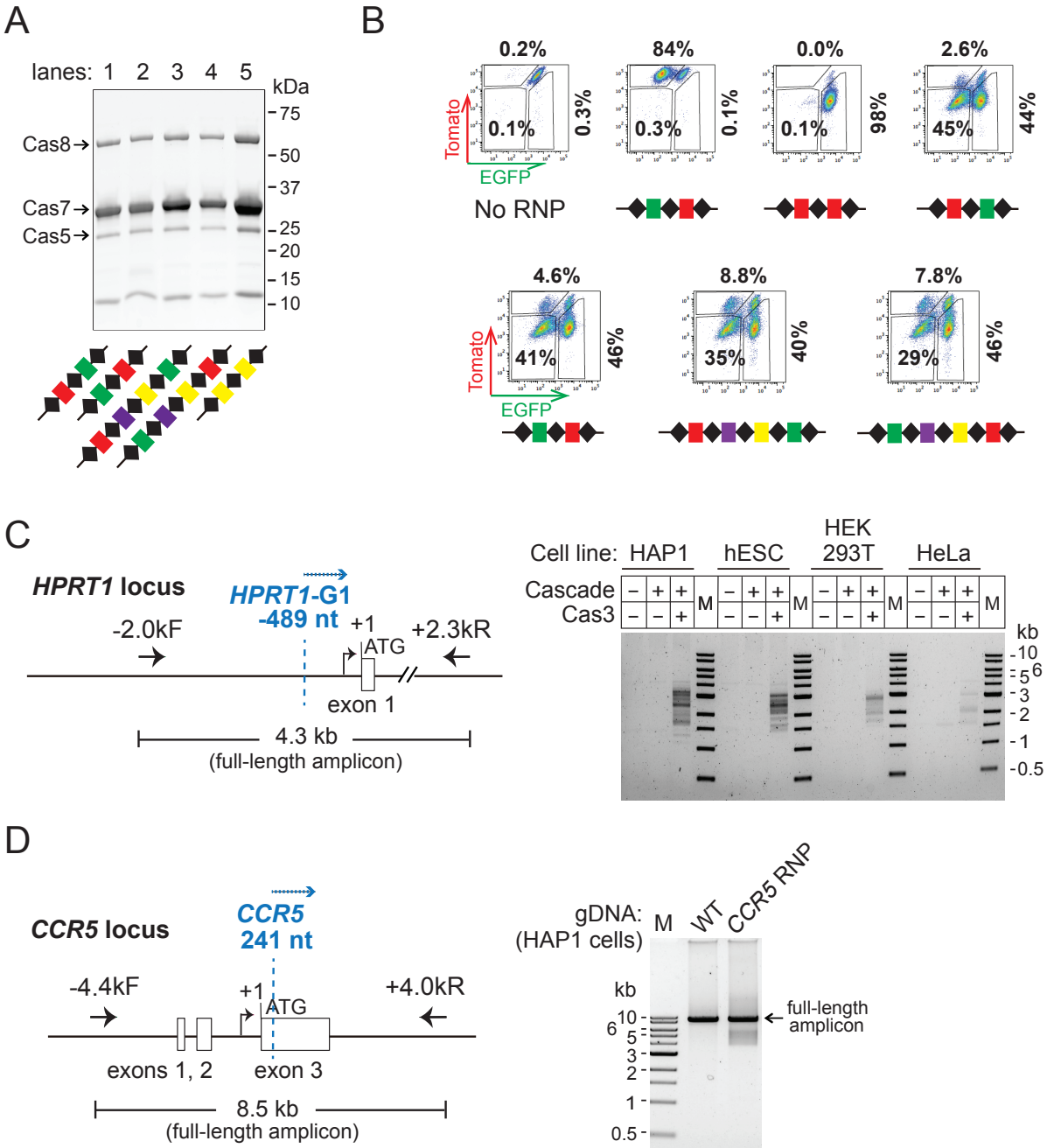
Fig. S1



5

6 **Figure S1. Cascade RNP and Cas3 protein titrations in human cell gene editing. Related**
7 **to Figure 2. (A)** RNP editing experiments in HAP1 cells with 50 pmol NlaCas3 and increasing
8 amount of *EGFP*-targeting NlaCascade. Cascade amount electroporated was titrated from 4.5
9 to 35 pmol. **(B)** RNP editing in HAP1 cells with 35 pmol *EGFP*-targeting NlaCascade and
10 increasing amount of Cas3. NlaCas3 electroporated was 0, 0.2, 0.8, 3.1, 12.5, and 50 pmol. The
11 editing efficiencies in (A) and (B) were measured and shown as in Figure 4B. (**C) (D)**
12 Representative flow cytometry plots from experiments in (A) and (B), with percentages of
13 EGFP- cells in total population shown on the top.

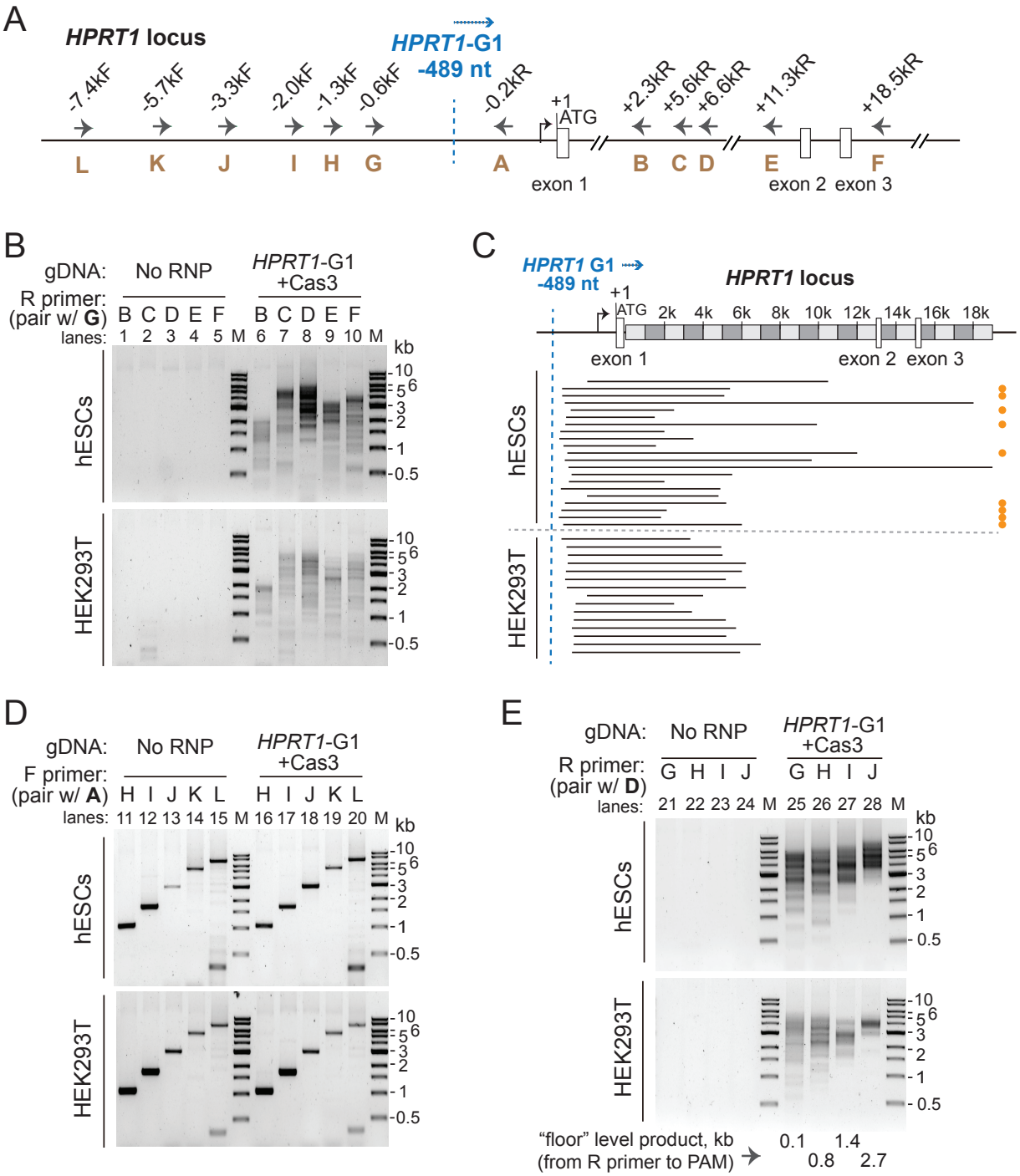
Fig. S2


A
lanes: 1 2 3 4 5
kDa
75
50
37
25
20
15
10
Cas8
Cas7
Cas5
B
0.2%
84%
0.0%
2.6%
Tomato
EGFP
0.1%
0.3%
0.3%
0.1%
0.1%
98%
45%
44%
No RNP
4.6%
8.8%
7.8%
41%
46%
35%
40%
29%
46%
C
HPRT1 locus
HPRT1-G1
-489 nt
-2.0kF
+1
ATG
+2.3kR
exon 1
4.3 kb
(full-length amplicon)
Cell line: HAP1 hESC HEK 293T HeLa
Cascade
Cas3
M
kb
10
6
5
3
2
1
0.5
D
CCR5 locus
CCR5
241 nt
-4.4kF
+1
ATG
+4.0kR
exons 1, 2
exon 3
8.5 kb
(full-length amplicon)
gDNA:
(HAP1 cells)
M
WT
CCR5 RNP
kb
full-length
amplicon

14

15 **Figure S2. NlaCascade-Cas3 RNP enables gene targeting in multiple human cell lines, at**
16 **the *HPRT1* or *CCR5* loci. Related to Figure 2. (A)** SDS-PAGE of purified NlaCascade
17 samples used for multiplexed editing in Figure 2E and for *CCR5* targeting in Figure S2A. The
18 spacer coloring scheme is as in Figure 2E. **(B)** Representative flow cytometry plots from an
19 experiment in Figure 2E, with percentages of EGFP-/tdTm+, EGPF+/ tdTm- or EGFP-/tdTm-
20 cells shown on the top, to the right or at the bottom, respectively. **(C)** Left, schematic of the
21 *HPRT1* locus. Big black arrows indicate annealing sites for two primers used in genomic PCR.
22 All positions indicated are relative to *HPRT1* translation start site (+1). Blue dashed line, the
23 recognition site (3rd nt of TTC PAM) for guide *HPRT1*-G1. The blue arrow indicates the inferred
24 direction of NlaCas3 translocation. Right, long-range PCR using genomic DNA extracted from
25 various human cell types (HAP1, hESCs, HEK293T, and Hela) edited with Cas3 and Cascade-
26 *HPRT1*. Smaller-than-full-length amplicons indicate large genomic deletions. M, size markers.
27 **(D)** Left, schematic of the *CCR5* locus. Big black arrows indicate annealing sites for two primers
28 for genomic PCR. All positions indicated are relative to *CCR5* translation start site (+1). Blue
29 dashed line, recognition site (3rd nt of TTC PAM) for guide *CCR5*-G2. The blue arrow indicates
30 the inferred direction of NlaCas3 translocation. Right, long-range PCR as described in (A), using
31 genomic DNA extracted from HAP1 cells edited with Cas3 and Cascade-*CCR*5. Smaller-than-
32 full-length amplicons indicate large genomic deletions resulted from *CCR5* targeting.

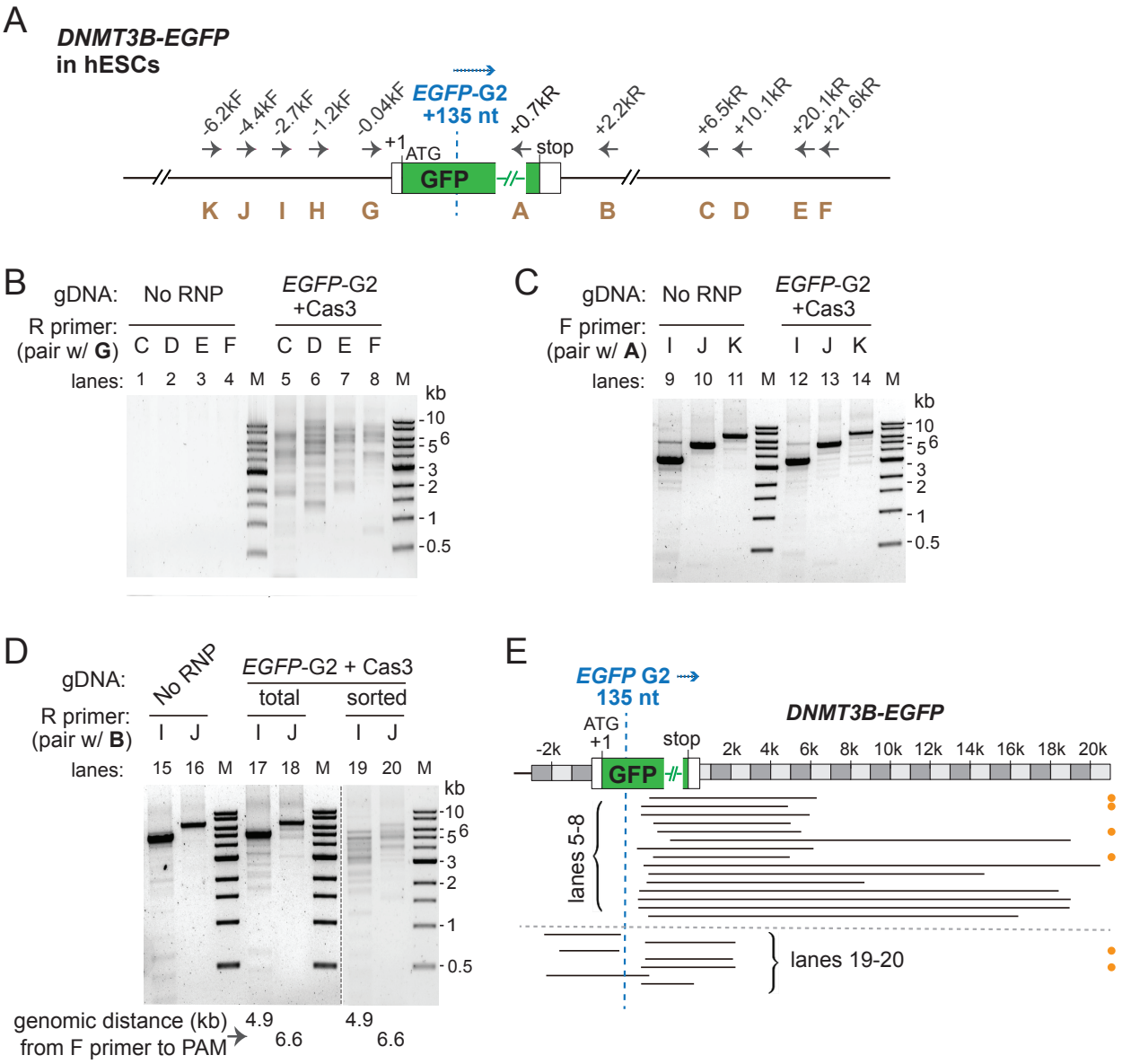
Fig. S3


A
HPRT1 locus
HPRT1-G1
-489 nt
-7.4kF
-5.7kF
-3.3kF
-2.0kF
-1.3kF
-0.6kF
-0.2kR
+1
ATG
+2.3kR
+5.6kR
+6.6kR
+11.3kR
+18.5kR
L K J I H G A B C D E F
exon 1
exon 2 exon 3
B
gDNA: No RNP HPRT1-G1 +Cas3
R primer: (pair w/ G) B C D E F B C D E F
lanes: 1 2 3 4 5 M 6 7 8 9 10 M
kb
10 5 6 3 2 1 0.5
hESCs
HEK293T
C
HPRT1 G1
-489 nt
HPRT1 locus
+1
ATG 2k 4k 6k 8k 10k 12k 14k 16k 18k
exon 1
exon 2 exon 3
hESCs
HEK293T
D
gDNA: No RNP HPRT1-G1 +Cas3
F primer: (pair w/ A) H I J K L H I J K L
lanes: 11 12 13 14 15 M 16 17 18 19 20 M
kb
10 5 6 3 2 1 0.5
hESCs
HEK293T
E
gDNA: No RNP HPRT1-G1 +Cas3
R primer: (pair w/ D) G H I J G H I J
lanes: 21 22 23 24 M 25 26 27 28 M
kb
10 5 6 3 2 1 0.5
hESCs
HEK293T
"floor" level product, kb (from R primer to PAM)
0.1 1.4
0.8 2.7

33

34 **Figure S3. NlaCRISPR-Cas3 creates targeted, large uni-directional genomic deletions in**
35 **hESC and HEK293T cells. Related to Figure 3. (A)** Schematic of *HPRT1* locus and annealing
36 sites of PCR primers used in (B), (D) and (E). All positions are relative to *HPRT1* translation
37 start site (+1). Blue dashed line, recognition site (3rd nt of TTC PAM) for guide *HPRT1*-G1. The
38 blue arrow indicates the inferred direction of NlaCas3 translocation. **(B)(D)(E)** Genomic lesion
39 analysis via long-range PCR, using primers amplifying regions downstream (B) and upstream
40 (D) of the CRISPR-targeted *HPRT1* site, or regions spanning both directions (E). Genomic DNA
41 samples used as PCR template were extracted from hESCs and HEK293T cells. Large, uni-
42 directional deletions were detected in PAM-proximal genomic region, from cells treated with
43 Cas3 and Cascade *HPRT1*-G1, but not the untreated control. Smaller-than-full-length amplicons
44 indicate large deletions. The lack of full-length product from the un-edited control is likely due to
45 a GC-rich region in exon 1 (~ 400 bp downstream of target site) that prevents amplification.
46 Smaller-than-full-length amplicons indicate large deletions. M, size markers. **(C)** *HPRT1* deletion
47 locations, revealed by TOPO cloning of pooled tiling PCRs from lanes 6-10 in (B) and Sanger
48 sequencing of randomly selected individual clones. The black lines denote deleted genomic
49 regions. Orange, green and the lack of dots on the right indicate Groups II, IV and I deletion
50 junctions as in Dolan *et al.*, 2019.

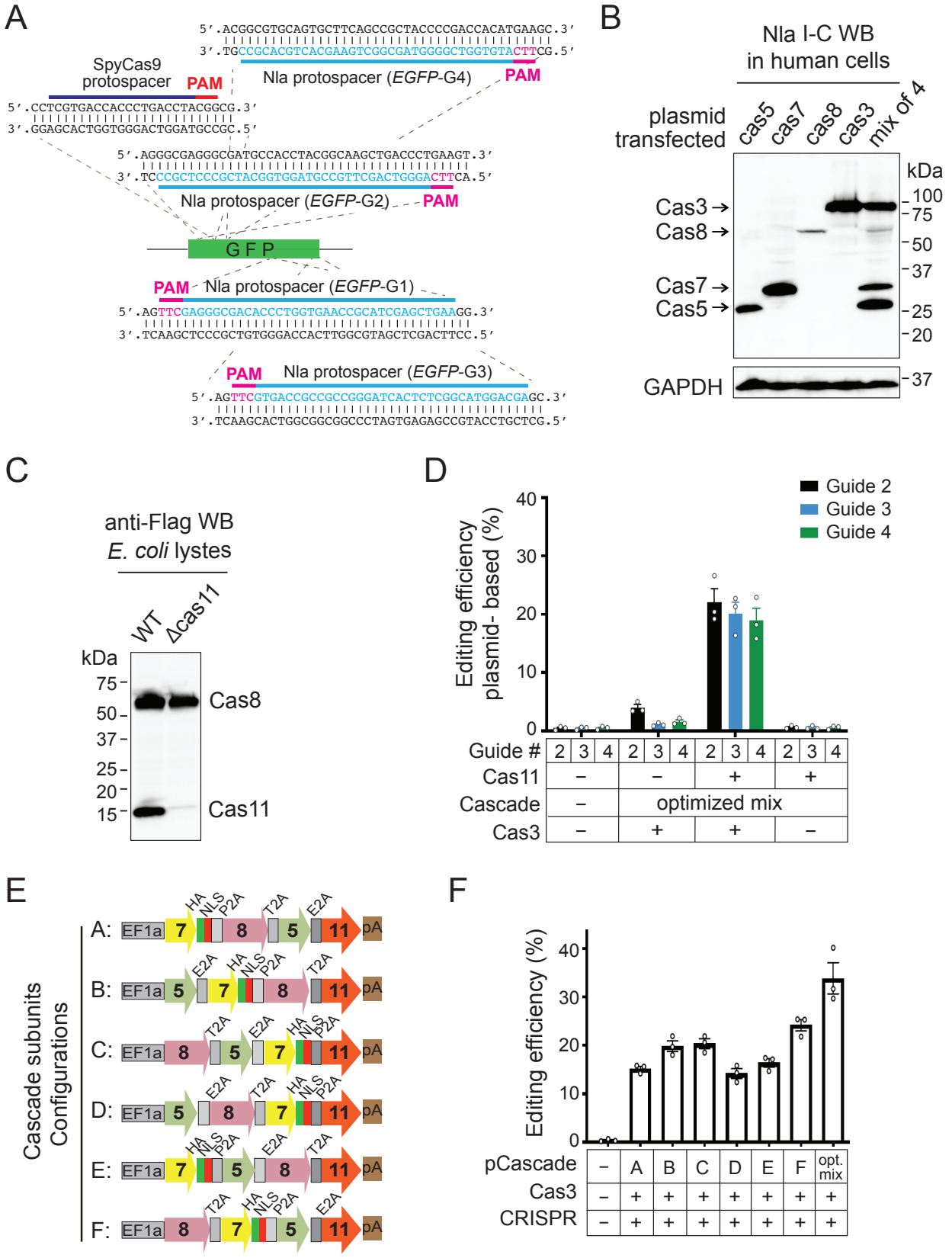
Fig. S4


A
DNMT3B-EGFP
in hESCs
-6.2kF
-4.4kF
-2.7kF
-1.2kF
-0.04kF
EGFP-G2
+135 nt
+0.7kR
+2.2kR
+6.5kR
+10.1kR
+20.1kR
+21.6kR
+1
ATG
stop
GFP
K J I H G A B C D E F
B
gDNA: No RNP EGFP-G2 +Cas3
R primer: (pair w/ G) C D E F C D E F
lanes: 1 2 3 4 M 5 6 7 8 M
kb
10
5 6
3
2
1
0.5
C
gDNA: No RNP EGFP-G2 +Cas3
F primer: (pair w/ A) I J K I J K
lanes: 9 10 11 M 12 13 14 M
kb
10
5 6
3
2
1
0.5
D
gDNA: No RNP EGFP-G2 + Cas3
total sorted
R primer: (pair w/ B) I J I J I J
lanes: 15 16 M 17 18 M 19 20 M
kb
10
5 6
3
2
1
0.5
genomic distance (kb) from F primer to PAM
4.9 6.6 4.9 6.6
E
EGFP G2
135 nt
DNMT3B-EGFP
ATG
+1
stop
-2k 2k 4k 6k 8k 10k 12k 14k 16k 18k 20k
GFP
lanes 5-8
lanes 19-20

51

52 **Figure S4. NlaCRISPR-Cas3 induces large deletions at the *DNMT3b* locus in hESCs.**
53 **Related to Figure 3. (A)** Schematic of the *DNMT3b-EGFP* locus in hESC reporter cells, with
54 annealing sites of PCR primers used in (B)-(E) indicated. All positions are relative to *EGFP*
55 translation start site (+1). Blue dashed line, recognition site (3rd nt of TTC PAM) for guide *EGFP-*
56 G2. The blue arrow indicates the inferred direction of NlaCas3 translocation. **(B)** Genomic lesion
57 analysis via long-range PCR, using primers amplifying regions downstream (B) or upstream (C)
58 of CRISPR-targeted *EGFP* site, or regions spanning both directions (D). Genomic DNA used as
59 PCR template was extracted from hESC reporter cells bearing *EGFP* and *tdTm* at the
60 endogenous *DNMT3b* locus. Large, uni-directional deletions were detected in the PAM-proximal
61 region, from cells edited with Cas3 and Cascade *EGFP-G2*, but not the "no RNP" control cells.
62 Smaller-than-full-length products indicate large deletions. M, size markers. Discontinuous lanes
63 from the same gel are separated by dashed grey line. **(E)** Deletion locations revealed by TOPO
64 cloning of pooled tiling PCRs from lanes 5-8 from (B) and 19-20 from (D). Randomly selected
65 clones are Sanger sequenced. The black lines indicate deleted genomic regions. Orange dots
66 and lack thereof on the right indicate Groups II and I deletion junctions as in Dolan *et al.*, 2019.
67 Note the existence of three bi-directional deletion events from PCR of lanes 19-20.

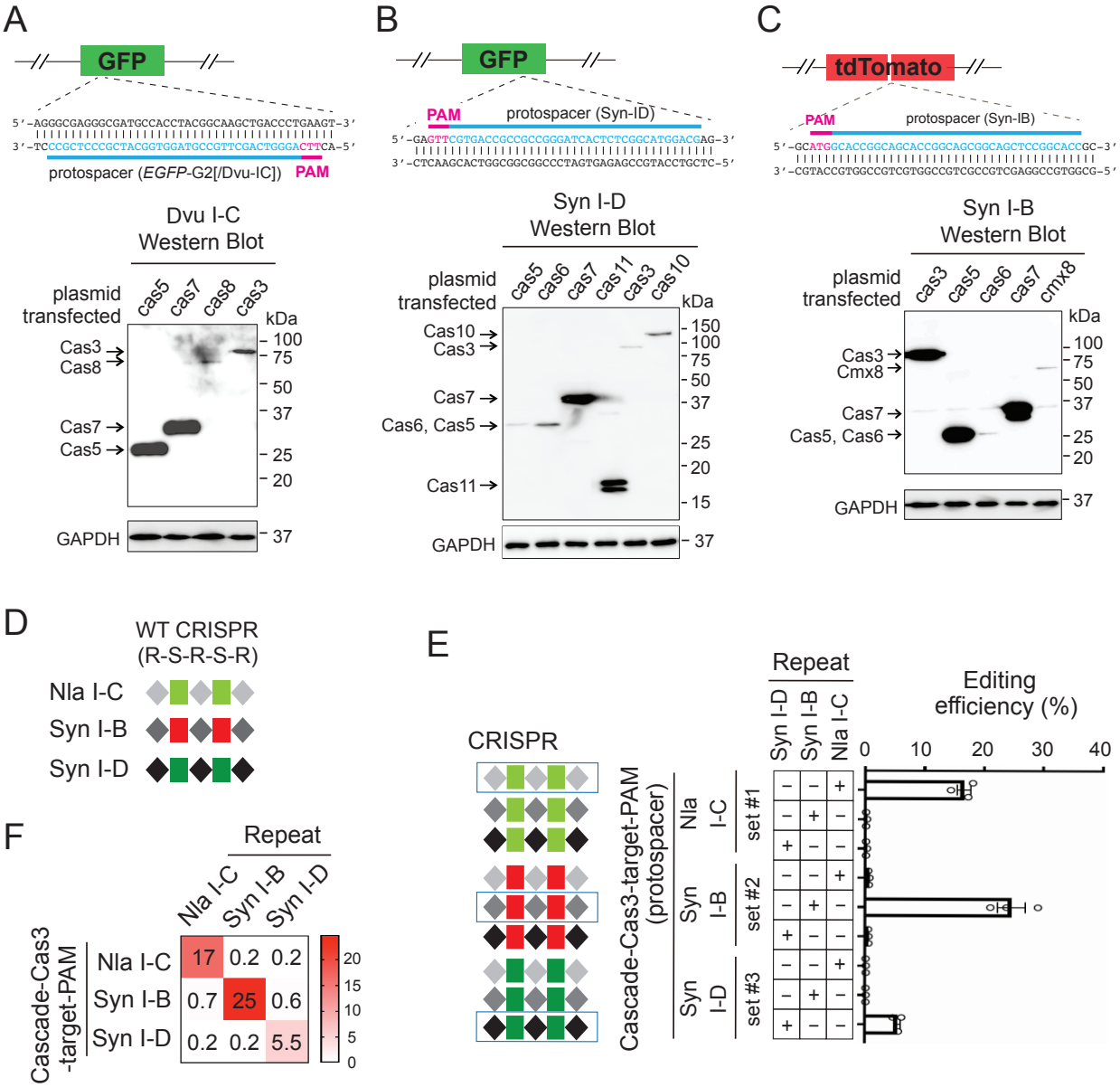
Fig. S5


A
SpyCas9 protospacer
PAM
Nla protospacer (EGFP-G4)
Nla protospacer (EGFP-G2)
GFP
Nla protospacer (EGFP-G1)
Nla protospacer (EGFP-G3)
B
Nla I-C WB in human cells
plasmid transfected
cas5
cas7
cas8
cas3
mix of 4
kDa
Cas3
Cas8
Cas7
Cas5
GAPDH
C
anti-Flag WB
E. coli lystes
WT
Δcas11
Cas8
Cas11
D
Guide 2
Guide 3
Guide 4
Editing efficiency plasmid- based (%)
Guide #
Cas11
Cascade
optimized mix
Cas3
E
Cascade subunits Configurations
EF1a
HA
NLS
P2A
T2A
E2A
bA
F
Editing efficiency (%)
pCascade
opt. mix
Cas3
CRISPR

68

69 **Figure S5. NlaCas11 is the missing component for efficient plasmid-based editing in**
70 **human cells. Related to Figures 4-5. (A**) Schematics of EGFP reporter and target sites for all
71 NlaCascades and SpyCas9. Protospacers are indicated in blue and corresponding PAMs in
72 magenta. **(B)** Anti-HA western blot detecting expression of all canonical *cas* genes of Nla I-C
73 system (*cas5*, *cas7*, *cas8* and *cas3*) after plasmid transfection into HAP1 cells. Bottom, GAPDH
74 is probed as loading control. Molecular weight markers (kDa) are indicated. **(C)** Nla I-C CRISPR
75 system indeed expresses a previously overlooked Cas11 from within *cas8*. Plasmids expressing
76 CRISPR and the *cascade* operon were co-transformed into *E. coli* BL21(DE3), and the resulting
77 strains subject to Western Blot. pCascade plasmids have a Flag-tag at the C-terminus of *cas8*.
78 Both Cas8 and Cas11 proteins were detectable by anti-Flag western from wt strain; whereas
79 Cas11 production was abolished by mutations disrupting the RBS and internal translation start
80 site. Molecular weight markers (kDa) are indicated. **(D)** The Nla *crispr-cas* plasmids depicted in
81 Figure 5A were transfected into HAP1 cells to evaluate editing efficiencies for *EGFP*-targeting
82 guides 2, 3, and 4. The results were plotted as the percentage of EGFP- cells in the total
83 population. Data are shown as mean ± SEM, n=3. **(E)** Schematics of the polycistronic constructs
84 tested in (F). The NLS, HA tag and regulatory elements are as described in Figure 4A. **(F)** Each
85 plasmid depicted in (E) was transfected into HAP1 cells along with the Cas3- and CRISPR-
86 encoding plasmids, and the gene editing efficiencies were assessed and shown as in Figure 4B.
87 Data are shown as mean ± SEM, n=3.

Fig. S6


A
GFP
5′-AGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGT-3′
3′-TCCCGCTCCCGCTACGGTGGATGCCGTTCGACTGGGACTTCA-5′
protospacer (EGFP-G2[/Dvu-IC])
PAM
Dvu I-C
Western Blot
plasmid transfected
cas5
cas7
cas8
cas3
kDa
Cas3
Cas8
Cas7
Cas5
100
75
50
37
25
20
GAPDH
37
B
GFP
PAM
protospacer (Syn-ID)
5′-GAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAG-3′
3′-CTCAAGCACTGGCGGCGGCCCTAGTGAGAGCCGTACCTGCTC-5′
Syn I-D
Western Blot
plasmid transfected
cas5
cas6
cas7
cas11
cas3
cas10
kDa
Cas10
Cas3
Cas7
Cas6, Cas5
Cas11
150
100
75
50
37
25
20
15
GAPDH
37
C
tdTomato
PAM
protospacer (Syn-IB)
5′-GCATGGCACCGGCAGCACCGGCAGCGGCAGCTCCGGCACCGC-3′
3′-CGTACCGTGGCCGTCGTGGCCGTCGCCGTCGAGGCCGTGGCG-5′
Syn I-B
Western Blot
plasmid transfected
cas3
cas5
cas6
cas7
cmx8
kDa
Cas3
Cmx8
Cas7
Cas5, Cas6
100
75
50
37
25
20
GAPDH
37
D
WT CRISPR
(R-S-R-S-R)
Nla I-C
Syn I-B
Syn I-D
E
CRISPR
Cascade-Cas3-target-PAM
(protospacer)
Nla I-C
Syn I-B
Syn I-D
set #1
set #2
set #3
Repeat
Syn I-D
Syn I-B
Nla I-C
Editing efficiency (%)
0
10
20
30
40
F
Repeat
Nla I-C
Syn I-B
Syn I-D
Cascade-Cas3-target-PAM
Nla I-C
Syn I-B
Syn I-D
17
0.2
0.2
0.7
25
0.6
0.2
0.2
5.5
20
15
10
5
0

88

Figure S6. Target sequence, protein expression and repeat specificity for orthogonal Dvu I-C, Syn I-D, and Syn I-B CRISPR systems. Related to Figures 6. (A-C) Top: Schematics of target sites used, with protospacers for Cascade RNPs in blue and corresponding PAMs in magenta. Bottom: anti-HA western blot detecting expression of all *cas* genes of the Dvu I-C (A), Syn I-D (B), and Syn I-B (C) systems, after transfection into HAP1 cells. GAPDH was probed as loading control. Molecular weight markers (kDa) are indicated. Bha I-C is not included because there are no epitope tags on its *cas* plasmids. **(D)** Schematics of wt CRISPR constructs for Nla I-C, Syn I-D, and Syn I-B editors. Light grey, dark grey, and black diamonds indicate CRISPR repeats of the I-C, I-B and I-D systems, respectively. Light green (*EGFP*-targeting), red (*tdTm*-targeting), and dark green (*EGFP*-targeting) rectangles indicate CRISPR spacers for the I-C, I-B, and I-D editors. **(E)** Repeat specificity suffices for CRISPR-Cas3 orthogonality in human cells. Mix-and-match experiment assaying Cas plasmids from three distinct type I systems paired with wt or chimeric CRISPR constructs. The actual repeat and spacer analyzed in each test were indicated, with schematics of the entire CRISPR array included on the left; the three wt CRISPRs without repeat swap were boxed. Genome editing was evaluated in HAP1 cells following plasmid transfection, and the efficiencies were plotted as in Figure 4B. Data are shown as mean ± SEM, n=3. **(F)** Heatmaps of gene editing efficiencies reported in (E).