Molecular Therapy
Original Article

CRISPR/Cas9 Editing of the Mutant Huntingtin Allele In Vitro and In Vivo

Alex Mas Monteys,[1] Shauna A. Ebanks,[1] Megan S. Keiser,[1] and Beverly L. Davidson[1,2]

[1]Raymond G. Perelman Center for Cellular and Molecular Therapeutics, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; [2]Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Huntington disease (HD) is a fatal dominantly inherited neurodegenerative disorder caused by CAG repeat expansion (>36 repeats) within the first exon of the *huntingtin* gene. Although mutant huntingtin (mHTT) is ubiquitously expressed, the brain shows robust and early degeneration. Current RNA interference-based approaches for lowering mHTT expression have been efficacious in mouse models, but basal mutant protein levels are still detected. To fully mitigate expression from the mutant allele, we hypothesize that allele-specific genome editing can occur via prevalent promoter-resident SNPs in heterozygosity with the mutant allele. Here, we identified SNPs that either cause or destroy PAM motifs critical for CRISPR-selective editing of one allele versus the other in cells from HD patients and in a transgenic HD model harboring the human allele.

INTRODUCTION

Huntington's disease (HD) is a fatal neurodegenerative disorder caused by CAG repeat expansion in the *huntingtin* (*HTT*) gene. Although huntingtin is ubiquitously expressed, the neuropathology of HD is characterized by early striatal atrophy followed by volume loss in other brain areas.[1,2] There is no cure for HD and treatments are focused on symptom management.[3] Earlier studies using genetically modified mouse models showed that HD-like phenotypes can be resolved if mutant huntingtin expression is eliminated, even at advanced disease stages,[4,5] suggesting that therapeutic strategies focused on eliminating mutant huntingtin expression will be highly beneficial. As examples, knockdown strategies using RNAi or antisense oligonucleotides, which reduce mutant huntingtin expression either alone or together with the normal huntingtin, are beneficial in various mouse models.[6–9] Other strategies, such as genome editing with zinc finger nucleases targeted to the CAG-repeat expansion region, have also shown promise.[10]

Genome editing with the recently discovered CRISPR/Cas9 system represents an exciting alternative for tackling dominantly inherited genetic disorders such as HD.[11–13] The most recent system advancements involve expressing Cas9 along with a single guide RNA molecule (sgRNA). When co-expressed, sgRNAs bind and recruit Cas9 to a specific genomic target sequence where it mediates a double-strand DNA (dsDNA) break, activating the dsDNA break repair machinery. Targeted gene deletions by non-homologous end joining (NHEJ) can be made when a pair of sgRNA/Cas9 complexes bind in proximity and produce dsDNA breaks.[13–15]

Given the potency and sequence specificity of the CRISPR/Cas9 targeting, and the fact that huntingtin is an important protein for several cellular functions,[16] the use of CRISPR/Cas9 to direct allele-specific genome editing is an attractive alternative to the partial reduction approach using ASOs or RNAi methods. Targeting specificity of the CRISPR/Cas9 complex is regulated by two different elements, first, the binding complementarity between the targeted genomic DNA sequence (genDNA) and the 20 nt-guiding sequence of the sgRNA, and, second, the presence of a protospacer-adjacent motif (PAM) juxtaposed to the genDNA/sgRNA complementary region.[11,13,17] While previous studies have shown that nucleotide mismatches at positions 1–10 on the sgRNA-target site interface are not well tolerated for cleavage, sequence context at this region is crucial to determine which nucleotide positions are more effective to influence cleavage.[11,14,17–19] However, the preservation of an intact PAM motif appears to be critical and genome wide studies searching for Cas9 off-target cleavage events demonstrate that mutations on the PAM motif result in an important reduction of cleavage efficacy.[20–24] Therefore, allele-specific gene editing could be achieved by taking advantage of prevalent SNPs that either eliminate or create a PAM sequence. In HD, polyglutamine repeat expansion occurs within exon-1 of *HTT*.[1] Because the main regulatory elements for *HTT* expression reside within the first two kilobase 5′ of the transcription start site,[25] SNP-dependent PAMs in heterozygosity with the mutation are natural CRISPR/Cas9 targets for allele-specific editing. We therefore screened genomic regions adjacent to *HTT* exon-1 to identify SNPs that were prevalent, and were within the critical position for CRISPR/Cas9- or CRISPR/Cpf1-directed editing, and tested their utility for allele-specific editing in HD patient cell lines and a mouse model expressing full length mutant human HTT.

Received 16 October 2016; accepted 11 November 2016;
http://dx.doi.org/10.1016/j.ymthe.2016.11.010.

Correspondence: Alex Mas Monteys, The Children's Hospital of Philadelphia, 5060 CTRB, 3501 Civic Center Boulevard, Philadelphia, PA 19104, USA.
E-mail: monteysam@email.chop.edu
Correspondence: Beverly L. Davidson, The Children's Hospital of Philadelphia, 5060 CTRB, 3501 Civic Center Boulevard, Philadelphia, PA 19104, USA.
E-mail: davidsonbl@email.chop.edu

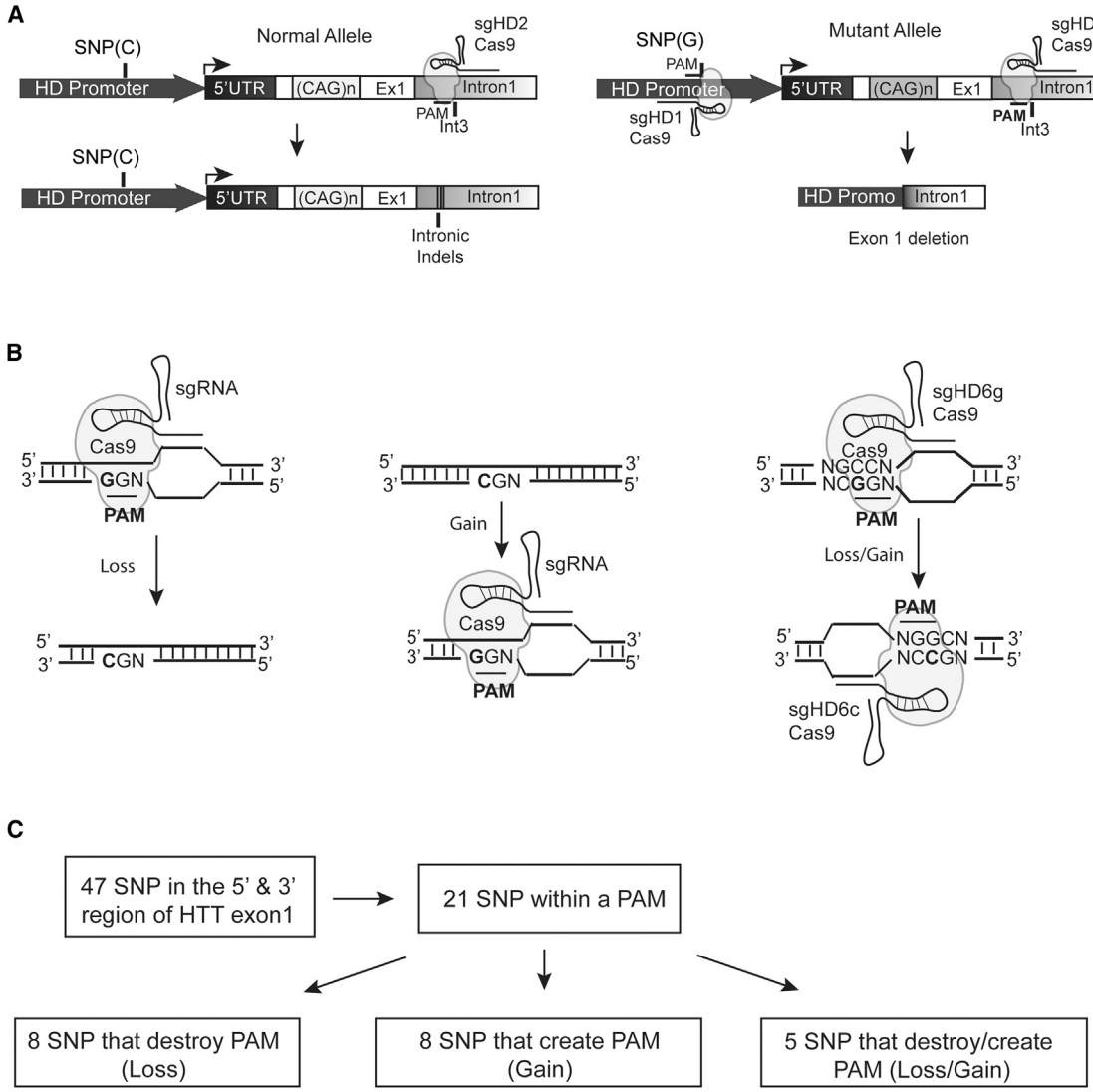


Figure 1. SNP-Dependent Editing for Huntington Disease Therapy

(A) Cartoon depicting the allele-specific editing strategy to abrogate mutant HTT expression. SNPs within PAM sequences upstream of *HTT* exon-1 permit specific targeted deletions of the mutant allele when present in heterozygosity. After DNA repair, mutant *HTT* exon-1 is deleted by a pair of sgRNA/Cas9 complexes binding upstream and downstream of exon-1 (right), whereas intronic indels could be generated by a single dsDNA break in the normal allele (left). (B) The nucleotide variation of a SNP within a PAM alters Cas9 recognition resulting in the loss (left), the gain (middle), or the simultaneous loss of a PAM in one DNA strand and the gain of a PAM on the opposite strand (right). (C) There are 21 out of 47 prevalent SNPs flanking *HTT* exon-1 that are located within predicted critical positions of a PAM sequence for the CRISPR/SpCas9 system analyzed. The minor frequency allele either mediates the loss (eight SNPs), gain (eight SNPs), or a loss/gain (five SNPs) of a PAM motif.

RESULTS

Screening SNP-Derived PAM Motifs in the *HTT* Locus

Our goal was to delete the mutant *HTT* allele using SNP-dependent PAMs flanking *HTT* exon-1 that, when present in heterozygosity, would tether the Cas9 protein to the mutant, but not the normal allele (Figure 1A). The CRISPR/SpCas9 system from *Streptococcus pyogenes* is the most widely used, and its PAM sequence (NRG, where N represents any nucleotide, R a purine, and G the conserved guanine) has been fully characterized.[11,13] SNPs present at the third PAM nucleotide position could generate, remove, or simultaneously do both in a strand specific way (Figure 1B). Using the NCBI website and the 1000 Genomes database, we identified 47 SNPs with a prevalence of more than 5%, located upstream (within ~5 kilobase, Promoter/5′ UTR) and downstream (6.5 kilobase, Intron1) of *HTT* exon-1. Of these, 21 were present at the conserved third nucleotide of the NRG PAM sequence of SpCas9 (Figures 1C and S1; Table S1). NAG PAMs were included in our screen, although SpCas9 recognition for NAG PAM is less efficient than NGG PAM.[26,27] Overall, the nucleotide variation caused the loss (eight SNPs), gain (eight SNPs), or simultaneously the loss in one DNA strand

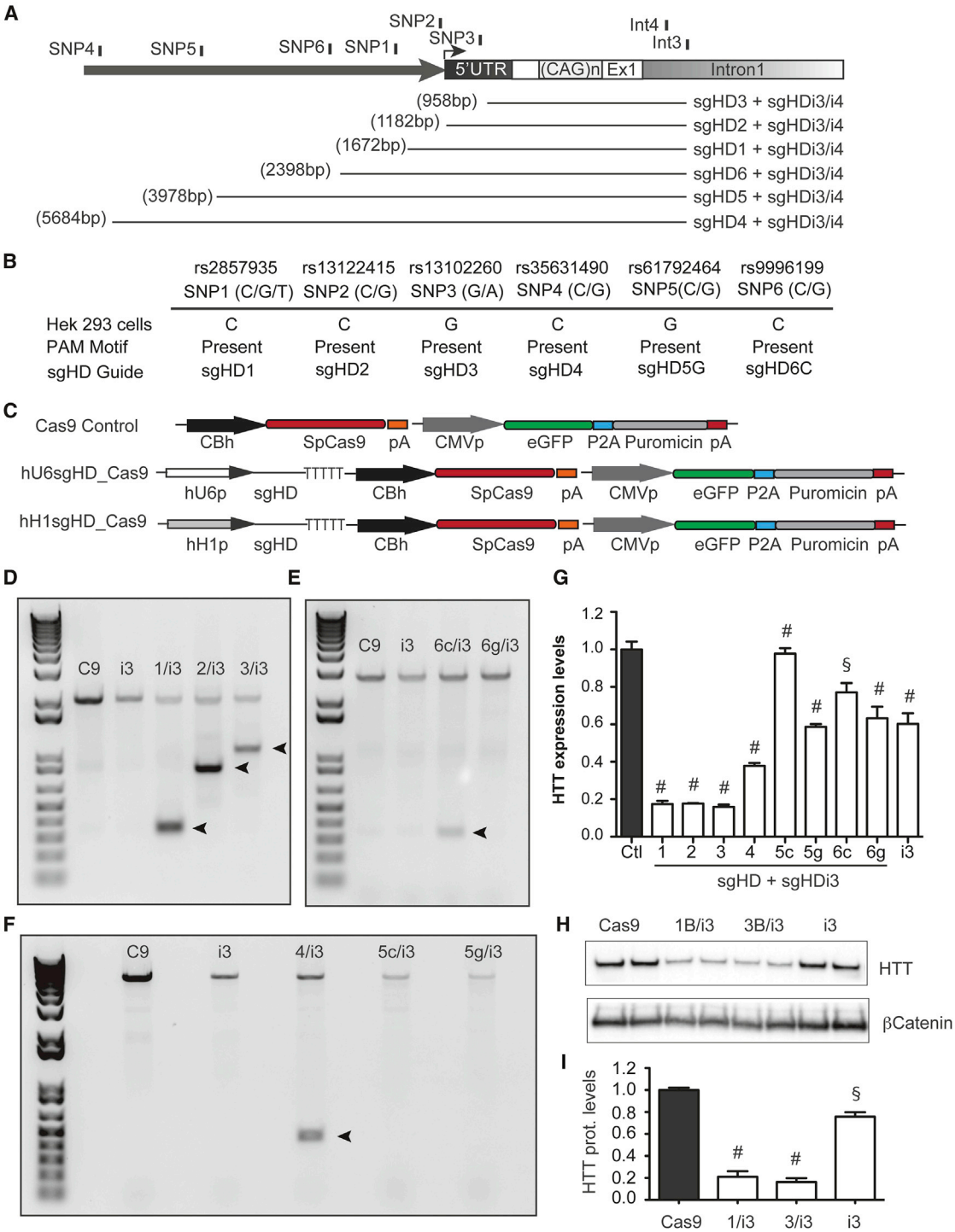


Figure 2. Cleavage of SNP-Dependent sgHD/SpCas9 Complexes in HEK293 Cells

(A) Cartoon depicting the relative position of the six prevalent SNP-dependent PAMs upstream of *HTT* exon-1 and two common PAMs within *HTT* intron-1. The estimated size of the targeted deleted sequence is indicated. (B) The genotype of the prevalent SNPs within the *HTT* promoter in HEK293 cells is shown. All SNPs were homozygous for the nucleotide variation and the PAM motif was present for the sgRNA indicated. (C) A diagram of the CRISPR expression systems transfected into HEK293 cells is shown. (D–F) A genomic PCR showing *HTT* exon-1-targeted deletion by sgRNA/SpCas9 pair complexes binding upstream and downstream of the target sequence is shown in the images. (G) RT-qPCR analysis of HTT mRNA levels in HEK293 cells transfected with sgHD/SpCas9 expression cassettes targeting upstream promoter SNPs and the

(legend continued on next page)

and a gain on the opposite strand (five SNPs, Loss/Gain) (Figure 1C; Table S1).

Experimental Validation of *HTT* Promoter SNP-Dependent PAM Motifs

We next developed single-guide RNAs (sgRNAs, all ≤20 nt) to six of the identified SNP-dependent PAMs upstream of *HTT* exon-1 to test as candidates for CRISPR/Cas9 cleavage in HEK293 cells. There were five SNPs that were located within the ~5 kilobase of the HTT promoter region (SNPs 1, 2, 4, 5, and 6) and one at the 5′ UTR near the HTT transcription start site (SNP3) (Figures 2A, 2B, and S2). These SNPs have a minor allele frequency of >10% in the general population, and the nucleotide variations cause the Loss or a Loss/Gain of the PAM motif (Table S2). Common sgRNAs were also designed to target sequences within *HTT* intron-1 (sgHDi3 and sgHDi4; Figures 2A and S2). The sgRNAs were cloned downstream of the hU6 or hH1 promoter, along with other elements as depicted (Figure 2C). HEK293 cells, which are homozygous for the targeting SNPs (Figure 2B), were transfected with SpCas9 and sgRNA expression plasmids and genomic deletion assessed. DNA products of the anticipated size were amplified in most of the sgRNA/SpCas9 pair complexes tested (Figures 2D–2F). Sanger sequencing of the small-amplified PCR products confirmed *HTT* exon-1 deletion and dsDNA repair (Figure S4). As expected, *HTT* remained intact in cells expressing SpCas9 or a single sgRNA sequence (sgHDi3; Figure 2, sgHD1, sgHD2, and sgHD3; Figure S5A) or co-expressing sgHDi3 with a sgRNA sequence for which a PAM sequence is absent in the *HTT* promoter (sgHD5c/i3 and sgHD6g/i3; Figure 2). We did not detect *HTT* exon-1 cleavage in cells transfected with sgHD5g/i3, in spite of the presence of the PAM. Both sgHD1 and sgHD5g have a 17 nt complementary sequence, yet sgHD1/i3 eliminated *HTT* exon-1, while sgHD5/i3 did not. Interestingly, sgHD1 has eight guanines, six cytosines, and one adenosine, whereas sgHD5g has four guanines, three cytosines, and three adenosines. This is consistent with earlier work showing a direct correlation between the sequence composition of the sgRNA complementary region to sgRNA activity, with the most active sequences enriched for guanine and cytosine and depleted of adenosine.[28]

HTT mRNA and protein levels were reduced in cells following editing, as determined by qPCR and western blot, respectively (Figures 2G–2I and S3). Reduction of HTT mRNA levels was greater in cells expressing sgRNA/SpCas9 complex pairs that generated small targeted deletions, suggesting that *HTT* exon-1 removal efficacy may be influenced by the distance between the two dsDNA breaks (compare sgHD1, 2, and 3 versus sgHD4 and 6) (Figure 2G). Also, our results corroborate previous studies showing preference of SpCas9 for NGG over NAG PAM sequences (compare sgHD1, 2, 3, and 4 [NGG] versus sgHD6c [NAG]) (Figure 2G).[26,27] Interestingly, cells expressing a single sgRNA sequence alone, or sgHDi3 in combination with sgHD6g or sgHD5g, also showed reduced HTT mRNA and protein levels, albeit not to as great an extent as those where *HTT* exon-1 was removed (Figures 2G–2I, S5B, and S5C). Because these sgRNA/Cas9 complexes did not remove HTT exon-1, it suggests that elements within the first intron (sgHDi3) and the promoter region (sgHD1, sgHD2, and sgHD3) might be disrupted as result of indels generated after DNA repair (Figures 2G–2I and S5D).

The generation of short N-terminal fragments as a result of mutant HTT protein cleavage is one of the pathogenic hallmarks of HD. Whereas toxicity of N-terminal fragments has been widely demonstrated, several studies suggest that truncated C-terminal fragments resulting from mutant HTT proteolysis may also contribute to HD pathogenesis.[29,30] Importantly, our data indicate that truncated C-terminal fragments are also eliminated in HEK293 cells edited with our most effective sgRNA sequences, as determined by qRT-PCR or western blot (Figures S6B–S6D).

Assessment of Editing Specificity in HD Human Fibroblasts

Next, we aimed to determine whether allele-specific editing could be achieved for a single allele using the SNP-dependent PAMs in the HTT promoter region. There were 23 lines of fibroblast cell lines from HD patients that were screened for SNP heterozygosity using direct Sanger sequencing of PCR amplified products. There were 11 lines that were heterozygous for SNP1; one line was heterozygous for SNP2, SNP4, and SNP6; and two lines were heterozygous for SNP3 and SNP5 (Table S3).

We focused on the sgHD1/i3 Cas9 complex pair, since it was one of the most active sgRNA/Cas9 pairs, generated a larger HTT promoter deletion than sgHD2/i3 and sgHD3/i3, and the SNP within the PAM was the most prevalent among the HD fibroblast lines tested and is present in heterozygosity for more than 20% in the population. Expression vectors for sgHD1/i3 and SpCas9, or SpCas9 only, were generated (Figure 3A) for testing in HD fibroblast cell lines. Two lines, ND31551 and ND33392, which are heterozygous for the SNP1 on opposite alleles, were chosen for specificity testing (Figure 3B; Table S4). PCR of genomic DNA showed target cleavage in cells transfected with plasmids expressing sgHD1/i3 and SpCas9 relative to those lacking sgRNAs (Figure 3C). Semiquantitative PCR for the normal and mutant HTT mRNAs showed target mRNA knockdown (Figures 3D–3F), which for ND31551 is the normal allele, and for ND33392 is the mutant allele. Western blot for protein confirmed allele-specific reduction of the target allele (Figures 3G and 3H).

common intronic sgHDi3 sequence is shown. All of the samples are normalized to human GAPDH, and the results are the mean ± SEM relative to cells transfected with plasmids containing the SpCas9 only control (n = 6 independent experiments; §p < 0.001, #p < 0.0001, and one-way ANOVA followed by a Bonferroni's post hoc). (H) sgHD1/i3/SpCas9, sgHD3/i3/SpCas9, and sgHDi3/SpCas9 expression cassettes were transfected into HEK293 cells, and endogenous HTT protein levels were determined after puromycin selection and expansion. Cells transfected with Cas9 only were used as a control and beta catenin served as a loading control. (I) The quantification of HTT protein levels after treatment with sgHD/SpCas9 complexes is shown. The data are the mean ± SEM relative to cells transfected with plasmids containing SpCas9 only control (n = 6 independent experiments; #p < 0.0001, §p < 0.001, and one-way ANOVA followed by Bonferroni's post hoc).

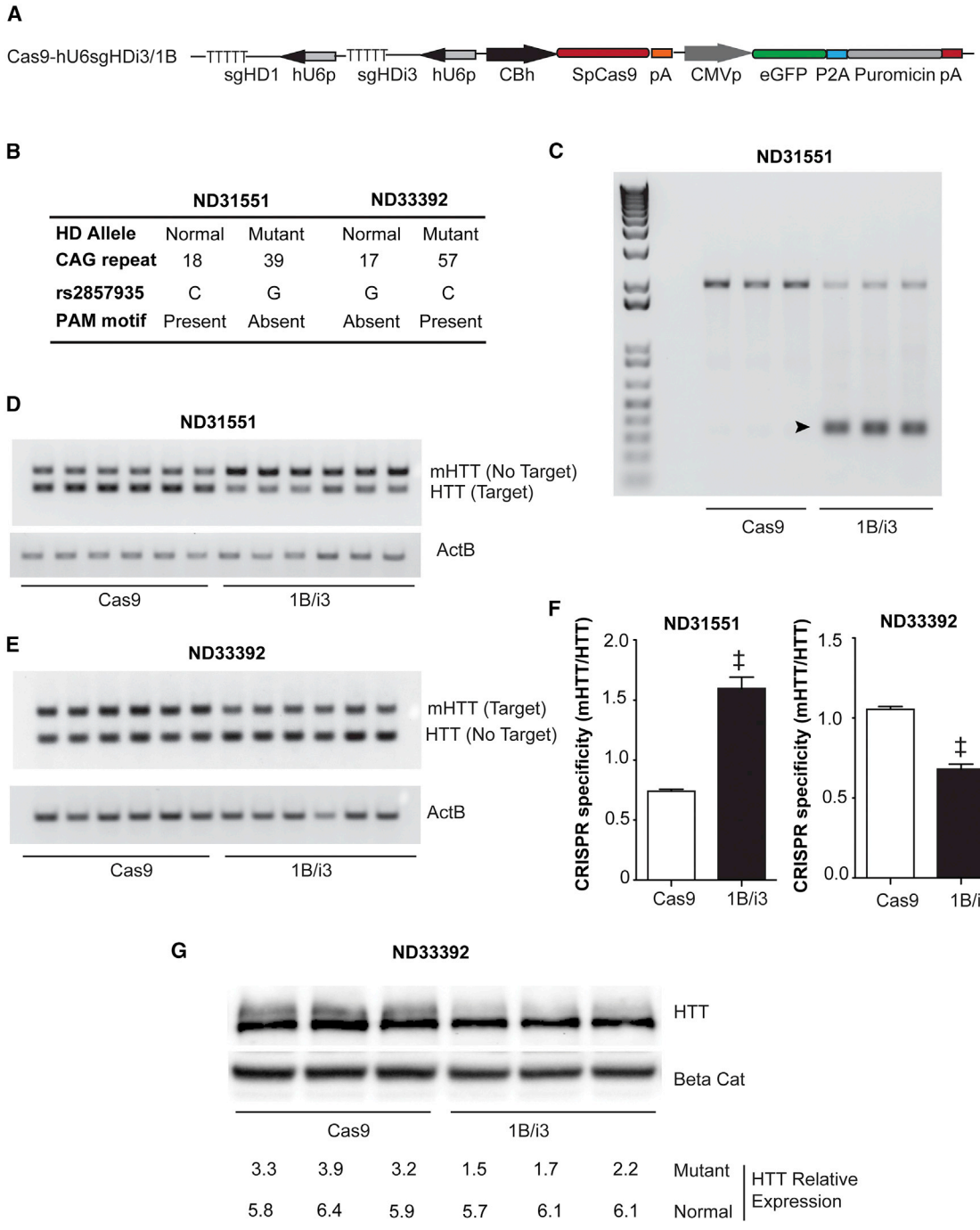


Figure 3. Assessment of Allele-Specific Cleavage in Human HD Fibroblasts

(A) Cartoon depicting the CRISPR expression plasmid used to co-express sgHD1 and sgHDi3 expression cassettes. SpCas9 and the selective reporter eGFP/puromycin expression cassettes present in the same plasmid are also shown. (B) ND31551 and ND33392 HD fibroblasts lines were selected to determine allele-specific deletion of *HTT*. CAG repeat length, nucleotide variation, and the allele location of the PAM motif are indicated in the image. (C) A representative genomic PCR showing *HTT* exon-1 deletion of DNA harvested from the electroporated ND31551 HD fibroblast cell line is shown in the image. The arrow indicates the expected PCR amplification product resulting from allele-specific deletion. (D and E) A semiquantitative PCR reaction showing the reduction of the targeted allele containing the conserved PAM sequence is shown in the image. For ND31551 fibroblasts, the PAM sequence is conserved in the normal allele, while for ND33392 fibroblasts, the PAM sequence is in the mutant allele. The expression levels are reduced only on the PAM-containing allele. (F) The quantification of mRNA reduction in treated HD fibroblasts is shown. The data show the ratio between mRNA levels of the mutant with respect to the normal allele, relative to cells electroporated with vectors expressing only the

(legend continued on next page)

A

	sgHD1	sgHD2	sgHD3	sgHDi3
Comp. (nt)	17	19	20	18
0 mm	1	1	1	1
1 mm	2	0	0	0
2 mm	32	1	0	2
3 mm	382	38	6	18
Total	416	40	7	21

B

	sgHD1	sgHD2	sgHD3	sgHDi3
Intergenic	71	11	2	6
Promoter	58	4	0	1
5'UTR	64	3	0	0
Intron	143	20	5	13
Exon	75	1	0	1
3'UTR	5	1	0	0
Total	416	40	7	21

C

5' GTTTAATTGAGTTGTCATATGTTAATAACGGTAT 3'
5' ATACCGTTATTAACATATGACAACTCAATTAAAC 3' (Rev Complement)

TGTCGCCCCGCTCCAGGCGTCGGTC----------GGCGGGAGTGGCGGGGCAGGGGGGG
TGTCGCCCCGCTCCAGGCGTCGG---ATACCGTTATTAACATATGACAACTCAATTAAAC--CGGCGGGAGTGGCGGGGCAGGGGGGG
TGTCGCCCCGCTCCAGGCGTCG--------GTTATTAACATATGACAACTCAATTAAAC---------CGGCGGGAGTGGCGGGGCAGGGGGGG
TGTCGCCCCGCTCCAGGCGTCGG-GTTTAATTGAGTTGTCATATGTTAATAACGGTATCTCGGCGGGAGTGGCGGGGCAGGGGGGG
TCGCCCCGCTCCAGGCGTCGG----------TATTAACATATGACAACTCAATTAAAC------CGGCGGGAGTGGCGGGGCAGGGGGGG
TGTCGCCCCGCTCCAGGCGTCGG----------ATTAACATATGACAACTCAATTAAAC----------CGGGAGTGGCGGGGCAGGGGGGG

HTT Promoter | ODN | HTT Intron1

D

Gene	Chromosome	Strand	Mismatches	Mismatch position	Location	Guide	Indels
SLC45A4	chr8	-	2	GCTCCA**t**GCG**g**CGGCGG	Exon	sgHD1	Undetected
CHRNA2	chr8	+	3	GCT**t**CAGGCGT**a**GG**g**GG	Exon	sgHD1	Undetected
TRIB1	chr8	+	3	GCTCC**g**GG**g**G**c**CGGCGG	Exon	sgHD1	Undetected
SMARCD1	chr12	+	3	GCTCCA**a**GCG**g**CGGCG**c**	Exon	sgHD1	Undetected
COX11	chr17	+	3	GCT**c**TTAGG**c**CGCC**g**CGG	Exon/5'UTR	sgHDi3	Undetected
STT3B	chr3	+	2	GC**a**CCAGGCG**g**CGGCGG	5'UTR	sgHD1	Undetected
NUP210	chr3	-	3	GCTTT**a**AGGA**t**GCCT**t**GG	Intron	sgHDi3	Undetected
DNAJC16	chr1	+	3	**a**CTTTTAGGA**g**GCCT**a**GG	Intron	sgHDi3	Undetected
LEPR	chr1	+	3	GC**a**TTTAGGA**g**GCCT**a**GG	Intron	sgHDi3	Undetected
CBFA2T3	chr16	+	1	GCTCCAGGCGT**g**GGCGG	Intron	sgHD1	Undetected
NAV2	chr11	+	2	GCTTTT**g**GGACGCCTC**a**G	Intron	sgHDi3	Undetected

Figure 4. Assessing Off-Target Activity of sgHD1 and sgHDi3/Cas9

(A) Table depicting the number of off-target sites for the most active sequences predicted to bind with 1, 2, or 3 mismatches. The nucleotide length of the complementary guide sequence is also indicated in the table. (B) Table highlighting the number of off-target sites binding at different genomic regions using the UCSC genome browser is shown. (C) The HD fibroblasts were electroporated with plasmids expressing sgHD1/i3 and SpCas9 along with an ODN sequence. The Sanger sequencing results showed the incorporation of the ODN sequence at the DNA cleavage site. A *HTT* promoter sequence and a *HTT* intron sequence outside the ODN sequence are also depicted. (D) Sanger sequencing results from 11 predicted off-target sites are shown. The gene name, chromosome position, DNA strand, number of mismatches and position within the guide, gene location, sgRNA sequence, and indel presence or absence are indicated.

Assessment of Off-Target Cleavage Sites in Edited HD Fibroblasts

Although truncated sgRNA sequences (<20 nt) are reported to have higher selectivity for the on-target site, any sgRNA/Cas9 complex can also generate unwanted dsDNA breaks at off-target sites that resemble the on-target sequence.[18,22,26,31] We used the Cas9-Off finder algorithm to predict the number of potential off-target sites for the most effective sgRNAs (sgHD1, sgHD2, sgHD3, and sgHDi3) and the UCSC genome browser for mapping their location in the human genome.[32] Our screen identified 416 sites for sgHD1, whereas 40, 21, and 7 off-targets are predicted for sgHD2, sgHDi3, and sgHD3, respectively (Figure 4A). Of note, sgHD1 has the shortest complementary sequence (17 nt), which could explain its higher frequency for genomic off targets. Importantly, all guides showed full complementarity only to *HTT* and more than 90% of the off-targets have three mismatches. As expected, they occur in

Cas9 control. The results are mean ± SEM relative to cells transfected with plasmids containing SpCas9 only control (n = 6; individual electroporations per group, ‡p < 0.01, Mann-Whitney t test). (G) A representative western blot showing allele-specific depletion (upper band of HTT doublet consisting of normal [lower] and expanded polyQ-containing proteins) after electroporation of HD fibroblasts with sgHD/SpCas9 expression vectors is shown in the image. The numbers below the image refer to relative HTT levels.

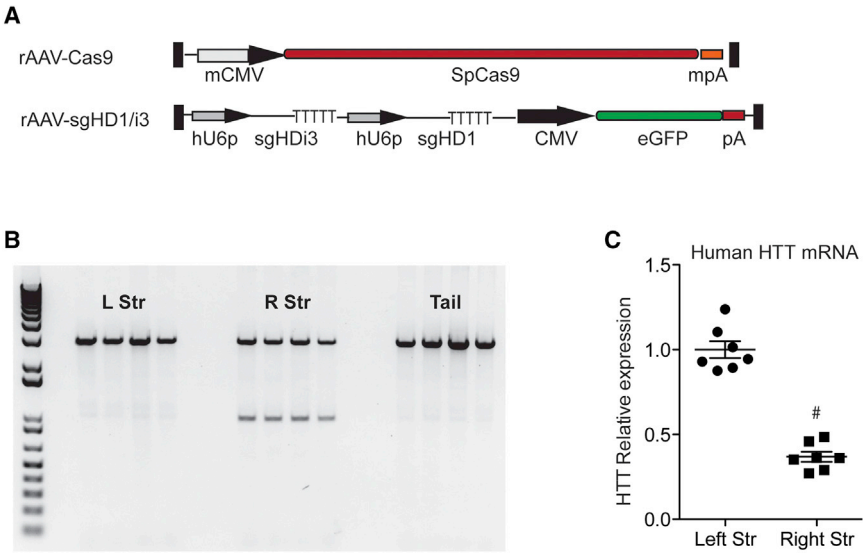


Figure 5. In Vivo Gene Editing of the Mutant *HTT* Allele

(A) Cartoon depicting rAAV shuttle vectors containing the SpCas9 and sgHD1/i3 expression cassettes (mCMV, minimal CMV promoter; mpA, minimal polyA and hU6p, human U6 promoter; pA, SV40 polyA). (B) PCR of isolated genomic DNA showing human *HTT* exon-1 targeted deletion after injection of vectors expressing SpCas9 and sgRNA sequences is shown. left striatum, LStr; right striatum, RStr. (C) qRT-PCR analysis of HTT mRNA levels in striatum samples harvested 3 weeks after SpCas9 and sgHD1/i3 delivery is shown. All of the samples were normalized to beta actin and the results are mean ± SEM relative to uninjected striatal samples (n = 7 animals per group, #p < 0.0001, and unpaired t test).

the promoter, 5′ UTR, exons, introns, 3′ UTR, and intergenic regions. The highest number was predicted within introns (Figure 4B). HD fibroblasts were electroporated with vectors expressing SpCas9 and sgHD1/i3, along with a short ODN sequence for mapping off-target dsDNA breaks.[22] As expected, the ODN sequence was incorporated within the *HTT* gene locus (Figure 4C), but it was not detected in any of the 11 top off-target sites tested (Figure 4D).

Editing *HTT* Exon-1 In Vivo

BacHD mice are transgenic for a modified human *HD* allele,[33] which fortuitously contains SNPs 1, 2, and 3. These mice were used to first evaluate the efficacy of mutant *HTT* editing in vivo at the genomic level. For this, recombinant AAVs (rAAVs) expressing either SpCas9 (rAAV.SpCas9) or the sgRNAs (rAAV.sgHD1/i3) were generated, which effectively delete human *HTT* exon-1 in vitro (Figures 5A and S7). Mice were injected on the right hemisphere with rAAV. SpCas9 plus rAAV.sgHD1B/i3. The left hemisphere was used as a control and left uninjected. Brains were harvested 3 weeks later and genomic DNA isolated. PCR amplification of genomic DNA demonstrates cleavage in the setting of Cas9 and sgRNA expressing AAVs only (Figure 5B). Accordingly, HTT mRNA levels were reduced on the right, but not the left, hemisphere in concordance with DNA cleavage (Figure 5C). Interestingly, mouse Htt mRNA levels were also reduced on the injected hemisphere, although to a lesser degree than the human *HTT* allele. Several binding sites for sgHD1 and sgHDi3 were identified within the mouse *Htt* locus. Three binding sites were identified for sgHD1, all within exon-1 and containing five mismatches. In contrast, a single binding site with two mismatches within the first intron was predicted for sgHDi3 (Figure S8). We hypothesize that indels within the intron caused by sgHDi3 may disrupt transcription factor binding sites and, similar to HEK293 cells expressing sgHDi3 alone (Figure 2G), modestly reduce mouse Htt expression.

Other Cas9 Systems for Silencing the *HTT* Allele

Finally, we aimed to broaden the approach we describe here for allele-specific editing of *HTT* to the other CRISPR/Cas9 and CRISPR/Cpf1 systems recently described. Thus, we screened which of the 47 SNPs flanking *HTT* exon-1 were contained within the conserved PAM nucleotide positions for the other systems. Engineered SpCas9 variants from *Streptococcus pyogenes* with altered PAM specificities have been generated (SpCas9_VQR, SpCas9_EQR, and SpCas9_VRER).[21] The SpCas9 VQR variant strongly recognizes sequences bearing the NGAN PAM and, with lower efficiency, those sites with a NGNG motif. SpCas9_EQR is more specific for an NGAG PAM. In contrast, SpCas9_VRER has a strong selectivity for a NGCG PAM sequence with no cleavage activity when this is varied. For SpCas9_VQR, the SNP could be positioned at the second or the third nucleotide of the NGAN PAM, as well as at the second and fourth nucleotide of the NGNG PAM sequence. In contrast, because of the selectivity of the SpCas9_EQR for NGAG and SpCas9_VRER for NGCG sequences, the SNP could be permitted at any position of their PAM (Table S1; Figure S1). The discovery of SaCas9 from *Staphylococcus aureus* has extended the number of CRISPR/Cas9 systems, with the advantage that a SaCas9-encoding transgene can be easily packaged into AAV viral vectors.[24] SaCas9 primarily recognizes a NNGRRT PAM, although dsDNA breaks are also observed at DNA targets adjacent to NNGRR motifs. For SaCas9, only those SNPs positioned at the third nucleotide of the PAM would allow for allele specificity (Table S1; Figure S1). A new Class 2 CRISPR system was recently identified that contains Cpf1 as effector protein to mediate dsDNA breaks.[23] Unlike Cas9 that recognizes a G-rich PAM motif, the Cpf1 PAM motif is T-rich. Currently, 16 Cpf1-family proteins have been characterized, but only the Cpf1 proteins from *Acidaminococcus* (AsCpf1) and *Lachnospiraceae* (LbCpf1) have shown robust DNA interference activity when expressed in mammalian cells. AsCpf1 has strong selectivity for a TTTN PAM and does not recognize any sequence variants. Therefore, SNPs present at any position

of the TTTN PAM could disrupt AsCpf1 recognition. In contrast, LbCpf1 recognizes multiple T-rich PAMs, albeit with different cleavage activity. Thus, for LbCpf1, only those SNPs where the variant nucleotide did not generate any other PAM that could be recognized above the LbCpf1 cleavage threshold activity could be considered for allele discrimination (Table S1; Figure S1).[23]

Overall, we identified 36 SNPs located within the specific PAM positions described above. Again, we found instances where the nucleotide variation caused the loss (12 SNPs), gain (11 SNPs), or a simultaneous loss in one DNA strand and a gain on the opposite strand (13 SNPs) (Table S1; Figure S1). Of special interests are the SNPs that generate a loss/gain, since CRISPR complexes could be designed for any of the two possible nucleotides linked to the mutant allele. Of note, we found instances where the same Cas9 protein would be predicted to target each nucleotide variation using a different sgRNA sequence or, alternatively, a different CRISPR effector protein could be used to target each nucleotide variant. Two interesting observations also arose from our screen. First, in the rs113331544 SNP, for which the minor allele contains a six-nucleotide insertion, the same PAM sequence is present on both alleles, but a different sgRNA sequence could be designed to tether SpCas9 to the mutant allele depending on the nucleotide variation. Second, for the rs28393280 and the rs28583447 SNPs, the nucleotide variation causes the gain of two PAM motifs on the same allele, one on the positive and the other on the negative DNA strand. Those SNPs could be appropriate for targeting with a nickase effector protein, which would efficiently generate on-target dsDNA breaks without detectable damage at potential off-target sites.[34]

DISCUSSION

Currently, reduction of HTT mRNA levels with RNAi and ASOs are the leading therapeutic options for HD.[9,35] However, it is unknown whether these treatments when administered to human patients will be as beneficial as observed in animal models. Here, we investigated the possibility of using the CRISPR/Cas9 technology to target and inactivate the mutant *HTT* allele. Targeted gene deletions can be generated when two sgRNA/Cas9 complexes cause dsDNA breaks followed by DNA repair.[13,15] Given the potency of CRISPR/Cas9 and the high likelihood of cleaving both *HTT* alleles, the role of HTT protein on important cellular functions,[16] and the fact that is unknown if complete loss of the huntingtin gene in human brain cells would also be tolerated as reported in adult mice,[36] allele-specific gene editing provides an important strategy to investigate. Indeed, recent studies demonstrated allele-specific editing by taking advantage of SNP positioned at these conserved PAM nucleotides.[37–39]

We designed guide RNAs that bind and tether SpCas9 to six prevalent SNPs located 5′ of *HTT* exon-1, which in combination with a guide binding within the first *HTT* intron, effectively eliminate expression of the HTT protein. We found that the distance between upstream and downstream guides influenced editing efficacy, as well as confirmed the SpCas9 preference in HD cell lines. Our studies also suggest that intronic transcription binding sites may effect *HTT* gene expression, since indels generated by SpCas9 within the first *HTT* intron modestly reduced gene expression. This is important when designing intronic guide sequences that are not allele specific, since expression of the normal allele could also be affected. Moreover, we demonstrate that by eliminating the HTT proximal promoter in addition to HTT exon-1, we not only terminate the production of toxic HTT-exon-1 proteins, but also truncated C-terminal HTT proteins.[30]

In HD fibroblast cell lines for which these SNPs are present in heterozygosity, we observed *HTT* exon-1 excision only on the allele where the nucleotide variation did not disrupt the PAM motif. Notably, SNP1 (rs2857935) has a prevalence of 22% among the human population. In the HD fibroblast lines tested here, we observed that 9 out of 11 were heterozygous for the SNP and the PAM was linked to the mutant allele. This raises the exciting possibility that this SNP is in linkage disequilibrium with the mutant allele in the general HD population and warrants further study.

While this manuscript was in preparation, Shin and colleagues published a similar study using CrispR/Cas9 to inactivate the mutant HTT allele.[39] While both studies demonstrate the efficacy of CRISPR/Cas9 to eliminate HTT exon-1, there are notable differences. First, Shin and colleagues generate a 44 kilobase deletion to inactivate HTT expression, whereas we create small-targeted deletions that are sufficient to terminate HTT expression. Second, we show that allele-specific HTT exon-1 deletion could be achieved using a single SNP-dependent PAM in the HTT promoter in combination with a common guide in intron 1, achieving elimination of N-terminal and C-terminal protein fragments. And third, we demonstrate for the first time, that sgRNA/Cas9 complexes are also effective in vivo in an HD mouse model. We found rAAV delivery of the sgRNA/SpCas9 complexes reduced human mutant HTT expression to 40% in treated hemispheres, a level of reduction known to provide benefit by RNAi or ASOs.[6,7,9]

The translation of this approach to humans will require delivery systems that allow for transient expression so as to reduce the risk of off-target consequences from lasting expression of the editing machinery, such as occurs in the setting of viral-mediated gene transfer systems. Additionally, the Cas9 protein is of bacterial origin and, if an immune response were elicited in human brain, edited cells could be effectively eliminated, mitigating the positive consequences of removing the mutant allele. Therefore, efforts should be extended to ensure that Cas9 and probably also the sgRNAs are transiently expressed.[40,41]

The importance of on-target selectivity is crucial when using gene-editing approaches. In our strategy, we used truncated sgRNA guides, which have been shown to minimize unintended dsDNA breaks.[42] We screened for potential off-targets from our guides using a combination of an in silico and PCR-based approach and did not detect disruption of the top predicted off-target genes. Importantly, additional tools with significant on-target selectivity such as the high fidelity Cas9 proteins and the Cas9 nickases, developed during the course of our study, could be tested for efficacy using our strategy.[34,43,44]

In summary, we developed and confirmed a strategy for allele-specific genome-editing of mutant *HTT* based on CRISPR/Cas9 technology that takes advantage of highly prevalent SNPs in the *HTT* locus for guiding mutant allele-specific cleavage and show its effectiveness in reducing the expression from mutant *HTT* alleles in human HD fibroblasts and mice brain.

MATERIALS AND METHODS

Prediction of SNP-Dependent PAM Motifs

SNPs with a prevalence of $\geq$5% located upstream (6.5 kilobase) and downstream (Intron1) of *HTT* exon-1 were obtained from the 1000 Genomes database using the NCBI variation viewer website (http://www.ncbi.nlm.nih.gov/variation/view/?q=HTT&filters=source:dbsnp&assm=GCF_000001405.25). To predict SNP-dependent PAM motifs, SNPs were screened against the consensus PAM sequences of *Streptococcus pyogenes* (SpCas9, NGG or NAG) and *Staphylococcus aureus* (SaCas9 NNGRRT) or the CRISPR/Cpf1 systems of *Acidaminococcus*, NTTT) and *Lachnospiraceae* (LbCpf1, heterogeneous PAMs). Only those SNPs positioned in a conserved nucleotide PAM position in which the nucleotide variation disrupted the consensus PAM were predicted as SNP-dependent PAM motifs.

Cell Culture and Transfection

HEK293 cells (obtained from CHOP Research Vector Core stock) were maintained in DMEM media containing 10% fetal bovine serum (FBS), 1% L-Glutamine, and 1% penicillin/streptomycin at 37°C with 5% CO_2. Cells were cultured in 24 well plates and transfected at 80%–90% confluence using Lipofectamine 2000 transfection reagent, according to the manufacturer's protocol. Human HD patient fibroblasts (obtained from Coriell Institute for Medical Research cell repository) were maintained on MEM media supplemented with 15% FBS, 1% MEM non-essential amino acids, 1% penicillin/streptomycin, and 1% L-Glutamine at 37°C with 5% CO_2. DNA transfection was done by electroporation using Invitrogen Neon transfection reagent using the electroporation conditions (ND31551: 1650V, 10 ms, and three pulses and ND33392: 1450V, 20 ms, and two pulses) and following the guidelines provided by manufacturer. Cells were not authenticated or tested for Mycoplasma by the investigators since they previously passed the quality controls of CHOP Research Vector Core and the Coriell Institute for Medical Research cell repository. None of the cells used in the study were listed in ICLAC database of commonly misidentified cell lines.

sgRNA and Cas9 Plasmid Construction

The plasmid pX330 containing the SpCas9 and sgRNA expression cassettes was kindly provided by Dr. Feng Zhang and used as a template for further modifications. To determine transfection efficacy and for selecting positive transfected cells, a CMV reporter cassette expressing eGFP/P2A/puromycin fusion protein was cloned downstream of the SpCas9 expression cassette. For all sgRNAs, the guide complementary sequences were cloned using a single cloning step with a pair of partially complementary oligonucleotides. The oligo pairs encoding the genomic complementary guide sequences were annealed and ligated into the BbsI cloning site upstream and in frame with the invariant scaffold of the sgRNA sequence. A hU6 or the hH1 Pol3 promoter was use to drive expression of the sgRNA sequences depending on the presence (hU6) or absence (hH1) of a guanine nucleotide at the transcription start site of the sgRNA sequence.

Genomic DNA Extraction, SNP Genotyping, and Genome Editing Analysis

Genomic DNA from HD fibroblast and HEK293 cell lines was extracted using a DNeasy Blood & Tissue Kit (QIAGEN) according to manufacturer's instructions. SNPs were genotyped by direct Sanger sequencing of PCR amplified products containing the SNPs and using the primers listed in Table S5. To determine which nucleotide variation of SNP1 (rs2857935) was linked to the normal or the mutant allele, the genomic sequence containing SNP1 and the CAG repeat was amplified by PCR and cloned into TOPO plasmids using the TOPO TA Cloning Kit and subsequently transformed into DH5alpha competent cells. Individual colonies were analyzed using Sanger sequencing to determine which nucleotide variant is associated with the normal or mutant allele. Deletions of *HTT* exon-1 were confirmed on genomic DNA samples by PCR, using primers binding outside the intervening segment cleaved by the sgRNA/SpCas9 complex pair (Table S5).

RNA Extraction, qRT-PCR, and SQ-PCR of HTT Expression Levels

Total RNA was extracted using TRIzol (Life Technologies) according to the manufacturer's protocol, with the exception of 1 μL Glycoblue (Life Technologies) in addition to the aqueous phase on the isopropanol precipitation step and a single wash with cold 70% ethanol. RNA samples were quantified by spectrophotometry and subsequently cDNAs were generated from 1 μg of total RNA with random hexamers (TaqMan RT reagents, Applied Biosystems). To determine human HTT expression levels in HD fibroblasts and HEK293 cells, we used TaqMan probes for human HTT and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) mRNAs obtained from Applied Biosystems. For determining human and mouse HTT expression levels in BACHD mice samples, we used TaqMan Probes for human HTT mRNA, mouse Htt mRNA, and mouse beta actin mRNA obtained from Applied Biosystems. Relative HTT gene expression was determined using the ddCt method. Allele-specific editing was determined by a Semiquantitative PCR amplification of the CAG repeat within *HTT* exon-1. RT-PCR experiments were carried out with cDNAs generated from 1 μg of total RNA and using 80 ng for PCR amplification. The RT-exponential phase was determined on 25–30 cycles to allow semiquantitative (SQ) comparison of cDNAs developed from identical reactions with BIOLASE Taq Polymerase (Bioline). A SQ-PCR reaction for ActinB (20 cycles) was used as a reference gene to determine loading differences between samples. The primers and conditions are shown in Table S5.

Huntingtin Western Blots

HEK293 cells were transfected with sgRNA/SpCas9 expression cassettes, selected for 2 days with puromycin (3 μM) and expanded until cells reached confluence. Then, cells were rinsed once with PBS and lysed with Passive Lysis Buffer (PBL, Promega). Protein

concentrations were determined using the DC protein assay (Bio-Rad) and 15 μg of protein loaded on a 3%–8% NuPAGE Tris-Acetate Gel (Novex Life Technologies). HD fibroblast cells were electroporated with sgRNA/SpCas9 expression cassettes, selected for 2 days with puromycin (2 μM), and expanded until cells reached confluence. Cells were then rinsed with iced-cold PBS, deattached, pelleted, snap froze, and lysed with SDP lysis buffer (50 mM Tris pH8.0, 150 mM NaCl, 1% NP40, 1× complete protease inhibitors, 1× phosphatase inhibitors, and 100 mM PMSF), followed by incubation on ice for 20 min with occasional vortexing. Debris was removed by centrifugation (15 min, 20,000 *g* 4°C) and the supernatant retained. Protein concentrations were determined using the DC protein assay (Bio-Rad). Samples (25 μg) were prepared for immunoblotting by denaturing the lysates in LDS sample buffer (Invitrogen) with 2× reducing agent (100 mM DTT, Invitrogen) and heating to 70°C for 10 min. Samples were resolved on a 10% low-Bis acrylamide gels (200:1 acrylamide:Bis) with Tris-glycine running buffer (25 mM Tris, 190 mM Glycine, and 0.1% SDS) containing 10.7 mM Beta mercaptoethanol. Gels were ran on ice for 40 min at 90 V through the stack, then at 190 V. Proteins were transferred o/n at 30 V and 4°C onto polyvinylidene fluoride (PVDF) membranes with NuPAGE Transfer Buffer (Invitrogen: 25 mM Bicine, 25 mM Bis-Tris, 1.025 mM EDTA, 5% MeOH, and pH7.2). Membranes were blocked with 5% milk in PBS-T and then blotted with a mouse anti-HTT (MAB2166. dilution: 1:5,000; Millipore) or rabbit anti beta-catenin (Ab2365. dilution: 1:5,000; Abcam) antibodies followed by horseradish peroxidase-coupled antibodies (Goat anti-mouse: 115-035-146. dilution: 1:10,000 or Goat anti-Rabbit: 111-035-144. dilution: 1:50,000; Jackson ImmunoResearch). Blots were developed with ECL Plus reagents (Amersham Pharmacia). HTT reduction was determined by densitometry (n = 3 independent experiments) of protein levels relative to beta-catenin with the ChemiDoc Imaging System (Bio-rad) and Image Lab analysis software.

rAAV Vector Design and Production

For in vivo studies, two different rAAV vectors were generated. One expressed SpCas9 and one the sgRNAs expression cassettes. Viruses were generated through the Research Vector Core at the Raymond G Perelman Center for Cellular and Molecular Therapeutics at The Children's Hospital of Philadelphia. SpCas9 was expressed under the control of a minimal cytomegalovirus immediate-early gene enhancer/promoter region (CMV promoter) and cloned upstream of a minimal poly A sequence (FBAAV-Cas9). The sgRNA expression cassettes were moved into an AAV shuttle plasmid containing an eGFP gene under the control of the CMV promoter and upstream of an SV40pA signal. All rAAV plasmid shuttles have AAV2 inverted terminal repeat sequences. RAAV vectors were produced by the standard calcium phosphate transfection method in HEK293 cells by using the AdHelper, AAV1 trans-packaging, and AAV shuttle plasmids. Vector titers were determined by RT-PCR and were 1×10^{13} vg/mL. Vector purity was also tested by silver stain.

Off-Target Analysis

Potential off-target loci for sgHD guide sequences in the human genome were determined using the Cas9-Off finder algorithm previously described by Bae et al.[32] To determine potential off-target sites, HD human fibroblasts were electroporated with an sgHD1/i3_Cas9 expression cassette and a 35 nt ODN sequence as described by Tsai et al.[22] At 1 week after transfection, genomic DNA was obtained and amplicons generated with Phusion polymerase using PCR primers flanking the potential site. Amplicons were subjected to Sanger sequencing to determine mutations in the cleavage site using specific primers, as well as cloned into TOPO-cloning system for sequence confirmation using 3–4 colonies/site.

Mouse Studies

Animal protocols were approved by The Children's Hospital of Philadelphia Institutional Animal Care and Use Committee. The 6-week-old male BacHD mice (FVB/N-Tg(HTT*97Q)IXwy/J transgenic mice) were obtained from Jackson Laboratories. Mice were housed in a temperature-controlled environment on a 12 hr light/dark cycle. Food and water were provided ad libitum. Mice were injected at 8 weeks of age with a combination 1:1 of rAAV2/1-SpCas9 virus and rAAV-hU6sgRNA/eGFP virus. For rAAV injections, mice were anesthetized with isoflurane and 5 μL of rAAV mixture injected unilaterally into the right striata at 0.2 μL/min (coordinates: +0.86 mm rostral to Bregma, −1.8 mm lateral to medial, and −2.5 mm ventral from brain surface). After 3 weeks, mice were anesthetized with a ketamine and xylazine mix and perfused with 18 mL of 0.9% cold saline mixed with 2 mL RNAlater (Ambion) solution. Brains were removed, blocked, and cut into 1-mm-thick coronal slices. Tissue punches from striata were taken using a tissue corer (1.4-mm in diameter; Zivic Instruments). All tissue punches were flash frozen in liquid nitrogen and stored at −80°C until use.

Statistical Analysis

All statistical analyses were performed using GraphPad Prism v5.0 software. For all in vitro and in vivo studies, the appropriate experimental sample size was confirmed by performing a power analysis with $\alpha = 0.05$ resulting on $\beta > 0.8$. Outlier samples were detected using the Grubb's test ($\alpha = 0.05$). Normal distribution of the samples was determined by using the Kolmogorov-Smirnov normality test. All data with normal distribution were analyzed using one-way ANOVA followed by a Bonferroni's post hoc or an unpaired t test. Otherwise, data without normal distribution were analyzed using a Mann-Whitney test as indicated. Statistical significance was considered *$p < 0.05$, ‡$p < 0.01$, §$p < 0.001$, and #$p < 0.0001$. All results are shown as the mean ± SEM.

SUPPLEMENTAL INFORMATION

Supplemental Information includes eight figures and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.ymthe.2016.11.010.

AUTHOR CONTRIBUTIONS

A.M.M. and B.L.D. developed the study, designed the experiments, and analyzed the data. A.M.M. and S.A.E. carried out CRISPR-Cas9 related experiments and analyzed data. M.S.K. performed the rAAV

injections and assisted with necropsies. A.M.M. and B.L.D. wrote the manuscript with input from all authors.

ACKNOWLEDGMENTS

Funding support was provided by Hoppy's Hope Foundation (A.M.M.), Philly Cure HD (B.L.D.), The Leslie Gehry Brenner Prize (B.L.D.), The Foerderer Grant for Excellence (A.M.M.), the NIH (NS084475 and NS076631; B.L.D.), and The Children's Hospital of Philadelphia Research Institute. The authors would like to thank Luis Tecedor for statistical consultation.

REFERENCES

1. Walker, F.O. (2007). Huntington's disease. Semin. Neurol. *27*, 143–150.
2. Hicks, R.R., Smith, D.H., Lowenstein, D.H., Saint Marie, R., and McIntosh, T.K. (1993). Mild experimental brain injury in the rat induces cognitive deficits associated with regional neuronal loss in the hippocampus. J. Neurotrauma *10*, 405–414.
3. Johnson, C.D., and Davidson, B.L. (2010). Huntington's disease: progress toward effective disease-modifying treatments and a cure. Hum. Mol. Genet. *19* (R1), R98–R102.
4. Yamamoto, A., Lucas, J.J., and Hen, R. (2000). Reversal of neuropathology and motor dysfunction in a conditional model of Huntington's disease. Cell *101*, 57–66.
5. Díaz-Hernández, M., Torres-Peraza, J., Salvatori-Abarca, A., Morán, M.A., Gómez-Ramos, P., Alberch, J., and Lucas, J.J. (2005). Full motor recovery despite striatal neuron loss and formation of irreversible amyloid-like inclusions in a conditional mouse model of Huntington's disease. J. Neurosci. *25*, 9773–9781.
6. Harper, S.Q., Staber, P.D., He, X., Eliason, S.L., Martins, I.H., Mao, Q., Yang, L., Kotin, R.M., Paulson, H.L., and Davidson, B.L. (2005). RNA interference improves motor and neuropathological abnormalities in a Huntington's disease mouse model. Proc. Natl. Acad. Sci. USA *102*, 5820–5825.
7. Boudreau, R.L., McBride, J.L., Martins, I., Shen, S., Xing, Y., Carter, B.J., and Davidson, B.L. (2009). Nonallele-specific silencing of mutant and wild-type huntingtin demonstrates therapeutic efficacy in Huntington's disease mice. Mol. Ther. *17*, 1053–1063.
8. Drouet, V., Perrin, V., Hassig, R., Dufour, N., Auregan, G., Alves, S., Bonvento, G., Brouillet, E., Luthi-Carter, R., Hantraye, P., and Déglon, N. (2009). Sustained effects of nonallele-specific Huntingtin silencing. Ann. Neurol. *65*, 276–285.
9. Kordasiewicz, H.B., Stanek, L.M., Wancewicz, E.V., Mazur, C., McAlonis, M.M., Pytel, K.A., Artates, J.W., Weiss, A., Cheng, S.H., Shihabuddin, L.S., et al. (2012). Sustained therapeutic reversal of Huntington's disease by transient repression of huntingtin synthesis. Neuron *74*, 1031–1044.
10. Garriga-Canut, M., Agustín-Pavón, C., Herrmann, F., Sánchez, A., Dierssen, M., Fillat, C., and Isalan, M. (2012). Synthetic zinc finger repressors reduce mutant huntingtin expression in the brain of R6/2 mice. Proc. Natl. Acad. Sci. USA *109*, E3136–E3145.
11. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.
12. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.
13. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.
14. 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.
15. Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. eLife *2*, e00471.
16. Zuccato, C., Valenza, M., and Cattaneo, E. (2010). Molecular mechanisms and potential therapeutical targets in Huntington's disease. Physiol. Rev. *90*, 905–981.
17. Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature *507*, 62–67.
18. Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat. Biotechnol. *31*, 822–826.
19. Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. Nat. Biotechnol. *32*, 677–683.
20. Anders, C., Niewoehner, O., Duerst, A., and Jinek, M. (2014). Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature *513*, 569–573.
21. 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.
22. Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P., et al. (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat. Biotechnol. *33*, 187–197.
23. 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.
24. 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.
25. Coles, R., Caswell, R., and Rubinsztein, D.C. (1998). Functional analysis of the Huntington's disease (HD) gene promoter. Hum. Mol. Genet. *7*, 791–800.
26. Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. *31*, 827–832.
27. Zhang, Y., Ge, X., Yang, F., Zhang, L., Zheng, J., Tan, X., Jin, Z.B., Qu, J., and Gu, F. (2014). Comparison of non-canonical PAMs for CRISPR/Cas9-mediated DNA cleavage in human cells. Sci. Rep. *4*, 5405.
28. Moreno-Mateos, M.A., Vejnar, C.E., Beaudoin, J.D., Fernandez, J.P., Mis, E.K., Khokha, M.K., and Giraldez, A.J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. Nat. Methods *12*, 982–988.
29. Ochaba, J., Lukacsovich, T., Csikos, G., Zheng, S., Margulis, J., Salazar, L., Mao, K., Lau, A.L., Yeung, S.Y., Humbert, S., et al. (2014). Potential function for the Huntingtin protein as a scaffold for selective autophagy. Proc. Natl. Acad. Sci. USA *111*, 16889–16894.
30. El-Daher, M.T., Hangen, E., Bruyère, J., Poizat, G., Al-Ramahi, I., Pardo, R., Bourg, N., Souquere, S., Mayet, C., Pierron, G., et al. (2015). Huntingtin proteolysis releases non-polyQ fragments that cause toxicity through dynamin 1 dysregulation. EMBO J. *34*, 2255–2271.
31. Cho, S.W., Kim, S., Kim, J.M., and Kim, J.S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat. Biotechnol. *31*, 230–232.
32. Bae, S., Park, J., and Kim, J.S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics *30*, 1473–1475.
33. Gray, M., Shirasaki, D.I., Cepeda, C., André, V.M., Wilburn, B., Lu, X.H., Tao, J., Yamazaki, I., Li, S.H., Sun, Y.E., et al. (2008). Full-length human mutant huntingtin with a stable polyglutamine repeat can elicit progressive and selective neuropathogenesis in BACHD mice. J. Neurosci. *28*, 6182–6195.
34. Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., and Zhang, F. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell *154*, 1380–1389.
35. McBride, J.L., Pitzer, M.R., Boudreau, R.L., Dufour, B., Hobbs, T., Ojeda, S.R., and Davidson, B.L. (2011). Preclinical safety of RNAi-mediated HTT suppression in the rhesus macaque as a potential therapy for Huntington's disease. Mol. Ther. *19*, 2152–2162.

36. Wang, G., Liu, X., Gaertig, M.A., Li, S., and Li, X.J. (2016). Ablation of huntingtin in adult neurons is nondeleterious but its depletion in young mice causes acute pancreatitis. Proc. Natl. Acad. Sci. USA *113*, 3359–3364.

37. Li, Y., Mendiratta, S., Ehrhardt, K., Kashyap, N., White, M.A., and Bleris, L. (2016). Exploiting the CRISPR/Cas9 PAM constraint for single-nucleotide resolution interventions. PLoS ONE *11*, e0144970.

38. Courtney, D.G., Moore, J.E., Atkinson, S.D., Maurizi, E., Allen, E.H., Pedrioli, D.M., McLean, W.H., Nesbit, M.A., and Moore, C.B. (2016). CRISPR/Cas9 DNA cleavage at SNP-derived PAM enables both in vitro and in vivo KRT12 mutation-specific targeting. Gene Ther. *23*, 108–112.

39. Shin, J.W., Kim, K.H., Chao, M.J., Atwal, R.S., Gillis, T., MacDonald, M.E., Gusella, J.F., and Lee, J.M. (2016). Permanent inactivation of Huntington's disease mutation by personalized allele-specific CRISPR/Cas9. Hum. Mol. Genet., Published online September 15, 2016. http://dx.doi.org/10.1093/hmg/ddw286.

40. Hendel, A., Bak, R.O., Clark, J.T., Kennedy, A.B., Ryan, D.E., Roy, S., Steinfeld, I., Lunstad, B.D., Kaiser, R.J., Wilkens, A.B., et al. (2015). Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. Nat. Biotechnol. *33*, 985–989.

41. Rahdar, M., McMahon, M.A., Prakash, T.P., Swayze, E.E., Bennett, C.F., and Cleveland, D.W. (2015). Synthetic CRISPR RNA-Cas9-guided genome editing in human cells. Proc. Natl. Acad. Sci. USA *112*, E7110–E7117.

42. Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat. Biotechnol. *32*, 279–284.

43. Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., and Joung, J.K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature *529*, 490–495.

44. Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. Science *351*, 84–88.

YMTHE, Volume *25*

Supplemental Information

CRISPR/Cas9 Editing of the Mutant Huntingtin Allele In Vitro and In Vivo

Alex Mas Monteys, Shauna A. Ebanks, Megan S. Keiser, and Beverly L. Davidson

Figure S1: List of SNP-dependent PAMs targeted by different CRISPR systems flanking *HTT* exon-1.

HD PROMOTER TARGETED SNPs

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs35631490	GTCTGCGTCAGGGTTTCCTTCTTTT[C/G]CAGCCCCACCCCGCGTGCATCCCAC	C	G

SpCas9_WT (PAM Motif: NRG)

Ref. Allele:

```
                                CGGGGTGGGGCGCACGTAGG
5′ GTCTGCGTCAGGGTTTCCTTCTTTTCCAGCCCCACCCCGCGTGCATCCCAC 3′
3′ CAGACGCAGTCCCAAAGGAAGAAAAGGTCGGGGTGGGGCGCACGTAGGGTG 5′
                           PAM
```

Minor Allele

```
5′ GTCTGCGTCAGGGTTTCCTTCTTTTGCAGCCCCACCCCGCGTGCATCCCAC 3′
3′ CAGACGCAGTCCCAAAGGAAGAAAACGTCGGGGTGGGGCGCACGTAGGGTG 5′
```

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:

```
                               TCGGGGTGGGGCGCACGTAG
5′ GTCTGCGTCAGGGTTTCCTTCTTTTCCAGCCCCACCCCGCGTGCATCCCAC 3′
3′ CAGACGCAGTCCCAAAGGAAGAAAAGGTCGGGGTGGGGCGCACGTAGGGTG 5′
                          PAM
```

Minor Allele

```
5′ GTCTGCGTCAGGGTTTCCTTCTTTTGCAGCCCCACCCCGCGTGCATCCCAC 3′
3′ CAGACGCAGTCCCAAAGGAAGAAAACGTCGGGGTGGGGCGCACGTAGGGTG 5′
```

SaCas9(NNGRRT/NNGRR):

Ref. Allele:

```
                                CGGGGTGGGGCGCACGTAGG
5′ GTCTGCGTCAGGGTTTCCTTCTTTTCCAGCCCCACCCCGCGTGCATCCCAC 3′
3′ CAGACGCAGTCCCAAAGGAAGAAAAGGTCGGGGTGGGGCGCACGTAGGGTG 5′
                           PAM
```

Minor Allele

```
5′ GTCTGCGTCAGGGTTTCCTTCTTTTGCAGCCCCACCCCGCGTGCATCCCAC 3′
3′ CAGACGCAGTCCCAAAGGAAGAAAACGTCGGGGTGGGGCGCACGTAGGGTG 5′
```

LbCpf1(TTTN/TTCN/CTTN/TCTN/ATTN/TCCN/…):
Ref. Allele:

```
                          PAM
5′ GTCTGCGTCAGGGTTTCCTTCTTTTCCAGCCCCACCCCGCGTGCATCCCAC 3′
3′ CAGACGCAGTCCCAAAGGAAGAAAAGGTCGGGGTGGGGCGCACGTAGGGTG 5′
                             GCCCCACCCCGCGTGCATCC
```

Minor Allele

```
5′ GTCTGCGTCAGGGTTTCCTTCTTTTGCAGCCCCACCCCGCGTGCATCCCAC 3′
3′ CAGACGCAGTCCCAAAGGAAGAAAACGTCGGGGTGGGGCGCACGTAGGGTG 5′
```

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs73086139	TCAAGGCCTCTTCTCTCTTCTCGGC[A/G]GGACAGGCACAGGCAGGTGGCCAGG	A	G

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:

```
5′ TCAAGGCCTCTTCTCTCTTCTCGGCAGGACAGGCACAGGCAGGTGGCCAGG 3′
3′ AGTTCCGGAGAAGAGAGAAGAGCCGTCCTGTCCGTGTCCGTCCACCGGTCC 5′
```

Minor Allele:

```
                         PAM
5′ TCAAGGCCTCTTCTCTCTTCTCGGCGGACAGGCACAGGCAGGTGGCCAGG 3′
3′ AGTTCCGGAGAAGAGAGAAGAGCCGCCTGTCCGTGTCCGTCCACCGGTCC 5′
      GGCCTCTTCTCTCTTCTCG
```

SpCas9_VRER (PAM Motif: NGCG)
Ref. Allele:

```
5′ TCAAGGCCTCTTCTCTCTTCTCGGCAGGACAGGCACAGGCAGGTGGCCAGG 3′
3′ AGTTCCGGAGAAGAGAGAAGAGCCGTCCTGTCCGTGTCCGTCCACCGGTCC 5′
```

Minor Allele:

```
                         PAM
5′ TCAAGGCCTCTTCTCTCTTCTCGGCGGACAGGCACAGGCAGGTGGCCAGG 3′
3′ AGTTCCGGAGAAGAGAGAAGAGCCGCCTGTCCGTGTCCGTCCACCGGTCC 5′
```

```
       GGCCTCTTCTCTCTTCTCG
```

SaCas9(NNGRRT/NNGRR):

```
Ref. Allele:
5′ TCAAGGCCTCTTCTCTCTTCTCGGCAGGACAGGCACAGGCAGGTGGCCAGG
3′ AGTTCCGGAGAAGAGAGAAGAGCCGTCCTGTCCGTGTCCGTCCACCGGTCC

Minor Allele:
                          PAM
5′ TCAAGGCCTCTTCTCTCTTCTCGGCGGGACAGGCACAGGCAGGTGGCCAGG
3′ AGTTCCGGAGAAGAGAGAAGAGCCGCCCTGTCCGTGTCCGTCCACCGGTCC
       GGCCTCTTCTCTCTTCTCG
```

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs73086140	GCCAGGTGTCATGCTTAGCTCCCCG[C/G]CCAGTGAGATTCTTTCATTTAACAA	C	G

SpCas9_WT (PAM Motif: NRG)

```
Ref. Allele:
                              TCACTCTAAGAAAGTAAATT
5′ GCCAGGTGTCATGCTTAGCTCCCCGCCCAGTGAGATTCTTTCATTTAACAA
3′ CGGTCCACAGTACGAATCGAGGGGCGGGTCACTCTAAGAAAGTAAATTGTT
                            PAM
Minor Allele:
                         PAM
5′ GCCAGGTGTCATGCTTAGCTCCCCGGCCAGTGAGATTCTTTCATTTAACAA
3′ CGGTCCACAGTACGAATCGAGGGGCCGGTCACTCTAAGAAAGTAAATTGTT
      AGGTGTCATGCTTAGCTCCC
```

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

```
Ref. Allele:
                               GGTCACTCTAAGAAAGTAAATT
5′ GCCAGGTGTCATGCTTAGCTCCCCGCCCAGTGAGATTCTTTCATTTAACAA
3′ CGGTCCACAGTACGAATCGAGGGGCGGTCACTCTAAGAAAGTAAATTGTT
                         PAM
Minor Allele:
```

```
5′ GCCAGGTGTCATGCTTAGCTCCCCGGCCAGTGAGATTCTTTCATTTAACAA
   CGGTCCACAGTACGAATCGAGGGGCCGGTCACTCTAAGAAAGTAAATTGTT
```

SpCas9_VRER (PAM Motif: NGCG)

Ref. Allele:

```
                              GGTCACTCTAAGAAAGTAAATT
5′ GCCAGGTGTCATGCTTAGCTCCCCGCCAGTGAGATTCTTTCATTTAACAA
3′ CGGTCCACAGTACGAATCGAGGGGCGGTCACTCTAAGAAAGTAAATTGTT
                           PAM
```

Minor Allele:

```
5′ GCCAGGTGTCATGCTTAGCTCCCCGGCCAGTGAGATTCTTTCATTTAACAA
   CGGTCCACAGTACGAATCGAGGGGCCGGTCACTCTAAGAAAGTAAATTGTT
```

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs61791259	GATTACAGGCACCTGCCACCATGCC[C/T]GGCTAATTTTTGTATTTTTAGTTGA	C	T

SaCas9 (PAM Motif: NNGRRT)

Ref. Allele:

```
                                GATTAAAAACATAAAAATCA
5′ GATTACAGGCACCTGCCACCATGCCCGGCTAATTTTTGTATTTTTAGTTGA
3′ CTAATGTCCGTGGACGGTGGTACGGCCGATTAAAAACATAAAAATCAACT
                           PAM
```

Minor Allele:

```
GATTACAGGCACCTGCCACCATGCCTGGCTAATTTTTGTATTTTTAGTTGA
CTAATGTCCGTGGACGGTGGTACGGACCGATTAAAAACATAAAAATCAACT
```

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs61792460	GAGGGTTTCATCTTGTTGGTCAGGC[A/G]GACTTGAACTCCTGACCTCAGGTGA	G	A

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:

```
                          PAM
GAGGGTTTCATCTTGTTGGTCAGGCGGACTTGAACTCCTGACCTCAGGTGA
CTCCCAAAGTAGAACAACCAGTCCGCCTGAACTTGAGGACTGGAGTCCACT
   GGTTTCATCTTGTTGGTCAG

Minor Allele:
GAGGGTTTCATCTTGTTGGTCAGGCAGACTTGAACTCCTGACCTCAGGTGA
CTCCCAAAGTAGAACAACCAGTCCGTCTGAACTTGAGGACTGGAGTCCACT
```

SpCas9_VRER (PAM Motif: NGCG)

```
Ref. Allele:
                          PAM
GAGGGTTTCATCTTGTTGGTCAGGCGGACTTGAACTCCTGACCTCAGGTGA
CTCCCAAAGTAGAACAACCAGTCCGCCTGAACTTGAGGACTGGAGTCCACT
   GGTTTCATCTTGTTGGTCAG

Minor Allele:
GAGGGTTTCATCTTGTTGGTCAGGCAGACTTGAACTCCTGACCTCAGGTGA
CTCCCAAAGTAGAACAACCAGTCCGTCTGAACTTGAGGACTGGAGTCCACT
```

SaCas9(NNGRRT/NNGRR)

```
Ref. Allele:
                           PAM
GAGGGTTTCATCTTGTTGGTCAGGCGGACTTGAACTCCTGACCTCAGGTGA
CTCCCAAAGTAGAACAACCAGTCCGCCTGAACTTGAGGACTGGAGTCCACT
  GGGTTTCATCTTGTTGGTCAG

Minor Allele:
GAGGGTTTCATCTTGTTGGTCAGGCAGACTTGAACTCCTGACCTCAGGTGA
CTCCCAAAGTAGAACAACCAGTCCGTCTGAACTTGAGGACTGGAGTCCACT
```

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs73086145	GCTGTTTTTTCTTTAGAATGGGGGA[C/G]GTTATCAGGCTCTACATGGTGTGTA	C	G

SpCas9_WT (PAM Motif: NRG)

```
Ref. Allele:
GCTGTTTTTTCTTTAGAATGGGGGACGTTATCAGGCTCTACATGGTGTGTA
```

CGACAAAAAAGAAATCTTACCCCCTGCAATAGTCCGAGATGTACCACACAT

Minor Allele

PAM

GCTGTTTTTTCTTTAGAATGGGG**GAG**GTTATCAGGCTCTACATGGTGTGTA
CGACAAAAAAGAAATCTTACCCCCTCCAATAGTCCGAGATGTACCACACAT
GTTTTTTCTTTAGAATGGGG

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs61792461	TCTACATGGTGTGTAGTCGGCTAGC[A/G]TGTTGTAAGCCTTTCCCTGTGTCAC	A	G

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:

TCTACATGGTGTGTAGTCGGCTAGCATGTTGTAAGCCTTTCCCTGTGTCAC
AGATGTACCACACATCAGCCGATCGTACAACATTCGGAAAGGGACACAGTG

Minor Allele

PAM

TCTACATGGTGTGTAGTCGGCTA**GCG**TGTTGTAAGCCTTTCCCTGTGTCAC
AGATGTACCACACATCAGCCGATCGCACAACATTCGGAAAGGGACACAGTG
ACATGGTGTGTAGTCGGCTA

SpCas9_VRER (PAM Motif: NGCG)

Ref. Allele:

TCTACATGGTGTGTAGTCGGCTAGCATGTTGTAAGCCTTTCCCTGTGTCAC
AGATGTACCACACATCAGCCGATCGTACAACATTCGGAAAGGGACACAGTG

Minor Allele

PAM

TCTACATGGTGTGTAGTCGGCTA**GCG**TGTTGTAAGCCTTTCCCTGTGTCAC
AGATGTACCACACATCAGCCGATCGCACAACATTCGGAAAGGGACACAGTG
ACATGGTGTGTAGTCGGCTA

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs61792462	CTGTGTCACAAGTGCTCATCTGGAA[C/G]AGGATTCTAATGACTGCCTGTGGCT	C	G

SpCas9_WT (PAM Motif: NRG)

Ref. Allele:

CTGTGTCACAAGTGCTCATCTGGAACAGGATTCTAATGACTGCCTGTGGCT
GACACAGTGTTCACGAGTAGACCTTGTCCTAAGATTACTGACGGACACCGA

Minor Allele

PAM

CTGTGTCACAAGTGCTCATCTGG**AAG**AGGATTCTAATGACTGCCTGTGGCT
GACACAGTGTTCACGAGTAGACCTTCTCCTAAGATTACTGACGGACACCGA
GTCACAAGTGCTCATCTGG

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:

CTGTGTCACAAGTGCTCATCTGGAACAGGATTCTAATGACTGCCTGTGGCT
GACACAGTGTTCACGAGTAGACCTTGTCCTAAGATTACTGACGGACACCGA

Minor Allele

PAM

CTGTGTCACAAGTGCTCATCTGGA**AGA**GGATTCTAATGACTGCCTGTGGCT
GACACAGTGTTCACGAGTAGACCTTCTCCTAAGATTACTGACGGACACCGA
GTCACAAGTGCTCATCTGGA

SaCas9(NNGRRT/NNGRR)

Ref. Allele:

CTGTGTCACAAGTGCTCATCTGGAACAGGATTCTAATGACTGCCTGTGGCT
GACACAGTGTTCACGAGTAGACCTTGTCCTAAGATTACTGACGGACACCGA

Minor Allele

PAM

CTGTGTCACAAGTGCTCATCTGG**AAGAG**GATTCTAATGACTGCCTGTGGCT
GACACAGTGTTCACGAGTAGACCTTCTCCTAAGATTACTGACGGACACCGA
GTCACAAGTGCTCATCTGG

LbCpf1(TTTN/TTCN/CTTN/TCTN/ATTN/TCCN/…)
Ref. Allele:
CTGTGTCACAAGTGCTCATCTGGAACAGGATTCTAATGACTGCCTGTGGCT
GACACAGTGTTCACGAGTAGACCTTGTCCTAAGATTACTGACGGACACCGA

Minor Allele
GACACAGTGTTCACGAGTAGACC
CTGTGTCACAAGTGCTCATCTGGAAGAGGATTCTAATGACTGCCTGTGGCT
GACACAGTGTTCACGAGTAGACC**TTCT**CCTAAGATTACTGACGGACACCGA
PAM

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs61090955	TCATTTTATGTGATTCCTTTCTAGA[A/G]GTACTACTCATTACTTCTGCTTGCA	A	G

SpCas9_WT (PAM Motif: NRG)
Ref. Allele:
TCATTTTATGTGATTCCTTTCTAGAAGTACTACTCATTACTTCTGCTTGCA
AGTAAAATACACTAAGGAAAGATCTTCATGATGAGTAATGAAGACGAACGT

Minor Allele
PAM
TCATTTTATGTGATTCCTTTCTA**GAG**GTACTACTCATTACTTCTGCTTGCA
AGTAAAATACACTAAGGAAAGATCTCCATGATGAGTAATGAAGACGAACGT
TTTTATGTGATTCCTTTCTA

LbCpf1(TTTN/TTCN/CTTN/TCTN/ATTN/TCCN/…)
Ref. Allele:
GTAAAATACACTAAGGAAAGA
TCATTTTATGTGATTCCTTTCTAGAAGTACTACTCATTACTTCTGCTTGCA
AGTAAAATACACTAAGGAAAGA**TCTT**CATGATGAGTAATGAAGACGAACGT
PAM
Minor Allele
TCATTTTATGTGATTCCTTTCTAGAGGTACTACTCATTACTTCTGCTTGCA

AGTAAAATACACTAAGGAAAGATCTCCATGATGAGTAATGAAGACGAACGT

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs77384845	TAGCTGAAGGAAGGACAGGGACTGT[C/T]ATACACTAGCTAAGAGGCAAACTGC	C	T

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:

```
                          ATGTGATCGATTCTCCG
TAGCTGAAGGAAGGACAGGGACTGTCATACACTAGCTAAGAGGCAAACTGC
ATCGACTTCCTTCCTGTCCCTGACAGTATGTGATCGATTCTCCGTTTGACG
                        PAM
```

Minor Allele

```
TAGCTGAAGGAAGGACAGGGACTGTTATACACTAGCTAAGAGGCAAACTGC
ATCGACTTCCTTCCTGTCCCTGACAATATGTGATCGATTCTCCGTTTGACG
```

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs61792464	TCCCTCATTCAGGTTGATGTCCTAA[C/G]CCCCAGAACCTCAGAATGGGATTGT	G	C

SpCas9_WT (PAM Motif: NRG)

Ref. Allele:

```
                            GGTCTTGGAGTCTTACCCTA
TCCCTCATTCAGGTTGATGTCCTAACCCCCAGAACCTCAGAATGGGATTGT
AGGGAGTAAGTCCAACTACAGGATTGGGGGTCTTGGAGTCTTACCCTAACA
                         PAM
Minor Allele           PAM
TCCCTCATTCAGGTTGATGTCCTAAGCCCCAGAACCTCAGAATGGGATTGT
AGGGAGTAAGTCCAACTACAGGATTCGGGGTCTTGGAGTCTTACCCTAACA
    TCATTCAGGTTGATGTCCT
```

LbCpf1(TTTN/TTCN/CTTN/TCTN/ATTN/TCCN/…)

Ref. Allele:

```
TCCCTCATTCAGGTTGATGTCCTAACCCCCAGAACCTCAGAATGGGATTGT
AGGGAGTAAGTCCAACTACAGGATTGGGGGTCTTGGAGTCTTACCCTAACA
```

Minor Allele

GGAGTAAGTCCAACTACAGG
TCCCTCATTCAGGTTGATGTCCTAAGCCCCAGAACCTCAGAATGGGATTGT
AGGGAGTAAGTCCAACTACAGG**ATTC**GGGGTCTTGGAGTCTTACCCTAACA
PAM

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs112435590	TCCATGCCAAGAAGGCAACAGAGAG[G/T]GCCAGGGAGACTGAAGTCATACCCT	G	T

SpCas9_WT (PAM Motif: NRG)
Ref. Allele:
PAM
TCCATGCCAAGAAGGCAACAGAG**AGG**GCCAGGGAGACTGAAGTCATACCCT
AGGTACGGTTCTTCCGTTGTCTCTCCCGGTCCCTCTGACTTCAGTATGGGA
ATGCCAAGAAGGCAACAGAG

Minor Allele
TCCATGCCAAGAAGGCAACAGAGAGTGCCAGGGAGACTGAAGTCATACCCT
AGGTACGGTTCTTCCGTTGTCTCTCACGGTCCCTCTGACTTCAGTATGGGA

SaCas9(NNGRRT/NNGRR)
Ref. Allele:
PAM
TCCATGCCAAGAAGGCAACAGAGAGGCCAGGGAGACTGAAGTCATACCCT
AGGTACGGTTCTTCCGTTGTCTCTCCCGGTCCCTCTGACTTCAGTATGGGA
TCCATGCCAAGAAGGCAACA
Minor Allele
PAM
TCCATGCCAAGAAGGCAACAGAGAGTGCCAGGGAGACTGAAGTCATACCCT
AGGTACGGTTCTTCCGTTGTCTCTCACGGTCCCTCTGACTTCAGTATGGGA
TCCATGCCAAGAAGGCAACA

==========

# Variant ID	Sequence Variation		Ref.Allele	Min. allele
rs111670395	CCCAGGTTCAAGCAATTCTGCCTCA[A/G]CCTCCGGAATAGCTGGGACTACAGG	G	A	

SpCas9_WT (PAM Motif: NRG)
Ref. Allele:
PAM

```
CCCAGGTTCAAGCAATTCTGCCTCAGCCTCCGGAATAGCTGGGACTACAGG
GGGTCCAAGTTCGTTAAGACGGAGTCGGAGGCCTTATCGACCCTGATGTCC
    GGTTCAAGCAATTCTGCCT

Minor Allele
CCCAGGTTCAAGCAATTCTGCCTCAACCTCCGGAATAGCTGGGACTACAGG
GGGTCCAAGTTCGTTAAGACGGAGTTGGAGGCCTTATCGACCCTGATGTCC
```

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs9996199	TGTGGCCTGGCTAAAGTAGGCTTTA[C/G]TGGGCTCCTCTCTGCCTGCATCACC	C	G

SpCas9_WT (PAM Motif: NRG)

```
Ref. Allele:
                           CCGAGGAGAGACGGACG
TGTGGCCTGGCTAAAGTAGGCTTTACTGGGCTCCTCTCTGCCTGCATCACC
ACACCGGACCGATTTCATCCGAAATGACCCGAGGAGAGACGGACGTAGTGG
                         PAM
Minor Allele
                       PAM
TGTGGCCTGGCTAAAGTAGGCTTTAGTGGGCTCCTCTCTGCCTGCATCACC
ACACCGGACCGATTTCATCCGAAATCACCCGAGGAGAGACGGACGTAGTGG
   GGCCTGGCTAAAGTAGGCTT
```

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs28431418	CTCCCCGCAGGGCTGTCCGGGTGAG[C/T]ATGGCTCTGGCCACGGGCCAGTGTG T	C	

SaCas9(NNGRRT/NNGRR)

```
Ref. Allele:
                     PAM
CTCCCCGCAGGGCTGTCCGGGTGAGTATGGCTCTGGCCACGGGCCAGTGTG
GAGGGGCGTCCCGACAGGCCCACTCATACCGAGACCGGTGCCCGGTCACAC
 TCCCCGCAGGGCTGTCCGG

Minor Allele         PAM
CTCCCCGCAGGGCTGTCCGGGTGAGCATGGCTCTGGCCACGGGCCAGTGTG
```

GAGGGGCGTCCCGACAGGCCCACTCGTACCGAGACCGGTGCCCGGTCACAC
TCCCCGCAGGGCTGTCCGG

==========

 # Variant ID **Sequence Variation** **Ref.Allele** **Min. allele**

rs2857935 GGGCGCAGGCCCATGCGGAAAGGAT[A/C/G]CCCCGCCGACGCCTGGAGCGGGGCG G C/T (Rev Strand)

SpCas9_WT (PAM Motif: NRG)

Ref. Allele:

GGCGGCTGCGGACCTCG

GGGCGCAGGCCCATGCGGAAAGGATCCCCCGCCGACGCCTGGAGCGGGGCG

CCCGCGTCCGGGTACGCCTTTCCTAGGGGGCGGCTGCGGACCTCGCCCCGC

PAM

Minor Allele

GGGCGCAGGCCCATGCGGAAAGGATGCCCCGCCGACGCCTGGAGCGGGGCG

CCCGCGTCCGGGTACGCCTTTCCTACGGGGCGGCTGCGGACCTCGCCCCGC

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:

GGGGCGGCTGCGGACCTCG

GGGCGCAGGCCCATGCGGAAAGGATCCCCCGCCGACGCCTGGAGCGGGGCG

CCCGCGTCCGGGTACGCCTTTCC**TAG**GGGGCGGCTGCGGACCTCGCCCCGC

PAM

Minor Allele

GGGCGCAGGCCCATGCGGAAAGGATGCCCCGCCGACGCCTGGAGCGGGGCG

CCCGCGTCCGGGTACGCCTTTCCTACGGGGCGGCTGCGGACCTCGCCCCGC

SaCas9(NNGRRT/NNGRR)

Ref. Allele: GGCTGCGGACCTCGCCCCG

GGGCGCAGGCCCATGCGGAAAGGATCCCCCGCCGACGCCTGGAGCGGGGCG

CCCGCGTCCGGGTACGCCTTTCCTAGGGGGCGGCTGCGGACCTCGCCCCGC

PAM

Minor Allele GGCTGCGGACCTCGCCCCG

GGGCGCAGGCCCATGCGGAAAGGATACCCCGCCGACGCCTGGAGCGGGGCG

CCCGCGTCCGGGTACGCCTTTCCTATGGGGCGGCTGCGGACCTCGCCCCGC

PAM

LbCpf1(TTTN/TTCN/CTTN/TCTN/ATTN/TCCN/…):
Ref. Allele:
PAM
GGGCGCAGGCCCATGCGGAAAGGATCCCCCGCCGACGCCTGGAGCGGGGCG
CCCGCGTCCGGGTACGCCTTTCCTAGGGGGCGGCTGCGGACCTCGCCCCGC
CCGCCGACGCCTGGAGCGGG
Minor Allele
GGGCGCAGGCCCATGCGGAAAGGATACCCCGCCGACGCCTGGAGCGGGGCG
CCCGCGTCCGGGTACGCCTTTCCTATGGGGCGGCTGCGGACCTCGCCCCGC

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs149624523	CCCCGCCCCGGCCTCGCCACGCCCC[C/T]ACCTCACCACGCCCCCCGCATCGCC	T	C

SaCas9(NNGRRT/NNGRR)
Ref. Allele:
CCCCGCCCCGGCCTCGCCACGCCCCTACCTCACCACGCCCCCCGCATCGCC
GGGGCGGGGCCGGAGCGGTGCGGGGATGGAGTGGTGCGGGGGGCGTAGCGG

Minor Allele
GAGTGGTGCGGGGGGCGTAG
CCCCGCCCCGGCCTCGCCACGCCCCCACCTCACCACGCCCCCCGCATCGCC
GGGGCGGGGCCGGAGCGGTGCGGGGGTGGAGTGGTGCGGGGGGCGTAGCGG
PAM

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs13122415	attacagtctcaccacgccccgtcc[C/G]CTCTCCGTTGAGCCCCGCGCCTTCG	C	G

SpCas9_WT (PAM Motif: NRG)
Ref. Allele:
GAGGCAACTCGGGGCGCGG
ATTACAGTCTCACCACGCCCCGTCCCCTCTCCGTTGAGCCCCGCGCCTTCG

TAATGTCAGAGTGGTGCGGGGCAGGGGAGAGGCAACTCGGGGCGCGGAAGC
PAM

Minor Allele
ATTACAGTCTCACCACGCCCCGTCCGCTCTCCGTTGAGCCCCGCGCCTTCG
TAATGTCAGAGTGGTGCGGGGCAGGCGAGAGGCAACTCGGGGCGCGGAAGC

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:
ATTACAGTCTCACCACGCCCCGTCCCCTCTCCGTTGAGCCCCGCGCCTTCG
TAATGTCAGAGTGGTGCGGGGCAGGGGAGAGGCAACTCGGGGCGCGGAAGC

Minor Allele
AGAGGCAACTCGGGGCG
ATTACAGTCTCACCACGCCCCGTCCGCTCTCCGTTGAGCCCCGCGCCTTCG
TAATGTCAGAGTGGTGCGGGGCAGGCGAGAGGCAACTCGGGGCGCGGAAGC
PAM

SpCas9_VRER (PAM Motif: NGCG)

Ref. Allele:
ATTACAGTCTCACCACGCCCCGTCCCCTCTCCGTTGAGCCCCGCGCCTTCG
TAATGTCAGAGTGGTGCGGGGCAGGGGAGAGGCAACTCGGGGCGCGGAAGC

Minor Allele
AGAGGCAACTCGGGGCG
ATTACAGTCTCACCACGCCCCGTCCGCTCTCCGTTGAGCCCCGCGCCTTCG
TAATGTCAGAGTGGTGCGGGGCAGGCGAGAGGCAACTCGGGGCGCGGAAGC
PAM

SaCas9(NNGRRT/NNGRR)

Ref. Allele:
GAGGCAACTCGGGGCG
ATTACAGTCTCACCACGCCCCGTCCCCTCTCCGTTGAGCCCCGCGCCTTCG
TAATGTCAGAGTGGTGCGGGGCAGGGGAGAGGCAACTCGGGGCGCGGAAGC
PAM

Minor Allele
ATTACAGTCTCACCACGCCCCGTCCGCTCTCCGTTGAGCCCCGCGCCTTCG
TAATGTCAGAGTGGTGCGGGGCAGGCGAGAGGCAACTCGGGGCGCGGAAGC

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs113331544	GGCCTTGCTGTGTGAGGCAGAACCT[-/GCGGGG]GCGGGGGCAGGGGCGGGCTGGTTCC	-	insGCGGGG

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:
PAM
GGCCTTGCTGTGTGAGGCAGAACCTGCGGGGGCAGGGGCGGGCTGGTTCC
CCGGAACGACACACTCCGTCTTGGACGCCCCCGTCCCCGCCCGACCAAGG
GCTGTGTGAGGCAGAACCT
Minor Allele
PAM
GGCCTTGCTGTGTGAGGCAGAACCTGCGGGGGCGGGGGCAGGGGCGGGCTGGTTCC
CCGGAACGACACACTCCGTCTTGGACGCCCCCGCCCCCGTCCCCGCCCGACCAAGG
GTGAGGCAGAACCTGCGGGG

SpCas9_VRER (PAM Motif: NGCG)
Ref. Allele:
PAM
GGCCTTGCTGTGTGAGGCAGAACCTGCGGGGGCAGGGGCGGGCTGGTTCC
CCGGAACGACACACTCCGTCTTGGACGCCCCCGTCCCCGCCCGACCAAGG
GCTGTGTGAGGCAGAACCT
Minor Allele
PAM
GGCCTTGCTGTGTGAGGCAGAACCTGCGGGGGCGGGGGCAGGGGCGGGCTGGTTCC
CCGGAACGACACACTCCGTCTTGGACGCCCCCGCCCCCGTCCCCGCCCGACCAAGG
GTGAGGCAGAACCTGCGGGG

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs13132932	TGTGTGAGGCAGAACCTGCGGGGGC[A/G]GGGGCGGGCTGGTTCCCTGGCCAGC	A	G

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:
TGTGTGAGGCAGAACCTGCGGGGGCAGGGGCGGGCTGGTTCCCTGGCCAGC
ACACACTCCGTCTTGGACGCCCCCGTCCCCGCCCGACCAAGGGACCGGTCG

Minor Allele
PAM
TGTGTGAGGCAGAACCTGCGGGGGCGGGGGCGGGCTGGTTCCCTGGCCAGC
ACACACTCCGTCTTGGACGCCCCCGCCCCCGCCCGACCAAGGGACCGGTCG
GAGGCAGAACCTGCGGGG

SpCas9_VRER (PAM Motif: NGCG)
Ref. Allele:
TGTGTGAGGCAGAACCTGCGGGGGCAGGGGCGGGCTGGTTCCCTGGCCAGC
ACACACTCCGTCTTGGACGCCCCCGTCCCCGCCCGACCAAGGGACCGGTCG

Minor Allele
PAM
TGTGTGAGGCAGAACCTGCGGGGGCGGGGGCGGGCTGGTTCCCTGGCCAGC
ACACACTCCGTCTTGGACGCCCCCGCCCCCGCCCGACCAAGGGACCGGTCG
GAGGCAGAACCTGCGGGG

SaCas9(NNGRRT/NNGRR)
Ref. Allele:
TGTGTGAGGCAGAACCTGCGGGGGCAGGGGCGGGCTGGTTCCCTGGCCAGC
ACACACTCCGTCTTGGACGCCCCCGTCCCCGCCCGACCAAGGGACCGGTCG

Minor Allele **PAM**
TGTGTGAGGCAGAACCTGCGGGGGCGGGGGCGGGCTGGTTCCCTGGCCAGC
ACACACTCCGTCTTGGACGCCCCCGCCCCCGCCCGACCAAGGGACCGGTCG
GTGAGGCAGAACCTGCGGGG

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs13102260	AGCGTCTGGGACGCAAGGCGCCGTG[A/G]GGGCTGCCGGGACGGGTCCAAGATG	G	A

SpCas9_WT (PAM Motif: NRG)
Ref. Allele:
PAM

AGCGTCTGGGACGCAAGGCGCCGTGGGGGCTGCCGGGACGGGTCCAAGATG
TCGCAGACCCTGCGTTCCGCGGCACCCCCGACGGCCCTGCCCAGGTTCTAC
   GTCTGGGACGCAAGGCGCCG

Minor Allele
AGCGTCTGGGACGCAAGGCGCCGTGAGGGCTGCCGGGACGGGTCCAAGATG
TCGCAGACCCTGCGTTCCGCGGCACTCCCGACGGCCCTGCCCAGGTTCTAC

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:
AGCGTCTGGGACGCAAGGCGCCGTGGGGGCTGCCGGGACGGGTCCAAGATG
TCGCAGACCCTGCGTTCCGCGGCACCCCCGACGGCCCTGCCCAGGTTCTAC

Minor Allele
PAM
AGCGTCTGGGACGCAAGGCGCCGTGAGGGCTGCCGGGACGGGTCCAAGATG
TCGCAGACCCTGCGTTCCGCGGCACTCCCGACGGCCCTGCCCAGGTTCTAC
   GTCTGGGACGCAAGGCGCCG

SaCas9(NNGRRT/NNGRR)
Ref. Allele:
PAM
AGCGTCTGGGACGCAAGGCGCCGTGGGGGCTGCCGGGACGGGTCCAAGATG
TCGCAGACCCTGCGTTCCGCGGCACCCCCGACGGCCCTGCCCAGGTTCTAC
 GCGTCTGGGACGCAAGGCGCCG

Minor Allele
AGCGTCTGGGACGCAAGGCGCCGTGAGGGCTGCCGGGACGGGTCCAAGATG
TCGCAGACCCTGCGTTCCGCGGCACTCCCGACGGCCCTGCCCAGGTTCTAC

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs10009935	CTCACTTGGGTCTTCCCTTGTCCTC[C/T]CGCGAGGGGAGGCAGAGCCTTGTTG	T	C

SpCas9_WT (PAM Motif: NRG)

Ref. Allele:

```
CTCACTTGGGTCTTCCCTTGTCCTCTCGCGAGGGGAGGCAGAGCCTTGTTG
GAGTGAACCCAGAAGGGAACAGGAGAGCGCTCCCCTCCGTCTCGGAACAAC


Minor Allele
                              GCTCCCCTCCGTCTCGG
CTCACTTGGGTCTTCCCTTGTCCTCCCGCGAGGGGAGGCAGAGCCTTGTTG
GAGTGAACCCAGAAGGGAACAGGAGGGCGCTCCCCTCCGTCTCGGAACAAC
                            PAM

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:
                               CGCTCCCCTCCGTCTCGG
CTCACTTGGGTCTTCCCTTGTCCTCTCGCGAGGGGAGGCAGAGCCTTGTTG
GAGTGAACCCAGAAGGGAACAGGAGAGCGCTCCCCTCCGTCTCGGAACAAC
                            PAM
Minor Allele
CTCACTTGGGTCTTCCCTTGTCCTCCCGCGAGGGGAGGCAGAGCCTTGTTG
GAGTGAACCCAGAAGGGAACAGGAGGGCGCTCCCCTCCGTCTCGGAACAAC

SaCas9(NNGRRT/NNGRR)
Ref. Allele:
CTCACTTGGGTCTTCCCTTGTCCTCTCGCGAGGGGAGGCAGAGCCTTGTTG
GAGTGAACCCAGAAGGGAACAGGAGAGCGCTCCCCTCCGTCTCGGAACAAC

Minor Allele
                              GCTCCCCTCCGTCTCGGAA
CTCACTTGGGTCTTCCCTTGTCCTCCCGCGAGGGGAGGCAGAGCCTTGTTG
GAGTGAACCCAGAAGGGAACAGGAGGGGCGCTCCCCTCCGTCTCGGAACAAC
                          PAM
```

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs58870770	TGAATGAGTTGTGGTTGCCAAGTAA[-/A]GTGGTGAACTTACGTGGTGATTAAT	A	delA

```
LbCpf1(TTTN/TTCN/CTTN/TCTN/ATTN/TCCN/…):
Ref. Allele:
CTTACTCAACACCAACGGTTC
```

```
TGAATGAGTTGTGGTTGCCAAGTAAAGTGGTGAACTTACGTGGTGATTAAT
ACTTACTCAACACCAACGGTTCATTTCACCACTTGAATGCACCACTAATTA
                      PAM


Minor Allele
TGAATGAGTTGTGGTTGCCAAGTAAGTGGTGAACTTACGTGGTGATTAAT
ACTTACTCAACACCAACGGTTCATTCACCACTTGAATGCACCACTAATTA
```

AsCpf1(TTTN)

```
Ref. Allele:
 CTTACTCAACACCAACGGTTC
TGAATGAGTTGTGGTTGCCAAGTAAAGTGGTGAACTTACGTGGTGATTAAT
ACTTACTCAACACCAACGGTTCATTTCACCACTTGAATGCACCACTAATTA
                      PAM
Minor Allele
TGAATGAGTTGTGGTTGCCAAGTAAGTGGTGAACTTACGTGGTGATTAAT
ACTTACTCAACACCAACGGTTCATTCACCACTTGAATGCACCACTAATTA
```

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs28571971	ACAGTAGGAGTTAGGAAGTACTCTG[C/G]TGCAGTTCAGGCCTTTCTCTTACCT	C	G

SpCas9_WT (PAM Motif: NRG)

```
Ref. Allele:           PAM
ACAGTAGGAGTTAGGAAGTACTCTGGTGCAGTTCAGGCCTTTCTCTTACCT
TGTCATCCTCAATCCTTCATGAGACCACGTCAAGTCCGGAAAGAGAATGGA
      GGAGTTAGGAAGTACTC
Minor Allele:
                           GTCAAGTCCGGAAAGAGAATGG
ACAGTAGGAGTTAGGAAGTACTCTGCTGCAGTTCAGGCCTTTCTCTTACCT
TGTCATCCTCAATCCTTCATGAGACGACGTCAAGTCCGGAAAGAGAATGGA
                         PAM
```

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs28583447	CAGTAGGAGTTAGGAAGTACTCTGG[C/T]GCAGTTCAGGCCTTTCTCTTACCTC	T	C

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:
CAGTAGGAGTTAGGAAGTACTCTGGTGCAGTTCAGGCCTTTCTCTTACCTC
GTCATCCTCAATCCTTCATGAGACCACGTCAAGTCCGGAAAGAGAATGGAG

Minor Allele: TCAAGTCCGGAAAGAGAATG
CAGTAGGAGTTAGGAAGTACTCTGGCGCAGTTCAGGCCTTTCTCTTACCTC
GTCATCCTCAATCCTTCATGAGACCGCGTCAAGTCCGGAAAGAGAATGGAG
PAM

SpCas9_VRER (PAM Motif: NGCG)
Ref. Allele:
CAGTAGGAGTTAGGAAGTACTCTGGTGCAGTTCAGGCCTTTCTCTTACCTC
GTCATCCTCAATCCTTCATGAGACCACGTCAAGTCCGGAAAGAGAATGGAG

Minor Allele: **PAM**
CAGTAGGAGTTAGGAAGTACTCTGGCGCAGTTCAGGCCTTTCTCTTACCTC
GTCATCCTCAATCCTTCATGAGACCGCGTCAAGTCCGGAAAGAGAATGGAG
GGAGTTAGGAAGTACTCTG

Minor Allele: TCAAGTCCGGAAAGAGAATG
CAGTAGGAGTTAGGAAGTACTCTGGCGCAGTTCAGGCCTTTCTCTTACCTC
GTCATCCTCAATCCTTCATGAGACCGCGTCAAGTCCGGAAAGAGAATGGAG
PAM

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs28468636	TCTCTTACCTCTCAGTATTCTATTT[C/G]CGATCTGGATGTGTCCCAGATGGCA	C	G

SpCas9_WT (PAM Motif: NRG)
Ref. Allele: TAGACCTACACAGGGTCTA
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT
PAM

Minor Allele
TCTCTTACCTCTCAGTATTCTATTTGCGATCTGGATGTGTCCCAGATGGCA

AGAGAATGGAGAGTCATAAGATAAACGCTAGACCTACACAGGGTCTACCGT

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT

Minor Allele **PAM**
TCTCTTACCTCTCAGTATTCTATTTGCGATCTGGATGTGTCCCAGATGGCA
AGAGAATGGAGAGTCATAAGATAAACGCTAGACCTACACAGGGTCTACCGT
ACCTCTCAGTATTCTATTT

SpCas9_VRER (PAM Motif: NGCG)
Ref. Allele:
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT

Minor Allele **PAM**
TCTCTTACCTCTCAGTATTCTATTTGCGATCTGGATGTGTCCCAGATGGCA
AGAGAATGGAGAGTCATAAGATAAACGCTAGACCTACACAGGGTCTACCGT
ACCTCTCAGTATTCTATTT

SaCas9(NNGRRT/NNGRR):
Ref. Allele: TAGACCTACACAGGGTCTA
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT
PAM

Minor Allele
TCTCTTACCTCTCAGTATTCTATTTGCGATCTGGATGTGTCCCAGATGGCA
AGAGAATGGAGAGTCATAAGATAAACGCTAGACCTACACAGGGTCTACCGT

LbCpf1(TTTN/TTCN/CTTN/TCTN/ATTN/TCCN/…):

Ref. Allele: **PAM**
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT
ATCTGGATGTGTCCCAGATG
Minor Allele

TCTCTTACCTCTCAGTATTCTATTTGCGATCTGGATGTGTCCCAGATGGCA
AGAGAATGGAGAGTCATAAGATAAACGCTAGACCTACACAGGGTCTACCGT

==========

 # Variant ID Sequence Variation Ref.Allele Min. allele
rs28564368 TATTCTATTTCCGATCTGGATGTGT[A/C]CCAGATGGCATTTGGTAAGAATATC C A

SpCas9_WT (PAM Motif: NRG)
Ref. Allele: TCTACCGTAAACCATTCTTA
TATTCTATTTCCGATCTGGATGTGTCCCAGATGGCATTTGGTAAGAATATC
ATAAGATAAAGGCTAGACCTACACAGGGTCTACCGTAAACCATTCTTATAG
PAM
Minor Allele
TATTCTATTTCCGATCTGGATGTGTACCAGATGGCATTTGGTAAGAATATC
ATAAGATAAAGGCTAGACCTACACATGGTCTACCGTAAACCATTCTTATAG

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele: GTCTACCGTAAACCATTCTT
TATTCTATTTCCGATCTGGATGTGTCCCAGATGGCATTTGGTAAGAATATC
ATAAGATAAAGGCTAGACCTACACAGGGTCTACCGTAAACCATTCTTATAG
PAM
Minor Allele
TATTCTATTTCCGATCTGGATGTGTACCAGATGGCATTTGGTAAGAATATC
ATAAGATAAAGGCTAGACCTACACATGGTCTACCGTAAACCATTCTTATAG

Ref. Allele: TCTACCGTAAACCATTCTTA
TATTCTATTTCCGATCTGGATGTGTCCCAGATGGCATTTGGTAAGAATATC
ATAAGATAAAGGCTAGACCTACACAGGGTCTACCGTAAACCATTCTTATAG
PAM
Minor Allele
TATTCTATTTCCGATCTGGATGTGTACCAGATGGCATTTGGTAAGAATATC
ATAAGATAAAGGCTAGACCTACACATGGTCTACCGTAAACCATTCTTATAG

LbCpf1(TTTN/TTCN/CTTN/TCTN/ATTN/TCCN/…):
Ref. Allele: **PAM**
TATTCTATTTCCGATCTGGATGTGTCCCAGATGGCATTTGGTAAGAATATC
ATAAGATAAAGGCTAGACCTACACAGGGTCTACCGTAAACCATTCTTATAG

```
                            AGATGGCATTTGGTAAGAATAT
Minor Allele
TATTCTATTTCCGATCTGGATGTGTACCAGATGGCATTTGGTAAGAATATC
ATAAGATAAAGGCTAGACCTACACATGGTCTACCGTAAACCATTCTTATAG
```

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs77173925	GAATAAATTATTCTAAAGGATGGAA[A/G]AACTTTTTGGATATTTGGAGAAATT	A	G

SpCas9_WT (PAM Motif: NRG)

```
Ref. Allele:
GAATAAATTATTCTAAAGGATGGAAAAACTTTTTGGATATTTGGAGAAATT
CTTATTTAATAAGATTTCCTACCTTTTTGAAAAACCTATAAACCTCTTTAA

Minor Allele           PAM
GAATAAATTATTCTAAAGGATGGAAGAACTTTTTGGATATTTGGAGAAATT
CTTATTTAATAAGATTTCCTACCTTCTTGAAAAACCTATAAACCTCTTTAA
    AAATTATTCTAAAGGATGG
```

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

```
Ref. Allele:
GAATAAATTATTCTAAAGGATGGAAAAACTTTTTGGATATTTGGAGAAATT
CTTATTTAATAAGATTTCCTACCTTTTTGAAAAACCTATAAACCTCTTTAA

Minor Allele             PAM
GAATAAATTATTCTAAAGGATGGAAGAACTTTTTGGATATTTGGAGAAATT
CTTATTTAATAAGATTTCCTACCTTCTTGAAAAACCTATAAACCTCTTTAA
      ATTATTCTAAAGGATGGAA
```

SaCas9(NNGRRT/NNGRR):

```
Ref. Allele:
GAATAAATTATTCTAAAGGATGGAAAAACTTTTTGGATATTTGGAGAAATT
CTTATTTAATAAGATTTCCTACCTTTTTGAAAAACCTATAAACCTCTTTAA

Minor Allele            PAM
GAATAAATTATTCTAAAGGATGGAAGAACTTTTTGGATATTTGGAGAAATT
CTTATTTAATAAGATTTCCTACCTTCTTGAAAAACCTATAAACCTCTTTAA
     AATTATTCTAAAGGATGG
```

AsCpf1(TTTN)
Ref. Allele:
ATTTAATAAGATTTCCTACC
GAATAAATTATTCTAAAGGATGGAAAAACTTTTTGGATATTTGGAGAAATT
CTTATTTAATAAGATTTCCTACCTTTTTGAAAAACCTATAAACCTCTTTAA
PAM

Minor Allele
GAATAAATTATTCTAAAGGATGGAAGAACTTTTTGGATATTTGGAGAAATT
CTTATTTAATAAGATTTCCTACCTTCTTGAAAAACCTATAAACCTCTTTAA

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs3905238	TTGTATCATGTCAATGTATTACTTA[C/T]GCAAAAATAATACATTAAAAAAAAT	A	G (Rev Strand)

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:
TTGTATCATGTCAATGTATTACTTATGCAAAAATAATACATTAAAAAAAAT
AACATAGTACAGTTACATAATGAATACGTTTTTATTATGTAATTTTTTTTA

Minor Allele TTTTATTATGTAATTTTTT
TTGTATCATGTCAATGTATTACTTACGCAAAAATAATACATTAAAAAAAAT
AACATAGTACAGTTACATAATGAATGCGTTTTTATTATGTAATTTTTTTTA
PAM

SpCas9_VRER (PAM Motif: NGCG)
Ref. Allele:
TTGTATCATGTCAATGTATTACTTATGCAAAAATAATACATTAAAAAAAAT
AACATAGTACAGTTACATAATGAATACGTTTTTATTATGTAATTTTTTTTA

Minor Allele TTTTATTATGTAATTTTTT
TTGTATCATGTCAATGTATTACTTACGCAAAAATAATACATTAAAAAAAAT
AACATAGTACAGTTACATAATGAATGCGTTTTTATTATGTAATTTTTTTTA
PAM

==========

 # Variant ID Sequence Variation Ref.Allele Min. allele

rs33950430 AGGTATTCACTAATTTTGAGTAACA[-/AACA]CTGCTCACAAAGTTTGGATTTTGGC AACA delAACA

LbCpf1(TTTN/TTCN/CTTN/TCTN/ATTN/TCCN/…):

Ref. Allele:

ATAAGTGATTAAAACTCATT

AGGTATTCACTAATTTTGAGTAACAAACACTGCTCACAAAGTTTGGATTTTGGC

TCCATAAGTGATTAAAACTCATTGTTTGTGACGAGTGTTTCAAACCTAAAACCG

PAM

Minor Allele

AGGTATTCACTAATTTTGAGTAACACTGCTCACAAAGTTTGGATTTTGGC

TCCATAAGTGATTAAAACTCATTGTGACGAGTGTTTCAAACCTAAAACCG

AsCpf1(TTTN)

Ref. Allele:

ATAAGTGATTAAAACTCATT

AGGTATTCACTAATTTTGAGTAACAAACACTGCTCACAAAGTTTGGATTTTGGC

TCCATAAGTGATTAAAACTCATTGTTTGTGACGAGTGTTTCAAACCTAAAACCG

PAM

Minor Allele

AGGTATTCACTAATTTTGAGTAACACTGCTCACAAAGTTTGGATTTTGGC

TCCATAAGTGATTAAAACTCATTGTGACGAGTGTTTCAAACCTAAAACCG

==========

 # Variant ID Sequence Variation Ref.Allele Min. allele

rs28377140 AGGCAATTAATACTTGCTTCTGGCA[C/G]TTTCTTATTCTCCTTCAGATTCCTA G C

SpCas9_WT (PAM Motif: NRG)

Ref. Allele: **PAM**

AGGCAATTAATACTTGCTTCTGGCAGTTTCTTATTCTCCTTCAGATTCCTA

TCCGTTAATTATGAACGAAGACCGTCAAAGAATAAGAGGAAGTCTAAGGAT

ATTAATACTTGCTTCTGG

Minor Allele AGAATAAGAGGAAGTCTAAGG

```
AGGCAATTAATACTTGCTTCTGGCACTTTCTTATTCTCCTTCAGATTCCTA
TCCGTTAATTATGAACGAAGACCGTGAAAGAATAAGAGGAAGTCTAAGGAT
                         PAM
================================================================================================================

# Variant ID      Sequence Variation                                                  Ref.Allele        Min. allele
rs3856973   ttaaaaataaaaataaGTTAACACT[C/T]GATTAACCCTGACATTTCCCTATCC     G     A (Rev strand)

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:              CTAATTGGGACTGTAAAGGG
TTAAAAATAAAAATAAGTTAACACTCGATTAACCCTGACATTTCCCTATCC
AATTTTTATTTTTATTCAATTGTGAGCTAATTGGGACTGTAAAGGGATAGG
                       PAM
Minor Allele
TTAAAAATAAAAATAAGTTAACACTTGATTAACCCTGACATTTCCCTATCC
AATTTTTATTTTTATTCAATTGTGAACTAATTGGGACTGTAAAGGGATAGG

SaCas9(NNGRRT/NNGRR):
Ref. Allele:                AATTGGGACTGTAAAGGG
TTAAAAATAAAAATAAGTTAACACTCGATTAACCCTGACATTTCCCTATCC
AATTTTTATTTTTATTCAATTGTGAGCTAATTGGGACTGTAAAGGGATAGG
                       PAM
Minor Allele
TTAAAAATAAAAATAAGTTAACACTTGATTAACCCTGACATTTCCCTATCC
AATTTTTATTTTTATTCAATTGTGAACTAATTGGGACTGTAAAGGGATAGG

LbCpf1(TTTN/TTCN/CTTN/TCTN/ATTN/TCCN/…):
Ref. Allele:
TTAAAAATAAAAATAAGTTAACACTCGATTAACCCTGACATTTCCCTATCC
AATTTTTATTTTTATTCAATTGTGAGCTAATTGGGACTGTAAAGGGATAGG

Minor Allele           PAM
TTAAAAATAAAAATAAGTTAACACTTGATTAACCCTGACATTTCCCTATCC
AATTTTTATTTTTATTCAATTGTGAACTAATTGGGACTGTAAAGGGATAGG
                         ATTAACCCTGACATTTCCCTA
================================================================================================================

# Variant ID      Sequence Variation                                                  Ref.Allele        Min. allele
```

rs4498089 AGTGTTAACTTATTTTTATTTTTAA[A/G]AAAATTGTTAAGGGCTTTCCAGCAA A G

SpCas9_WT (PAM Motif: NRG)
Ref. Allele:
AGTGTTAACTTATTTTTATTTTTAAAAAAATTGTTAAGGGCTTTCCAGCAA
TCACAATTGAATAAAAATAAAAATTTTTTTAACAATTCCCGAAAGGTCGTT

Minor Allele **PAM**
AGTGTTAACTTATTTTTATTTTTAAGAAAATTGTTAAGGGCTTTCCAGCAA
TCACAATTGAATAAAAATAAAAATTCTTTTAACAATTCCCGAAAGGTCGTT
   GTTAACTTATTTTTATTTTT

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:
AGTGTTAACTTATTTTTATTTTTAAAAAAATTGTTAAGGGCTTTCCAGCAA
TCACAATTGAATAAAAATAAAAATTTTTTTAACAATTCCCGAAAGGTCGTT

Minor Allele **PAM**
AGTGTTAACTTATTTTTATTTTTAAGAAAATTGTTAAGGGCTTTCCAGCAA
TCACAATTGAATAAAAATAAAAATTCTTTTAACAATTCCCGAAAGGTCGTT
   GTTAACTTATTTTTATTTTTA

SaCas9(NNGRRT/NNGRR):
Ref. Allele:
AGTGTTAACTTATTTTTATTTTTAAAAAAATTGTTAAGGGCTTTCCAGCAA
TCACAATTGAATAAAAATAAAAATTTTTTTAACAATTCCCGAAAGGTCGTT

Minor Allele **PAM**
AGTGTTAACTTATTTTTATTTTTAAGAAAATTGTTAAGGGCTTTCCAGCAA
TCACAATTGAATAAAAATAAAAATTCTTTTAACAATTCCCGAAAGGTCGTT
   GTTAACTTATTTTTATTTTT
AsCpf1(TTTN)
Ref. Allele:
 CACAATTGAATAAAAATAAAA
AGTGTTAACTTATTTTTATTTTTAAAAAAATTGTTAAGGGCTTTCCAGCAA
TCACAATTGAATAAAAATAAAAATTTTTTTTAACAATTCCCGAAAGGTCGTT

PAM

Minor Allele

AGTGTTAACTTATTTTTATTTTTAAGAAAATTGTTAAGGGCTTTCCAGCAA

TCACAATTGAATAAAAATAAAAATTCTTTTAACAATTCCCGAAAGGTCGTT

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs57666989	CTGCAAGCTCCGCTTCCCGAGTTCA[C/T]GCCATTCTCCTGCCTCAGTCTCCCA	C	T

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele: TAAGAGGACGGAGTCAGAGGG

CTGCAAGCTCCGCTTCCCGAGTTCACGCCATTCTCCTGCCTCAGTCTCCCA

GACGTTCGAGGCGAAGGGCTCAAGTGCGGTAAGAGGACGGAGTCAGAGGGT

PAM

Minor Allele

CTGCAAGCTCCGCTTCCCGAGTTCATGCCATTCTCCTGCCTCAGTCTCCCA

GACGTTCGAGGCGAAGGGCTCAAGTACGGTAAGAGGACGGAGTCAGAGGGT

SpCas9_VRER (PAM Motif: NGCG)

Ref. Allele: TAAGAGGACGGAGTCAGAGG

CTGCAAGCTCCGCTTCCCGAGTTCACGCCATTCTCCTGCCTCAGTCTCCCA

GACGTTCGAGGCGAAGGGCTCAAGTGCGGTAAGAGGACGGAGTCAGAGGGT

PAM

Minor Allele

CTGCAAGCTCCGCTTCCCGAGTTCATGCCATTCTCCTGCCTCAGTCTCCCA

GACGTTCGAGGCGAAGGGCTCAAGTACGGTAAGAGGACGGAGTCAGAGGGT

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs10006129	tcttgatctcctgacctcgtcatcc[C/G]ccgaccttgtgatccgcccacctcg	G	C

SpCas9_WT (PAM Motif: NRG)

Ref. Allele:

TCTTGATCTCCTGACCTCGTCATCCGCCGACCTTGTGATCCGCCCACCTCG

AGAACTAGAGGACTGGAGCAGTAGGCGGCTGGAACACTAGGCGGGTGGAGC

Minor Allele CTGGAACACTAGGCGGGTGG
TCTTGATCTCCTGACCTCGTCATCCCCCGACCTTGTGATCCGCCCACCTCG
AGAACTAGAGGACTGGAGCAGTAGGGGGCTGGAACACTAGGCGGGTGGAGC
PAM

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:
TCTTGATCTCCTGACCTCGTCATCCGCCGACCTTGTGATCCGCCCACCTCG
AGAACTAGAGGACTGGAGCAGTAGGCGGCTGGAACACTAGGCGGGTGGAGC

Minor Allele CTGGAACACTAGGCGGGTGG
TCTTGATCTCCTGACCTCGTCATCCCCCGACCTTGTGATCCGCCCACCTCG
AGAACTAGAGGACTGGAGCAGTAGGGGGCTGGAACACTAGGCGGGTGGAGC
PAM

SpCas9_VRER (PAM Motif: NGCG)
Ref. Allele: CTGGAACACTAGGCGGGTGG
TCTTGATCTCCTGACCTCGTCATCCGCCGACCTTGTGATCCGCCCACCTCG
AGAACTAGAGGACTGGAGCAGTAGGCGGCTGGAACACTAGGCGGGTGGAGC
PAM
Minor Allele
TCTTGATCTCCTGACCTCGTCATCCCCCGACCTTGTGATCCGCCCACCTCG
AGAACTAGAGGACTGGAGCAGTAGGGGGCTGGAACACTAGGCGGGTGGAGC

SaCas9(NNGRRT/NNGRR):
Ref. Allele:
TCTTGATCTCCTGACCTCGTCATCCGCCGACCTTGTGATCCGCCCACCTCG
AGAACTAGAGGACTGGAGCAGTAGGCGGCTGGAACACTAGGCGGGTGGAGC

Minor Allele CTGGAACACTAGGCGGGTGG
TCTTGATCTCCTGACCTCGTCATCCCCCGACCTTGTGATCCGCCCACCTCG
AGAACTAGAGGACTGGAGCAGTAGGGGGCTGGAACACTAGGCGGGTGGAGC
PAM

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs28696693	GGTAATTTTTGTATTTTTAGTAGAG[A/G]TGGGGTTTTGCCATGATGAGCAGGC	A	G

SpCas9 (PAM Motif: NRG)

Ref. Allele:

GGTAATTTTTGTATTTTTAGTAGAGATGGGGTTTTGCCATGATGAGCAGGC
CCATTAAAAACATAAAAATCATCTCTACCCCAAAACGGTACTACTCGTCCG

Minor Allele **PAM**

GGTAATTTTTGTATTTTTAGTAGAGGTGGGGTTTTGCCATGATGAGCAGGC
CCATTAAAAACATAAAAATCATCTCCACCCCAAAACGGTACTACTCGTCCG
ATTTTTGTATTTTTAGTAG

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele: **PAM**

GGTAATTTTTGTATTTTTAGTAGAGATGGGGTTTTGCCATGATGAGCAGGC
CCATTAAAAACATAAAAATCATCTCTACCCCAAAACGGTACTACTCGTCCG
TTTTTGTATTTTTAGTAG

Minor Allele

GGTAATTTTTGTATTTTTAGTAGAGGTGGGGTTTTGCCATGATGAGCAGGC
CCATTAAAAACATAAAAATCATCTCCACCCCAAAACGGTACTACTCGTCCG

==========

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs28393280	GGATTTTGAATGCGGAACCAACTGC[A/G]CTTGTTGAACTCTGCTAAGTATAAC	A	G

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:

GGATTTTGAATGCGGAACCAACTGCACTTGTTGAACTCTGCTAAGTATAAC
CCTAAAACTTACGCCTTGGTTGACGTGAACAACTTGAGACGATTCATATTG

Minor Allele: **PAM**

GGATTTTGAATGCGGAACCAACTGCGCTTGTTGAACTCTGCTAAGTATAAC
CCTAAAACTTACGCCTTGGTTGACGCGAACAACTTGAGACGATTCATATTG

ATTTTGAATGCGGAACCAAC

SpCas9_VRER (PAM Motif: NGCG)

Ref. Allele:

GGATTTTGAATGCGGAACCAACTGCACTTGTTGAACTCTGCTAAGTATAAC
CCTAAAACTTACGCCTTGGTTGACGTGAACAACTTGAGACGATTCATATTG

Minor Allele: **PAM**

GGATTTTGAATGCGGAACCAACTGCGCTTGTTGAACTCTGCTAAGTATAAC
CCTAAAACTTACGCCTTGGTTGACGCGAACAACTTGAGACGATTCATATTG
ATTTTGAATGCGGAACCAAC

Minor Allele: ACAACTTGAGACGATTCAT

GGATTTTGAATGCGGAACCAACTGCGCTTGTTGAACTCTGCTAAGTATAAC
CCTAAAACTTACGCCTTGGTTGACGCGAACAACTTGAGACGATTCATATTG
PAM

Figure S2: List of sgRNA sequences designed targeting prevalent SNP-dependent PAM upstream *HTT* exon-1, and within *HTT* intron-1.

sgRNA sequences targeting SNP-dependent PAM:

Sequence sgHD1

Complementary guide: 5'GCTCCAGGCGTCGGCGG 3' (sgHD1) (n=17 nt)
sgRNA expression cassette: hU6 promoter

Major Allele (C): Have a PAM motif on the Positive strand (NGG)
GTCGCCCCGCTCCAGGCGTCGGCGGGGGATCCTTTCCGCATGGGCCTGCGCC
CAGCGGGGCGAGGTCCGCAGCCGCCCCCTAGGAAAGGCGTACCCGGACGCGG
GCTCCAGGCGTCGGCGG

Minor Allele (G): Disruption of the PAM motive.
GTCGCCCCGCTCCAGGCGTCGGCGGGGCATCCTTTCCGCATGGGCCTGCGCC
CAGCGGGGCGAGGTCCGCAGCCGCCCCGTAGGAAAGGCGTACCCGGACGCGG

Sequence sgHD2

Complementary guide: 5'GGCGCGGGGCTCAACGGAG 3' (sgHD2) (n=19 nt)
sgRNA expression cassette: hU6 promoter

Major Allele (C): Has a PAM motif on the Negative strand (NGG)
GAGGCAACTCGGGGCGCGG
TTACAGTCTCACCACGCCCCGTCCCCTCTCCGTTGAGCCCCGCGCCTTC
AATGTCAGAGTGGTGCGGGGCAGGGGAGAGGCAACTCGGGGCGCGGAAG

Minor Allele (G): Disruption of the PAM motive
TTACAGTCTCACCACGCCCCGTCCGCTCTCCGTTGAGCCCCGCGCCTTC
AATGTCAGAGTGGTGCGGGGCAGGCGAGAGGCAACTCGGGGCGCGGAAG

Sequence sgHD3
Complementary guide: 5'GTCTGGGACGCAAGGCGCCG3' (sgHD3) (n=20 nt)
sgRNA expression cassette: hU6 promoter

Major Allele (G): Has a PAM motif on the Positive strand (NGG) Use U6 promoter
GCGTCTGGGACGCAAGGCGCCGTGGGGGCTGCCGGGACGGGTCCAAGAT
CGCAGACCCTGCGTTCCGCGGCACCCCCGACGGCCCTGCCCAGGTTCTA
GTCTGGGACGCAAGGCGCCG

Minor Allele (A): Loss PAM motif on the Positive strand (NGG) Use U6 promoter
GCGTCTGGGACGCAAGGCGCCGTGAGGGCTGCCGGGACGGGTCCAAGAT
CGCAGACCCTGCGTTCCGCGGCACTCCCGACGGCCCTGCCCAGGTTCTA

Sequence sgHD4
Complementary guide: 5' GATGCACGCGGGGTGGGGC 3' (n=19 nt)
sgRNA expression cassette: hU6 promoter

Major Allele: Has a PAM motif on the negative strand (NGG))
3' CGGGGTGGGGCGCACGTAG 5'
5' TCTGCGTCAGGGTTTCCTTCTTTTCCAGCCCCACCCCGCGTGCATCCCA 3'
3' AGACGCAGTCCCAAAGGAAGAAAAGGTCGGGGTGGGGCGCACGTAGGGT 5'

Minor Allele: SNP produces a Loss of the PAM motif.
5' TCTGCGTCAGGGTTTCCTTCTTTTGCAGCCCCACCCCGCGTGCATCCCA 3'
3' AGACGCAGTCCCAAAGGAAGAAAACGTCGGGGTGGGGCGCACGTAGGGT

Sequence sgHD5g and sgHD5c

Complementary guide: 5'ATTCAGGTTGATGTCCT 3' (sgHD5g) (n=17 nt)
Complementary guide: 5'ATCCCATTCTGAGGTTCTGG 3 (sgHD5c) (n=20 nt)
sgRNA expression cassette: hH1 promoter

Major Allele: Has a PAM motif on the Positive strand (NAG)

```
5'CCCTCATTCAGGTTGATGTCCTAAGCCCCAGAACCTCAGAATGGGATTG 3'
3'  GGGAGTAAGTCCAACTACAGGATTCGGGGTCTTGGAGTCTTACCCTAAC 5'
       ATTCAGGTTGATGTCCT
```

Minor Allele: generates a PAM motif on the Negative strand (NAG)

```
                              GGTCTTGGAGTCTTACCCTA
5'CCCTCATTCAGGTTGATGTCCTAACCCCCAGAACCTCAGAATGGGATTG 3'
3'GGGAGTAAGTCCAACTACAGGATTGGGGGTCTTGGAGTCTTACCCTAAC 5'
```

Sequence sgHD6c and sgHD6g

Complementary guide: 5'GCAGGCAGAGAGGAGCC 3' (sgHD6c) (n=17 nt)
Complementary guide: 5'GCCTGGCTAAAGTAGGCTT 3 (sgHD6g) (n=19 nt)
sgRNA expression cassette: hU6 promoter

Major Allele (C): Have a PAM motif on the negative strand (NAG)

```
                                 CCGAGGAGAGACGGACG
5' GTGGCCTGGCTAAAGTAGGCTTTACTGGGCTCCTCTCTGCCTGCATCAC 3'
3' CACCGGACCGATTTCATCCGAAATGACCCGAGGAGAGACGGACGTAGTG 5'
```

Minor Allele (G): Generates a PAM motif on the Positive strand (NAG)

```
5' GTGGCCTGGCTAAAGTAGGCTTTAGTGGGCTCCTCTCTGCCTGCATCAC 3'
```

3’ CACCGGACCGATTTCATCCGAAATCACCCGAGGAGAGACGGACGTAGTG 5’
GCCTGGCTAAAGTAGGCTT

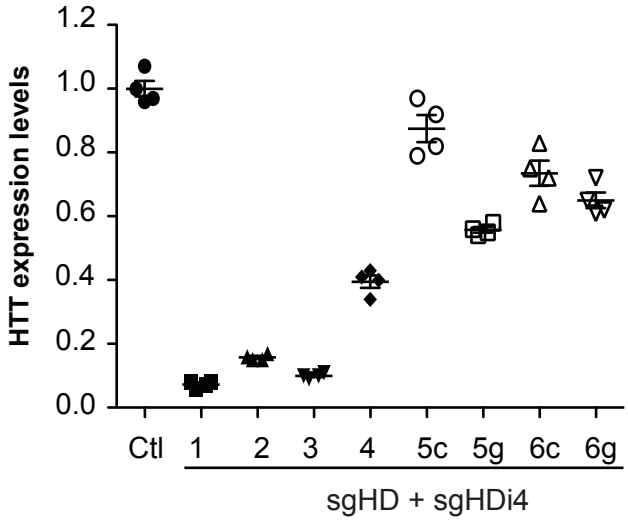
Sequence sgHDi3
Complementary guide: 5’GCTTTTAGGACGCCTCGG 3’ (sgHDi3) (n=18 nt)
sgRNA expression cassette: hU6 promoter

5’ AATGCTTTTAGGACGCCTCGGCGGGAGTGGCGGGGCAGGGGGGGGGCG 3
3’ TTACGAAAATCCTGCGGAGCCGCCCTCACCGCCCCGTCCCCCCCCCGC 5’
GCTTTTAGGACGCCTCGG

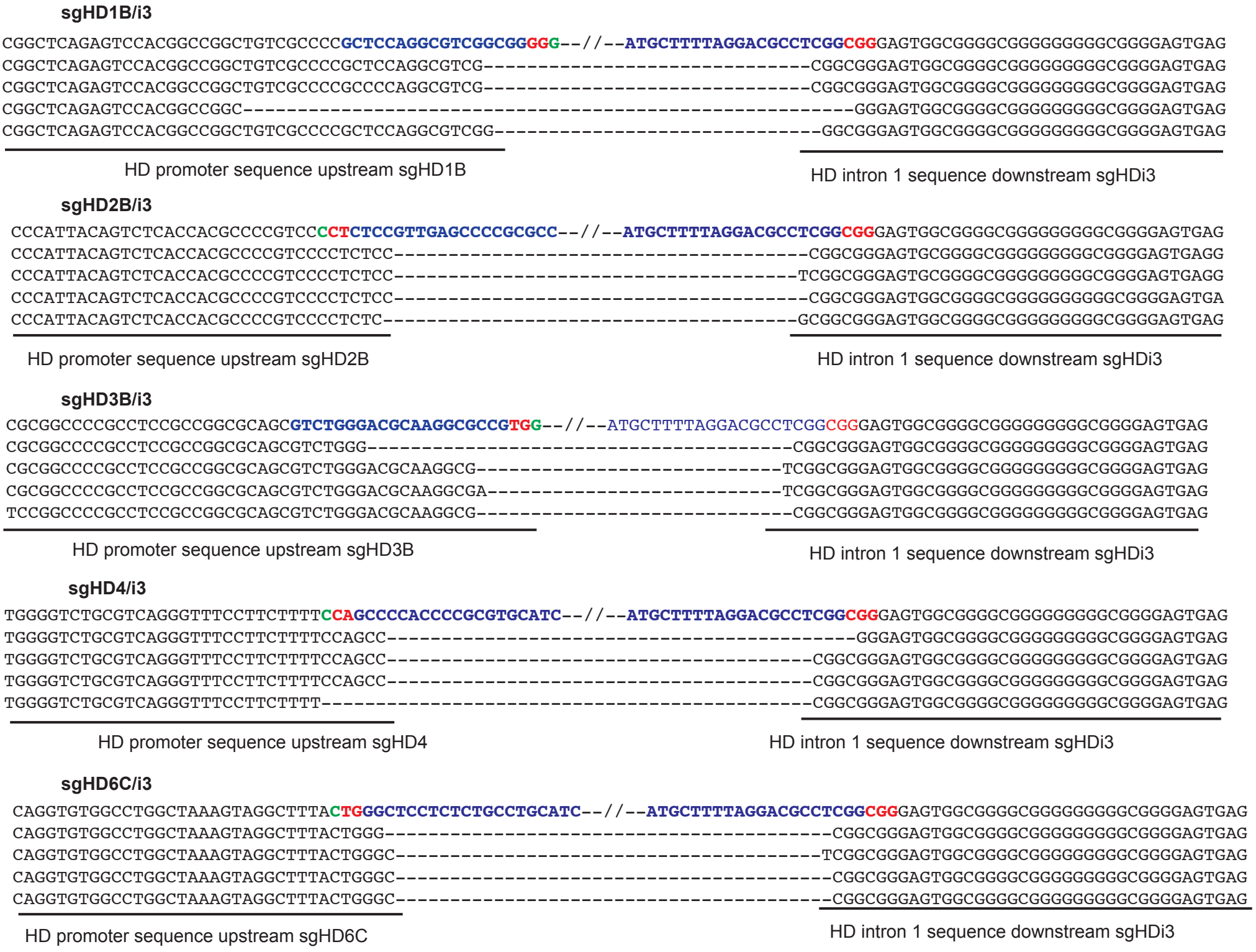
Sequence sgHDi4
Complementary guide: 5’ GCGGGACACTTCGAGAGG 3’ (sgHDi4) (n=18 nt)
sgRNA expression cassette: hU6 promoter

5’ GGCGCGGGACACTTCGAGAGGAGGCGGGGTTTGGAGCTGGAGAGATGT 3’
3’ CCGCGCCCTGTGAAGCTCTCCTCCGCCCCAAACCTCGACCTCTCTACA 5’
GCGGGACACTTCGAGAGG

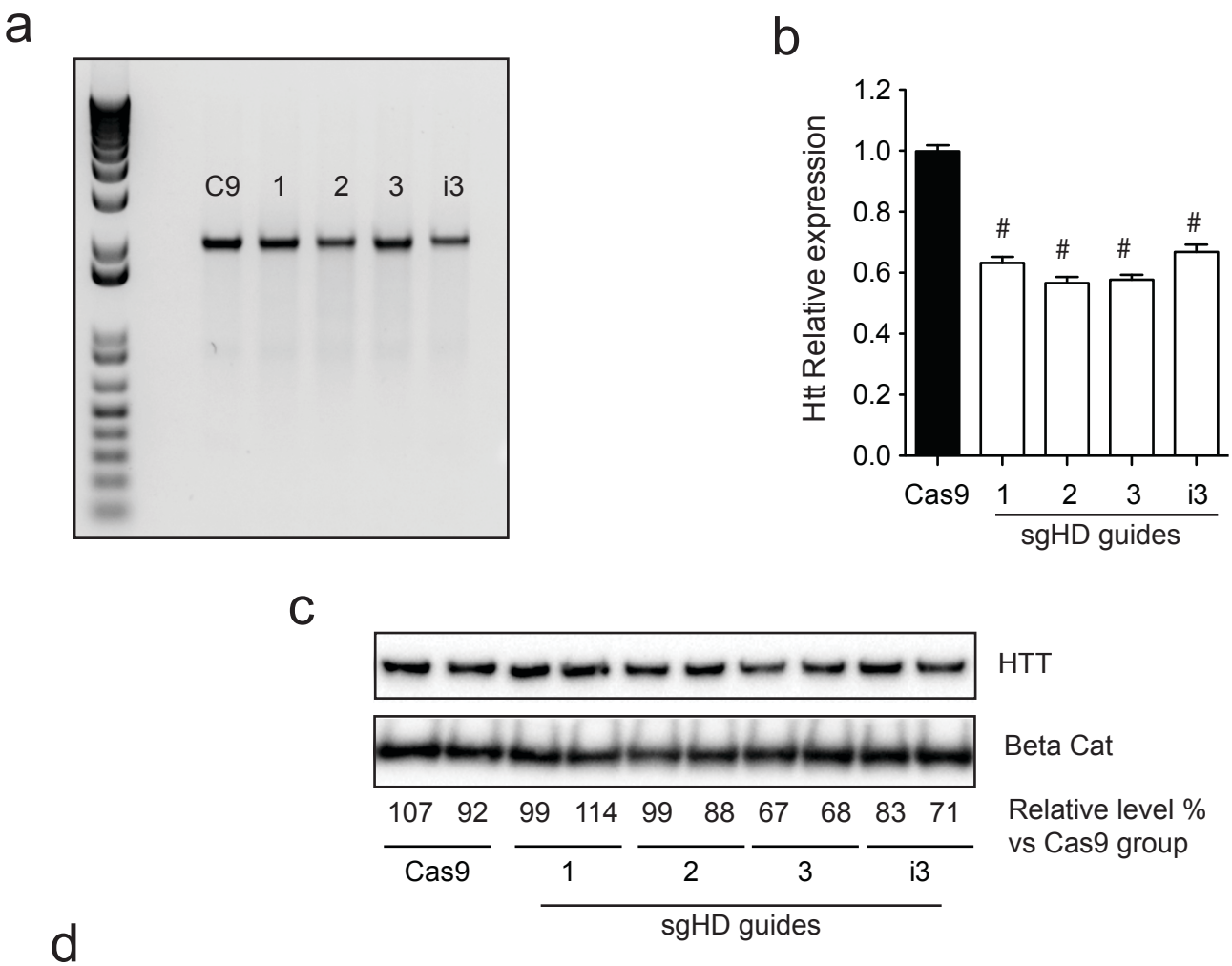
Supplementary Figure 3: Efficacy of SNP-dependent sgHD/SpCas9 complexes in HEK 293 cells using a common intronic sgHDi4 sequences.



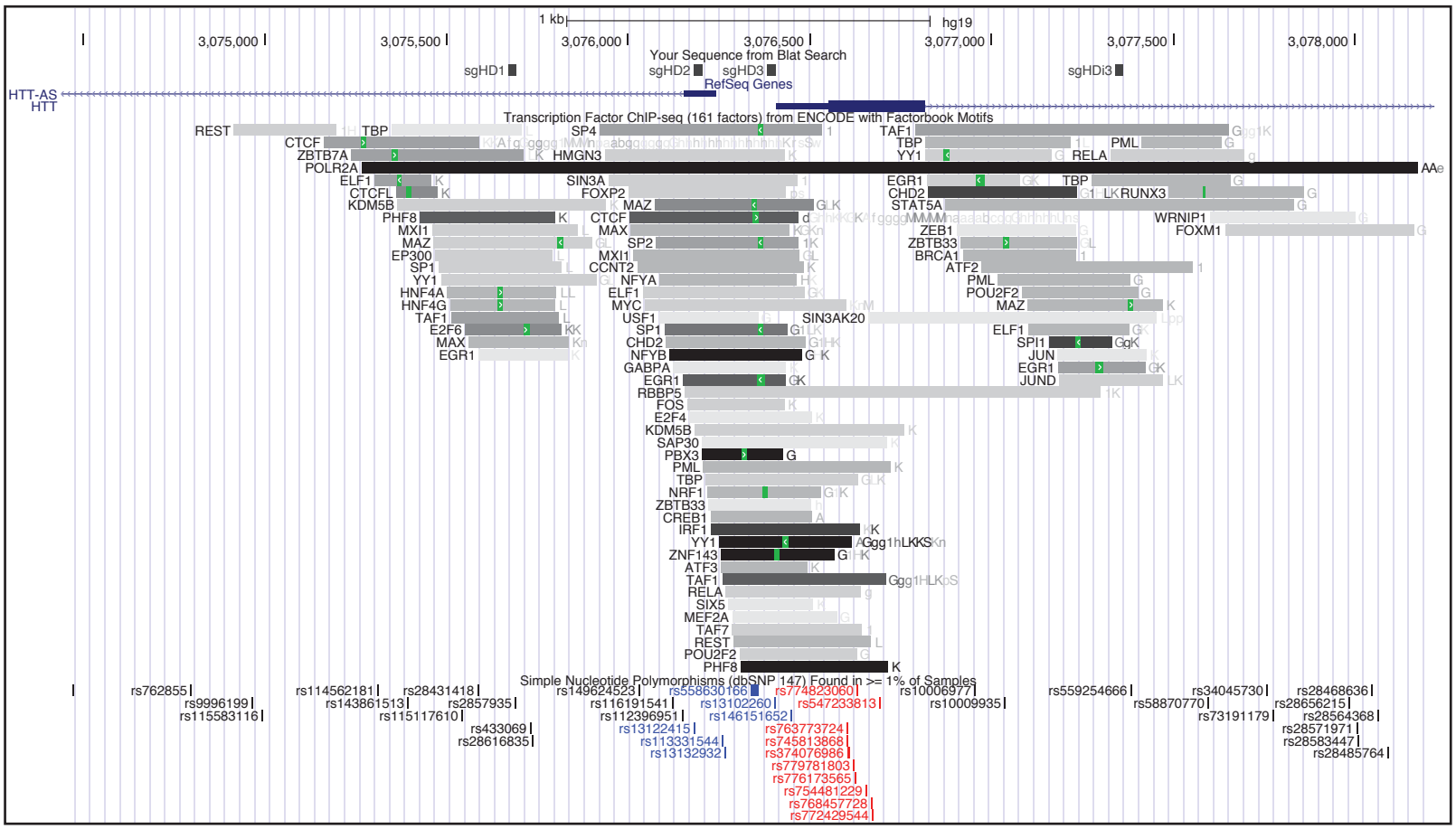
Supplementary Figure 4: Sanger sequencing of PCR amplified products after of HTT exon-1 editing.



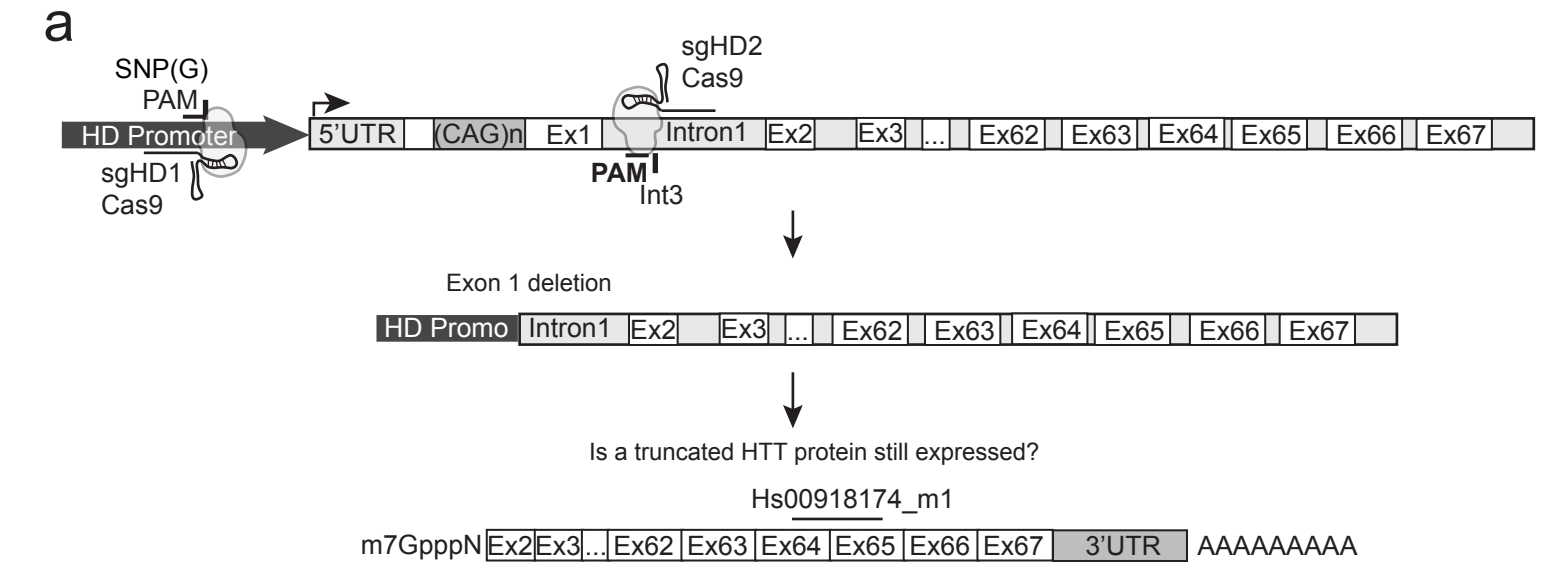
Supplementary Figure 5: Cleavage of single sgHD/SpCas9 complexes in HEK 293 cells.

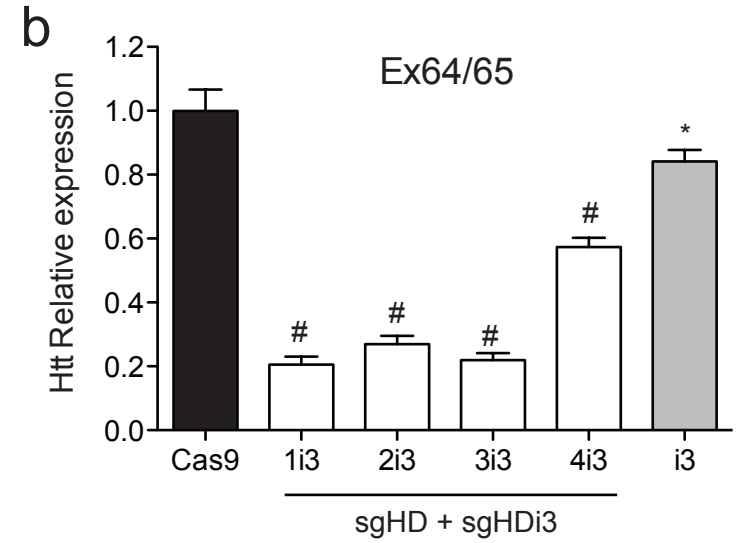


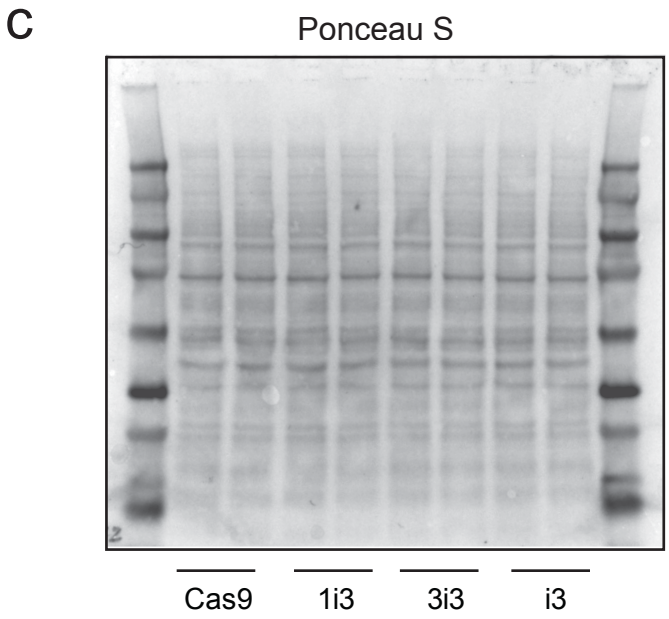
d

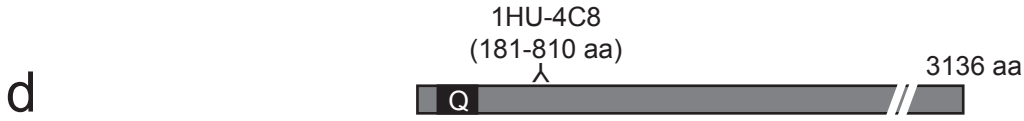


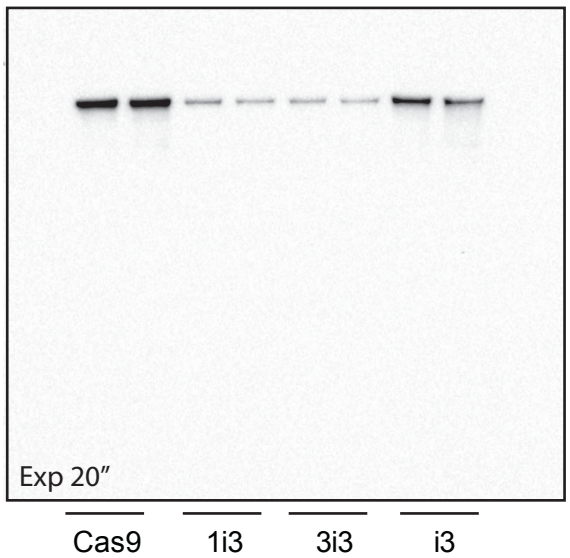
Supplementary Figure 6: C-terminal HTT products are not generated after HTT exon-1 deletion in HEK293 cells.

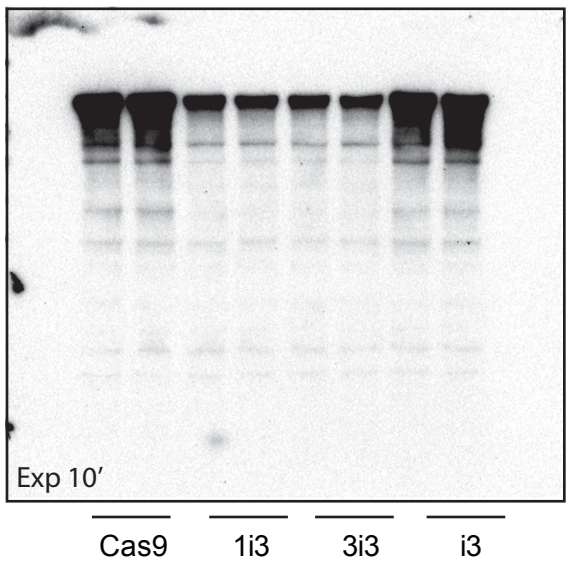




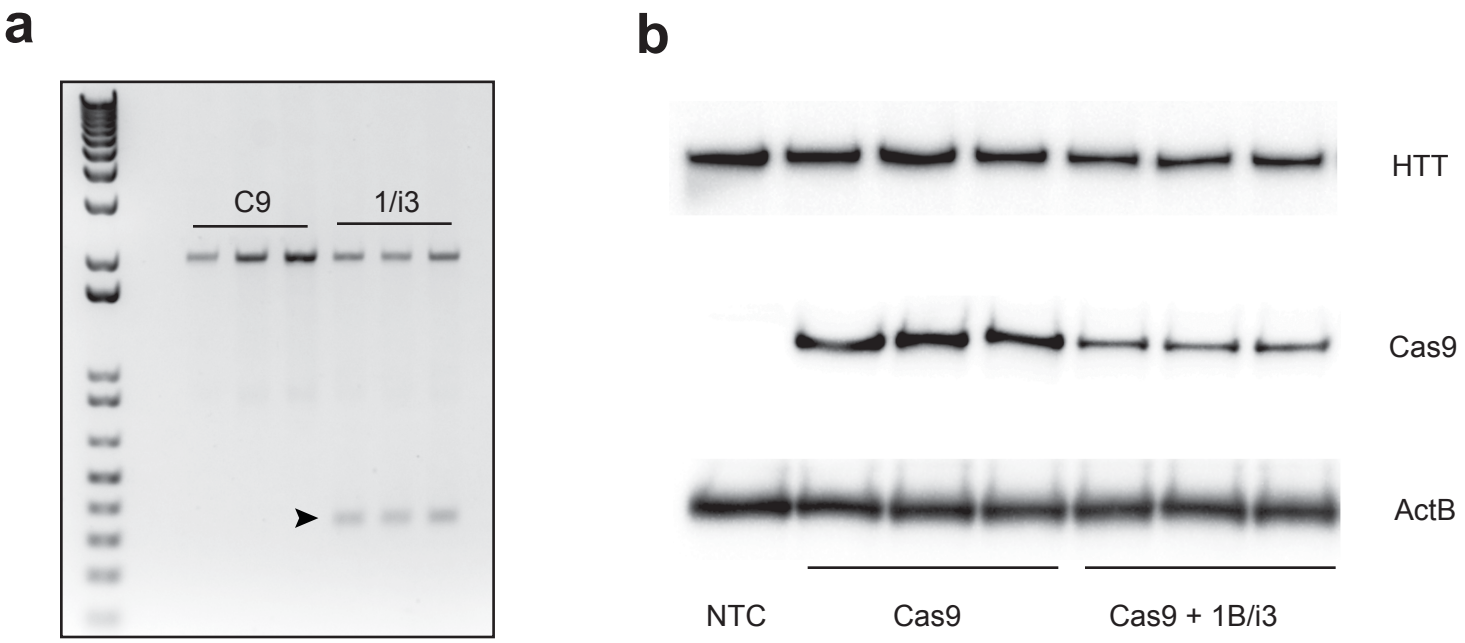






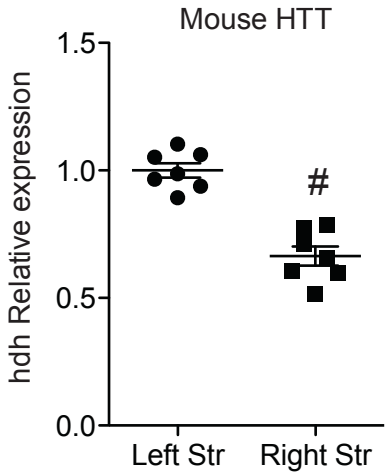


Supplementary Figure 7: In vitro testing of rAAV2/1 SpCas9 and rAAV2/1 vectors in HEK 293 cells.



Supplementary Figure 8: In vivo gene editing of endogenous mouse HTT allele.

a



b

```
sgHDi3:  GCTTTTAGGACGCCTCGGNRG
i3OT1:   GCTTTTAGGAgGtCTCGGCGG

sgHD1:   GCTCCAGGCGTCGGCGGNRG
1m5-OT1: GgcggAGGCGgCGGCGGCGG
1m5-OT2: GgTtgAGGCGgaGGCGGCGG
1m5-OT3: GgaCCtGGCagCGGCGGTGG
```

c

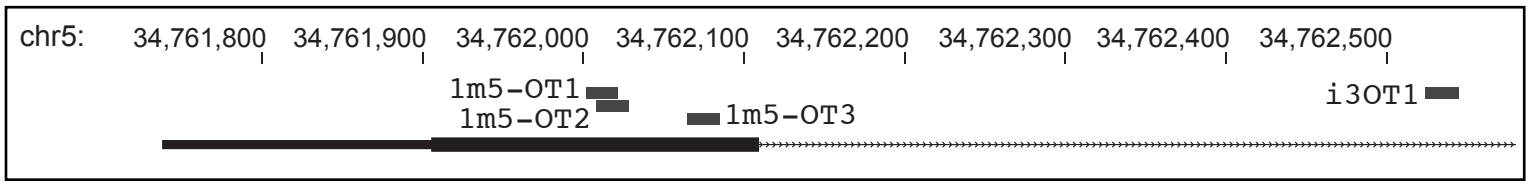


Table S1: List of the 36 prevalent SNPs flanking *HTT* exon-1 within the critical nucleotide position of the different CRISPR systems analyzed.

			1000 Genome DATA		
# Variant ID	Location	Sequence Variation	Ref.Allele	Min. Allele	MAF >0.05
rs35631490	**Promoter**	**GTCTGCGTCAGGGTTTCCTTCTTTT[C/G]CAGCCCCACCCCGCGTGCATCCCAC**	**C**	**G**	**0.1074**
rs73086139	Promoter	TCAAGGCCTCTTCTCTCTTCTCGGC[A/G]GGACAGGCACAGGCAGGTGGCCAGG	A	G	0.0659
rs73086140	Promoter	GCCAGGTGTCATGCTTAGCTCCCCG[C/G]CCAGTGAGATTCTTTCATTTAACAA	C	G	0.0962
rs113541600	Promoter	TAGGAACCTTATTTCTCTCTCGCTC[-/T]TTTTTTTTTTTTTGAGACAGAGTCT	T	delT	0.2091
rs61791259	Promoter	GATTACAGGCACCTGCCACCATGCC[C/T]GGCTAATTTTTGTATTTTTAGTTGA	C	T	0.0661
rs61792460	Promoter	GAGGGTTTCATCTTGTTGGTCAGGC[A/G]GACTTGAACTCCTGACCTCAGGTGA	G	A	0.0966
rs73086144	Promoter	GCCACTGCGCCTTCATCTCTCTTCT[A/G]TGTATGTGTACGCTGTTTTTTCTTT	G	A	0.0899
rs73086145	Promoter	GCTGTTTTTTCTTTAGAATGGGGGA[C/G]GTTATCAGGCTCTACATGGTGTGTA	C	G	0.0899
rs61792461	Promoter	TCTACATGGTGTGTAGTCGGCTAGC[A/G]TGTTGTAAGCCTTTCCCTGTGTCAC	A	G	0.0901
rs61792462	Promoter	CTGTGTCACAAGTGCTCATCTGGAA[C/G]AGGATTCTAATGACTGCCTGTGGCT	C	G	0.0579
rs61090955	Promoter	TCATTTTATGTGATTCCTTTCTAGA[A/G]GTACTACTCATTACTTCTGCTTGCA	A	G	0.0865
rs77384845	Promoter	TAGCTGAAGGAAGGACAGGGACTGT[C/T]ATACACTAGCTAAGAGGCAAACTGC	C	T	0.0984
rs10011412	Promoter	agctgaaggaaggacagggactgtC[A/G]TACACTAGCTAAGAGGCAAACTGCT	A	G	0.1158
rs61792464	**Promoter**	**TCCCTCATTCAGGTTGATGTCCTAA[C/G]CCCCAGAACCTCAGAATGGGATTGT**	**G**	**C**	**0.1372**
rs112435590	Promoter	TCCATGCCAAGAAGGCAACAGAGAG[G/T]GCCAGGGAGACTGAAGTCATACCCT	G	T	0.0881
rs111670395	Promoter	CCCAGGTTCAAGCAATTCTGCCTCA[A/G]CCTCCGGAATAGCTGGGACTACAGG	G	A	0.0663
rs762855	Promoter	TTGAGAAGGACAGCAGAGAAACAGC[C/T]GTTAGTtcccagttcttgggaggct	G	A (Rev strand)	0.4828
rs9996199	**Promoter**	**TGTGGCCTGGCTAAAGTAGGCTTTA[C/G]TGGGCTCCTCTCTGCCTGCATCACC**	**C**	**G**	**0.1575**
rs143861513	Promoter	TCCGGGGCGCTGCGCTGGGACCGAT[-/G]GGGGGGCGCCAGGCCTGTGGACACC	-	insG	0.0599
rs28431418	Promoter	CTCCCCGCAGGGCTGTCCGGGTGAG[C/T]ATGGCTCTGGCCACGGGCCAGTGTG	T	C	0.1438
rs2857935	**Promoter**	**GGGCGCAGGCCCATGCGGAAAGGAT[A/C/G]CCCCGCCGACGCCTGGAGCGGGGCG**	**G**	**C/T (Rev Strand)**	**0.226**
rs28616835	Promoter	GCGCCCGCGCTCGGCGCCCCCTCCA[C/T]GGCCCCGCCCCGTCCATGGCCCCGT	C	T	0.0839
rs149624523	Promoter	CCCCGCCCCGGCCTCGCCACGCCCC[C/T]ACCTCACCACGCCCCCCGCATCGCC	T	C	0.0927
rs13122415	**Promoter**	**attacagtctcaccacgccccgtcc[C/G]CTCTCCGTTGAGCCCCGCGCCTTCG**	**C**	**G**	**0.1082**
rs113331544	5'UTR	GGCCTTGCTGTGTGAGGCAGAACCT[-/GCGGGG]GCGGGGGCAGGGGCGGGCTGGTTCC	-	insGCGGGG	0.0851
rs13132932	5'UTR	TGTGTGAGGCAGAACCTGCGGGGGC[A/G]GGGGCGGGCTGGTTCCCTGGCCAGC	A	G	0.0727
rs13102260	**5'UTR**	**AGCGTCTGGGACGCAAGGCGCCGTG[A/G]GGGCTGCCGGGACGGGTCCAAGATG**	**G**	**A**	**0.1581**
rs10009935	Intron 1	CTCACTTGGGTCTTCCCTTGTCCTC[C/T]CGCGAGGGGAGGCAGAGCCTTGTTG	T	C	0.0877
rs58870770	Intron 1	TGAATGAGTTGTGGTTGCCAAGTAA[-/A]GTGGTGAACTTACGTGGTGATTAAT	A	delA	0.0845
rs34045730	Intron 1	GAGGTGTACATTTTACCAGTATTCC[A/T]GTCAGGCTTGCCAGAATACGGGGGG	A	T	0.0589
rs28656215	Intron 1	GGAAGTCTGTGTGTCGAGTGTACAG[C/T]AGGAGTTAGGAAGTACTCTGGTGCA	T	C	0.1544
rs28571971	Intron 1	ACAGTAGGAGTTAGGAAGTACTCTG[C/G]TGCAGTTCAGGCCTTTCTCTTACCT	G	C	0.0839
rs28583447	Intron 1	CAGTAGGAGTTAGGAAGTACTCTGG[C/T]GCAGTTCAGGCCTTTCTCTTACCTC	T	C	0.0839
rs28468636	Intron 1	TCTCTTACCTCTCAGTATTCTATTT[C/G]CGATCTGGATGTGTCCCAGATGGCA	C	G	0.0839
rs28564368	Intron 1	TATTCTATTTCCGATCTGGATGTGT[A/C]CCAGATGGCATTTGGTAAGAATATC	C	A	0.0839
rs28485764	Intron 1	GATGGCATTTGGTAAGAATATCTCT[A/G]TTAAGACTGATTAATTTTTAGTAAT	G	A	0.0845
rs77173925	Intron 1	GAATAAATTATTCTAAAGGATGGAA[A/G]AACTTTTTGGATATTTGGAGAAATT	A	G	0.0843
rs3905238	Intron 1	TTGTATCATGTCAATGTATTACTTA[C/T]GCAAAAATAATACATTAAAAAAAAT	A	G (Rev Strand)	0.4443
rs33950430	Intron 1	AGGTATTCACTAATTTTGAGTAACA[-/AACA]CTGCTCACAAAGTTTGGATTTTGGC	AACA	delAACA	0.0835
rs28377140	Intron 1	AGGCAATTAATACTTGCTTCTGGCA[C/G]TTTCTTATTCTCCTTCAGATTCCTA	G	C	0.1064
rs3856973	Intron 1	ttaaaaataaaaataGTTAACACT[C/T]GATTAACCCTGACATTTCCCTATCC	G	A (Rev strand)	0.4081
rs4498089	Intron 1	AGTGTTAACTTATTTTTATTTTTAA[A/G]AAAATTGTTAAGGGCTTTCCAGCAA	A	G	0.2917
rs112353753	Intron 1	GACATGCACTGCCATGCCTGGGTAA[-/T]TTTTTTTTTTTCCCCCGAGACGGAG	T	delT	0.1154
rs57666989	Intron 1	CTGCAAGCTCCGCTTCCCGAGTTCA[C/T]GCCATTCTCCTGCCTCAGTCTCCCA	C	T	0.1424
rs10006129	Intron 1	tcttgatctcctgacctcgtcatcc[C/G]ccgaccttgtgatccgcccacctcg	G	C	0.0845
rs28696693	Intron 1	GGTAATTTTTGTATTTTTAGTAGAG[A/G]TGGGGTTTTGCCATGATGAGCAGGC	A	G	0.0839
rs28393280	Intron 1	GGATTTTGAATGCGGAACCAACTGC[A/G]CTTGTTGAACTCTGCTAAGTATAAC	A	G	0.0717

	SpCas9_WT (NRG)						SpCas9_VQR/EQR (NGAN_NGNG/NGAG)					
	Reference Allele			Minor allele			Reference Allele			Minor allele		
# Variant ID	PAM Motif	Strand	PAM Ref>Min	PAM Motif	Strand	PAM Ref>Min	PAM Motif	Strand	PAM Ref>Min	PAM Motif	Strand	PAM Ref>Min
rs35631490	TGG	Negative	Loss	-	-	-	GGAA	Negative	Loss	-	-	-
rs73086139	-	-	-	-	-	-	-	-	-	GGCG	Positive	Gain
rs73086140	GGG	Negative	Loss	CGG	Positive	Gain	GGCG	Negative	Loss	-	-	-
rs113541600	-	-	-	-	-	-	-	-	-	-	-	-
rs61791259	-	-	-	-	-	-	-	-	-	-	-	-
rs61792460	-	-	-	-	-	-	GGCG	Positive	Loss	-	-	-
rs73086144	-	-	-	-	-	-	-	-	-	-	-	-
rs73086145	-	-	-	GAG	Positive	Gain	-	-	-	-	-	-
rs61792461	-	-	-	-	-	-	-	-	-	AGCG	Positive	Gain
rs61792462	-	-	-	AAG	Positive	Gain	-	-	-	AGAG	Positive	Gain
rs61090955	-	-	-	GAG	Positive	Gain	-	-	-	-	-	-
rs77384845	-	-	-	-	-	-	TGAC	Negative	Loss	-	-	-
rs10011412	-	-	-	-	-	-	-	-	-	-	-	-
rs61792464	GGG	Negative	Loss	AAG	Positive	Gain	-	-	-	-	-	-
rs112435590	AGG	Positive	Loss	-	-	-	-	-	-	-	-	-
rs111670395	CAG	Positive	Loss	-	-	-	-	-	-	-	-	-
rs762855	-	-	-	-	-	-	-	-	-	-	-	-
rs9996199	CAG	Negative	Loss	TAG	Positive	Gain	-	-	-	-	-	-
rs143861513	-	-	-	-	-	-	-	-	-	-	-	-
rs28431418	-	-	-	-	-	-	-	-	-	-	-	-
rs2857935	GGG	Negative	Loss	-	-	-	GGAT	Negative	Loss	-	-	-
rs28616835	-	-	-	-	-	-	-	-	-	-	-	-
rs149624523	-	-	-	-	-	-	-	-	-	-	-	-
rs13122415	AGG	Negative	Loss	-	-	-	GGGG	Positive	Loss	-	-	-
rs113331544	-	-	-	-	-	-	-	-	-	GGCG	Positive	Gain
rs13132932	-	-	-	-	-	-	-	-	-	GGCG	Positive	Gain
rs13102260	TGG	Positive	Loss	-	-	-	-	-	-	TGAG	Positive	Gain
rs10009935	-	-	-	CGG	Negative	Gain	CGAG	Negative	Loss	-	-	-
rs58870770	-	-	-	-	-	-	-	-	-	-	-	-
rs34045730	-	-	-	-	-	-	-	-	-	-	-	-
rs28656215	-	-	-	-	-	-	-	-	-	-	-	-
rs28571971	TGG	Positive	Loss	CAG	Negative	Gain	-	-	-	-	-	-
rs28583447	-	-	-	-	-	-	-	-	-	TGCG	Negative	Gain
rs28468636	CGG	Negative	Loss	-	-	-	-	-	-	TGCG	Positive	Gain
rs28564368	GGG	Negative	Loss	-	-	-	GGAC/TGGG	Neg/Neg	Loss	-	-	-
rs28485764	-	-	-	-	-	-	-	-	-	-	-	-
rs77173925	-	-	-	AAG	Positive	Gain	-	-	-	AGAA	Positive	Gain
rs3905238	-	-	-	-	-	-	-	-	-	TGCG	Negative	Gain
rs33950430	-	-	-	-	-	-	-	-	-	-	-	-
rs28377140	CAG	Positive	Loss	AAG	Negative	Gain	-	-	-	-	-	-
rs3856973	-	-	-	-	-	-	CGAG	Negative	Loss			
rs4498089	-	-	-	AAG	Positive	Gain	-	-	-	AGAA	Positive	Gain
rs112353753	-	-	-	-	-	-	-	-	-	-	-	-
rs57666989	-	-	-	-	-	-	GGCG	Negative	Loss	-	-	-
rs10006129	-	-	-	GGG	Negative	Gain	-	-	-	CGGG	Negative	Gain
rs28696693	-	-	-	AGG	Positive	Gain	AGAT	Positive	Loss	-	-	-
rs28393280	-	-	-	-	-	-	-	-	-	TGCG	Positive	Gain

	SpCas9_VRER (NGCG)						SaCas9 (NNGRRT/NNGRR)					
	Reference Allele			Minor allele			Reference Allele			Minor allele		
# Variant ID	PAM Motif	Strand	PAM Ref>Min	PAM Motif	Strand	PAM Ref>Min	PAM Motif	Strand	PAM Ref>Min	PAM Motif	Strand	PAM Ref>Min
rs35631490	-	-	-	-	-	-	TGGAA	Negative	Loss	-	-	-
rs73086139	-	-	-	GGCG	Positive	Gain	-	-	-	GCGGG	Positive	Gain
rs73086140	GGCG	Negative	Loss	-	-	-	-	-	-	-	-	-
rs113541600	-	-	-	-	-	-	-	-	-	-	-	-
rs61791259	-	-	-	-	-	-	CCGGG	Negative	Loss	-	-	-
rs61792460	GGCG	Positive	Loss	-	-	-	GCGGA	Positive	Loss	-	-	-
rs73086144	-	-	-	-	-	-	-	-	-	-	-	-
rs73086145	-	-	-	-	-	-	-	-	-	-	-	-
rs61792461	-	-	-	AGCG	Positive	Gain	-	-	-	-	-	-
rs61792462	-	-	-	-	-	-	-	-	-	AAGAG	Positive	Gain
rs61090955	-	-	-	-	-	-	-	-	-	-	-	-
rs77384845	-	-	-	-	-	-	-	-	-	-	-	-
rs10011412	-	-	-	-	-	-	-	-	-	-	-	-
rs61792464	-	-	-	-	-	-	-	-	-	-	-	-
rs112435590	-	-	-	-	-	-	GAGAG	Positive	-	GAGAGT	Positive	Gain
rs111670395	-	-	-	-	-	-	-	-	-	-	-	-
rs762855	-	-	-	-	-	-	-	-	-	-	-	-
rs9996199	-	-	-	-	-	-	-	-	-	-	-	-
rs143861513	-	-	-	-	-	-	-	-	-	-	-	-
rs28431418	-	-	-	-	-	-	GTGAGT	Positive	Loss	GTGAG	Positive	-
rs2857935	-	-	-	-	-	-	CGGGG	Negative	-	CGGGGT	Negative	Gain
rs28616835	-	-	-	-	-	-	-	-	-	-	-	-
rs149624523	-	-	-	-	-	-	-	-	-	GTGGG	Negative	Gain
rs13122415	-	-	-	AGCG	Negative	Gain	AGGGG	Negative	Loss	-	-	-
rs113331544	-	-	-	GGCG	Positive	Gain	-	-	-	-	-	-
rs13132932	-	-	-	GGCG	Positive	Gain	-	-	-	GCGGG	Positive	Gain
rs13102260	-	-	-	-	-	-	TGGGG	Positive	Loss	-	-	-
rs10009935	-	-	-	-	-	-	-	-	-	CGGGA	Negative	Gain
rs58870770	-	-	-	-	-	-	-	-	-	-	-	-
rs34045730	-	-	-	-	-	-	-	-	-	-	-	-
rs28656215	-	-	-	-	-	-	-	-	-	-	-	-
rs28571971	-	-	-	-	-	-	-	-	-	-	-	-
rs28583447	-	-	-	GGCG/TGCG	Pos/Neg	Gain	-	-	-	-	-	-
rs28468636	-	-	-	TGCG	Positive	Gain	CGGAA	Negative	Loss	-	-	-
rs28564368	-	-	-	-	-	-	-	-	-	-	-	-
rs28485764	-	-	-	-	-	-	-	-	-	-	-	-
rs77173925	-	-	-	-	-	-	-	-	-	AAGAA	Positive	Gain
rs3905238	-	-	-	TGCG	Negative	Gain	-	-	-	-	-	-
rs33950430	-	-	-	-	-	-	-	-	-	-	-	-
rs28377140	-	-	-	-	-	-	-	-	-	-	-	-
rs3856973	-	-	-	-	-	-	TCGAGT	Negative	Loss	-	-	-
rs4498089	-	-	-	-	-	-	-	-	-	AAGAA	Positive	Gain
rs112353753	-	-	-	-	-	-	-	-	-	-	-	-
rs57666989	GGCG	Negative	Loss	-	-	-	-	-	-	-	-	-
rs10006129	GGCG	Negative	Loss	-	-	-	-	-	-	GGGGG	Negative	Gain
rs28696693	-	-	-	-	-	-	-	-	-	-	-	-
rs28393280	-	-	-	TGCG/AGCG	Pos/Neg	Gain	-	-	-	-	-	-

	LbCpf1 (TTTN/TTCN/CTTN/TCTN/ATTN/TCCN/TTGN/GTTN/CCCN/CCTN...)						AsCpf1(TTTN)					
	Reference Allele			Minor allele			Reference Allele			Minor allele		
# Variant ID	PAM Motif	Strand	PAM Ref>Min	PAM Motif	Strand	PAM Ref>Min	PAM Motif	Strand	PAM Ref>Min	PAM Motif	Strand	PAM Ref>Min
rs35631490	TCCA	Positive	Loss	-	-	-	-	-	-	-	-	-
rs73086139	-	-	-	-	-	-	-	-	-	-	-	-
rs73086140	-	-	-	-	-	-	-	-	-	-	-	-
rs113541600	-	-	-	-	-	-	-	-	-	-	-	-
rs61791259	-	-	-	-	-	-	-	-	-	-	-	-
rs61792460	-	-	-	-	-	-	-	-	-	-	-	-
rs73086144	-	-	-	-	-	-	-	-	-	-	-	-
rs73086145	-	-	-	-	-	-	-	-	-	-	-	-
rs61792461	-	-	-	-	-	-	-	-	-	-	-	-
rs61792462	-	-	-	TCTT	Negative	Gain	-	-	-	-	-	-
rs61090955	-	-	-	-	-	-	-	-	-	-	-	-
rs77384845	-	-	-	-	-	-	-	-	-	-	-	-
rs10011412	-	-	-	-	-	-	-	-	-	-	-	-
rs61792464	-	-	-	-	-	-	-	-	-	-	-	-
rs112435590	-	-	-	-	-	-	-	-	-	-	-	-
rs111670395	-	-	-	-	-	-	-	-	-	-	-	-
rs762855	-	-	-	-	-	-	-	-	-	-	-	-
rs9996199	-	-	-	-	-	-	-	-	-	-	-	-
rs143861513	-	-	-	-	-	-	-	-	-	-	-	-
rs28431418	-	-	-	-	-	-	-	-	-	-	-	-
rs2857935	TCCC	Positive	Loss	-	-	-	-	-	-	-	-	-
rs28616835	-	-	-	-	-	-	-	-	-	-	-	-
rs149624523	-	-	-	-	-	-	-	-	-	-	-	-
rs13122415	-	-	-	-	-	-	-	-	-	-	-	-
rs113331544	-	-	-	-	-	-	-	-	-	-	-	-
rs13132932	-	-	-	-	-	-	-	-	-	-	-	-
rs13102260	-	-	-	-	-	-	-	-	-	-	-	-
rs10009935	-	-	-	-	-	-	-	-	-	-	-	-
rs58870770	-	-	-	-	-	-	TTTA	Negative	Loss	-	-	-
rs34045730	-	-	-	-	-	-	-	-	-	-	-	-
rs28656215	-	-	-	-	-	-	-	-	-	-	-	-
rs28571971	-	-	-	-	-	-	-	-	-	-	-	-
rs28583447	-	-	-	-	-	-	-	-	-	-	-	-
rs28468636	TCCG	Positive	Loss	-	-	-	-	-	-	-	-	-
rs28564368	TCCC	Positive	Loss	-	-	-	-	-	-	-	-	-
rs28485764	-	-	-	-	-	-	-	-	-	-	-	-
rs77173925	-	-	-	-	-	-	TTTT	Negative	Loss	-	-	-
rs3905238	-	-	-	-	-	-	-	-	-	-	-	-
rs33950430	TTTG	Negative	Loss	-	-	-	TTTG	Negative	Loss	-	-	-
rs28377140	-	-	-	-	-	-	-	-	-	-	-	-
rs3856973	-	-	-	-	-	-	-	-	-	-	-	-
rs4498089	-	-	-	-	-	-	TTTA	Negative	Loss	-	-	-
rs112353753	-	-	-	-	-	-	-	-	-	-	-	-
rs57666989	-	-	-	-	-	-	-	-	-	-	-	-
rs10006129	-	-	-	-	-	-	-	-	-	-	-	-
rs28696693	-	-	-	-	-	-	-	-	-	-	-	-
rs28393280	-	-	-	-	-	-	-	-	-	-	-	-

Supplementary Table 2: List of prevalent SNPs upstream human HTT exon-1.

				Allele frequency			
SNP ID	**Variant ID**	**Location**	**SNP**	**Reference**	**1000G MAF**	**Strand**	**PAM**
SNP4	rs35631490	3,071,679	C/G	C=0.8926	G = 0.1074	+	Loss
SNP5	rs61792464	3,073,385	G/C	G=0.8628	C = 0.1372	+	Gain/Loss
SNP6	rs9996199	3,074,965	C/G	C=0.8425	G = 0.1575	+	Gain/Loss
SNP1	rs2857935	3,075,691	C/G/T	C=0.7710	G= 0.2260	-	Loss
SNP2	rs13122415	3,076,181	C/G	C=0.8918	G = 0.1082	+	Loss
SNP3	rs13102260	3,076,405	G/A	G=0.8419	A = 0.1581	+	Loss

Supplementary Table 3: SNP genotyping of 23 HD fibroblast lines.

HD Fibroblast	CAG repeat	SNP1 rs2857935 (C>G/T) 0.226	SNP2 rs13122415 (C>G) 0.1082	SNP3 rs13102260 (G>A) 0.1581	SNP4 rs35631490 (C>G) 0.1074	SNP5 rs61792464 (C>G) 0.1372	SNP6 rs9996199 (C>G) 0.1575
GM04723	CAG:72	**C/G**	C	G	C	G	C
GM04869	CAG:50	C	C	G	C	G	C
GM04767	CAG:47	**C/G**	C	G	C	G	C
GM04887	ND	C	C	G	C	G	C
GM04849	ND	C	C	G	C	G	C
GM04689	CAG:46	C	C	G	C	G	C
ND29801	CAG:40	C	C	G	C	G	C
ND29970	CAG:40	**C**	C	**G/A**	C	**G/C**	**G/C**
ND30013	CAG:43	C	C	G	C	G	C
ND30015	CAG:41	**C/G**	C	G	C	G	C
ND30016	CAG:41	**C/G**	C	G	C	G	C
ND30047	CAG:41	**C/G**	C	G	C	G	C
ND30259	CAG:38	C	C	G	C	G	C
ND30422	CAG:40	**C/G**	C	G	C	G	C
ND30626	CAG:41	**C/G**	C	G	C	G	C
ND30967	CAG:43	C	C	G	C	G	C
ND31038	CAG:44	**C/G**	C	G	C	G	C
ND31551	CAG:39	**C/G**	**C/G**	**G/A**	**C/G**	G	**G/C**
ND31846	CAG:40	C	C	G	C	G	C
ND33392	CAG:57	**C/G**	C	G	C	G	C
ND33947	CAG:40	**C/G**	C	G	C	G	C
ND40536	CAG:66	C	C	G	C	G	C
ND40534	CAG:46/26	**C**	C	G	C	G	C

Supplementary Table 4: Genotyping of the nucleotide variation in the mutant and normal allele of HD fibroblast lines heterozygotic for SNP1.

a

			Huntingtin Allele			
HD Fibroblast	**CAG repeat**	**SNP**	**Normal**	**Mutant**	**Targeted**	**Family**
GM04723	CAG:72/17	C/G	G	**C**	Mutant	691
GM04767	CAG:47/18	C/G	G	**C**	Mutant	691
ND30015	CAG:41/20	C/G	G	**C**	Mutant	NINDS3749
ND30016	CAG:41/21	C/G	G	**C**	Mutant	NINDS3749
ND30047	CAG:41/18	C/G	G	**C**	Mutant	NINDS3753
ND30422	CAG:40/18	C/G	G	**C**	Mutant	NINDS3751
ND30626	CAG:41/21	C/G	G	**C**	Mutant	NINDS3752
ND31038	CAG:44/19	C/G	**C**	G	Normal	NINDS3752
ND31551	CAG:39/18	C/G	**C**	G	Normal	Unknown
ND33392	CAG:57/17	C/G	G	**C**	Mutant	NINDS4250
ND33947	CAG:40/18	C/G	G	**C**	Mutant	Unknown

b

SNP1: rs2857935 (C/G/T) PAM: Loss

Major Allele (C): PAM motif on the Positive strand (NGG)

Minor Allele (G): Disruption of the PAM motive.

Minor Allele (T): Disruption of the PAM motive.

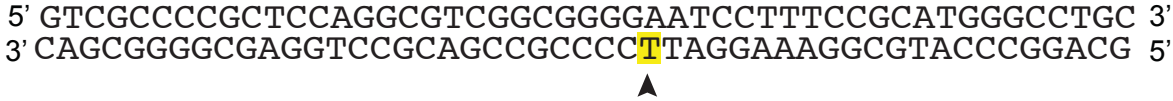


Table S5

Oligos to generate guide sequences:

Name:	Sequence
PosCRPAS1	caccGCTCCAGGCGTCGGCGG
NegCRPAS1	aaacCCGCCGACGCCTGGAGC
PosCRPAS2	caccGGCGCGGGGCTCAACGGAG
NegCRPAS2	aaacCTCCGTTGAGCCCCGCGCC
PosCRPAS3	caccGTCTGGGACGCAAGGCGCCG
NegCRPAS3	aaacCGGCGCCTTGCGTCCCAGAC
PosCRPAS4	caccGATGCACGCGGGGTGGGGC
NegCRPAS4	aaacGCCCCACCCCGCGTGCATC
PosCRPAS5G	tcccATTCAGGTTGATGTCCT
NegCRPAS5G	aaacAGGACATCAACCTGAAT
PosCRPAS5C	tcccATCCCATTCTGAGGTTCTGG
NegCRPAS5C	aaacCCAGAACCTCAGAATGGGAT
PosCRPAS6C	caccGCAGGCAGAGAGGAGCC
NegCRPAS6C	aaacGGCTCCTCTCTGCCTGC
PosCRPAS6G	caccGCCTGGCTAAAGTAGGCTT
NegCRPAS6G	aaacAAGCCTACTTTAGCCAGGC
PosCRi3	caccGCTTTTAGGACGCCTCGG
NegCRI3	aaacCCGAGGCGTCCTAAAAGC
PosCRi4	caccGCGGGACACTTCGAGAGG
NegCRI4	aaacCCTCTCGAAGTGTCCCGC

Primers to assess cleavage

Name:	Sequence
Fwd1SNP	5’-GAC CAC GCG CAT TCT CT -3’
Fwd4SNP	5'- GGA AAC AGG ACA GAT GAA GGAG- 3'
Fwd5SNP	5'- CAG CTC AGA CGG AAG TGT ATT T-3'
Fwd6SNP	5'-CTC CCA AGA ACT GGG AAC TAA C 3'
Rev3Cleavage	5'-ACC ACC GTG ATC ATG AAC TAA A-3'

Name:	Sequence
Fwd HTT	TCGGTGCAGCGGCTCCTC
Rev HTT	ATGGCGACCCTGGAAAAGCTG
FwdActB	TTCGCGGGCGACGATGC
RevActB	CGTACATGGCTGGGGTGTTG