OXFORD

ORIGINAL ARTICLE

Permanent inactivation of Huntington's disease mutation by personalized allele-specific CRISPR/Cas9

Jun Wan Shin[1,2], Kyung-Hee Kim[1,2], Michael J. Chao[1,2], Ranjit S. Atwal[1,2], Tammy Gillis[1], Marcy E. MacDonald[1,2,3], James F. Gusella[1,3,4] and Jong-Min Lee[1,2,3,*]

[1]Molecular Neurogenetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA, [2]Department of Neurology, Harvard Medical School, Boston, MA, USA, [3]Medical and Population Genetics Program, The Broad Institute of M.I.T. and Harvard, Cambridge, MA, USA and [4]Department of Genetics, Harvard Medical School, Boston, MA, USA

*To whom correspondence should be addressed at: Molecular Neurogenetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA 02114, USA. Email: jlee51@mgh.harvard.edu

Abstract

A comprehensive genetics-based precision medicine strategy to selectively and permanently inactivate only mutant, not normal allele, could benefit many dominantly inherited disorders. Here, we demonstrate the power of our novel strategy of inactivating the mutant allele using haplotype-specific CRISPR/Cas9 target sites in Huntington's disease (HD), a late-onset neurodegenerative disorder due to a toxic dominant gain-of-function CAG expansion mutation. Focusing on improving allele specificity, we combined extensive knowledge of huntingtin (*HTT*) gene haplotype structure with a novel personalized allele-selective CRISPR/Cas9 strategy based on Protospacer Adjacent Motif (PAM)-altering SNPs to target patient-specific CRISPR/Cas9 sites, aiming at the mutant *HTT* allele-specific inactivation for a given diplotype. As proof-of-principle, simultaneously using two CRISPR/Cas9 guide RNAs (gRNAs) that depend on PAM sites generated by SNP alleles on the mutant chromosome, we selectively excised ~44 kb DNA spanning promoter region, transcription start site, and the CAG expansion mutation of the mutant *HTT* gene, resulting in complete inactivation of the mutant allele without impacting the normal allele. This excision on the disease chromosome completely prevented the generation of mutant *HTT* mRNA and protein, unequivocally indicating permanent mutant allele-specific inactivation of the HD mutant allele. The perfect allele selectivity with broad applicability of our strategy in disorders with diverse disease haplotypes should also support precision medicine through inactivation of many other gain-of-function mutations.

Introduction

Huntington's disease (HD; OMIM # 143100) (1–3) is one of many genetic disorders, in which a mutation causes disease by a dominant effect of the mutant protein (4). The HD mutation involves expansion of a CAG repeat in the huntingtin gene (*HTT*; OMIM # 613004) that results in an elongated polyglutamine tract in the huntingtin protein. Because this mutation has occurred independently many times in humans, it resides on many different DNA haplotypes in

the HD population, in each case with a distinct array of associated alleles at polymorphic sites in *HTT* (5,6). Regardless of surrounding DNA haplotype, CAG repeat expansions longer than 35 elicit characteristic clinical symptoms, including involuntary movements, cognitive decline and psychiatric disturbance in a fully dominant fashion (2,7). Age at onset of motor signs and age at death are both determined primarily by the size of the expanded repeat (7–11), and through genome-wide association analysis, we recently

discovered genetic loci significantly associated with the difference between observed age at onset of motor signs and that expected based upon the CAG repeat length of HD subjects (12). Although this gain-of-function mutation has been known for more than 20 years (1), and huntingtin has been implicated in many biological processes (2,13,14), effective mechanism-based treatments have yet to be developed.

The recognition in dominant disorders like HD that the presence of the mutant protein is the trigger of disease has stimulated the pursuit of gene silencing as a potential therapeutic avenue (15). For example, RNA-mediated interference or antisense oligonucleotide (ASO)-mediated silencing have emerged as direct solutions to counter the CAG triplet repeat expansion 'polyglutamine' diseases and amyotrophic lateral sclerosis, based on therapeutic efficacy in mouse and rat models (16–18). A major concern is the impact of the treatment on the normal allele and the potential for causing damage to the patient through loss of normal protein activity. Consequently, allele-specific targeting of polymorphic differences between the mutant and normal mRNAs has been attempted in HD model systems, where both ASO and single-stranded RNA strategies have demonstrated the feasibility of preferentially reducing the levels of the mutant allele (15,19,20). The potential applicability of such allele-specific strategies is enhanced in HD by the extensive genetic analysis that has been performed for *HTT* and its haplotypes, including in some cases full-sequence determination (21). This knowledge of the sequence diversity of HD disease chromosomes also provides the opportunity to extend beyond mRNA-lowering strategies to capitalize directly on the greater number of variants found in genomic DNA. Importantly, loss of one copy of *HTT* due to inheritance of balanced translocation chromosome did not generate typical HD symptoms (22), supporting that 1) HD is not caused by haploinsufficiency, and 2) one functional copy of *HTT* is sufficient to maintain cells' integrity. Thus, permanent and selective inactivation of the mutant *HTT* DNA could represent an alternative precision medicine approach to HD and provide a model for application of such a strategy to other dominant disorders. In this study, we developed a strategy to target only the mutant *HTT* DNA, using a powerful CRISPR/Cas9 gene editing technology (23) to create pre-specified inactivating deletion mutation exclusively on the mutant allele.

We first revealed individual DNA variations whose alleles generate or eliminate Protospacer Adjacent Motif (PAM) sequences on the eight most frequent *HTT* gene haplotypes. We then identified pairs of PAM sequences that are present on the mutant chromosome haplotype but absent from the normal chromosome haplotype in a given HD patient. As a proof-of-principle, we used personalized CRISPR/Cas9 strategy to target two patient-specific PAM sites simultaneously to eliminate the promoter region, transcription start site and the CAG expansion mutation of the mutant allele without altering the normal allele. This CRISPR/Cas9 strategy using pairs of custom-designed mutant haplotype-specific PAM-altering SNP variations provides the means to inactivate mutant alleles from the source in a completely allele-specific manner, and with appropriate knowledge of background haplotype, can be used to permanently inactivate any gain-of-function mutations in the human genome.

Results

HTT gene haplotype-specific CRISPR PAM sites generated by PAM-altering SNPs

Our extensive characterization of many haplotypes carrying the HD mutation points to the need in such disorders for a personalized approach to mutant allele-specific strategies to eliminate production of the toxic mutant huntingtin protein. Because we have previously defined individual polymorphic sites that distinguish the common haplotypes in the HD and normal populations (6,21), we reasoned that we could achieve optimal selectivity to mutant allele inactivation using a CRISPR/Cas9 strategy targeting DNA variations that create PAM sequences on the mutant allele. We focused on SNPs altering the CRISPR PAM sites as our basis of allele discrimination since CRISPR *SpCas9* nuclease tolerates single base mismatch between the target sequence and crRNA (24). We have elaborated a strategy for predictable haplotype-specific inactivation that relies on two gRNAs to simultaneously target locations with PAM-altering variations and thereby to generate a pre-designed knockout deletion only on the mutant chromosome.

We first generated a comprehensive map of PAM-altering variations in the *HTT* gene region by cataloging those in 1000 Genomes Project data: 2381 SNPs and 35 indels either generate or eliminate at least one PAM sequence from various bacterial strains (Supplementary Materials, Figs S1 and S2; Tables S1 and S2). In addition, 81 indel sites are flanked by nearby PAM sites (Supplementary Materials, Fig. S3, Table S3) that may also permit allele discrimination through differential hybridization of CRISPR crRNA. The plethora of potential discriminating sites is exemplified by the PAM sequence currently used most commonly (i.e., NGG from *Streptococcus pyogenes* Cas9). 1003 NGG PAM-altering SNPs are present among the 8 most common *HTT* disease haplotypes (Supplementary Materials, Fig. S4, Table S4), which collectively account for approximately 90% of disease chromosomes in HD subjects of European ancestry (6,21). These polymorphic NGG PAM sequences are widely distributed across the *HTT* gene region in a haplotype-specific fashion, providing many personalized allele-selective targets for a CRISPR/Cas9 strategy, as illustrated by pair-wise comparisons in Figure 1 where target PAM sites are represented by circles whose size increases with the frequency differences with which they can be expected to discriminate between two haplotypes in a given diplotype.

Allele-specificity of PAM-altering SNP-based CRISPR/Cas9

For a proof-of-principle experiment, we specifically examined the potential for discrimination of disease and normal chromosomes based on variants present in the most frequent diplotype in the European HD population, namely hap.01/hap.08, accounting for approximately 9% of HD individuals. There are 9 locations where all 47 hap.01 chromosomes from the 1000 Genomes Project have an NGG PAM site that is absent from any of 382 hap.08 chromosomes (Fig. 1, red arrow). Conversely, 16 sites have NGG PAM sequences on hap.08 chromosomes that are not on hap.01 chromosomes (Fig. 1, black arrow). As the goal of this proof-of-principle test was to inactivate *HTT* on the hap.01 chromosome in primary fibroblast cells from an HD patient (Coriell ID, GM01169) with an expanded CAG allele on hap.01 and a normal CAG allele on hap.08, we chose targets among the 9 NGG PAM sites specific to hap.01.

We chose a dual-gRNA approach because of its capacity to create large genomic deletions between the two targeted sites (25). Although many dual-gRNA combinations would be expected to generate intragenic exon-crossing deletions, we chose two hap.01-specific NGG PAM sites (rs1313774 in the upstream promoter region for gRNA 1; rs16843804 in intron 3 for gRNA 2)
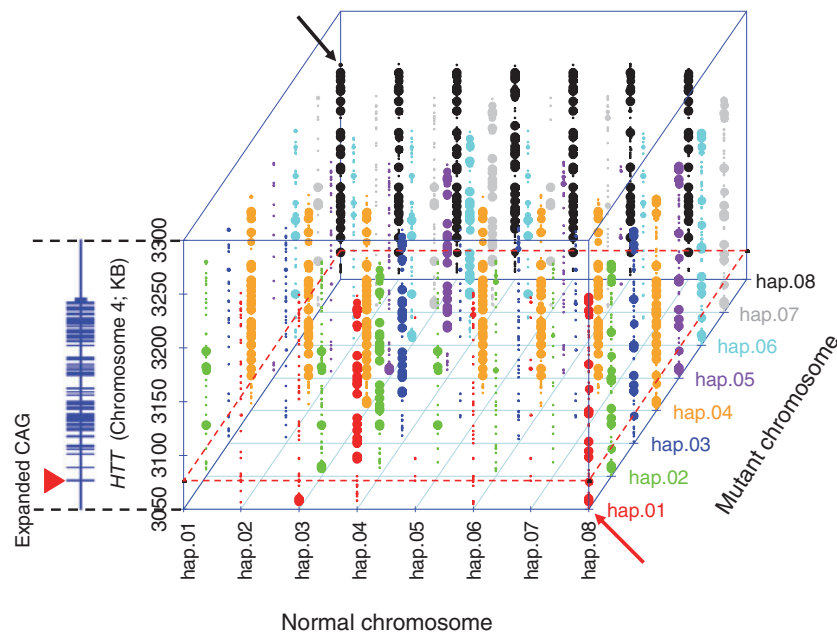
Figure 1. A comprehensive landscape of haplotype-specific NGG PAM sites. For 984 SNPs whose reference or alternative alleles generate NGG PAM sites (chr4: 3042474-3315873; hg19 coordinate), we identified haplotype-specific PAM sites. For a given *HTT* haplotype, 1000 Genomes Project chromosomes do not have entirely identical sequences. Therefore, for a given *HTT* haplotype and NGG PAM site, we calculated the percentage of chromosomes with the NGG PAM sequence. We then compared haplotypes in a pair-wise fashion to identify optimal sites with the maximum presence/absence on disease (diagonal-axis) and normal chromosomes (horizontal-axis) to target the disease chromosome in any given combination of haplotypes. Each PAM site on the disease chromosome is represented by a dot whose size reflects the difference in its frequency on the disease and normal haplotype of the pair. The large dots are exclusively on the disease chromosome (i.e., 100% differential). The vertical-axis represents the chromosomal location of the PAM site variant with an adjacent schematic diagram of *HTT* (RefSeq, NM_002111). A red dotted boundary line indicates the location of the CAG repeat (a filled red triangle on the Y-axis) relative to PAM sites in each haplotype. Each mutant haplotype is colour-coded. For example, red dots represent the differences between hap.01 and other chromosomes, in cases where hap.01 is the mutant chromosome. Red and black arrows point to hap.01/hap.08 and hap.08/hap.01 combinations, respectively.

(Fig. 2) to selectively inactivate the mutant allele by excising the promoter region, transcription start site and the expanded CAG repeat in the exon 1 along with exons 2 and 3 from the mutant copy. Importantly, extensive knowledge of the *HTT* gene haplotype structure also permitted us to choose two SNPs near the chosen PAM sites (rs1313773, monitoring SNP 1; rs2024115, monitoring SNP 2) for closely monitoring the specificity of gRNAs with PAM-altering SNPs (Fig. 2, Supplementary Materials, Figs S5 and S6, Table S5). Primary fibroblast cells derived from a HD patient with hap.01/hap.08 diplotype (GM01169) were transfected with gRNA-expressing CRISPR/Cas9 vectors singly or simultaneously, and the allele specificity of genome modification was judged by sequence analysis of PCR products generated from the pool of transfected cells. In most cases, the single gRNA transfection generated out-of-frame mutations on the mutant allele (Table 1) except a small proportion of normal allele modification by gRNA 2 (Supplementary Material, Fig. S7). The simultaneous use of two gRNAs successfully generated various modifications approximating the expected large 44 kb deletion. All such dual gRNA-mediated large deletions occurred exclusively on the mutant chromosome (Fig. 3), demonstrating perfect allele specificity in comparison to the less complete specificity of the single gRNA approach due to the possible reactivity of Cas9 with non-canonical PAM (26).

Permanent inactivation of the CAG expansion mutation by mutant haplotype-specific CRISPR/Cas9

Co-transfection of patient fibroblast cells with 2 different CRISPR/Cas9 vectors expressing 2 gRNAs and subsequent sequencing analysis indicated that all large deletion occurred on the mutant haplotype, indicating that each modified allele represents a cell with an knockout mutation in the mutant allele. From four independent transfection experiments, we obtained 4 independent targeted single cell clones (TSCC) with mutant allele-inactivated. As anticipated, all targeted single cell clones completely lacked expanded CAG repeats in the DNA (Supplementary Material, Fig. S8). To fully characterize the DNA modification and functional consequences generated by the PAM-altering SNP-based allele-specific CRISPR/Cas9, we selected the first targeted single cell clone (TSCC 1). As shown in Figure 4, untreated patient-derived fibroblasts (GM01169) revealed heterozygous genotypes for PAM sites and monitoring SNPs, confirming the presence of both hap.01 mutant allele and hap.08 normal alleles (Fig. 4A). In contrast, a representative clonal line derived by the dual-gRNA approach showed only the allele characteristic of the hap.08 normal chromosome (Fig. 4B). Additional PCR assays designed to detect the targeted mutant allele revealed genotypes for hap.01 at the monitoring sites, with a large deletion of 44,741 nucleotides with no local inversion relative to the reference sequence (Fig. 4C). The excised mutant DNA was not integrated elsewhere in the genome because 1) parental line's heterozygous genotypes of SNPs located in the excised region became homozygous in the targeted clonal line, and 2) PCR assays to detect an inversion did not reveal such an event in the targeted clonal line (Fig. 5). Furthermore, completely consistent with our hypothesis, the targeted clone (Fig. 6A) lacked any expanded CAG repeat in DNA (Fig. 6B) and RNA (Fig. 6C), and showed significantly reduced total huntingtin protein levels (Fig. 6D) due to complete loss of mutant
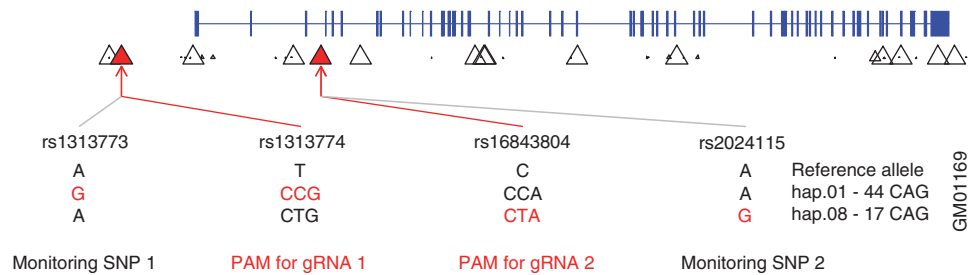
Figure 2. Allele-specific PAM sites and monitoring SNPs on the hap.01/hap.08 diplotype. The most common *HTT* haplotypes on mutant and normal chromosomes are hap.01 and hap.08, respectively. Thus, we tested our novel strategy using primary fibroblast cells derived from a HD patient (Coriell cell line ID, GM01169) with the hap.01 and hap.08 haplotype combination. This HD patient carries 44 and 17 CAGs on hap.01 and hap.08, respectively. The percentage of PAM-containing chromosomes from the 1000 Genomes Project analysis was used to identify hap.01-specific PAM sites. Each triangle represents the location of an NGG PAM site and its frequency difference between hap.01 and hap.08 in the 1000 Genomes Project data. Subsequently, two NGG PAM sites that 1) are present only on the hap.01 haplotype, 2) enclose a region spanning the promoter and the HD mutation, and 3) are flanked by nearby SNPs with different genotypes for hap.01 and hap.08 were chosen to test our strategy. SNPs rs1313774 and rs16843804 were used for the NGG PAM sequence of gRNA 1 and gRNA 2, respectively. Successful targeting by gRNA1 and gRNA2 would excise approximately 44 kb of DNA including the promoter region, transcription start site, and the first 3 exons of *HTT* including the CAG expansion mutation. Two monitoring SNPs, rs1313774 and rs2024115, were used to evaluate the allele specificity of CRISPR/Cas9. Genotypes of those 4 SNP sites were determined by PCR assay followed by sequencing. Phased genotypes were independently validated in a surrogate trio samples (Supplementary Material, Fig. S5).

Table 1. Modifications of DNA by single gRNA

Cells were transfected with a single vector containing gRNA 1 (with rs1313774 NGG PAM site), and DNA samples (collected after 5 days) were amplified for MiSeq analysis of total reads to determine the types of modification of target sites on the mutant and normal chromosomes. Mutant and normal alleles represented in the sequence reads were identified by the 'G' and 'A' alleles of monitoring SNP 1 (i.e., rs1313773), respectively. Recovered w/o modification and recovered with modification indicate that the sequence was unaltered or altered, respectively, by the CRISPR/Cas9. Number of reads and % indicate the number of unique MiSeq sequence reads and the proportion of such reads, respectively. Similarly, cells were transfected with a single vector containing gRNA 2 (with rs16843804 NGG PAM site), and DNA samples were amplified for MiSeq analysis. For this region, mutant and normal alleles were identified by the A and G alleles of the monitoring SNP 2 (i.e., rs2024115), respectively.

gRNA 1

Allele	Description	Number of reads (%)
Mutant	Recovered w/o modification	28,538 (46.5)
	Recovered with modification	861 (1.4)
Normal	Recovered w/o modification	31,942 (52.1)
	Recovered with modification	0 (0)

gRNA 2

Allele	Description	Number of reads (%)
Mutant	Recovered w/o modification	16,753 (29.7)
	Recovered with modification	8,824 (15.7)
Normal	Recovered w/o modification	30,103 (53.4)
	Recovered with modification	671 (1.2)

huntingtin protein (Fig. 6E), confirming that the hap.01-specific dual gRNA CRISPR targeting based on PAM-altering SNPs induced the excision of the expected large DNA segment from the mutant chromosome, ablating production of mutant *HTT* mRNA (Fig. 6G).

Applications of various combinations of allele-specific gRNAs to different HD patients' cell lines

We then evaluated whether different combinations of allele-specific PAM-altering gRNAs could excise the CAG expansion mutation while leaving normal allele intact using additional

independent HD cell lines. We tested 1) a neural precursor cell (NPC) line from an induced pluripotent stem (iPS) cell line (CS97iHD-180n1) (27,28) derived from a Coriell fibroblast line GM09197, and 2) an iPS cell line derived from a Coriell fibroblast line GM04723. An NPC line of GM09197 carries hap.01 and hap.08 as expanded (CAG 180) and normal chromosome (CAG 18), respectively. An iPS cell line of GM04723 carries hap.03 and hap.08 as expanded (CAG 69) and normal chromosome (CAG 15), respectively. Among other possible gRNA combinations, we tested 3 sets of gRNAs: 1) gRNA1 and gRNA2 (Fig. 2) to target hap.01 in the NPC line from GM09197, 2) gRNA3 and gRNA4 to target hap.01 mutant chromosome in the NPC line from GM09197 and 3) gRNA 3 and gRNA 2 to target mutant hap.03 chromosome of an iPS cell line from GM04723. gRNA3 and gRNA4 depend on PAM sites on the mutant chromosomes generated by SNPs rs2857935 and rs7659144, respectively. As shown in Supplementary Material, Table S4, the reference allele of rs2857935 and the reference allele of rs7659144 generate NGG PAM sites on hap.01 and hap.03, not on hap.08 and therefore can be used to excise the CAG expansion mutation in samples with hap.01/hap.08 or hap.03/hap.08 diplotype. Applications of combinations of two allele-specific gRNAs simultaneously to an NPC line and an iPS cell line resulted in excision of the target region as evidenced by correctly sized PCR amplification (Fig. 7). Subsequently, the allele-specificity was judged focusing on an NPC line of GM09197/gRNA 1+ gRNA2 set (Fig. 7A) because 1) monitoring SNP sites are not available and 2) PAM sites are part of excised regions for other sets (i.e. Fig. 7B and C). Co-transfection of an NPC line with gRNA 1 and gRNA 2, and subsequent MiSeq analysis revealed that only mutant chromosomes were modified by allele-specific CRISPR, generating expected large deletions exclusively on the mutant chromosomes as evidenced by sequences of monitoring sites 1 and 2 (Supplementary Material, Fig. S9). Together, these additional data strongly support the wide applicability and allele-specificity of our novel approach.

Discussion

Gene silencing approaches to lower the levels of gain-of-function mutant proteins have demonstrated some therapeutic benefit in animal models and are being moved forward to human trials (15–18). However, the degree to which the toxic
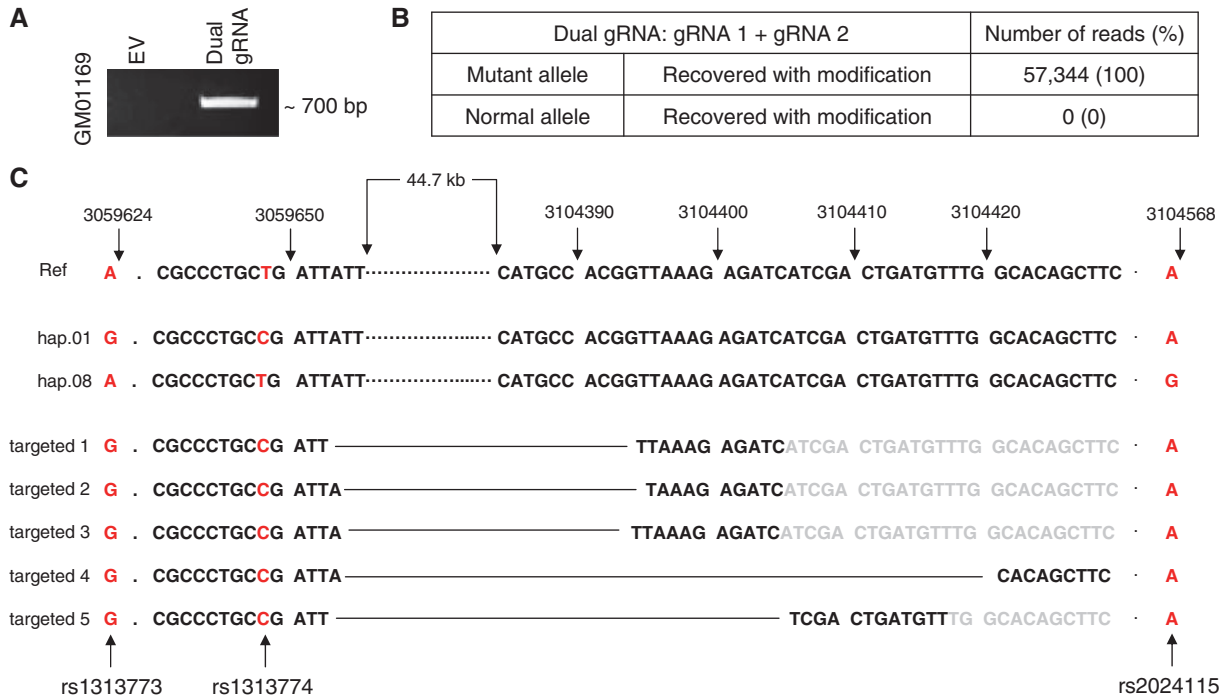
Figure 3. Mutant allele-specific modifications recovered after dual gRNA CRISPR/Cas9 targeting. (**A**) Cells were treated simultaneously with gRNA 1 and gRNA 2. PCR assays were performed on the transfected pool to determine whether the ∼44 kb deletion had occurred. Based upon the assay design, a ∼700 bp PCR product would be generated only if the 44 kb were removed from the genomic DNA. EV and Dual gRNA represent transfection of empty CRISPR/Cas9 vector and co-transfection of gRNA1 and gRNA2 CRISPR/Cas9 vectors, respectively. (**B**) The 700 bp PCR product in Figure 3A was extracted from the gel and subjected to MiSeq sequence analysis to determine allele-specificity and to characterize modified alleles. (**C**) 5 different modifications were recovered by MiSeq analysis in a representative transfection experiment. All of them carried the 'G', 'C' and 'A' alleles at rs1313773 (Monitoring SNP 1), rs1313774 (PAM site 1) and rs2024115 (monitoring SNP 2), respectively, indicating that the mutant chromosome (hap.01) was targeted simultaneously by the two gRNA. rs16843804 (PAM site 2) was part of the excised region, and therefore, not detected by MiSeq analysis. All alleles with modifications indicated a ∼44 kb large deletion of the mutant chromosome. Breakpoints are connected by the black solid line in each targeted allele. Letters in black and grey represent observed and expected sequences, respectively.
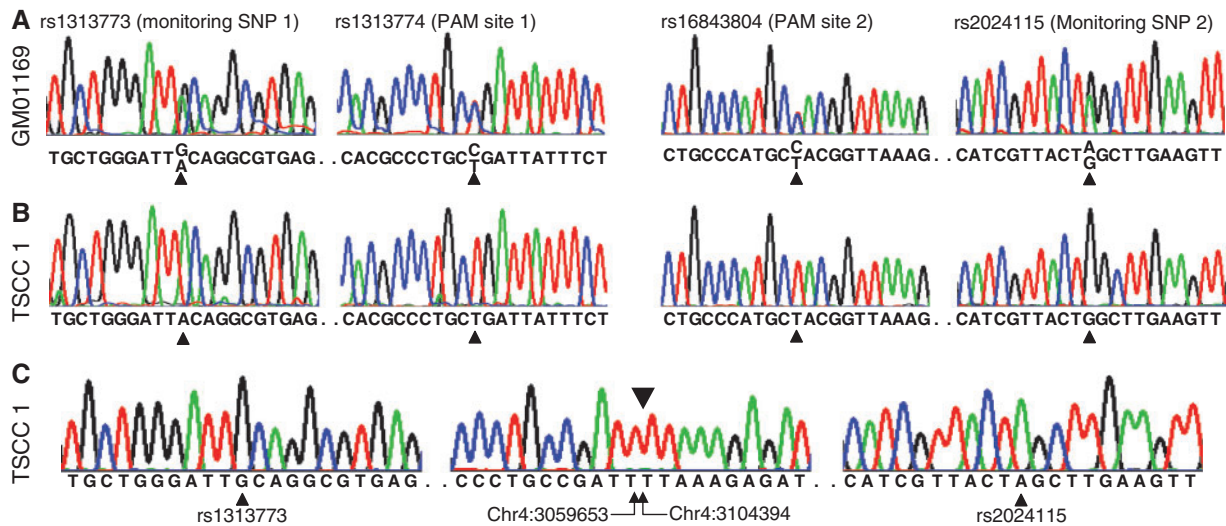
Figure 4. Characterization of the targeted mutant allele. (**A**) Genotypes of untreated cells at rs1313773 (monitoring SNP 1), rs1313774 (PAM site 1), rs16843804 (PAM site 2), and rs2024115 (Monitoring SNP 2) were determined by Sanger sequencing analyses. All genotypes were heterozygous, as expected for the hap.01/hap.08 diplotype. (**B**) A targeted single cell clone (TSCC 1) was genotyped for same variations to confirm the large deletion and to reveal its chromosome of origin. As anticipated, PCR assays revealed only the normal hap.08 allele at all 4 SNP sites, indicating that the deletion occurred on the hap.01 disease chromosome, removing binding sites for assay PCR primers. (**C**) DNA sequencing of the ∼700 bp products revealed that it was derived from the targeted disease chromosome as evidenced by the presence of hap.01-specific alleles at 2 monitoring sites and an ∼44 kb genomic deletion (i.e., from chr4:3059653 to chr4:3104394). The breakpoint junction is marked above the central sequence by an inverted triangle. Genomic coordinates are given below.

Parental line

5'...AGGGTTAATC$_\text{A}^\text{G}$AGTGTTAACT..TCTCACTGGG$_\text{G}^\text{A}$TGAACTAGCA...3'

Targeted clone

5'...AGGGTTAATCAAGTGTTAACT..TCTCACTGGGGTGAACTAGCA...3'

rs3856973   **rs2285086**

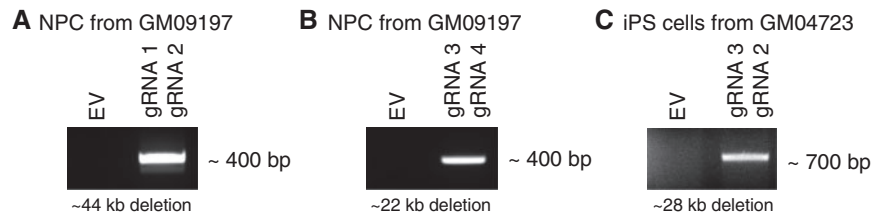**Figure 5.** Consequences of the excised region. To determine whether the ∼44 kb excised DNA from chromosome 4 was incorporated into any other region of the genome, we PCR amplified and genotyped two SNPs that are informative in terms of distinguishing hap.01 from hap.08. Haplotype-discriminating SNPs rs3856973 and rs2285086 are in intron 1 and intron 2 of *HTT*, respectively; hap.01 and hap.08 carry the 'G' and 'A' alleles for rs3856973, and the 'A' and 'G' alleles for rs2285086, respectively. If these segments of the excised ∼44 kb DNA fragment were incorporated elsewhere in the genome without inversion, genotyping assays would show the presence of both alleles. Untreated cells showed heterozygous genotypes for both haplotype-discriminating SNPs, but the targeted cell clone exhibited only the alleles diagnostic of the normal *HTT* segment. In addition, primers designed to detect inversion of the excised region did not generate PCR products (data, not shown), indicating that the segment was not incorporated in the genome with an inverted configuration.

Figure 6. Outcomes of dual gRNA-mediated mutant allele-specific CRISPR/Cas9 targeting. A targeted single cell clone (TSCC 1) was fully characterized to understand the outcomes of mutant allele-specific CRISPR/Cas9 targeting. Untreated cells (GM01169) were used as a reference sample. (**A**) A deletion of approximately 44 kb DNA in the TSCC 1 was confirmed by a PCR analysis. (**B**) Removal of the expanded CAG repeat from the targeted hap.01 chromosome was confirmed by PCR assay of DNA from the untreated cells and a targeted clone. (**C**) Absence of *HTT* RNA with the expanded CAG repeat as a result of our targeting was confirmed by RT-PCR analysis of RNA from the untreated parental line and a targeted clone. (**D**) Total huntingtin protein levels (normal and mutant) were determined by western blot analysis using a pan huntingtin antibody (MAB 2166) in the untreated parental line and targeted cell clone. (**E**) Western blot analysis was performed using mutant huntingtin-specific antibody (1F8) to determine the expression levels of mutant huntingtin. (**F**) Loading of equal amount of protein samples was confirmed by beta-actin (ACTB) antibody. (**G**) A summary of the CRISPR/Cas9-targeting induced DNA structure is shown. The normal haplotype is intact and its RNA production is unhampered. However, the targeted disease chromosome now lacks the promoter region, transcription start site, the first exon with the mutation, and two additional exons and did not produce mRNA with expanded CAG.

Figure 7. Application of allele-specific dual gRNA-mediated CRISPR/Cas9 to independent cell lines. Different combinations of allele-specific gRNAs based on PAM-altering SNPs were tested on independent HD cell lines. We tested 1) an NPC line from an iPS cell line (CS97iHD-180n1) (27,28) derived from a Coriell fibroblast GM09197, and 2) an iPS cell line derived from a Coriell fibroblast GM04723. An NPC line of GM09197 carries hap.01 and hap.08 as expanded (CAG 180) and normal chromosome (CAG 18), respectively. An iPS line of GM04723 carries hap.03 and hap.08 as expanded (CAG 69) and normal chromosome (CAG15), respectively. Among other possible gRNA combinations, we tested 3 sets of gRNAs: (**A**) gRNA 1 and gRNA 2 (Fig. 2) to target hap.01 mutant allele in the NPC line, (**B**) gRNA 3 and gRNA 4 to target hap.01 mutant chromosome in the NPC line, and (**C**) gRNA 3 and gRNA 2 to target mutant hap.03 chromosome of an iPS cells. gRNA 3 and gRNA 4 depend on PAM sites on the mutant chromosomes, generated by SNPs rs2857935 and rs7659144, respectively (Supplementary Material, Table S4). After co-transfection of two CRISPR/Cas9 vectors expressing two gRNAs/puromycin selection, PCR assays were performed using sets of primers described in the method section. Expected size of excision for each experiment is shown under each gel picture. Based upon the PCR assay designs, each PCR products would be generated only if the large deletion occurs. EV represents an empty vector control transfection.

protein must be reduced to eliminate any disease-producing effect is not known in any of these disorders. On the other hand, mRNA-lowering approaches that are not sufficiently allele-selective risk deleterious consequences due to insufficient normal protein activity. For example, with respect to *HTT*, huntingtin function is thought to be essential for the development and maintenance of mature neurons, but one copy of the gene is sufficient for normal development and viability (22,29–31). Consequently, the maximal therapeutic benefit is expected when the mutant allele is completely silenced while expression from the normal allele is left intact. In addition, the hurdle of balancing allelic effects of gene silencing at the RNA level may require repeated treatments which could increase the risk of complications in individuals with chronic progressive neurodegenerative diseases. These considerations make mutant allele-specific DNA targeting approaches highly promising, because the therapeutic modification of DNA represents a permanent change and therefore does not require repeated treatments. With some beneficial effects of RNA-lowering approaches in HD model systems, establishing a personalized DNA targeting strategy with perfect mutant allele specificity to avoid any permanent inactivation of the normal allele is critically important for the development of effective and safe therapeutics. The alternative approach that we propose is to directly target the DNA to completely inactivate only the mutant allele from the source, as we demonstrate that a personalized dual-gRNA CRISPR/Cas9 strategy of utilizing PAM-altering variants can achieve this complete selectivity to prevent the generation of mutant *HTT* mRNA.

Our novel strategy to improve allele specificity, not focusing on increasing targeting efficiency, is significant for a number of reasons. Firstly, perfect allele specificity was achieved by simultaneously using two gRNAs targeting mutant allele PAM sites generated by SNP variations. Partial allele specificity may be obtained by targeting an SNP allele by crRNA hybridization (23,32). However, we hypothesize that better allele discrimination can be achieved by using PAM-altering SNPs as we demonstrated near perfect and perfect allele specificity by single gRNA and dual gRNA approaches, respectively. Non-canonical PAM sequences were able to generate DNA breaks as previously reported (26) and as shown in our data. However, a dual gRNA approach overcame the limitation of non-canonical PAM-mediated CRISPR/Cas9 targeting, allowing perfect allele specificity. Very low levels of DNA breaks may occur at a single gRNA site of normal alleles, but such non-allele-specific modifications

at single site in the promoter region or in an intron may be tolerated because those changes do not impact the transcription of the normal allele dramatically. The use of two gRNAs simultaneously may increase the levels of off-targeting and lower the targeting efficiency as observed in our data, but effort in the field to improve the specificity and efficacy of CRISPR/Cas9 systems is expected to significantly improve the target specificity and efficiency of the CRISPR/Cas9 systems (15,33). Secondly, our strategy of eliminating the promoter region, transcription start site and the disease mutation completely prevented the production of mRNA from the mutant allele, and therefore is not dependent on nonsense-mediated decay mechanisms for gene inactivation. Depending on the types of modifications generated by CRISPR/Cas9, nonsense-medicated decay may not be activated (e.g. in-frame deletion), or even truncated protein may be produced potentially leading to even worse disease outcomes (34,35). The problem of unpredictable modification in DNA is a serious problem when CRISPR/Cas9 is directly applied to humans. However, this potential problem can be solved by eliminating the essential elements required for transcription of the mutant allele as demonstrated here. Disease alleles by point mutations may be distinguished from normal alleles by crRNA hybridization only when nearby PAM sites are present (32,36). In addition, rare cases of disease mutations directly alter PAM sequences permitting allele discrimination (37,38). Even in challenging situations where the PAM sites are not present near the disease mutations or disease mutations do not directly generate PAM sites, our strategy of eliminating promoter and transcription start site in an allele-specific fashion will successfully be applied to knock out disease alleles regardless of the types and positions of the mutations as long as haplotypes of genes are characterized. Our approach targeting disease allele-carrying haplotype guarantees the inactivation of the mutant allele regardless of the types of the modifications, providing a versatile mechanism for completely shutting down any gene of choice. Lastly, our novel approach of using PAM-altering SNPs and nearby DNA variations provides highly powerful experimental tools for directly monitoring specificity of allele-selective CRISPR/Cas9 reagents. When haplotype-informative SNPs are used, determination of targeting efficiency, allele specificity and possibility of genome-integration are quite straightforward. In addition, when a single cell clones with an inactivated mutant allele have to be established, this advantage make the screening assays highly feasible. DNA polymorphisms that are widespread in humans make our strategy powerful and broadly

applicable, permitting the application of our concept to virtually any places in the genome.

With prior detailed delineation of necessary haplotypes, this strategy can be applied to a wide variety of devastating dominant gain-of-function disorders; some notable examples are listed in Table 2, along with the array of polymorphisms that affect only NGG PAM sites which provide a rich collection for targeting sites using our dual-gRNA approach. In our proof-of-principle experiment, we chose not only to inactivate the mutant *HTT* allele, but in doing so, to remove the mutation itself to eliminate any opportunity for expressing a potentially toxic truncated polyglutamine protein. A similar consideration could be applied, where appropriate, in the design of therapeutic approaches in these other dominant disorders. Critical challenges have to be addressed before applying CRISPR strategies to humans. Nonetheless, the promise of this strategy to completely inactivate only the mutant gene dictates that it now be tested in animal model systems to define the safety of the approach and to establish optimal mechanisms for delivery and efficiency in order to support eventual clinical trials.

Materials and Methods

Identification of PAM-altering variations and mapping PAM sites to *HTT* haplotypes

DNA variations reported by the 1000 Genomes Project (phase 1, release 3) were analysed to identify 1) SNPs and indels that either create or eliminate PAM sequences, and 2) indel polymorphisms flanked by PAM sequences for various bacterial strains. This analysis was focused on *HTT* (chromosome 4: 3042474-3315873, hg19 assembly) to avoid targeting of regulatory or coding regions of other genes. Subsequently, 1000 Genomes Project chromosomes were classified into the *HTT* haplotypes that we have developed previously (6,21). For each haplotype and PAM site, the percentage of PAM-containing chromosomes out of all chromosomes was calculated to generate a comprehensive map of PAM sites on each *HTT* haplotype. PAM sites were compared

in a pair-wise manner to reveal haplotype-specific PAM sites for personalized allele-selective targeting in any given diplotype.

Selection of haplotype-specific PAM sites to specifically inactivate mutant *HTT*

Since hap.01 and hap.08 are the most common *HTT* haplotypes on HD disease and normal chromosomes, respectively (21), we tested our strategy in the cells of an HD patient with the hap.01/hap.08 diplotype. Sites of the most commonly used PAM sequence (NGG from *Streptococcus pyogenes* Cas9) were compared between hap.01 and hap.08 chromosomes in the 1000 Genomes Project data. We aimed at completely inactivating the mutant allele using two mutant allele-specific CRISPR/Cas9 vectors. Therefore, we identified two NGG PAM sites that are: 1) present on hap.01 (representing the mutant chromosome), but absent from hap.08 (representing the normal chromosome), 2) flanking the promoter region and the HD CAG expansion mutation in the first exon and 3) flanked by nearby SNPs to monitor allele specificity (monitoring SNPs). Genotypes of PAM sites, monitoring SNPs and other haplotype-informative SNPs were confirmed by Sanger sequencing.

Generation of vector co-expressing gRNA and Cas9

CRISPR/Cas9 vectors were prepared specifying two different gRNAs at regions requiring the chosen PAM sites: gRNA 1 (chr4:3059651-3059670) targets upstream of the promoter region and gRNA 2 (chr4:3104392-3104411) targets an intron 3 region. lentiCRISPRv2 plasmid was a purchased from Addgene (plasmid # 52961). gRNA sequences for target sites were designed and these sequences were cloned into the vector according to the protocol on the GeCKO website (http://genome-engineering.org/gecko/). These vectors produce both crRNA and tracerRNA as a contiguous molecule. Sequences of crRNAs are provided in Supplementary Material, Figure S6. Quality and potential off-target sites of gRNAs were evaluated using the Optimized CRISPR Design website (http://crispr.mit.edu/).

Table 2. NGG PAM-altering SNPs in selected genes associated with dominant disease. We searched NGG PAM-altering SNPs in a limited selection of genes responsible for dominant disease outcomes in order to evaluate the breadth of applicability of our approach. For each locus, the longest transcript from the RefSeq database was chosen to represent the gene. Subsequently, a flanking region of 10 kb was added at both ends of the transcript region to search for NGG PAM-altering SNPs in 1000 Genomes Project data (phase 1, release 3; minor allele frequency > 1% in all populations). The numbers of NGG PAM sites generated by either reference alleles or alternative alleles in the region, representing a repertoire of allele-specific CRISPR/Cas9 target sites across each gene, were counted

Disease	Gene	Region (hg19)	Number of SNPs with MAF > 1%	NGG PAM sites generated by reference alleles	NGG PAM sites generated by alternative alleles
Alzheimer disease	*APP*	chr21:27242860-27553446	1329	288	196
Amyotrophic lateral sclerosis	*SOD1*	chr21:33021934-33051243	84	24	18
Frontotemporal dementia	*MAPT*	chr17:43961747-44115699	972	214	207
Frontotemporal dementia	*C9orf72*	chr9:27536542-27583864	266	47	45
Parkinson disease	*LRRK2*	chr12:40608812-40773086	754	138	105
Dentatorubral-pallidoluysian atrophy	*ATN1*	chr12:7023625-7061484	133	42	24
Myotonic dystrophy 1	*DMPK*	chr19:46262966-46295815	94	38	10
Myotonic dystrophy 2	*CNBP*	chr3:128876657-128912810	96	26	9
Spinocerebellar ataxia Type 1	*ATXN1*	chr6:16289342-16771721	1884	426	276
Spinocerebellar ataxia Type 2	*ATXN2*	chr12:111880017-112047480	370	78	53
Spinocerebellar ataxia Type 3	*ATXN3*	chr14:92514895-92582965	416	94	61
Spinocerebellar ataxia Type 6	*CACNA1A*	chr19:13307255-13627274	1543	408	263
Spinocerebellar ataxia Type 7	*ATXN7*	chr3:63840232-63999136	419	75	67
Spinocerebellar ataxia Type 17	*TBP*	chr6:170853420-170891958	147	33	21

Cell culture, transfection and establishment of targeted single cell clones

Untransformed primary fibroblast cells derived from a male Caucasian HD patient were obtained from the Coriell Cell Repositories. Cells (GM01169, https://catalog.coriell.org/0/Sections/Search/Sample_Detail.aspx?Ref=GM01169) carry hap. 01 and hap.08 haplotypes based on phasing and haplotype analysis (6). Fibroblast cells were cultured in Dulbecco's modified eagle medium (DMEM, Gibco) supplemented with 20% FBS (Gibco) and incubated at 37 °C in a humidified chamber with 5% CO_2. Cells were transfected with both plasmid vectors simultaneously or singly. Approximately, 1.2 million cells were transfected with 4 µg of vector by electroporation (Amaxa Nucleofector I, Lonza) using the Basic Primary Fibroblasts Nucleofector Kit (Lonza). Forty-eight hours after transfection, cells were treated with 1.5 µg/ml of puromycin (Invitrogen) for 72 h to enrich transfected cells. To obtain single cell clones, a small fraction of cells were then transferred to new plates; ~1000 single cells were seeded onto a 10 of 100 mm culture dishes by limiting dilution and incubated for 2 weeks. Subsequently, each well-isolated visible colony was picked and independently cultured for validation analysis. Four independent targeted clones whose mutant *HTT* alleles were selectively inactivated were obtained from 4 independent transfection experiments. One representative clone was further fully characterized.

Extraction of genomic DNA, total RNA and cDNA synthesis

Genomic DNA was isolated from fibroblast cells using DNeasy Blood & Tissue Kit (Qiagen) according to the manufacturer's instructions. Total RNA was extracted from cells using RNeasy Plus Mini Kits (Qiagen). The quality and quantity of extracted RNA were measured using NanoDrop (Thermo Scientific). cDNA was synthesized from 50 ng of total RNA using SuperScript III Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol.

Polymerase chain reaction (PCR)

PCR reactions were performed by using Q5 High-Fidelity DNA Polymerase (NEB) or AccuPrime GC-Rich DNA Polymerase (Invitrogen) for high GC content PCR. The primer sets used in PCR experiments are provided in Supplementary Material, Table S5. The reaction for Q5 High-Fidelity DNA Polymerase was performed as follows: initial denaturation for 3 min at 98 °C, then 35 cycles of denaturation for 30 s at 98 °C, annealing for 30 s at 64 °C and extension for 30 s at 72 °C, then the final extension for 2 min at 72 °C. The reaction for AccuPrime GC-Rich DNA Polymerase was performed as follows: initial denaturation for 3 min at 95 °C, then 30 cycles of denaturation for 30 s at 95 °C, annealing for 30 s at 59 °C and extension for 30 s at 72 °C, then the final extension for 10 min at 72 °C. PCR amplicons were resolved in 2% agarose gel (Invitrogen), and extracted from the gel using QIAquick Gel Extraction Kit (Qiagen) for subsequent analysis.

Deep sequencing of PCR products by MiSeq and western blot analysis

Deep sequencing of PCR products was performed by the CCIB DNA Core at the Massachusetts General Hospital (Boston, MA). Briefly, upon ligation of Illumina adaptors and a unique identifier to the amplicons, paired-end sequencing (2x150b) was performed on the Illumina MiSeq system. Analysis-ready data were provided by the core facility and further summarized. Mutant huntingtin protein levels were determined by Western blot analysis. 15 microgram of whole cell lysate was resolved by 4% Tris-Glycine gel. A transferred membrane was probed by mutant huntingtin-specific 1F8 antibody or pan huntingtin antibody (MAB2166 from EMD Millipore). Equal loading was confirmed by beta-actin (ACTB) protein.

Application of allele-specific dual gRNA-mediated CRISPR/Cas9 to an NPC and an iPS cell line from independent HD subjects

An NPC cell line was generated from an iPS cell line (CS97iHD-180n1) by a STEMdiff protocol using Neural Induction Medium (STEMCELL Technologies). The parental iPS cell line CS97iHD-180n1 was derived from Coriell fibroblast GM09197 (https://catalog.coriell.org/0/Sections/Search/Sample_Detail.aspx?Ref=GM09197&Product=CC) (28). This NPC cell line was maintained on poly-L-ornithine/laminin coated plates in media (70% DMEM, 30% Hams F12, 1X B27 Supplement, 20 ng/ml FGF, 20 ng/ml EGF and 5 µg/ml Heparin) with 5% CO_2 at 37 °C. An iPS cell line was generated from Coriell fibroblast GM04723 (https://catalog.coriell.org/0/Sections/Search/Sample_Detail.aspx?Ref=GM04723&Product=CC) by the Harvard Stem Cell Institute iPS Core Facility (http://ipscore.hsci.harvard.edu/) (39). iPS cell lines were cultured on matrigel-coated plates with mTeSR™1 (STEMCELL Technologies) with 5% CO_2 at 37 °C. The *HTT* gene haplotype of these cell lines were determined by haplotyping phasing as described elsewhere (6). The same gRNAs (i.e., gRNA 1 and gRNA 2; Fig. 2) were tested on the NPC line. In addition, two additional gRNAs were tested on an NPC and an iPS cell line. The gRNA 3 targets chr4:0375669-3075688 (hg19 assembly), and depends on the PAM site generated by rs2857935. This gRNA was used in combinations to treat both NPC and iPS lines. The gRNA 4 targets chr4:3098324-3098343, and depends on the PAM site generated by rs7659144. This gRNA was used in combinations to test the NPC line. In summary, the NPC line was treated with either 1) gRNA 1 + gRNA 2, or 2) gRNA 3 + gRNA 4, and the iPS line was treated with gRNA 3 + gRNA 2 (Fig. 7). NPC and iPS cell lines were transfected with 2 µg of CRISPR vectors containing gRNAs by an electroporation method (Amaxa Nucelofector I, Lonza) using Human Stem Cell Nucleofector Kit 1 (Lonza). Twenty-four hour post-transfection, iPSCs and NPCs were then treated with 0.5 µg/ml of puromycin for 48 h or 2 µg/ml of puromycin for 72 h, respectively. Subsequently, genomic DNA was extracted (Qiagen) for PCR assays to confirm large deletion. Excision of the expanded CAG repeat by gRNA 1 + gRNA 2 would generate approximately 400 bp PCR amplification when the following primers were used; forward primer, GTATTTTTAGTAGAGACGGGGTTTC and reverse primer, ATTCCTCACAGCACATCTCT. Excision of the expanded CAG repeat by gRNA 3 + gRNA 4 was predicted to generate approximately 400 bp PCR product when amplified by the following two primers: forward primer, TGAGTATGGCTCTGGCCACG and reverse primer, AAATGGAATCCAGAGTTACCAGAAGGG. The excision of CAG expansion mutation by application of gRNA 3 and gRNA 2 to the iPS cell line was expected to generate approximately 700 bp PCR product when the following primers were used; forward primer, CTTCTCGCTGCACTAATCAC and reverse primer, ATTCCTCACAGCACATCTCT. Amplified DNA from an NPC cell line/gRNA

1+gRNA 2 experiment was cleaned up using ExoSAP-IT (Affymetrix) (Fig. 7A), and was subjected to MiSeq analysis to detect modified alleles as described in the method section.

Software package for data analysis

All data analyses were performed using R (version, 3.0.2).

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

Funding

References

1. HDCRG., (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell*, **72**, 971–983.
2. Bates, G.P., Dorsey, R., Gusella, J.F., Hayden, M.R., Kay, C., Leavitt, B.R., Nance, M., Ross, C.A., Scahill, R.I., Wetzel, R., *et al.* (2015) Huntington disease. *Nat. Rev. Dis. Primers*, **23**, 15005.
3. Huntington, G. (1872) On chorea. *Med. Surg. Rep.*, **26**, 320–321.
4. Orr, H.T and Zoghbi, H.Y. (2007) Trinucleotide repeat disorders. *Annu. Rev. Neurosci.*, **30**, 575–621.
5. Warby, S.C., Montpetit, A., Hayden, A.R., Carroll, J.B., Butland, S.L., Visscher, H., Collins, J.A., Semaka, A., Hudson, T.J. and Hayden, M.R. (2009) CAG expansion in the Huntington disease gene is associated with a specific and targetable predisposing haplogroup. *Am. J. Hum. Genet.*, **84**, 351–366.
6. Lee, J.M., Gillis, T., Mysore, J.S., Ramos, E.M., Myers, R.H., Hayden, M.R., Morrison, P.J., Nance, M., Ross, C.A., Margolis, R.L., *et al.* (2012) Common SNP-based haplotype analysis of the 4p16.3 Huntington disease gene region. *Am. J. Hum. Genet.*, **90**, 434–444.
7. Lee, J.M., Ramos, E.M., Lee, J.H., Gillis, T., Mysore, J.S., Hayden, M.R., Warby, S.C., Morrison, P., Nance, M., Ross, C.A., *et al.* (2012) CAG repeat expansion in Huntington disease determines age at onset in a fully dominant fashion. *Neurology*, **78**, 690–695.
8. Andrew, S.E., Goldberg, Y.P., Kremer, B., Telenius, H., Theilmann, J., Adam, S., Starr, E., Squitieri, F., Lin, B., Kalchman, M.A., *et al.* (1993) The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat. Genet.*, **4**, 398–403.
9. Duyao, M., Ambrose, C., Myers, R., Novelletto, A., Persichetti, F., Frontali, M., Folstein, S., Ross, C., Franz, M., Abbott, M., *et al.* (1993) Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nat. Genet.*, **4**, 387–392.
10. Snell, R.G., MacMillan, J.C., Cheadle, J.P., Fenton, I., Lazarou, L.P., Davies, P., MacDonald, M.E., Gusella, J.F., Harper, P.S. and Shaw, D.J. (1993) Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. *Nat. Genet.*, **4**, 393–397.
11. Keum, J.W., Shin, A., Gillis, T., Mysore, J.S., Abu Elneel, K., Lucente, D., Hadzi, T., Holmans, P., Jones, L., Orth, M., *et al.* (2016) The HTT CAG-Expansion Mutation Determines Age at Death but Not Disease Duration in Huntington Disease. *Am. J. Hum. Genet.*, **98**, 287–298.
12. Genetic Modifiers of Huntington's Disease (GeM-HD). Consortium. (2015) Identification of genetic factors that modify clinical onset of Huntington's disease. *Cell*, **162**, 516–526.
13. Cattaneo, E., Zuccato, C. and Tartari, M. (2005) Normal huntingtin function: an alternative approach to Huntington's disease. *Nat. Rev. Neurosci.*, **6**, 919–930.
14. Ross, C.A. and Tabrizi, S.J. (2011) Huntington's disease: from molecular pathogenesis to clinical treatment. *The Lancet. Neurol.*, **10**, 83–98.
15. Keiser, M.S., Kordasiewicz, H.B. and McBride, J.L. (2015) Gene suppression strategies for dominantly inherited neurodegenerative diseases: lessons from Huntington's disease and spinocerebellar ataxia. *Hum. Mol. Genet.*, **15**, R53–64.
16. Bonini, N.M. and La Spada, A.R. (2005) Silencing polyglutamine degeneration with RNAi. *Neuron*, **48**, 715–718.
17. 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.
18. Smith, R.A., Miller, T.M., Yamanaka, K., Monia, B.P., Condon, T.P., Hung, G., Lobsiger, C.S., Ward, C.M., McAlonis-Downes, M., Wei, H., *et al.* (2006) Antisense oligonucleotide therapy for neurodegenerative disease. *J. Clin. Invest.*, **116**, 2290–2296.
19. Carroll, J.B., Warby, S.C., Southwell, A.L., Doty, C.N., Greenlee, S., Skotte, N., Hung, G., Bennett, C.F., Freier, S.M. and Hayden, M.R. (2011) Potent and selective antisense oligonucleotides targeting single-nucleotide polymorphisms in the Huntington disease gene/allele-specific silencing of mutant huntingtin. *Mol. Ther.*, **19**, 2178–2185.
20. Yu, D., Pendergraff, H., Liu, J., Kordasiewicz, H.B., Cleveland, D.W., Swayze, E.E., Lima, W.F., Crooke, S.T., Prakash, T.P. and Corey, D.R. (2012) Single-stranded RNAs use RNAi to potently and allele-selectively inhibit mutant huntingtin expression. *Cell*, **150**, 895–908.
21. Lee, J.M., Kim, K.H., Shin, A., Chao, M.J., Abu Elneel, K., Gillis, T., Mysore, J.S., Kaye, J.A., Zahed, H., Kratter, I.H., *et al.* (2015) Sequence-Level Analysis of the Major European Huntington Disease Haplotype. *Am. J. Hum. Genet.*, **97**, 435–444.
22. Ambrose, C.M., Duyao, M.P., Barnes, G., Bates, G.P., Lin, C.S., Srinidhi, J., Baxendale, S., Hummerich, H., Lehrach, H., Altherr, M., *et al.* (1994) Structure and expression of the Huntington's disease gene: evidence against simple inactivation due to an expanded CAG repeat. *Somat. Cell Mol. Genet.*, **20**, 27–38.
23. Hsu, P.D., Lander, E.S. and Zhang, F. (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, **157**, 1262–1278.
24. 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.

25. Byrne, S.M., Ortiz, L., Mali, P., Aach, J. and Church, G.M. (2015) Multi-kilobase homozygous targeted gene replacement in human induced pluripotent stem cells. *Nucleic Acids Res.*, **43**, e21.

26. 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.

27. Consortium, H.D.i. (2012) Induced pluripotent stem cells from patients with Huntington's disease show CAG-repeat-expansion-associated phenotypes. *Cell Stem Cell*, **11**, 264–278.

28. Mattis, V.B., Tom, C., Akimov, S., Saeedian, J., Ostergaard, M.E., Southwell, A.L., Doty, C.N., Ornelas, L., Sahabian, A., Lenaeus, L., *et al.* (2015) HD iPSC-derived neural progenitors accumulate in culture and are susceptible to BDNF withdrawal due to glutamate toxicity. *Hum. Mol. Genet.*, **24**, 3257–3271.

29. Dragatsis, I., Levine, M.S. and Zeitlin, S. (2000) Inactivation of Hdh in the brain and testis results in progressive neurodegeneration and sterility in mice. *Nat. Genet.*, **26**, 300–306.

30. Duyao, M.P., Auerbach, A.B., Ryan, A., Persichetti, F., Barnes, G.T., McNeil, S.M., Ge, P., Vonsattel, J.P., Gusella, J.F., Joyner, A.L., *et al.* (1995) Inactivation of the mouse Huntington's disease gene homolog Hdh. *Science*, **269**, 407–410.

31. Zeitlin, S., Liu, J.P., Chapman, D.L., Papaioannou, V.E. and Efstratiadis, A. (1995) Increased apoptosis and early embryonic lethality in mice nullizygous for the Huntington's disease gene homologue. *Nat. Genet.*, **11**, 155–163.

32. Smith, C., Abalde-Atristain, L., He, C., Brodsky, B.R., Braunstein, E.M., Chaudhari, P., Jang, Y.Y., Cheng, L. and Ye, Z. (2015) Efficient and allele-specific genome editing of disease loci in human iPSCs. *Mol. Ther.*, **23**, 570–577.

33. 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.

34. Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C., Lawton, M., Trottier, Y., Lehrach, H., Davies, S.W., *et al.* (1996) Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell*, **87**, 493–506.

35. Sathasivam, K., Neueder, A., Gipson, T.A., Landles, C., Benjamin, A.C., Bondulich, M.K., Smith, D.L., Faull, R.L., Roos, R.A., Howland, D., *et al.* (2013) Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease. *Proc. Natl Acad. Sci. USA*, **110**, 2366–2370.

36. Yoshimi, K., Kaneko, T., Voigt, B. and Mashimo, T. (2014) Allele-specific genome editing and correction of disease-associated phenotypes in rats using the CRISPR-Cas platform. *Nat. Commun.*, **5**, 4240.

37. Bakondi, B., Lv, W., Lu, B., Jones, M.K., Tsai, Y., Kim, K.J., Levy, R., Akhtar, A.A., Breunig, J.J., Svendsen, C.N., *et al.* (2016) In Vivo CRISPR/Cas9 Gene Editing Corrects Retinal Dystrophy in the S334ter-3 Rat Model of Autosomal Dominant Retinitis Pigmentosa. *Mol. Ther.*, **24**, 556–563.

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. Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., *et al.* (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science*, **318**, 1917–1920.