

Research Article

Targeting N-Terminal Huntingtin with a Dual-sgRNA Strategy by CRISPR/Cas9

Junjiao Wu,^{1,2} Yu Tang ,¹ and Chun-Li Zhang ,^{3,4}

¹National Clinical Research Center for Geriatric Disorders, Department of Geriatrics, Xiangya Hospital, Central South University, Changsha, Hunan Province, China

²Department of Rheumatology and Immunology, Xiangya Hospital, Central South University, Changsha, Hunan Province, China

³Department of Molecular Biology, University of Texas Southwestern Medical Center, Dallas, TX, USA

⁴Hamon Center for Regenerative Science and Medicine, University of Texas Southwestern Medical Center, Dallas, TX, USA

Correspondence should be addressed to Yu Tang; tangyu-sky@163.com and Chun-Li Zhang; chun-li.zhang@utsouthwestern.edu

Received 30 April 2019; Revised 30 July 2019; Accepted 5 September 2019; Published 16 November 2019

Academic Editor: Marco Fichera

Copyright © 2019 Junjiao Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Huntington's disease (HD) is an autosomal dominant progressive neurodegenerative disorder, caused by a CAG/polyglutamine (polyQ) repeat expansion in the *Huntingtin* (HTT) gene. The polyQ tract is located in and transcribed from N-terminal HTT of exon 1. HTT is a large multifaceted protein, which participates in a range of cellular functions. Previous studies have shown that truncated HTT, which lacks N-terminus, retains specific functions that can produce neuroprotective benefits. It gives an insight that it is possible to repair HD by removing deleterious N-terminal HTT with CRISPR/Cas9, without compromising functions of remaining HTT peptides. To successfully generate functional truncated HTT proteins, an alternative downstream ATG start codon that is capable of initiating truncated HTT expression is required. In this study, we searched all possible in-frame ATGs before exon 7 and demonstrated that one of them can indeed initiate the downstream GFP expression in plasmids. We then tried to remove endogenous N-terminal HTT with an optimized dual-sgRNA strategy by CRISPR/Cas9; however, we cannot detect obvious traits of truncated HTT expression. Our results suggest that noncanonical ATGs of N-terminal HTT may not be effective in the genomic context, as in the construct context. Nevertheless, our study examined the therapeutic efficacy of downstream noncanonical ATGs for protein translation and also provided an optimized dual-sgRNA strategy for further genome manipulation of the HTT gene.

1. Introduction

Huntington's disease (HD) is an autosomal dominant neurodegenerative disease caused by the expanded CAG tract resided in the first exon of the *Huntingtin* (HTT) gene [1, 2]. Basically, the pathogenic mutant HTT consists of more than 35 CAG repeats, which are then translated into polyglutamine (polyQ) proteins [3, 4]. Although HTT is ubiquitously expressed in the body, expanded polyQ proteins may form intracellular aggregates and preferentially cause the loss of medium spiny neurons (MSNs) in the striatum, via a gain of toxic function [5]. The CAG tract length determines HTT propensity for aggregation and

toxicity and is inversely correlated with age of onset in HD patients [6].

HTT is a multiple conformation protein of normally 3144 amino acids and has conserved N-terminal sequences [7, 8]. By binding with other proteins, HTT is considered to be multifaceted that is essential for a spectrum of cellular functions, such as embryonic development, antiapoptotic pathway, BDNF modulation, ciliogenesis, autophagy, and vesicular transport [7]. N-terminal HTT containing the first 17 amino acids (N17) and the polyQ tract can be formed by proteolytic cleavage and is one of the most widely studied HTT peptides. HTT proteolysis in patient brains is the cascade of various cleavage events. Numerous proteases,

including caspases, calpains, cathepsins, matrix metalloproteinases, and aspartic proteases, that cleave HTT have been reported [9–13].

The N17 domain, as an evolutionarily conserved domain among vertebrates, functions as a nuclear export signal (NES) [14] and is necessary for nuclear exclusion of small mutant HTT fragments, thereby modifying nuclear pathogenesis and disease severity [15]. The N17 domain forms an amphipathic α -helical structure that signals endoplasmic reticulum (ER) stress [16] and also senses reactive oxygen species (ROS) stress, specifically through direct oxidation of a highly conserved methionine at position 8 [17]. Moreover, the polyQ tract can act as a flexible domain, allowing the flanking domains to come into close spatial proximity to form proper functional conformations [18]; however, this flexibility may be impaired with expanded polyQ tracts. Importantly, multiple posttranslational modifications, such as phosphorylation, ubiquitination, acetylation, and sumoylation, have been identified in the N17 domain, that may affect the clearance of HTT and its subcellular localization [19–21]. Thus, N-terminal HTT is considered to be important in regulating protein aggregation, subcellular localization, and cytotoxicity. Despite of this fact, other modification sites in the remaining HTT protein have been demonstrated to possess pivotal cellular functions as well [7]. For instance, sequential N-terminal proteolysis of HTT also generates a C-terminal fragment, which induces cytotoxicity via abnormal vacuolation of ER and increases ER stress [22]. Interestingly, without a whole stretch of N-terminus, however, HTT may also retain some beneficial functions [23, 24]. For example, Liu et al. reported that a truncated HTT lacking the first 237 amino acids spanning exon 1~6 can rescue the HTT deficiency-induced neuritic degeneration, which largely depends on the dynein-binding region known to be involved in intracellular trafficking [24].

So far, different strategies have been employed to produce therapeutic effects in both cellular and animal models of HD, including traditional avenues such as RNA interference (RNAi) [25–27] and antisense oligonucleotides (ASOs) [28–31], pharmacological treatments [32, 33], and the emerging genome editing by CRISPR/Cas9 both in vitro and in vivo [34–39]. A recent study by Southwell et al. developed ASOs targeting HD single-nucleotide polymorphisms (SNPs) that can selectively suppress mutant HTT expression in the mouse HD model [31]. They also achieved preclinical efficacy data of those HTT ASOs delivered to cynomolgus monkeys via lumbar puncture [31]. Similarly, two other recent studies have used an allele-selective CRISPR/Cas9 strategy based on protospacer adjacent motif (PAM) sites created by altering SNPs, thereby selectively inactivating mutant HTT allele for a given diplotype [34, 35]. As this is a quite novel personalized strategy, it requires a comprehensive analysis of patients' HTT haplotype patterns. The nonallele-selective inactivation of HTT has also been achieved by using a pair of sgRNAs flanking CAG repeats, in a transgenic HD mouse model [36], as well as in patient-derived fibroblasts with the help of Cas9 nickase [37]. CRISPR/Cas9 was also employed to correct the HTT mutation in an HD-induced pluripotent stem cell

(iPSC) line, which can rescue the HD phenotypes such as impaired neural rosette formation, increased susceptibility to growth factor withdrawal, and deficits in mitochondrial respiration [38]. Notably, CRISPR/Cas9 can permanently disrupt the targeted genes, rendering it to be a more efficient therapeutic approach than those with continuous delivery of drugs for treatment. However, it also brings about the caveats that the treatment cannot be reversed in the event of adverse effects.

Taking advantage of the higher efficiency by the CRISPR/Cas9 system, we hypothesize that we may remove a length of N-terminal HTT to abolish the neurotoxicity stressed by polyQ aggregations and meanwhile keep the functions of probably generated truncated HTT after CRISPR targeting. To successfully obtain this type of endogenous truncated HTT, it is required that an alternative downstream ATG start codon will be capable of initiating truncated HTT translation, albeit at lower levels. Therefore, in this study, we tried to search all possible in-frame ATGs within the first 6 exons that coded for 249 amino acids, which were then cloned to fuse with GFP in plasmids, and we demonstrated that one of them can indeed start downstream translation. However, after removing endogenous N-terminal HTT by an optimized dual-sgRNA targeting, we failed to detect obvious traits of truncated HTT expression. Our results suggest that noncanonical ATGs of N-terminal HTT may not be effective in the genomic context, as in the construct context.

2. Materials and Methods

2.1. Cloning and Sequencing. Constructs of E3a-, E3b-, E4-, and E6-GFP were generated based on the pcDNA3.1(+) vector using Gibson Assembly Master Mix (New England Biolabs, MA, USA). Specifically, different lengths of HTT fragments were amplified from a human cDNA library with 21 CAGs located in N-terminus. E3a Δ -, E3b Δ -, E4 Δ -, and E6 Δ -GFP plasmids were then, respectively, constructed by PCR using primers flanking exon 2 and 5'UTR from their corresponding plasmids and finally self-ligated into circle plasmids. The length of 5'UTR in all constructs is 145 bp, which is exactly the full-length of 5'UTR in the NM_002111 transcript. All pcDNA constructs were finally confirmed by Sanger sequencing. For genome editing, lentiCRISPRv2 (Addgene #52961) was used as the backbone, and sgHTT sequences were designed with the help of CHOPCHOP website (<http://chopchop.cbu.uib.no/>). In brief, sgRNA oligos were annealed and inserted into the BsmBI digested vector. Junctions of exon 1/2 or targeted genomic regions were amplified by PCR and ligated into the pMD20T vector (Takara Bio Inc., CA, USA) for T-A cloning. White colonies were picked out for plasmid extraction and Sanger sequencing. Sequences of sgHTT and PCR primers are listed in Table 1.

2.2. Cell Culture. HEK293 cells were maintained in the DMEM medium (Thermo Fisher Scientific, MA, USA) supplemented with 10% fetal bovine serum (Corning, NY,

TABLE 1: Sequences of sgRNAs and PCR primers.

Sequences (5'-3')	
<i>sgRNA sequences</i>	
sgHTT-1	GAGTCGGCCCCGAGGCCTCCG
sgHTT-2	GCCTCCGGGGACTGCCGTGCC
sgHTT-3	GCGGGGACTGCCGTGCCGGGC
sgHTT-4	GTGCCGGGCGGGAGACCCCA
sgHTT-5	GTGAGGAAGCTGAGGAGGCGG
sgHTT-6	GGCTGAGGAAGCTGAGGAGG
sgHTT-7	GGCGGCGGCTGAGGAAGCTG
sgHTT-8	GAGCAGCGGCTGTGCCTGCCG
sgHTT-9	GAGCGGCTCCTCAGCCACAGC
<i>T7E1 assay and junction PCR</i>	
HTT-T7E1-F	CCAGCCATTGGCAGAGTCCG
HTT-T7E1-R	TTGCTGGGTCACTCTGTCTCTG
<i>Quantitative PCR</i>	
HTT-qF	TGACGCAGAGTCAGATGTCAG
HTT-qR	CCGAGGGGCACCATTCTTTT
GAPDH-qF	CGGGATTGTCTGCCCTAATTAT
GAPDH-qR	GCACGGAAGGTCACGATGT
<i>Sequencing of junctions</i>	
HTT-5'UTR/e1-F	CGGGAGACCGGTGGCTGAGG
HTT-e4/e5-R	CCGAGGGGCACCATTCTTTT

USA) and 1% penicillin/streptomycin at 37°C and 5% CO₂. For transfection, passage and seed HEK293 cells at a density of 2×10^5 cells per well of a 6-well plate. The same amount of 0.5 μ g constructed pcDNA plasmids were then transfected by the polyethylenimine (PEI) reagent. GFP signals and protein aggregations were observed 24 hrs later. For CRISPR targeting, HEK293 cells were transfected with different sgRNA pairs (0.5 μ g/0.5 μ g mixture) by the PEI reagent. After 48 hrs, collect HEK293 cells for indel analysis by the T7E1 assay. For clonal cell line selection, passage targeted cells with 0.25% trypsin and further dilute and seed 6000 single cells in a 10 cm dish. Single colonies were picked out under a microscopy and further examined by junction PCR. PCR primers are listed in Table 1.

2.3. T7E1 Assay. The cleavage efficiency of dual sgHTT was examined by the T7E1 assay accordingly to the previously reported method [40]. In brief, transfected cells were digested and spun down at 250 g for 4 min and later lysed with DNA lysis buffer (100 mM Tris-HCl (pH 8.0), 25 mM EDTA (pH 8.0), 100 mM NaCl, 1% Triton X-100, and 50 ng/ μ l RNase A) containing 0.5 mg/ml Proteinase K at 55°C for overnight. DNA lysates were used directly as templates for genomic PCR. T7E1 primers are listed in Table 1. PCR products were then melted and reannealed to form heteroduplexes, which were then digested by T7E1 (New England Biolabs) and electrophoresed in 2% agarose gels. Calculate the fraction of the PCR product cleaved (f_{cut}) by using the following formula: $f_{cut} = (b + c)/(a + b + c)$, where a is the integrated intensity of the undigested PCR product and b and c are the integrated intensities of each cleavage product. Indel occurrence was estimated with the following formula: indel (%) = $100 \times (-\sqrt{1 - f_{cut}})$.

2.4. Gene Expression. Gene expression was performed accordingly to the previously reported method [41]. Total mRNAs were isolated from cultured cells using TRIzol (Invitrogen, Carlsbad, CA, USA) and were reverse transcribed into cDNAs using a SuperScript III kit (Invitrogen). Quantitative PCR was then performed using SYBR GreenER SuperMix (Invitrogen) on an ABI 7900HT thermocycler (Applied Biosystems). The PCR program consisted of 2 min at 50°C and 10 min at 95°C, followed by 40 cycles of 15 sec at 95°C and 45 sec at 62°C. Relative expression levels were analyzed using the $2^{-\Delta\Delta Ct}$ algorithm normalized to the housekeeping gene GAPDH. Primer sequences are listed in Table 1.

2.5. Western Blot. Total proteins were extracted from cultured cells and homogenized mice brain tissues, respectively, with RIPA lysis buffer. Proteins were then separated by SDS-PAGE with Tris-glycine gels (8% for HTT; 15% for GFP) and transferred onto PVDF membranes (Millipore, MA, USA). Nonspecific sites were blocked in 5% nonfat milk for 1 hr at room temperature. The blots were then probed at 4°C for overnight with primary antibodies against the following proteins: HTT (1:1 000; D7F7, #5656, cell signaling), GFP (1:5 000; GFP-1020, Aves Labs), and α -tubulin (1:10 000; T5168, Sigma). After washing with PBS for 3 times, secondary antibodies conjugated with horseradish peroxidase (1:25 000) from Jackson ImmunoResearch (711-035-152, 715-035-150, and 703-035-155) were incubated, respectively, with blots for 2 hrs at room temperature. Protein bands were finally visualized with the Pierce ECL substrate (Thermo Fisher Scientific).

2.6. Statistical Analysis. Results were presented as means \pm SD values. Statistical significance ($p < 0.05$) was determined using Student's t -test.

3. Results and Discussion

3.1. Capability of Initiating Translation by Noncanonical ATGs. Functional truncated HTT is lack of N-terminal 237 amino acids, which are translated from within the first 6 exons [24]. To figure out any possible alternative ATG start codon that may initiate its downstream protein translation and produce functional truncated HTT, we first searched all in-frame ATGs before exon 7. We then identified four candidate ATGs, among which, two are located on exon 3 (E3a-ATG and E3b-ATG), one on exon 4 (E4-ATG), and one on exon 6 (E6-ATG) (Figure 1(a)). Interestingly, these candidate ATGs mostly conform to the consensus Kozak rule NNN(A/G)NNATGG for eukaryotic cells, suggestive of strong initiation potentials (Figure 1(a)).

To test which of these ATGs can indeed start downstream protein translation, we then constructed different lengths of HTT fragments harboring candidate ATGs, respectively, and fused them with GFP in plasmids (Figure 1(b)). Constructs of E3a-, E3b-, E4-, and E6-GFP included exon 1 and thus preferentially used E1-ATG to start their translation. However, exon 1 and E1-ATG were all

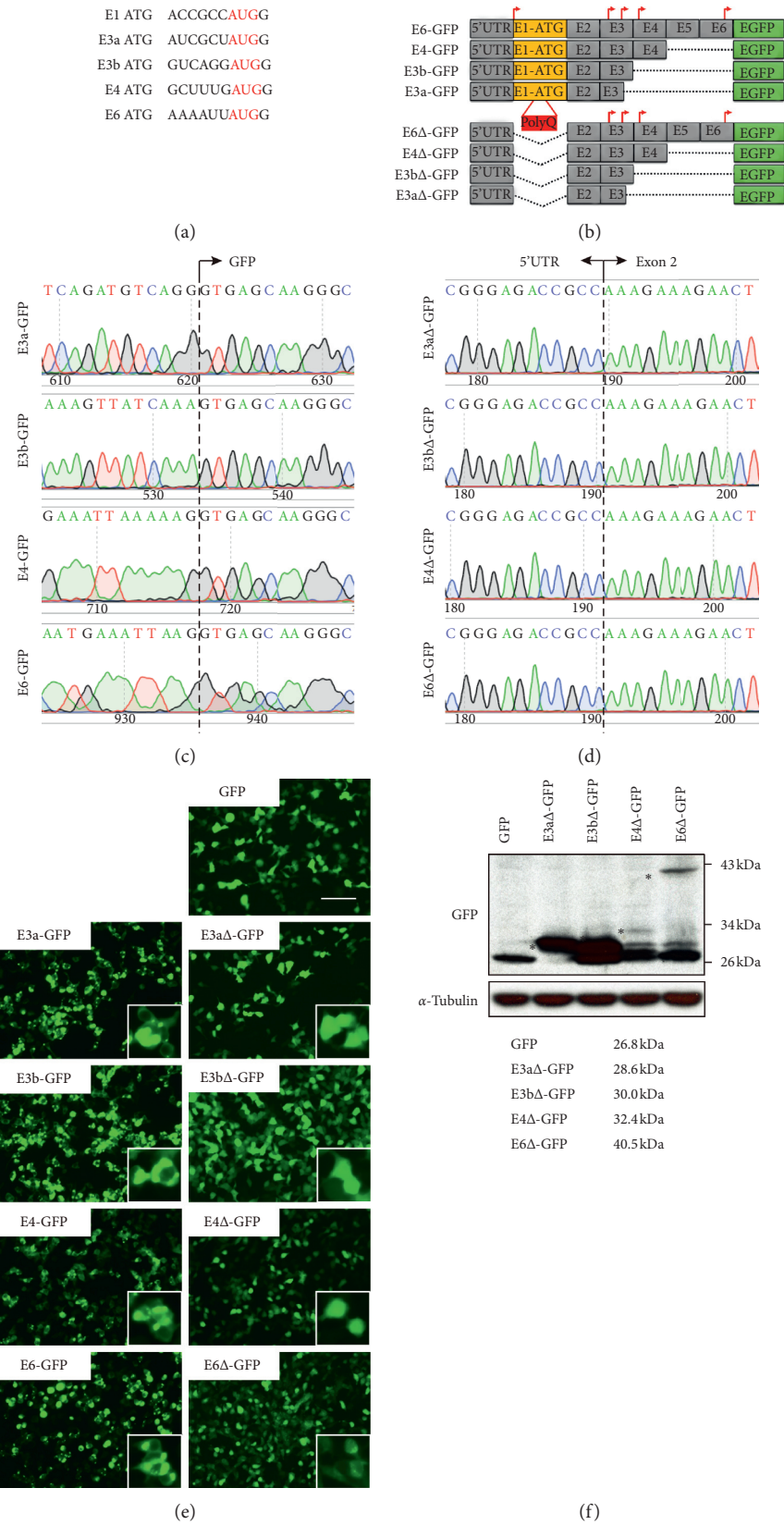


FIGURE 1: Capability of initiating translation by noncanonical in-frame ATGs. (a) Candidate in-frame ATGs before exon 7 of HTT. (b) Constructs of HTT fragments fused with GFP. (c, d) Sanger sequencing results of junctions in plasmid constructs. (e) Transfection of constructs in HEK293 cells. Scale bar, 100 μm. (f) Western blot analysis of fused proteins. Asterisks represented the E3-ATG initiated translation products.

deleted in E3a Δ -, E3b Δ -, E4 Δ -, and E6 Δ -GFP constructs (Figure 1(b)). All constructs retained the 5' untranslated region (5' UTR) of HTT to maintain the upstream regulatory machinery. Sanger sequencing of those constructs confirmed that ATG codons of GFP were all removed (Figure 1(c)) and that 5'UTR and exon 2 were seamlessly ligated in truncated constructs (Figure 1(d)).

After transfected into HEK293 cells with the same amount, constructs containing exon 1 did produce strong GFP signals and also led to extensive polyQ aggregations in the cytoplasm (Figure 1(e)). This is in line with previous studies that the NES located on the N17 region of exon 1 exported HTT from the nucleus to cytoplasm [14]. Interestingly, all of the truncated constructs (E3a Δ -, E3b Δ -, E4 Δ -, and E6 Δ -GFP) can produce GFP signals evenly in the nucleus and cytoplasm, as those in the GFP-alone control group (Figure 1(e)). Furthermore, western blot analysis showed that different lengths of HTT fragments were successfully translated and fused with downstream GFP although the productivities of E4 Δ -GFP and E6 Δ -GFP were greatly decreased (Figure 1(f)). Based on the sizes of protein bands (28.6 kDa for E3a Δ -GFP, 30.0 kDa for E3b Δ -GFP, 32.4 kDa for E4 Δ -GFP, and 40.5 kDa for E6 Δ -GFP), we demonstrated that E3a-ATG is responsible for the translation of those fused proteins (Figure 1(f) and Figure 2). The unspecific bands were probably derived from proteolytic cleavage of translated peptides. Taken together, these results suggest that alternative in-frame ATGs can initiate downstream GFP expression in plasmids.

3.2. Targeting N-Terminal HTT with a Dual-sgRNA Strategy.

N-Terminal HTT is transcribed from exon 1 just after 5'UTR, containing the conserved N17 region, polyQ, and polyproline (polyP) tract. In light of the initiation capability of downstream ATGs in plasmids, we proposed to endogenously delete the polyQ tract from N-terminal HTT by CRISPR/Cas9, in hope of maintaining the functions of remaining HTT peptides.

First, we designed a range of sgRNAs flanking the polyQ region, with 4 upstream sgRNAs and 5 downstream sgRNAs (Figure 3(a) and Table 1). We then used different combinations of sgRNA pairs to transfect cells and later examine their combined cleavage efficiencies by the T7E1 assay. Although this analysis would be a mixture of independent cleavage events induced by each single sgRNA and deletion events induced by the dual sgRNAs, we can nevertheless figure out the sgRNA pairs with higher cleavage efficiencies. Of those, upstream sgHTT-1 and sgHTT-4 combined with downstream sgRNAs were shown to be highly efficient in cutting genomic DNA than those with sgHTT-2 and sgHTT-3 (Figure 3(b)). To retain 5'UTR regulatory elements at the most and avoid the possible regulatory interference by uncut polyP sequences, we thus chose sgHTT-4/9 as the optimized sgRNA pair. We then transfected HEK293 cells and later screened out 12 single colonies. We did junction PCR and finally selected one cell colony (sgHTT-4/9 #7), which is shown to be homozygous, for further analysis (Figure 3(c)). We then sequenced #7

gRNA and found that polyQ was seamlessly deleted from N-terminal HTT (Figure 3(d)).

Since the downstream sgHTT target sites are adjacent to the exon 1/intron 1 boundary, this may disrupt the exon/intron junction by indels and then affect the RNA splicing process. A previous study has demonstrated that targeting the junction of HTT exon 1/intron 1 with CRISPR/Cas9 can reduce half of the mRNA level in the YAC128 mouse model [39]. We thus harvested mRNAs from cell colonies and reverse-transcribed them into cDNAs for integrity analysis. We sequenced #7 cDNA and found that the junction of exon 1/exon 2 was still intact, suggesting that no indels occurred in this region (Figure 3(d)). We next did quantitative PCR, and the results showed a ~23% reduction of the HTT level in colony #7 and a ~54% reduction in colony #10 (Figure 3(e)). This result echoes the heterozygous property of colony #10 but, meanwhile, documents that dual-sgHTT targeting in this colony may destroy HTT transcripts by certain mechanisms such as nonsense-mediated decay (NMD) that will eliminate mRNAs containing premature stop codons [42]. Moreover, sgHTT-9 targeting may have produced indels to disrupt the exon 1/intron 1 junction in colony #10, finally dampening the transcription.

We hence extracted proteins and performed western blot to examine the HTT expression. We can observe full-length HTT band and also the cleaved fragments in WT HEK293 cells (Figure 3(f)). Unexpectedly, in the sgHTT-4/9-#7 cell colony, there were no obviously visible HTT bands including any possible truncated HTT bands, suggesting the ineffective or extremely weak initiation capability of E3a-ATG in this endogenous condition (Figure 3(f)). Note that the primary antibody against HTT was produced by immunizing animals with a synthetic peptide corresponding to residues surrounding Pro1220 of human HTT protein, and it is unlikely that truncated HTT proteins cannot be recognized by this primary antibody.

Collectively, our results suggest that downstream non-canonical ATGs may not be effective in the genomic context, which is quite different from their capability in plasmid constructs. Several factors might be accounted for this discrepancy, leading to a failed proof of concept to exploit the possible functions of truncated HTT. First, transient transfection of HTT-exons-GFP constructs basically results in really high-level expression of proteins although we only transfected small amounts of plasmids. Also, CMV is a very strong promoter in constructs that may accelerate this massive expression and lead to the formation of cytoplasmic polyQ aggregates. Notably, the translated WT polyQs are within truncated peptides but not within the full-length HTT protein. In this truncated context, the conformation of translated peptides harboring WT polyQ might be changed and leads to form aggregates more easily. On the contrary, transfection of the E3a Δ -GFP construct with this strong promoter only produced small amount of proteins. This partially suggests that E3a-ATG is indeed far less efficient than the canonical E1-ATG start codon. Similarly, in the native genomic context, noncanonical E3a-ATG also has far lower efficiency to produce enough truncated proteins to be discoverable by the western blot assay, and in this case, the

E6Δ::GFP: 40.5 kDa	1	MELFLLCSDDAESDVRMVADECLNKV[KALMDSNLPRLQLELYKEIKKNGAPRSLRAALW
E4Δ::GFP: 32.4 kDa	1	MELFLLCSDDAESDVRMVADECLNKV[KALMDSNLPRLQLELYKEIK-----
E3bΔ::GFP: 30.0 kDa	1	MELFLLCSDDAESDVRMVADECLNKV[-----
E3aΔ::GFP: 28.6 kDa	1	MELFLLCSDDAESDVR[-----
GFP: 26.8 kDa	1	-----
E6Δ::GFP: 40.5 kDa	61	RFAELAHLVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMASFGNFANDNEI
E4Δ::GFP: 32.4 kDa	48	-----
E3bΔ::GFP: 30.0 kDa	28	-----
E3aΔ::GFP: 28.6 kDa	17	-----
GFP: 26.8 kDa	1	-----
E6Δ::GFP: 40.5 kDa	121	KVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWP
E4Δ::GFP: 32.4 kDa	48	KVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWP
E3bΔ::GFP: 30.0 kDa	28	KVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWP
E3aΔ::GFP: 28.6 kDa	17	--VSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWP
GFP: 26.8 kDa	1	--VSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWP
E6Δ::GFP: 40.5 kDa	181	LVTTLTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL
E4Δ::GFP: 32.4 kDa	108	LVTTLTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL
E3bΔ::GFP: 30.0 kDa	88	LVTTLTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL
E3aΔ::GFP: 28.6 kDa	76	LVTTLTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL
GFP: 26.8 kDa	60	LVTTLTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL
E6Δ::GFP: 40.5 kDa	241	VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA
E4Δ::GFP: 32.4 kDa	168	VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA
E3bΔ::GFP: 30.0 kDa	148	VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA
E3aΔ::GFP: 28.6 kDa	136	VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA
GFP: 26.8 kDa	120	VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA
E6Δ::GFP: 40.5 kDa	301	DHYQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYK*
E4Δ::GFP: 32.4 kDa	228	DHYQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYK*
E3bΔ::GFP: 30.0 kDa	208	DHYQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYK*
E3aΔ::GFP: 28.6 kDa	196	DHYQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYK*
GFP: 26.8 kDa	180	DHYQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYK*

FIGURE 2: Translation products of HTT : GFP constructs initiated by E3a-ATG. Translation ORFs were aligned using EMBL-EBI multiple sequence alignment program (T-Coffee, <http://www.ebi.ac.uk/Tools/msa/tcoffee/>) and viewed by BoxShade Server (https://embnet.vital-it.ch/software/BOX_form.html).

potency of truncated HTT would also be far below therapeutic requirements. Second, the complexity of endogenous regulatory system flanking N-terminal HTT is really different from the simple well-known plasmid elements. A more comprehensive analysis of endogenous regulatory elements and their orchestration need to be considered.

By using the knock-in technology with the CRISPR system, further studies may perform gene engineering to keep the strong E1-ATG after 5'UTR that directly ligates with downstream in-frame HTT sequences excluding CAG

repeats. Alternatively, it might be possible to drive translation from the E3a-ATG by adding some upstream regulatory elements. And for sure, pathogenic N-terminal HTT can also be replaced by a WT fragment with the similar knock-in method to correct the mutations. The concern is that the efficacy of knock-in mediated by homologous recombination (HR) is much lower than that by non-homologous end joining (NHEJ). For in vivo gene targeting or gene therapy, it may not be optimal to deliver a stretch of DNA donor together with CRISPR vectors. In this study, we

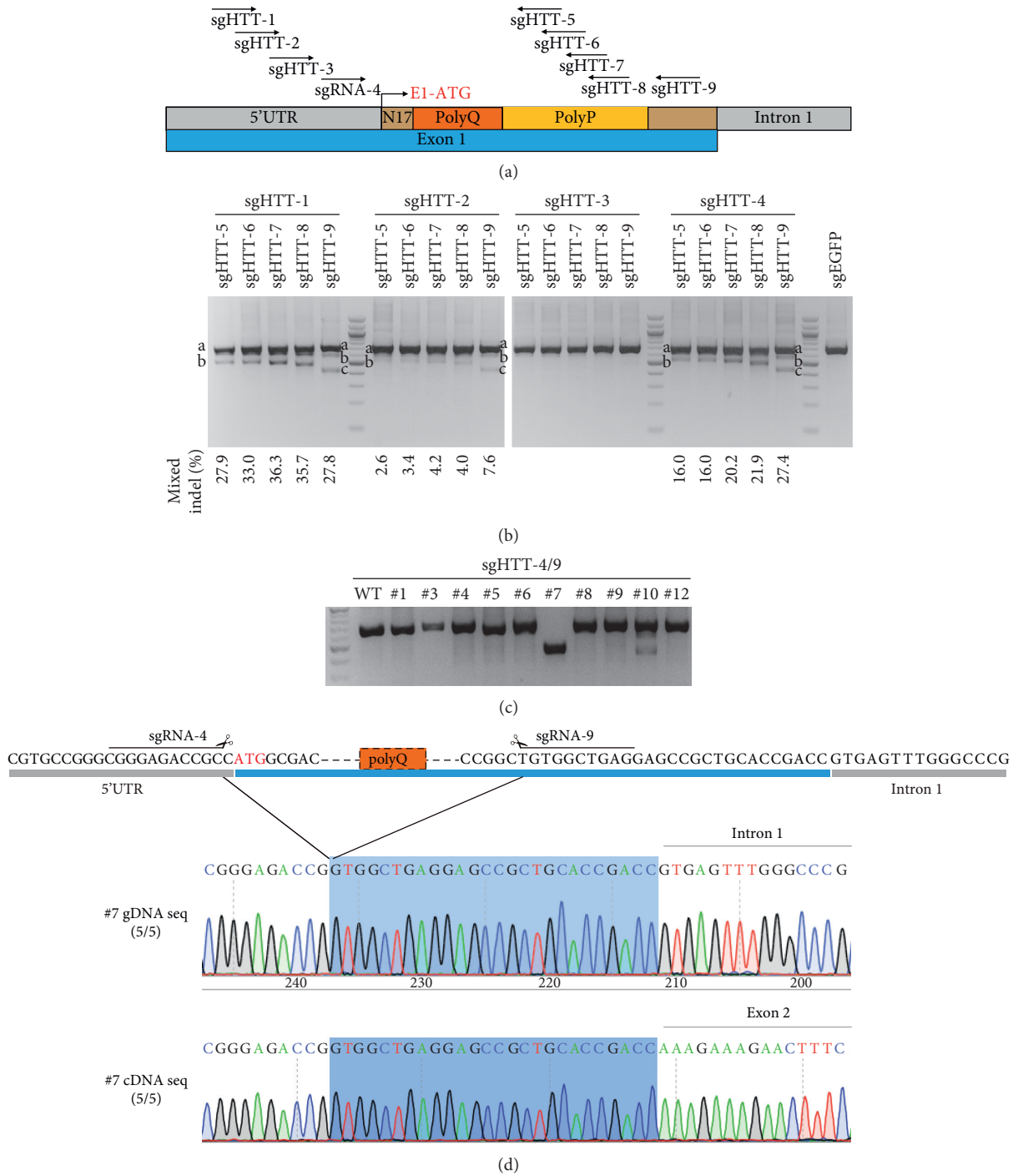


FIGURE 3: Continued.

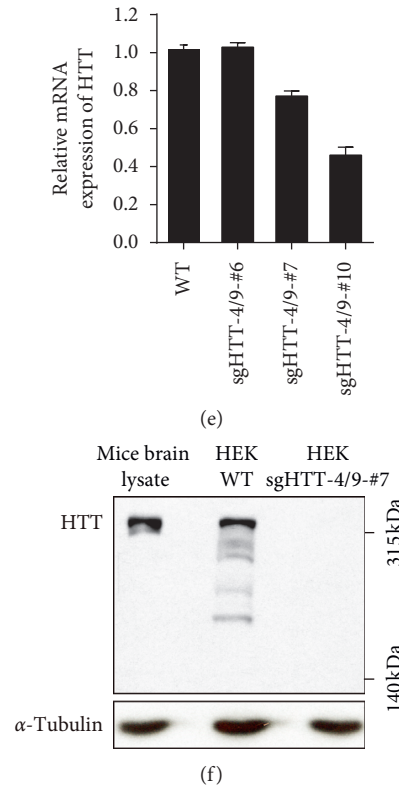


FIGURE 3: Targeting N-terminal HTT with a dual-sgRNA strategy. (a) Design of sgRNAs flanking the polyQ region. (b) Detection of cleavage efficiencies with different combinations of sgRNA pairs by the T7E1 assay. Mixed indels were calculated based on the fractions of PCR products. “a” represents the undigested PCR product, and “b” and “c” represent cleavage products. (c) Single colonies screening by junction PCR after sgHTT-4/9 targeting. (d) Sanger sequencing results of gDNA and cDNA junctions in the sgHTT-4/9-#7 cell colony. (e) Measurement of HTT levels by quantitative PCR. (f) Western blot detection of full-length HTT and possible truncated HTT proteins.

originally aimed to figure out noncanonical ATGs that may initiate truncated HTT translation, in which case we can turn detrimental HTT directly into functional ones only by NHEJ mechanisms. Since E3a-ATG appears not to be strong enough in the genomic context, these knock-in strategies currently may not be promising for in vivo gene therapy. However, these generated knock-in cells would be helpful for future studies in dissecting truncated HTT functions. Overall, our study examined the therapeutic efficacy of downstream noncanonical ATGs for HTT translation and also provided an optimized dual-sgRNA strategy for further genome manipulation of the HTT gene.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

YT and CLZ conceived this study. JW and YT performed all the experiments. JW and YT prepared and revised the

manuscript. All authors reviewed and approved the final manuscript.

Acknowledgments

This study was supported by grants from the National Natural Science Foundation of China (No. 81801200) and Hunan Provincial Natural Science Foundation of China (No. 2019JJ40476).

References

- [1] H. M. Yang, S. Yang, S. S. Huang, B. S. Tang, and J. F. Guo, “Microglial activation in the pathogenesis of Huntington’s disease,” *Frontiers in Aging Neuroscience*, vol. 9, p. 193, 2017.
- [2] J. M. Burgunder, “Translational research in Huntington’s disease: opening up for disease modifying treatment,” *Translational Neurodegeneration*, vol. 2, no. 1, p. 2, 2013.
- [3] E. Cattaneo, C. Zuccato, and M. Tartari, “Normal huntingtin function: an alternative approach to Huntington’s disease,” *Nature Reviews Neuroscience*, vol. 6, no. 12, pp. 919–930, 2005.
- [4] H. Jiang, Y. M. Sun, Y. Hao et al., “Huntingtin gene CAG repeat numbers in Chinese patients with Huntington’s disease and controls,” *European Journal of Neurology*, vol. 21, no. 4, pp. 637–642, 2014.
- [5] I. Han, Y. You, J. H. Kordower, S. T. Brady, and G. A. Morfini, “Differential vulnerability of neurons in Huntington’s disease:

- the role of cell type-specific features,” *Journal of Neurochemistry*, vol. 113, no. 5, pp. 1073–1091, 2010.
- [6] S. E. Andrew, Y. Paul Goldberg, B. Kremer et al., “The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington’s disease,” *Nature Genetics*, vol. 4, no. 4, pp. 398–403, 1993.
 - [7] F. Saudou and S. Humbert, “The biology of huntingtin,” *Neuron*, vol. 89, no. 5, pp. 910–926, 2016.
 - [8] X. Feng, S. Luo, and B. Lu, “Conformation polymorphism of polyglutamine proteins,” *Trends in Biochemical Sciences*, vol. 43, no. 6, pp. 424–435, 2018.
 - [9] R. K. Graham, Y. Deng, E. J. Slow et al., “Cleavage at the caspase-6 site is required for neuronal dysfunction and degeneration due to mutant huntingtin,” *Cell*, vol. 125, no. 6, pp. 1179–1191, 2006.
 - [10] J. Gafni, E. Hermel, J. E. Young, C. L. Wellington, M. R. Hayden, and L. M. Ellerby, “Inhibition of calpain cleavage of huntingtin reduces toxicity,” *Journal of Biological Chemistry*, vol. 279, no. 19, pp. 20211–20220, 2004.
 - [11] A. Lunkes, K. S. Lindenberg, L. Ben-Haiem et al., “Proteases acting on mutant huntingtin generate cleaved products that differentially build up cytoplasmic and nuclear inclusions,” *Molecular Cell*, vol. 10, no. 2, pp. 259–269, 2002.
 - [12] J. P. Miller, J. Holcomb, I. Al-Ramahi et al., “Matrix metalloproteinases are modifiers of huntingtin proteolysis and toxicity in Huntington’s disease,” *Neuron*, vol. 67, no. 2, pp. 199–212, 2010.
 - [13] A. L. Southwell, C. W. Bugg, L. S. Kaltenbach et al., “Perturbation with intrabodies reveals that calpain cleavage is required for degradation of huntingtin exon 1,” *PLoS One*, vol. 6, no. 1, Article ID e16676, 2011.
 - [14] Z. Zheng, A. Li, B. B. Holmes, J. C. Marasa, and M. I. Diamond, “An N-terminal nuclear export signal regulates trafficking and aggregation of Huntingtin (Htt) protein exon 1,” *Journal of Biological Chemistry*, vol. 288, no. 9, pp. 6063–6071, 2013.
 - [15] X. Gu, J. P. Cante, E. R. Greiner et al., “N17 Modifies mutant huntingtin nuclear pathogenesis and severity of disease in HD BAC transgenic mice,” *Neuron*, vol. 85, no. 4, pp. 726–741, 2015.
 - [16] R. S. Atwal, J. Xia, D. Pinchev, J. Taylor, R. M. Eband, and R. Truant, “Huntingtin has a membrane association signal that can modulate huntingtin aggregation, nuclear entry and toxicity,” *Human Molecular Genetics*, vol. 16, no. 21, pp. 2600–2615, 2007.
 - [17] L. F. DiGiovanni, A. J. Mocle, J. Xia, and R. Truant, “Huntingtin N17 domain is a reactive oxygen species sensor regulating huntingtin phosphorylation and localization,” *Human Molecular Genetics*, vol. 25, no. 18, pp. 3937–3945, 2016.
 - [18] N. S. Caron, C. R. Desmond, J. Xia, and R. Truant, “Polyglutamine domain flexibility mediates the proximity between flanking sequences in huntingtin,” *Proceedings of the National Academy of Sciences*, vol. 110, no. 36, pp. 14610–14615, 2013.
 - [19] J. S. Steffan, N. Agrawal, J. Pallos et al., “SUMO modification of Huntingtin and Huntington’s disease pathology,” *Science*, vol. 304, no. 5667, pp. 100–104, 2004.
 - [20] L. M. Thompson, C. T. Aiken, L. S. Kaltenbach et al., “IKK phosphorylates huntingtin and targets it for degradation by the proteasome and lysosome,” *The Journal of Cell Biology*, vol. 187, no. 7, pp. 1083–1099, 2009.
 - [21] X. Cong, J. M. Held, F. DeGiacomo et al., “Mass spectrometric identification of novel lysine acetylation sites in huntingtin,” *Molecular & Cellular Proteomics*, vol. 10no. 10, Article ID M111.009829, 2011.
 - [22] M. T. El-Daher, E. Hangen, J. Bruyere et al., “Huntingtin proteolysis releases non-polyQ fragments that cause toxicity through dynamin 1 dysregulation,” *The EMBO Journal*, vol. 34, no. 17, pp. 2255–2271, 2015.
 - [23] G. Wang, X. Liu, M. A. Gaertig, S. Li, and X.-J. Li, “Ablation of huntingtin in adult neurons is nondeleterious but its depletion in young mice causes acute pancreatitis,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 12, pp. 3359–3364, 2016.
 - [24] X. Liu, C. E. Wang, Y. Hong et al., “N-terminal huntingtin knock-in mice: implications of removing the N-terminal region of huntingtin for therapy,” *PLoS Genetics*, vol. 12, no. 5, Article ID e1006083, 2016.
 - [25] E. L. Pfister, L. Kennington, J. Straubhaar et al., “Five siRNAs targeting three SNPs may provide therapy for three-quarters of Huntington’s disease patients,” *Current Biology*, vol. 19, no. 9, pp. 774–778, 2009.
 - [26] M. Takahashi, S. Watanabe, M. Murata et al., “Tailor-made RNAi knockdown against triplet repeat disease-causing alleles,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 50, pp. 21731–21736, 2010.
 - [27] S. Aguiar, B. van der Gaag, and F. A. B. Cortese, “RNAi mechanisms in Huntington’s disease therapy: siRNA versus shRNA,” *Translational Neurodegeneration*, vol. 6, p. 30, 2017.
 - [28] H. B. Kordasiewicz, L. M. Stanek, E. V. Wancewicz et al., “Sustained therapeutic reversal of Huntington’s disease by transient repression of huntingtin synthesis,” *Neuron*, vol. 74, no. 6, pp. 1031–1044, 2012.
 - [29] J. B. Carroll, S. C. Warby, A. L. Southwell et al., “Potent and selective antisense oligonucleotides targeting single-nucleotide polymorphisms in the Huntington disease gene/allele-specific silencing of mutant huntingtin,” *Molecular Therapy*, vol. 19, no. 12, pp. 2178–2185, 2011.
 - [30] M. E. Østergaard, A. L. Southwell, H. Kordasiewicz et al., “Rational design of antisense oligonucleotides targeting single nucleotide polymorphisms for potent and allele selective suppression of mutant huntingtin in the CNS,” *Nucleic Acids Research*, vol. 41, no. 21, pp. 9634–9650, 2013.
 - [31] A. L. Southwell, H. B. Kordasiewicz, D. Langbehn et al., “Huntingtin suppression restores cognitive function in a mouse model of Huntington’s disease,” *Science Translational Medicine*, vol. 10no. 461, Article ID eaar3959, 2018.
 - [32] M. Yu, Y. Fu, Y. Liang et al., “Suppression of MAPK11 or HIPK3 reduces mutant Huntingtin levels in Huntington’s disease models,” *Cell Research*, vol. 27, no. 12, pp. 1441–1465, 2017.
 - [33] C. Pidgeon and H. Rickards, “The pathophysiology and pharmacological treatment of Huntington disease,” *Behavioural Neurology*, vol. 26, no. 4, pp. 245–253, 2013.
 - [34] J. W. Shin, K. H. Kim, M. J. Chao et al., “Permanent inactivation of Huntington’s disease mutation by personalized allele-specific CRISPR/Cas9,” *Human Molecular Genetics*, vol. 25, no. 20, pp. 4566–4576, 2016.
 - [35] A. M. Monteys, S. A. Ebanks, M. S. Keiser, and B. L. Davidson, “CRISPR/Cas9 editing of the mutant huntingtin allele in vitro and in vivo,” *Molecular Therapy*, vol. 25, no. 1, pp. 12–23, 2017.
 - [36] S. Yang, R. Chang, H. Yang et al., “CRISPR/Cas9-mediated gene editing ameliorates neurotoxicity in mouse model of Huntington’s disease,” *Journal of Clinical Investigation*, vol. 127, no. 7, pp. 2719–2724, 2017.
 - [37] M. Dabrowska, W. Juzwa, W. J. Krzyzosiak et al., “Precise excision of the CAG tract from the huntingtin gene by Cas9 nickases,” *Frontiers in Neuroscience*, vol. 12, p. 75, 2018.

- [38] X. Xu, Y. Tay, B. Sim et al., "Reversal of phenotypic abnormalities by CRISPR/Cas9-mediated gene correction in Huntington disease patient-derived induced pluripotent stem cells," *Stem Cell Reports*, vol. 8, no. 3, pp. 619–633, 2017.
- [39] N. Kolli, M. Lu, P. Maiti et al., "CRISPR-Cas9 mediated gene-silencing of the mutant huntingtin gene in an in vitro model of Huntington's disease," *International Journal of Molecular Sciences*, vol. 18, no. 4, p. 754, 2017.
- [40] Z. Hu, Z. Shi, X. Guo et al., "Ligase IV inhibitor SCR7 enhances gene editing directed by CRISPR-Cas9 and ssODN in human cancer cells," *Cell & Bioscience*, vol. 8, no. 1, p. 12, 2018.
- [41] Y. Tang, M. L. Liu, T. Zang et al., "Direct reprogramming rather than iPSC-based reprogramming maintains aging hallmarks in human motor neurons," *Frontiers in Molecular Neuroscience*, vol. 10, p. 359, 2017.
- [42] S. Brogna and J. Wen, "Nonsense-mediated mRNA decay (NMD) mechanisms," *Nature Structural & Molecular Biology*, vol. 16, no. 2, pp. 107–113, 2009.