Cas9 Nickase-Mediated Contraction of CAG/CTG Repeats *in Vivo* is Accompanied by Improvements in Huntington's Disease Pathology

Alvaro Murillo[1], Melanie Alpaugh[2], Ruban Rex Peter Durairaj[1], Emma L. Randall[1], Meghan Larin[1,3], Laura Heraty[1], Alys N. Aston[1], Alysha S. Taylor[1], Alex Mas Monteys[4], Nina Stöberl[5], Aeverie E. R. Heuchan[1], Pascale Aeschlimann[1], Kyle Fears[1], Christian Landles[6], Georgina F Osborne[6], Antoine Mangin[1], Soumyasree Bhattacharyya[1], Emmanouil Metzakopian[7], Nicholas D. Allen[5], Jack Puymirat[8], Gillian P Bates[6], Beverly L. Davidson[4,9], Francesca Cicchetti[10], Mariah J. Lelos[5], and Vincent Dion[1,11]*

1: UK Dementia Research Institute at Cardiff University, Hadyn Ellis Building, Maindy Road, Cardiff, United Kingdom, CF24 4HQ.

2: Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada N1G 2W1.

3: Current address: AstraZeneca R&D 1 Francis Crick Avenue Cambridge, UK CB2 0AA.

4: Raymond G. Perelman Center for Cellular and Molecular Therapeutics, Children's Hospital of Philadelphia, Philadelphia, 19104, USA.

5: School of Biosciences, Cardiff University, The Sir Martin Evans Building, Museum Avenue, Cardiff, CF10 3AX.

6: Department of Neurodegenerative Disease, Huntington's Disease Centre, Queen Square Institute of Neurology, University College London, London, WC1N 3BG, UK.

7: UK Dementia Research Institute, University of Cambridge, Clifford Albutt building, Cambridge biomedical campus, Cambridge, CB2 0AH, UK.

8: Centre de Recherche du CHUQ site LOEX and Department of Neurological Sciences, Hôpital Enfant-Jésus.

9: Department of Pathology, University of Pennsylvania, Philadelphia, 19104, USA.

10: Centre de Recherche du CHU de Québec, Axe Neurosciences; Département de Psychiatrie et Neurosciences, Université Laval, Québec, QC, Canada.

11: Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Maindy Road, Cardiff, UK, CF24 4HQ.

*: corresponding author: dionv@cardiff.ac.uk

One sentence summary: The Cas9 nickase contracts CAG/CTG repeats at multiple disease loci in patient-derived cells and improves molecular and behavioral phenotypes in HD.

Keywords: CRISPR; Cas9 nickase, gene editing, off-target mutations, Huntington's disease, myotonic dystrophy type 1, R6/1.

Abstract:

Expanded CAG/CTG repeats cause over 15 different diseases that all remain without a disease-modifying treatment. Because repeat length accounts for most of the variation in disease severity, contracting them presents an attractive therapeutic avenue. Here, we show that the CRISPR-Cas9 nickase targeted to CAG/CTG repeats leads to efficient contractions in Huntington's disease patient-derived neurons and astrocytes, and in myotonic dystrophy type 1 patient-derived neurons. The approach is allele-selective and free of detectable off-target mutations. Striatal injection of the Cas9 nickase in a mouse model for Huntington's disease using adeno-associated viral vectors led to contractions in over half the infected cells. Upon injection, we observed a reduction in the number of inclusion bodies, improved transcriptome, and ameliorated locomotion. The effects were greater than expected from the contractions induced and suggest that non-cell autonomous mechanisms may be involved. Our results provide the proof-of-concept that correction of CAG/CTG repeats can improve Huntington's disease phenotypes *in vivo*.

67 **Introduction**

68 The heterozygous expansion of CAG/CTG repeats at 15 different loci in the genome
69 causes clinically distinct intractable neurodegenerative and neuromuscular disorders[1],
70 including Huntington's disease (HD) and myotonic dystrophy type 1 (DM1)[2,3]. They
71 are all currently without a disease modifying treatment[4,5]. The size of the expanded
72 repeat tract explains up to 60% of the variation in the age at disease onset in HD[6],
73 with longer repeats leading to more severe phenotypes. This is compounded by repeat
74 expansions accumulating over time in somatic tissues, especially in those cell types
75 that are most affected by the disease[7–9]. Moreover, DNA repair factors involved in
76 somatic expansion have been identified as disease modifiers[10–14]. These observations
77 suggest that somatic expansion is a major driver of disease pathogenesis and that
78 preventing expansions in relevant cell types, or, better, inducing contractions, may
79 provide a much-needed therapeutic avenue.

80 Several studies have used gene editing *in vivo* at the *DMPK* and *HTT* loci, where
81 repeat expansions cause DM1 and HD, respectively[15–25]. All these approaches mutate
82 the repeat locus, either by excising and producing uncontrollable insertions, deletions,
83 and rearrangements or by mutating the repeat tract using base editors. A small
84 molecule has also been developed that induces contractions of the repeat[26]. Editing
85 efficiencies ranged from 2% to 54% of the targeted cells. Molecular and behavioural
86 rescues have not been consistently tested, but some studies reported partial
87 rescues[15,22].

88 Importantly, despite patients having a heterozygous expansion, typical gene editing
89 approaches mutate both alleles, which may have undesirable consequences[4,27,28].
90 Three studies have circumvented this problem by designing sgRNAs that target a
91 specific single-nucleotide polymorphism distinguishing the two *HTT* alleles[16,21,29].
92 However, this approach reduces the total number of individuals who stand to benefit
93 since the most common variant is found in only 30% of HD individuals of European
94 ancestry[29].

95 Altogether, these results highlight the need to correct - rather than mutate - the
96 expanded allele. Gene editing approaches need to be efficient, allele-selective, and
97 applicable to as many individuals as possible, regardless of disease. Double-strand
98 breaks, which are highly mutagenic, should be avoided to reduce the risk of off-target

99 mutations. Importantly, behavioural and molecular pathology must be improved by the
100 treatment.

101

102 We showed previously that the Cas9D10A nickase could induce contractions of
103 ectopic CAG repeats when targeted directly to the repeat tract in immortalised cells
104 using one single-guide RNA (sgRNA) in a transcription-dependent and replication-
105 independent manner[30–32]. Here, we harnessed the Cas9D10A nickase to correct
106 expanded CAG/CTG repeats and provide the *in vivo* proof-of-concept for the safe
107 correction of expanded repeats as a therapeutic avenue.

108 **Results**

109 *Nickase-induced contractions in patient-derived cells.*

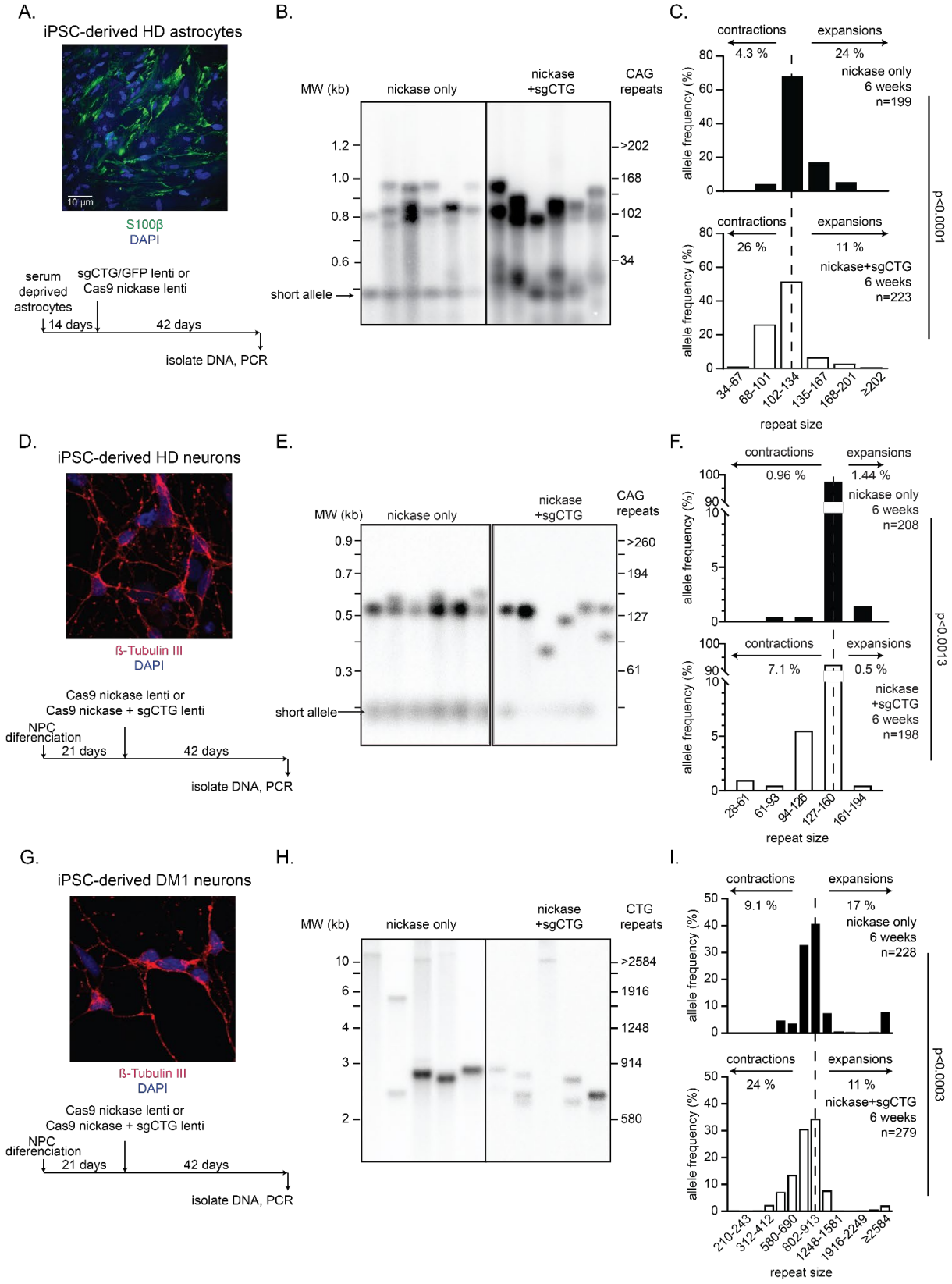
110 We first investigated the possibility of contracting CAG/CTG repeats in disease-
111 relevant cells. Astrocytes contribute considerably to HD pathogenesis[33–37]. We
112 differentiated HD human induced pluripotent stem cells (iPSCs) (CS09iHD109-n1),
113 harbouring 130 to 140 CAG/CTG repeats at the *HTT* locus[38], into astrocyte progenitors
114 and then to mature S100β$^+$ astrocytes (Fig. 1a and Supplementary Fig. 1a). We
115 transduced them with a lentivirus expressing Cas9D10A (Supplementary Table 1). A
116 second lentivirus expressed the sgRNA against the repeat tract (sgCTG) and
117 contained a GFP cassette to monitor silencing of the vector. We found that GFP and
118 Cas9D10A expression was sustained over 42 days (Supplementary Fig. 1bc). We
119 assessed repeat size distribution using small-pool PCR (Fig. 1b)[39] and observed a 6-
120 fold increase in the proportion of contracted alleles in the dual treated cultures. We
121 also saw a >2-fold decrease in the number of large alleles (>134 CAGs) compared to
122 populations transduced with Cas9D10A only (P≤0.0001, Fig. 1c). On average, the
123 cells treated with both Cas9D10A and sgCTG had 20 fewer repeats than the cells
124 treated with Cas9D10A only. These results suggest that HD iPSC S100β$^+$ astrocytes
125 undergo robust contractions when exposed to both the Cas9D10A and sgCTG.

126 In HD, the driving factor of pathogenesis is neuronal death, which occurs first in the
127 striatum with medium spiny neurons, followed by neurons in the cortex[40]. We
128 differentiated the same HD iPSCs into neurons (Fig. 1d)[41]. Our cultures were 99% pure
129 as judged by β-tubulin staining (Supplementary Fig. 1d). We exposed the neurons to the
130 Cas9D10A and sgCTG for 21 or 42 days and performed small-pool PCRs
131 (Supplementary Fig. 1ef, Fig. 1e). We found an increase in the frequency of alleles
132 shorter than 127 CAGs over time, reaching a 7-fold increase after 42 days and a 3-
133 fold decrease in the frequency of alleles with over 160 repeats (Fig. 1e-f, P≤0.0013).
134 We confirmed the results using single-molecule real-time (SMRT) long read
135 sequencing (Supplementary Fig. 2). These results suggest that the Cas9D10A can
136 contract expanded CAG/CTG repeats in HD iPSC-derived neurons.

137 The advantage of our design is that the repeat tract itself is targeted and, therefore, it
138 should induce contractions at multiple disease loci. Although DM1 patients display

139 debilitating myotonia and muscle wasting, the brain is extensively affected, including
140 the cortex[42]. We generated cortical neurons from an iPSC line that harbours an
141 average of 996 CTGs at the *DMPK* locus, with some alleles exceeding 2500 CTGs
142 (Supplementary Fig. 1g, Fig. 1gh). Although it was not sufficient to lead to repeat sizes
143 in the non-pathogenic range, the cells treated with both Cas9D10A and sgCTG had
144 166 fewer repeats on average than the cells expressing only Cas9D10A (Fig. 1i). We
145 conclude that Cas9D10A induces contractions in at least two distinct disease loci
146 without changing the Cas9D10A construct or the sgRNA.

147 One important question is: what is the minimum expansion length that remains a
148 substrate for Cas9D10A? We have shown previously that repeats of at least 101
149 CAG/CTGs contract readily to non-pathogenic sizes (i.e., 35 CAGs or lower for HD
150 and below 50 for DM1) but that 42 repeats remain stable[30]. To determine the threshold
151 with more precision, we used a HD patient-derived lymphoblastoid cell line that
152 harbours 60 CAG repeats at the expanded *HTT* locus and delivered both Cas9D10A
153 and sgCTG via lentiviruses (Supplementary Table 1). Forty-two to 56 days after
154 transduction of the sgCTG virus, we isolated the cells and performed SMRT
155 sequencing. We found that cells transduced with both viruses contracted readily
156 (Supplementary Fig. 3), suggesting that the threshold for contraction with Cas9D10A
157 is between 42 and 60 repeats.



158

159 **Figure 1.** Cas9D10A-induced contractions in HD and DM1 iPSC-derived cells. A) Top: Representative
160 confocal image of HD iPSC-derived astrocytes stained with S100β and DAPI. Bottom: timeline of the
161 experiments. B) Representative small-pool PCR blot showing contractions in HD iPSC-derived
162 astrocytes that were either transduced with the Cas9D10A only or both the Cas9D10A and the sgCTG
163 for 42 days. C) Quantification of the small-pool PCR blots for HD iPSC-derived astrocytes. D) Top:

164 Representative confocal image of HD iPSC-derived cortical neurons stained with β-Tubulin III and
165 DAPI. Bottom: timeline of the experiments. E) Representative small-pool PCR blot showing contractions
166 in HD iPSC-derived cortical neurons that were transduced with Cas9D10A only or both Cas9D10A and
167 the sgCTG for 42 days. F) Quantification of the small-pool PCR blots for HD iPSC-derived cortical
168 neurons. G) Top: Representative confocal image of DM1 iPSC-derived cortical neurons stained with β-
169 Tubulin III and DAPI. Bottom: timeline of the experiments. H) Representative small-pool PCR blot
170 showing contractions in DM1 iPSC-derived cortical neurons that were transduced with Cas9D10A only
171 or both Cas9D10A and the sgCTG for 42 days. I) Quantification of the small-pool PCR blots for DM1
172 iPSC-derived cortical neurons. 'n' is the number of alleles counted in the small-pool PCR experiments.
173 Scale bar = 10 µm. Dashed lines in panels C, F, and I indicate modal repeat size.

174

175 *Off-target mutations remain at background levels in the presence of both Cas9D10A*
176 *and sgCTG*

177 An important safety concern in the context of gene editing is the generation of off-
178 target mutations[43]. Most approaches mitigate this issue by carefully designing the
179 sgRNA and adjusting the target sequence accordingly. Here, this is not possible as
180 we target the repeat tract itself. To address whether our approach led to unwanted off-
181 target mutations, we first used PCR-free whole genome sequencing using Illumina
182 short-read sequencing. We used the same HD iPSC-derived neurons cultured without
183 transduction for 42 days and compared them with cultures that were transduced with
184 either Cas9D10A only or Cas9D10A and the sgCTG. We obtained at least $5.5x10^9$
185 reads per treatment, including $>3.9x10^8$ reads per condition spanning 1830 genes that
186 contain a targetable CAG/CTG repeat (see methods, Supplementary Table 2). The
187 frequency of mutations was not different between the transduced and the non-
188 transduced cells (P>0.31, Table 1, Supplementary Fig. 4). Similar results were found
189 using the HD iPSC-derived astrocytes (P>0.41, Supplementary Figure 4, Table 1). In
190 total, we found 8 mutations overlapping with CAG/CTG repeats, but they were not
191 enriched in the samples treated with Cas9D10A and sgCTG (Fisher's exact test
192 P>0.26, Supplementary Table 2). These results suggest that the expression of
193 Cas9D10A together with sgCTG did not induce mutations above the background seen
194 in untreated cells.

195 Second, we looked for rearrangements at the non-expanded *HTT* allele 42 days after
196 treatment in HD iPSC-derived neuronal cultures using the targeted long-read
197 sequencing data presented above (Supplementary Fig. 2). We used Repeat
198 Detector[44] to determine repeat size and Sniffles2[45] to detect rearrangements in this
199 dataset. These analyses suggested that one sample transduced with Cas9D10A and

200 sgCTG had an insertion downstream of the repeat tract at a frequency of 0.107.
201 However, we could not identify this rearrangement with the Integrative Genomics
202 Viewer[46] or by manually looking for the rearranged alleles in the circular consensus
203 sequences. Moreover, we found no rearrangements in the other 3 samples treated
204 with both Cas9D10A and sgCTG, which came from the same neuronal epithelium, but
205 were differentiated separately from neuronal progenitor cells. Thus, rearranged alleles
206 were not detectable in this sample. We conclude that this was a false negative and
207 that mutations are rare, if present at all, in the non-expanded alleles upon a 42-day
208 exposure to Cas9D10A and sgCTG.

209 **Table 1.** Analysis of mutations in HD iPSC-derived cortical neurons and astrocytes.

		CAG/CTG repeat genes			Whole Genome		
Cell type	Treatment	# of mutations*	# of reads $x10^8$	P-value§	# of mutations*	# of reads $x10^9$	P-value§
Neurons	-	51±15	4.06	-	1558±260	5.6	-
	Cas9D10A	60±18	4.17	0.52	1806±396	5.7	0.23
	Cas9D10A + sgCTG	39±15	3.98	0.31	1512±261	5.5	0.99
Astrocytes	Cas9D10A	26±22	1.53	0.41	1819±793	2.1	0.52
	Cas9D10A + sgCTG	13±4	1.33		1361±308	1.9	

210 *: The number of mutations obtained per sample (8 samples/treatment for the neuronal cultures and 3 samples/treatment for the astrocytic
211 cultures). No difference was seen between 21 and 42 days in astrocytes and therefore the data were combined for this analysis.

212 §: The number of mutations found at the selected 1830 genes and genome-wide were normalised to the total reads per sample and indicated as
213 the number of mutations normalised to the number of reads (Supplementary Figure 4). Differences in normalised mutations between neuronal
214 cultures conditions were calculated by one-way ANOVA with a post-hoc Tukey's multiple comparison test, except for the whole genome data
215 from astrocytes where we used a Student's t-test.

216 *Expression of AAV-delivered GFP and Cas9D10A in vivo*

217 Next, we established whether we could deliver Cas9D10A and sgCTG directly to the
218 striatum of mice using recombinant adeno-associated viral vectors (AAV) serotype 9.
219 The whole CRISPR system is too large to be packaged into a single AAV and therefore
220 we used a two-vector system with one AAV harbouring Cas9D10A and the other
221 expressing both sgCTG and GFP (Fig. 2ab, Supplementary Table 3).

222 When injected in R6/1 mice two days after birth (P2), GFP was detectable throughout
223 the striatum along with some expression in the cortex 6 weeks after injection
224 (Supplementary Fig. 5a). Expression was still present 3 months post-injection without
225 a noticeable drop in expression over time (Supplementary Fig. 5b). AAV9 exhibited
226 strong tropism to $NeuN^+$ neurons and expressed GFP well (Supplementary Fig. 5cd),
227 consistent with previous work[47]. Furthermore, using qPCR, we found that there were
228 26 ± 6 copies of the sgCTG/GFP AAV genome per cell in striata, on average, one
229 month after injection (Supplementary Fig. 5e). These data demonstrate that AAV9 can
230 infect striatal neurons in R6/1 mice and that GFP is robustly expressed over several
231 months.

232 Cas9D10A expression, however, proved more challenging. We tested four different
233 AAV constructs for Cas9 expression (Supplementary Table 3) that were previously
234 shown to induce edits *in vivo*[21,48,49] and introduced the D10A mutation. Of the four
235 constructs injected in wild type and R6/1 mice, one (version 1 - v1) did not express
236 any detectable Cas9D10A, v2 and v3 were expressed 2 weeks post-injection but
237 silenced later (Supplementary Fig. 6a-f). The last construct, v4, containing a miniCMV
238 promoter, was still expressed 2 months after injection in adults, albeit at lower levels
239 (Supplementary Fig. 6g). When we injected P2 mice, we could detect Cas9D10A for
240 up to 5 months, but expression decreased over time (Supplementary Fig. 6gh). These
241 data suggest that Cas9D10A expression wanes over time at a rate heavily dependent
242 on the construct used. We continued with Cas9D10A v4 to test whether the nickase
243 can induce contractions *in vivo*.

244

245 *Large contractions after delivery of Cas9D10A and sgCTG to neonatal mice*

246 Next, we aimed to determine whether Cas9D10A could induce contractions *in vivo*.
247 Contractions in the mouse brain are measured against an ongoing rate of somatic
248 expansion. SMRT sequencing of the expanded repeats showed, as expected, that
249 there was a small but progressive increase in the modal repeat size over time,
250 accompanied by a bimodality beyond 10 weeks of age (Supplementary Fig. 2e). This
251 is due to different cell types having different rates of expansion, with the medium spiny
252 neurons having the largest repeat sizes[50]. The average copy number of Cas9D10A 1
253 month after injection of R6/1 mice at P2 was 0.3 ± 0.15 copies per cell on average
254 (Supplementary Fig. 6i). Therefore, the efficiency of contractions is capped at 26% on
255 average in these experiments, as predicted by a Poisson distribution.

256 Repeat size in P2-injected mice showed fewer expansions in the expanding medium
257 spiny neurons as well as an accumulation of contractions over time (Fig. 2d-h).
258 Specifically, short alleles of 12 to 17 CAGs, which are well within the non-pathogenic
259 range, appeared 5 months post-injection (Supplementary Fig. 7a-f; Fig. 2d-h). These
260 large contractions were seen in all 3 mice sequenced at this late time point. The
261 instability index[51] showed a slower increase over time in the hemisphere injected with
262 both Cas9D10A and sgCTG compared to Cas9D10A alone (Supplementary Fig. 8a,
263 P=0.0004), which was accompanied by a more negative contraction index
264 (Supplementary Fig. 8b, P=0.0002), and a lower expansion index (Supplementary Fig.
265 8c, P=0.005). We further found that the average repeat size was reduced compared
266 to mice injected only with Cas9D10A, especially after 4 and 5 months (Supplementary
267 Fig. 8d, P<0.001). We observed a rise in the percentage of short alleles in the samples
268 treated with both Cas9D10A and sgCTG from 5.1 ± 1.5% 1-month post-injection to
269 15.7 ± 7.6% 5 months after the injections (Supplementary Fig. 8d, Fig. 2i, P=0.035).
270 Given that we have not sorted the infected cells, the expected maximum was 26%.
271 We conclude that most cells that express Cas9D10A together with sgCTG in the
272 striatum of R6/1 mice harbor contractions after 5 months.

273
274

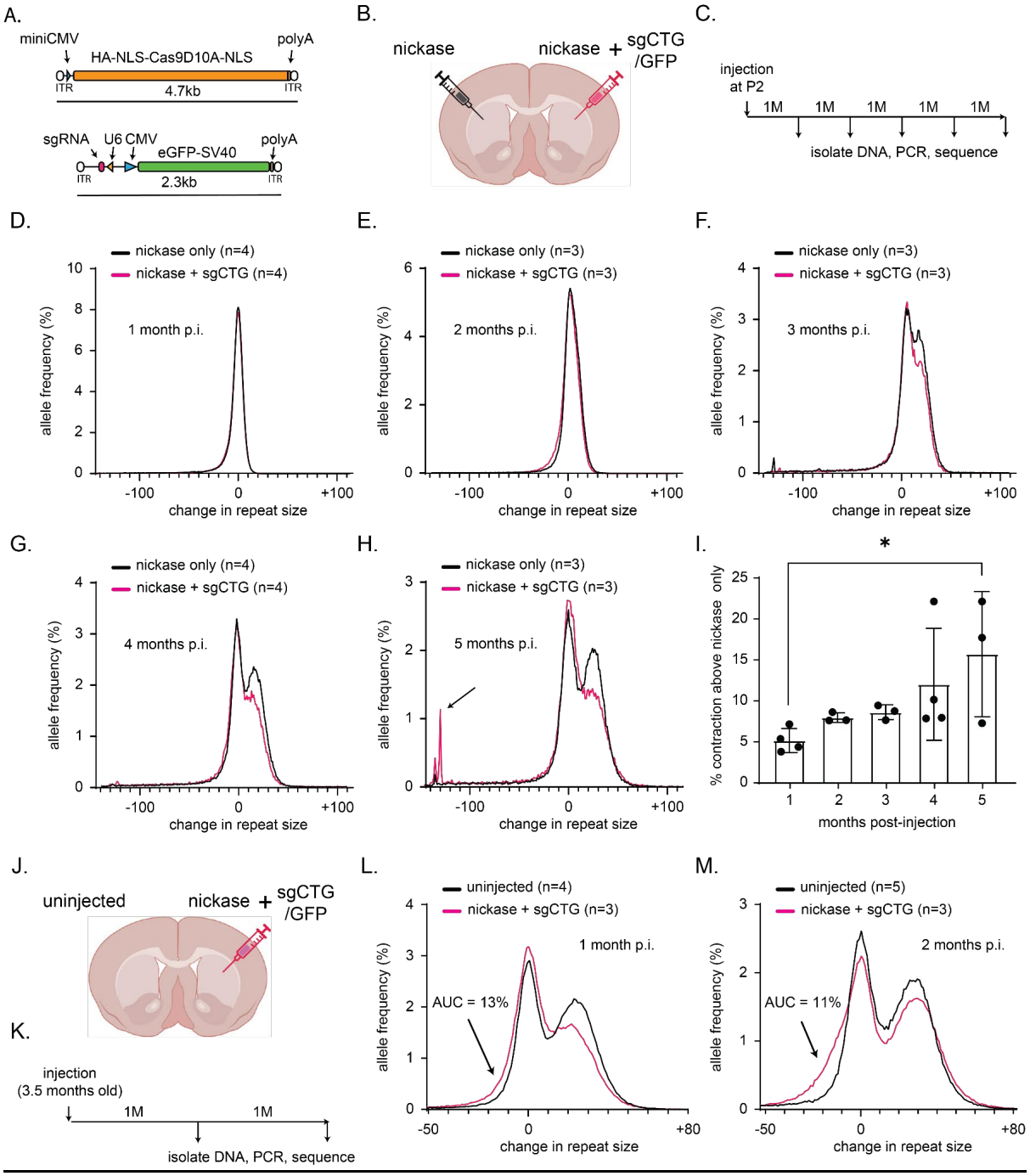


Figure 2. CAG repeat contraction in P2- and adult injected R6/1 animals. A) Schematic of the AAV constructs used. B) Schematic of the experimental approach. C) Timeline of the experiment. P2 = 2 days after birth. D) Aggregated graphs of repeat size distribution in hemispheres of R6/1 mice injected with Cas9D10A only (black) versus Cas9D10A + sgCTG (magenta) 1 month, E) 2 months, F) 3 months, G) 4 months and H) 5 months post-injection. Arrow points to large contractions with 17 remaining CAGs. Note that the graphs are normalised to the modal peak size found in each hemisphere (see methods). The mice used in these experiments had a modal repeat size ranging from 138 to 151 CAGs. A nonparametric Wilcoxon matched-pairs signed rank test showed no significant difference in repeat size in animals sacrificed 1 to 3 months post-injection. After 4 and 5-months, the difference was significant (P value <0.0001 (****)). I) Difference in the area under the curve between the Cas9D10A + sgCTG treatment and Cas9D10A alone (*P<0.05 using a one-way ANOVA). Error bars represent ± standard deviation. Number of animals: Cas9D10A only; 1 month n=4, 2 months n=3; 3 months n=3; 4 months

288 n=4; 5 months n=3; Cas9D10A + sgCTG; 1 month n=4, 2 months n=3; 3 months n=3; 4 months n=4; 5
289 months n=3). J) Schematic of the experimental approach for injections in adult mice. K) Timeline of the
290 experiments. L) Aggregated graphs of repeat size distribution in the striatum of uninjected adult R6/1
291 mice (black) compared to striata injected with Cas9D10A + sgCTG (magenta) 1 month post-injection.
292 Unpaired, nonparametric Kolmogorov-Smirnov test returned a significant difference in cumulative
293 distributions. (P <0.023 (*)). M) Same as L, but 2 months after injection (P<0.0001 (****)). The mice
294 used in these experiments had a modal repeat size ranging from 138 to 145 CAGs. Number of animals:
295 uninjected 1-month post-injection n=4, 2 months post-injection n=5; Cas9D10A + sgCTG; 1 month post-
296 injection n=3, 2 months post-injection n=3.

297

298 *Contractions in the R6/1 adult striatum upon Cas9D10A and sgCTG delivery*

299 Next, we asked whether contractions also occurred in adult mice upon treatment with
300 the Cas9D10A. We repeated the experiments, injecting 3.5 months old adult R6/1
301 mice with both Cas9D10A and sgCTG and comparing them with age-matched un-
302 injected controls. We tracked the instability at 1 and 2 months post-injection (Fig. 2j-
303 m). We found robust infection efficiencies with an average of 464 ± 217 and 663 ± 92
304 AAV DNA copies per cell when injecting AAV-sgCTG-GFP and AAV-Cas9D10A,
305 respectively, 1 month post-injection (Supplementary Fig. 6jk). Consequently, we saw
306 contractions *in vivo* 1 and 2 months post-injection (Supplementary Fig. 7g-i; Fig. 2lm)
307 with the instability index being lower and the contraction index decreasing over time,
308 indicative of an increase in the number and/or a decrease in the size of the short alleles
309 (Supplementary Fig. 8e-g). We conclude that our gene editing regimen induces
310 contractions in adult R6/1 mice after symptom onset.

311 *Optimisation of the AAV constructs for robust expression in vivo*

312 We next optimised the Cas9D10A construct further by using a cell-type specific codon
313 optimization as described (patent WO2023105212A1). We found that the best
314 construct for expression in the striatum amongst the ones tested was a skeletal muscle
315 codon optimised Cas9D10A (v5 - Supplementary Table 3) used together with an
316 optimised sgRNA scaffold containing an extra loop and a point mutation that improves
317 editing efficiency[52] injected at a ratio of 1:2 (Supplementary Fig. 10&11). Under these
318 conditions, injecting 2-month-old animals yielded an average of 330 ± 92 and 1020 ±
319 173 AAV genomes per cell for the Cas9D10A v5 and the sgCTG-Mut+5 AAVs
320 (Supplementary Fig. 12), respectively, 3 months post-injection. However, presumably
321 because of Cas9D10A silencing, immunofluorescence data suggest that 19% of

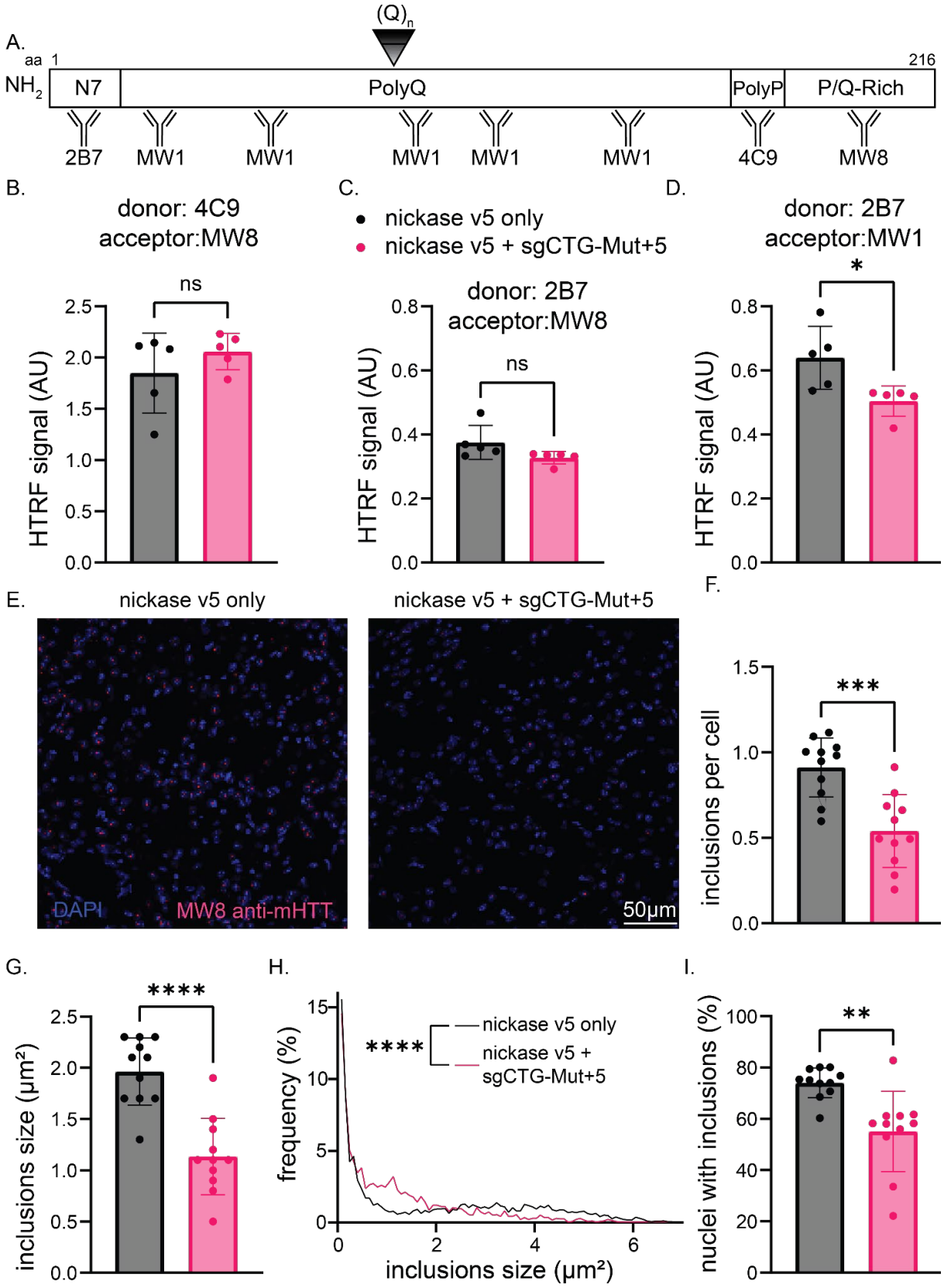
322 neurons expressing GFP also expressed Cas9D10A (Supplementary Fig. 13). We
323 used these conditions to assess the impact of contractions *in vivo*.

324

325 *CAG repeat contractions translate to shorter polyglutamine stretches.*

326 We assessed whether contractions in the CAG/CTG repeat of the m*HTT* gene could
327 be observed in the protein. To this end, we used Homogeneous Time Resolved
328 Fluorescence (HTRF, Figure 3a)[53]. This assay uses striatal extracts together with
329 antibody combinations to assess how close two epitopes are from each other within
330 the mHTT protein. To establish the method in R6/1 mice, we first determined levels of
331 soluble and aggregated mHTT. As determined for other HD mouse models[53], we found
332 that soluble mHTT decreases over the life span of the R6/1 mice and aggregated
333 mHTT increases with age (Supplementary Fig. 14a-c). We then compared the levels
334 of the various forms of mHTT in the mice injected with either Cas9D10A v5 alone or
335 together with the sgCTG-Mut+5. We found no change in the amounts of aggregating
336 (4C9 and MW8 antibodies) and soluble (2B7 and MW8 antibodies) mHTT (Fig 3bc).
337 However, when we used the 2B7 antibody together with MW1, which targets the
338 polyglutamine tract, we observed a reduction in signal (Fig. 3d, Supplementary Table
339 4), suggesting that fewer polyglutamines were found after the treatment with
340 Cas9D10A *in vivo*.

341



342

343 **Figure 3.** CRISPR-Cas9D10A injection reduces polyQ length and aggregation of mHTT *in vivo*. A)
344 Schematic scaled diagram of a mHTT with 150 glutamines showing the relative location of HTT
345 antibodies. Four antibodies detect mHTT exon 1: 2B7 maps to within the first 17 amino acids, MW1
346 detects expanded polyQ tracts, 4C9 detects the proline-rich region that lies between the two proline

347 repeats in human HTT and MW8 acts as a neo-epitope antibody to the C-terminus of HTT. B) 4C9-Tb
348 with MW8-d2 were used to track changes in aggregated mHTT, C) 2B7-Tb with MW8-d2 were used to
349 track soluble mHTT and D) 2B7-Tb with MW1d2 pair for polyglutamine size assay by HTRF in striatal
350 lysates from R6/1 mice at 5 months of age (3 months post-injection). (*P<0.05 using a t-test). Error bars
351 represent ± standard deviation. Number of animals: 5 per group). E) representative z-projection images
352 of striatal areas from 5-months-old R6/1 animal brains sacrificed 3 months post-injection of Cas9D10A
353 only or Cas9D10A and sgCTG-Mut+5 and stained with DAPI (blue) and mutant polyglutamine inclusions
354 (red). F) Analysis of total number of inclusions normalised to the total number of nuclei per image in
355 striatal regions injected with Cas9D10A and the sgCTG-Mut+5 in comparison with their littermates that
356 received Cas9D10A only. G) Two-dimensional analysis of the average aggregate size in treated
357 animals. H) Distribution frequency of all analysed inclusions. Kolmogorov-Smirnov test to compare
358 inclusion size distributions was used. P value <0.0001 (****)(n = 4009 and 2303 inclusions in Cas9D10A
359 v5-injected only and Cas9D10A v5 + sgCTG-Mut+5 injected animals, respectively). I) Quantification of
360 the number of mHTT positive nuclei in mice injected with Cas9D10A only or together with the sgCTG-
361 Mut+5. Data are mean ± SD. fgi Student t-test were used (**P<0.01; ***P<0.001; ****P<0.0001). (n= 3-
362 4 images per mouse, between 1000 and 1900 cells were assessed per mouse, with 3 mice per
363 condition).

364 *Cas9 nickase and sgRNA delivery in vivo reduces the levels of mHTT inclusions*

365 A molecular hallmark of HD in R6/1 mice is the presence of mHTT aggregates in the
366 striatum[54]. To determine whether the injection of both the Cas9D10A and sgCTG-
367 Mut+5 AAVs could improve this phenotype, we performed immunofluorescence using
368 the specific MW8 antibody[55–57]. We found that the number of inclusions per cell as well
369 as the size of the remaining inclusions decreased by over 40% in the striatum of mice
370 injected with both AAVs compared to the ones injected with only Cas9D10A (Fig. 3e-
371 h). Importantly, fewer cells in the striatum had inclusions (Fig. 3i). These results were
372 specific for the injected region since a similar analysis in cortical areas of the same
373 mice did not show any differences between groups (Supplementary Fig. 15d-f). We
374 conclude that the injection of both the Cas9D10A and the sgCTG-Mut+5 AAVs
375 decreases the burden of mHTT inclusions in the striatum. This is consistent with
376 Cas9D10A slowing the transition from diffuse aggregates to inclusions suggested by
377 the HTRF analysis.

378 *Nickase treatment partially mitigated gene expression alteration in the striatal medium*
379 *spiny neurons*

380 Dysregulation of transcription is a hallmark of HD pathology[58,59]. Thus, we aimed to
381 determine whether the nickase could improve transcriptome dysregulation. To this
382 end, we performed single-nuclei RNA sequencing of striatal samples from both wild
383 type and R6/1 littermates 3 months after injecting them with either the Cas9D10A v5
384 alone or together with sgCTG-Mut+5 AAV. We first identified the medium spiny neuron

385 population (Supplementary Fig. 16a-c, Fig. 4a) and determined the effect of the
386 treatment by comparing wild type mice injected with only the Cas9D10A v5 AAV to
387 those injected with both AAVs. We found that the treatment dysregulated only 100
388 genes in wild type mice (adjusted p-value = 0.01, fold change = 1.5; Supplementary
389 Table 5, Supplementary Fig. 16d; Fig. 4b). Similar results were found using a
390 pseudobulk approach (Supplementary Table 6, Supplementary Fig. 16e).

391 The effect of expressing mHTT led to more pronounced changes, R6/1 injected with
392 Cas9D10A alone showed 651 dysregulated genes in medium spiny neurons
393 compared to wild type littermates with the same treatment (Fig 4c, Supplementary
394 Table 5). The pathways affected included dopaminergic synapse, calcium signaling,
395 ribosomes, and multiple pathways of neurodegeneration, as expected from other HD
396 models[58,59]. Similar results were obtained when we used a pseudo-bulk approach in
397 medium spiny neurons with ion channels and neurotransmitter receptors being most
398 affected (Supplementary Fig. 16f). These changes were comparable to those seen in
399 two other HD mouse models[58,60] (Supplementary Fig. 17).

400 Next, we compared transcriptome alterations caused by the expression of both the
401 Cas9D10A and sgRNA-Mut+5 in medium spiny neurons in R6/1 mice (Fig. 4d). We
402 found that the treatment led to 198 dysregulated genes, with 94% (186) of these being
403 down-regulated. We assessed whether these changes rescued dysregulated genes
404 caused by the presence of the HD transgene. To do so, we looked at the subset of
405 genes that were significantly upregulated due to the R6/1 genotype (401 genes) and
406 that were also found expressed in R6/1 medium spiny neurons from mice injected with
407 both Cas9D10A and sgRNA-Mut+5 AAVs (362 genes). We found that in this subset,
408 15% (53 genes) showed a significantly increased expression ($P<0.01$,
409 $|\log_2 FC<0.585|$), in the animals treated with both AAVs. (Fig. 4e). GO cellular
410 component analysis showed that neuron to neuron synapse, dense core granule and
411 synaptic vesicles genes were specifically rescued (Fig. 4f). These changes in gene
412 expression were not due to a change in the number of medium spiny neurons or glia
413 present in the dual injected R6/1 animals (Supplementary Fig. 16b) or to off-target
414 binding of the Cas9 nickase as the misregulated genes were not enriched for genes
415 containing potential off-target sites (Supplementary Table 5). We conclude that our
416 gene editing treatment can mitigate some of the transcriptome alterations in the
417 striatum of R6/1 mice.

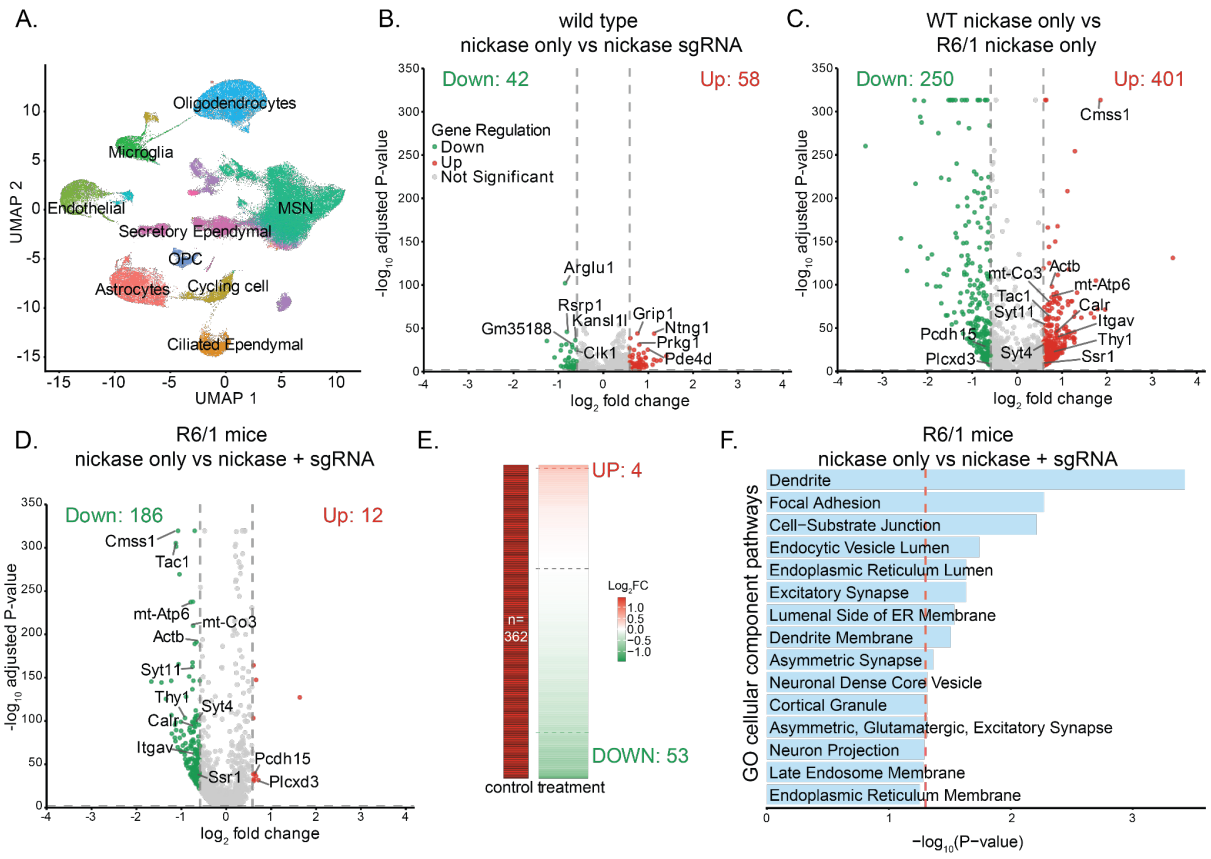


Figure 4. Cas9D10A co-injection with a sgCTG-Mut+5 against the repeat tract mitigates genes upregulated in HD medium spiny neurons 3-months post-injection. A) U-MAP plot of all nuclei across striatal regions from disease and control mice showing all major cell types in the brain (MSNs: medium spiny neurons). B) Volcano plot of differentially expressed genes (DEGs) in medium spiny neurons between wild type (WT) animals injected with Cas9D10A v5 only versus Cas9D10A v5 + sgCTG-Mut+5. C) Volcano plot for DEGs in medium spiny neurons between WT versus R6/1 animals injected with Cas9D10A v5 only. D) Volcano plot comparing DEGs in CRISPR-treated R6/1 medium spiny neurons relative to Cas9D10A v5-injected R6/1 medium spiny neurons. Number of animals: WT Cas9D10A v5 only n=3, WT Cas9D10A v5 + sgCTG-Mut+5 n=3, R6/1 Cas9D10A v5 only n=3, R6/1 Cas9D10A v5 + sgCTG-Mut+5 n=3. E) Heatmap of the changes in expression due to the co-injection of the Cas9 Cas9D10A v5 and sgCTG-Mut+5 specifically in the genes that are upregulated in the R6/1 medium spiny neurons compared to WT medium spiny neurons shown in C. The threshold for significance for differentially expressed genes was an adjusted $P < 0.01$ and an absolute log2 fold change > 0.585. F) Gene ontology (GO) cellular component pathways terms for DEGs which are altered by CRISPR injection in R6/1 medium spiny neurons comparing the treated mice with and without both Cas9D10A and sgCTG-Mut+5.

Intrastriatal Cas9 nickase and sgRNA injections improve motor performance.

We investigated whether the Cas9D10A treatment could improve motor performance. Thus, we injected the striata bilaterally of 2-month-old wild type and R6/1 animals with the Cas9D10A v5 alone or together with sgCTG-Mut+5 (Fig. 5a) and tested them one

441 month after injection. We found significant behavioural impairments in R6/1 mice
442 expressing Cas9D10A only, versus wild type mice with the same treatment, including
443 the balance beam and accelerating as well as fixed rotarod (Supplementary Fig. 18a-
444 d, Supplementary Tables 7 to 10), as expected from previous work[61,62]. The wild type
445 mice injected with both AAVs were not significantly different from those injected with
446 only one AAV in these tests. When we performed an open field test (Fig. 5b), we found
447 that wild type mice injected with both AAVs displayed less rearing than those injected
448 with only the Cas9D10A AAV, but they remained unaffected in all other measures
449 (Supplementary Fig. 19a-g). These results suggest that the treatment with both AAVs
450 has little effect in wild type mice.

451 When comparing the R6/1 mice expressing Cas9D10A alone in an open field test, we
452 found significant behavioural impairments compared to wild type littermates injected
453 with the same AAV. The R6/1 mice showed a 46% reduction in the distance travelled
454 and velocity as well as a 31% reduction in the total time in movement ($P \leq 0.004$, Fig.
455 5c-e). When comparing R6/1 mice expressing both Cas9D10A and sgRNA to those
456 that received only the Cas9D10A AAV, we found an improvement of 61% in distance
457 travelled and velocity as well as a 35% increase in the total time in movement in R6/1
458 mice injected with both AAVs compared with those injected with only one AAV
459 ($P \leq 0.004$, Fig. 5c-e). The R6/1 mice injected with both AAVs were not significantly
460 different from the WT animals. Amelioration of the locomotor phenotype in dual-
461 injected R6/1 mice was not due to changes in motor coordination or strength because
462 parameters tested on the rotarod and balance beam were not affected by the injection
463 of the CRISPR system in R6/1 mice (Supplementary Fig. 18; Supplementary Fig.
464 19fg). We conclude that the striatal injection of the Cas9D10A v5 together with sgCTG-
465 Mut+5 improved gross locomotor phenotypes as measured by an open field test in
466 R6/1 mice.

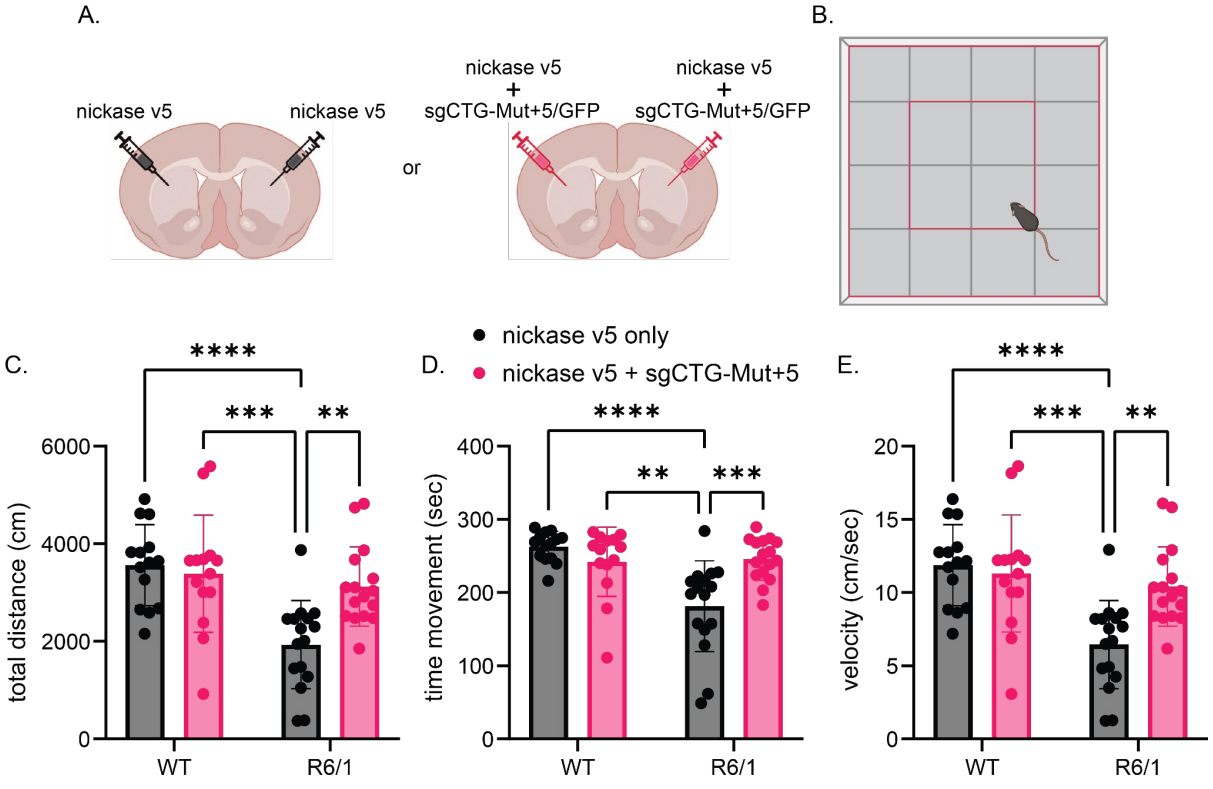


Figure 5. Injection of the Cas9D10A targeting the CAG repeats improves motor phenotypes in R6/1 HD mice in open field test. A) Schematic of the experimental approach with wild type (WT) and R6/1 animal bilaterally injected with the Cas9D10A v5 only or the Cas9D10A v5 and the sgCTG-Mut+5 AAVs. B) Open field schematic showing criteria adopted for defining arena areas; total area 80x80; external area of the arena 40x40 total; internal arena of the arena 40x40 in the centre of the arena (red square). C) Total distance travelled (WT Cas9D10A v5 only vs R6/1 Cas9D10A v5 only: P≤0.0001; R6/1 Cas9D10A v5 only vs R6/1 Cas9D10A v5 together with sgCTG-Mut+5: P≤0.004). D) Total time in movement (WT Cas9D10A v5 only vs R6/1 Cas9D10A v5 only: P≤0.0001; R6/1 Cas9D10A v5 only vs R6/1 Cas9D10A v5 together with sgCTG-Mut+5: P≤0.0005), and E) velocity (WT Cas9D10A v5 only vs R6/1 Cas9D10A v5 only: P≤0.0001; R6/1 Cas9D10A v5 only vs R6/1 Cas9D10A v5 together with sgCTG-Mut+5: P≤0.004) in WT and R6/1 mice injected with the Cas9D10A only or the Cas9D10A with sgCTG-Mut+5. Data presented is the average of two trials per mouse with the standard deviation between groups (for individual trials, see Supplementary Fig. 20). N of animals =14-16 per group. (**P<0.01; ***P<0.001; ****P<0.0001). Statistical analyses are in Supplementary Table 7.

483 **Discussion**

484 Here we provide evidence that Cas9D10A can induce contractions when targeted to
485 expanded CAG/CTG repeats at different disease loci, cell types, and *in vivo*.
486 Cas9D10A did not increase off-target mutation frequencies above background and
487 contractions only occurred on the expanded allele. To our knowledge, the data
488 presented here is the first example of a gene editing approach using a single Cas9
489 and sgRNA pair capable of correcting the mutation that causes clinically distinct
490 disorders in a safe and efficient manner. *In vivo*, nickase-mediated contractions
491 translated to fewer glutamines in the human mHTT fragment and were accompanied
492 by the improvement of molecular and locomotor phenotypes of HD mice.
493
494 Importantly, contraction efficiency improved over time *in vivo*, suggesting that
495 contracted alleles will accumulate for as long as expression of both Cas9D10A and
496 sgCTG is maintained. Thus, the duration of the expression is expected to be a rate
497 limiting factor for the editing efficiency and it is unknown how long expression would
498 be needed for clinically relevant improvements in patients. On the other hand, the
499 propensity of Cas9D10A to be silenced *in vivo* means that self-inactivating
500 approaches, which involve the generation of double-stranded breaks[16], may not be
501 needed, potentially improving safety.
502
503 Although we speculate that contractions drive much of the improvements that we
504 observed, other mechanisms are likely to be also at play. For instance, it may be that
505 some of the changes in gene expression induced by the injection of the AAVs may be
506 beneficial to HD pathology. Another possibility is that the reduction in somatic
507 expansion that accompanies contractions plays a role. This is attractive because
508 several drivers of somatic expansion impact the age at disease onset in HD[11,12].
509 However, reducing expansions when the repeat tract is already large, as in this mouse
510 model, may have only limited effect on disease onset and progression[63,64]. In these
511 cell types, improvements may require contractions beyond simply prevention of
512 somatic expansion. We also considered that the presence of Cas9D10A bound to DNA
513 would change expression of CAG/CTG-containing genes, but found no evidence to
514 support this hypothesis from the snRNA-seq data. Intriguingly, the size of the mHTT
515 inclusions were reduced uniformly across striatal sections and not only in cells that
516 were infected by both AAVs, suggesting a potential non-cell autonomous effect of

517 contractions on the surrounding cells. Such an effect may potentiate the improvements
518 induced by the treatment. This is in line with previous observations that *HTT* silencing
519 in neurons improved astrocyte phenotypes[65]. Moreover, Oura *et al*[66] used a chimeric
520 mouse approach to show that as few as 20% of edited cells, all of which had the same
521 repeat excision, led to an improvement in HTT inclusion pathology, weight, and
522 clasping phenotypes compared to R6/2 HD mice without any edited cells. Regardless
523 of the exact mechanism, our data suggest that expressing Cas9D10A and a sgRNA
524 against the repeat tract itself leads to significant improvements in the molecular and
525 the locomotor phenotypes.
526
527 One limitation of our work is the use of two different AAVs to deliver Cas9D10A and
528 sgCTG. This is common to other gene editing approaches as well[23,24]. One way
529 around this is to use smaller Cas9 orthologs, for example the Cas9 from *S. aureus*[67],
530 as has been done *in vivo* for the *DMPK* locus[19,20]. Unexpectedly, this did not lead to
531 higher levels of editing compared to a two-vector system[19,20]. Therefore, the size of
532 the Cas9 enzyme and its delivery in the brain remain important challenges.
533
534 It is noteworthy that upon expression of the Cas9D10A we induced exclusively repeat
535 contractions, rather than a collection of uncontrolled insertions and deletions[15–25,66,68–
536 70]. Our approach is as efficient as excising the repeat tract, while remaining more
537 precise and avoiding unwanted on-target and off-target mutations. This is of note
538 because other approaches that involve base editing of the repeat tract targeted all
539 CAG/CTG repeats, including non-expanded repeats[23,24]. The Cas9 nickase approach
540 is therefore more precise and produces fewer unnecessary mutations.
541
542 A gene editing approach would have several advantages over more traditional
543 approaches, including antisense oligonucleotides and microRNAs, which reduce the
544 amount of *HTT* or *DMPK* mRNA[5,71,72]. Specifically, our gene editing approach would
545 permanently improve the most proximal cause of the disease by reducing the size of
546 the expanded CAG/CTG repeat. Recently, a small molecule that targets DNA
547 secondary structures formed by expanded CAG/CTG repeats was shown to induce
548 contractions *in vivo*[26,73]. However multiple stereotactic injections or continuous supply
549 of the drug over 16 weeks were required for the effect to be significant. A gene editing
550 approach would potentially require only a single dose. Crucially, from a patient benefit

551 standpoint, it is imperative that we develop multiple approaches to maximise the
552 chances of delivering successful drugs into the clinic.
553

554 **Methods**

555 *iPSC cultures and differentiation conditions*

556 For the HD iPSC-derived astrocytes, we obtained CS09iHD109-n1[38] iPSCs from the
557 Cedars-Sinai Medical Center's David and Janet Polak Foundation Stem Cell Core
558 Laboratory. We found the expanded *HTT* allele to be around 135 CAG repeats and
559 the short one having 19 units. Cultures of iPSCs were tested regularly and found to be
560 negative for the presence of mycoplasma using a service provided by Eurofins.
561 IPSCs were grown in plates coated with Matrigel in E8 Flex Medium at 37°C, 5% CO_2.
562 For astrocyte differentiation iPSCs grown in E8 medium were supplemented with Rho-
563 associated protein kinase inhibitor (RI) at approximately 60% confluency for 24 h.
564 Neural progenitors (NPCs) were first derived by culture in SLI medium (advanced
565 DMEM F-12 supplemented with 1x PenStrep, 1x Glutamax, 1% Neurobrew (w/o)
566 vitamin A, 10 µM SB-431542, 200 nM LDN-193189, 1.5 µM IWR-1) for 8 days. On day
567 8, cells were treated with RI for 1 h and split 1:4 into NB medium (advanced DMEM F-
568 12 1x PenStrep, 1x Glutamax, 2% Neurobrew (w/o) vitamin A) supplemented with RI
569 for 24 h. From day 14 to 16, 20 µM PluriSIn was added to the medium to eliminate any
570 undifferentiated iPSCs. On day 16, differentiated NPCs were frozen. Astrocyte
571 progenitor cells (APCs) were differentiated from these NPCs as previously
572 described[74]. Six million NPCs were thawed into NB medium containing 20 µM PluriSIn.
573 After 2 days, the medium was changed to NF (advanced DMEM F-12 containing 1x
574 PenStrep, 1x Glutamax, 2% Neurobrew with vitamin A, 20 ng ml^{-1} fibroblast growth
575 factor 2 "Improved Sequence" (FGF-2 IS) and cells were allowed to reach confluency.
576 They were then passaged three times in NEL medium (advanced DMEM F-12 with 1x
577 PenStrep, 1x Glutamax, 2% Neurobrew with vitamin A, 20 ng ml^{-1} epidermal growth
578 factor 20 ng ml^{-1} leukaemia inhibitory factor). CD44 positive cells were collected using
579 the DB FACSAria Fusion and maintained into NEF medium (advanced DMEM F-12
580 supplemented with 1x PenStrep, 1x Glutamax, 2% Neurobrew with vitamin A, 20 ng
581 ml^{-1} epidermal growth factor, 20 ng ml^{-1} FGF-2 IS). To mature astrocytes, APCs were
582 seeded onto Matrigel coated 6 well plates at 200,000 cells per well with STEMdiff™
583 Astrocyte Maturation medium (STEMCELL Technologies). After 2 weeks, the media
584 was changed to advanced DMEM F-12 with 2% FBS for another 2 weeks. Once
585 mature, astrocytes were plated in advanced DMEM F-12 medium without FBS for
586 treatment with Cas9D10A.

587 CS09iHD109-n1 and DM1 iPSCs (GM24559, obtained from the Coriell Institute) were
588 differentiated to cortical neurons[74] by first bringing the cells to confluency and
589 maintaining them in neuronal induction medium (0.5x advanced DMEM F-12 and 0.5x
590 Neurobasal medium supplemented with 25 U ml^{-1} PenStrep, 2.5 µg ml^{-1} insulin
591 solution human, 50 µM 2-Mercaptoethanol, 1x MEM non-essential amino acids
592 solution, 500 µM sodium pyruvate solution, 1x N2 supplement, 1x B27 supplement,
593 1mM L-Glutamine, 10 µM SB-431542, 1 µM dorsomorphin dihydrochloride) for 12
594 days. The resulting neuroepithelium was disaggregated and plated in a new precoated
595 2x Matrigel well with neuronal maintenance media (NMM; 50x advanced DMEM F-12
596 and 50x Neurobasal medium supplemented with 25 U ml^{-1} PenStrep, 2.5µg ml^{-1}
597 insulin, 50 µM 2-Mercaptoethanol, 1x MEM non-essential amino acids, 500 µM sodium
598 pyruvate solution, 1x N2 supplement, 1x B27 supplement, and 1 mM L-Glutamine)
599 also containing 20 ng ml^{-1} human FGF-2 IS. Once neuronal rosettes were visible,
600 Neural Rosette Selection Reagent (Stem Cell Tech) was added to the culture for 1 h
601 at 37°C for NPC selection. NPCs in suspension were collected and plated. NPC were
602 maintained in NMM+FGF-2 IS. 50,000 live cells cm^{-2} and 30,000 live cells cm^{-2} of HD
603 and DM1 NPC, respectively, were plated for neuronal differentiation in wells pre-
604 coated with Laminin from Engelbreth-Holm-Swarm murine sarcoma basement
605 membrane (Merck). Cells were plated in NMM and half the media was replaced twice
606 a week. Neuronal cultures were kept for 3 weeks for maturation before being used for
607 lentiviral transduction.

608

609 Mature neuronal and astrocytic cultures were transduced with lentiviral vectors at a
610 multiplicity of infection of 5. The lentiviruses drive the expression of the Cas9D10A
611 together with a blasticidin resistance gene, BSD, alone or together with the sgCTG
612 cassette (Supplementary Table 1). Transduced cells were selected using blasticidin
613 (5 µg ml^{-1}) for 7 days. Cells were collected with scrapers 3 and 6 weeks after the
614 infection to isolate DNA and measure repeat size.

615

616 *Lymphoblastoid cultures*

617 Lymphoblastoid cell line from a Huntington's disease patient (GM03620, obtained from
618 the Coriell Institute) was cultured in RPMI with 2mM L-Glutamine + 15%FBS media.

619 Cell cultures were transduced with lentiviral vectors expressing Cas9D10A (pLenti-
620 EF1alpha-Cas9D10A nickase-Blast; Supplementary Table 1) at a multiplicity of
621 infection of 10 along with 5 µg ml^{-1} polybrene. Transduced cells were selected using
622 blasticidin (10 µg ml^{-1}) for a minimum of 14 days. Then, cells were verified to be
623 expressing Cas9D10A at multiple timepoints throughout to measure nickase protein
624 expression via western blot (Supplementary Fig. 3a) and further transduced with a
625 lentiviral vector expressing the sgCTG (pLV(gRNA)-CMV-eGFP:T2A:Hygro-
626 U6(sgCTG); Supplementary Table 1) at a multiplicity of infection 20 + 5 µg ml^{-1}
627 polybrene on day 0 of the experiment. Transduced cells were selected using
628 hygromycin (50 µg ml^{-1}) for 28 days. Cells were collected 42 and 56 days after the
629 infection with the sgCTG to isolate DNA and measure repeat size using long-read
630 sequencing.

631

632 *Immunostaining*

633 Mature astrocytes were plated onto CELLSTAR 96-well tissue culture plates (Greiner
634 Bio-One Ltd.), fixed with 4% paraformaldehyde (PFA), washed with PBS,
635 permeabilized with 0.4% Triton-X in PBS, washed with PBS, and then blocked for 1 h
636 at RT with 0.2% Triton-X in PBS + 1% normal goat serum (NGS, Merck). Antibodies
637 used and their concentration is found in Supplementary Table 11. Cells were
638 counterstained with DAPI (1 µg ml^{-1}) before imaging with the Opera Phenix High
639 Content Screening System (Perkin Elmer).

640 NPCs were plated on 13 mm coverslips for neuronal differentiation. Neurons were
641 fixed in 4% PFA. Cells were washed in PBS, blocked, and permeabilized for 1 h in 1%
642 BSA, 0.1% Triton X-100, and 4% NGS in PBS, and incubated with primary antibody in
643 1% BSA, 0.1% Triton X-100, and 1% NGS in PBS overnight at 4°C. After washing with
644 PBS, cells were incubated with appropriate Alexa-conjugated secondary antibody in
645 1% BSA, 0.1% Triton X-100 in PBS for 1 h at room temperature. Cells were
646 counterstained with DAPI Staining Reagent (1 µg ml^{-1}; Chemometec) and after several
647 washes in PBS, coverslips were mounted onto SuperFrost Plus slides (Thermo Fisher
648 Scientific) with Fluoromount aqueous mounting medium (Merck) and imaged with the
649 Leica SP8 confocal microscope. Images were captured using LAS X software and
650 were analysed with Fiji software.

651

652 *Flow cytometry*

653 Mature cultures were resuspended by Accutase (Fisher Scientific) and fixed with PFA
654 2%, Saponin 0,1%, in PBS. Fixation step was inactivated using 1.25 M glycine. Cells
655 were washed/blocked with 500 µl of blocking solution (0.5% Saponin; 0.5% BSA; in
656 PBS) and incubated with primary antibody in 300 µl 0.5% Saponin; 0.5% BSA; 1%
657 NGS in PBS for 30 min at 4°C. After washing with a cleaning buffer (0.1% Saponin;
658 0.2% BSA; in PBS), cells were incubated with appropriate Alexa-conjugated
659 secondary antibody in the cleaning buffer for 30 min at 4°C. Cells were counterstained
660 with DAPI Staining Reagent (1 ng ml^{-1}; Fisher Scientific). Finally, cells were
661 resuspended in 0.5% BSA and EDTA 5 mM in PBS. Samples were run on the BD
662 FACSAria Fusion and analysed using FlowJo v10 software.

663

664 *Small-pool PCR*

665 Small-pool PCR was done as previously described[30,31], but using primers listed in
666 Supplementary Table 12. Genomic DNA was extracted with the NucleoSpin Tissue
667 DNA extraction kit (Macherey-Nagel). PCRs were run using between 10-500 pg of
668 DNA measured by Quant-iT Qubit dsDNA HS Assay Kit (ThermoFisher) to find the
669 amount of amplifiable alleles for each sample. The products were run on 1.5% agarose
670 gels in 1X TAE buffer and transferred onto a GeneScreen Hybridization Transfer
671 Membrane. The hybridization was done at 48°C (for *HTT* blots) or 51°C (for *DMPK*
672 blots) in ULTRAhyb Ultrasensitive Hybridization buffer (Fisher Scientific) with a ^{32}P-
673 labelled oligo containing 10 CAG repeats (oVIN-100, Supplementary Table 12). The
674 membranes were exposed to a phosphor screen and scanned using Bio-Rad PMI
675 Personal Molecular Imager. Images were captured using Quantity One 1-D Analysis
676 Software (Basic V4.6.6) and analysed with Fiji version 1.53f51. Blinded allele counts
677 were completed for each membrane based on allele frequency calculated by Poisson
678 distribution as described[31]. Each band was sized using the ladder run on either side
679 of the gel. For the *HTT* locus in astrocytes, we excluded repeats below 34 repeats so
680 as to not conflate contractions with the non-expanded allele. For the *DMPK* locus,
681 small-pool PCRs, sizing was further divided between fragment sizes of 2 kb and 3 kb
682 by fitting an exponential decay of the distance travelled by the ladder bands. This

683 allowed us to subdivide this region of over 580 repeats into three bins, where most of
684 the contractions were observed. Differences between conditions were determined
685 using a Mann-Whitney U Test calculated by GraphPad Prism software (v.10.0.0).

686 *Potential off-target identification*

687 To assess potential off-target effects of our CRISPR-Cas9 system targeting CAG/CTG
688 repeats in the human genome (GRCh38/hg38), we used the CasOFFinder web tool[75].
689 We used two different queries CTGCTGCTGCTGCTGCTG and
690 CTGCTGCTGCTGCTGCTGC to capture potential binding. All analyses were
691 conducted using the 5'-NRG-3' PAM sequence, allowing up to 2 mismatches. This
692 resulted in 1846 potential off-target genes genome-wide. Coverage analysis of those
693 genes revealed 16 genes with very low coverage, which were filtered out, leading to a
694 total of 1830 potential off-target loci (Supplementary Table 2). The same queries were
695 used on the mouse genome (mm10). We found 3132 potential off-targets
696 (Supplementary Table 5).

697

698 *Whole genome sequencing*

699 Genomic DNA was extracted from cells using the Macherey-Nagel NucleoSpin Tissue
700 DNA mini kit and made into dual-indexed single stranded DNA libraries using 'Illumina
701 DNA PCR-free Prep kit, Tagmentation' (Illumina). Quantification of the libraries was
702 done using the KAPA Illumina Universal Library Quant kit (Roche). The sequencing
703 was performed on an Illumina NovaSeq 6000 using the v1.5 S4 reagent kit, 300 cycles
704 using XP 4-lane splitter kit. Together the 33 samples generated 1600Gb of raw data.

705 Demultiplexing was done using bcl2fastq conversion software (Illumina) and the
706 paired-end sequences were aligned to GRCh38/hg38 with the Burrows-Wheeler
707 Aligner (BWA)[76] version bwa/0.7.17 and the mem algorithm. Subsequently, the
708 Sequence Alignment/Map (SAM)[77] files produced were sorted via the Picard
709 (picard/2.27.5 - https://broadinstitute.github.io/picard/). These files were then
710 converted to binary alignment/map (BAM) format and indexed using SAMtools
711 (samtools/1.21)[77].

712 To assess sequence coverage, genome-wide and potential off-target candidate
713 regions, Mosdepth[78] (0.3.11) was applied to all BAM files. The resulting coverage bed

714 files were combined to create an average coverage using an in-house python script
715 (available on https://github.com/DionLab/Murillo-et-al). Read groups were attributed to
716 the BAM files with the Picard tool, which also facilitated the subsequent analysis for
717 single nucleotide polymorphisms using bcftools[79] version 1.21 with htslib version 1.22.

718 To capture rare mosaic variants, we used an approach similar to one used by Dong et
719 al[80]. Variant calling was performed on the whole genome as well as candidate off-
720 target regions specified in the BED file using Bcftools mpileup[77]. Variants with an allele
721 frequency of at least 0.05 and DP2+DP3 ≥ 2. The output was compressed to BCF
722 format following the standard bcftools workflow for each sample. Variants were further
723 filtered for quality using the bcftools filter, excluding those with the following criteria:
724 "QUAL < 30 or DP < 10 or MQ<40 or VAF ≥ 0.2, calculated as (DP4[2] +DP4[3]) /
725 (DP4[0]+DP4[1]+DP4[2]+DP4[3])". Filtered results from all VCF files for 33 samples
726 with 4 lanes were subsequently compiled into a single CSV (Supplementary Table 2)
727 format using a custom Python script for further analysis. Only variants that were found
728 in more than one lane were considered and we removed the ones found at day 0 as
729 they were already in the population before the treatments. We ran this analysis on
730 both the 1830 potential off-targets and the whole genome separately.

731 *Long read sequencing of repeats*

732 Expanded repeats were sequenced as before using SMRT HiFi amplicon
733 sequencing[44] (see Supplementary Table 12 for primers). Mouse samples were
734 mechanically disaggregated in Nucleospin Tissue DNA extraction T1 buffer using a
735 mini hand-held homogenizer (ThermoFisher). Genomic DNA was extracted with the
736 NucleoSpin Tissue DNA extraction kit (Macherey-Nagel). Three PCRs (3 x 50 µl
737 volume) per sample were combined and cleaned with AMPure PB beads according to
738 PacBio's barcoded overhang adapter protocol (Pacific Biosciences #101-791-700).
739 Barcoded sample pool purity was analysed on the 5200 Fragment Analyser (Agilent)
740 and final library concentration determined using Invitrogen Qubit 1X HS dsDNA kit.
741 Sample pools were sequenced using the Pacific Biosciences Sequel IIe.

742 Precise CAG/CTG repeat sizes were determined using Repeat Detector[44] (version
743 1.0.15eb445). Unaligned reads were assessed using the restrictive profile with a
744 repeat size range of [0–250]. For each analysis, the –-with-revcomp option was

745 enabled and data was output to a histogram (-o histogram option). Obtained
746 histograms were plotted using GraphPad Prism software (v.10.0.0). As the sequences
747 are unaligned when we determine repeat size, the CCSs containing between 0 and 5
748 CAG/CTGs are removed from analysis as they include a large proportion of truncated
749 PCR products and other sequences that do not align with the *HTT* locus.

750 *Alignment of long reads*

751 The repeat regions were aligned with the NGMLR aligner using the *HTT* region from
752 GRCh38 containing 19 CAG repeats as reference sequence. The resulting SAM files
753 were converted into BAM and .bai index files using Picard and Samtools, respectively.
754 The bam files were checked for percentage of mapped reads using Samtools flagstat
755 command. The indexed BAM files were then visualised using the Integrative Genomics
756 Viewer (IGV 2.16.2)[46] to assess the alignment quality and to confirm the presence of
757 the expected CAG repeat count. Sniffles2[45] was run on the aligned BAM files for
758 identifying rearrangements.

759 *Area under the curve, repeat size indices, and comparisons*

760 To calculate the difference in the area under the curve (Supplementary Fig. 9), we
761 used the histograms generated by Repeat Detector. The total read counts were
762 normalised for each sample and adjusted such that the mode is set to 0. Then we
763 subtracted the frequency of reads of the treated samples from those in the control
764 conditions for each repeat size. The positive values were then added together to
765 produce the difference in the area under the curve. It represents the frequency of
766 alleles that has changed compared to the control samples. We also generated delta
767 plots to compare the differences in repeat size between two treatments. To do so,
768 allele frequencies were normalized to the number of total reads in the samples and
769 grouped into 5-repeat bins. Then the relative frequencies of the experimental samples
770 were deducted from that of the control sample. Statistical distribution analysis using a
771 matched-pair Wilcoxon or unpaired Kolmogorov-Smirnov tests was performed without
772 binning. The instability index was calculated as previously described for images[51], but
773 using normalised read frequencies and a 5% threshold. The expansion and
774 contraction indices were calculated such that expansion index + contraction index =
775 instability index.

776

Mouse housing

C57BL/6J R6/1 mice[81] (B6.Cg-Tg(HDexon1)61Gpb/J) were maintained as hemizygotes with *ad libitum* access to food and water and maintained in a temperature and humidity-controlled environment on a 12 h dark/light cycle. Health and wellbeing of the animals were monitored daily and their weight were checked every week in a T1 biosafety level 2 animal facility. All experimental procedures done in Cardiff followed protocols in accordance with the United Kingdom Animals (Scientific Procedures) Act of 1986. All experimental procedures performed on R6/1 mice were approved by Cardiff University Animal Welfare and Ethical Review Board and carried out under Home Office Licenses P49E8C976 and PP7595333. We did not go beyond 2 months in these experiments because the R6/1 mice become markedly unwell beyond this age and to continue would have risked unnecessary suffering. Stereotactic injections done in wild type mice presented in Supplementary Fig. 6a-d were done at the Center de Recherche du CHU de Québec (Québec, QC, Canada). All procedures on these animals were completed in accordance with the guidelines of the Canadian Council on Animal Care and were approved by the Comité de Protection des Animaux du CRCHUQ-UL under protocol number 2019-49, CHU-17-106.

Stereotactic injections

The R6/1 mice used here were of both sexes and had between 138 and 151 CAG/CTG repeats. Mice were anaesthetised with isoflurane and placed in a stereotactic frame. In P2 animals, a syringe was used to penetrate the skin and skull. In adult (3.5 months old) mice, the scalp was shaved, a longitudinal incision was made to expose the skull surface, and 2 burr holes were drilled above the infusion sites. A 4 µl viral suspension ($4.8x10^{10}$ vg of Cas9D10A v4 and sgCTG/GFP AAVs in 1:1 ratio) was stereotactically injected into the striatum of adult mice according to the Paxinos and Franklin mouse brain atlas (AP +0.8; ML +-1.8; DV -2.2 mm from bregma). 2-months old animals were injected in the same adult coordinates with a 4 µl viral suspension ($7.98x10^{10}$ vg of Cas9D10A v5 and sgCTG-Mut+5/GFP AAVs in 1:2 ratio). P2 mice were stereotactically injected with 1 µl ($1.44x10^{10}$vg of Cas9D10A v4 and sgCTG/GFP AAVs in 1:1 ratio) virus suspension into the striatum (AP +2.3; ML +-1.4; DL 1.8 mm from lambda). Hamilton syringes 701 N 10 µl and 5 µl FIX NDL (26S/51/3) were used for the infusion in adult and post-natal mice respectively. The infusion rate was 200 nL

809 min^{-1}, and the needle remained in place for 5 min after infusion for vector absorption
810 before removal. In adult mice, the injection site was closed with stitches, and mice
811 recovered in incubators. For the wild type animals injected with the different constructs
812 found in Supplementary Fig. 6a-d, each construct was injected together with the
813 sgCTG AAV9 into the cortex and the striatum (AP: +/- 1.6, MD: +/- 1.4, DV -0.75) using
814 a total of $6x10^{10}$ vg, at a ratio of Cas9D10A to sgCTG of 1:1, divided equally between
815 the two hemispheres. Animals were between 8 and 9 months of age.

816 *Behavioural tests*

817 Male and female mice were injected at 2 months of age and tested 1 month post-
818 injection. Mice were handled for 1 minute per day for 5 days each before the first
819 behavioural test. For motor coordination assessment, mice were placed on a rotarod
820 with a fixed speed of 12 rpm and 24 rpm (two total trials, 3 hours apart) and the time
821 of the first fall and the total number of falls were recorded until the sum of latencies to
822 fall reached a total of 60 s per trial. For accelerated rotarod, animals were placed on
823 the rod with a constant increase from 5 to 40 rpm over the 5-minute trial. A balance
824 beam test was used to evaluate fine motor coordination and balance. Mice were
825 placed at one end of the beam, and the time to reach an escape box containing nesting
826 material and located in the opposite end was recorded. The beam dimensions were:
827 L80 cm, W0.5–1.5 cm, H34–54 cm with an incline of 17°. The house dimensions were:
828 L11 cm, W11 cm, H10 cm. Mice were allowed to rest for 3 hours before the next trial
829 with a total of two trials. Muscular strength was measured by placing the mice on top
830 of a wire rectangle of approximately 20 cm × 21 cm, surrounded by tape to prevent
831 mice walking off the edge, and after a light shaking so mice gripped the wires, the lid
832 was turned upside down. Latency to the first fall was recorded within the total 60 s of
833 the test. Open field test was conducted by positioning the animal at the centre of the
834 white open field arena (80 cm x 80 cm x 30 cm). The locomotive behaviour was
835 monitored for a duration of 5 minutes with the EthoVision tracking software. All
836 experiments were conducted during the light phase and were performed and analysed
837 in a blinded manner.

838

839

840

841 *Perfusion and immunostaining of mouse tissues*

842 R6/1 animals were deeply anaesthetised with Dolethal (Vetoquino) and transcardially
843 perfused with 4% PFA in PBS. Brains were removed, 2 h post-fixed in 4% PFA, and
844 stored in 30% sucrose in PBS at 4°C. Wild type animals presented in Supplementary
845 Figure 6a-d were anaesthetised the same way, but the perfusion was done with PBS.
846 30 µm coronal sections were cut with a freezing sliding microtome and stored at −20°C
847 in cryoprotectant solution (30% ethylene glycol, 30% glycerol, 20 mM PBS) until
848 processing. Free-floating brain sections were washed in PBS and an antigen retrieval
849 step was performed. For IBA1, GFAP and HA antibodies immunostaining, tissue
850 sections were incubated for 30 minutes at 75°C in sodium citrate buffer pH 6. For
851 GFPs, NeuN and DARPP-32 immunostaining, tissue sections were boiled in citrate
852 buffer pH 6 for 2 minutes. Brain sections were then washed in PBS and blocked and
853 permeabilized for 1 h in 1% BSA, 0.2% Triton X-100, and 4% NGS in PBS for IBA1
854 and GFAP antibodies; 1 h in 1% BSA, 0.3% Triton X-100, and 4% NGS in PBS for
855 NeuN and DARPP-32 antibodies; 1 h in 1% BSA, 0.5% Triton X-100, and 4% NGS in
856 PBS for HA antibody. Then, floating sections were incubated with primary antibodies
857 in 1% BSA, 0.1% Triton X-100, and 1% NGS in PBS overnight at 4°C. After washing
858 with PBS, sections were incubated with appropriate Alexa-conjugated secondary
859 antibody in 1% BSA, 0.1% Triton X-100 in PBS for 1 h at room temperature. Tissues
860 were stained with DAPI and after several washes in PBS, sections were mounted onto
861 SuperFrost Plus slides (Thermo Fisher Scientific), air-dried, and coverslipped with
862 Fluoromount aqueous mounting medium (Merck). Leica SP8 confocal (LAS X
863 software) and Evos FL Auto 2 (Invitrogen EVOS FL Auto 2.0 Imaging System)
864 microscopes were used to capture images.

865 Quantification of the inclusion bodies was done using the total number of cells or
866 nuclei, as measured by DAPI staining, taking 3 to 4 images from 3 different mice at
867 several different antero-posterior and dorso-ventral coordinates in the striatum. The
868 reported decreases were calculated as 100% minus the percentage of the Cas9D10A
869 treatment alone. We reported the results the same way for the size of the inclusion
870 bodies and their intensities. All image capture and analysis were done where we are
871 blind to the treatment done. Total cells and inclusions quantification were performed
872 with ImageJ (v2.3.051).

873 *Western blotting*

874 To assess Cas9 levels in astrocytic and lymphoblastoid cultures, proteins were
875 extracted using RIPA buffer supplemented with cOmplete^{TM} EDTA-free protease
876 inhibitors. Mouse brain tissues were mechanically disaggregated in RIPA buffer
877 supplemented with cOmplete^{TM} EDTA-free protease inhibitors using a mini hand-held
878 homogenizer (ThermoFisher). Samples were incubated for 30 min in a rotator at 4°C
879 and soluble fractions were collected by centrifugation (30 min; 13000g at 4°C). Protein
880 concentrations were determined using Pierce BCA protein assay (ThermoFisher). 10
881 to 30 µg protein was loaded on 4-12% Bis-Tris Plus precast gels (ThermoFisher) and
882 run in 1X MES buffer (ThermoFisher) alongside Bio-Rad Kaleidoscope molecular
883 weight marker. Proteins were transferred onto a nitrocellulose membrane using the
884 Bolt system (ThermoFisher). The results were imaged on the Odyssey Imager (LI-
885 COR Biosciences).

886 *HTRF assay*

887 HTRF assay was performed as previously described[53,82]. In summary, a 5% (w/v) total
888 protein homogenate was prepared in ice-cold bioassay buffer (PBS, 1% Triton-X-100)
889 with complete protease inhibitor cocktail tablets (Roche), by homogenizing three times
890 for 30 s in lysing matrix D tubes at 6.5 m s^{-1} (MP Biomedicals) in a Fast-Prep-24TM
891 instrument (MP Biomedicals). Lysates were snap frozen and used for assays the
892 following day. Tissue homogenates to a final volume of 10 µL were pipetted in triplicate
893 into a 384-well proxiplate (Greiner Bio-One). Antibody concentrations used were 1 ng
894 of donor per well and 40 ng of acceptor per well. For HTRF assays, antibodies were
895 added per well in 5 µL HTRF detection buffer [50 mM NaH2PO4, 0.2 M KF, 0.1%
896 bovine serum albumin, 0.05% Tween-20] with complete protease inhibitor cocktail
897 tablets (Roche). Plates were incubated for 3 h on an orbital shaker (250 rpm) at room
898 temperature, before reading on an EnVision (Revvity) plate reader using optimised
899 HTRF detection parameters as described previously[82] .

900 *Single-nuclei RNA sequencing and analysis*

901 <u>Library preparation and sequencing:</u> Nuclear isolation was performed in a Genomics
902 Chromium platform from 12 striatal samples. Droplet-based snRNA sequencing
903 libraries were prepared at the UK DRI Single Cell and Spatial Omics Facility using the

904 Chromium Next GEM Single Cell 3' Kit v3.1 (10x Genomics, Pleasanton CA) according
905 to the manufacturer's protocol. The library was sequenced at UCL genomic facility
906 using NovaSeq 6000 S4 v1.5 (200 Cycles) with configuration 28-10-10-90.

907 Processing, quality control, and filtering: Raw FASTQ files were processed using Cell
908 Ranger v8.0.0 [83](10x Genomics) on a high-performance computing cluster. A custom
909 mouse reference genome was created by adding GFP, Cas9D10A, and human
910 huntingtin exon1 (*huHTT*) transgene sequences to the standard mouse genome
911 (GRCm39). The data were processed using Seurat v5.0 in R[84–87]. Quality control
912 metrics were calculated including mitochondrial gene percentage using the
913 PercentageFeatureSet() function with the pattern "^mt-". Cells were filtered to retain
914 those with 200-5000 detected genes and less than 5% mitochondrial reads. Sex
915 information was mapped to each sample and converted to numeric format for
916 downstream regression.

917 Normalization and batch effect correction: Unique Molecular Identifier (UMI) counts
918 from each cell were normalized using SCTransform with the mitochondrial gene
919 percentage, total UMI count (nCount_RNA), and sex regressed out using
920 vars.to.regress[88]. Principal component analysis was performed on the normalized data
921 using RunPCA() with default parameters, and the first 20 principal components were
922 selected for downstream analysis. Batch effect correction was performed using
923 Harmony integration implemented through the RunHarmony() function[89]. Integration
924 was performed on the SCTransform-normalized data using sample identity (orig.ident)
925 as the batch variable with theta = 1. The integrated representation was saved as
926 "pca_integrated" for subsequent clustering and visualization.

927 Cell clustering, visualization, and identification: Cell clustering was performed using
928 the integrated principal components with the FindNeighbors() and FindClusters()
929 functions in Seurat[84–87,90]. A resolution of 0.5 was used for Louvain clustering. UMAP
930 (Uniform Manifold Approximation and Projection) dimensionality reduction was
931 performed using RunUMAP() on both unintegrated and integrated PCA embeddings
932 with default parameters.

933 Cell type annotation was performed using a curated list of marker genes for 13 cell
934 types[58]. Direct and indirect medium spiny neurons were pooled for the subsequent

935 analyses. Module scores were calculated for each cell type using AddModuleScore()
936 with the respective marker gene sets. Automatic cell type assignment was performed
937 by calculating average module scores for each cluster and assigning the cell type with
938 the highest average score using Seurat v5.0 [84–88,90].

939 Single-nuclei differential expression analysis: Cell type-specific differential expression
940 analysis was performed using FindMarkers() implemented in Seurat v5.0 [84–88,90]. Prior
941 to analysis, SCTransform models were prepared using PrepSCTFindMarkers() to
942 ensure compatibility across sample subsets.

943 Differential expression analysis was conducted using Seurat's FindMarkers function
944 with the MAST framework. Analysis was restricted to genes expressed in at least 10%
945 of cells in either comparison group, with logfc.threshold = 0.

946 Pseudo-bulk differential expression analysis: We performed pseudobulk differential
947 expression (DE) analysis separately on each annotated cell type using DESeq2
948 (v1.48.2)[91]. For each annotated cell type, raw UMI counts from the Seurat object were
949 aggregated by its *orig.ident* to generate sample-level pseudobulk count matrices,
950 treating each replicate as a separate unit for statistical testing. Genes with less than
951 10 UMI counts in all samples or expressed in less than two replicates per group were
952 excluded. Differential expression testing was performed using a design formula of ~
953 sample_group. For each comparison, DESeq2 was utilized to estimate dispersion and
954 fit negative binomial generalized linear models. This resulted in csv files with
955 normalized counts and differential gene expression statistics for all genes
956 (Supplementary Tables 5 and 6).

957 Filtering and pathway enrichment analysis: The threshold for significance for
958 differentially expressed genes were adjusted p-value < 0.01 and absolute $\log_2$ fold
959 change > 0.585 . Pathway enrichment analysis was performed on these filtered gene
960 sets using the enrichR package. The Gene Ontology Cellular Component (2023
961 version) was used and the enrichment results were ranked by p-value and combined
962 score, with specific tracking of Huntington's disease-related pathways across all
963 comparisons. For KEGG pathway analysis, Tubb3 emerged as the most stable
964 housekeeping gene calculated by the coefficient of variation and used as
965 housekeeping gene for DGE normalization. We used a targeted gene panel from the

966 KEGG Huntington's Disease pathway (hsa05016, MSigDB M13486), which is
967 available through the Molecular Signatures Database.

968 Comparison with previously published datasets: To compare gene expression at the
969 single-nucleus level against published datasets, we first subdivided our MSN
970 population into the direct (dMSN) and indirect (iMSN) medium spiny neurons. For each
971 subtype, we extracted lists of significantly differentially expressed genes comparing
972 R6/1 mice and wildtype littermates treated with the Cas9D10A AAV only. These were
973 compared to differentially expressed genes reported by Lee et al[58] and Lim et al[92] for
974 R6/2 mice as well as compared to zQ175 mice[58]. We determined the direction of gene
975 expression changes by using the sign of the $\log_2$ fold-change in each dataset and
976 assessed whether they were concordant with our observations. A Fisher's exact test
977 was used to determine whether the number of common / divergent genes were
978 different between datasets (Supplementary Fig. 17).

979 Off-target enrichment analysis: We extracted the set of genes overlapping Cas9D10A
980 potential off-target loci in the mouse genome. We used a Fisher's exact test
981 (Supplementary Table 5) to determine whether off-target containing genes were more
982 likely to be differentially regulated compared to the total number of dysregulated
983 genes. We used the pseudobulk medium spiny neurons for this analysis
984 (Supplementary Fig. 16e-g).

985 *Q-PCR*

986 Quantitative real-time PCR (qPCR) was performed using the Applied Biosystems
987 QuantStudio 7 Flex Real-Time PCR system and analysed with the QuantStudio Real-
988 Time PCR Software (v1.7 Thermo Fisher). Briefly, qPCR was assayed in a total
989 volume of 10 µl reaction mixture containing the ready-to-use FastStart Universal SYBR
990 Green Master (ROX) (Thermo Fisher) and 0.5 µM of forward and reverse mix
991 appropriate primers (Supplementary Table 12). All qPCR reactions were run in
992 triplicates. Mean cycle threshold (Ct) values for each reaction were recorded and the
993 relative DNA copies were calculated and normalised to Actin (2^(ActinCt-geneCt)).

994 *Statistics*

995 Total number of alleles amplified by Small-pool PCR was calculated using Poisson
996 distribution as described[30]. We determined P-values for small-pool PCR using a Mann-

997 Whitney nonparametric U-test comparing control and Cas9D10A + sgCTG conditions.
998 We determined statistical significance between the Cas9D10A only and Cas9D10A +
999 sgCTG HD iPSC-derived neuronal populations shown in Supplementary Fig. 2cd
1000 using a Student's t-test. A nonparametric paired Wilcoxon test was used to compare
1001 two hemispheres from the same animal and a nonparametric unpaired Kolmogorov-
1002 Smirnov test when comparing samples from different animals. Delta plots used
1003 aggregated data visualization purposes only, not for statistical testing. A Student's t-
1004 test was also used HTRF and aggregate measures. Furthermore, differences in
1005 instability indices over time were calculated using a one-way ANOVA with a post-hoc
1006 Tukey's multiple comparison test. We used two-way ANOVA to estimate differences
1007 on treatment and time, followed by a post-hoc Tukey's multiple comparison test. The
1008 same test was used for behavioral analysis. For repeat size comparison, in allele
1009 frequency or delta plot analysis, after long-read sequencing two tests were used. For
1010 comparison of repeat instability between striatal samples within the same animal a
1011 paired nonparametric Wilcoxon signed rank test was used. For comparison of repeat
1012 sizes between different animals and in the lymphoblastoid experiments, an unpaired
1013 nonparametric Kolmogorov-Smirnov test was used. Statistical analyses were done
1014 using GraphPad Prism software version 10.0.0 and were always two-tailed tests using
1015 a significance cut off at P=0.05.

1016 *Data availability statement*

1017 The CCSs and snRNA-seq sequences are available from SRA
1018 (https://www.ncbi.nlm.nih.gov/sra) PRJNA1077893.

1019 *Code availability statement*

1020 Custom scripts are available at https://github.com/DionLab/Murillo-et-al.

1021

1022 **Acknowledgements**
1023 Mouse image (Fig. 5b) and brain slice (Fig. 2bj&5a) were created with BioRender.com
1024 (agreement numbers are MK28EFZJ1S and JM28EFZ3G9, respectively). We would
1025 like to thank Andrew Jefferson for training and for keeping the fluorescence
1026 microscopes in excellent working order and Joanne Morgan and the core team at the
1027 Centre for Neuropsychiatric Genetics and Genomics at Cardiff University for support
1028 with sequencing. Also, we would also like to thank Daria Gavriouchkina and Modesta
1029 Blunskyte from the UK DRI Single Cell and Spatial Omics Facility for their help in single
1030 nuclei RNA-seq library preparation. We thank the UCL genomic facility staff for the
1031 snRNA-seq sequencing. We thank Vanessa Wheeler and Geneviève Gourdon for
1032 critical reading of the manuscript. We also thank Sarah Langley for advice with the
1033 snRNA-seq data analysis and Sandeep Sundara Rajan for designing the Cas9D10A
1034 v5. This research was undertaken using the supercomputing facilities at Cardiff
1035 University operated by Advanced Research Computing at Cardiff (ARCCA) on behalf
1036 of the Cardiff Supercomputing Facility and the HPC Wales and Supercomputing Wales
1037 (SCW) projects. We acknowledge the support of the latter, which is part-funded by the
1038 European Regional Development Fund (ERDF) via the Welsh Government. We also
1039 thank the University of Lausanne for providing some plasmids used in this study.
1040
1041 **Funding**
1042 This work is supported by the UK Dementia Research Institute [award numbers UK
1043 DRI-TAP2022FA2, DRI-TAP2022FA14, and UK DRI-3204] through UK DRI Ltd,
1044 principally funded by the UK Medical Research Council. VD is further supported by a
1045 professorship from the Academy of Medical Sciences (AMSPR1\1014) and the
1046 Moondance Foundation Laboratories. This project is part of the UCL Neurogenetic
1047 Therapies Programme (to V.D.), generously funded by The Sigrid Rausing Trust. It
1048 was partially supported by a grant from Huntington Society of Canada (to FC, VD, and
1049 Vanessa Wheeler) and by the Myotonic Dystrophy Foundation (to VD, JP, and
1050 Geneviève Gourdon). NS was funded by the European Union's Horizon 2020 research
1051 and innovation programme awarded to NDA under the Marie Skłodowska-Curie Grant
1052 Agreement No. 813851. GPB is supported by the CHDI Foundation. A.M. holds a
1053 Future Leaders in Neuroscience Research Award from the Neuroscience and Mental
1054 Health Innovation Institute at Cardiff University. MJL is funded by an MRC Programme
1055 grant (MR/T033428/1).

Author contributions

AMurillo: Conceptualization, Formal Analysis, Investigation, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing

MA: Conceptualization, Formal Analysis, Investigation, Methodology, Visualization, Writing – review & editing

RRPD: Formal Analysis, Visualization, Writing – review & editing

ELR: Methodology, Writing – review & editing

MLarin: Conceptualization, Formal Analysis, Investigation, Methodology, Visualization, Writing – review & editing

LH: Methodology, Writing – review & editing

ANA: Methodology, Writing – review & editing

AST: Formal Analysis, Visualization, Writing – review & editing

AMM: Resources, Writing – review & editing

NS: Investigation, Resources, Writing – review & editing

AERH: Formal Analysis, Writing – review & editing

PA: Methodology, Writing – review & editing

KF: Methodology, Writing – review & editing

CL: Methodology, Writing – review & editing

GFO: Methodology, Writing – review & editing

AMangin: Methodology, Writing – review & editing

SB: Methodology, Writing – review & editing

EM: Resources, Supervision, Writing – review & editing

NDA: Resources, Funding acquisition, Supervision, Writing – review & editing

JP: Conceptualization, Supervision, Funding Acquisition, Writing – review & editing

GPB: Resources, Supervision, Writing – review & editing

BLD: Resources, Supervision, Writing – review & editing

FC: Conceptualization, Funding acquisition, Supervision, Writing – review & editing

MLelos: Conceptualization, Funding acquisition, Supervision, Writing – review & editing

VD: Conceptualization, Funding acquisition, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing

Conflict of interests

V.D. declares that he has had a research contract with Pfizer Inc unrelated to this work in the past 5 years. B.L.D. serves on the advisory board of Latus Bio, Seamless Bio, and Carbon Biosciences and has sponsored research unrelated to this work from Latus Bio. E.M. is an inventor listed under patent WO2023105212A1, cited herein and V. D, A. Murillo, A. Mangin, and E.M. are named inventors on patent application

EP25387174.3, which covers systems and compositions. All other authors declare no competing interests.

References

1. Khristich, A. N. & Mirkin, S. M. On the wrong DNA track: Molecular mechanisms of repeat-mediated genome instability. *J Biol Chem* **295**, 4134–4170 (2020).
2. The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**, 971–983 (1993).
3. Fu, Y. H. *et al.* An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* **255**, 1256–1258 (1992).
4. Tabrizi, S. J. *et al.* Potential disease-modifying therapies for Huntington's disease: lessons learned and future opportunities. *Lancet Neurol* **21**, 645–658 (2022).
5. Pascual-Gilabert, M., Artero, R. & López-Castel, A. The myotonic dystrophy type 1 drug development pipeline: 2022 edition. *Drug Discov Today* **28**, 103489 (2023).
6. McAllister, B. *et al.* Timing and Impact of Psychiatric, Cognitive, and Motor Abnormalities in Huntington Disease. *Neurology* **96**, e2395–e2406 (2021).
7. Morales, F. *et al.* Individual-specific levels of CTG•CAG somatic instability are shared across multiple tissues in myotonic dystrophy type 1. *Human Molecular Genetics* **32**, 621–631 (2023).
8. Kennedy, L. *et al.* Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. *Human Molecular Genetics* **12**, 3359–3367 (2003).
9. Mätlik, K. *et al.* Cell-type-specific CAG repeat expansions and toxicity of mutant Huntingtin in human striatum and cerebellum. *Nat Genet* 1–12 (2024) doi:10.1038/s41588-024-01653-6.
10. Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium. Identification of

1121 Genetic Factors that Modify Clinical Onset of Huntington's Disease. *Cell* **162**, 516–
1122 526 (2015).
1123 11. Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium. CAG Repeat Not
1124 Polyglutamine Length Determines Timing of Huntington's Disease Onset. *Cell* **178**,
1125 887-900.e14 (2019).
1126 12. Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium. Genetic
1127 modifiers of somatic expansion and clinical phenotypes in Huntington's disease
1128 highlight shared and tissue-specific effects. *Nat Genet*
1129 https://doi.org/10.1038/s41588-025-02191-5 (2025) doi:10.1038/s41588-025-
1130 02191-5.
1131 13. Bettencourt, C. *et al.* DNA repair pathways underlie a common genetic mechanism
1132 modulating onset in polyglutamine diseases. *Annals of Neurology* **79**, 983–990
1133 (2016).
1134 14. McAllister, B. *et al.* Exome sequencing of individuals with Huntington's disease
1135 implicates FAN1 nuclease activity in slowing CAG expansion and disease onset. *Nat*
1136 *Neurosci* **25**, 446–457 (2022).
1137 15. Yang, S. *et al.* CRISPR/Cas9-mediated gene editing ameliorates neurotoxicity in
1138 mouse model of Huntington's disease. *J Clin Invest* **127**, 2719–2724 (2017).
1139 16. Merienne, N. *et al.* The Self-Inactivating KamiCas9 System for the Editing of CNS
1140 Disease Genes. *Cell Rep* **20**, 2980–2991 (2017).
1141 17. Duarte, F. *et al.* Limitations of Dual-Single Guide RNA CRISPR Strategies for the
1142 Treatment of Central Nervous System Genetic Disorders. *Human Gene Therapy* **34**,
1143 958–974 (2023).
1144 18. Duarte, F. *et al.* Semi-automated workflows to quantify AAV transduction in various

1145 brain areas and predict gene editing outcome for neurological disorders. *Molecular*
1146 *Therapy - Methods & Clinical Development* **29**, 254–270 (2023).
1147 19. Lo Scrudato, M. *et al.* Genome Editing of Expanded CTG Repeats within the Human
1148 DMPK Gene Reduces Nuclear RNA Foci in the Muscle of DM1 Mice. *Molecular*
1149 *Therapy* **27**, 1372–1388 (2019).
1150 20. Cardinali, B. *et al.* Time-controlled and muscle-specific CRISPR/Cas9-mediated
1151 deletion of CTG-repeat expansion in the DMPK gene. *Molecular Therapy - Nucleic*
1152 *Acids* **27**, 184–199 (2022).
1153 21. Monteys, A. M., Ebanks, S. A., Keiser, M. S. & Davidson, B. L. CRISPR/Cas9 Editing of
1154 the Mutant Huntingtin Allele In Vitro and In Vivo. *Mol Ther* **25**, 12–23 (2017).
1155 22. Yan, S. *et al.* Cas9-mediated replacement of expanded CAG repeats in a pig model
1156 of Huntington's disease. *Nat. Biomed. Eng* **7**, 629–646 (2023).
1157 23. Choi, D. E. *et al.* Base editing strategies to convert CAG to CAA diminish the
1158 disease-causing mutation in Huntington's disease. *eLife* **12**, RP89782 (2024).
1159 24. Matuszek, Z. *et al.* Base editing of trinucleotide repeats that cause Huntington's
1160 disease and Friedreich's ataxia reduces somatic repeat expansions in patient cells
1161 and in mice. *Nat Genet* 1–15 (2025) doi:10.1038/s41588-025-02172-8.
1162 25. Izzo, M. *et al.* Muscle-specific gene editing improves molecular and phenotypic
1163 defects in a mouse model of myotonic dystrophy type 1. *Clin Transl Med* **15**, e70227
1164 (2025).
1165 26. Nakamori, M. *et al.* A slipped-CAG DNA-binding small molecule induces
1166 trinucleotide-repeat contractions in vivo. *Nat Genet* **52**, 146–159 (2020).
1167 27. Bečanović, K. *et al.* A SNP in the HTT promoter alters NF-κB binding and is a
1168 bidirectional genetic modifier of Huntington disease. *Nat Neurosci* **18**, 807–816

1169 (2015).

1170 28. Van Raamsdonk, J. M. *et al.* Loss of wild-type huntingtin influences motor
1171 dysfunction and survival in the YAC128 mouse model of Huntington disease.
1172 *Human Molecular Genetics* **14**, 1379–1392 (2005).

1173 29. Fang, L. *et al.* Haplotyping SNPs for allele-specific gene editing of the expanded
1174 huntingtin allele using long-read sequencing. *Human Genetics and Genomics*
1175 *Advances* **4**, 100146 (2023).

1176 30. Cinesi, C., Aeschbach, L., Yang, B. & Dion, V. Contracting CAG/CTG repeats using
1177 the CRISPR-Cas9 nickase. *Nat Commun* **7**, 13272 (2016).

1178 31. Aeschbach, L. & Dion, V. Minimizing carry-over PCR contamination in expanded
1179 CAG/CTG repeat instability applications. *Sci Rep* **7**, 18026 (2017).

1180 32. Larin, M. *et al.* Cas9 Nickase-mediated contractions of CAG/CTG repeats are
1181 transcription-dependent and replication-independent. *NAR Molecular Medicine*
1182 ugae013 (2024) doi:10.1093/narmme/ugae013.

1183 33. Jy, S. *et al.* Expression of mutant huntingtin in glial cells contributes to neuronal
1184 excitotoxicity. *The Journal of cell biology* **171**, (2005).

1185 34. Garcia, V. J. *et al.* Huntington's Disease Patient-Derived Astrocytes Display
1186 Electrophysiological Impairments and Reduced Neuronal Support. *Frontiers in*
1187 *Neuroscience* **13**, (2019).

1188 35. Ah, J. *et al.* Frequency of nuclear mutant huntingtin inclusion formation in neurons
1189 and glia is cell-type-specific. *Glia* **65**, (2017).

1190 36. Faideau, M. *et al.* In vivo expression of polyglutamine-expanded huntingtin by
1191 mouse striatal astrocytes impairs glutamate transport: a correlation with
1192 Huntington's disease subjects. *Human Molecular Genetics* **19**, 3053–3067 (2010).

1193 37. Estrada-Sánchez, A. M. & Rebec, G. V. Corticostriatal dysfunction and glutamate
1194 transporter 1 (GLT1) in Huntington's disease: Interactions between neurons and
1195 astrocytes. *Basal Ganglia* **2**, 57–66 (2012).
1196 38. HD iPSC Consortium, T. Induced Pluripotent Stem Cells from Patients with
1197 Huntington's Disease Show CAG-Repeat-Expansion-Associated Phenotypes. *Cell*
1198 *Stem Cell* **11**, 264–278 (2012).
1199 39. Monckton, D. G., Wong, L. J., Ashizawa, T. & Caskey, C. T. Somatic mosaicism,
1200 germline expansions, germline reversions and intergenerational reductions in
1201 myotonic dystrophy males: small pool PCR analyses. *Hum Mol Genet* **4**, 1–8 (1995).
1202 40. Bates, G. P. *et al.* Huntington disease. *Nat Rev Dis Primers* **1**, 15005 (2015).
1203 41. Bunting, E. L. *et al.* Antisense oligonucleotide-mediated MSH3 suppression reduces
1204 somatic CAG repeat expansion in Huntington's disease iPSC-derived striatal
1205 neurons. *Sci Transl Med* **17**, eadn4600 (2025).
1206 42. De Serres-Bérard, T., Pierre, M., Chahine, M. & Puymirat, J. Deciphering the
1207 mechanisms underlying brain alterations and cognitive impairment in congenital
1208 myotonic dystrophy. *Neurobiol Dis* **160**, 105532 (2021).
1209 43. Han, H. A., Pang, J. K. S. & Soh, B.-S. Mitigating off-target effects in CRISPR/Cas9-
1210 mediated in vivo gene editing. *J Mol Med* **98**, 615–632 (2020).
1211 44. Taylor, A. S. *et al.* Repeat Detector: versatile sizing of expanded tandem repeats and
1212 identification of interrupted alleles from targeted DNA sequencing. *NAR Genomics*
1213 *and Bioinformatics* **4**, lqac089 (2022).
1214 45. Smolka, M. *et al.* Detection of mosaic and population-level structural variants with
1215 Sniffles2. *Nat Biotechnol* 1–10 (2024) doi:10.1038/s41587-023-02024-y.
1216 46. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat Biotechnol* **29**, 24–26 (2011).

1217 47. Marco, S., Murillo, A. & Pérez-Otaño, I. RNAi-Based GluN3A Silencing Prevents and
1218 Reverses Disease Phenotypes Induced by Mutant huntingtin. *Mol Ther* **26**, 1965–
1219 1972 (2018).
1220 48. Suzuki, K. *et al.* In vivo genome editing via CRISPR/Cas9 mediated homology-
1221 independent targeted integration. *Nature* **540**, 144–149 (2016).
1222 49. Bubeck, F. *et al.* Engineered anti-CRISPR proteins for optogenetic control of
1223 CRISPR–Cas9. *Nat Methods* **15**, 924–927 (2018).
1224 50. Kovalenko, M. *et al.* Msh2 Acts in Medium-Spiny Striatal Neurons as an Enhancer of
1225 CAG Instability and Mutant Huntingtin Phenotypes in Huntington's Disease Knock-
1226 In Mice. *PLOS ONE* **7**, e44273 (2012).
1227 51. Lee, J.-M. *et al.* A novel approach to investigate tissue-specific trinucleotide repeat
1228 instability. *BMC Systems Biology* **4**, 29 (2010).
1229 52. Dang, Y. *et al.* Optimizing sgRNA structure to improve CRISPR-Cas9 knockout
1230 efficiency. *Genome Biology* **16**, 280 (2015).
1231 53. Landles, C. *et al.* Mutant huntingtin protein decreases with CAG repeat expansion:
1232 implications for therapeutics and bioassays. *Brain Communications* **6**, fcae410
1233 (2024).
1234 54. DiFiglia, M. *et al.* Aggregation of Huntingtin in Neuronal Intranuclear Inclusions and
1235 Dystrophic Neurites in Brain. *Science* **277**, 1990–1993 (1997).
1236 55. Bayram-Weston, Z., Jones, L., Dunnett, S. B. & Brooks, S. P. Comparison of mHTT
1237 Antibodies in Huntington's Disease Mouse Models Reveal Specific Binding Profiles
1238 and Steady-State Ubiquitin Levels with Disease Development. *PLOS ONE* **11**,
1239 e0155834 (2016).
1240 56. Ko, J., Ou, S. & Patterson, P. H. New anti-huntingtin monoclonal antibodies:

1241 implications for huntingtin conformation and its binding proteins. *Brain Research*
1242 *Bulletin* **56**, 319–329 (2001).
1243 57. Landles, C. *et al.* Proteolysis of mutant huntingtin produces an exon 1 fragment that
1244 accumulates as an aggregated protein in neuronal nuclei in Huntington disease. *J*
1245 *Biol Chem* **285**, 8808–8823 (2010).
1246 58. Lee, H. *et al.* Cell Type-Specific Transcriptomics Reveals that Mutant Huntingtin
1247 Leads to Mitochondrial RNA Release and Neuronal Innate Immune Activation.
1248 *Neuron* **107**, 891-908.e8 (2020).
1249 59. Handsaker, R. E. *et al.* Long somatic DNA-repeat expansion drives
1250 neurodegeneration in Huntington's disease. *Cell* **188**, 623-639.e19 (2025).
1251 60. Lim, R. G. *et al.* Huntington disease oligodendrocyte maturation deficits revealed by
1252 single-nucleus RNAseq are rescued by thiamine-biotin supplementation. *Nat*
1253 *Commun* **13**, 7791 (2022).
1254 61. Murillo, A. *et al.* Cas9 nickase-mediated contraction of CAG/CTG repeats at
1255 multiple disease loci. 2024.02.19.580669 Preprint at
1256 https://doi.org/10.1101/2024.02.19.580669 (2024).
1257 62. Brooks, S. P. *et al.* Longitudinal analysis of the behavioural phenotype in R6/1
1258 (C57BL/6J) Huntington's disease transgenic mice. *Brain Research Bulletin* **88**, 94–
1259 103 (2012).
1260 63. Gray, M. *et al.* Full-Length Human Mutant Huntingtin with a Stable Polyglutamine
1261 Repeat Can Elicit Progressive and Selective Neuropathogenesis in BACHD Mice. *J.*
1262 *Neurosci.* **28**, 6182–6195 (2008).
1263 64. Aldous, S. G. *et al.* A CAG repeat threshold for therapeutics targeting somatic
1264 instability in Huntington's disease. *Brain* **147**, 1784–1798 (2024).

1265 65. Gangwani, M. R. *et al.* Neuronal and astrocytic contributions to Huntington's
1266 disease dissected with zinc finger protein transcriptional repressors. *Cell Reports*
1267 **42**, 111953 (2023).
1268 66. Oura, S. *et al.* Precise CAG repeat contraction in a Huntington's Disease mouse
1269 model is enabled by gene editing with SpCas9-NG. *Commun Biol* **4**, 1–13 (2021).
1270 67. Ran, F. A. *et al.* In vivo genome editing using Staphylococcus aureus Cas9. *Nature*
1271 **520**, 186–191 (2015).
1272 68. Wang, Y. *et al.* Therapeutic Genome Editing for Myotonic Dystrophy Type 1 Using
1273 CRISPR/Cas9. *Molecular Therapy* **26**, 2617–2630 (2018).
1274 69. van Agtmaal, E. L. *et al.* CRISPR/Cas9-Induced (CTG·CAG)n Repeat Instability in the
1275 Myotonic Dystrophy Type 1 Locus: Implications for Therapeutic Genome Editing.
1276 *Mol Ther* **25**, 24–43 (2017).
1277 70. Mosbach, V. *et al.* Resection and repair of a Cas9 double-strand break at CTG
1278 trinucleotide repeats induces local and extensive chromosomal deletions. *PLOS*
1279 *Genetics* **16**, e1008924 (2020).
1280 71. Estevez-Fraga, C., Tabrizi, S. J. & Wild, E. J. Huntington's Disease Clinical Trials
1281 Corner: August 2023. *Journal of Huntington's Disease* **12**, 169–185 (2023).
1282 72. Mathews, E. W. *et al.* Suppression of Huntington's Disease Somatic Instability by
1283 Transcriptional Repression and Direct CAG Repeat Binding. *Nat Commun* **16**, 10009
1284 (2025).
1285 73. Hasuike, Y. *et al.* CAG repeat-binding small molecule improves motor coordination
1286 impairment in a mouse model of Dentatorubral-pallidoluysian atrophy. *Neurobiol*
1287 *Dis* **163**, 105604 (2022).
1288 74. Shi, Y., Kirwan, P. & Livesey, F. J. Directed differentiation of human pluripotent stem

1289 cells to cerebral cortex neurons and neural networks. *Nat Protoc* **7**, 1836–1846
1290 (2012).
1291 75. Bae, S., Park, J. & Kim, J.-S. Cas-OFFinder: a fast and versatile algorithm that
1292 searches for potential off-target sites of Cas9 RNA-guided endonucleases.
1293 *Bioinformatics* **30**, 1473–1475 (2014).
1294 76. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler
1295 transform. *Bioinformatics* **25**, 1754–1760 (2009).
1296 77. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008
1297 (2021).
1298 78. Pedersen, B. S. & Quinlan, A. R. Mosdepth: quick coverage calculation for genomes
1299 and exomes. *Bioinformatics* **34**, 867–868 (2018).
1300 79. Li, H. A statistical framework for SNP calling, mutation discovery, association
1301 mapping and population genetical parameter estimation from sequencing data.
1302 *Bioinformatics* **27**, 2987–2993 (2011).
1303 80. Dong, Y. *et al.* Genome-Wide Off-Target Analysis in CRISPR-Cas9 Modified Mice and
1304 Their Offspring. *G3 Genes|Genomes|Genetics* **9**, 3645–3651 (2019).
1305 81. Mangiarini, L. *et al.* Exon 1 of the HD Gene with an Expanded CAG Repeat Is
1306 Sufficient to Cause a Progressive Neurological Phenotype in Transgenic Mice. *Cell*
1307 **87**, 493–506 (1996).
1308 82. Landles, C. *et al.* Development of novel bioassays to detect soluble and aggregated
1309 Huntingtin proteins on three technology platforms. *Brain Commun* **3**, fcaa231
1310 (2021).
1311 83. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells.
1312 *Nat Commun* **8**, 14049 (2017).

1313 84. Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell
1314 analysis. *Nat Biotechnol* **42**, 293–304 (2024).
1315 85. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-
1316 1902.e21 (2019).
1317 86. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell
1318 transcriptomic data across different conditions, technologies, and species. *Nat*
1319 *Biotechnol* **36**, 411–420 (2018).
1320 87. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-
1321 3587.e29 (2021).
1322 88. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell
1323 RNA-seq data using regularized negative binomial regression. *Genome Biology* **20**,
1324 296 (2019).
1325 89. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with
1326 Harmony. *Nat Methods* **16**, 1289–1296 (2019).
1327 90. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction
1328 of single-cell gene expression data. *Nat Biotechnol* **33**, 495–502 (2015).
1329 91. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and
1330 dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).
1331 92. Lim, R. G. *et al.* Huntington disease oligodendrocyte maturation deficits revealed by
1332 single-nucleus RNAseq are rescued by thiamine-biotin supplementation. *Nat*
1333 *Commun* **13**, 7791 (2022).
1334 93. Mali, P. *et al.* RNA-Guided Human Genome Engineering via Cas9. *Science* **339**, 823–
1335 826 (2013).
1336 94. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries

for CRISPR screening. *Nat Methods* **11**, 783–784 (2014).

95. Ruiz Buendía, G. A. *et al*. Three-dimensional chromatin interactions remain stable upon CAG/CTG repeat expansion. *Sci Adv* **6**, eaaz4012 (2020).

96. Dion, V., Lin, Y., Hubert, L., Waterland, R. A. & Wilson, J. H. Dnmt1 deficiency promotes CAG repeat expansion in the mouse germline. *Hum Mol Genet* **17**, 1306–1317 (2008).

Supplementary Material

Cas9 Nickase-Mediated Contraction of CAG/CTG Repeats *in Vivo* is Accompanied by Improvements in Huntington's Disease Pathology

Alvaro Murillo[1], Melanie Alpaugh[2], Ruban Rex Peter Durairaj[1], Emma L. Randall[1], Meghan Larin[1,3], Laura Heraty[1], Alys N. Aston[1], Alysha S. Taylor[1], Alex Mas Monteys[4], Nina Stöberl[5], Aeverie E. R. Heuchan[1], Pascale Aeschlimann[1], Kyle Fears[1], Christian Landles[6], Georgina F Osborne[6], Antoine Mangin[1], Soumyasree Bhattacharyya[1], Emmanouil Metzakopian[7], Nicholas D. Allen[5], Jack Puymirat[8], Gillian P Bates[6], Beverly L. Davidson[4,9], Francesca Cicchetti[10], Mariah J. Lelos[5], and Vincent Dion[1,11]*

1: UK Dementia Research Institute at Cardiff University, Hadyn Ellis Building, Maindy Road, Cardiff, United Kingdom, CF24 4HQ.

2: Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada N1G 2W1.

3: Current address: AstraZeneca R&D 1 Francis Crick Avenue Cambridge, UK CB2 0AA.

4: Raymond G. Perelman Center for Cellular and Molecular Therapeutics, Children's Hospital of Philadelphia, Philadelphia, 19104, USA.

5: School of Biosciences, Cardiff University, The Sir Martin Evans Building, Museum Avenue, Cardiff, CF10 3AX.

6: Department of Neurodegenerative Disease, Huntington's Disease Centre, Queen Square Institute of Neurology, University College London, London, WC1N 3BG, UK.

7: UK Dementia Research Institute, University of Cambridge, Clifford Albutt building, Cambridge biomedical campus, Cambridge, CB2 0AH, UK.

8: Centre de Recherche du CHUQ site LOEX and Department of Neurological Sciences, Hôpital Enfant-Jésus: LOEX, CHU de Québec-Université Laval Research Center, Quebec City, QC, Canada.

9: Department of Pathology, University of Pennsylvania, Philadelphia, 19104, USA.

10: Centre de Recherche du CHU de Québec, Axe Neurosciences; Département de Psychiatrie et Neurosciences, Université Laval, Québec, QC, Canada.

11: Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Maindy Road, Cardiff, UK, CF24 4HQ.

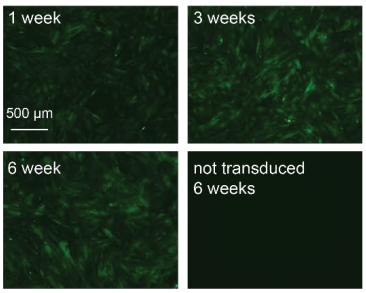
*: corresponding author: dionv@cardiff.ac.uk

A.

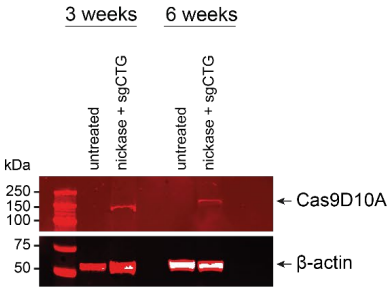
HD iPSC-derived astrocytes

time post transduction	no. nuclei	Ki67⁺ cells	% of Ki67⁺ cells
3 weeks	423	1	0.24
6 weeks	868	2	0.23

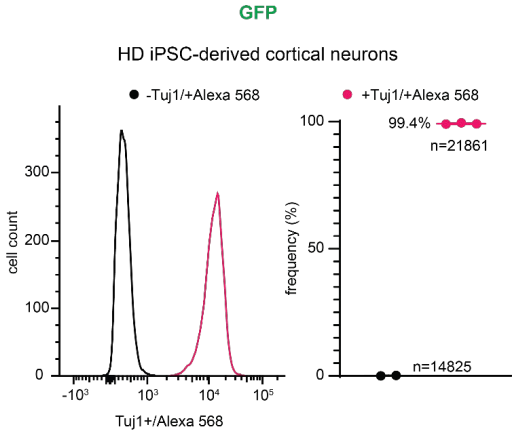
B.



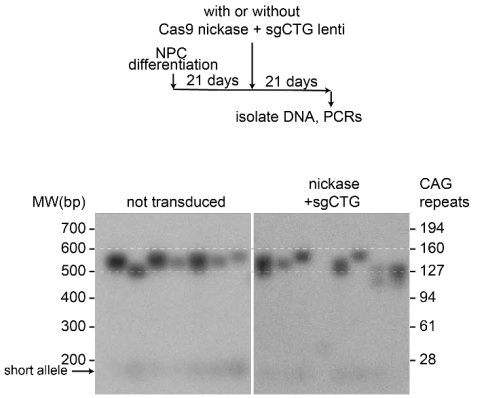
C.



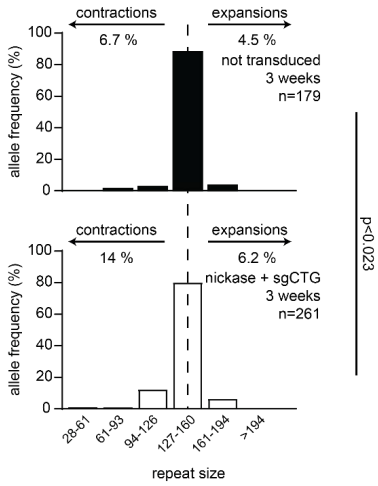
D.



E.



F.



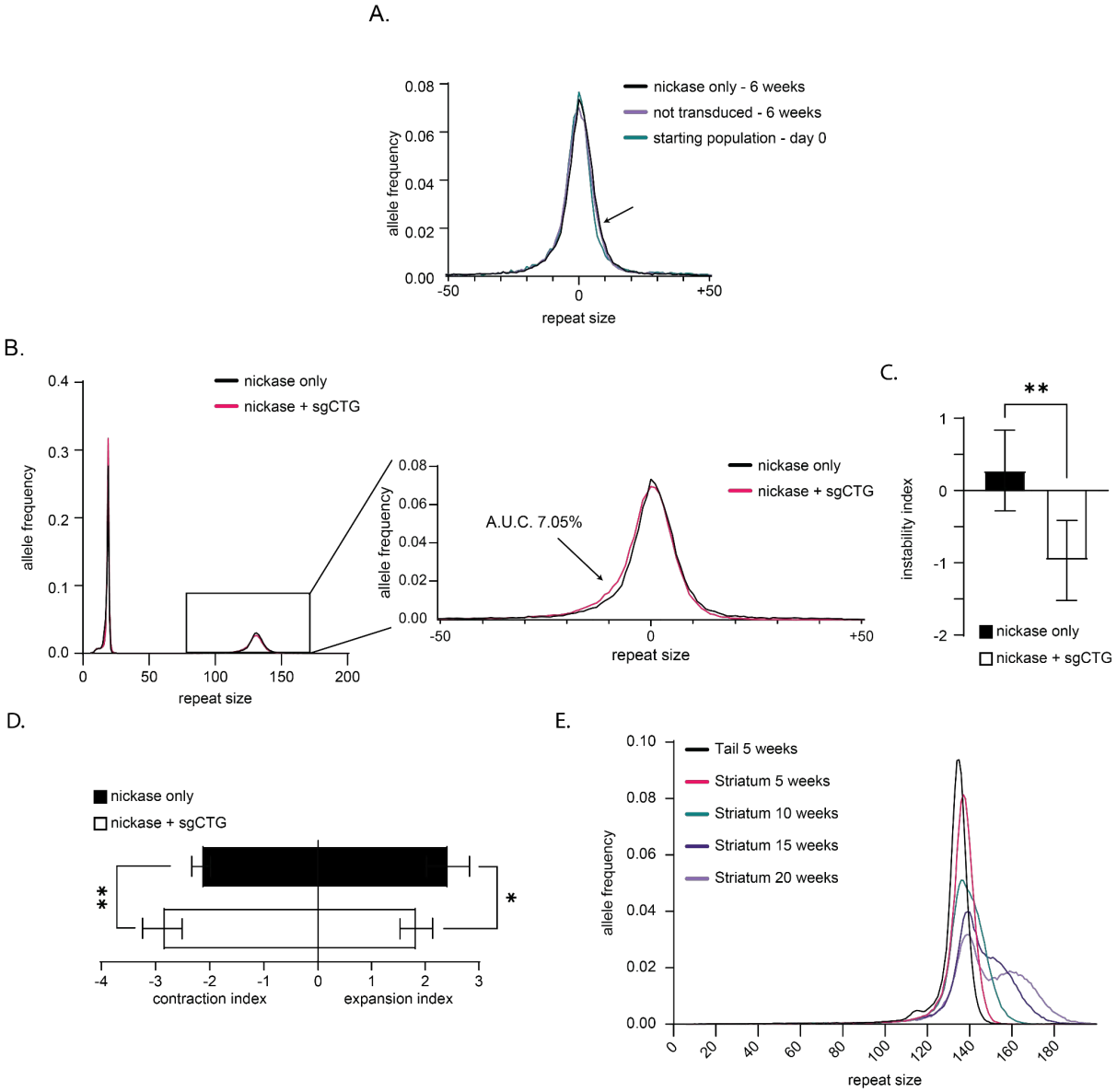
G.

DM1 iPSC-derived cortical neurons

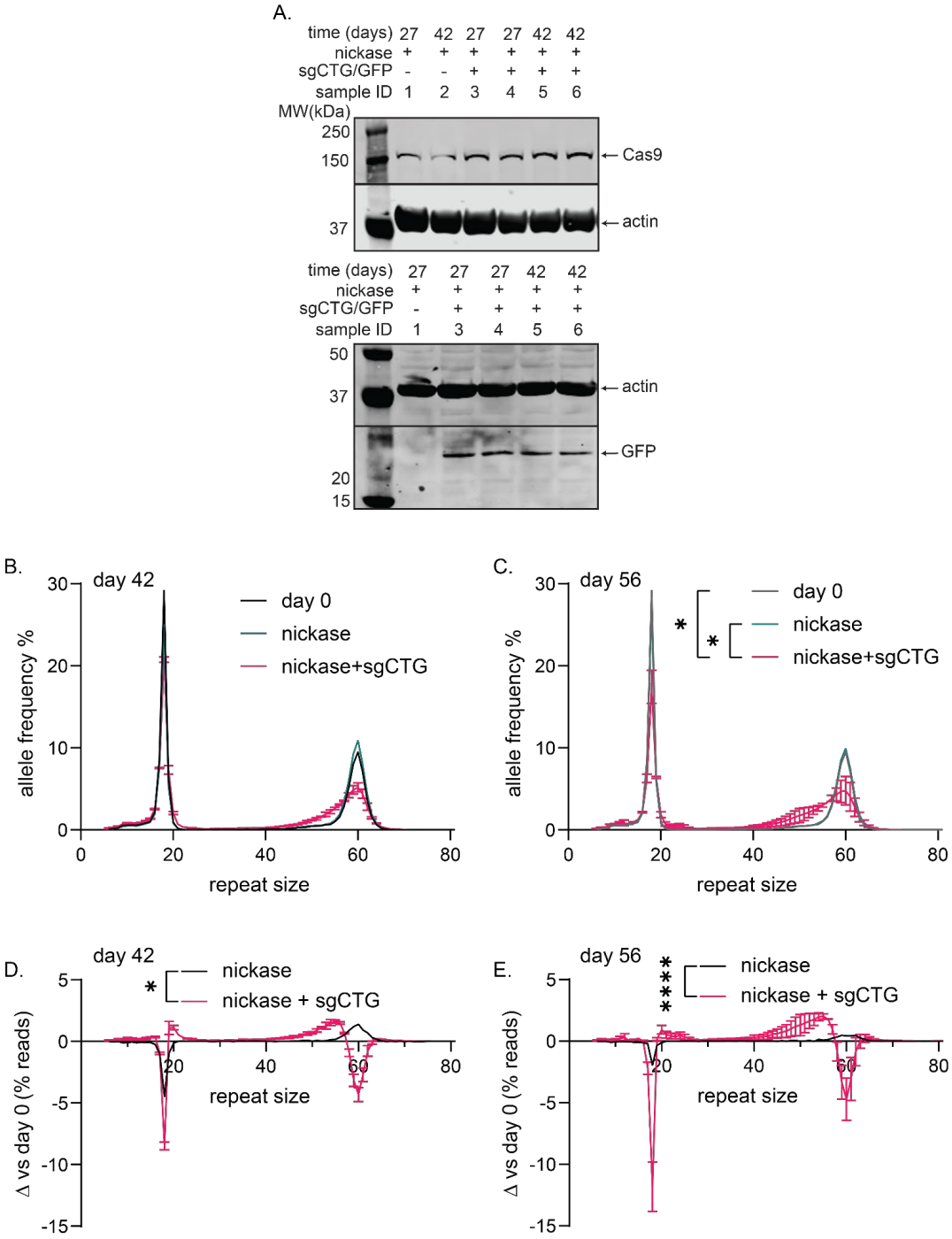
no. nuclei	Tuj1⁺	Ki67⁺ cells	% of Tuj1⁺ cells	% of Ki67⁺ cells
1148	1128	19	98 %	1.7 %

Supplementary Figure 1: CAG/CTG repeat contractions in HD iPSC-derived astrocytes and neurons. A) HD iPSC-derived astrocyte cultures were sorted using CD44, ensuring that they were all of an astrocytic lineage. We found little replication in these cultures over 21 and 42 days after transduction with Cas9D10A and sgCTG as measured by Ki67 staining. B) Lentivirus-encoded GFP expression on the sgCTG vector over 42 days42 days time course in HD iPCS-derived S100β⁺ astrocytes (pLV[gRNA]-CMV-EGFP-U6>{sgCTG}). C) Cas9D10A expression in HD iPSC-derived S100β⁺ astrocytes over 21 and 42 days from (pLenti-EF1alpha-Cas9D10A nickase-Blast). D) Characterisation of the HD iPSC-derived cortical neurons by the expression of Tuj1 as a marker for neurons. Left: Representative flow cytometry histogram of cultures stained with β-tubulin III (Tuj1; magenta) and only with the secondary antibody (black), suggesting low variability in Tuj1 expression and high quality neuronal cultures Right: quantification of 2 to 3 independent experiments. N = number of cells. E) Timeline and representative small pool PCR

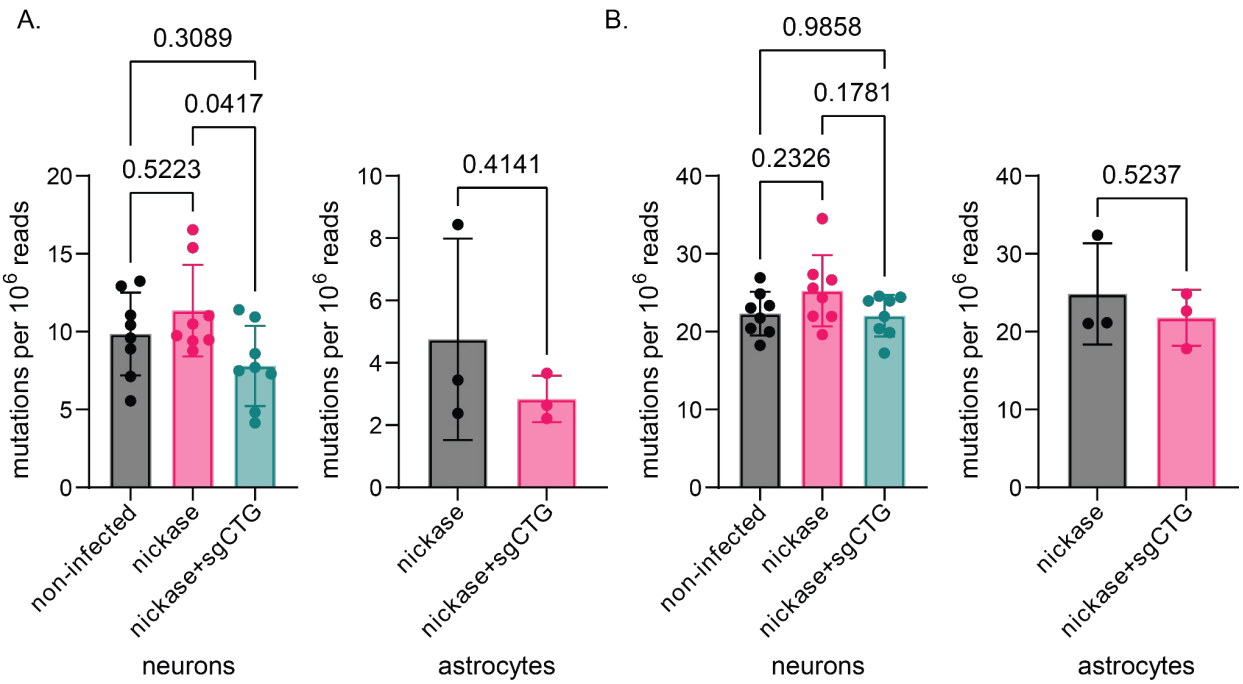
1399 blot of HD iPSC-derived neuronal cultures grown in the presence of both Cas9D10A and sgCTG or
1400 untreated for 21 days. F) Quantification of the small pool PCRs. P-value determined using a Mann-
1401 Whitney U test comparing non-transduced and Cas9D10A + sgCTG. G) DM1 iPSC-derived neuronal
1402 cultures stained for Ki67, as a marker of proliferation, and Tuj1, as a marker of neuronal cell type.



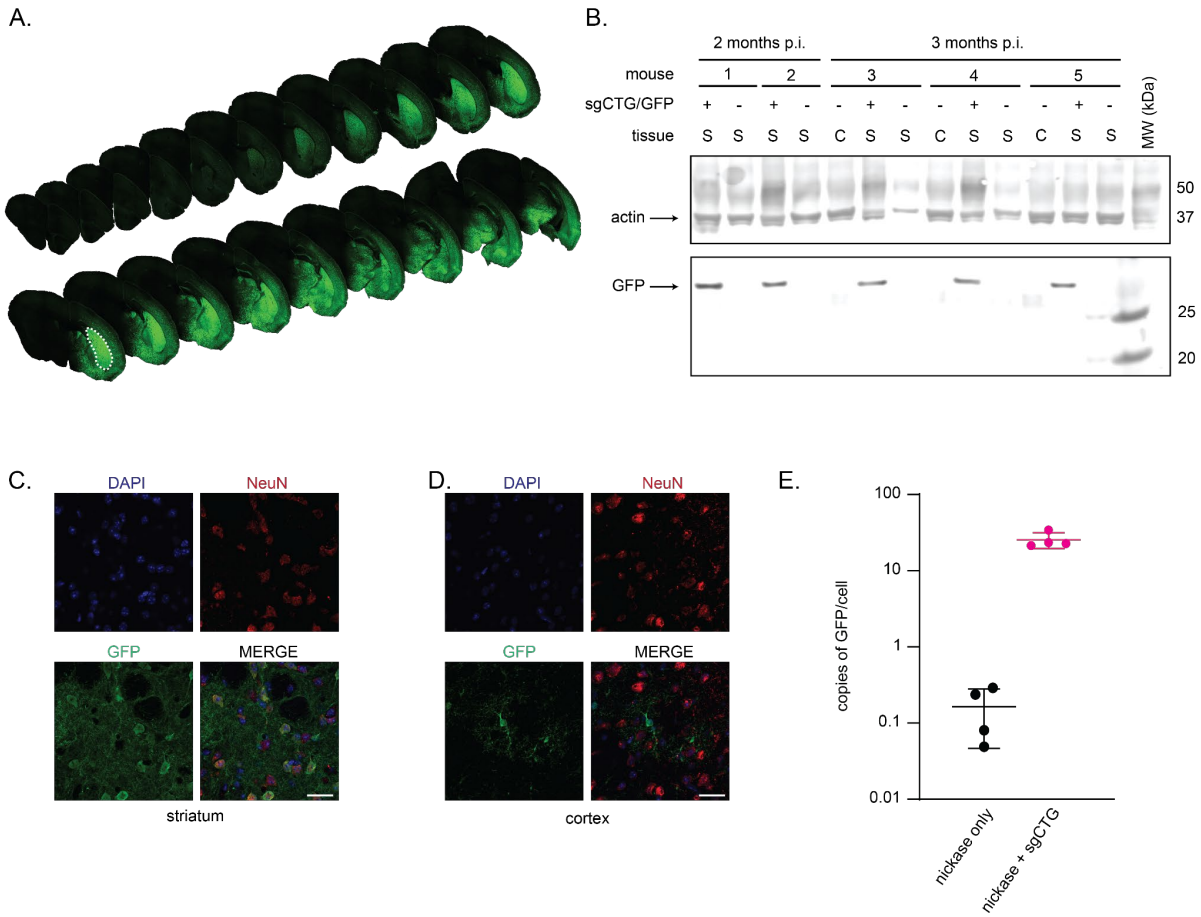
Supplementary Figure 2: SMRT sequencing reproduces expected patterns of repeat size distribution. A) SMRT sequencing of DNA from HD iPSC-derived cortical neurons showing changes in repeat size between cultures of HD-derived cortical neurons that have not been transduced (n=4), transduced with Cas9D10A only (n=4), compared to the starting repeat size distribution (n=2). Arrow indicates the accumulation of expansions over the 42 days period. The calculated AUC difference for expansions between day 0 and non-transduced cells at 42 days was 5.6 % whereas between D0 Cas9D10A only the AUC was 7.8%. B) SMRT sequencing of DNA from HD iPSC-derived cortical neurons showing both the non-pathogenic and the expanded allele (left) and the change in repeat size between the cells treated with Cas9D10A only versus those treated with both Cas9D10A and sgCTG for 42 days (n=4 per condition). Analysis of the difference in the area under the curve (AUC) shows that 7% of the cells transduced with both Cas9D10A and sgCTG had shorter alleles above those treated with Cas9D10A only. This is comparable to small-pool PCR data. C) Instability indices of the HD iPSC-derived cortical neurons treated with Cas9D10A only (black boxes) and Cas9D10A + sgCTG (white boxes) (n=4 per condition). **:P<0.01 calculated using a Student's t-test test. D) Same as in (C) but plotting both the expansion and contraction indices. *: P<0.05. E) Representative example time course analysis using SMRT sequencing showing the expected increase in the frequency of longer alleles over several weeks in the R6/1 striatum. These data show that the approach can measure somatic expansions readily.



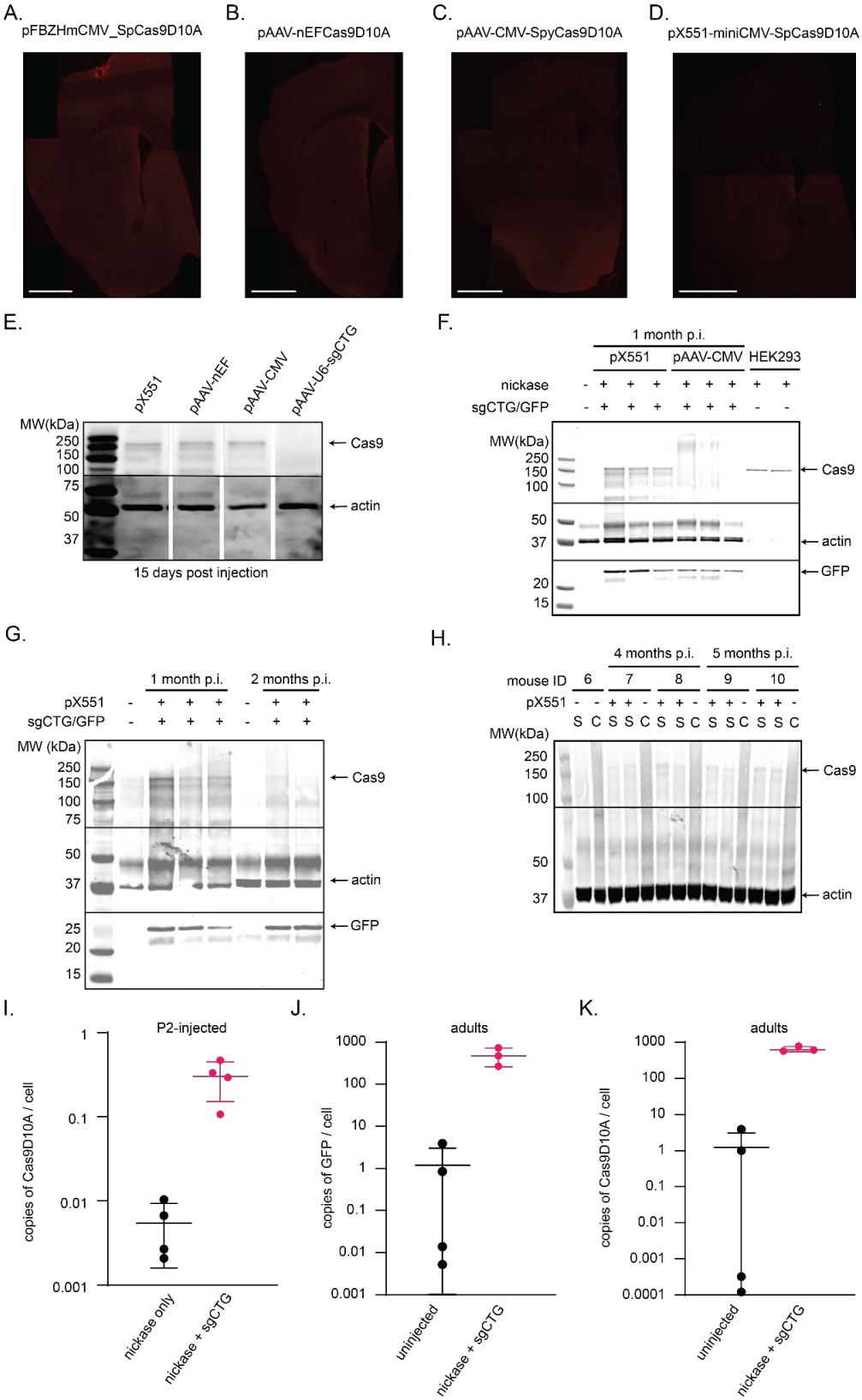
Supplementary Figure 3. As few as 60 CAG/CTG repeats are a target for contraction by the Cas9D10A. A) Western blot of Cas9D10A and GFP expression from lymphoblastoid cultures transduced with pLenti-EF1alpha-Cas9D10A nickase-Blast and pLV(gRNA)-CMV-eGFP:T2A:Hygro-U6(sgCTG) at the indicated post-infection days. Full blots are found in Supplementary Fig. 21. B) Aggregated graphs of repeat size distribution in lymphoblastoid cultures with Cas9D10A only (magenta) and Cas9D10A + sgCTG (blue) at 42 days of treatment versus day 0 (black). C) Same as B at 56 days after the treatment time point. D-E) Delta plots (see methods) of the data presented in B and C showing the difference at 42 days (D) and 56 days (E) after continuous treatment. This is compared to the starting repeat size distribution at D0. Kolmogorov-Smirnov tests to compare cumulative distributions were used in all graphs. P value <0.05 (*); <0.0001 (****). Data are mean ± SD (D0 n = 1; D42 and D56 Cas9D10A only n=1; D42 and D56 Cas9D10A + sgCTG = 2).



Supplementary Figure 4. Analysis of mutations in HD iPSC-derived cells. This figure provides a graphical representation of the data listed in Table 1. A) The number of mutations found at the selected 1830 genes were normalised to the total reads per sample and indicated as the number of mutations normalised to 10^6 reads. B) The number of mutations found genome-wide were normalised to the total reads per sample and indicated as the number of mutations normalised to 10^6 reads. Differences in normalised mutations between neuronal cultures conditions were calculated by one-way ANOVA with a post-hoc Tukey's multiple comparison test, except for the whole genome data from astrocytes where we used a Student's t-test.



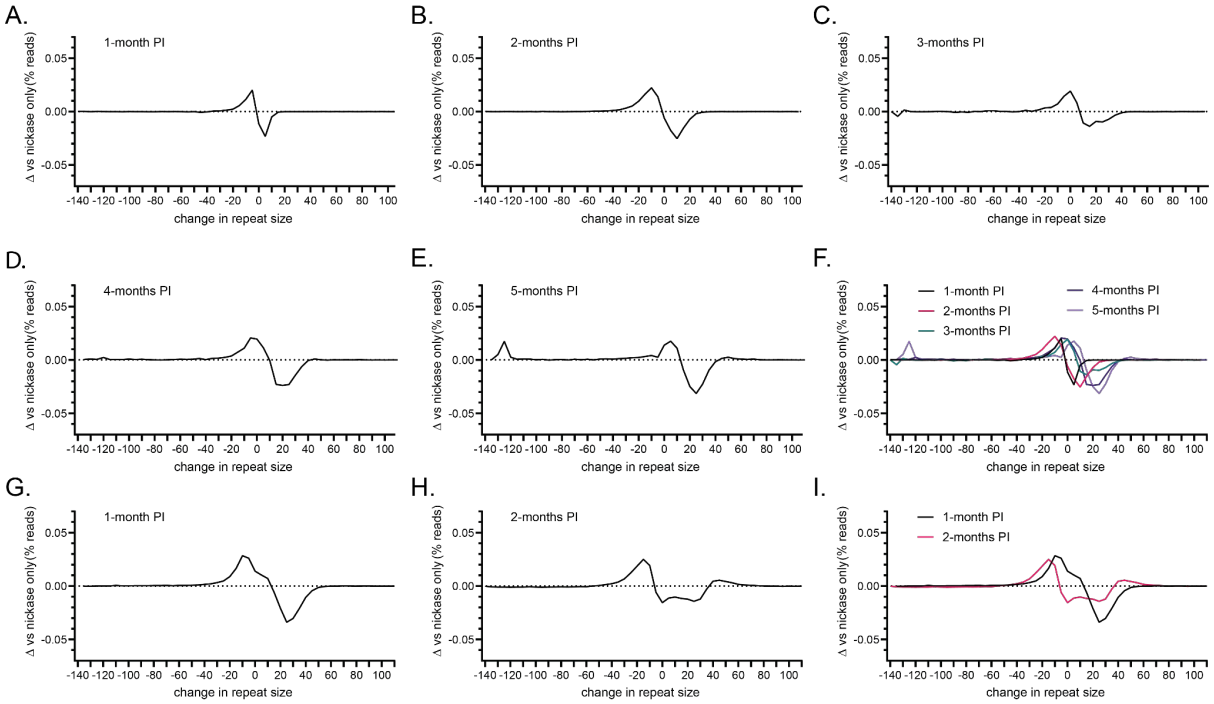
Supplementary Figure 5: GFP expression in P2-injected mice. A) Coronal sections through the mouse brain showing GFP expression in the injected hemisphere 6 weeks post-injection (p.i.) throughout the striatum (dashed lines) and into much of the cortex. B) GFP and actin expression 2 and 3 months p.i. in striatum (S) and cerebellum (C). C) Representative example of GFP colocalization with NeuN+ neurons in the striatum. D) GFP expression in the cortex did not colocalize with NeuN, but rather the GFP positive cells had an astrocytic morphology. Scale bar = 25 µm. E) qPCR results for the number of GFP DNA found inCas9D10A-only or Cas9D10A + sgCTG-injected striatum, showing that there were, on average, 26 ± 6 copies of the AAV per cell 1 month after injection n=4 per condition.



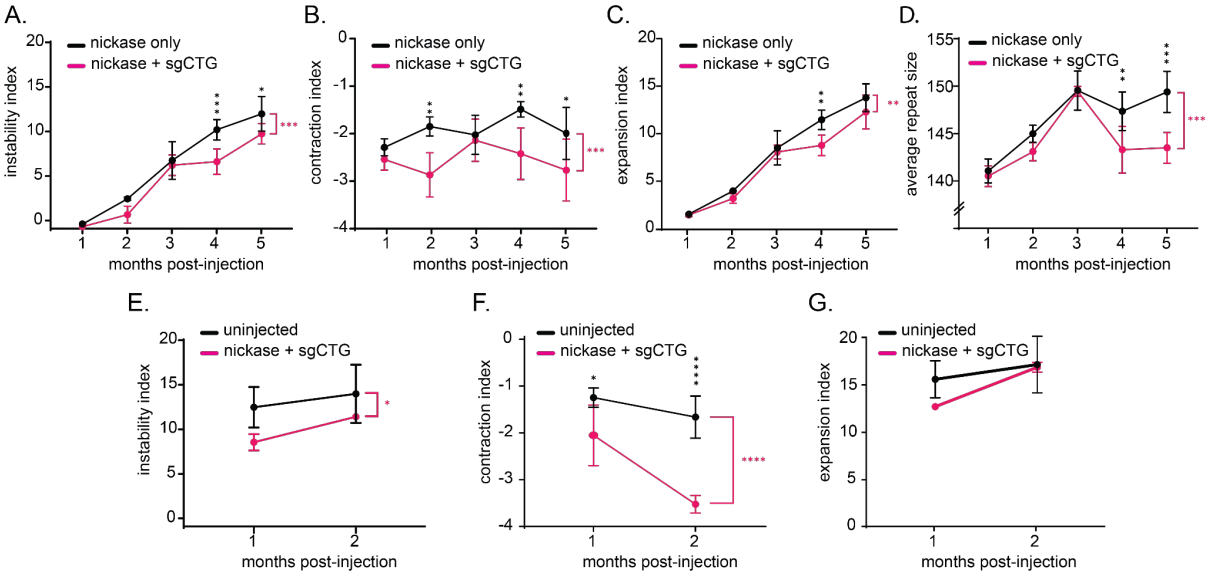
Supplementary Figure 6: Cas9D10A expression *in vivo*. A) Immunofluorescence 1 month post-injection of Cas9D10A expressed from pFBZHmCMV_SpCas9D10A AAV9 injected into wild type striatum. B) Same as A, but using pAAV-nEFCas9D10A. C) Same as A but using pAAV-CMV-SpyCas9D10A. D) Same as A but using pX551-miniCMV-SpCas9D10A. Scale bars = 1000 µm E) Western blot of Cas9D10A expression from the indicated AAVs in the striatum of adult wild type mice 15 days post-injection (p.i.). This gel was spliced for ease of comparison. Full blots are found in

1459 Supplementary Fig. 21. F) Western blot of Cas9D10A expression in the striatum of adult R6/1 mice
1460 injected with the indicated AAVs 1 month post-injection. G) Western blot of Cas9D10A expression in
1461 the striatum of adult R6/1 mice from an AAV9 packaged using pX551-miniCMV-SpCas9D10A, 1 and 2
1462 months after injection. H) Cas9D10A expression in the striatum of P2-injected R6/1 mice using an AAV9
1463 packaged from pX551-miniCMV-SpCas9D10A, 4 and 5 months post-injection. I) qPCR data quantifying
1464 the number of Cas9D10A AAV genomes present compared to the amount of actin. This was done using
1465 striatum DNA of P2-injected R6/1 mice 1 month after injection in mice injected with Cas9D10A AAV
1466 only (black) or both AAVs (pink). Each dot is a different mouse, n=4 per condition. J) Same as (I), but
1467 mice were injected at 3.5 months of age and analysed one month later for the GFP AAV. K) Same as
1468 J, but the data shown is for the number of Cas9D10A AAV DNA copies (Number of animals: uninjected
1469 n=4; Cas9D10A + sgCTG n=3).

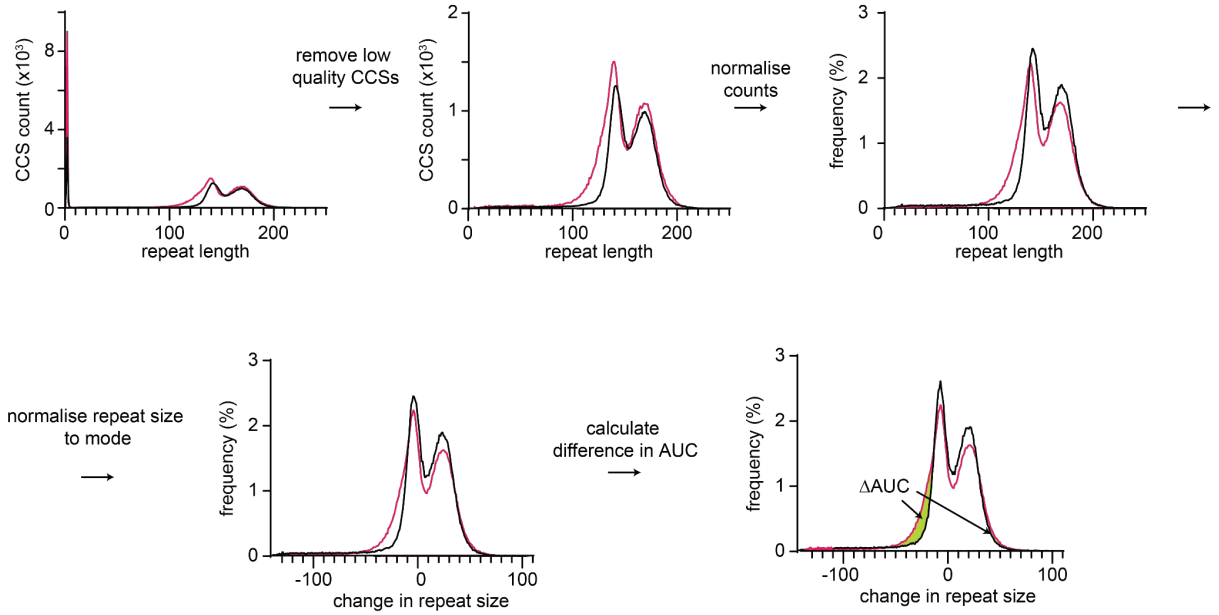
1470



Supplementary Figure 7. Nickase treatment induces contractions that accumulate over time. A) Aggregated delta plot of repeat size distribution in hemispheres of P2 R6/1 mice injected with Cas9D10A + sgCTG subtracting Cas9D10A only repeat size distribution (see methods) at 1 month, B) 2 months, C) 3 months, D) 4 months and E) 5 months. F) Combined time points to show changes over time. Delta-Plot comparing 1M vs 4M and 1vs 5M PI were compared using an unpaired, nonparametric Kolmogorov-Smirnov test. P value <0.0001 (****). G) Aggregated delta plot of repeat size distribution in hemispheres of adult R6/1 mice injected with Cas9D10A + sgCTG subtracting noninjected repeat size distribution (see methods) at 1 month and H) 2 months post-injection. I) Combined time points to show changes over time. Delta plots of 1 month and 2 months post-injection were compared using an unpaired, nonparametric Kolmogorov-Smirnov test . P value <0.002 (**).

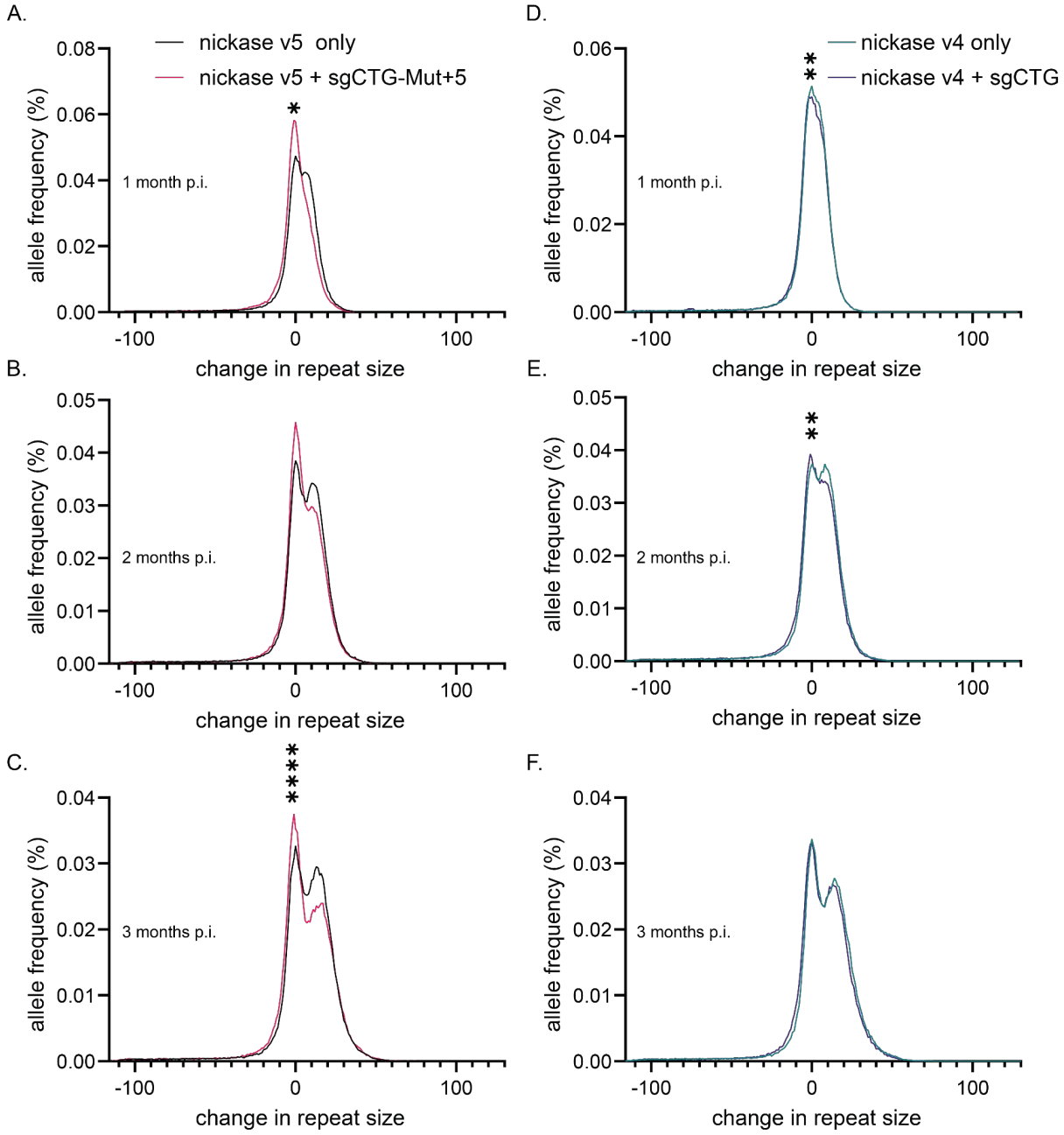


Supplementary Figure 8. CRISPR-CasD10A reduces instability and increases contraction index. A) Instability index of P2 injected mice treated with both Cas9D10A and sgCTG (magenta) versus Cas9D10A only (black). A two-way ANOVA showed an effect of the treatment (P<0.001, magenta asterisks), with a post-hoc Tukey's multiple comparison test showing significant differences 4 and 5 months post-injection (*P<0.05, ***P<0.001, black asterisks). B) The same data as in A, but showing only the contraction index (more negative values indicating more contractions). *P<0.05, **P<0.01, ***P<0.001. C) Same as in F but only showing the expansion index.**P<0.01. D) The average repeat size over time. **P<0.01, ***P<0.001 using a two-way ANOVA. Error bars represent standard deviation. Number of animals: Cas9D10A only; 1 month n=4, 2 months n=3; 3 months n=3; 4 months n=4; 5 months n=3; Cas9D10A + sgCTG; 1 month n=4, 2 months n=3; 3 months n=3; 4 months n=4; 5 months n=3). E) Instability index of the adult mice treated with both Cas9D10A and sgCTG (magenta) versus uninjected (black). A two-way ANOVA showed an effect of the treatment *P<0.05. F) The same data as E, but showing only the contraction index (more negative values indicating more/larger contractions). Magenta asterisks: Two way ANOVA, ****P<0.0001, post-hoc Tukey's multiple comparison test at individual time points (black asterisks) *P<0.05, ****P<0.0001. G) Same as in E but only showing the expansion index. Error bars represent ± standard deviation. Number of animals: uninjected 1 month post-injection n=4, 2 months post-injection n=5; Cas9D10A + sgCTG; 1 month post-injection n=3, 2 months post-injection n=3.

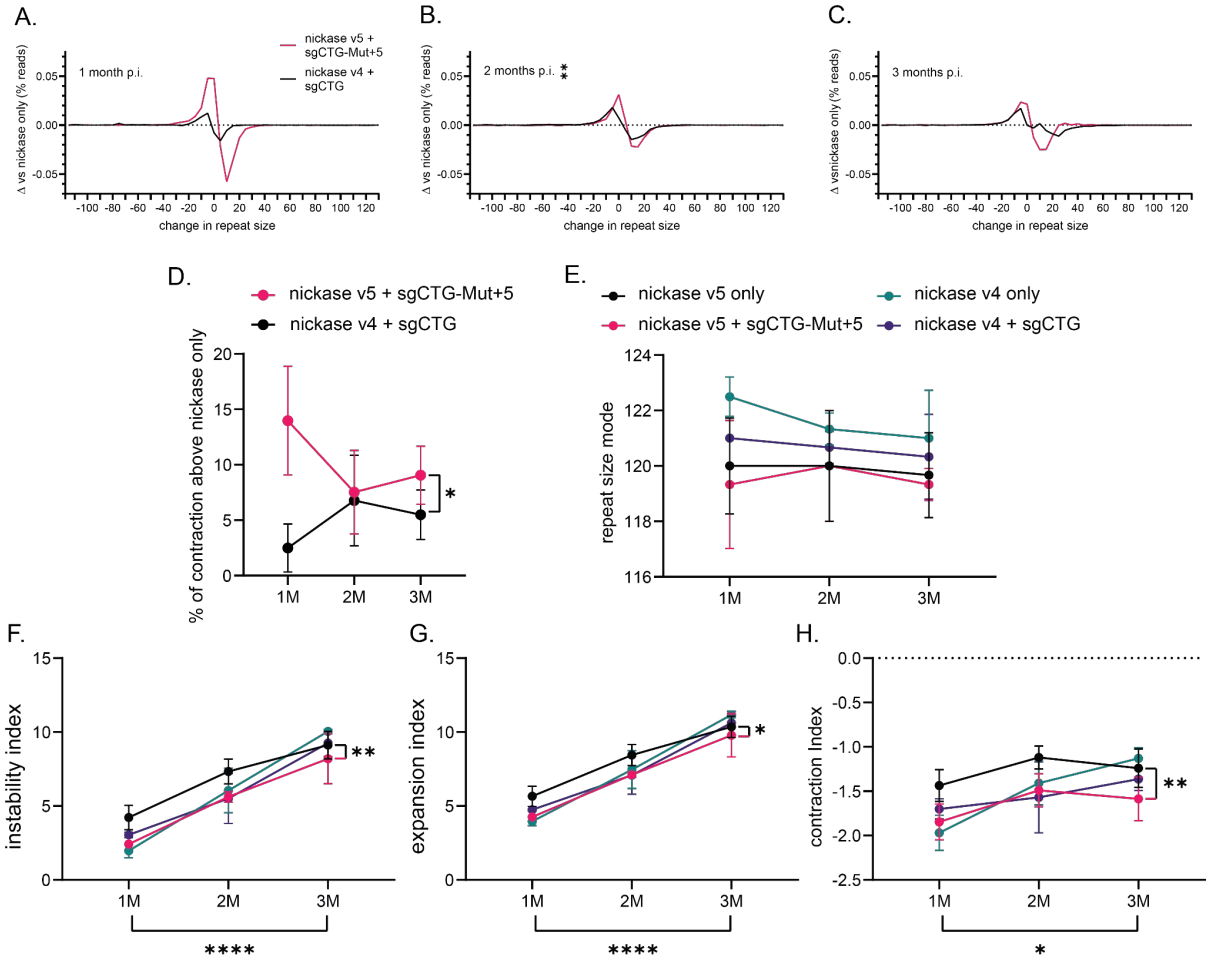


1501

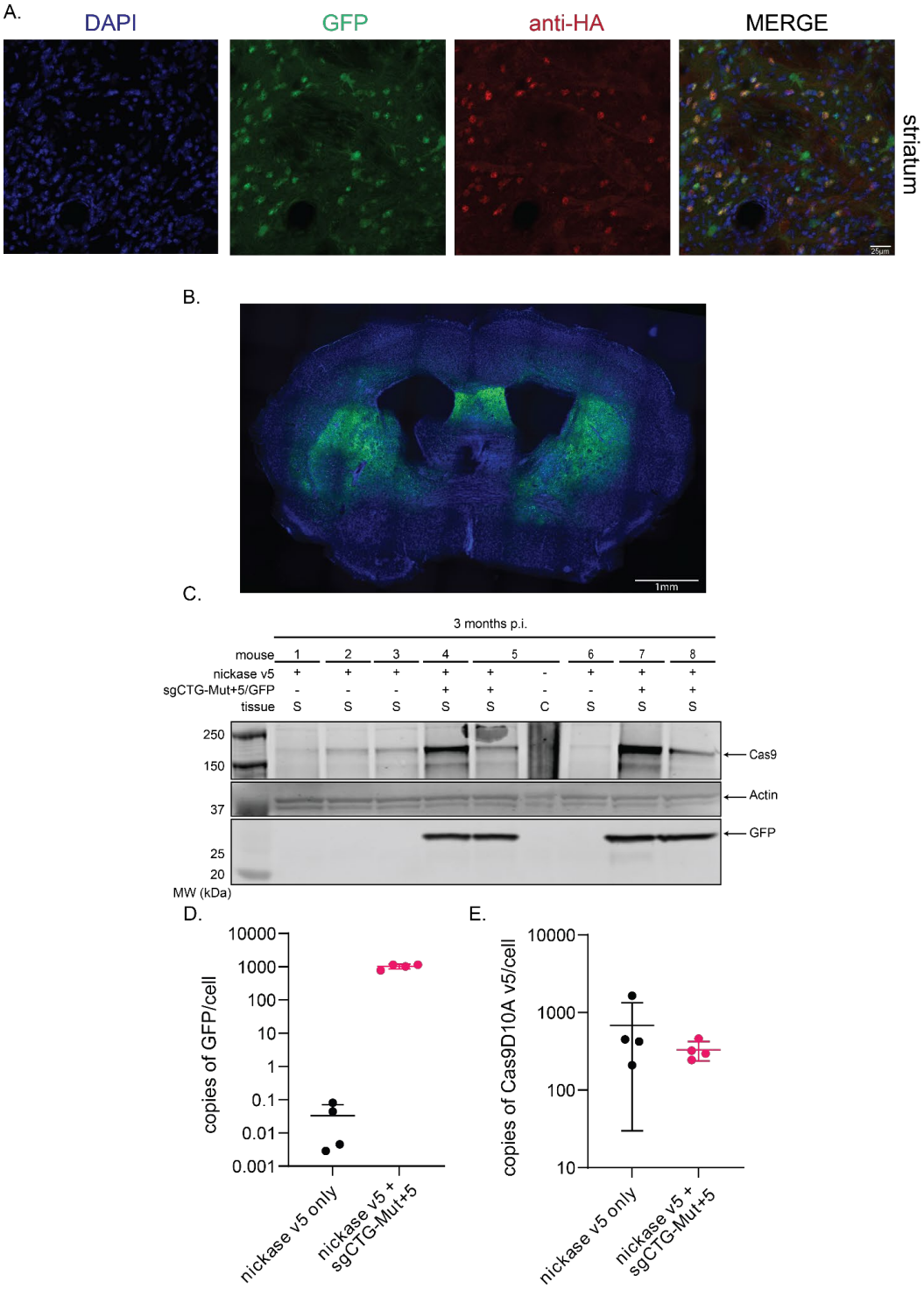
1502 **Supplementary Figure 9:** Explanation of how the difference in the area under the curve was calculated.
1503 This is an example dataset. Repeat Detector is used to determine repeat sizes. Then repeat size
1504 between 0 and 5 CAGs are removed as they are low quality reads that do not align to the *HTT* locus
1505 and truncated PCR products. Then the graphs are normalised from CCS counts and for the modal
1506 repeat size of each sample. This prevents taking small expansions in the untreated samples to inflate
1507 the number of contractions. The difference in the area under the curve is calculated for both shorter and
1508 longer alleles.



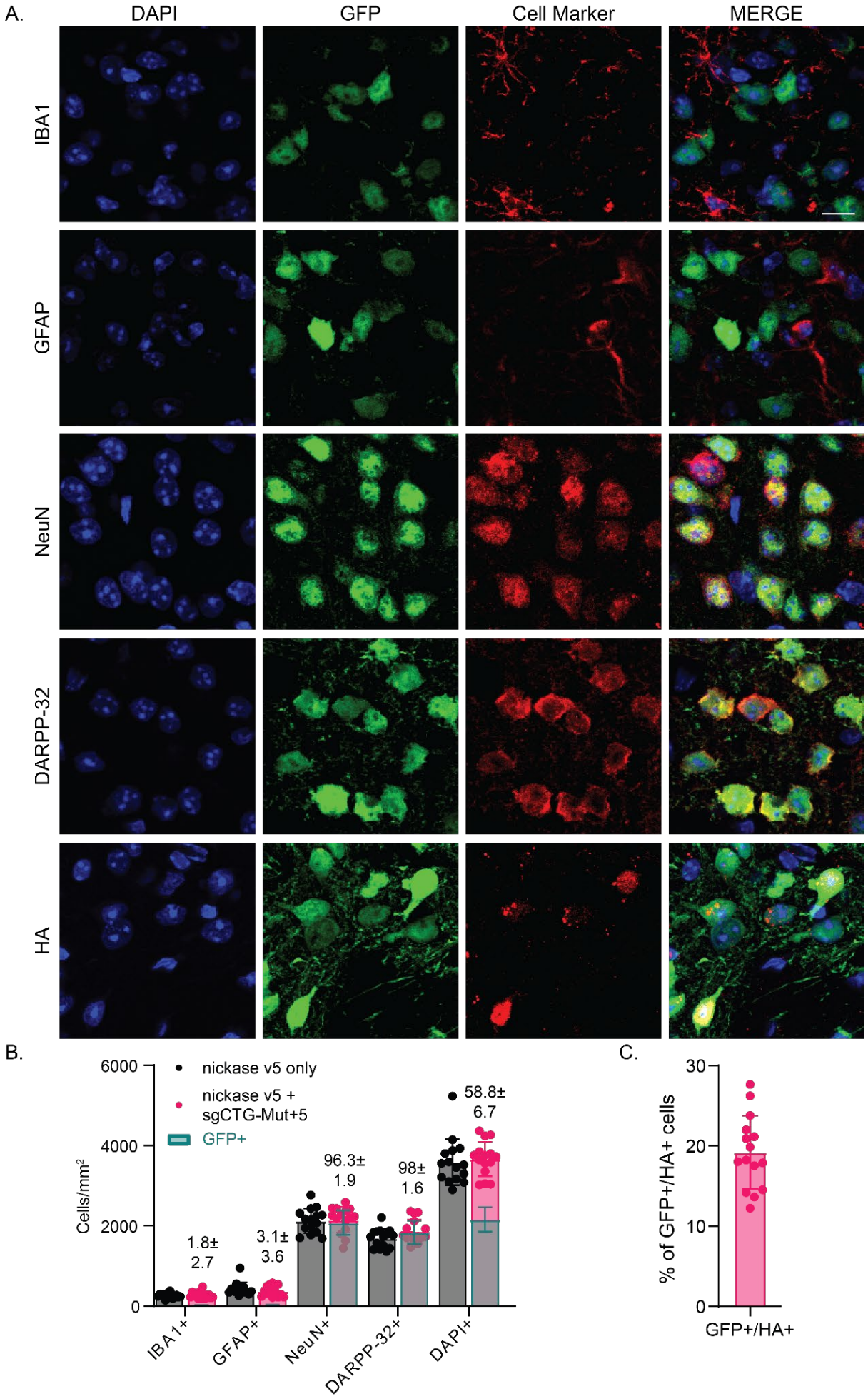
Supplementary Figure 10. Instability of CAG repeats in adult-injected mice over time.A) Aggregated graphs of repeat size distribution in the striatum of adult R6/1 mice injected with the Cas9D10A v5 only (black) compared to striata injected with Cas9D10A v5 + sgCTG-Mut+5 (magenta) 1 month post-injection. B) Same as A but 2 and C) 3 months-post-injection. D) Aggregated graphs of repeat size distribution in the striatum of adult R6/1 mice injected with the Cas9D10A v4 only (green) compared to striata injected with Cas9D10A v4 + sgCTG (blue) 1 month post-injection.E) Same as D but 2 and F) 3 months-post-injection. Note that the graphs are normalized to the modal peak size found in each mouse (see methods). T test, paired nonparametric test Wilcoxon matched-pairs signed rank test to compare cumulative distributions (*P value <0.05; **P value <0.01; ****P value <0.0001). Number of animals: Cas9D10A v5 only; 1 month n=3, 2 months n=3; 3 months n=3; Cas9D10A v5 + sgCTG-Mut+5; 1 month n=3, 2 months n=3; 3 months n=3); Cas9D10A v4 only; 1 month n=2, 2 months n=3; 3 months n=3; Cas9D10Av4 + sgCTG-Mut+5; 1 month n=2, 2 months n=3; 3 months n=3).



Supplementary Figure 11. Optimizing the AAV cargos. 2-month-old R6/1 mice were injected in the striatum with the Cas9D10A v5 + sgCTG-Mut+5 ratio 1:2 or Cas9 v4 + sgCTG ratio 1:1. The other striatum received the respective Cas9D10A AAV only. A) Delta plot analysis revealed significant differences in the contractions induced between treatments at 1, 2 (B) or 3 (C) months post-injection. Unpaired nonparametric Kolmogorov-Smirnov test was used to compare cumulative distributions (**P<0.01) D) Area under the curve comparing the Cas9D10A v5 + sgCTG-Mut+5 treatment with Cas9D10A v5 alone (*P<0.05 using a one-way ANOVA). E) Analysis of the modal CAG repeat size showed no difference between groups by 2-way ANOVA test. F) Instability index of the mice treated versus Cas9 only. A two-way ANOVA showed an effect of the treatment reducing repeats only in the Cas9v5 and sgCTG-Mut+5 treated mice. G) Same as in F, but showing only the expansion index and H) contraction index. D-H) Cohorts show an effect of the timing as expected by the two-way ANOVA analysis (P<0.0001) . Two way ANOVA, *P<0.05, **P<0.0, ****P<0.0001, post-hoc Tukey's multiple comparison test. Error bars represent ± standard deviation. Number of animals: indicated in Supplementary Fig. 10.

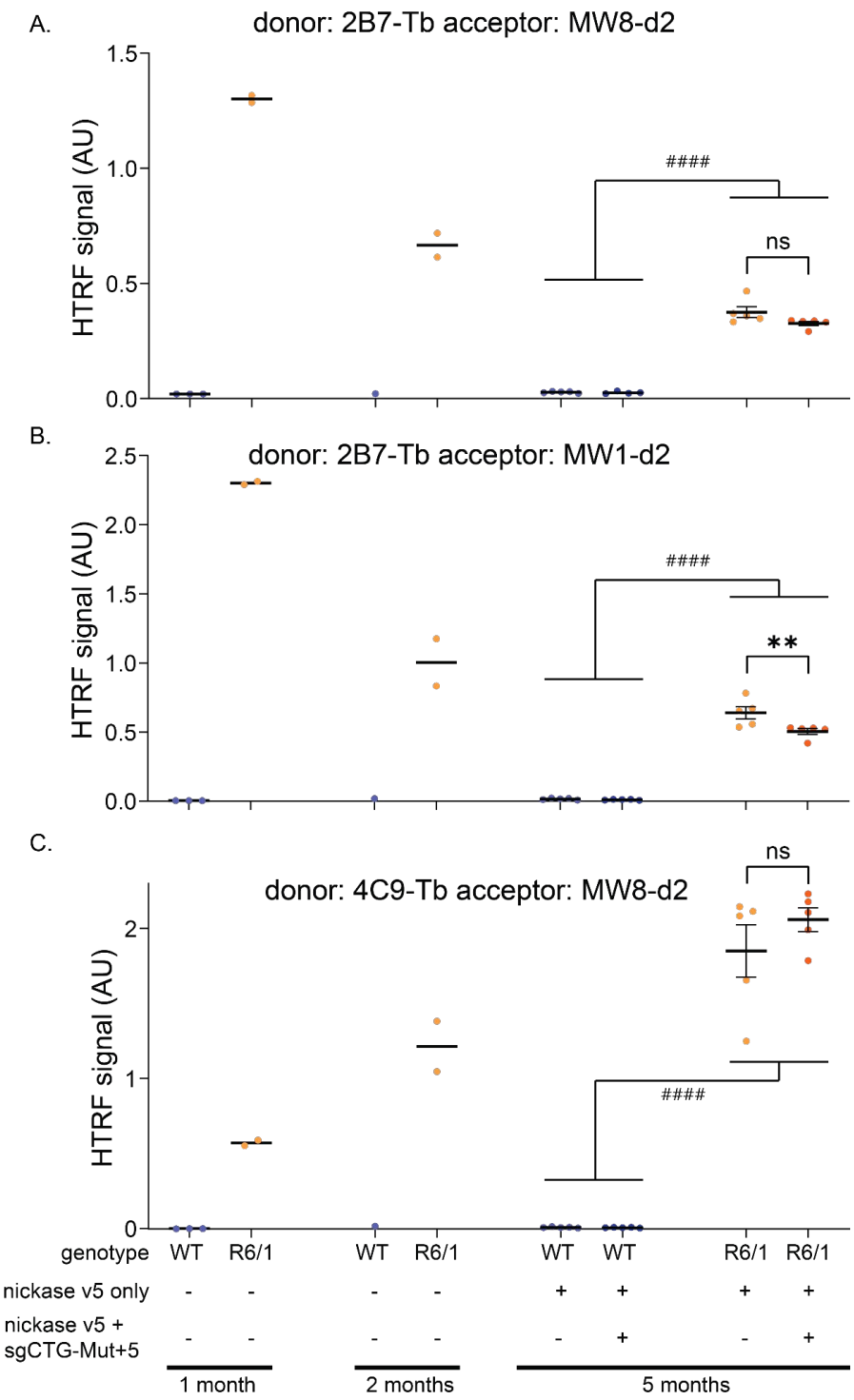


Supplementary Figure 12. Cas9 nickase optimized and scAAV-sgCTG-Mut+5 express in mouse brain. A) Representative image of 2-months-old wild type injected mouse sacrificed 15 days PI. Immunofluorescence against Cas9OPTI with the anti-HA antibody reveal co-expression with GFP express by the AAV genome. B) 3-months post-injected brain analysis showed that a single injection of AAV9 resulted in widespread and efficient transduction of GFP detected by IF with anti-GFP antibody C) Western blot shows high levels of Cas9D10A expression as well as GFP 3-month PI. Full blots are found in Supplementary Fig. 21. D) DNA extracted from striatal samples were used to determine the number of Cas9OPTi (D) and GFP (E) AAV genome copies by qPCR in control and treated mouse samples. Data are mean ± SD (n = 4 animals per condition).

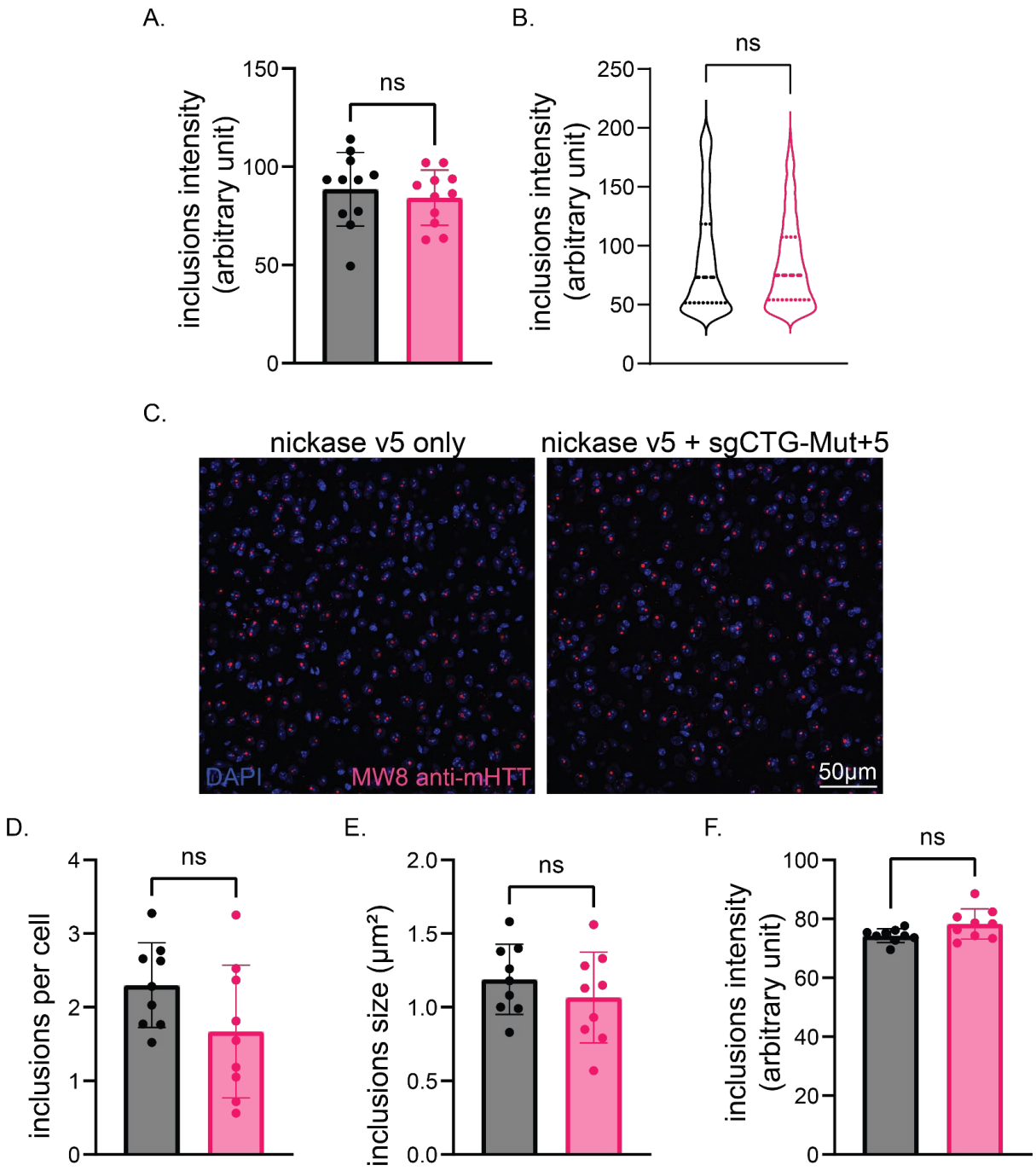

A.
DAPI
GFP
Cell Marker
MERGE
IBA1
GFAP
NeuN
DARPP-32
HA
B.
nickase v5 only
nickase v5 + sgCTG-Mut+5
GFP+
Cells/mm²
6000
4000
2000
0
1.8± 2.7
3.1± 3.6
96.3± 1.9
98± 1.6
58.8± 6.7
IBA1+
GFAP+
NeuN+
DARPP-32+
DAPI+
C.
% of GFP+/HA+ cells
30
20
10
0
GFP+/HA+

1549 **Supplementary Figure 13.** Infection rates of sgCTG-Mut+5/GFP and Cas9D10A v5 AAVs injected in
1550 the mouse striatum . A) Representative images of 5-months-old R6/1+ injected with the nickase v5 and
1551 sgCTG-Mut+5/GFP mouse sacrificed 3 months post-injection. Immunofluorescence with the anti-GFP
1552 antibody reveals co-expression with other cell type markers: Iba1 (microglia), GFAP (astrocytes), NeuN
1553 (neurons), DARPP-32 (medium spiny neurons). Cas9D10A was detected using an anti-HA antibody.
1554 Scale bar: 10µm. B) Quantification of striatal cells positive for a given marker in R6/1 mice injected with
1555 the Cas9D10A only compared with those injected with the Cas9D10A and sgCTG-Mut+5 show no
1556 significant differences in the total cell types or in total cell per mm^2 (DAPI+cells). Quantification of
1557 transduced cells with the sgCTG-Mut+5/GFP (green bars) shows a tropism for neurons of the AAV9.
1558 Percentages indicate the proportion of GFP+ for each cell marker. C) Quantification of striatal cells
1559 positive for GFP that co-express the nickase detected with the anti-HA antibody (19.2±4.6). Data are
1560 mean ± SD (n = 3 animals per group, average of 298 cells per image from 5 striatal fields at different
1561 antero-posterior and dorso-ventral levels were assessed per mouse).

1562

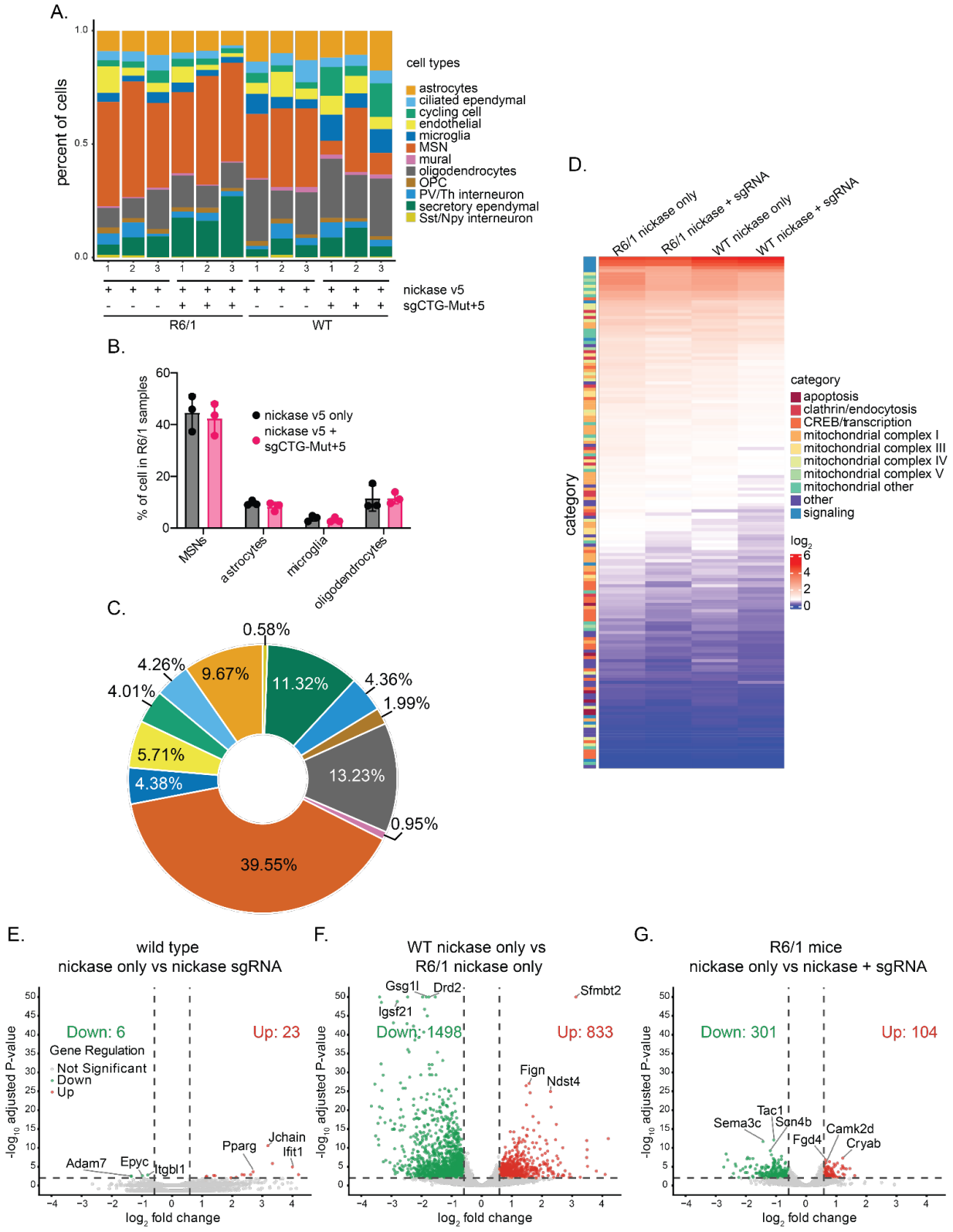


Supplementary Figure 14. mHTT expression analysis by HTRF assay over the time. Cortical tissues from 1 and 2-months-old control wild type (WT) and R6/1 were used to validate the assay in the R6/1 mouse HD model. The expression of mHTT was compared with striatal tissue samples from experimental animals in Fig. 3b-d. A) 4C9-Tb with MW8-d2 was used to track aggregated mutant HTT, B) 2B7-Tb with MW8d2 were used to track changes in soluble mutant HTT and C) 2B7-Tb with MW1-d2 were used for polyglutamine size assay by HTRF in striatal lysates from R6/1 mice at 1, 2 and 5 months of age. Data are mean ± SD (n 1–5 animals per group). 2-way ANOVA test was run in 5-months old animals (see Supplementary Table 4) with a post-hoc Tukey's multiple comparison test showing significant differences between WT and R6/1 animals (####P<0.0001, black hash) and between R6/1 injected with Cas9D10A only versus dual AAVs injection (**P<0.01, black asterisks).



Supplementary Figure 15. Cas9 nickase brain injection improves mHTT aggregation pathology in striatal areas but not in uninjected cortical areas. A) Average intensity of inclusions between groups and B) intensity distribution shows no differences between animal groups, suggesting that the remaining foci, although smaller, remain inclusions (4009 and 2303 inclusions assessed in total in Cas9 only and Cas9 + sgCTG injected animals respectively). C) Representative z-projection images of layer 2-3 cortical areas from the same preparations as in Fig. 3e of 5-months-old R6/1 animal brains sacrificed 3 months post-injection of the Cas9D10A only or Cas9D10A and sgCTG and stained with DAPI (blue) and mutant polyglutamine inclusions (red). mHTT analysis in cortical areas revealed no significant differences between R6/1 animal groups in number of inclusions per cell (D), inclusions sizes (E) or

1585 intensity (F). Data are mean ± SD from a Student t-test. (n = 3-4 images per mouse, ~900-1600 cells
1586 assessed total per mouse, with n = 3 mice per condition).

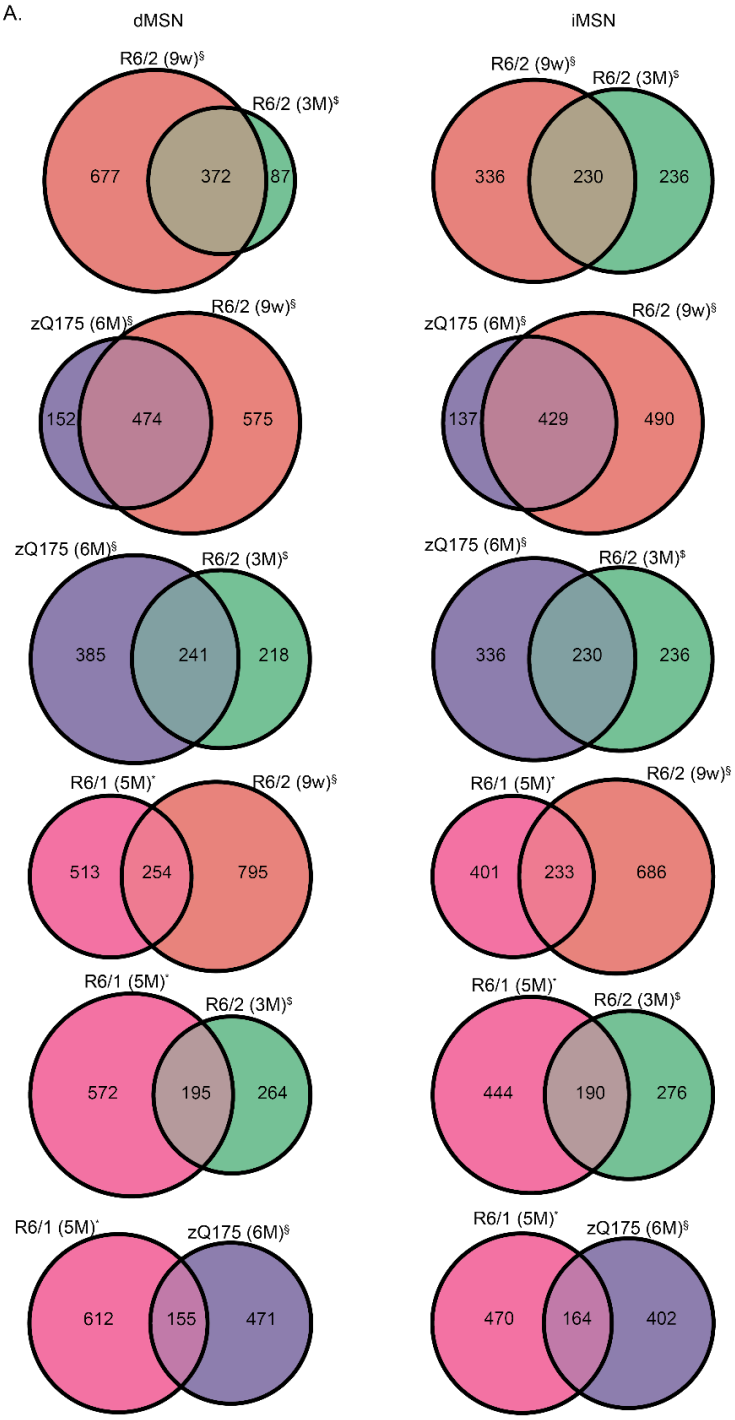


Supplementary Figure 16. CRISPR injection mitigates disease-associated transcription pathology in HD medium spiny neurons and pseudo-bulk differential expression analysis. A) Cell type frequency distribution found in striatal tissue per sample condition. B) Cell types distribution in R6/1 mice injected with the Cas9D10A only compared with those injected with the Cas9D10A and sgCTG-Mut+5. C) Cell type proportion distribution for all samples combined. D) Heatmap of normalized gene expression showing a select subset of DEGs in HD-relevant pathways. The DEGs (rows) are colour-coded on the

1594 right by the direction of gene expression in the specified region (right columns). Horizontal bars in the
1595 left column are colour-coded for pathway gene related categories in HD. E) Volcano plot for
1596 differentially expressed genes (DEGs) in medium spiny neurons (MSNs) with a pseudo-bulk differential
1597 expression analysis between wild type animals injected with Cas9D10A v5 only versus Cas9D10A v5
1598 + sgCTG-Mut+5, F) WT versus R6/1 animals injected with Cas9D10A v5 only and G) CRISPR-treated
1599 R6/1 MSNs relative to Cas9D10A v5-injected R6/1. Number of animals: WT Cas9D10A v5 only n=3,
1600 WT Cas9D10A v5 + sgCTG-Mut+5 n=3, R6/1 Cas9D10A v5 only n=3, R6/1 Cas9D10A v5 + sgCTG-
1601 Mut+5 n=3.

1602

A.

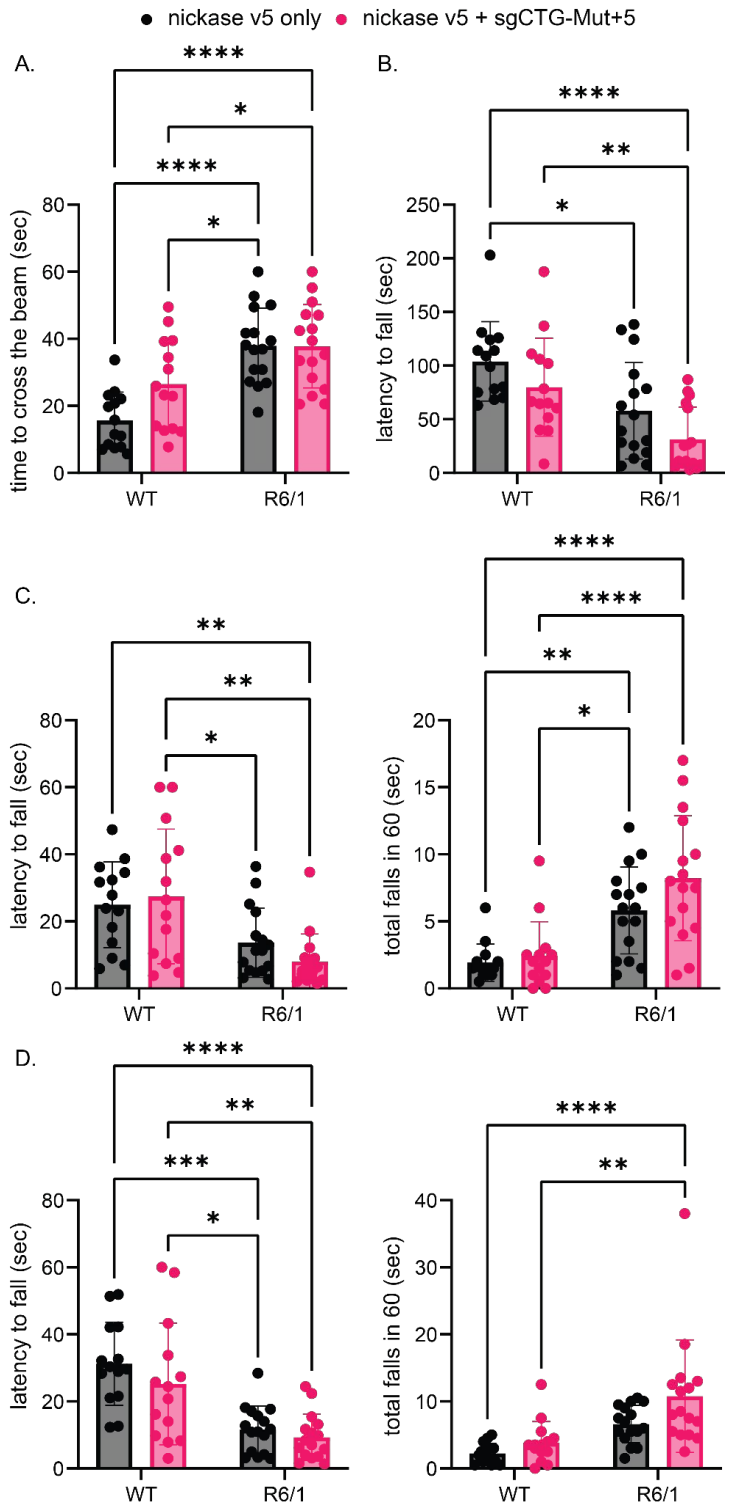


B.

Comparison	Pathway	Number of DEGs 1st transgenic mice vs WT	Number of DEGs 2nd transgenic mice vs WT	Number of overlapping genes	Fisher test	Percentage in same direction
R6/2 (9w)§ vs R6/2 (3M)§	Direct	1049	459	372	<0.0001	99.7
R6/2 (9w)§ vs R6/2 (3M)§	Indirect	566	466	230	0.0918	100
zQ175 (6M)§ vs R6/2 (9w)§	Direct	626	1049	474	<0.0001	100
zQ175 (6M)§ vs R6/2 (9w)§	Indirect	566	919	429	<0.0001	100
zQ175 (6M)§ vs R6/2 (3M)§	Direct	626	459	241	0.0050	99.2
zQ175 (6M)§ vs R6/2 (3M)§	Indirect	566	466	230	0.0918	97.8
R6/1 (5M)* vs R6/2 (9w)§	Direct	767	1049	254	0.0020	98.8
R6/1 (5M)* vs R6/2 (9w)§	Indirect	634	919	233	0.0005	99.1
R6/1 (5M)* vs R6/2 (3M)§	Direct	767	459	195	<0.0001	100
R6/1 (5M)* vs R6/2 (3M)§	Indirect	634	466	190	0.0101	100
R6/1 (5M)* vs zQ175 (6M)§	Direct	767	626	155	0.1151	100
R6/1 (5M)* vs zQ175 (6M)§	Indirect	634	566	164	0.3827	99.4

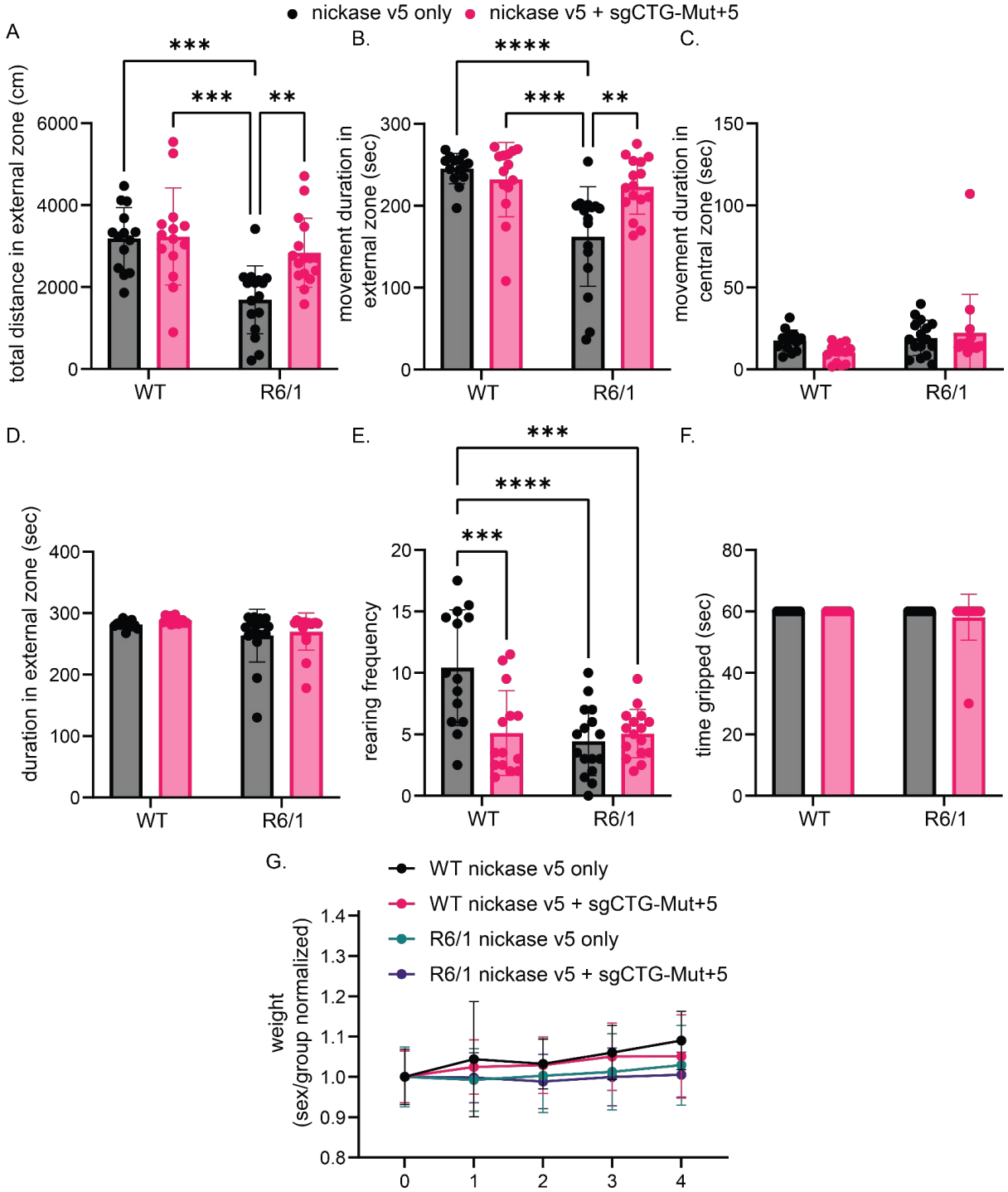
Supplementary Figure 17. Differential expression analysis in medium spiny neurons (MSNs) from R6/1 mice injected with the Cas9D10A only show similar alterations as other HD animal models. A) Venn diagrams with the number of genes altered compared to wild type animals. Diagrams show the

comparison of D1+ MSNs of the direct pathway (dMSNs) and D2+ MSNs of the indirect pathway (iMSNs) of the indicated mouse models and ages from: *: animals from this study; §: animals from Lee et al[58] . $: animals from Lim et al[92] . B) Table summarizing differentially expressed genes (DEGs) in MSNs from previous Venn diagrams. Fisher's exact test was run using total DEGs and overlapping genes to assess similarities between animal models.

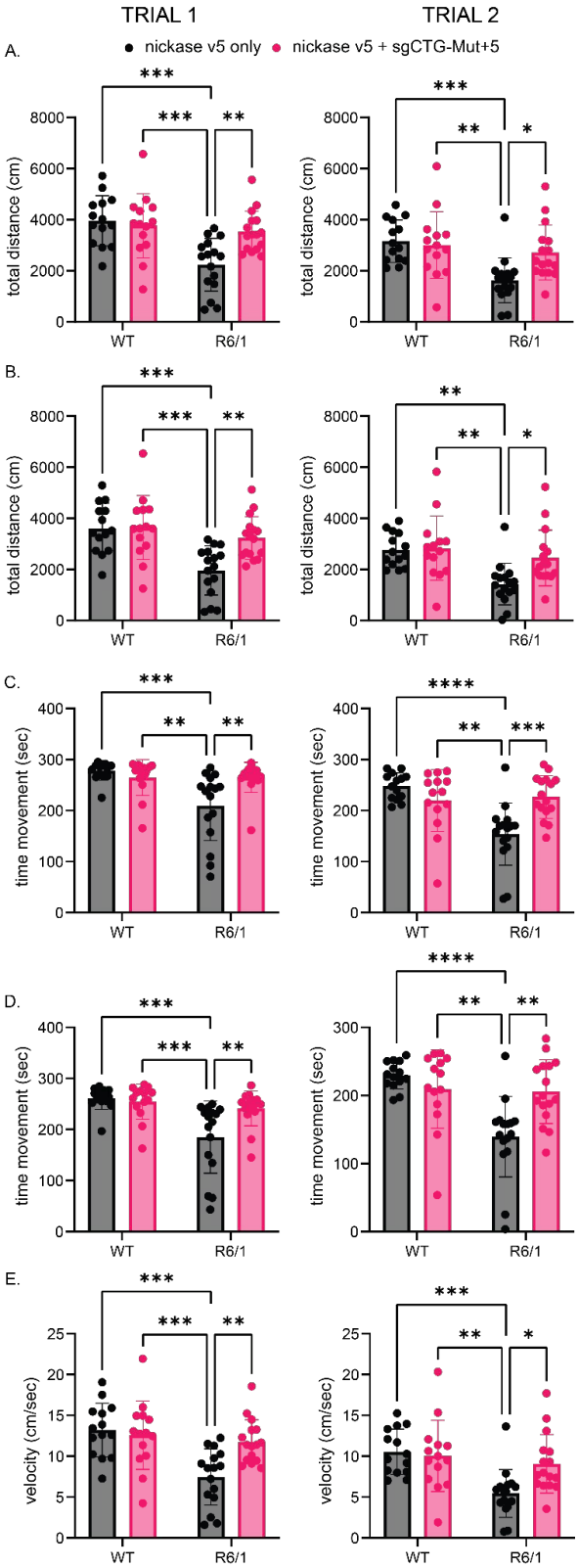


Supplementary Figure 18. Treated R6/1 animals do not improve in motor coordination tests. A) Average of the timing needed for the animals to cross the beam was measured and no differences induced by the treatment were found. B) We measured the first fall timing in the accelerated rotarod and no improvement after the treatment was found in R6/1. C) First fall timing (left) and total number of falls (right) from a fixed rotarod set at 12 rpm for 3-month-old mice of the indicated genotypes that received bilateral intrastriatal injections of Cas9D10A only and the Cas9D10A and the sgCTG-Mut+5

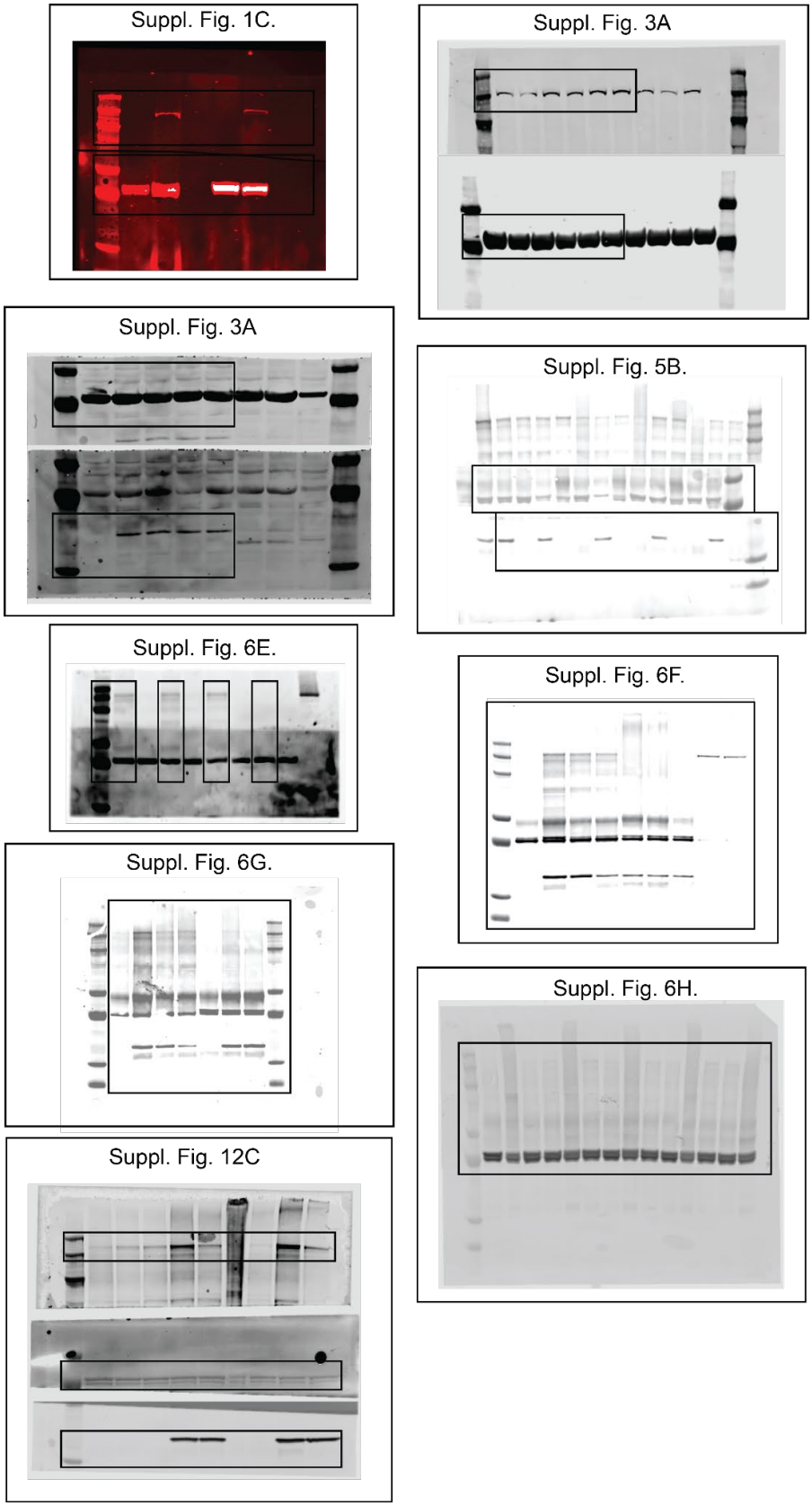
1620 at 2 months of age. D) Same as C) with the rotarod set at 24rpm. All data are an average of two trials.
1621 N of animals=14-16 per group. Statistical analyses are in Supplementary Table 8.



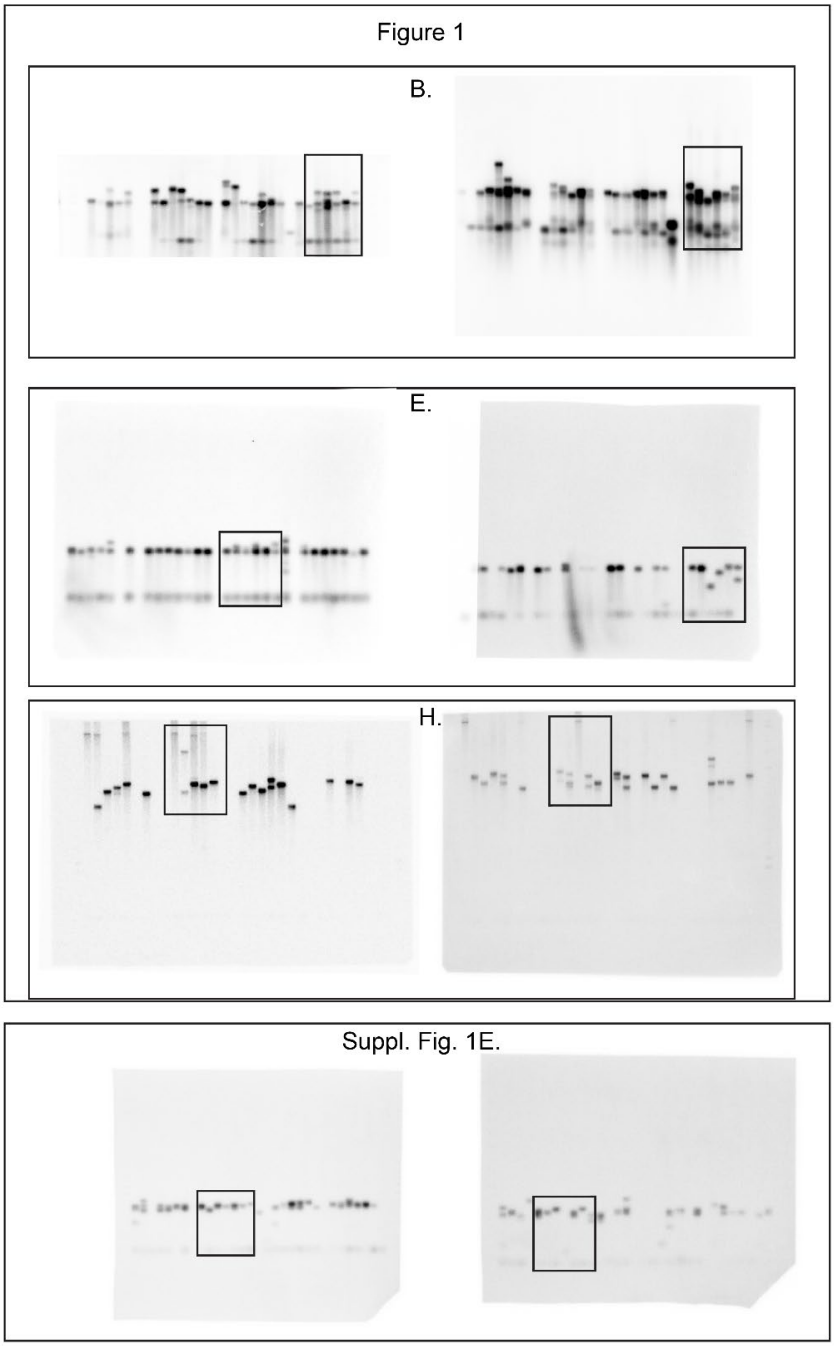
Supplementary Figure 19. Injection with the CRISPR system does not alter or rescue exploratory and anxiety parameters. A) Average of total distance traveled in the external part of the arena. B) Average time in movement on the external part of the arena. C) Average of the total movement time in the central part of the arena. D) Total movement and stop time in the external part of the arena. E) Exploratory behavioural measured by the rearing frequency shows no differences between R6/1 animal groups but it is reduced in wild type (WT) animals. F) Animal muscular strengths are not affected by the injections. G) Weight of the animals were measured before the injection and every week after. Normalization per gender and groups shows no differences caused by the treatments. N of animals=14-16 per group. Statistical analyses are in Supplementary Table 9.



Supplementary Figure 20. Locomotor activity is stable between Open Field trials. Same data as in Fig. 5c-e and Supplementary Fig. 19ab showing differences between 2 performed open field trials; first trial left column and second trial right column. A) Total distance moved in the 5 min test. B) Same as A but total distance in the external part of the arena. C) Same as A but with the total movement time. D) Same as C but total movement time in the external part of the arena. E) Mean velocity in the total arena. N of animals=14-16 per group. Statistical analyses are in Supplementary Table 10.



Supplementary Figure 21: Unaltered full western blots membranes. Black boxes indicate where the blots were cropped.



Supplementary Figure 22: Unaltered full small-pool PCR membranes from Fig. 1beh and Supplementary Fig. 1e. Black boxes indicate where the membranes were cropped.

1645 **Supplementary Table 1**: Plasmids used herein.

Name	Description	Addgene number	Reference
pcDNA3.3-TOPO - Cas9D10A nickase	Cas9_D10A expression plasmid. Also harbors a G418 resistance gene. Used as a positive control for Cas9D10A detection in HEK293 cells using western blotting.	41816	[93]
pPN10 – sgCTG	Expresses the sgRNA against $(CUG)_6C$ from a U6 promoter. Also contains a Puromycin resistance gene.	114385	[30]
pLenti-EF1alpha-Cas9D10A nickase-Blast	Expresses Cas9D10A together with a blasticidin resistance gene. Suitable for lentivirus packaging. Used with the HD lymphoblastoid cell line, iPSC-derived astrocytes and iPSC-derived neurons.	63593	[94]
pLV(gRNA)-CMV-eGFP-U6(sgCTG)	Contains a eGFP gene along with a U6 promoter driving the sgCTG (target sequence: $(CUG)_6C$). Used with the HD iPSC-derived astrocytes.	216732	This study
pLV(gRNA)-CMV-eGFP:T2A:Hygro-U6(sgCTG)	Contains a eGFP gene with a T2A sequence for hygromycin resistant gene along with a U6 promoter driving the sgCTG (target sequence: $(CUG)_6C$). Used with the HD lymphoblastoid cell line.	-	This study
pLenti-U6-sgCTG-CAG-Cas9D10A nickase-Blast	Expresses Cas9D10A under a CAG promoter together with a blasticidin resistance gene, along with a U6 promoter driving the sgCTG (target sequence: $(CUG)_6C$). Suitable for lentivirus packaging. Used in the iPSC-derived neurons.	216733	This study
pAAV-U6-sgCTG-CMV-GFP	AAV plasmid containing sgCTG (sgCTG (target sequence: $(CUG)_6C$) driven by a U6 promoter and eGFP under the control of the CMV promoter. Used for the *in vivo* experiments.	216734	This study
pFBZHmCMV_SpCas9D10A	AAV plasmid containing spCas9. Plasmid contains a CMV promoter, and	-	This study

	SpCas9 was cloned upstream of a minimal poly A sequence.		
pX551-miniCMV-SpCas9D10A nickase	Expresses SpCas9-D10A nickase from a miniCMV promoter. For AAV packaging. Modified from pX551-miniCMV-SpCas9, which was a gift from Alex Hewitt (Addgene #107031)	216735	This study
pAAV-nEF-SpCas9D10A nickase	Expresses Cas9-D10A nickase from a nEF promoter. For AAV packaging. Derived from pAAV-nEF-SpCas9 (Addgene #87115)[48].	216736	This study
pAAV-CMV-SpCas9D10A nickase	Expresses Cas9-D10A nickase from a CMVd1 promoter. For AAV packaging. Derived from pAAV-CMV-SpCas9 (Addgene #113034)[49]	216737	This study
scAAV-U6-sgCTG-Mut+5-CMV-GFP	AAV plasmid containing sgCTG (sgCTG (target sequence: $(CUG)_6$) with Mut+5 optimization driven by a U6 promoter and eGFP under the control of the CMV promoter. Used for the *in vivo* experiments	-	This study
pAAV_miniCMV_HA_Cas9D10A_Muscle codon Optimized	Expresses SpCas9-D10A nickase muscle optimization with 3xHA tag sequences under a miniCMV promoter. For AAV packaging. Modified from 216735	-	This study

1646

1647

Supplementary Table 2: Whole Genome Sequencing of off-target mutations found in HD iPSc-derived cells and in genes containing a potential off-target site. See Excel sheets attached.

1653 **Supplementary Table 3.** CRISPR system AAV construct versions used in this study

Name	Other names	Description	Plasmid in table 3	Used in figures:	Addgene number	Reference
Cas9D 10A v1	Nickase v1	No optimization of spCas indicated	pFBZHmCMV_SpCas9D10A	Supplementary Fig. 6a	-	This study
Cas9D 10A v2	Nickase v2	No optimization of spCas indicated	pX551-miniCMV-SpCas9D10A nickase	Suppl Fig6A-F	216735	This study
Cas9D 10A v3	Nickase v3	Optimization of nuclear transport of spCas9.	pAAV-nEF-SpCas9D10A nickase	Suppl Fig6A-E	216736	This study
Cas9D 10A v4	Nickase v4	human codon–optimized spCas9 from F Ann Ran et al., 2013	pAAV-CMV-SpCas9D10A nickase	Fig 2;Suppl Fig5-11	216737	This study
Cas9D 10A v5	Nickase v5	spCas9D10A mouse muscle codon optimization	pAAV_miniCMV_HA_Cas9D10A_Muscle codon Optimized	Fig 3-5;Suppl Fig10-19	-	This study
sgCTG-1	Guide-1	Express sgCTG (sgCTG (target sequence: $(CUG)_6C$) with an original tracrRNA sequence	pAAV-U6-sgCTG-CMV-GFP	Fig 2;Suppl Fig5-11	216734	This study
sgCTG-2	Guide-2	sgCTG-1 adapted for higher efficiency following Ying Dang et al., 2015 modification of the ”Mut+5”. (target sequence: (CUG)6)	scAAV-U6-sgCTG-Mut+5-CMV-GFP	Fig 3-5;Suppl Fig10-19	-	This study

1654

1655 **<u>Supplementary Table 4:</u>** statistics of HTRF tests Supplementary Figure 14

HTRF soluble assay 2B7-tb with MW8-d2 **2-way-ANOVA post-hoc Tukey's multiple comparison test**	
Genotype	F (1, 15) = 572.8, P<0.0001
Treatment	F (1, 15) = 3.375, P=0.0861
Genotype X Treatment	F (1, 15) = 2.917, P=0.1083
HTRF polyglutamine assay 2B7-tb with MW1-d2; 2-way-ANOVA	
Genotype	F (1, 16) = 522.9, P<0.0001
Treatment	F (1, 16) = 522.9, P<0.0001
Genotype X Treatment	F (1, 16) = 522.9, P<0.0001
HTRF aggregate assay 4C9-tb with MW8-d2; 2-way-ANOVA	
Genotype	F (1, 16) = 413.3, P<0.0001
Treatment	F (1, 16) = 1.176, P=0.2942
Genotype X Treatment	F (1, 16) = 1.204, P=0.2888

1656

1657

Supplementary Table 5: snRNA-seq extended data. Including cell cluster module score, single cell DE output & pathways, pathway files, heatmaps selected genes, cell maker lists, RStudio package versions, and the list of potential off-target genes and differentially expressed genes in those potential off-targets. See Excel sheets attached.

Supplementary Table 6: snRNA-seq extended data of pseudo-bulk DE output. Including pseudo-bulk DE output of all cell clusters.

1665 **<u>Supplementary Table 7:</u>** statistics of behavioral tests figure 5

Open Field-Total distance moved **2-way-ANOVA post-hoc Tukey's multiple comparison test**	
Genotype	F (1, 56) = 15.03, P=0.0003
Treatment	F (1, 56) = 4.351, P=0.0416
Genotype X Treatment	F (1, 56) = 7.791, P=0.0072
Open Field-Total time in movement; 2-way-ANOVA	
Genotype	F (1, 56) = 12.07, P=0.0010
Treatment	F (1, 56) = 3.792, P=0.0565
Genotype X Treatment	F (1, 56) = 14.37, P=0.0004
Open Field-Velocity; 2-way-ANOVA	
Genotype	F (1, 56) = 15.02, P=0.0003
Treatment	F (1, 56) = 4.352, P=0.0415
Genotype X Treatment	F (1, 56) = 7.781, P=0.0072

1666

1667

Supplementary Table 8: statistics of behavioral tests Supplementary Fig. 18

Balance beam; 2-way-ANOVA post-hoc Tukey's multiple comparison test	
Genotype	F (1, 56) = 31.34, P<0.0001
Treatment	F (1, 56) = 3.198, P=0.0791
Genotype X Treatment	F (1, 56) = 3.375, P=0.0715
Accelerated rotarod-latency to fall; 2-way-ANOVA	
Genotype	F (1, 56) = 21.05, P<0.0001
Treatment	F (1, 56) = 6.043, P=0.0171
Genotype X Treatment	F (1, 56) = 0.01428, P=0.9053
Fixed rotarod-12rpm latency to fall; 2-way-ANOVA	
Genotype	F (1, 56) = 19.82, P<0.0001
Treatment	F (1, 56) = 0.2038, P=0.6534
Genotype X Treatment	F (1, 56) = 1.351, P=0.2501
Fixed rotarod-12rpm total falls in 60 seconds; 2-way-ANOVA	
Genotype	F (1, 56) = 33.01, P<0.0001
Treatment	F (1, 56) = 2.979, P=0.0899
Genotype X Treatment	F (1, 56) = 1.282, P=0.2624
Fixed rotarod-24rpm latency to fall; 2-way-ANOVA	
Genotype	F (1, 56) = 34.25, P<0.0001
Treatment	F (1, 56) = 1.891, P=0.1746
Genotype X Treatment	F (1, 56) = 0.3699, P=0.5455
Fixed rotarod-24rpm total falls in 60 seconds; 2-way-ANOVA	
Genotype	F (1, 56) = 20.00, P<0.0001
Treatment	F (1, 56) = 5.205, P=0.0264
Genotype X Treatment	F (1, 56) = 1.018, P=0.3173

Supplementary Table 9: statistics of behavioral tests Supplementary Fig. 19

Open Field-Total distance moved external arena **2-way-ANOVA post-hoc Tukey's multiple comparison test**	
Genotype	F (1, 56) = 16.09, P=0.0002
Treatment	F (1, 56) = 6.451, P=0.0139
Genotype X Treatment	F (1, 56) = 5.434, P=0.0234
Open Field-Total time in movement in external arena; 2-way-ANOVA	
Genotype	F (1, 56) = 16.85, P=0.0001
Treatment	F (1, 56) = 4.598, P=0.0364
Genotype X Treatment	F (1, 56) = 11.12, P=0.0015
Open Field-Total distance moved central arena; 2-way-ANOVA	
Genotype	F (1, 56) = 8.612e-005, P=0.9926
Treatment	F (1, 56) = 3.445, P=0.0687
Genotype X Treatment	F (1, 56) = 7.220, P=0.0095
Open Field-Total time external arena; 2-way-ANOVA	
Genotype	F (1, 56) = 6.713, P=0.0122
Treatment	F (1, 56) = 0.8767, P=0.3531
Genotype X Treatment	F (1, 56) = 0.001078, P=0.9739
Open Field-Rearing Frequency; 2-way-ANOVA	
Genotype	F (1, 56) = 12.31, P=0.0009
Treatment	F (1, 56) = 7.451, P=0.0085
Genotype X Treatment	F (1, 56) = 11.94, P=0.0011
Muscular strength; 2-way-ANOVA	
Genotype	F (1, 56) = 0.8711, P=0.3547
Treatment	F (1, 56) = 0.8711, P=0.3547
Genotype X Treatment	F (1, 56) = 0.8711, P=0.3547
Weight; 3-way-ANOVA	
Genotype	F (1, 56) = 0.9942, P=0.3230
Treatment	F (1, 56) = 0.1308, P=0.7189
Timing	F (4, 224) = 9.325, P<0.0001
Genotype X Treatment	F (1, 56) = 0.008928, P=0.9251
Genotype X Timing	F (4, 224) = 3.608, P=0.0071
Treatment X Timing	F (4, 224) = 1.184, P=0.3186
Genotype X Treatment X Timing	F (4, 224) = 0.4213, P=0.7932

Supplementary Table 10: statistics of behavioral tests Supplementary Fig. 20

Open Field-Total distance moved 1st trial 2-way-ANOVA post-hoc Tukey's multiple comparison test	
Genotype	F (1, 56) = 13.70, P=0.0005
Treatment	F (1, 56) = 4.285, P=0.0431
Genotype X Treatment	F (1, 56) = 7.788, P=0.0072
Open Field-Total distance moved 2nd trial; 2-way-ANOVA	
Genotype	F (1, 56) = 11.51, P=0.0013
Treatment	F (1, 56) = 3.052, P=0.0861
Genotype X Treatment	F (1, 56) = 5.368, P=0.0242
Open Field-Total distance moved external arena 1st trial; 2-way-ANOVA	
Genotype	F (1, 56) = 15.53, P=0.0002
Treatment	F (1, 56) = 6.155, P=0.0161
Genotype X Treatment	F (1, 56) = 5.672, P=0.0207
Open Field-Total distance moved external arena 2nd trial; 2-way-ANOVA	
Genotype	F (1, 56) = 11.46, P=0.0013
Treatment	F (1, 56) = 4.660, P=0.0352
Genotype X Treatment	F (1, 56) = 3.510, P=0.0662
Open Field-Total time in movement 1st trial; 2-way-ANOVA	
Genotype	F (1, 56) = 9.515, P=0.0032
Treatment	F (1, 56) = 3.709, P=0.0592
Genotype X Treatment	F (1, 56) = 9.453, P=0.0033
Open Field-Total time in movement 2nd trial; 2-way-ANOVA	
Genotype	F (1, 56) = 11.43, P=0.0013
Treatment	F (1, 56) = 2.975, P=0.0901
Genotype X Treatment	F (1, 56) = 15.52, P=0.0002
Open Field-Total time in movement external arena 1st trial; 2-way-ANOVA	
Genotype	F (1, 56) = 14.60, P=0.0003
Treatment	F (1, 56) = 4.493, P=0.0385
Genotype X Treatment	F (1, 56) = 7.184, P=0.0096
Open Field-Total time in movement external arena 2nd trial; 2-way-ANOVA	
Genotype	F (1, 56) = 13.81, P=0.0005
Treatment	F (1, 56) = 3.337, P=0.0731
Genotype X Treatment	F (1, 56) = 11.62, P=0.0012
Open Field-Velocity 1st trial; 2-way-ANOVA	
Genotype	F (1, 56) = 13.69, P=0.0005
Treatment	F (1, 56) = 4.287, P=0.0430
Genotype X Treatment	F (1, 56) = 7.776, P=0.0072
Open Field-Velocity 2nd trial; 2-way-ANOVA	
Genotype	F (1, 56) = 11.49, P=0.0013
Treatment	F (1, 56) = 3.053, P=0.0861
Genotype X Treatment	F (1, 56) = 5.365, P=0.0242

1677 **Supplementary Table 11:** Antibodies used in this study.

1678

Antibody	Reference	Species	Clonality	Method and concentration
Anti-HA-Tag (C29F4)	Cell Signalling-3724T	Rabbit	Monoclonal	Immunofluorescence (1:500)
Anti- Anti-huntingtin Hybridoma MW8	11719267 11792860	Mouse	Monoclonal	Immunofluorescence (1:1000)
Anti-S100 beta antibody [EP1576Y] - Astrocyte Marker	ab52642	Rabbit	Monoclonal	Immunofluorescence (1:100)
Anti-Beta III Tubulin, Cy3 Conjugate Antibody	AB15708C3	Rabbit	Polyclonal	Immunofluorescence (1:1000)
Anti-CRISPR/Cas9 monoclonal antibody 4G10	C15200216-100	Mouse	Monoclonal	Western blotting (1:1000)
Anti-β-Actin	A5441	Mouse	Monoclonal	Western blotting (1:2000)
Anti-Green Fluorescent Protein	MAB3580	Mouse	Monoclonal	Western blotting (1:1000)
Anti-Green Fluorescent Protein	ZRB1097	Rabbit	Monoclonal	Immunofluorescence (1:1000)
Anti-NeuN clone A60	MAB377	Mouse	Monoclonal	Immunofluorescence (1:500)
Anti-iba1	09-19741	Rabbit	Polyclonal	Immunofluorescence (1:2000)

1679

Anti-GFAP	MAB360	Mouse	Monoclonal	Immunofluorescence (1:1000)
Anti-Green Fluorescent Protein	AB16901	Chicken	Polyclonal	Immunofluorescence (1:1000)
Anti-DARPP-32	2306T	Rabbit	Monoclonal	Immunofluorescence (1:500)
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 488	A11008	Goat	Polyclonal	Immunofluorescence (1:1000)
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 680	A21058	Goat	Polyclonal	Western blotting (1:1000)
Anti-β-Tubulin III antibody clone SDL.3D10	T8660	Mouse	Monoclonal	FACS (1:500)
Anti-human CD44 Antibody, FITC Conjugate Antibody	130-113-896	Mouse	Monoclonal	FACS (1:500)
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 568	A-11031	Goat	Polyclonal	FACS (1:1000)
Anti-HTT aa 1-17 2B7	CHDI Foundation	Mouse	Monoclonal	HTRF (1ng well^{-1})
Anti-polyQ MW1	CHDI Foundation	Mouse	Monoclonal	HTRF (40ng well^{-1})
Anti-HTT aa 51-71 4C9	CHDI Foundation	Mouse	Monoclonal	HTRF (1ng well^{-1})

Anti-HTT end at proline MW8	CHDI Foundation	Mouse	Monoclonal	HTRF (40ng well^{-1})

1680

1681

1682 **Supplementary Table 12:** Primers used herein.

Primer	Sequence (5′ to 3’)	Application	Reference
oVIN-1333	CCGCTCAGGTTCTGCTTTTA	Amplification of the *HTT* locus in astrocytes for small-pool PCR.	[95]
oVIN-1334	CAGGCTGCAGGGTTACCG	Amplification of the *HTT* locus in astrocytes for small-pool PCR.	[95]
oVIN-1345	ATGAAGGCCTTCGAGTCCCTCAAGTCCTTC	Amplification of the *HTT* locus in neurons for small-pool PCR.	This study
oVIN-1347	CGGCTGAGGCAGCAGCGGCTGT	Amplification of the *HTT* locus in neurons for small-pool PCR.	This study
oVIN-1251	GAGCGTGGGTCTCCGCCCAG	Amplification of the *DMPK* locus for small pool PCR.	[95]
oVIN-1252	CACTTTGCGAACCAACGATA	Amplification of the *DMPK* locus for small pool PCR.	[95]

oVIN-100	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	Small-pool PCR probe for the repeats	[96]
oVIN-1321	GCCTCCCTTACCATGCAGT	Amplifies human *HTT* locus for PacBio amplicon sequencing.	This study
oVIN-3490	ACAAGGGAAGACCCAAGTGA	Amplifies human *HTT* locus for PacBio amplicon sequencing.	This study
oVIN2612	AAGGGCATCGACTTCAAGGA	Amplification of GFP by qPCR	This study
oVIN2613	GGCGGATCTTGAAGTTCACC	Amplification of GFP by qPCR	This study
oVIN2622	CGATAAGAACCTGCCCAACG	Amplification of SpCas9v4 by qPCR	This study
oVIN2623	GGTTGGTCTTGAACAGCAGG	Amplification of SpCas9 by qPCR	This study
oVIN3321	GGAGACGGGGTCACCCACAC	Amplification of mouse β-actin for qPCR	This study
oVIN3322	AGCCTCAGGGCATCGGAACC	Amplification of mouse β-actin for qPCR	This study

oVIN3766	CCTCACCTTCAGAATCCCGT	Amplification of SpCas9v5 by qPCR	This study
oVIN3767	CCTTGTCCACAACCTCCTCG	Amplification of SpCas9v5 by qPCR	This study

1683