

CRISPR/Cas9-Mediated Excision of ALS/FTD-Causing Hexanucleotide Repeat Expansion in *C9ORF72* rescues major disease mechanisms *in vivo* and *in vitro*

Katharina E. Meijboom^{1,2}, Abbas Abdallah¹, Nicholas P. Fordham¹, Hiroko Nagase¹, Tomás Rodríguez⁵, Carolyn Kraus⁵, Tania F. Gendron³, Gopinath Krishnan², Rustam Esanov⁴, Nadja S. Andrade⁴, Matthew J. Rybin⁴, Melina Ramic⁴, Zachary D. Stephens⁸, Alireza Edraki⁵, Meghan T. Blackwood¹, Aydan Kahriman², Nils Henninger², Jean-Pierre A. Kocher⁸, Michael Benatar⁶, Michael H. Brodsky⁷, Leonard Petrucelli³, Fen-Biao Gao², Erik J. Sontheimer⁵, Robert H. Brown², Zane Zeier^{4,5*}, Christian Mueller^{1*}

¹ Horae Gene Therapy Center, University of Massachusetts Medical School, Worcester, MA 01605, USA

² Department of Neurology, University of Massachusetts Medical School, Worcester, MA 01605, USA

³ Department of Neuroscience, Mayo Clinic, Jacksonville, FL 32224, USA

⁴ Department of Psychiatry & Behavioral Sciences, Center for Therapeutic Innovation, University of Miami Miller School of Medicine, Miami, FL 33136, USA

⁵ RNA Therapeutics Institute and Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA 01605, USA

⁶ Department of Neurology, University of Miami Miller School of Medicine, Miami, FL, 33136, USA

⁷ Department of Molecular, Cell and Cancer Biology, University of Massachusetts Medical School, Worcester, MA, 01605, USA

⁸ Department of Quantitative Health Sciences. Mayo Clinic, Rochester, 55905, USA

* These authors contributed equally to this work

Zane Zeier: zzeier@med.miami.edu

Christian Mueller: Christian.mueller4@sanofi.com

Abstract

A hexanucleotide repeat expansion (HRE) consisting of GGGGCC₂₄₊ in the *C9ORF72* gene is the most common genetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). Both are fatal neurodegenerative diseases with no current approved treatments that significantly slow disease progression or extend life expectancy. Several hypotheses have emerged to explain how this HRE causes neuronal death, including *C9ORF72* haploinsufficiency, sequestration of RNA-binding proteins in the nucleus, and production of dipeptide repeat proteins. In the present study we used a CRISPR/Cas9 gene-editing approach to remove the HRE from the *C9ORF72* genomic locus, designing guide RNAs (gRNAs) flanking the HRE, and delivered Cas9 and gRNAs using adeno-associated virus serotype 9 (AAV9) vectors. Here, we demonstrate successful excision of the HRE in *C9ORF72* in primary cortical neurons and brains of three mouse models containing the *C9ORF72* expanded HRE (ranging from 500-600 repeats) as well as in patient-derived iPSC motor neurons and brain organoids (450 repeats). This resulted in a reduction of RNA foci, poly-dipeptides and haploinsufficiency, the major hallmarks of C9-ALS/FTD, making this an extremely attractive therapeutic approach to these diseases.

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease that affects upper and lower motor neurons in the brain and spinal cord. It is characterized, primarily, by progressive muscle weakness that ultimately results in respiratory failure. ALS is a debilitating disease that is invariably fatal with most affected individuals succumbing within 3-5 years of symptom onset ¹. The 2011 discovery of a hexanucleotide repeat expansion (HRE) in the non-coding region of the Chromosome 9 Open Reading Frame 72 (*C9ORF72*) gene revealed this to be the most common known cause of both hereditary (40%) and apparently-sporadic (5-6%) ALS cases ²⁻⁴. In healthy individuals the size of the hexanucleotide sequence is less than 24 repeats, but affected individuals have expansion mutations that can number in the thousands of repeats ^{5,6}. The HRE is also the most common genetic cause of frontotemporal dementia (FTD), a related neurodegenerative disease with multiple clinical manifestations including disinhibited behavior, language disorders, loss of executive control, and motor symptoms ⁷. Both ALS and FTD are aggressive diseases with no current treatments that significantly slow disease progression or extend life expectancy ⁸.

The *C9ORF72* gene locus produces three protein-coding RNA transcript variants, V1, V2 and V3 through alternative splicing and transcription start site use ². V1 is translated into a 222 amino acid protein, while V2 and V3 encode the predominant 481 amino acid *C9ORF72* protein ². *C9ORF72* interacts with endosomes and is required for normal vesicle trafficking, autophagy induction and lysosomal biogenesis in diverse cell types, including motor neurons ^{9,10}. The HRE gives rise to three pathological hallmarks of *C9ORF72* ALS: 1). It impairs expression, leading to *C9ORF72* haploinsufficiency that compromises neuronal viability ^{10,11}. 2). Sense and antisense transcription of the *C9ORF72* HRE produces G₄C₂ and C₄G₂ transcripts that accumulate in cell nuclei and sequester RNA-binding proteins, resulting in RNA foci ^{2,12-19}. 3). Sense and antisense HRE

transcripts are translated via an abnormal mechanism into toxic poly-dipeptides, producing glycine-arginine (GR), glycine-proline (GP), and glycine-alanine (GA) poly-dipeptides in the sense direction and proline-alanine (PA), proline-arginine (PR), and GP in the anti-sense direction^{20–25}. These aggregation-prone poly-dipeptides have been found in the brains and spinal cords of C9ORF72 ALS and FTD patients^{17,25,26} and are toxic in cell culture^{27–30} and animal models^{31–37}. Furthermore, there is a direct form of cooperative pathogenesis between these gain- and loss-of-function mechanisms in C9ORF72, where C9ORF72 haploinsufficiency impairs clearance of poly-dipeptides, making motor neurons hypersensitive to the toxic effects of the poly-dipeptides^{10,11}. Despite these advances in elucidating the molecular pathology of the HRE, there are no approved therapies that have been developed based on this knowledge. Existing treatments do not directly target the pathological hallmarks of C9ORF72 ALS/FTD and are very modestly beneficial⁸.

The field of gene therapy and gene editing offer potential treatment modalities for C9-ALS/FTD pathology by attenuating C9ORF72-related toxicity at the genetic level. Among these, one attractive approach is to physically excise the repeat expansion mutation from the C9ORF72 genomic locus using gene editing tools such as the CRISPR/Cas9 system. The Cas9 endonuclease produces double-strand breaks in genomic DNA at specific sequences that are determined by complementarity with a guide RNA (gRNA) and the presence of a protospacer-adjacent motif (PAM)^{38,39}. This versatile genome editing technology is currently being tested in human clinical trials for treatment of sickle cell disease, transthyretin amyloidosis, and vision loss by Leber's congenital amaurosis^{40–42}.

Here, we developed an AAV9-mediated CRISPR/Cas9 gene-editing approach to remove the GGGGCC_n HRE from its genomic locus using gRNAs flanking the HRE. To thoroughly evaluate

the efficiency of gene editing and the ability to confer gene editing *in vivo*, we employed both human and murine derived C9ORF72-related model systems. We utilized patient-derived iPSCs, induced motor neurons (iMNs) and brain organoids with the endogenous mutation as well as primary neurons with expanded repeats from three different C9BAC transgenic mouse models. Furthermore, we tested our gene editing approach *in vivo* in three different C9BAC transgenic mouse models and show that HRE excision results in a dramatic reduction of RNA foci and toxic poly-dipeptides – both gain-of-function pathologies caused by the HRE – while rescuing transcript and protein levels.

2. Results

To excise the HRE from *C9ORF72*, we designed gRNAs flanking the repeat sequence. Since the HRE is located within an intronic region, our aim was to excise the repeat without affecting exon sequences. However, only 35 base pairs (bp) separate the HRE from exon 2, complicating our strategy of targeting intronic sequence flanking the HRE. Therefore, as an alternative strategy, we used gRNAs targeting intron 2, thereby excising the HRE as well as exon 2, a non-coding exon present only in V2 mRNAs, hypothesizing that exon 2 excision would not greatly affect overall *C9ORF72* mRNA and protein levels. Four gRNAs were designed flanking the HRE and exon 2, with two gRNAs targeting the 3' DNA sequence downstream of the HRE and two targeting the upstream, 5' sequence (Fig. 1a).

To evaluate gRNA efficiencies, HEK293T cells were co-transfected with a plasmid expressing *S. pyogenes* Cas9 (SpCas9) and another plasmid expressing two gRNA combinations flanking the repeat in groups of gRNAs 1 and 3, gRNAs 1 and 4, gRNAs 2 and 3, or gRNAs 2 and 4, respectively. Cell lysates were collected 48 hours after transfection, genomic DNA was isolated,

and the targeted region amplified by PCR with primers flanking the HRE sequence (Fig. 1a,b). In HEK293T cells, *C9ORF72* contains only three G₄C₂ repeats. Given the lower GC content from this small number of repeats, DNA polymerase is able to amplify the sequence yielding a 523bp band in unedited *C9ORF72*. Excision of the target sequence and subsequent repair of the DNA template produces a shorter amplicon of 249bps for gRNA1,3, 246bps for gRNA1,4, 322bps for gRNA2,3, and 321bps for gRNA2,4. Cells transduced with a plasmid expressing only GFP and untreated cells were used as experimental controls, producing only one amplicon of 523bps, confirming the presence of unedited template sequences in non SpCas9 and gRNA treated cells (Fig. 1b). In contrast, *C9ORF72* editing was observed for all four combinations of gRNAs, as indicated by the presence of smaller amplicons resolved by electrophoresis. Genome editing was very efficient using gRNA2,3 and gRNA2,4 as indicated by the presence of bright bands at 322bp and 321bp, respectively, with little amplification of 523bp products that are produced from unedited template DNA.

We next estimated the percentage of segmental deletions after editing of HEK293T cells using Sanger-sequencing, to test which gRNA pairs were most effective. With gRNA2,3 and gRNA2,4, we achieved over 80% editing efficiency (Fig. 1c). Insertion and deletion (InDel) analysis for gRNA2,3-mediated genome editing demonstrated 65.4% of correct editing and 23.2% excision of the HRE plus an insertion of 1 nucleotide (Fig. 1d). For gRNA2,4, 22.6% correct editing was achieved, with editing of the HRE plus an insertion of 1 or 2 nucleotides in 66% of editing events (Fig. 1d). Since these small insertions are located within the intron, they do not change amino acid composition of the *C9ORF72* protein and are unlikely to change transcription levels. Since editing efficiencies with Sanger-sequencing are only estimates because of amplification bias and to further investigate on-target genomic rearrangements after CRISPR Cas9 editing, we assessed the rate of

intended HRE deletions or other types of genomic rearrangements with the UDiTaS unidirectional sequencing approach ⁴³, using a single locus-specific primer and a transposon specific primer containing unique molecular identifiers for library construction. UDiTaS analysis at the gRNA2,3 and gRNA2,4 edited loci revealed primarily excision of the HRE. As observed by the amplicon deep sequencing, all groups, including control groups, demonstrated a 2.6-4.9% InDel rate. Additional analysis showed these mostly consisted of 6 bp deletions, which was likely due to the HEK293T cells containing a different number of hexanucleotide repeats than the hg38 assembly we used for analysis. Inversions accounted for <0.4% and was not significantly higher than in the control groups and there was no evidence of further on-target genomic rearrangements. The SpCas9 endonucleases can generate unintended editing at off-target sites throughout the genome ⁴⁴, particularly at sites with sequence homology to gRNAs. Since unintended editing is a potential safety concern, we performed an *in silico* analyses of the human genome (Supplementary table 1) to identify off-target sequences with less than three mismatches with gRNA target sequences. For coding sequences, we identified 2 mismatch targets for gRNA2 and gRNA3 and three mismatch targets for gRNA4. Using amplicon deep sequencing, we interrogated genomic DNA from edited HEK293 cells, but found no evidence of off-target editing at near-cognate sequences (Supplementary Fig. 1), indicating high guide-specificity of genome editing. Therefore, combinations gRNA2,3, and gRNA2,4, and SpCas9 were packaged in separate AAV9 vectors and used in all subsequent *in vitro* and *in vivo* experiments.

While we achieved highly efficient editing in HEK293T cells, they have only 3 G₄C₂ repeats. Therefore, we next tested the editing system in mouse primary cortical neurons from a C9ORF72 BAC111 transgenic line with 600 repeats, a more disease-relevant *in vitro* model system ⁴⁵. BAC111 mice were crossed with B6J.129(Cg)-Gt(Rosa)26Sor^{tm1.1(CAG-cas9*,-EGFP)Fz}/J to create

het:het C9⁺/Cas9⁺ embryos where Cas9 is endogenously expressed in neurons harboring the HRE expansion. Primary cortical neurons were then harvested at embryonic day 15 (E15) and transduced with AAV9 expressing gRNA2,3 or gRNA2,4 at day *in vitro* (DIV) 4 at a multiplicity of infection (MOI) of 50,000 (Fig. 2a). For these experiments, a gRNA targeting the *Rosa* locus was packaged in AAV9 and used as a control since this sequence is not naturally expressed in mice or humans. Since these motor neurons express *C9ORF72* with an expanded repeat (600 repeats), the Q5 DNA polymerase used for PCR is not expected to amplify through this GC-rich region. Therefore, PCR amplification with primers flanking the HRE will only occur if the G₄C₂ repeats are excised during gene editing. Following PCR, we observed no DNA amplification in untreated primary neurons or those treated with AAV9-Rosa. Alternatively, amplification of template DNA isolated from neurons transduced with AAV9-gRNA2,3 and AAV9-gRNA2,4 produced a strong band at around 320 bps, demonstrating efficient excision of the HRE mutation in neurons (Fig. 2b).

A limitation of this experiment is that we used neurons that endogenously express Cas9. While this is more efficient, it is also less clinically relevant because Cas9 does not need to be introduced exogenously using AAV vectors. Therefore, to assess editing efficiency in non-Cas9 expressing primary neurons, we used AAV vectors to introduce Cas9 and gRNA expression in primary neurons from heterozygous BAC111 C9⁺/Cas9⁻ embryos. As in neurons with endogenous Cas9 expression, we observed efficient excision of the HRE mutation in neurons transduced with AAV-Cas9 and either AAV9-gRNA2,3 or AAV9-gRNA2,4, but not in neurons treated with buffer (PBS) or AAV9-Rosa (Fig. 2b). The BAC111 mouse model harbors 6–8 copies of exon 1–6 of the human *C9ORF72* gene, including 600 G₄C₂ repeats, and recapitulates the molecular pathology of C9-ALS/FTD since mice develop RNA foci and toxic dipeptides⁴⁵. This model is in contrast with C9-

ALS/FTD patients, who have only one copy of the *C9ORF72* gene with a pathological expansion⁴⁶. Therefore, we validated our system using another transgenic mouse model (C9-500) developed by Dr. Laura Ranum at the University of Florida⁴⁷. The C9-500 mouse expresses only one complete copy of the human *C9ORF72* gene with 500 G₄C₂ repeats and does not express Cas9 endogenously. Using the same experimental approach as before, we transduced primary neurons from C9-500 transgenic mice with AAV9-SpCas9 and AAV9-gRNA2,3 or AAV9-gRNA2,4. Again, we observed efficient excision of the HRE as indicated by efficient amplification of the *C9ORF72* target sequence (Fig. 2b).

Having observed efficient removal of the HRE, we next sought to determine whether genome editing ameliorates two pathogenic gain-of-function hallmarks of C9-ALS/FTD: RNA foci and production of toxic poly-dipeptide proteins. To this end, primary neurons from BAC111 and C9-500 transgenic mice were treated with AAV9-gRNA2,3 or AAV9-gRNA2,4 and AAV9-SpCas9 (Fig. 2a). To evaluate poly-dipeptide production, we used a sensitive immunoassay⁴⁸ to quantify levels of GP, an abundant poly-dipeptide synthesized from expanded sense and antisense *C9ORF72* transcripts. Both gRNA combinations resulted in a significant ~50% reduction of GP poly-dipeptides in both neuronal models (Fig. 2c). Next, we labeled RNA foci by fluorescent *in situ* hybridization (FISH) staining (Fig. 2d) and binned nuclei according to the percentage of observed foci in BAC111 and C9-500 primary neurons (1-5 foci, 6-10 foci and 11+ foci). Six days after AAV9-gRNA2,3 and AAV9-gRNA2,4 treatment, the percentage of nuclei with foci was markedly decreased, from 64% in the untreated PBS condition to 17% (AAV9-gRNA2,3) and 16% (AAV9-gRNA2,4) in the treated conditions in C9-500 primary neurons (Fig. 2e). Similarly, in BAC111 primary neurons, the percentage of nuclei with foci decreased after treatment from 60% to 38% (AAV9-gRNA2,3) and 29% (AAV9-gRNA2,4). Curiously, editing resulted consistently

in enlarged foci in BAC111 primary neurons, but not C9-500 primary neurons (Fig. 2d). To exclude the possibility that the formation of large foci is caused by the accumulation of excised DNA, slides with primary neurons were treated with DNase or RNase prior to analysis by FISH. We found that DNase resulted in degradation of genomic DNA but labeled HRE foci remained. In contrast, treatment with RNase degraded all foci, including large foci (data not shown). These findings indicate that excised HRE DNAs do not contribute to the formation of HRE foci and that enlarged foci observed in BAC111 primary neurons is a consequence their having more copies of the expanded *C9ORF72* transgene relative to C9-500 neurons. Collectively, these results demonstrate that both gRNA pairs acting with CRISPR/Cas9 nuclease correctly excised the HRE and thereby reduced levels of RNA foci and poly-dipeptide proteins *in vitro*; that is, they are both effective in reducing gain-of-function effects in neurons containing the human *C9ORF72* HRE.

One proposed mechanism by which the HRE leads to pathology is through reduced expression of *C9ORF72* protein through epigenetic repression⁴⁹⁻⁵¹, inhibition of RNA processing^{12,52,53}, or changes in relative expression of RNA variants⁵⁴. Through alternative splicing and transcription start site utilization, three protein-coding *C9ORF72* transcripts are produced, encoding two isoforms of the *C9ORF72* protein². Transcript variants 2 (NM_018325.3) and 3 (NM_001256054.1) encode a 481 amino acid protein, whereas variant 1 (NM_145005.5) encodes a shorter 222 amino acid protein². The V2 variant constitutes the majority (~80%) of *C9ORF72* transcripts in the brain whereas V1 and V3 represent 20% and less than 3%, respectively⁵⁵. Given the evidence that *C9ORF72* haploinsufficiency may contribute to pathology, we intentionally designed gRNA pairs to exclusively target non-protein-coding DNA. In addition, to determine whether gene editing restores normal expression levels of *C9ORF72* transcripts, we used digital droplet PCR (ddPCR). Given the location of gRNA2,3 and gRNA2,4, we were unable to design

ddPCR probes to specifically target V2; however, we could indirectly infer V2 levels using probes for V1, V3 and a probe that amplifies all three variants. We extracted RNA from treated and untreated C9-500 primary neurons and normalized expression levels to TATA-box Binding Protein (TBP). Gene editing increased V1 transcription while a significant decrease in V3 expression was observed (Fig. 2f). Since V3 expression makes up for less than 3% of all C9 transcripts⁵⁵, there appeared to be little effect on total transcriptional output from the locus with no significant effect observed using a probe that detects all three transcripts (Fig. 2f). Likewise, levels of the larger, dominantly expressed C9ORF72 protein isoform were unchanged in edited C9-500 primary neurons relative to unedited neurons (Fig. 2g). Together, these results demonstrate that excision of the HRE does not significantly impact C9ORF72 transcription or protein levels, a consequence that could have deleterious effects.

Having established the editing efficacy using AAV9-gRNA2,3 and AAV9-gRNA2,4 in different C9ORF72 primary neurons from multiple transgenic mouse models, we next evaluated their efficacy *in vivo* using three transgenic mouse models: BAC111/Cas9, BAC112, and C9-500 mice. Since BAC111/Cas9 mice express Cas9 endogenously, the performance of gRNAs can be evaluated more effectively. Since BAC112 mice have multiple copies of the expanded C9ORF72 gene but do not express Cas9, we were able to evaluate the efficiency of gene editing using a dual vector system. Since the C9-500 mouse has only one copy of the HRE expansion, it is more clinically-relevant - from a gene editing perspective. While none of these mouse models present the progressive neurodegeneration seen in ALS and FTD patients, they all exhibit gain-of-function pathological features seen in patients, such as RNA foci, and RAN translation resulting in the presence of toxic poly-dipeptides. Young adult BAC111 mice (aged 2–3-months) were injected bilaterally in the striatum with AAV9-gRNA2,3 or AAV9-gRNA2,4, (7×10^9 vector genomes (VG)

of the gRNA vectors) while BAC112 and C9-500 mice additionally received 2×10^{10} VG of AAV9-Cas9. Eight weeks following injection, mice were sacrificed and striatal tissues from both brain hemispheres were harvested (Fig. 3a). To evaluate gene editing events, we again used a PCR assay with primers flanking the HRE sequence, whereby amplification only occurs in the presence of edited DNA but not for templates with the long GC-rich HRE sequence (BAC111: 600 repeats, BAC112: 550 repeats, C9-500: 500 repeats). As observed in primary neurons, gRNA2,3 and gRNA2,4-edited DNA produced the expected amplicon, while equivalent amounts of DNA from PBS and AAV9-Rosa/AAV9-Cas9 treated animals did not (Fig. 3b). These results document successful *in vivo* HRE excision in three different transgenic mouse models of *C9ORF72* ALS and FTD.

Having observed excision of the HRE, we next sought to determine whether this attenuates toxic gain-of-function pathologies *in vivo*. As before, we used ELISA and FISH to quantify poly-dipeptide and HRE RNA foci levels, respectively. We found levels of poly-GP²⁴, to be significantly decreased in the striatum of C9-500 mice infused with AAV9-SpCas9 and either AAV9-gRNA2,3 or AAV9-gRNA2,4, but not in mice treated with PBS (Fig. 3c). Although poly-GP is abundant and commonly used as a biomarker for overall poly-dipeptide production, it is less toxic than other poly-dipeptides. The arginine-containing poly-dipeptides, poly-GR and poly-PR, are particularly cytotoxic²⁸. The currently-identified mechanisms contributing to poly-GR toxicity include impaired protein translation^{56,57}, formation of stress granules²⁹, inhibition of DNA repair⁵⁸ and reduced mitochondrial functionality³⁷. To validate that HRE excision reduces levels of arginine-containing poly-dipeptides, we evaluated the effects of HRE excision on poly-GR levels. We found a significant and robust reduction of poly-GR levels after treatment with AAV9-

gRNA2,3 and gRNA2,4 in the striatum of BAC111/Cas9 (50% reduction) and in C9-500 mice (60% reduction) (Fig. 3d).

Next, we quantified the percentage of nuclei with HRE RNA foci and assigned nuclei to one of several bins: 1-5 foci, 6-10 foci and 11+ foci. For C9-500 mice, the number of nuclei in the striatum with foci was significantly and robustly reduced from 86% in the untreated condition to 20-33% in the treated conditions. As in the BAC111/Cas9 mice, the number of striatal nuclei with foci was reduced from 44% in controls to 17-18% in the treated animals. Moreover, fewer foci per cell were detected in treated animals, where both gRNA combinations are potent inhibitors of G₄C₂ RNA foci (Fig. 3e, f). As in primary neurons, V-all transcript levels and C9ORF72 long isoform protein levels remained unchanged after treatment. Collectively, these results strongly suggest that removal of the HRE corrects toxic gain-of-function pathologies (RNA foci and poly-dipeptides) *in vivo* and are unlikely to induce further haploinsufficiency (Fig. 3g, h).

As proof-of-mechanism, we delivered Cas9 and gRNAs by AAV. Because of the large size of SpCas9 (4.1kb) we delivered SpCas9 and the gRNA pairs separately in two AAV9 vectors in C9-500 mice. Since gene editing will only occur when a cell is transduced with both the nuclease and gRNAs, we stained both using RNAScope (Fig. 4a) and counted how many cells were transduced by both vectors after striatal injections in C9-500 mice. We found that most cells in the striatum were transduced by both AAVs, 58% for gRNA2,3 and 69% for gRNA2,4 treated C9-500 mice (Fig. 4b).

To investigate editing efficiencies after treatment with SpCas9 and gRNA2,3 or gRNA2,4 in C9-500 mice, we first used PacBio No-Amp targeted sequencing to sequence through the editing sites and the GC-rich HRE⁵⁹. After two runs with limited numbers of on-target reads of mouse striatal

DNA samples, we opted to use UDiTaS to measure HRE excision and investigate on-target rearrangements and AAV integrations. UDiTaS analysis at the gRNA2,3 and gRNA2,4 edited loci revealed 15% and 23% HRE excision respectively, without large InDels (Fig. 4d). As to be expected⁶⁰, we found high levels of AAV integration at the gRNA2,3 and gRNA2,4 editing site (Fig. 4d). Since UDiTaS captures 140bp from each end of the amplicons we could not distinguish between AAV integrations paired with a removal of the HRE, and AAV integrations in which the HRE remained intact. To resolve this, we resorted to analyzing the PacBio No-Amp data on similarly treated samples and determined that 78.6% of AAV integrations were paired with removal of the HRE (data not shown). This level of HRE excision supports the previous data in which we found a >50% (Fig. 3c-f) reduction of RNA foci and RAN translation, and >50% of striatal cells were transduced with Cas9 and gRNAs in treated C9-500 mice (Fig. 4b). Interestingly, the percentage of inversions was higher in treated C9-500 brains, than in HEK293T cells (Fig. 1e), however, due to the relatively short UDiTaS read length we were not able to see if these inversions still had the HRE or if only the 3' end of the excised sequence was inverted and back integrated.

To investigate the effects of HRE excision with AAV9-gRNA2,3 and gRNA2,4 on C9ORF72 pathology in human motor neurons, we used an iPSC line from an C9-ALS/FTD patient carrying a pathogenic *C9ORF72* HRE or a control line derived from an individual with fragile x syndrome with a repeat expansion in the *FMR1* gene⁶¹. Both iPSC lines were derived from lymphocytes and used to produce induced motor neurons (iMNs), as we previously reported⁶¹. The resulting iMN cultures were co-transduced with AAV9-SpCas9 and either AAV9-gRNA2,3 or AAV9-gRNA2,4. After seven days, genomic DNA was harvested and analyzed for genome editing using PCR and primers spanning the HRE. Notably, C9-ALS/FTD patient-derived cells harbor one expanded *C9ORF72* allele and one allele with a WT, unexpanded G₄C₂ repeat sequence. Transgenic mice

used in this study do not have WT human *C9ORF72* alleles, ; the transgene harbors only the expanded HRE sequence. Therefore, unlike cells from transgenic mice, in human patient-derived cells the PCR assay will produce a single amplicon of 523 bps due to amplification of the WT allele template and absence of amplification through the expanded allele. Healthy individuals, or patients with fragile x syndrome, have two WT *C9ORF72* alleles and will therefore produce the same 523 bp amplicon in the PCR assay. Since the gRNAs 2,3 and 2,4 will induce excision of the G₄C₂ repeats within WT alleles as well as excision of the HRE in expanded alleles, the result is a shift from 523 to 320 bps when editing occurs in either patient or control cell lines. We found that co-transduction of C9-ALS/FTD iMNs resulted in a shift from amplification of a 523bp product to the 320bp product in cells treated with AAV9-SpCas9 and either gRNA pair, but not in C9-ALS/FTD or fragile X syndrome control iMNs without transduction with AAV9-SpCas9 (Fig. 5a). However, based on this PCR assay alone, it is not possible to determine whether gene editing occurs at WT alleles, expanded alleles or both. To address this, iPSC cultures were co-transduced with AAV9-SpCas9 and AAV9-gRNA2,4 and serially plated at low density and single colonies manually selected to obtain clonal iPSC cell populations. Analysis of several clones indicated efficient editing events as indicated by the presence of both 523 and 320 bp bands (Fig. 5b). To ensure clonality of iPSC lines, colonies were sub-cloned by titration and manual selection. From clone #11, we selected three sub-clones (4, 6 and 11) to exemplify possible outcomes: no editing, editing of the WT allele alone, or editing of both alleles. Using PCR, clone 11-6 was found to have no editing as indicated by a 523 bp band while clones 11-4 and 11-11 were found to have editing as indicated by the presence of a 320 bp band (Fig. 5c). To determine whether gene editing of the expanded allele occurred in sub-clones, we used a repeat-primed PCR to amplify the HRE sequence followed by capillary electrophoresis to resolve amplicons as we have done previously

⁶¹. This analysis confirmed that deletion of the HRE had occurred for sub-clone 11-11, but not sub-clone 11-4, as indicated by the presence of a prototypical sawtooth pattern resulting from amplification of the HRE (Fig. 5d). PacBio No-Amp sequencing ⁵⁹ confirmed 100% excision of the HRE in expanded alleles in sub-clone 11-11, while the unexpanded WT allele was not edited (Fig. 5e). To assess the quality of the editing, we performed UDiTaS analysis on the sub-clone 11-11 and we found no genomic rearrangements or AAV integrations at the editing locus (Fig. 5f).

To evaluate the effect of complete HRE excision, we measured levels of GP poly-dipeptides in the parent *C9ORF72* ALS/FTD iPSC line and sub-clone 11-11. We found a near complete elimination of GP poly-dipeptides in the gRNA2,4-edited 11-11 sub-clone relative to untreated parent iPSCs (Fig. 5g). Next, we investigated whether AAV9 SpCas9/gRNA2,4 mediated gene editing could instigate HRE excision in *C9ORF72* ALS/FTD cerebral brain organoids. Cerebral brain organoids are 3D structures cultured from iPSCs that better recapitulate the physical and diverse cell-type composition of the human brain, as compared to 2D iMN cultures. Cultures of *C9ORF72* iPSCs were used to produce brain organoids as described previously ⁶² and transduced with an AAV9-GFP expression vector to visualize transduction (Fig. 5h). Next, organoids were co-transduced with AAV9-SpCas9/gRNA2,4 at DIV 5, DIV 17 or both DIV 5 and 17. At DIV 32, organoids were harvested, and genomic DNA was isolated for PCR analysis to detect genome editing events. Successful editing of brain organoids occurred in gRNA2,4-treated organoids after early treatment at DIV5, late treatment at DIV17 and with treatment at both time points (Fig. 5i). Two rounds of transduction produced the most robust band shift from 523 to 320 bps - indicating more efficient gene editing than was achieved by transduction at the 17-day time point.

Since the HRE is located in the promoter region of V2, the most highly expressed *C9ORF72* variant, one of the potential outcomes of excising the HRE is that reduced *C9ORF72* transcription

efficiency might exacerbate C9ORF72 haploinsufficiency. Alternatively, since HRE excision decreases the physical distance between V2 and the V1/V3 promoter, HRE excision could potentially increase V2 expression. In the brain of *C9ORF72* BAC transgenic mice and primary neurons derived from them, *C9ORF72* transcript and protein levels remained largely unchanged after treatment (Fig. 2f-g, 3g-h). However, this could be due to incomplete transduction of tissues, abnormal promoter activity resulting from genomic insertion of the human sequence into the mouse genome, or cross-reactivity of the C9ORF72 antibody with mouse *c9orf72* proteins. Therefore, we took advantage of the iPSC system, in which the HRE exists in its natural genomic location in a human cell. In this model, purified sub-clones can be compared without the potential confounds of incomplete transduction, inefficient editing or presence of mouse *c9orf72*. To determine whether HRE excision alters transcription rates, RNAs and proteins were isolated from parent *C9ORF72* iMNs or the AAV9-gRNA2,4-edited 11-11 isogenic iMNs. We found a significant increase of almost 50% in *C9ORF72* transcription using primers that detect all three variants and a trend towards increased V1 and V3 levels (Fig. 5j). Western analysis confirmed increased expression of C9ORF72 protein: levels of the long isoform, translated from V2 and V3, were robustly and significantly ($p < 0.05$) increased in gene-corrected iMNs as compared to the parent line (Fig. 5k). Similarly, the short C9ORF72 isoform, translated from V1, was increased by 40% (Fig. 5k, $p = 0.08$). Together, these results suggest that excision of the HRE does not exacerbate and indeed rescues haploinsufficiency.

Figure 1: HEK293T data

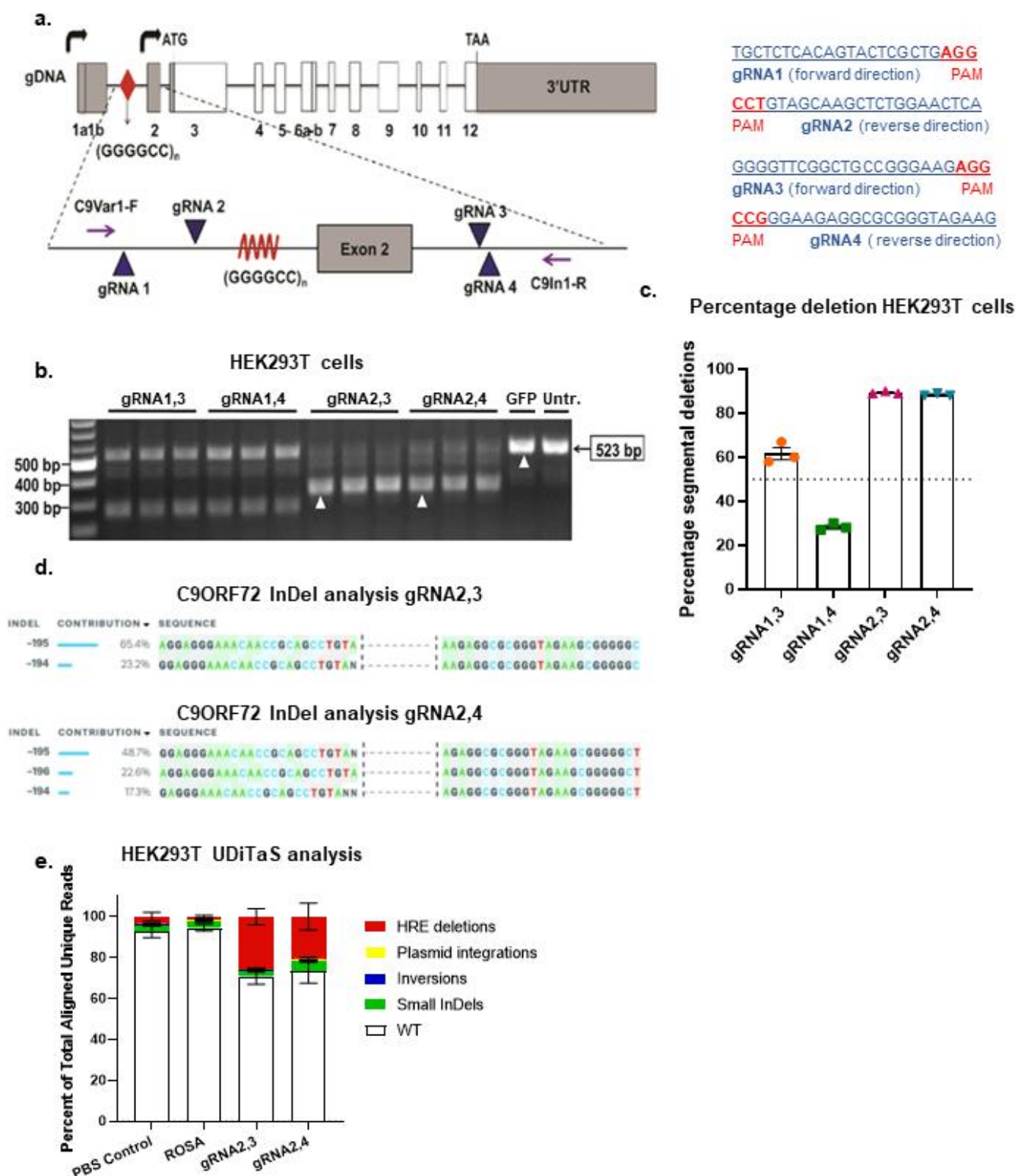


Figure 1. Screening of gRNAs flanking the (GGGGCC)_n HRE for excision, in HEK293T cells. **a.** Human *C9ORF72* gene showing (GGGGCC)_n repeats (red diamond) and the locations and sequences of guide RNAs 1, 2, 3 & 4 in addition to primers C9Var1-F and C9In1-R in relation to exon 2 and the (GGGGCC)_n repeat. **b.** Screening of gRNAs in HEK293T cells. The CRISPR-targeted *C9ORF72* HRE locus was PCR-amplified by C9Var1-F and C9In1-R primers and PCR products were visualized by agarose gel electrophoresis. The expected unedited PCR product size is 523 bp; edited PCR product size for 1-3 and 1-4 is ~250 bp; edited PCR product size for 2,3 and 2,4 is ~320 bp. **c.** Percentage of segmental deletions of the HRE in gRNA 1-3, 1-4, 2,3, 2,4 treated HEK293T cells measured after Sanger-sequencing. **d.** InDel analysis of gRNA2,3 and 2,4 after Sanger-sequencing, using ICE. **e.** Composition of edited alleles at gRNA2,3 and gRNA2,4 editing loci in *C9ORF72* in HEK293T cells by UDiTaS analysis. The graphs show the percentage HRE deletions, InDels, plasmid insertions, and inversions. Mean ±SEM, *n* = 4 per group.

Figure 2: *in vitro* data

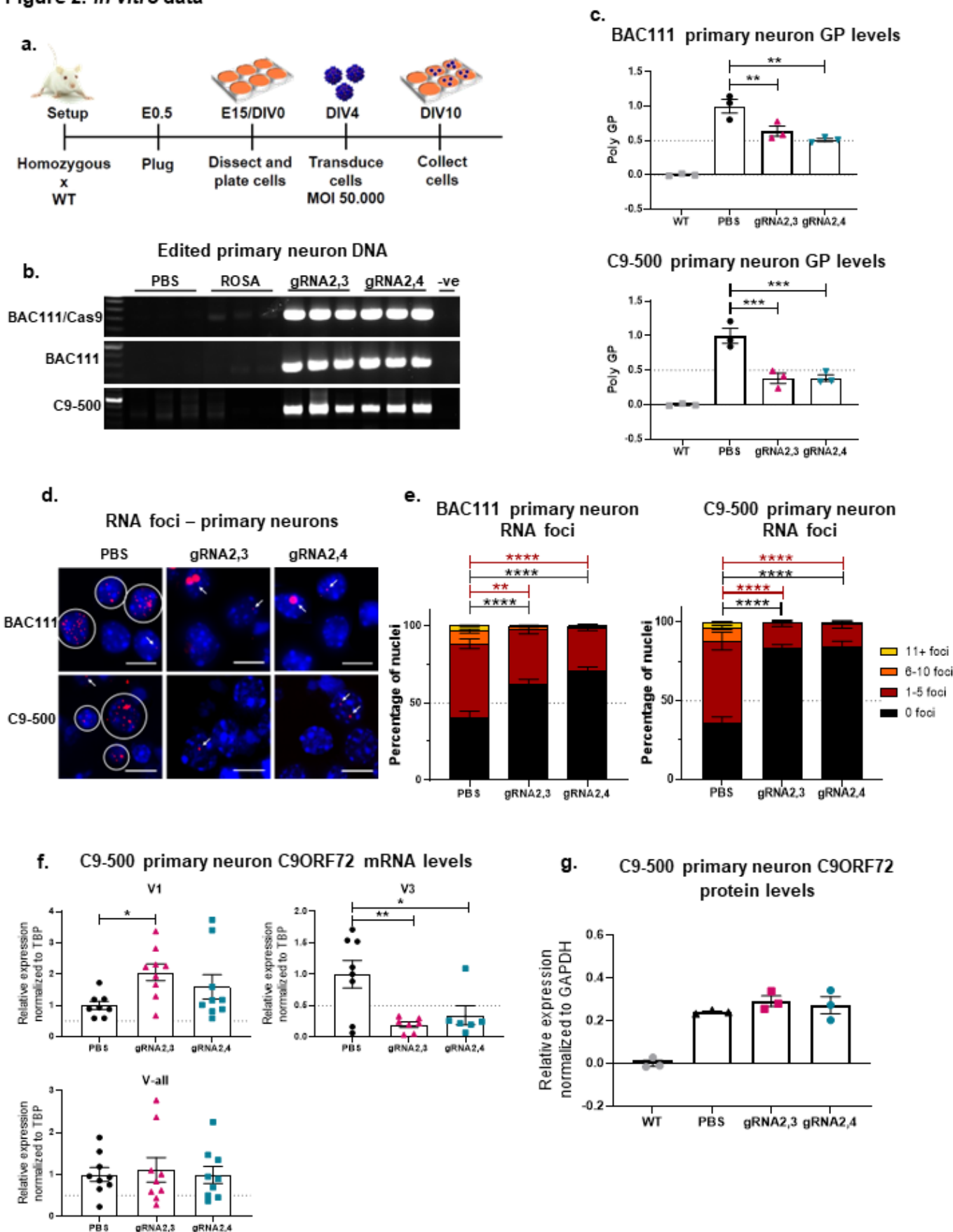


Figure 2. *In vitro* gene editing of embryonic primary cortical neurons results in a decrease of pathological hallmarks in C9ORF72 ALS/FTD. **a.** Schematic of harvest and AAV9-delivered Cas9 and gRNA treatment of different C9ORF72 BAC transgenic mouse cortical primary neurons. Homozygous C9ORF72 BAC transgenic mice were crossed with WT mice, resulting in heterozygous mice. To generate BAC111/Cas9 mice, homozygous BAC111 mice were crossed with homozygous B6J.129(Cg)-*Gt(Rosa)26Sor^{tm1.1(CAG-cas9*,-EGFP)Fz}*/J resulting in heterozygosity of both BAC111 C9ORF72 and Cas9. Cells were harvested at embryonic day (E) 15 and transduced with AAV9-gRNA2,3, AAV9-gRNA2,4 and controls at day *in vitro* (DIV) 4 and collected at DIV10 for further analyses. BAC111 and C9-500 primary neurons were treated with both AAV9-SpCas9 and AAV9-gRNA2,3 or AAV9-gRNA2,4 and BAC111/Cas9 with only AAV9-gRNA2,3 or AAV9-gRNA2,4. **b.** Genomic DNA of treated BAC111/Cas9, BAC111 and C9-500 primary neurons was amplified with primers flanking gRNA seed sequences, and PCR products were resolved by electrophoresis. The expected band for gRNA2,3 and gRNA2,4 edited DNA is 680 bps. No band is expected for unedited DNA. **c.** Relative expression of poly-GP in BAC111 and C9-500 primary neurons treated with AAV9-Cas9, and AAV9-gRNA2,3 or AAV9-gRNA2,4, assayed by sandwich immunoassay. Data represent mean $\pm$ SEM, $n = 3$ per experimental group, one-way ANOVA, Dunnett, * $p < 0.05$, ** $p < 0.01$. **d.** Representative images of sense RNA foci visualized by fluorescent *in situ* hybridization (FISH) in BAC111/Cas9 and C9-500 primary neurons treated with AAV9-gRNA2,3 and AAV9-gRNA2,4 for six days. Scale bars represent 10 μ m. **e.** Quantification of sense foci in treated and untreated BAC111/Cas9 and C9-500 primary neurons. Bar graph represents average percent of nuclei containing 0, 1-5, 6-10 and 11+ foci. Three independent experiments and > 600 nuclei counts per sample. Two-way ANOVA, Tukey: ** $p < 0.01$, **** $p < 0.0001$. **f.** ddPCR analysis of mRNA expression levels of variants V1, V3 and all variants jointly (V-all) after AAV9-gRNA2,3 and AAV9-gRNA2,4 treatment in C9-500 primary neurons. Expression levels were normalized to TBP and compared to the PBS-treated control. Mean $\pm$ SEM, $n = 8-9$, one-way ANOVA, Dunnett, * $p < 0.05$, ** $p < 0.01$. **g.** Densitometric quantification of the C9ORF72 long isoform in C9-500 primary neurons treated with AAV9-gRNA2,3 and AAV9-gRNA2,4. C9 protein levels were normalized against total protein levels. Levels from primary neurons from wild type mice were averaged and subtracted as a background signal. Mean $\pm$ SEM, $n = 3$, one-way ANOVA, Dunnett, differences between PBS and AAV9-gRNA2,3, AAV9-gRNA2,4 treated cells were not significant.

Figure 3: *in vivo* data – disease mechanisms

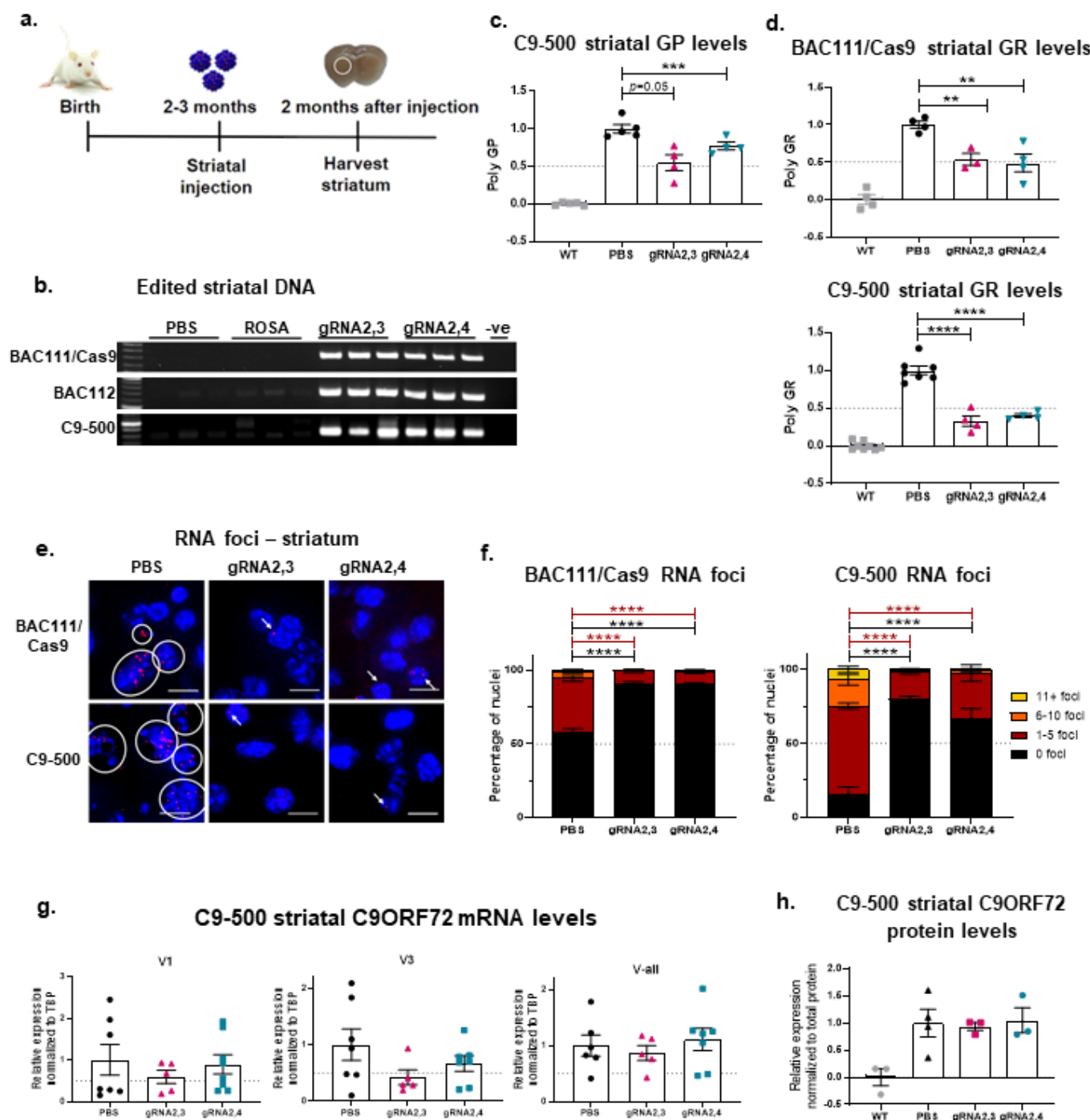


Figure 3. *In vivo* gene editing results in a decrease of RNA foci and poly-dipeptides in different C9 BAC transgenic mouse models. **a.** Schematic of treatment and harvest of AAV9-delivered Cas9 and gRNA treatment of different C9ORF72 BAC transgenic mouse models. Mice were treated at 2–4 months of age with bilateral striatal injections, and the striatums were harvested

8 weeks after treatment for further analyses. BAC111, BAC112 and C9-500 mice were treated with both AAV9-SpCas9 and AAV9-gRNA2,3 or AAV9-gRNA2,4 and BAC111/Cas9 mice with only AAV9-gRNA2,3 or AAV9-gRNA2,4. **b.** Genomic DNA of treated BAC111/Cas9, BAC112 and C9-500 mouse striatums was amplified with primers flanking gRNA seed sequences, and PCR products were resolved by electrophoresis. The expected band for gRNA2,3 and gRNA2,4 edited DNA is 680 bps. No band is expected for unedited DNA. AAV9-Rosa treatment and PBS treatment were used as a control. Genomic DNA was amplified with primers flanking gRNA seed sequences and PCR products resolved by electrophoresis. The expected band for gRNA2,3 and 2,4 edited DNA is 680 bps. No band is expected for unedited DNA. **c.** Relative expression of poly-GP in the striatums of treated C9-500 mice. Data represent mean $\pm$ SEM, $n = 4-5$ per experimental group, one-way ANOVA, Dunnett, $p = 0.09$, ** $p < 0.01$. **d.** Relative expression of poly-GR in the striatums of treated BAC111/Cas9 and C9-500 mice. Data represent mean $\pm$ SEM, $n = 4-8$ per experimental group, one-way ANOVA, Dunnett, **** $p < 0.0001$. **e.** Representative images of sense RNA foci visualized by FISH in BAC111/Cas9 and C9-500 striatums treated with AAV9-gRNA2,3, AAV9-gRNA2,4 and PBS. Scale bars represent 10 μ m. **f.** Quantification of sense foci in treated BAC111/Cas9 and C9-500 mice striatums. Bar graph represents average percent of nuclei containing 0, 1-5, 6-10 and 11+ foci. $n = 3-4$ mice per group, and > 600 nuclei were counted per sample. Two-way ANOVA, Tukey: **** $p < 0.0001$. **g.** ddPCR analysis of mRNA expression levels of variants V1, V3 and all variants jointly (V-all) after gRNA2,3 and gRNA2,4 treatment in C9-500 mice. Expression levels were normalized to TBP and compared to the PBS-treated control. Mean $\pm$ SEM, $n = 5-8$, one-way ANOVA, there were no significant differences between treated and untreated groups. **h.** Densitometric quantification of the C9ORF72 long isoform in treated C9-500 mice. C9 protein levels were normalized against total protein levels. Levels from wild type mice were averaged and subtracted as a background signal. Mean $\pm$ SEM, $n = 3-4$, one-way ANOVA, Dunnett, differences between PBS and gRNA2,3, gRNA2,4 treated cells were not significant.

Figure 4: *in vivo* data - editing

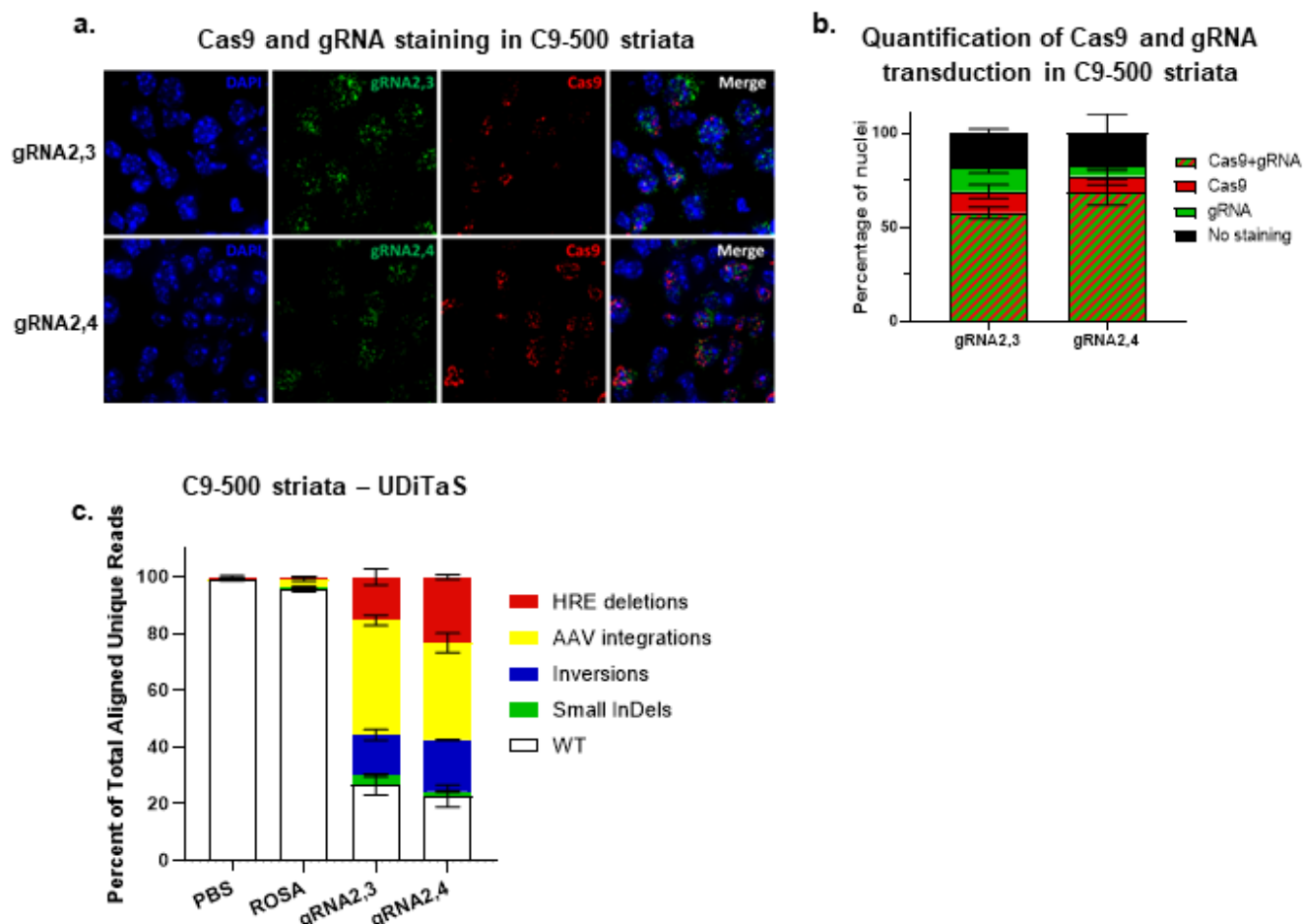


Figure 4. *In vivo* gene editing quantity and quality in C9 BAC transgenic mouse model. **a.** Representative 64x images of C9-500 striata labeled with RNAscope *in situ* hybridization for Cas9 (red) and GFP (green) expression and nuclei (DAPI, blue) after AAV9-Cas9 and AAV9-gRNA2,3 or AAV9-gRNA2,4 treatment. The gRNA vectors expressed GFP. **b.** Percentage of nuclei stained with Cas9, GFP, Cas9 and GFP, and no staining of AAV9-Cas9 and AAV9-gRNA2,3 or AAV9-gRNA2,4 treated C9-500 mice. **c.** Percentage of reads in which the HRE is excised after gRNA2,3 and gRNA2,4 editing in C9-500 striatal DNA measured by No-Amp PacBio. Mean $\pm$ SEM, $n = 3$ per group. **d.** Composition of edited alleles at gRNA2,3 and gRNA2,4 editing loci in C9ORF72 in C9-500 striata by UDiTaS analysis. The graphs show the percentage of HRE deletions, small InDels, AAV insertions, and inversions. Mean $\pm$ SEM, $n = 4$ per group.

Figure 5: iPSC data

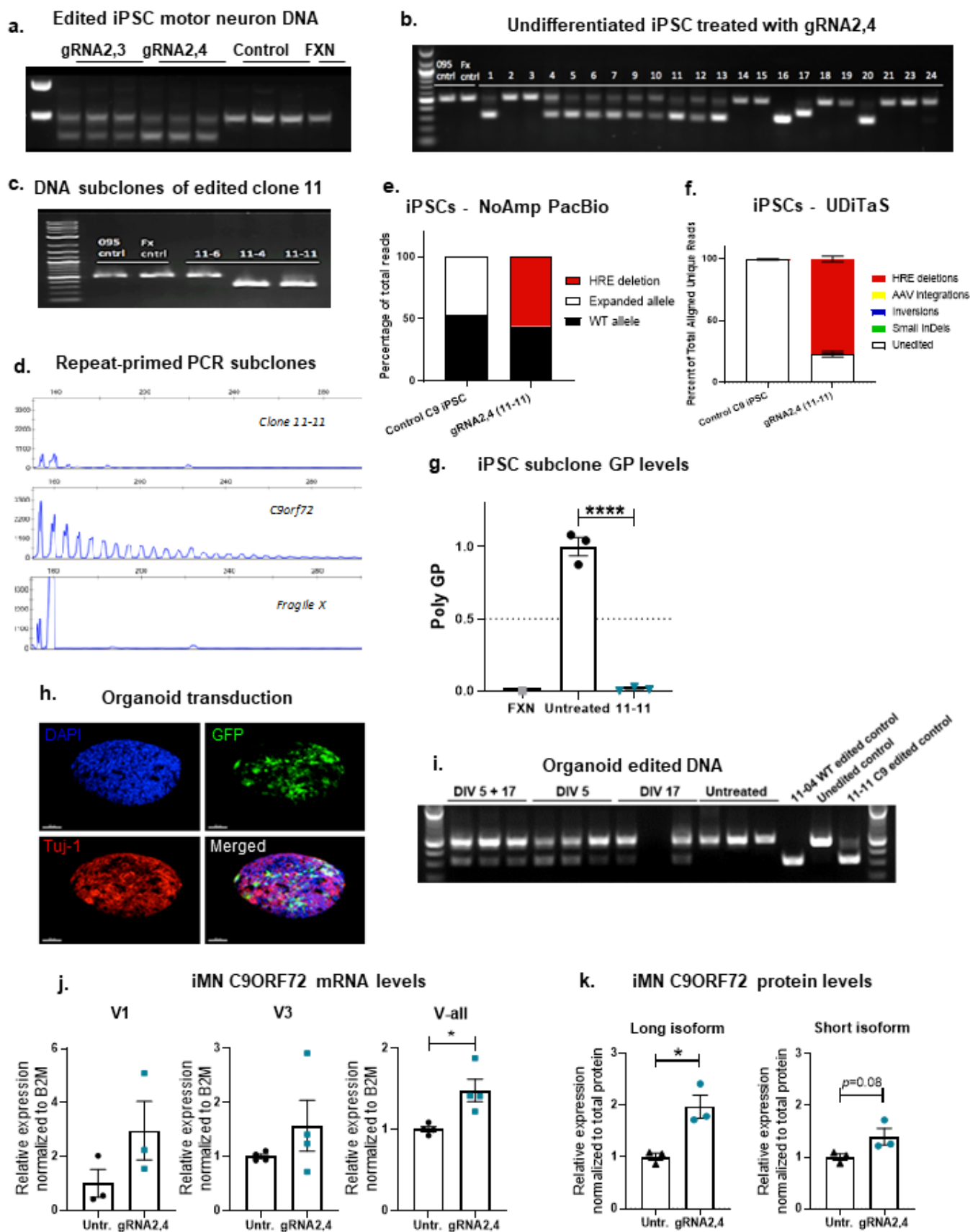


Figure 5. Genome editing of the hexanucleotide repeat (G₄C₂)_x expansion in C9ORF72 iPSCs and iMNs. **a.** C9ORF72 patient iPSC motor neurons transduced with Cas9 and gRNA2,3 and gRNA2,4. Genomic DNA was amplified with primers flanking gRNA seed sequences and PCR products resolved by electrophoresis. DNA isolated from Fragile X syndrome iPSC neurons (FXN) were used as a control. **b.** Genome editing of undifferentiated C9ORF72 iPSCs with gRNA2,4. iPSCs were co-transduced with AAV Cas9 and gRNA2,4, individual colonies (1-24) were manually selected and analyzed for genome editing by PCR using primers spanning the HRE. **c.** Clone 11 selected from iPSCs transduced with Cas9 AAV and gRNA2,4 was subcloned to generate lines 11-6, 11-4 and 11-11. Analysis of genome editing revealed editing for subclones 11-4 and 11-11 but not 11-6 as indicted by a shift from 521 to 321 base pair product using primers that span the HRE. **d.** Capillary electrophoresis fragment analysis of amplicons produced by repeat primed PCR of the C9ORF72 HRE for subclone 11-6,11-4, 11-11, and FGX and untreated patient iPSCs. **e.** Relative expression of poly-GP in the striatums of treated C9-500 mice. Data represent mean $\pm$ SEM, $n = 4-3$ per experimental group, one-way ANOVA, Dunnett, **** $p < 0.0001$. **f.** Representative confocal Z-stack images of organoid transduction. Organoids were stained for DAPI (blue) and immunolabeled with antibodies against Tuj-1 (red) and GFP (green). **g.** Genomic DNA was isolated and amplified by PCR using primers spanning the C9ORF72 expansion after transduction by Cas9 and gRNA2,4 AAV vectors at 5 days, 5 and 17 days, 17 days or untransduced. DNA from edited iPSC subclones 11-04, 11-11 and DNA from one unedited line was used as controls. **h.** ddPCR analysis of mRNA expression levels of variants V1, V3 and all variants jointly (V-all) in gRNA2,4 edited 11-11 isogenic iMNs. Expression levels were normalized to B2M and compared to untreated control patient iMNs. Mean $\pm$ SEM, $n = 3$, two-tailed Student's t -test, * $p < 0.05$. **i.** Densitometric quantification of the C9ORF72 long isoform and short isoform in gRNA2,4 edited 11-11 isogenic iMNs. C9 protein levels were normalized against total protein levels. Mean $\pm$ SEM, $n = 3$, two-tailed Student's t -test, * $p < 0.05$. **j.** Percentage of reads in which the HRE is excised in gRNA2,4 edited 11-11 isogenic iPSCs measured by No-Amp PacBio. **k.** Composition of edited alleles at the gRNA2,4 editing loci in gRNA2,4 edited 11-11 isogenic iPSCs by UDiTaS analysis. The graphs show the percentage of HRE deletions, small InDels, AAV insertions, and inversions.

3. Discussion

In this study we successfully excised the ALS/FTD-causing HRE in the first intron of *C9ORF72*, using AAV9 to deliver CRISPR SpCas9 nuclease and two gRNAs flanking the HRE to patient cells, primary neurons and three mouse models. This single intervention attenuated the three major disease mechanisms at play in C9-ALS/FTD. We (1) suppressed production of poly-GP dipeptides in edited isogenic patient cells and both poly-GP and poly-GR in primary neurons and mice, (2) reduced levels of nuclear sense RNA foci in two *C9ORF72* BAC transgenic mouse models and primary neurons, and (3) increased of *C9ORF72* mRNA transcript and protein levels by 50% in edited isogenic patient cells, thereby countering HRE-induced haploinsufficiency. These findings are consistent with, but extend, other reports demonstrating positive benefit of excising the HRE in *C9ORF72*. Selvaraj *et al.*⁶³ attenuated AMPAR-mediated excitotoxicity in patient isogenic cell lines edited by CRISPR/Cas9 to excise the HRE. The Talbot group used CRISPR/Cas9-mediated HDR to replace the pathogenic expansion in patient cells with a donor template carrying a normal repeat size of (G₄C₂)₂ to create an isogenic iMN line, and thereby reduced the abundance of RNA foci, poly-dipeptides and methylation levels of *C9orf72*⁶⁴. Both studies elegantly documented that excision of the HRE using CRISPR/Cas9 rescued of components of C9ORF72 pathology, but both employed transfection to edit the patient cells^{63,64}. By contrast, we have developed a method for AAV-mediated genome editing in patient-derived iPSC model systems and an AAV-mediated delivery system that we tested in multiple C9BAC transgenic models including those without endogenous Cas9-expression. Efforts by others using Cas9-expressing mice are less comprehensive and less relevant to developing patient therapies⁶⁵. Moreover, our study confirmed suppression of poly-dipeptide expression using assays for both poly-GP and poly-GR.

We found that excision of the HRE may rescue *C9ORF72* haploinsufficiency, a widely documented aspect of C9-ALS/FTD pathology^{49,54,66–68}. Because there are multiple proposed roles for *C9ORF72* in a range of cell types, it has reasonably been proposed that a reduction in *C9ORF72* protein levels is deleterious. Such functions for *C9ORF72* include: membrane trafficking in autophagy⁶⁹, lysosomes⁷⁰, and endosomes⁶⁹; immune homeostasis in microglia and macrophages^{71,72}, and assembly of stress granules.⁷³ Consistent with these observations, *C9orf72* knock-out mice exhibit immune dysfunction and weight loss, but no motor neuron degeneration or impairment^{71,72,74}, suggesting that haploinsufficiency alone does not cause a C9-ALS/FTD phenotype. Similarly, *C9ORF72* BAC-transgenic mice expressing a transgene with expanded HRE while also expressing endogenous *c9orf72*, do not develop a motor phenotype^{75–77}. It is striking that in *C9orf72* knockout mice that also express an expanded HRE, gain-of-function pathology is exacerbated¹¹, resulting in a motor phenotype^{11,78}. Despite strong evidence for the role of haploinsufficiency in *C9ORF72* pathology, most published therapeutic approaches for C9-ALS/FTD, such as microRNA^{55,79} antisense oligonucleotide strategies^{80,81}, and antibody therapies⁸² focus on reducing HRE gain-of-function but not rescue of haploinsufficiency. Our present study is, to our knowledge, the first to demonstrate an *in vivo* therapy delivering gRNAs as well as Cas9, that ameliorates both toxic acquired functions of mutant HRE (RNA foci, poly-dipeptides) and *C9ORF72* haploinsufficiency.

AAV has many advantages and is presently regarded as the safest and most effective gene therapy vector for the CNS, with potentially a long-lived therapeutic effect following a single treatment. Currently, three AAV-based therapies are approved by the FDA for the treatment of lipoprotein lipase deficiency (Glybera), Leber congenital amaurosis (Luxturna) and spinal muscular atrophy (SMA) (Zolgensma). Despite this, in this study we employed AAV as merely a delivery tool for a

proof-of-mechanism. A non-viral mRNA mediated approach would be better suited for clinical translation as AAV promotes long-term expression, which is unnecessary for CRISPR/Cas9-mediated editing in the CNS. Furthermore, long-term expression of an active genome editing system with AAV present safety concerns, such as: 1). increased genomic instability and off-target editing due to the sustained activity of an active nuclease ⁸³; 2). immune responses triggered by long-term expression of a bacterial protein (Cas9) ⁸⁴; 3). high AAV genome integration levels at the CRISPR/Cas9-induced double strand break ⁶⁰, which we also detected in our *in vivo* experiments (Fig. 4d). Although some of these issues can be addressed by using self-inactivating AAV vectors ⁸⁵ or tuneable AAV vectors ⁸⁶, self-integration of AAV genome fragments remains a concern. Delivery of Cas9 and gRNA mRNA by lipid nanoparticles is currently tested in a clinical trial for the treatment of transthyretin amyloidosis. Encouraging preliminary data shows high levels of on-target editing and no serious adverse events ⁸⁷. Safety profiles, delivery efficiencies and rapid advancements in lipid nanoparticle development make it a preferred choice to deliver Cas9 and gRNAs to excise the HRE in C9ORF72.

CRISPR technology has expanded the reach of gene editing; formerly only a versatile research tool, this technology is becoming a powerful method for the treatment of human genetic disease. Here, we have demonstrated that gRNA2,3 and gRNA2,4 in combination with SpCas9 are capable of excising the ALS/FTD-causing HRE in C9ORF72 in three mouse models, primary neurons, and patient cells. This excision attenuated the three major C9-ALS/FTD adverse, pathologic elements ascribed to the expanded repeats in C9ORF72: toxic poly-dipeptides, RNA foci, and haploinsufficiency. These proof-of-concept findings support the view that CRISPR methodologies should be explored as an excellent strategy to halt or slow disease progression in ALS, FTD and ALS/FTD patients, and potentially other disorders caused by expanded domains of repeated DNA.

4. Materials and Methods

gRNA design and cloning

The CRISPRseek Bioconductor package⁸⁸ was used to identify CRISPR/Cas9 target sites flanking the HRE in C9orf72 intron 1 and intron 2 (Fig. 1a). Two gRNAs from each intron with low predicted off-target activity⁸⁹ were selected for further analysis. To confirm high activity at these target sites, guide sequences were cloned into the pX330 sgRNA expression plasmid⁹⁰ (Addgene plasmid 42230), and the full target site was cloned into pM427, a GFP reporter plasmid for target site cleavage⁹¹. High nuclease activity was confirmed by induction of GFP-positive cells following co-transfection of these plasmids into HEK293T cells (data not shown). Next, the guides were cloned in pairs into the 3' end of a CB-GFP encoding AAV proviral plasmid, each under the control of a U6 promoter. Two plasmids pAAV-CB-GFP-C9gR-2,3 and pAAV-CB-GFP-C9gR-2,4 (most efficient at editing HEK293T cells) were packaged into rAAV9 vectors. AAV9-pU1a-SpCas9-RBGpA vector was used to deliver and express SpCas9 (Addgene, 121507). To measure InDels in C9ORF72 after editing in HEK293T cells, cells were harvested 72 hr after transfection and genomic DNA was extracted (Gentra Puregene Tissue Kit, Qiagen). The C9ORF72 locus was amplified by PCR, Sanger-sequenced (Genewiz), and analyzed by the ICE web-based interface (Synthego). InDels up to 20bp were permitted. To measure off-target effects in predicted targets, loci were amplified by PCR and sequenced by CRISPR Amplicon Sequencing (MGH, USA). Next, paired-end amplicon sequencing datasets of potential off-target editing sites were analyzed using CRISPResso2⁹². FASTQ files were trimmed for 3' adapters and low-quality base calls were analyzed in Batch mode with r1 and r2 inputs. "Guide" parameter (-g) specified the sgRNA binding site projected by the *in silico* analysis, not the near-cognate on-target guide sequence. Cleavage offset was set to -3. Editing quantification ignored 15bp from read ends, ignoring substitutions (--

ignore_substitutions) to exclude erroneous base calls. The AMACR locus was analyzed using 150bp, single-end reads (-r1 flag only) in the manner described above.

Animal experiments

All mouse experiments were conducted at UMass Medical School following protocols approved by the Institutional Review Board. The University of Massachusetts Medical School Institutional Animal Care and Use Committee approved all experiments involving animals. Three transgenic mouse models were used in this work, each harboring the human *C9ORF72* HRE and varying degrees of flanking sequence and copy numbers: BAC111, BAC112 (Dr. Robert Baloh, Novartis, Basel, formerly Cedars Sinai, Los Angeles), and C9-500 (Dr. Laura Ranum, U Florida Genetics Institute, Gainesville). BAC111 mice were generated in our laboratory and express approximately 600 repeat motifs within a truncated human *C9ORF72* gene (from exons 1-6) and heterozygous mice express 6-8 copies and are described in greater details in Peters et. al, 2015⁴⁵. BAC112 mice express human *C9ORF72* transgene with 550 repeat motifs and heterozygous mice have 16-20 copies of the transgene⁹³. The C9-500 *C9ORF72* BAC transgenic mouse model also expresses the whole human *C9ORF72* gene and heterozygous mice have one copy of the gene, with an expansion of 500 repeats⁴⁷. For all experiments, heterozygous *C9ORF72* BAC transgenic mice were used. These were generated by crossing homozygous *C9ORF72* mice with WT mice. To generate BAC111/Cas9 mice, homozygous BAC111 mice were crossed with homozygous B6J.129(Cg)-*Gt(Rosa)26Sor^{tm1.1}(CAG-cas9*,-EGFP)^{Fezh}/J* (obtained from the Jackson laboratory) resulting in heterozygosity of both BAC111 *C9ORF72* and Cas9.

Stereotaxic striatal brain injections

Adult male mice between ages of 8 to 16 weeks were anesthetized using Ketamine/Xylazine. Mice were then restrained on 'Just for mouse stereotaxic instrument' (Stoelting). Two holes were drilled

bilaterally at coordinates X= +/-2mm Y=+1mm, Z= -2.5mm, applying Bregma as a point of origin. Using a 10ul Hamilton Syringe 10µl, gastight model, 1701 RN and a removable 33-gauge Small Hub Needle (Hamilton Company), 5 µL were injected on each side at a rate of 0.5µL/min. Mice were injected with 2×10^{10} vector genomes (VG) of AAV9-Cas9 and 7×10^9 VP of the gRNA vectors or Rosa control per striatal side, which is 3.8×10^{10} VG in total for both striata per mouse and 1.65×10^{12} VG/kg. Animals were then administered three doses of pain medication Ketophen and Buprenorphine: directly after waking up from anesthesia and at 24- and 48-hours post-surgery. Tissues were harvested eight weeks after injection for use in DNA, RNA, and protein analysis. Prior to harvest, mice were subjected to cardiac perfusion with phosphate buffered solution (PBS). 3mm striatal punches were freshly taken using a disposable Biopsy punch with plunger (Integra) and then flash-frozen in liquid nitrogen and stored at -80°C. For FISH, brains were frozen in Optimal Cutting Temperature (OCT) compound and cryosectioned.

Cell culture

For plasmid screening, HEK293T cells were seeded (1.2×10^5 cells) in a 24-well plate, maintained with DMEM media (11995-073; Gibco), 10% Fetal Bovine Serum FBS (F-0926 Sigma-Aldrich) and 1% Penicillin- Streptomycin (30-001-CI; Corning). The cells were transfected after 24h, using jetPRIME Reagent (Polyplus) and 1µg of each plasmid DNA. Forty-eight hours post-transfection, cells were rinsed with PBS and collected for further analyses.

For primary neuron culture, homozygous C9ORF72 mice were crossed to wild type mice to produce heterozygous progeny. Embryos were harvested at day 16 with cortices subsequently dissected in 4°C Hank's Buffered Salt Solution (HBSS, Cellgro) under dissection microscope. The cortices were stored in HBSS until processing. Cells were then dissociated for 15 min in 3.5ml of 0.05% Trypsin (Invitrogen), with gentle inversion at 7.5 min intervals. After 15 mins, Trypsin

dissociation was inactivated by addition of 7.5ml of Dulbecco's Modified Eagle Medium (DMEM, Gibco-ThermoFisher) + 10% Fetal Bovine Serum (FBS) and homogenized via pipetting. This mixture was then added to a new 50ml falcon tube through a 40um filter. An additional 10ml of DMEM + 10% FBS was added through the filter to collect any remaining cells. Cells were spun down via centrifugation, and the pelleted cells resuspended via pipetting in 20ml of pre-warmed (37°C) Neurobasal medium (Gibco-Thermofisher). After counting, Neurobasal was added to desired seeding concentration and supplemented with 20ul/ml B27 (Invitrogen), 10ul/ml Pen/Strep (Invitrogen), and 10ul/ml Glutamax (Invitrogen). The cells were seeded on 6-well plates (Falcon/Olympus) at 1,000,000 cells/well in a volume of 2ml per well. Cells were also seeded on 4 chamber CC2 glass slides (LAB-TEK) at 300,000 cells/chamber in a volume of 800ul per chamber. Both the plates and chamber slides with slightly swirled immediately after seeding in order to prevent clumping. The plates were pre-coated with poly-d-Lysine (Invitrogen) at room temperature overnight, washed three times with ddH₂O and dried before seeding. Cells were treated on DIV 4 post-plating by removing half of the media and replacing it with new media (pre-warmed and supplemented with B27, Pen/Strep, and Glutamax at previously described concentrations) containing the experimental vectors (AAV9-Cas9: MOI:30.000, AAV9-gRNA2,3 and AAV9-gRNA2,4 vector: MOI:20.000, with a total MOI:50.000). On day 10 the cells were harvested for additional experiments or fixed for FISH staining.

iPSC cell lines and motor neuron differentiation

Patient-derived iPSCs were generated and differentiated into motor neurons as previously described⁹⁴. Briefly, iPSCs were maintained in Matrigel-coated flasks and mTeSR® medium (StemCell Technologies). For neural induction, iPSCs were dissociated and transferred to low attachment (poly-HEMA) flasks and cultured in suspension with neuroinduction medium for 21

days. Neurospheres were then seeded onto poly-L-ornithine/Laminin (PLO) coated plates to promote attachment and maintained in neuronal maturation medium for an additional 21 days.

AAV transduction of C9ORF72 iPSCs

Matrigel coated 12-well plates were seeded with undifferentiated iPSCs (3.5×10^5 cells/well). The following day, cells were co-transduced in a volume of 250 μ L/well with AAV9-SpCas9 and AAV9-gRNA2,4 vectors at a multiplicity of 3×10^5 and 1×10^5 AAV genome copies per cell, respectively. After six hours, an additional 1mL of medium was added to each well without removing viral vectors and maintained for three days at 37°C and 5% CO₂ in a humidified incubator. Transduced iPSCs were collected using gentle dissociated reagent and seeded at low density in Matrigel-coated 6-well plates. After three days, single colonies were manually selected, expanded, and analyzed for genome editing. To ensure clonality of iPSC lines, cultures were sub-cloned by titration and manual selection.

AAV transduction of differentiated C9ORF72 iPSC motor neurons

After 42 days of neurodifferentiation, iPSC motor neurons were transferred to 12-well PLO coated plates (3.5×10^5 cells/well) and allowed to recover for seven days. Neuronal cultures were co-transduced in a volume of 250 μ L/well with AAV9-SpCas9 and AAV9-gRNA2,3 or AAV9-gRNA2,4 vectors at a multiplicity of 3×10^5 and 1×10^5 AAV genome copies per cell, respectively. After six hours, an additional 1mL of medium was added to each well without removing viral vectors and maintained for seven days at 37°C and 5% CO₂ in a humidified incubator.

Generation and AAV transduction of brain organoids

Cultures of C9ORF72 iPSCs were dissociated, transferred to low attachment flasks, and maintained in brain organoid induction medium (1X DMEM F12, 1X N2 Supplement, 10 mg/ml

Heparin, 1X Penicillin/Streptomycin, 1X MEM Non-Essential Amino Acids, 1X Glutamax, 1 μ M CHIR99021, 1 μ M SB-431542, and 100 μ g/ml Primocin) similar to previously reported methods (PMID 27118425). After 5 or 17 days, embryoid bodies were transduced with 15 μ L of AAV9-Cas9 (5.5×10^{12} genome copies/mL) and 5 μ L AAV9-gRNA2,4 (8.5×10^{12} genomes/mL) in a volume of 500 μ L for 30 minutes. An additional 3mLs of medium used to suspend and transfer embryoid bodies to a 12-well plate. On day 32, organoids were harvested by centrifugation and genomic DNA was isolated for analysis of genome editing.

Droplet Digital PCR (ddPCR)

Droplet Digital PCR (ddPCR) analysis was conducted on samples from *in vitro* (primary neurons) and *in vivo* (mouse striatal injection) experimentation. Frozen tissue samples were homogenized using a TissueLyser II (Qiagen) before a total RNA extraction with Trizol (Life Technologies, USA) in accordance with the TriZol reagent user guide protocol (Invitrogen). For *in vitro* samples, 1ml of Trizol was directly added to frozen harvested primary neurons, homogenized via pipetting, with RNA subsequently extracted following the previously described TriZol reagent protocol. Reverse transcription to cDNA was performed using a High capacity RNA-to-cDNA Kit (Applied Biosystems-Thermofisher) in accordance with their quick reference guide protocol. ddPCR was then performed on the cDNA to assess the absolute concentration of *C9ORF72* transcripts. Master mixes for ddPCR were generated containing 11 μ L/sample of ddPCR Supermix (no dUTP) (Bio-Rad, USA), 1.1 μ L/sample of *C9ORF72* transcript-specific FAM labeled probe, 1.1 μ L/sample of TATA-box Binding Protein (TBP) specific HEX labeled probe as a housekeeping probe, and 8.8 μ L of cDNA diluted in ddH₂O. Sample droplets were generated by loading 20 μ L of sample and 70 μ L of ddPCR oil in to a DG8 cartridge (Bio-Rad, USA) and placing the cartridge in a QX100 droplet generator (Bio-Rad, USA). 40 μ L of the generated droplets were added to each well of a ddPCR

plate and placed in a PCR thermocycler under the following conditions: 95°C for 10mins, 40 cycles of {*95°C for 30 sec, *55°C for 1 min}, *98°C for 10 mins, hold at 4°C. (*) denotes a 30% RAMP. Following PCR, the ddPCR plate was read in a QX100 Droplet Digital PCR reader (Bio-Rad, USA) and analyzed using QuantaSoft software.

Western Blot

Radioimmunoprecipitation Assay (RIPA) buffer and cOmplete mini-proteinase inhibitors (1 tablet per 10 ml extraction solution, Roche) was used to homogenize striatal tissue and cells. Equal amounts of protein lysates, as measured by BCA Assay (Thermo Scientific, USA), were run on Novex 12% Tris-Glycine gels (Life Technologies, USA) using Tris-Glycine SDS running buffer (Invitrogen, USA). Protein was transferred to nitrocellulose membranes using an i-Blot transfer device (Invitrogen, USA). Membranes were blocked for 1 hour at room temperature with Odyssey Blocking Buffer (LiCor, USA) before being probed overnight with primary antibodies against C9ORF72 (1:1000, 22637-1-AP, Thermo Scientific, USA). Infrared labeled secondary antibodies were applied, and blots visualized. The band intensity was quantified using the Odyssey Infrared imaging system (LiCor, USA).

Poly-GP and Poly-GR immunoassay

Both primary neuron and mouse tissues were lysed in Radioimmunoprecipitation Assay (RIPA) buffer and a cOmplete protease inhibitor (Roche). Tissues were lysed using a hand lyser. Then, samples were centrifuged at 16,000 x g for 20 min and supernatant collected. Total protein concentration of lysates was determined by a BCA assay (Thermo Scientific, USA). Poly-GP and Poly-GR levels in lysates were measured using a previously described sandwich immunoassay that utilizes Meso Scale Discovery (MSD) electrochemiluminescence detection technology^{37,48}. Lysates were diluted in tris-buffered saline (TBS) and tested using equal concentrations for each

sample in duplicate wells. Response values corresponding to the intensity of emitted light upon electrochemical stimulation of the assay plate using the MSD QUICKPLEX SQ120 were acquired and background-corrected, using the average response from non-transgenic mice.

RNA fluorescence in situ hybridization (FISH)

Primary cortical neurons grown on 4-chamber CC2 glass slides (LAB-TEK) or cryosections (20µm) were fixed in 4% PFA for 20 min. Samples were then washed with DEPC-treated PBS containing 0.1% Tween-20 for 20 min five times and prehybridized for 1 hour at 55°C in 2x DEPC-treated saline sodium citrate (SSC) containing 40% formamide, 0.1% Tween-20, 50% of dextran sulfate, 50µg/ml heparin and 1mg/ml heat-denatured salmon sperm. Samples were hybridized overnight at 55°C in pre-hybridization buffer containing a 1:500 dilution of 1µM sense GGCCCC₄ DNA probe with Cy3 5' end tag. After hybridization, the samples were washed in a pre-warmed mix 1:1 of hybridization buffer and 2X SSC for 30 min at 55°C. Followed by 2 washes in pre-warmed 2X SSC for 30 min at 55°C, 2 washes with 0.2 X SSC for 30 min at 55°C, and 2 washes in 1X PBS+ 0.1% Tween-20 for 30 min, at room temperature. Nuclei were stained by 4',6-diamidino-2-phenylindole (DAPI, 1:10,000) for 5 min, and were then incubated in 0.5% Sudan Black in 70% ethanol for 7 min to reduce autofluorescence. Lab Vision PermaFluor (Thermo Scientific Shandon, USA) was used to mount glass coverslips. For quantification, z-stacks of the cultured primary neurons and external pyramidal layer of brain tissues were acquired by a Leica DM5500 B microscope with a Leica DFC365 FX camera. A total magnification of 100.8X (63X oil lens * 1.6X camera) was used to generate the images, which were maximally projected, and foci were counted manually in +300 cells per slide.

RNAscope

RNAscope *in situ* hybridization (*ish*) Fluorescent Multiplex Assay was performed as instructed by Advanced Cell Diagnostics (ACD). Brains were embedded in OCT, frozen and sectioned at 20µm. Sections were fixed in 4% PFA for 1 hour and dehydrated in 50% ethanol (5 min), 70% ethanol (5 min) and 100% ethanol (10 min) at room temperature. The slides were air dried and protease IV reagent was added to each section for 30 min. Slides were washed in PBS at room temperature and treated with a mixture of Channel 1 (GFP, 409011, ACD, USA) and Channel 3 probes (SpCas9, 533931-C3, ACD, USA) at a ratio of 50:1 dilution for two hours at 40°C. Following probe incubation, the slides were washed two times in 1X RNAscope wash buffer and submerged in AMP-1 reagent for 30 minutes at 40°C. Washes and amplification were repeated using AMP-2, AMP-3 and AMP-4 ALT A (channel 1 = 488nm, channel 3 = 647nm) reagents with a 15-min, 30-min, and 15-min incubation period, respectively. Sections were counterstained with DAPI for 30 sec at room temperature and cover-slipped using Lab Vision PermaFluor (Thermo Scientific Shandon, USA). Pictures of the staining were acquired by a Leica DM5500 B microscope with a Leica DFC365 FX camera. A total magnification of 100.8X (63X oil lens * 1.6X camera) was used to generate the images.

PacBio No-Amp

Genomic DNA was extracted using the Monarch® HMW DNA Extraction Kit (New England Biolabs, T3050S and T3060S) using manufacture's protocol for iPSC samples, and a modified manufacture's protocol for striatal tissues. Shortly, DNA was extracted from striatal tissues using the 'low input' protocol. Tissues were homogenized by pestle in 300 µl lysis buffer and 10 µl Proteinase K and then incubated at 56°C for 10 minutes in a thermal mixer with 1400 RPM agitation and 35 minutes without agitation. Afterwards, an additional 5 µl Proteinase K was added

to the samples and incubated for 45 minutes without agitation. 5 µl RNase A was added and mixed by inverting 10 times and incubated for 10 minutes at 56°C at 1400 RPM. Then, samples were centrifuged at maximum speed for 3 minutes and supernatant was transferred to a new tube and incubated on ice for 5 minutes. Next, 150 µl of Protein Separation Solution was added and mixed by inverting for 1 minute. This was followed by maximum centrifugation at maximum speed for 20 minutes. The remainder of the DNA extraction was executed following the manufacture's protocol. After DNA extraction, striatal DNA was reprecipitated using 0.1 volume 3M Sodium acetate and 2.5 volume ice cold 100% Ethanol at -20°C overnight. The next day, samples were centrifuged at maximum speed at 4°C for 30 minutes and pellets were washed twice with 0.5ml ice cold 75% Ethanol, spinning at 4°C for 10 mins each time and resuspended in EB.

PacBio No-AMP was conducted following PacBio's 'No-Amp Targeted Sequencing Utilizing the CRISPR-Cas9 System' manual. After inspecting the quality of the genomic DNA using a Femto Pulse system and Nanodrop, the SMRTbell library was prepared as follows: To reduce the amount of off-target molecules in the final SMRTbell library, the DNA (3 µg per sample) was dephosphorylated to prevent fragment ends from participating in the ligation reaction following the CRISPR-Cas9 Digestion step. Then, the dephosphorylated DNA was digested with Cas9 and gRNAs adjacent to the region of interest (as published in the manual, C9orf72.DC.1: TTGGTATTTAGAAAGGTGGT; and C9orf72.DC.2: GGAAGAAAGAATTGCAATTA), after which the DNA was purified using AMPure PB beads. Next, a capture adapter was ligated to the blunt ends and captured using AMPure PB beads. Then, failed ligation products and gDNA fragments were removed with a nuclease treatment. Finally, an additional treatment of the SMRTbell library with trypsin was used to remove enzymes during the AMPure PB bead purification step.

For sequence preparation, a standard PacBio sequencing primer was annealed to the enriched SMRTbell templates, and a Sequel II DNA polymerase 2.2 was bound to the primer-annealed SMRTbell templates. This complex was purified with AMPure PB beads and loaded on to a SMRT Cell and sequenced on a Sequel II System using an immobilization time of 4h and movie time of 30h.

HiFi long reads were produced using the standard demultiplexing and circular consensus tools in SMRT Link v10.2, with the exception that the "--all" input option was specified during circular consensus such that a read was output for every ZMW. For each sample, the reads were then aligned to human reference GRCh38 using the pbmm2 aligner and off-target reads were removed. Reads spanning the targeted repeat region were assessed for the presence or absence of the inserted repeat sequence in one of two ways based on whether or not they spanned the entirety of the repeat region reference coordinates (chr9: 27,573,485 - 27,573,546). For reads spanning the region, repeat lengths were determined based on the number of inserted bases in the region (via parsing CIGAR strings and extracting the corresponding read sequence). For reads only anchored on one side, and thus are soft-clipped in the repeat region, a manual inspection of the clipped sequence was performed to identify whether or not it contained the inserted repeat sequence. Editing efficiencies were then determined from the fraction of reads without the inserted repeat sequence.

Table 1. PacBio No-Amp number of reads

	iPSC		C9-500 Striatum		
	C9 control	gRNA2,4 11-11	PBS	gRNA2,3	gRNA2,4
#1	180	140	59	67	59
#2			90	56	80
#3			80	54	59

UDiTaS library preparation and analysis

Transposome was assembled according to published protocol (Giannoukos et al, 2018⁴³). Briefly, oligonucleotides containing unique molecular identifiers (UMIs) and i5 barcodes obtained from IDT were annealed and loaded into purified TN5. Library preparation was performed using a nested PCR approach as described previously in Ibraheim et al, 2021⁸⁵. Briefly, 1 µL of transposome was used to tagment 200 ng of genomic DNA at 55°C for 7 minutes. After tagmentation, 1 µL of 0.2% SDS was added to quench any remaining enzyme. The tagmented genomic DNA was used for the first round of PCR for library preparation with a common primer (located within the transposon) and a C9ORF72 specific nested primer for 10 cycles. 0.9x AMPure beads were used to purify the products. All of the products from the initial PCR were used in the second round of PCR with common primer and internal C9ORF72 primer for 15 cycles. 0.9x AMPure beads were used to clean up the products for the third round of PCR which used UDiTaS_i5_common and i7 barcode primers for 20 cycles. A final purification was performed using 0.5x/0.35x for double size selection to yield a size distribution of ~200-1000 bp in the final library. The libraries were sequenced on an Illumina MiniSeq for 2 x 150 cycles.

The analysis was performed using the UDiTaS v1.0 workflow described in Giannoukos et al, 2018⁴³ with modifications to allow compatibility with the MiniSeq platform⁸⁵. Paired end libraries were demultiplexed with bcl2fastq (Illumina), masking UMIs (12 nts 1-9). The un-demultiplexed UMIs were stored in a single FASTQ masking R1, I1, I2 nts 10-17. The demultiplexed R1 and R2 FASTQs were paired via UMIs with fastq-pair hash lookup tool⁹⁵. Gzipped FASTQs were added to individual UDiTaS compatible sample directories. Reference FASTA and 2-bit files were obtained from UCSC table browser⁹⁶. The analysis⁴³ was performed inside umasstr/UDiTaS docker container, skipping the demultiplexing step.

Table 2. UDiTaS number of reads

	HEK293T				iPSC		C9-500 Striatum			
	PBS	ROSA	gRNA2,3	gRNA2,4	C9 control	gRNA2,4 11-11	PBS	ROSA	gRNA2,3	gRNA2,4
#1	2220	318	8149	2999	34192	76114	18327	16670	47943	70315
#2	863	1142	4819	5697	45784	78449	29500	19244	41089	67761
#3	405	616	4575	3335	41522	16100	31847	25003	56987	67457
#4	448	628	8201	3651			30028	20677	31497	48168

Table 3. Primers used for UDiTaS. The red sequence represents 5p index sequence and the blue sequence represents the UMI.

3742_UDiTaS_i5_ common	AATGATACGGCGACCACCGAGATCTACAC
3895_Uditas_g2Nested_C9ORF72	TCTCTCCCCACTACTTGCTCTCACAGTACTC
3743_C9ORF72_gRNA2_i7	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGCTGAGGGTGA ACAAGAAAAGACCTG
UDiTaS adapter top i501	AATGATACGGCGACCACCGAGATCTACACTATAGCCTNNWNNWN NNTCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
UDiTaS adapter top i502	AATGATACGGCGACCACCGAGATCTACACATAGAGGCNNWNNWN NNTCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
UDiTaS adapter top i504	AATGATACGGCGACCACCGAGATCTACACGGCTCTGANNWNNWN NNTCGTCGGCAGCGTCAGATGTGTATAAGAGACAG

5. Acknowledgements

We thank the Mutagenesis Core Facility at UMASS Medical School for the design of gRNAs and *in silico* off-target effects analyses. We further thank G. Gao, J. Xie, and Q. Su at UMass Medical School who supervised the vector production and D. Wang who provided us with the SpCas9 plasmid. Furthermore, we are grateful to Eric Wieben and the Genome Analysis Core at the Mayo Clinic for performing the PacBio No-AMP experiments. Special thanks to Melany Hirsch for her assistance with the experiments. R.H.B. and C.M. received support from NINDS NS088689. Additionally, R.H.B. was funded by R01 NS111990 and R01 NS104022, the ALS Association, the Angel Fund for ALS Research, the Pierre L. de Bourgknecht ALS Research Foundation, ALS Finding A Cure, ALSOne, the Cellucci Fund, and the Max Rosenfeld ALS Research Fund. R37NS057553 and R01NS101986 and the Target ALS Foundation grant were rewarded to F.B.G.,

6. Author contributions

K.E.M. led design, analysis, interpretation of the study, and prepared the manuscript. C.M. conceived of the project and supervised all aspects of its execution and analysis. Testing of individual guideRNAs in HEK293T cells was executed by H.N. Cloning and prepping of constructs were done by K.E.M, H.N., A.A, M.B. PCRs were performed by K.E.M., H.N., A.A. Primary neuron experiments, mouse surgeries and FISH staining were executed by K.E.M., A.A., N.P.F. Experiments involving transcript levels were performed by K.E.M., N.P.F., M.T.B. Western blots were done by K.E.M. Poly-dipeptide levels were detected by K.E.M., G.K., T.F.G. iPSC, iMN and brain organoid experiments were executed by Z.Z., R.E., N.S.A., M.J.R., and M.R. InDel and off-target analyses were performed by K.E.M., A.E., T.R. PacBio No-Amp analyses were executed by Z.D.S and K.E.M. UDiTaS analyses were performed by T.R., C.K. and K.E.M.

Lastly, R.H.B., Z.Z., E.J.S., F.B.G. and C.M. have aided with scientific input, and N.P.F, Z.Z., R.H.B., F.B.G. M.B., E.J.S. have aided with editing.

7. Competing interests

C.M. is an inventor on the patent for this work and may be entitled to royalty payments in the future.

8. References

1. Wijesekera, L. C. & Leigh, P. N. Amyotrophic lateral sclerosis. *Orphanet J. Rare Dis.* **4**, 3 (2009).
2. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **72**, 245–256 (2011).
3. Majounie, E. *et al.* Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* **11**, 323–30 (2012).
4. Renton, A. E. *et al.* A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **72**, 257–268 (2011).
5. Van Mossevelde, S., van der Zee, J., Cruts, M. & Van Broeckhoven, C. Relationship between C9orf72 repeat size and clinical phenotype. *Current Opinion in Genetics and Development* vol. 44 117–124 (2017).
6. Iacoangeli, A. *et al.* C9orf72 intermediate expansions of 24–30 repeats are associated with ALS. *Acta Neuropathol. Commun.* **7**, 115 (2019).
7. Devenney, E. *et al.* Frontotemporal Dementia Associated With the C9ORF72 Mutation: A Unique Clinical Profile. *JAMA Neurol.* **71**, 331–339 (2014).
8. Nowicka, N., Juranek, J., Juranek, J. K. & Wojtkiewicz, J. Risk factors and emerging therapies in amyotrophic lateral sclerosis. *International Journal of Molecular Sciences* vol. 20 (2019).

9. Amick, J. & Ferguson, S. M. C9orf72: At the intersection of lysosome cell biology and neurodegenerative disease. *Traffic* vol. 18 267–276 (2017).
10. Shi, Y. *et al.* Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat. Med.* **24**, 313–325 (2018).
11. Zhu, Q. *et al.* Reduced C9ORF72 function exacerbates gain of toxicity from ALS/FTD-causing repeat expansion in C9orf72. *Nat. Neurosci.* **23**, 615–624 (2020).
12. Donnelly, C. J. *et al.* RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. *Neuron* **80**, 415–428 (2013).
13. Lagier-Tourenne, C. *et al.* Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc. Natl. Acad. Sci. U. S. A.* **110**, (2013).
14. Mizielinska, S. *et al.* C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol. (Berl.)* **126**, 845–857 (2013).
15. Mori, K. *et al.* HnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. *Acta Neuropathol. (Berl.)* **125**, 413–423 (2013).
16. Sareen, D. *et al.* Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci. Transl. Med.* **5**, 208ra149 (2013).
17. Zu, T. *et al.* RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc. Natl. Acad. Sci.* **110**, E4968–E4977 (2013).
18. Gendron, T. F., Belzil, V. V., Zhang, Y. J. & Petrucelli, L. Mechanisms of toxicity in C9FTLD/ALS. *Acta Neuropathologica* vol. 127 359–376 (2014).
19. Česník, A. B. *et al.* Nuclear RNA foci from C9ORF72 expansion mutation form paraspeckle-like bodies. *J. Cell Sci.* **132**, (2019).
20. Cleary, J. D. & Ranum, L. P. W. Repeat associated non-ATG (RAN) translation: new starts in microsatellite expansion disorders. *Curr. Opin. Genet. Dev.* **26**, 6–15 (2014).
21. Fry, M. & Usdin, Karen. *Human nucleotide expansion disorders*. (Springer, 2006).

22. Orr, H. T. & Zoghbi, H. Y. Trinucleotide Repeat Disorders. *Annu. Rev. Neurosci.* **30**, 575–621 (2007).
23. Zu, T. *et al.* Non-ATG-initiated translation directed by microsatellite expansions. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 260–265 (2011).
24. Gendron, T. F. *et al.* Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol. (Berl.)* **126**, 829–844 (2013).
25. Ash, P. E. A. *et al.* Unconventional Translation of C9ORF72 GGGGCC Expansion Generates Insoluble Polypeptides Specific to c9FTD/ALS. *Neuron* **77**, 639–646 (2013).
26. Mori, K. *et al.* The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTL/ALS. *Science* **339**, 1335–1338 (2013).
27. May, S. *et al.* C9orf72 FTL/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathol. (Berl.)* **128**, 485–503 (2014).
28. Wen, X. *et al.* Antisense Proline-Arginine RAN Dipeptides Linked to C9ORF72-ALS/FTD Form Toxic Nuclear Aggregates that Initiate In Vitro and In Vivo Neuronal Death. *Neuron* **84**, 1213–1225 (2014).
29. Sun, Y., Eshov, A., Zhou, J., Isiktas, A. U. & Guo, J. U. C9orf72 arginine-rich dipeptide repeats inhibit UPF1-mediated RNA decay via translational repression. *Nat. Commun.* **11**, 3354 (2020).
30. Kwon, I. *et al.* Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* **345**, 1139–1145 (2014).
31. Mizielinska, S. *et al.* C9orf72 repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science* 1192–1194 (2014) doi:10.1126/science.1256800.
32. West, R. J. H. *et al.* Co-expression of C9orf72 related dipeptide-repeats over 1000 repeat units reveals age- and combination-specific phenotypic profiles in *Drosophila*. *Acta Neuropathol. Commun.* **8**, 158 (2020).
33. Zhang, Y.-J. *et al.* Heterochromatin anomalies and double-stranded RNA accumulation underlie C9orf72 poly(PR) toxicity. *Science* **363**, (2019).

34. Zhang, Y.-J. *et al.* C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nat. Neurosci.* **19**, 668–677 (2016).
35. Zhang, Y.-J. *et al.* Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. *Nat. Med.* **24**, 1136–1142 (2018).
36. Yang, D. *et al.* FTD/ALS-associated poly(GR) protein impairs the Notch pathway and is recruited by poly(GA) into cytoplasmic inclusions. *Acta Neuropathol. (Berl.)* **130**, 525–535 (2015).
37. Choi, S. Y. *et al.* C9ORF72-ALS/FTD-associated poly(GR) binds Atp5a1 and compromises mitochondrial function in vivo. *Nat. Neurosci.* **22**, 851–862 (2019).
38. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
39. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).
40. Single Ascending Dose Study in Participants With LCA10 - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT03872479>.
41. A Safety and Efficacy Study Evaluating CTX001 in Subjects With Severe Sickle Cell Disease - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT03745287>.
42. Intellia Therapeutics. *Phase 1 Two-Part (Open-label, Single Ascending Dose (Part 1) and Open-label, Single Dose Expansion (Part 2)) Study to Evaluate Safety, Tolerability, Pharmacokinetics, and Pharmacodynamics of NTLA-2001 in Patients With Hereditary Transthyretin Amyloidosis With Polyneuropathy (ATTRv-PN)*. <https://clinicaltrials.gov/ct2/show/NCT04601051> (2020).
43. Giannoukos, G. *et al.* UDiTaSTM, a genome editing detection method for indels and genome rearrangements. *BMC Genomics* **19**, 212 (2018).
44. Zhang, X.-H., Tee, L. Y., Wang, X.-G., Huang, Q.-S. & Yang, S.-H. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Mol. Ther. Nucleic Acids* **4**, e264 (2015).

45. Peters, O. M. *et al.* Human C9ORF72 Hexanucleotide Expansion Reproduces RNA Foci and Dipeptide Repeat Proteins but Not Neurodegeneration in BAC Transgenic Mice. *Neuron* **88**, 902–909 (2015).
46. Leko, M. B. *et al.* Molecular Mechanisms of Neurodegeneration Related to C9orf72 Hexanucleotide Repeat Expansion. *Behav. Neurol.* **2019**, (2019).
47. Liu, Y. *et al.* C9orf72 BAC Mouse Model with Motor Deficits and Neurodegenerative Features of ALS/FTD. *Neuron* **90**, 521–534 (2016).
48. Su, Z. *et al.* Discovery of a Biomarker and Lead Small Molecules to Target r(GGGGCC)-Associated Defects in c9FTD/ALS. *Neuron* **83**, 1043–1050 (2014).
49. Jackson, J. L. *et al.* Elevated methylation levels, reduced expression levels, and frequent contractions in a clinical cohort of C9orf72 expansion carriers. *Mol. Neurodegener.* **15**, 7 (2020).
50. Xi, Z. *et al.* The C9orf72 repeat expansion itself is methylated in ALS and FTLD patients. *Acta Neuropathol. (Berl.)* **129**, 715–727 (2015).
51. Esanov, R. *et al.* A C9ORF72 BAC mouse model recapitulates key epigenetic perturbations of ALS/FTD. *Mol. Neurodegener.* **12**, 46–46 (2017).
52. Lee, Y.-B. *et al.* Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep.* **5**, 1178–1186 (2013).
53. Simón-Sánchez, J. *et al.* The clinical and pathological phenotype of C9ORF72 hexanucleotide repeat expansions. *Brain J. Neurol.* **135**, 723–735 (2012).
54. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **72**, 245–256 (2011).
55. Martier, R. *et al.* Artificial MicroRNAs Targeting C9orf72 Can Reduce Accumulation of Intra-nuclear Transcripts in ALS and FTD Patients. *Mol. Ther. Nucleic Acids* **14**, 593–608 (2019).
56. Hayes, L. R., Duan, L., Bowen, K., Kalab, P. & Rothstein, J. D. C9orf72 arginine-rich dipeptide repeat proteins disrupt karyopherin-mediated nuclear import. *eLife* **9**, e51685 (2020).

57. Moens, T. G. *et al.* C9orf72 arginine-rich dipeptide proteins interact with ribosomal proteins in vivo to induce a toxic translational arrest that is rescued by eIF1A. *Acta Neuropathol. (Berl.)* **137**, 487–500 (2019).
58. Andrade, N. S. *et al.* Dipeptide repeat proteins inhibit homology-directed DNA double strand break repair in C9ORF72 ALS/FTD. *Mol. Neurodegener.* **15**, 13 (2020).
59. Ebbert, M. T. W. *et al.* Long-read sequencing across the C9orf72 ‘GGGGCC’ repeat expansion: implications for clinical use and genetic discovery efforts in human disease. *Mol. Neurodegener.* **13**, 46 (2018).
60. Hanlon, K. S. *et al.* High levels of AAV vector integration into CRISPR-induced DNA breaks. *Nat. Commun.* **10**, 4439 (2019).
61. Esanov, R. *et al.* C9orf72 promoter hypermethylation is reduced while hydroxymethylation is acquired during reprogramming of ALS patient cells. *Exp. Neurol.* **277**, 171–177 (2016).
62. Qian, X. *et al.* Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. *Cell* **165**, 1238–1254 (2016).
63. Selvaraj, B. T. *et al.* C9ORF72 repeat expansion causes vulnerability of motor neurons to Ca²⁺-permeable AMPA receptor-mediated excitotoxicity. *Nat. Commun.* **9**, (2018).
64. Ababneh, N. A. *et al.* Correction of amyotrophic lateral sclerosis related phenotypes in induced pluripotent stem cell-derived motor neurons carrying a hexanucleotide expansion mutation in C9orf72 by CRISPR/Cas9 genome editing using homology-directed repair. *Hum. Mol. Genet.* **29**, 2200–2217 (2020).
65. Piao, X. *et al.* Dual-gRNA approach with limited off-target effect corrects C9ORF72 repeat expansion in vivo. *Sci. Rep.* **12**, 5672 (2022).
66. Gijselinck, I. *et al.* A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: A gene identification study. *Lancet Neurol.* **11**, 54–65 (2012).

67. van Blitterswijk, M. *et al.* Novel clinical associations with specific C9ORF72 transcripts in patients with repeat expansions in C9ORF72. *Acta Neuropathol. (Berl.)* **130**, 863–876 (2015).
68. Waite, A. J. *et al.* Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiol. Aging* **35**, 1779.e5-1779.e13 (2014).
69. Farg, M. A. *et al.* C9ORF72, implicated in amyotrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Hum. Mol. Genet.* **23**, 3579–3595 (2014).
70. Sullivan, P. M. *et al.* The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. *Acta Neuropathol. Commun.* **4**, 51 (2016).
71. Atanasio, A. *et al.* C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production and glomerulonephropathy in mice. *Sci. Rep.* **6**, 23204–23204 (2016).
72. O’Rourke, J. G. *et al.* C9orf72 is required for proper macrophage and microglial function in mice. *Science* **351**, 1324–1329 (2016).
73. Maharjan, N., Künzli, C., Buthey, K. & Saxena, S. C9ORF72 Regulates Stress Granule Formation and Its Deficiency Impairs Stress Granule Assembly, Hypersensitizing Cells to Stress. *Mol. Neurobiol.* **54**, 3062–3077 (2017).
74. Burberry, A. *et al.* Loss-of-function mutations in the C9ORF72 mouse ortholog cause fatal autoimmune disease. *Sci. Transl. Med.* **8**, 347ra93 (2016).
75. Peters, O. M. *et al.* Human C9ORF72 Hexanucleotide Expansion Reproduces RNA Foci and Dipeptide Repeat Proteins but Not Neurodegeneration in BAC Transgenic Mice. *Neuron* **88**, 902–909 (2015).
76. O’Rourke, J. G. *et al.* C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. *Neuron* **88**, 892–901 (2015).
77. 029099 - FVB/NJ-Tg(C9orf72)500Lpwr/J. <https://www.jax.org/strain/029099>.
78. Shao, Q. *et al.* C9orf72 deficiency promotes motor deficits of a C9ALS/FTD mouse model in a dose-dependent manner. *Acta Neuropathol. Commun.* **7**, (2019).

79. Martier, R. *et al.* Targeting RNA-Mediated Toxicity in C9orf72 ALS and/or FTD by RNAi-Based Gene Therapy. *Mol. Ther. Nucleic Acids* **16**, 26–37 (2019).
80. Donnelly, C. J. *et al.* RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. *Neuron* **80**, 415–428 (2013).
81. Jiang, J. *et al.* Gain of toxicity from ALS/FTD-linked repeat expansions in C9ORF72 is alleviated by antisense oligonucleotides targeting GGGGCC-containing RNAs. *Neuron* **90**, 535–550 (2016).
82. Nguyen, L. *et al.* Antibody Therapy Targeting RAN Proteins Rescues C9 ALS/FTD Phenotypes in C9orf72 Mouse Model. *Neuron* **105**, 645-662.e11 (2020).
83. Zuris, J. A. *et al.* Efficient Delivery of Genome-Editing Proteins In Vitro and In Vivo. *Nat. Biotechnol.* **33**, 73–80 (2015).
84. Nelson, C. E. *et al.* Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. *Nat. Med.* **25**, 427–432 (2019).
85. Ibraheim, R. *et al.* Self-inactivating, all-in-one AAV vectors for precision Cas9 genome editing via homology-directed repair in vivo. *Nat. Commun.* **12**, 6267 (2021).
86. Zhong, G. *et al.* A reversible RNA on-switch that controls gene expression of AAV-delivered therapeutics in vivo. *Nat. Biotechnol.* **38**, 169–175 (2020).
87. Gillmore, J. D. *et al.* CRISPR-Cas9 In Vivo Gene Editing for Transthyretin Amyloidosis. *N. Engl. J. Med.* **385**, 493–502 (2021).
88. Zhu, L. J., Holmes, B. R., Aronin, N. & Brodsky, M. H. CRISPRseek: a bioconductor package to identify target-specific guide RNAs for CRISPR-Cas9 genome-editing systems. *PloS One* **9**, e108424 (2014).
89. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
90. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).

91. Wilson, K. A., Chateau, M. L. & Porteus, M. H. Design and Development of Artificial Zinc Finger Transcription Factors and Zinc Finger Nucleases to the hTERT Locus. *Mol. Ther. Nucleic Acids* **2**, e87 (2013).
92. Clement, K. *et al.* CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).
93. O'Rourke, J. G. *et al.* C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. *Neuron* **88**, 892–901 (2015).
94. Zeier, Z. *et al.* Bromodomain inhibitors regulate the C9ORF72 locus in ALS. *Exp. Neurol.* **271**, 241–250 (2015).
95. Edwards, J. A. & Edwards, R. A. *Fastq-pair: efficient synchronization of paired-end fastq files*. 552885 <https://www.biorxiv.org/content/10.1101/552885v1> (2019) doi:10.1101/552885.
96. Karolchik, D. *et al.* The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* **32**, D493-496 (2004).