Title: Antisense oligonucleotide-mediated MSH3 suppression reduces somatic CAG repeat expansion in Huntington's patient iPSC-derived striatal neurons

Authors: Emma L. Bunting[1†], Jasmine Donaldson[1†], Sarah A. Cumming[2], Jessica Olive[1], Elizabeth Broom[1], Mihai Miclăuş[3,4], Joseph Hamilton[1], Matthew Tegtmeyer[5], Hien T. Zhao[6], Jonathan Brenton[3,7], Won-Seok Lee[5,8], Robert E. Handsaker[5,8], Susan Li[6], Brittany Ford[6], Mina Ryten[3,7], Steven McCarroll[5,8], Holly B. Kordasiewicz[6], Darren G. Monckton[2], Gabriel Balmus[3,4], Michael Flower[1], Sarah J. Tabrizi[1*]

Affiliations:

[1] Huntington's Disease Centre and Department of Neurodegenerative Disease, UCL Queen Square Institute of Neurology; and UK Dementia Research Institute, UCL, London, UK

[2] School of Molecular Biosciences, College of Medical, Veterinary and Life Sciences, University of Glasgow; Glasgow, UK

[3] UK Dementia Research Institute at the University of Cambridge and Department of Clinical Neurosciences, University of Cambridge, Cambridge Biomedical Campus, Cambridge, CB2 0AH, UK.

[4] Department of Molecular Neuroscience, Transylvanian Institute of Neuroscience, 400191, Cluj-Napoca, Romania

[5] Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard; Cambridge, MA, USA

[6] Ionis Pharmaceuticals; Carlsbad, CA, USA

[7] Department of Genetics and Genomic Medicine, Great Ormond Street Institute of Child Health; London, UK

[8] Department of Genetics, Harvard Medical School; Boston, MA, USA

[†] These authors contributed equally

[*] Corresponding author: s.tabrizi@ucl.ac.uk

One Sentence Summary: Therapeutic lowering of *MSH3* by ASO treatment mediates a dose-dependent reduction in CAG repeat expansion in HD iPSC-derived neurons.

Abstract: Expanded Huntington's disease (HD) associated *HTT*-CAG alleles are genetically unstable and continue to somatically expand throughout life, driving disease onset. MSH3, a DNA mismatch repair protein, modifies HD onset and progression by driving CAG repeat expansion. *MSH3* is relatively tolerant of loss-of-function variation in humans, making it an ideal therapeutic target. Here, we show that an *MSH3*-targeting antisense oligonucleotide (ASO) effectively enters and engages its target in HD patient induced pluripotent stem cell (iPSC)-derived striatal neurons. ASO treatment led to a dose-dependent reduction of MSH3, and subsequent stalling of CAG repeat expansion. Maximal knock-down also effectively slowed CAG repeat expansion in FAN1$^{-/-}$ neurons, which otherwise have an accelerated expansion rate. Bulk RNA sequencing showed a safe profile for *MSH3* lowering in the iPSC-derived neuronal context. Finally, we created a novel humanised *MSH3* knock-in mouse model and demonstrated effective *in vivo* modulation of human *MSH3* RNA, providing an important tool for future translation of *MSH3*-targeting compounds. Our study validates the use of ASO-mediated *MSH3* lowering to delay onset and slow disease progression in HD patients, which we show can prevent *HTT* CAG repeat expansion in an HD patient-derived model system.

Main Text:

INTRODUCTION

Huntington's disease (HD) is a fatal neurodegenerative disorder caused by an expanded CAG repeat in the huntingtin gene (*HTT*), with longer inherited alleles initiating an earlier age at onset (AAO) *(1)*. Expanded CAG repeat tracts are unstable intergenerationally *(2-8)* and somatically *(5, 6, 9-12)*, with somatic cells showing a bias towards expansion in a length, time and tissuedependent manner *(6, 10, 11, 13-16)*. The timing of HD onset depends on the length and the purity of the inherited CAG repeat *(17)*. Length can be affected by modifiers responsible for maintaining DNA stability *(18, 19)*, while purity can be affected by modifiers interrupting the CAG tract, most notably by CAA *(15, 19, 20)*. Both classes of modifiers affect somatic expansion rates in blood, which in turn correlate with clinical outcomes *(15)*. These observations support the idea that HD pathogenesis involves two sequential components. Firstly, the inheritance of an expanded repeat (≥36 CAG), which upon transcription, replication, or repair, forms secondary structures such as hairpins that are substrates for further expansion. Secondly, somatic expansion of the CAG repeats to a length above a toxic pathogenic threshold, which triggers neuronal damage, dysfunction and eventual neurodegeneration that drives the clinical HD features *(21)*. It is now widely accepted that therapeutic strategies to reduce somatic expansion could delay onset and slow progression in HD patients.

MSH3 variants have been implicated as modifiers of HD by several genome-wide association studies *(19, 22, 23)*. MSH3 is a mismatch repair (MMR) protein, which, together with MSH2, forms the MutSβ complex. Although its canonical function is to protect DNA integrity by mainly repairing insertion-deletion loops *(24-26)*, collective evidence in multiple repeat expansion disorder mouse models, human cell models, and patients suggest it exerts its modifier effect through somatic expansion via the error-prone repair of abnormal DNA structures at CAG repeats. Lowering MSH3 reduces repeat expansion in mouse models of HD *(27-30)*, as well as other trinucleotide repeat disorders such as myotonic dystrophy type 1 (DM1; CTG expansions in *DMPK*) *(31, 32)* and fragile X syndrome (FXS; CGG expansions in *FMR1*) *(33)*. This is also seen in human cell models of CTG *(34-36)* and GAA *(37)* expansions. In DM1 patients, *MSH3* SNP rs26279, which results in an amino acid change in the ATPase domain, significantly affected somatic instability. In HD patients, the AAO-hastening *MSH3* haplotype 5AM1 (tagSNP: rs701383) was associated with *cis*-eQTLs for increased *MSH3* expression in blood and was enriched in the top 25% of HD gene carriers with the highest proportion of somatic expansions *(19)*. Another study identified the leading imputed SNP from the *MSH3/DHFR* signal in the TRACK-HD GWAS to be associated with a three-repeat allele (*MSH3 3a*) of a nona-nucleotide tandem repeat sequence within exon 1 of *MSH3*. This variant was associated with delayed AAO, reduced somatic expansion in both HD and DM1, and slower disease progression in HD *(38)*. Although a functional difference between the *MSH3 3a* allele and the most common *MSH3 6a* allele cannot be ruled out, whole blood RNA-Sequencing (RNA-Seq) indicated that homozygous *MSH3 3a* carriers had reduced *MSH3* expression *(38)*, suggesting this contributes, at least in part, to the reduced somatic expansion observed for individuals carrying these alleles. The same variant, *MSH3 3a*, was later implicated as a modifier of X-linked Dystonia Parkinsonism (XDP; CCCTCT expansions within a *SVA* retrotransposon insertion in *TAF1*), similarly delaying disease onset *(39)*. Finally, seven HD patients have been identified with heterozygous loss-of-function mutations in *MSH3*, all of whom had delayed onset and/or less severe cognitive/motor signs. For those with onset data available, their average age at motor onset was delayed by 10.6 years *(40, 41)*, ultimately validating the efficacy of MSH3 depletion in the target HD patient population. Together, these data highlight MSH3 lowering as a powerful therapeutic strategy for repeat expansion disorders.

Unlike many other DNA repair genes implicated as modifiers of HD, *MSH3* can tolerate a high degree of loss-of-function variation *(42, 43)*, making it an ideal therapeutic target. The ratio of observed/expected loss-of-function variants in *MSH3* is 0.912 *(44)*, which is well above the threshold of 0.6, which typically indicates significant intolerance of loss-of-function. Concordant with this, *Msh3*-deficient mice are not significantly more tumour-prone than their wild-type counterparts, and they demonstrate a normal lifespan *(45, 46)*. Mutations in other members of the MMR network typically cause genetic colorectal cancer syndromes in patients *(47-55)*, whereas *MSH3* mutations are generally well-tolerated *(44)*. This excludes some extremely rare cases, reported only in 16 individuals to date, whereby germline biallelic *MSH3* mutations were shown to associate with colorectal adenomas, which if left untreated can lead to cancer *(56-60)*. The lower risk of carcinogenesis in the absence of MSH3 compared to most other MMR proteins is likely explained by functional redundancy for the repair of base mismatches with the more abundant MSH6 *(61, 62)*. Additionally, MSH3's specific substrate, small insertion-deletion loops, are less oncogenic. Importantly, no links between MSH3 dysfunction and CNS-derived tumours have been made in humans *(63)*. This is supported by recent HD mouse data whereby injection of *MSH3* siRNAs, which are primarily retained in the CNS, resulted in no detectable impact on CNS mononucleotide repeat instability at 12 weeks post-injection *(29)*. Together, these data suggest that lowering MSH3 in the brain appears safe.

ASOs are short nucleic acids, chemically modified to bind RNA by Watson-Crick base pairing, leading to subsequent degradation. They have been shown to be effective in numerous animal models and patients with a wide range of diseases *(64, 65)*, with six so far having been approved for neurological disease *(66, 67)*. Given the technical challenges of targeting gene expression in patient brains, particularly in deep-brain structures like the striatum *(61, 62)*, it is important to investigate the degree of *MSH3* lowering that may be required to have a clinical benefit. This study examines the effect of *MSH3* lowering on CAG repeat instability in 125 CAG iPSC-derived medium spiny neuron (MSN)-enriched cultures. The effect of knock-out of FAN1, the most significant modifier of HD AAO *(19)*, on the somatic instability outcomes of *MSH3* lowering is also explored. In addition, to further characterise the transcriptional signature of *MSH3* lowering, its effects on the global transcriptome are examined through both the bulk and single-nuclei approaches. Finally, we describe the generation of a partially humanised knock-in *MSH3* mouse model, which will be important for evaluating human *MSH3*-targeting compounds *in vivo*.

RESULTS

MSH3 **ASO is efficiently taken up by HD patient-derived striatal neurons, resulting in dose-dependent knockdown.**

Ionis Pharmaceuticals Inc. designed approximately 230 gapmer ASOs complementary to human *MSH3* mRNA, which aimed to silence gene expression by recruiting RNAseH1-mediated degradation. ASOs were screened in A-431 cells at 2 μM, using qRT-PCR to assess *MSH3* expression following 24-hour ASO treatment. The top six *MSH3* ASOs were selected for further evaluation based on their potent activity against *MSH3* suppression in A-431 cells (table S1, fig. S1). The most potent of these ASOs, herein referred to as the *MSH3* ASO, was taken forward for further *in vitro* and *in vivo* studies.

Striatal MSNs are the most vulnerable cell-type in HD brains, and thus we sought to assess the effect of MSH3 lowering in post-mitotic neurons with physiological relevance. We differentiated HD patient-derived iPSCs originally carrying 125 CAG repeats in exon 1 of *HTT* to MSN-enriched striatal cultures using the Arber *et al.* 2015 *(68)* protocol (Fig. 1A). Consistent with previous reports, immunocytochemistry showed cultures were 86% ± 4 neuronal (MAP2+), 67% ± 13 striatal neuron identity (MAP2+, FOXP1+) and 30% ± 4 MSNs (MAP2+, FOXP1+, DARPP32+) (Fig. 1A)*(68, 69)*. MSN-enriched striatal cultures were treated at day 36 with the *MSH3* ASO for 1 week. To determine the cellular uptake of the *MSH3* ASO, these cultures were probed with an antibody that selectively binds to the phosphorothioate backbone of the ASO, demonstrating efficient gymnotic uptake by striatal neurons (FOXP1+, MAP2+, Fig. 1B). A dose titration was also performed, determining the doses required to achieve ~ 25%, ~ 50% and maximal non-toxic MSH3 knockdown to be 0.022, 0.26 and 3 μM, respectively (Fig. 1C, IC50 = 0.255, $R^2 = 0.75$). In addition, the effect of MSH3 lowering on the abundance of other MutS components (MSH2 and MSH6) was examined, with no changes detected (Fig. 1D). To assess the effect of long-term *MSH3* ASO treatment on MSH3 lowering in our MSN-enriched striatal cultures, we quantified *MSH3* transcript and protein over time. We observed continued dose-dependent knockdown of both measures (Fig. 1E, IC50 = 0.137, $R^2 = 0.75$, Fig. 1F) and found no changes in cell viability at 9 or 15 weeks of treatment (fig. S2). Together, these data show that the selected *MSH3* ASO successfully lowers MSH3 in human HD-striatal neurons without overt toxicity.

MSH3 lowering results in a dose-dependent reduction of CAG repeat expansion

To predict the degree of MSH3 knockdown required to slow somatic CAG repeat expansion by 50%, which may inform future drug trials, we measured the effect that long-term dose titration of MSH3 in MSN-enriched striatal cultures had on CAG repeat length over time. Cells were treated at day 36 and cultured continuously in ASO-containing media for 15 weeks, with samples collected every 3 weeks to assess CAG repeat length. We observed a dose-dependent slowing of CAG repeat expansion following *MSH3* ASO treatment with maximal MSH3 lowering ablating repeat expansions (β = -0.023 CAG/week ± 0.015, p = < 0.001)**,** whilst the scrambled (SCR) control ASO did not affect CAG repeat expansion rate (Fig. 2, A to C, fig. S3).

To examine the effects in cells completely lacking MSH3, we used CRISPR to generate 125 CAG $MSH3^{-/-}$ iPSC clones (Fig. 2D, fig. S4). Similar to the effects of the *MSH3* ASO, genetic *MSH3* knock-out did not lead to changes in MSH2 or MSH6 protein abundance (Fig. 2E). These clones were differentiated into MSN-enriched cultures, alongside those derived from 125 CAG unedited iPSCs, which had undergone the CRISPR editing process with no editing (unedited 125 CAG

CRISPR control), and CAG repeat expansion was measured for 12 weeks post-baseline (Fig. 2, F to H, fig.S5). While the CAG repeat in unedited neurons expanded over time (Figure 2, F to H, β = 0.089 CAG/week ± 0.01), *MSH3*$^{-/-}$ neurons showed ablation of CAG expansions and net contractions of the repeat, with a linear regression trend line that had a significantly non-zero slope (Fig. 2, F to H, β = -0.033 CAG / week ± 0.01, $p < 0.001$).

To determine the dose-dependence of the effect, MSH3 protein abundance was then plotted against CAG repeat instability rates for all *MSH3* ASO treatments and *MSH3*$^{-/-}$ (Fig. 2I). The fitted curve predicted that ~ 41% MSH3 lowering is sufficient to reduce the rate of somatic expansion by 50%, and that ~ 83% lowering completely prevents net expansions. MSH3 lowering beyond this point appeared to cause net contractions. These data show that lowering MSH3 via an ASO treatment can reduce CAG repeat expansion in a dose-dependent manner in MSN-enriched striatal cultures, with complete loss of MSH3 ablating repeat expansions and leading to net contractions of the CAG repeat over time.

Maximal ASO-mediated MSH3 lowering prevents repeat expansions on a FAN1 deficient background

FAN1 is the most significant genetic modifier of AAO in HD, with variants advancing onset by up to 5.2 years *(18, 19)*. Its expression has been shown to protect against CAG repeat expansion in HD patient cell models, with loss of FAN1 doubling *HTT* CAG expansion rates *(70, 71)*. Owing to the small absolute change in CAG length over the 15-week period in our 125Q MSN system (~ 1.2 CAGs total), we additionally sought to establish the effect of MSH3 lowering in a highly unstable system. If MSH3 lowering successfully slows repeat expansion in *FAN1*$^{-/-}$ cells, it could be an effective therapy even in individuals with high polygenic risk from genetic variants that promote somatic expansion. CAG expansion assays were therefore repeated in MSN-enriched cultures generated from a 125 CAG *FAN1*$^{-/-}$ line (fig. S4) treated with the maximal dose of the *MSH3* ASO (3 µM). For *FAN1*$^{-/-}$ MSN-enriched striatal cultures, linear regression analyses determined Vehicle (β = 0.254 CAG/week ± 0.06, p = < 0.001, Fig. 3B) and 3 µM SCR ASO slopes (β = 0.336 CAG/week ± 0.11, p< 0.001, Fig. 3B) to be significantly non-zero, demonstrating repeat expansion over time (Fig. 3, A, to C, fig. S6). Conversely, the 3 µM *MSH3* ASO condition was not significantly different to a zero slope (β = -0.054 CAG/week ± 0.08, p = 0.78, Fig. 3B) – demonstrating maximal MSH3 lowering prevented repeat expansions, even in a system highly biased towards expansions (Fig. 3, A to C). These experiments were performed alongside MSN-enriched striatal cultures derived from 125 CAG unedited CRISPR control iPSCs, which had undergone the CRISPR process without editing. Treatment of these cultures with 3 µM *MSH3* ASO replicated the data presented in Fig. 2A, showing *MSH3* ASO treatment ablates CAG repeat expansions (β = -0.02 CAG/week ± 0.02, Fig. 3, D to F, fig. S7). Together, these data demonstrate that MSH3 lowering can stop repeat expansion, even on a FAN1 deficient background. Large repeat expansions have been observed in HD patient brain *(10)*, even before onset, and are more common in HD patients with early onset *(14)*. However, these events are difficult to detect by bulk fragment length analysis, where there is bias towards amplification and detection of shorter and commoner alleles, and relative insensitivity to rare alleles. Small-pool PCR was therefore used to detect rare, large expansion or contraction events of the *HTT* mutant allele. *FAN1*$^{-/-}$ MSN-enriched striatal neurons treated with Vehicle or SCR ASO exhibited expansions up to 403 CAG repeats and contractions down to 120 CAG repeats (Fig. 3G). Although differences between treatment conditions did not reach significance, quantifying the frequency of these rare expansion events reflected the same pattern as in bulk fragment analysis measures. That is, large expansions were observed at similar frequencies with SCR ASO and Vehicle treatments, but *MSH3* ASO treatment tended to suppress their occurrence (Vehicle vs *MSH3* ASO: MD = 0.324 expansions per lane ± 0.155, $p = 0.0692$, Fig. 3H). Large expansions were also detectable in unedited MSN-enriched striatal neurons (fig. S8A), mirroring what is seen in HD patient post-mortem mortem brains *(10, 14)*. However, over the 12-week period assessed, these events occurred infrequently, and hence differences between conditions were not measurable for the unedited cells (fig. S8B). Collectively, these findings highlight the relevance of patient iPSC-derived neurons as a model for studying the biology of CAG repeat dynamics, reflecting the patterns observed in patient brains, and suggest that MSH3 may exert additional influence on the occurrence of large, rare expansion events.

Transcriptional changes associated with *MSH3* lowering show no dysregulation of cancer or DNA repair pathways

To assess the transcriptional signature of MSH3 lowering, we performed bulk RNA-Seq on neurons treated for 2 weeks post-baseline with Vehicle, SCR ASO (3 µM), and three *MSH3* ASOs (3 µM): *MSH3* ASO-1, which was used in all prior experiments, *MSH3* ASO-2 and *MSH3* ASO-3 (see methods section for sequence details). As expected, *MSH3* abundance was significantly reduced by all three *MSH3* ASOs compared to Vehicle, while the SCR ASO showed no effect (Fig. 4A). A stepwise reduction in *MSH3* read counts was observed following treatment with ASO-3 (reduced to 20% ± 9.0 of Vehicle, $p < 0.001$), ASO-2 (reduced to 10.2% ± 1.9 of Vehicle, $p < 0.001$) and ASO-1 (reduced to 3.9% ± 1.5 of Vehicle, $p < 0.001$). To compare transcriptional signatures of the three *MSH3* ASOs with SCR ASO and Vehicle control, we undertook differential gene expression analyses, pathway enrichment and Weighted Gene Co-expression Network Analysis (WGCNA).

There were 5290 differentially expressed genes (DEGs) between SCR ASO treated and Vehicle conditions, 2740 of which were downregulated (fig. S9A) suggestive of ASO chemistry effects.

To control for this and examine the impact of MSH3 reduction, we therefore compared each of the *MSH3* ASO treatments to SCR ASO. There were only two DEGs, both downregulated, between SCR and *MSH3* ASO-3, one of which being *MSH3* (Fig. 4C). Treatment with *MSH3* ASO-2 resulted in 53 DEGs, 48 of which were downregulated (Fig. 4D), and *MSH3* ASO-1 treatment resulted in 781 DEGs, 583 of which were downregulated (Fig. 4E). The full list of DEGs is provided in Data File S1, including that for each *MSH3* ASO in comparison to Vehicle (fig. S9, B to D).

We then performed Gene Ontology (GO) enrichment analyses and grouped enriched pathways into functional nodes to visualise these findings. The effects of ASO chemistry (SCR ASO vs Vehicle), revealed dysregulation of pathways that converged on RNA and protein metabolism, cellular localisation, and structural organisation (fig. S9E). GO enrichment analyses for *MSH3* ASO-2 and *MSH3* ASO-3 either returned no significant hits (False Discovery Rate (FDR) > 0.05), or there were insufficient DEGs to perform such analysis. The effect of *MSH3* ASO-1 compared to SCR ASO largely converged on terms related to cell development and structure and neuronal projection (Fig. 4F). The full details of the enriched GO terms are provided in Data File S2. In addition, we assessed whether *MSH3* lowering by these ASOs influenced DNA damage response (DDR) or cancer pathways. We found that these pathways were not significantly disrupted by *MSH3* lowering (Fig. 4G), even with ASO-1 achieving more than a 95% reduction in *MSH3*.

Next, we performed Weighted Gene Co-expression Network Analysis (WGCNA), a method that identifies clusters (modules) of genes with highly correlated expression patterns. This approach allows for the detection of potential co-regulated or functionally related genes by grouping them into modules in a holistic and unbiased manner (Fig. 4H, Data File S3). Compared to Vehicle, SCR ASO treatment correlated positively with the expression pattern of the Blue module, which showed enrichment for ribosomal and mitochondrial pathways. When compared to Vehicle, all three *MSH3* ASOs replicated this effect. Additionally, all four ASOs upregulated the Brown module, which was enriched for pathways related to the plasma membrane and structural activity. Although this upregulation of the Brown module was only statistically significant after FDR correction for ASO-1 (p = 0.033) and ASO-2 (p =0.0067), the consistent transcriptional pattern across all four ASO molecules suggests a common effect attributable to ASO chemistry.

Comparing *MSH3* ASOs to SCR ASO treatment, the only association was with the Magenta module, which was downregulated specifically by *MSH3* ASO-1. This module was enriched for synaptic membrane and morphogenesis. As this effect was not seen with other similarly potent *MSH3* ASOs, we considered whether an ASO sequence-specific effect, such as off-target knockdown, could be driving it. We examined the gene expression profiles of the predicted off-target loci for *MSH3* ASO-1; within our dataset, 9 of the 18 predicted off-target loci were annotated (fig. S10). Of those, two were downregulated; *ATG5* (p = 8.89E-28) and *EYS* (p = 1.98E-05). These genes are members of pathways involved in cell projection, cell development, system development and cell differentiation, suggesting they are likely responsible for the downregulation of the Magenta module.

Cell-type specific transcriptional profiling of *MSH3* ASO treated cultures using single nucleus RNA-Seq

We next undertook single nucleus RNA-Sequencing (snRNA-Seq) to investigate whether the *MSH3* ASO (*MSH3* ASO-1) had cell-type specific effects or affected the composition of cells in our MSN cultures. We compared 125 CAG MSNs treated with either *MSH3* ASO (3 μM) or SCR ASO (3 μM) for 12 weeks post-baseline. To reduce the complexity of this large dataset, whilst preserving its intrinsic structure, we used Uniform Manifold Approximation and Projection

(UMAP). We then used Leiden unsupervised clustering to identify groups of similar cells based on their gene expression profiles. Five clusters were identified, all of which were present in similar proportions in both *MSH3* and SCR ASO treatment groups (Fig. 5A). No clusters were specific to either treatment group, suggesting *MSH3* ASO treatment did not affect culture composition, either through influencing cell fate or causing toxicity (Fig. 5, B and C). *MSH3* ASO treatment reduced *MSH3* abundance (Fig. 5D) and resulted in fewer cells expressing *MSH3* in all clusters (fig. S11), suggesting it entered and engaged its target in all cell-types within the culture.

Differential expression analysis in this snRNA-Seq data replicated findings from bulk RNA-Seq. *MSH3* ASO-1 treatment was once again associated with significant downregulation of pathways involved in synaptic structures and neuronal development ($p < 0.01$, Fig. 5, E and F).

To determine whether differential expression was cell-state specific, each cell was scored by its expression of the 20 genes most downregulated by *MSH3* ASO treatment. This involved summing their expression and comparing to the sum of 5000 control genes. Scores were consistently lower for all *MSH3* ASO-treated clusters, compared to SCR ASO treated clusters (Fig. 5G). This suggests the downregulation of synaptic genes within our MSN-enriched cultures, which we believe to be off-target driven, was not cell-state specific.

***MSH3* lowering throughout the central nervous system after *MSH3* ASO injection into partially humanised *MSH3* knock-in mice**

Even when ASOs share the same chemistry and target gene, safety and efficacy can vary greatly depending on targeting sequence, as highlighted by our RNA-sequencing data. This underscores the need for a humanised mouse model for the most informative *in vivo* testing. We generated *MSH3* knock-in mice humanised for exons 2 to 10 under the endogenous mouse *Msh3* promoter (Fig. 6A), which can assist the future development of *MSH3*-targeting molecules. As proof-ofconcept, we tested *MSH3* ASO-1, our most potent tool compound. Bolus intracerebral ventricular (ICV) injection of ASOs into the lateral ventricles is an efficient method of ASO delivery *(72)*. *MSH3* knock-in mice were treated with *MSH3* ASO-1 via ICV, and tissues were harvested 14 days later (Fig. 6B). The *MSH3* ASO resulted in a significant dose-dependent reduction of *MSH3* mRNA across all regions tested, with the maximal dose of 30 µg leading to a 49% reduction in the dissected cortex ($p < 0.001$), a 46% reduction in the striatum ($p < 0.001$), a 77% reduction in the brainstem ($p < 0.001$) and a 74% reduction in the spinal cord ($p < 0.001$) (Fig. 6C). ICV injection with Vehicle (PBS) or a scrambled ASO showed no change in *MSH3* expression (Supplementary Figure 12A). Thus, ICV injection of the *MSH3* ASO reduced *MSH3* mRNA in multiple brain regions, including the striatum which shows vulnerability in HD pathogenesis. The ASO was well tolerated at the doses tested, as evidenced by normal home cage behaviour and no motor abnormalities at the two-week necropsy timepoint. Staining of the whole brain and spinal cord revealed no sign of neurodegeneration as assessed by calbindin expression (fig. S12B). These data show *in vivo* target engagement of the ASO demonstrating that >75% suppression of *MSH3* is well-tolerated *in vivo*, and, establishing a foundation for future translation of *MSH3*-targeting ASOs.

DISCUSSION

Establishing the degree of MSH3 lowering required for therapeutic benefit is vital to inform future clinical development strategies in patients. Our study explores this important question within the context of human HD patient-derived cells, which we address using a human *MSH3* targeting ASO (*MSH3* ASO-1). Using MSN-enriched striatal cultures derived from the 125 CAG HD iPSC model, we demonstrate that the *MSH3* ASO is successfully taken up into HD striatal neurons, and that MSH3 lowering leads to a dose-dependent reduction in somatic expansion. That is, 41% MSH3 lowering is sufficient to slow the somatic expansion rate by half, and 83% MSH3 lowering is predicted to completely halt the process. Knockdown beyond this point appears to result in net contractions of the *HTT* CAG repeat. Maximal MSH3 lowering (89%) with the *MSH3* ASO was effective at halting bulk somatic expansions, even in a more rapidly expanding *FAN1*$^{-/-}$ system and appeared to have an additional effect on reducing the frequency of rare, large expansion events.

The slowing effect of MSH3 lowering on somatic expansion has been previously well described in rodents and engineered cells *(27-29, 33)*. However, only four previous studies have investigated the dose dependence of this effect in HD, all of which were conducted using mouse models *(27, 29, 30, 73)*. We replicate these findings in a human system, whereby somatic expansion is slowed in a dose-dependent manner with MSH3 lowering. Additionally, we expand the scope compared to these previous studies by including five different degrees of MSH3 knockdown, thereby improving the granularity in the dose titration with relevance for clinical trial design. Moreover, we demonstrate that MSH3 lowering remains effective at ablating somatic expansions even in a highly somatically unstable system (125 CAG *FAN1*$^{-/-}$ MSN-enriched cultures), suggesting the treatment should be effective even in individuals with high polygenic risk with regards to somatic CAG repeat instability.

Understanding the mechanism(s) of contraction is of therapeutic interest, and the data in this study may provide some insights. We have showed in an HD neuronal model that lack of MSH3 expression leads to an apparent contraction bias, consistent with what has been reported during intergenerational transmission of disease-associated repeats in multiple MSH3-deficient transgenic mice, including DM1, HD, FRDA (Friedrich's ataxia) and FXS *(27, 31, 33, 74)*. Biochemical studies in Lovo cell extracts support these findings, showing CTG•CAG repeat contractions to be enhanced in the absence of MutSβ *(75)*. Similarly, *Msh2* deficiency resulted in an increase in both intergenerational and somatic contractions in a DM1 mouse model *(76)*. In two distinct human cell models harbouring CTG repeat expansions of varying length (CTG-22 and CTG-800), contractions were still permitted in *MSH3*-deficient cells *(35, 36)*. Finally, using a novel hTERT-immortalised retinal pigment epithelial cell (RPE1) model of CAG repeat instability, *MSH3*$^{-/-}$ showed contractions of the repeat over time *(77)*. Collectively, the data presented here, alongside these earlier findings, suggest that CAG repeat contraction, but not expansion, can continue in the absence of MutSβ.

Bulk RNA-Seq revealed significant ASO chemistry- and ASO sequence-specific effects on the global transcriptome, and importantly a safe *MSH3* lowering profile with no dysregulation of DNA repair or oncogenic pathways. The delivery of ASOs themselves, regardless of target, was positively associated with modules enriched for ribosomal and mitochondrial processes, as well as plasma membrane and structural activity. When controlling for these ASO chemistry effects by comparing to SCR ASO; two *MSH3*-targeting ASOs showed no significantly disrupted pathways, while one was negatively associated with a module enriched for synaptic membrane and morphogenesis, likely attributable to off-target effects. These 3-10-3 cEt-modified gapmer phosphorothioate ASOs were chosen as tool compounds for their enhanced potency, though these notable off-target effects related to chemistry and sequence highlight that further optimisation would be necessary for their use as human therapeutics. Though this ASO (*MSH3* ASO-*1*) was also the most potent ASO against *MSH3*, the disparity in module dysregulation between this ASO and *MSH3* ASO-2 is unlikely to be explained by non-significant differences in *MSH3* transcripts between the two treatments, particularly as there is no evidence elsewhere in the literature to suggest that MSH3 has a role in regulating the expression of synaptic genes. Although two weeks of MSH3 modulation may be insufficient to produce long-term effects on genome dysregulation and should be interpreted cautiously, it is noteworthy that RNA-seq analysis of the zQ175 striatum has also shown no effect of MSH3 deficiency, even at six months of age *(30)*. SnRNA-Seq replicated the effects of *MSH3* ASO-1 treatment, demonstrating significant downregulation of pathways involved in synaptic structures and neuronal development. Notably, *MSH3* ASO treatment reduced *MSH3* abundance in all clusters, suggesting it was taken up successfully in all cells within the culture. Moreover, the wider spread effects of treatment with this ASO did not appear to be cell-type specific.

This paper describes the generation of a partially humanised knock-in *MSH3* mouse model, representing an excellent tool for future development of *MSH3*-targeting therapies. The *MSH3* ASO showed excellent *in vivo* pharmacology and efficient knockdown of *MSH3* in multiple brain regions, validating that human *MSH3* can be modulated *in vivo* with a therapeutic modality. It demonstrated 46% knockdown of *MSH3* in the striatum compared to Vehicle treated animals, which based on previous studies *(29)*, we would expect to have a substantial effect on somatic CAG repeat expansion. Since the focus of this study was to investigate the biology of MSH3 lowering rather than to identify a lead clinical candidate, it was beyond the scope to evaluate the dose-response or duration of effect in the novel *MSH3* knock-in mouse, though this work will be important in future clinical development. Future studies crossing the humanised *MSH3* knock-in mice to HD mice would permit the study of the effects of MSH3 knockdown on phenotype. However, it is worth noting that in previous studies of MSH3 knock-out HD mice, whilst there was ablation of repeat expansion and correction of histopathology, no phenotypic change was seen *(27, 29, 30)*. This may be because to show a behavioural phenotype within the relatively short lifespan of a mouse, these models include very long CAG repeats that already exceed cellular toxic thresholds, meaning ablation of further repeat instability would have limited impact on phenotype.

While our study advances the understanding of MSH3 lowering in modulating CAG repeat instability in Huntington's disease, certain limitations highlight opportunities for future research. Like previous reports in similar model systems *(40)*, we observed small absolute changes in CAG length over the period assessed in our patient iPSC-derived neurons, likely owing to the post-mitotic nature of our system. While this is reflective of the human brain, in which CAG expansions accumulate gradually over decades, it limits the ability to detect subtle effects between conditions over relatively short periods of time, particularly when further compounded by natural biological variability between differentiations. In the current study, we modulated MSH3 abundance for two weeks, which limits our ability to fully assess the long-term effects of MSH3 lowering on genome dysregulation and transcriptional regulation. However, evidence from animal models indicates that long-term MSH3 knockout rescues HD without significant off-target changes. Therefore, the significant dose-dependent changes observed with MSH3 lowering in our study underscore the potency of this intervention. Despite these constraints, the significant dose-dependent changes observed with MSH3 lowering underscore the potency of this intervention. In addition, we employed small-pool PCR to assess the effect of MSH3 lowering on large, rare expansion and contraction events. While this provides valuable insights into repeat instability patterns in our cultures, small-pool PCR is a semi-quantitative technique which typically lacks power to detect statistically significant differences between conditions. This is largely owing to the limited numbers of DNA templates included per reaction, the limited number of reactions that can be feasibly run, and the rare nature of the events of interest. To further elucidate the effect of MSH3 lowering on these large, rare events, future studies could employ single molecule barcoded amplicon sequencing, a high throughput technique that would permit comparison of a larger number of rare alleles across conditions.

Despite these limitations, human genetics studies have already validated the therapeutic potential of MSH3 lowering. An *MSH3* polymorphism that causes a subtle ~ 10% reduction in expression, was associated with reduced somatic expansion and slower progression of both clinical and brain imaging measures, as well as a 1.05-year delay in onset *(22)*. Moreover, heterozygous loss-of-function mutations in *MSH3* led to a 10.6-year delay in onset *(40)*. Drug discovery programmes that are based on validated human genetics such as this have a greater chance of success *(78)*, and here the human genetics suggest that targeting somatic expansion is critically important. The challenge now lies in translating these therapies into meaningful clinical trials. The new HD integrated staging system *(79)* supports such trials to be conducted in the earliest phases of the disease, before clinical manifestations, to allow somatic expansion to be targeted prior to overt toxicity. However, this will also need to be balanced with selecting a disease stage at which an effect would be measurable over the course of a clinical trial.

Therapeutic strategies targeting pathogenic processes downstream of the mutant huntingtin protein have so far proved ineffective. The existence of genetic modifiers, which are primarily involved in CAG repeat expansion *(19)*, proves disease course can be modified through interventions targeting this process, which represents the earliest pathogenic event in HD. MSH3, based on its requirement for repeat expansion and relative tolerance of loss-of-function, is the prime therapeutic candidate. Targeting *MSH3* lowering to the CNS, for example through an intrathecally delivered ASO, would mitigate any concerns about systemic complications. Our comprehensive analysis, including repeat instability and deep transcriptomic profiling, and extending from cell to animal models, underscores the potential of *MSH3* ASO treatment in HD and in other repeat expansion disorders to which MSH3 contributes.

MATERIALS AND METHODS

Study design

This study was designed to identify the degree of MSH3 lowering required to alter somatic instability in human MSNs, the most vulnerable cell-type in HD brains. MSN-enriched cultures were generated from QS5.1 cells, a HD patient iPSC line with 125 CAG repeats in *HTT*, using the Arber *et al.* 2015 *(68)* differentiation protocol. Following differentiation, MSN cultures were treated continuously from day 36 (referred to as baseline) with *MSH3* ASO, SCR ASO, or Vehicle control (PBS). CRISPR-Cas9 was used in the same cell line to generate *MSH3* knock-out (*MSH3*$^{-/-}$) and *FAN1* knock-out (*FAN1*$^{-/-}$) lines. CAG repeat length was monitored by fragment analysis and small-pool PCR, measuring changes from baseline over time. Effect on global transcriptome was assessed using both bulk and snRNA-Seq. Finally, a humanised *MSH3* knock-in mouse model was developed, which demonstrated that human *MSH3* can be modulated *in vivo* with a therapeutic modality.

ASO sequence and structure

Our study utilised 3-10-3 constrained ethyl (cEt) gapmer phosphorothioate ASOs, which includes the addition of 2’4’-Constratined-2’-O-ethyl chemical modification, commonly used to increase stability, enhance binding affinity and hence increase potency. *MSH3* ASO-1, selected as our primary *MSH3* targeting ASO owing to its potency, had the sequence GCTTTTGGACCGTTTG and targeted *MSH3* exon 4. *MSH3* ASO-2 (AGAAGATAGCTGGTAG, exon 7) and *MSH3* ASO-3 (GTATATTGATGTAGGT, intron 8) were included as additional *MSH3*-targeting ASOs in our RNA sequencing experiment. Our SCR ASO, ID 549148, had the sequence GGCTACTACGCCGTCA, with the same 3-10-3 cEt chemical structure as the *MSH3* ASOs.

A-431 cell culture condition

A-431 is a human epidermoid carcinoma cell line purchased from ATCC®. A-431 cells were cultured in DMEM growth medium (DMEM 10% FBS, 50 units/mL penicillin and 50 µg/mL streptomycin) at 37°C and 10% CO_2. Cells were subcultured or trypsinised for plating when they reached 80% confluency. Cells were trypsinised, counted and diluted to 110,000 cells per mL in room temperature growth medium before adding 100 µL of the cell suspension to the wells of collagen I coated 96-well culture plate. Immediately after plating the cells, *MSH3* ASOs (table S1) diluted in water were added to the appropriate wells. The culture plate was incubated in a humidified incubator at 37°C and 10% CO_2. After 48 hours, the cells were washed once with PBS before lysing with guanidine isothiocyanate-containing buffer for RNA isolation and analysis. For each treatment condition, duplicate wells were tested.

Human iPSC maintenance and differentiation

QS5.1 125 CAG iPSCs were derived from blood samples drawn from a 7-year-old female paediatric HD patient after full informed consent from the parent, and ethical approval from the local ethics committee with GCP compliance. iPSCs were generated by Sendai reprogramming at Censo Biotechnologies. They were tested and found to be karyotypically normal even after prolonged culture over one year. These cells originally had a repeat sequence containing 125 uninterrupted *HTT* CAG repeats, followed by a single CAACAG cassette, and they exhibit repeat expansion in culture. Prior to differentiation, iPSCs were maintained in E8+ Flex medium and were passaged at ~ 80% confluency using ReLeSR^{TM} at a 1:6 ratio – typically every 3 to 5 days. Differentiation into MSN-enriched neuronal cultures was achieved following the 36-day protocol as described in Arber *et al.* 2015 *(68)*, which involves the promotion of lateral ganglionic eminence fate via activin A treatment, followed by terminal differentiation into striatal neurons. Prior to terminal differentiation, the cells were passaged onto Geltrex-coated plates to ensure they remained adherent during long-term culture. Cultures were treated continuously with ASOs from day 36 of the differentiation process, when they are known to express striatal neuronal markers. Neurons were not passaged after this stage, but entire wells were harvested every 3-4 weeks to monitor CAG repeat instability.

CRISPR-Cas9 gene editing

Genetic editing of *MSH3* and *FAN1* was accomplished by electroporation of 125 CAG iPSCs with Cas9 protein and synthetic gRNA complexes using the P3 Primary Cell 4-D Nucleofector™ X kit L (Lonza, cat #V4XP-3024). An hour before nucleofection, iPSCs were changed into E8 Flex+ media with 10 µM ROCK inhibitor. The gRNA complex was assembled by annealing a target-specific cRNA with an ATTO™ 550-tagged tracrRNA, which was then incubated with Cas9 protein at room temperature (RT) for 20 minutes. Cells were dissociated with Accutase® for 10 minutes at 37°C, spun down and resuspended in Lonza P3 solution with the gRNA/Cas9 complex. Cells were transferred to Nucleocuvettes™, electroporated using the 4D-Nucleofector™ (P3 primary cell buffer; program CA137), then reseeded into 6-well plates. After overnight incubation, cells were detached with Accutase® and FAC-sorted to collect ATTO™ 550-positive cells. These were seeded sparsely into 10 cm dishes and colonies manually picked 3 to 5 days later using an in-hood microscope. The *MSH3* exon 1 or *FAN1* exon 2 loci of successfully isolated clones were sequenced to identify clones with knock-out edits (*MSH3* Fw: CCATGTCTCGCCGGAAG; *MSH3* Rev: GCCTATCTCCTCAAACGGAAG: *FAN1* Fw: GTCAGAAGGGAAACCTCCTG; *FAN1* Rv: TGTGGACTAGAACCGGCAAA). The clones were confirmed as *MSH3* or *FAN1* knock-out by western blotting.

qRT-PCR of iPSC-derived neuronal samples

RNA was extracted with RNeasy Mini Plus (Qiagen) kit and SuperScript™ IV First-Strand Synthesis System (Thermo Fisher Scientific, cat #18091050) was used to transcribe RNA to cDNA, with both steps performed according to manufacturer's instructions. qRT-PCRs were prepared using 20 ng template in 15 µL reactions with TaqMan Fast Advanced mastermix (Thermo Fisher Scientific) with exon spanning TaqMan assays to *MSH3* (*MSH3*: Hs00989003_m1) and run on a QuantStudio5 machine (Thermo Fisher Scientific). The PCR cycling conditions used were: 95°C for 40 seconds, followed by 40 cycles of 95°C for 15 seconds and 60°C for 30 seconds. All reactions were run in triplicate, and three endogenous reference genes (*ATP5B*: Hs00969569_m1, *EIF4A2*: Hs00756996_g1, *UBC*: Hs00824723_m1) were included for normalisation purposes. Differences in gene expression were assessed using the ΔΔCt method.

Western blotting of iPSC-derived neuronal samples

Protein was extracted by incubating in RIPA buffer (supplemented with 2x protease inhibitor and 250 U/mL benzonase) before 4°C centrifugation at 20,000 g for 15 minutes. The supernatant was then processed using the NuPage^{TM} gel electrophoresis system (Invitrogen). Protein was transferred onto PVDF membrane using the XCell II^{TM} Blot Module and blocked in fluorescent blocking buffer (Rockland, cat #MB-070) for 1 hour. Primary antibodies: (Mouse α-MSH3 (1:500), BD Biosciences cat #611390; Rabbit α-MSH2 (1:1000), Cell Signaling Technologies cat #D24B5; Mouse α-MSH6 (1:1000), BD Biosciences cat #610919; Rabbit α-Actin (1:10,000), Abcam cat #ab8227) were prepared in the same buffer, and incubated overnight at 4°C. Membranes were washed three times for 10 minutes in 1% Tween in D-PBS, and probed with IRDye$^{®}$-conjugated secondary antibodies at 1:10,000 (Donkey α-mouse 680, LI-COR cat #926-68072; Donkey α-rabbit 800, LI-COR cat #926-32213) for 1 hour at RT. Membranes were imaged on the LI-COR Odyssey^{TM} CLx imaging system and analysed with Image Studio^{TM} Lite software.

Immunostaining of iPSC-derived neuronal samples

Cells at day 28 to 30 of the differentiation protocol were seeded on CellCarrier-96 Ultra microplates (Perkin Elmer) at a density of 20,000 cells/well. Plates were fixed at the indicated time-points using 4% PFA. Cells were permeabilised with 0.2% Triton X-100 for 15 minutes and blocked for 1 hour with D-PBS containing 1% goat serum and 1% BSA. After aspiration, primary antibodies (Rabbit α-DARPP32 (1:200), Abcam cat #ab40802; Mouse α-FOXP1 (1:1000), Abcam cat #ab32010; Chicken α-MAP2 (1:1000), Abcam cat #ab5392; Rabbit α-Sox2 (1:200), Abcam cat #ab97959; Mouse α-Ki67 (1:500); BD Biosciences cat #550609) were added in 1% BSA in D-PBS and incubated at 4°C overnight. After D-PBS washes, cells were incubated with Alexa Fluor-conjugated secondary antibodies at 1:10,000 (Goat α-rabbit 488, Abcam cat #a11008; Goat α-mouse 568, Abcam cat #a11031; Goat α-chicken 647, Abcam cat #a21449) (RT; 1 hour maximum), washed again before adding Hoechst 33342 (1:10,000, Thermo Fisher Scientific cat# 62249, washed again three times, then kept in 100 µL D-PBS with 0.02% sodium azide until being processed on the Opera Phenix™ high-content imaging system. Confocal images were analysed using Columbus™ software, involving the generation of custom modular scripts to define cellular regions. A 'secondary antibody only' control was used to set fluorescence thresholds.

Fragment analysis *HTT* CAG repeat sizing

DNA was extracted with QIAamp DNA Mini kit (Qiagen) and bulk measurement of the *HTT* exon 1 CAG repeat length for each sample was carried out by fragment analysis. In brief, 100 ng of DNA was used for amplification of *HTT* exon 1 by 6-FAM labelled PCR primers and the amplicons were resolved by capillary electrophoresis on an ABI 3730xl genetic analyzer (Thermo Fisher Scientific) with the MapMarker 1000 ROX size standard (Eutogentec). Primers used were HTT_exon_1_F:6FAM-CCTTCGAGTCCCTCAAGTCCTT;

HTT_exon_1_R:CGGCTGAGGCAGCAGCGGCTGT). PCR conditions were 95°C for 10 minutes, followed by 30 cycles of 95°C for 30 seconds and 58°C for 30 seconds and 72°C for 90 seconds, and a final cycle of 72°C for 7 minutes.

The size standard was aligned using GeneMapper (Thermo Fisher Scientific). The mode, median, standard deviation and instability index of the pathogenic repeat were calculated as previously described using a custom program, available at https://michaelflower.org *(70, 71)*. Instability index *(12, 80)* was calculated relative to the baseline timepoint (day 36 of differentiation, day 0 of treatment) for each differentiation and genotype. Quality control removed samples that were poorly amplified (<40 RFU) and outlying technical replicates (>1.5 standard deviations from the mean). The primary instability metric used was instability index relative to control, which is a measure of skewness of the repeat length distribution relative to the mean repeat length of the baseline (control) samples. It was calculated as the sum of the change in repeat length, multiplied by proportional peak area, for each peak in the distribution. To compare CAG instability between groups, the lmer R package was used to generate a linear mixed effects model ANCOVA with random intercept and random slope for time within differentiation: lmer(instability_metric ~ treatment * time + (1 + time | differentiation)). This model allows for random intercepts for each differentiation and random slopes of instability over time, and accounts for the variability in baseline CAG length and its trajectory over time within each differentiation. To explore differences between groups in change of instability index over time, pairwise comparisons were performed using the emtrends function from the emmeans R package, which estimates marginal means for linear trends. These comparisons were adjusted for multiple testing using the Bonferroni method.

Small-pool PCR CAG repeat sizing

As previously described *(10, 81, 82)*, this involved the serial dilution of genomic DNA into multiple PCRs, such that the input for each contained ~ 500 pg DNA (*i.e.*, a ‘small-pool’). Limiting the number of input DNA templates reduces PCR competition with the non-expanded allele. PCR products were then resolved by agarose gel electrophoresis, transferred onto a nylon membrane, and hybridised with a radiolabelled probe specific to CTG•CAG repeats (DM56). Allele repeat lengths were estimated using GelAnalyzer software. Quantifications were carried out by counting the number of rare events outside the range detected by fragment analysis, and normalising to number of lanes. Differences between conditions were assessed by one-way ANOVA.

Bulk RNA-Seq

Data preprocessing

A total of 63,187 transcripts were initially identified. Quality control measures were applied, filtering out transcripts with low read counts *(83)*. Specifically, transcripts with fewer than 10 reads in over 90% of samples were deemed unreliable and were excluded from the analysis *(84, 85)*. This resulted in a refined dataset of 17,667 genes. The excluded genes did not enrich for DNA repair functions (fig. S13), ensuring the removal process did not inadvertently filter out genes relevant to this study’s focus.

DESeq2 Analysis

The DESeq2 R package was used to perform differential expression analysis *(85)*. The filtered counts were used to create a DESeqDataSet object, specifying the experimental design and conditions. The data were normalised using the DESeq2 algorithm, which estimates size factors to account for differences in library sizes. Differential expression analysis was performed using the DESeq function, which calculates log2 fold changes and p-values for each gene. Multiple testing correction was applied to control the FDR using the Benjamini-Hochberg procedure. Genes with an adjusted p-value (FDR) below 0.05 were considered differentially expressed.

Gene Ontology Enrichment Analysis

The list of differentially expressed genes obtained from the DESeq2 analysis was used as the input for the GO enrichment analysis, using all of the 17,667 genes considered in our analyses as background. GO enrichment analysis was performed using ShinyGO *(86)*. Enriched GO terms were identified based on an FDR cut-off < 0.05 and only root GO terms were retained. Enriched pathways were grouped using the goSlim function from the GSEABase package, using the goslim_agr subset (https://geneontology.org/docs/download-ontology/), and presented in nodal pathway plots. DNA repair and cancer gene sets from GO and KEGG (Data File S3) and those from NCG (http://network-cancer-genes.org) were specifically assessed for enrichment using clusterProfiler.

WGCNA Analysis

WGCNA package in R was used to construct a scale-free network using the samples in table S2 *(84)*. Sample clustering identified five outlying samples (Supplementary Figure 14A), which were removed from both WGCNA and differential gene expression analyses. A soft thresholding power of 6 was selected, balancing scale independence and connectivity (fig. S14B). The network cluster dendrogram is provided in fig. S14C.

Traits were defined as pairwise comparisons of treatments, and these were associated with modules in a linear regression analysis with each module eigengene. P-values were FDR adjusted for multiple comparisons.

To determine the biological roles of genes within each module, enrichment analyses were performed using the anRichment *(87)* and clusterProfiler *(88)* tools with GO and KEGG gene sets. The full pathway enrichment analysis for each module is provided in Data File S4. Hub genes were identified by calculating the module membership (MM) and gene significance (GS) for each gene. Hub genes are characterised by high MM, indicating strong association with a specific module, and high GS, signifying a significant correlation with the trait of interest *(84)*.

snRNA-Seq

Harvested neurons were lysed in EZ lysis buffer (Sigma Aldrich) using a Dounce homogeniser. The lysate was filtered through a 40 µm cell strainer, and the nuclei were isolated as a pellet in a density gradient separation with 30% iodixanol. The nuclei were washed with 3% BSA in PBS supplemented with RNAse inhibitor (0.2 U/µL, Biosearch Technologies), followed by additional filtration through a 40 µm strainer. snRNA-Seq libraries were generated from these nuclei using the 10X Chromium Single Cell 3’ Reagents Kits v3.1 with dual index and sequenced on a NextSeq 500/550 high output kit v2.5 (Illumina) with paired-end reads (28 cycles for read 1, 10 cycles for index 1 and index 2, respectively, and 90 cycles for read 2). Raw sequence files were then aligned and prepared following the Drop-Seq workflow *(89)*. Raw counts were normalised for library size and log1p transformed prior to downstream analysis. Human reads were aligned to GRCh38 and filtered for high quality mapped reads (MQ 10). Leiden clustering was performed across resolutions in UMAP space (0.2,0.4,0.6) and then visualised to determine the dimensionality of the data. LEIDEN_0.2 was used for the downstream analysis. For the metagene analysis, summed expression of genes downregulated in expression as a result of *MSH3* ASO treatment were divided by a random control set of 1000 genes.

MSH3 knock-in mice

The *MSH3* knock-in mouse used in this study is a constitutive line generated via homologous recombination in collaboration with Taconic Biosciences (Fig. 6A). Specifically, mouse genomic sequence from codon K54 in exon 2 to codon G449 in exon 10 from ENSMUST000000185852 *Msh3* transcript was replaced with its human counterpart ENST00000265081.7 *MSH3* transcript. The positive selection markers, Neomycin resistance (NeoR) and Puromycin resistance (PuroR), were flanked by flippase recognition target (FRT) and F3 sites, respectively, and inserted into introns 2 and 9. The targeting vector was generated using BAC clones from the mouse C57BL/6J PRCI-23 and human PRCI-11 BAC libraries and was transfected into the Taconic Biosciences C57BL/6N Tac ES cell line. Homologous recombinant clones were isolated using double positive (NeoR and PuroR) and negative (Thymidine kinase – Tk) selections. The humanised allele was obtained after flippase (FLP)-mediated removal of the selection markers. The chimeric mouse-human *MSH3* gene was expressed under the control of endogenous mouse *Msh3* promoter, and thus should recapitulate the expression pattern of the mouse *Msh3* gene. *MSH3* knock-in mice of mixed sex at 2 to 3 months of age were used in these experiments. Studies were approved by IONIS IACUC protocol 1176.

ICV injection method

Mice were anesthetised with 3% isoflurane gas anesthesia in either air or oxygen carrier gas using an induction box. The animals were placed in a stereotaxic apparatus. Anesthesia was maintained with 2% isoflurane by a nose cone fitted to the stereotaxic instrument. For each animal a ~ 1.0 cm slightly off center (to the right) incision was made in the scalp. The subcutaneous tissue and periosteum were scraped from the skull with a dry sterile cotton tipped applicator. For intracerebroventricular (ICV) bolus injections the coordinates 0.3 mm anterior to bregma, 1.0 mm right lateral, and -3.0 mm ventral were used. 10 µL of injection solution was injected by hand at injection rates of approximately 1 μL/second with a 3-minute wait following completion of injection. The incision was sutured closed using one horizontal mattress stitch with 5-O Ethilon suture. The animals were then allowed to recover from the anesthesia in their home cage. All animals were randomly assigned to each condition, and technicians were blinded to treatment conditions.

qRT-PCR of mouse tissues

Each brain region and spinal cord sample was homogenised in 1 mL of Trizol reagent (Thermo Fisher Scientific) and total RNA was extracted using the Qiagen mini-RNA purification kit (Qiagen) according to the manufacturer's protocol. After purification, the RNA samples were subjected to real-time qRT-PCR analysis using the Life Technologies ABI QuantStudio 7 Flex Sequence Detection System (Applied Biosystems Inc). Briefly, qRT-PCR reactions were run with the AgPath- ID One-Step qRT-PCR Kit (Thermo Fisher Scientific, Waltham, MA) reagents and the following primer probe sets: *MSH3* Fw CAAGCCTGCAGCCAGTAG; Rv GCTGTATTCAAAATAAATGTTATCCATCC: *Ppia* Fw TCGCCGCTTGCTGCA; Rv ATCGGCCGTGATGTCGA. Target *MSH3* mRNA was then normalised to *Ppia*, a ubiquitously expressed housekeeping gene, and presented relative animals treated with Vehicle. Differences between conditions were analysed by two-way ANOVA.

Statistical analyses

All data were analysed using R (v4.3.1).

MSH3 ASO dose titration curves were fitted using a y ~ log(x) function. IC50 was calculated by two methods. Firstly a dose-response model was fitted using the 'drm' function from the drc package, and the IC50 was extracted using the ED function. Secondly we used linear interpolation, which involved performing a linear regression on log-transformed dose-response data to predict response values over a range of doses. Using the approx function, we interpolated the dose corresponding to a 50% response level.

To compare CAG instability rates across groups, we used a linear mixed effects model (lmer(instability_metric ~ treatment * time + (1 + time | differentiation))) and performed pairwise comparisons with emtrends, adjusting for multiple testing with Bonferroni (see fragment analysis section for more details).

The effects of treatment on protein levels in western blots and instability in small-pool PCR were assessed using one-way ANOVA followed by Bonferroni-corrected post hoc pairwise comparisons.

To evaluate the effect of treatments on protein levels across different tissues in the *in vivo* dose titration experiments, we performed a two-way ANOVA followed by post hoc pairwise comparisons. The two-way ANOVA was conducted using the aov function with the following model: level ~ treatment * tissue. Post hoc pairwise comparisons were performed using the emmeans function, and the resulting p-values were Bonferroni adjusted for multiple testing.

References:

1. M. E. MacDonald, C. M. Ambrose, M. P. Duyao, R. H. Myers, C. Lin, L. Srinidhi, G. Barnes, S. A. Taylor, M. James, N. Groot, H. MacFarlane, B. Jenkins, M. A. Anderson, N. S. Wexler, J. F. Gusella, G. P. Bates, S. Baxendale, H. Hummerich, S. Kirby, M. North, S. Youngman, R. Mott, G. Zehetner, Z. Sedlacek, A. Poustka, A.-M. Frischauf, H. Lehrach, A. J. Buckler, D. Church, L. Doucette-Stamm, M. C. O'Donovan, L. Riba-Ramirez, M. Shah, V. P. Stanton, S. A. Strobel, K. M. Draths, J. L. Wales, P. Dervan, D. E. Housman, M. Altherr, R. Shiang, L. Thompson, T. Fielder, J. J. Wasmuth, D. Tagle, J. Valdes, L. Elmer, M. Allard, L. Castilla, M. Swaroop, K. Blanchard, F. S. Collins, R. Snell, T. Holloway, K. Gillespie, N. Datson, D. Shaw, P. S. Harper, A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**, 971-983 (1993).

2. N. A. Aziz, M. J. van Belzen, I. D. Coops, R. D. M. Belfroid, R. A. C. Roos, Parent-of-origin differences of mutant HTT CAG repeat instability in Huntington's disease. *Eur J Med Genet* **54**, e413-418 (2011).
3. M. Duyao, C. Ambrose, R. Myers, A. Novelletto, F. Persichetti, M. Frontali, S. Folstein, C. Ross, M. Franz, M. Abbott, J. Gray, P. Conneally, A. Young, J. Penney, Z. Hollingsworth, I. Shoulson, A. Lazzarini, A. Falek, W. Koroshetz, D. Sax, E. Bird, J. Vonsattel, E. Bonilla, J. Alvir, J. Bickham Conde, J. H. Cha, L. Dure, F. Gomez, M. Ramos, J. Sanchez-Ramos, S. Snodgrass, M. de Young, N. Wexler, C. Moscowitz, G. Penchaszadeh, H. MacFarlane, M. Anderson, B. Jenkins, J. Srinidhi, G. Barnes, J. Gusella, M. MacDonald, Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nature Genetics* **4**, 387-392 (1993).
4. M. E. MacDonald, G. Barnes, J. Srinidhi, M. P. Duyao, C. M. Ambrose, R. H. Myers, J. Gray, P. M. Conneally, A. Young, J. Penney, Gametic but not somatic instability of CAG repeat length in Huntington's disease. *Journal of Medical Genetics* **30**, 982 (1993).
5. H. Telenius, B. Kremer, Y. P. Goldberg, J. Theilmann, S. E. Andrew, J. Zeisler, S. Adam, C. Greenberg, E. J. Ives, L. A. Clarke, M. R. Hayden, Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. *Nature Genetics* **6**, 409-414 (1994).
6. V. C. Wheeler, W. Auerbach, J. K. White, J. Srinidhi, A. Auerbach, A. Ryan, M. P. Duyao, V. Vrbanac, M. Weaver, J. F. Gusella, A. L. Joyner, M. E. MacDonald, Length-Dependent Gametic CAG Repeat Instability in the Huntington's Disease Knock-in Mouse. *Human Molecular Genetics* **8**, 115-122 (1999).
7. V. C. Wheeler, F. Persichetti, S. M. McNeil, J. S. Mysore, S. S. Mysore, M. E. MacDonald, R. H. Myers, J. F. Gusella, N. S. Wexler, T. US–Venezuela Collaborative Research Group, Factors associated with HD CAG repeat instability in Huntington disease. *Journal of Medical Genetics* **44**, 695 (2007).
8. C. Zühlke, O. Rless, B. Bockel, H. Lange, U. Thies, Mitotic stability and meiotic variability of the (CAG)n repeat in the Huntington disease gene. *Human Molecular Genetics* **2**, 2063-2067 (1993).
9. K. E. De Rooij, P. A. M. De Koning Gans, R. A. C. Roos, G.-J. B. Van Ommen, J. T. Den Dunnen, Somatic expansion of the $(CAG)^n$ repeat in Huntington disease brains. *Human Genetics* **95**, 270-274 (1995).
10. L. Kennedy, E. Evans, C. M. Chen, L. Craven, P. J. Detloff, M. Ennis, P. F. Shelbourne, Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. *Human Molecular Genetics* **12**, 3359-3367 (2003).
11. L. Kennedy, P. F. Shelbourne, Dramatic mutation instability in HD mouse striatum: does polyglutamine load contribute to cell-specific vulnerability in Huntington’s disease? *Human Molecular Genetics* **9**, 2539-2544 (2000).
12. R. Mouro Pinto, L. Arning, J. V. Giordano, P. Razghandi, M. A. Andrew, T. Gillis, K. Correia, J. S. Mysore, D.-M. Grote Urtubey, C. R. Parwez, S. M. von Hein, H. B. Clark, H. P. Nguyen, E. Förster, A. Beller, S. Jayadaev, C. D. Keene, T. D. Bird, D. Lucente, J.-P. Vonsattel, H. Orr, C. Saft, E. Petrasch-Parwez, V. C. Wheeler, Patterns of CAG repeat instability in the central nervous system and periphery in Huntington’s disease and in spinocerebellar ataxia type 1. *Human Molecular Genetics* **29**, 2551-2567 (2020).
13. P. F. Shelbourne, C. Keller-McGandy, W. L. Bi, S.-R. Yoon, L. Dubeau, N. J. Veitch, J. P. Vonsattel, N. S. Wexler, N. Arnheim, S. J. Augood, The US-Venezuela Collaborative

Research Group, Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. *Human Molecular Genetics* **16**, 1133-1142 (2007).

14. M. Swami, A. E. Hendricks, T. Gillis, T. Massood, J. Mysore, R. H. Myers, V. C. Wheeler, Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. *Human molecular genetics* **18**, 3039-3047 (2009).
15. M. Ciosi, A. Maxwell, S. A. Cumming, D. J. Hensman Moss, A. M. Alshammari, M. D. Flower, A. Durr, B. R. Leavitt, R. A. C. Roos, P. Holmans, L. Jones, D. R. Langbehn, S. Kwak, S. J. Tabrizi, D. G. Monckton, A genetic association study of glutamine-encoding DNA sequence structures, somatic CAG expansion, and DNA repair gene variants, with Huntington disease clinical outcomes. *EBioMedicine* **48**, 568-580 (2019).
16. N. J. Veitch, M. Ennis, J. P. McAbney, P. F. Shelbourne, D. G. Monckton, Inherited CAG·CTG allele length is a major modifier of somatic mutation length variability in Huntington disease. *DNA Repair (Amst)* **6**, 789-796 (2007).
17. D. R. Langbehn, M. R. Hayden, J. S. Paulsen, Predict-HD Investigators of the Huntington Study Group, CAG-repeat length and the age of onset in Huntington disease (HD): A review and validation study of statistical approaches. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **153B**, 397-408 (2010).
18. GeM-HD, Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. *Cell* **162**, 516-526 (2015).
19. GeM-HD, CAG Repeat Not Polyglutamine Length Determines Timing of Huntington's Disease Onset. *Cell* **178**, 887-900.e814 (2019).
20. G. E. B. Wright, J. A. Collins, C. Kay, C. McDonald, E. Dolzhenko, Q. Xia, K. Bečanović, B. I. Drögemöller, A. Semaka, C. M. Nguyen, B. Trost, F. Richards, E. K. Bijlsma, F. Squitieri, C. J. D. Ross, S. W. Scherer, M. A. Eberle, R. K. C. Yuen, M. R. Hayden, Length of Uninterrupted CAG, Independent of Polyglutamine Size, Results in Increased Somatic Instability, Hastening Onset of Huntington Disease. *Am J Hum Genet* **104**, 1116-1126 (2019).
21. E. P. Hong, M. E. MacDonald, V. C. Wheeler, L. Jones, P. Holmans, M. Orth, D. G. Monckton, J. D. Long, S. Kwak, J. F. Gusella, J.-M. Lee, Huntington's Disease Pathogenesis: Two Sequential Components. *Journal of Huntington's Disease* **10**, 35-51 (2021).
22. D. J. H. Moss, A. F. Pardinas, D. Langbehn, K. Lo, B. R. Leavitt, R. Roos, A. Durr, S. Mead, P. Holmans, L. Jones, S. J. Tabrizi, Identification of genetic variants associated with Huntington's disease progression: a genome-wide association study. *The Lancet. Neurology* **16**, 701-711 (2017).
23. J.-M. Lee, Y. Huang, M. Orth, T. Gillis, J. Siciliano, E. Hong, J. S. Mysore, D. Lucente, V. C. Wheeler, I. S. Seong, Z. L. McLean, J. A. Mills, B. McAllister, S. V. Lobanov, T. H. Massey, M. Ciosi, G. B. Landwehrmeyer, J. S. Paulsen, E. R. Dorsey, I. Shoulson, C. Sampaio, D. G. Monckton, S. Kwak, P. Holmans, L. Jones, M. E. MacDonald, J. D. Long, J. F. Gusella, Genetic modifiers of Huntington disease differentially influence motor and cognitive domains. *The American Journal of Human Genetics* **109**, 885-899 (2022).
24. R. R. Iyer, A. Pluciennik, V. Burdett, P. L. Modrich, DNA Mismatch Repair: Functions and Mechanisms. *Chemical Reviews* **106**, 302-323 (2006).

25. F. Palombo, I. Iaccarino, E. Nakajima, M. Ikejima, T. Shimada, J. Jiricny, hMutSβ, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Current Biology* **6**, 1181-1184 (1996).
26. J. Genschel, S. J. Littman, J. T. Drummond, P. Modrich, Isolation of MutSβ from Human Cells and Comparison of the Mismatch Repair Specificities of MutSβ and MutSα*. *Journal of Biological Chemistry* **273**, 19895-19901 (1998).
27. E. Dragileva, A. Hendricks, A. Teed, T. Gillis, E. T. Lopez, E. C. Friedberg, R. Kucherlapati, W. Edelmann, K. L. Lunetta, M. E. MacDonald, V. C. Wheeler, Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiol Dis* **33**, 37-47 (2009).
28. S. Tomé, K. Manley, J. P. Simard, G. W. Clark, M. M. Slean, M. Swami, P. F. Shelbourne, E. R. M. Tillier, D. G. Monckton, A. Messer, C. E. Pearson, MSH3 Polymorphisms and Protein Levels Affect CAG Repeat Instability in Huntington's Disease Mice. *PLOS Genetics* **9**, e1003280 (2013).
29. D. O'Reilly, J. Belgrad, C. Ferguson, A. Summers, E. Sapp, C. McHugh, E. Mathews, A. Boudi, J. Buchwald, S. Ly, D. Moreno, R. Furgal, E. Luu, Z. Kennedy, V. Hariharan, K. Monopoli, X. W. Yang, J. Carroll, M. DiFiglia, N. Aronin, A. Khvorova, Di-valent siRNA-mediated silencing of MSH3 blocks somatic repeat expansion in mouse models of Huntington’s disease. *Molecular Therapy* **31**, 1661-1674 (2023).
30. S. G. Aldous, E. J. Smith, C. Landles, G. F. Osborne, M. Cañibano-Pico, I. M. Nita, J. Phillips, Y. Zhang, B. Jin, M. B. Hirst, C. L. Benn, B. C. Bond, W. Edelmann, J. R. Greene, G. P. Bates, A CAG repeat threshold for therapeutics targeting somatic instability in Huntington's disease. *Brain* **147**, 1784-1798 (2024).
31. L. Foiry, L. Dong, C. Savouret, L. Hubert, H. te Riele, C. Junien, G. Gourdon, Msh3 is a limiting factor in the formation of intergenerational CTG expansions in DM1 transgenic mice. *Human genetics* **119**, 520-526 (2006).
32. W. van den Broek, M. R. Nelen, D. G. Wansink, M. M. Coerwinkel, H. T. Riele, P. Groenen, B. Wieringa, Somatic expansion behaviour of the (CTG)(n) repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. *Human Molecular Genetics* **11**, 191-198 (2002).
33. X. N. Zhao, D. Kumari, S. Gupta, D. Wu, M. Evanitsky, W. Yang, K. Usdin, Muts beta generates both expansions and contractions in a mouse model of the Fragile X-associated disorders. *Human Molecular Genetics* **24**, 7087-7096 (2015).
34. A.-M. M. Gannon, A. Frizzell, E. Healy, R. S. Lahue, MutS beta and histone deacetylase complexes promote expansions of trinucleotide repeats in human cells. *Nucleic Acids Research* **40**, 10324-10333 (2012).
35. R. Nakatani, M. Nakamori, H. Fujimura, H. Mochizuki, M. P. Takahashi, Large expansion of CTG•CAG repeats is exacerbated by MutSβ in human cells. *Sci. Rep.* **5**, (2015).
36. N. Keogh, K. Y. Chan, G.-M. Li, R. S. Lahue, MutSβ abundance and Msh3 ATP hydrolysis activity are important drivers of CTG•CAG repeat expansions. *Nucleic Acids Research* **45**, 10068-10078 (2017).
37. A. Halabi, S. Ditch, J. Wang, E. Grabczyk, DNA mismatch repair complex MutSβ promotes GAA·TTC repeat expansion in human cells. *Journal of Biological Chemistry* **287**, 29958-29967 (2012).

38. M. Flower, V. Lomeikaite, M. Ciosi, S. Cumming, F. Morales, K. Lo, D. Hensman Moss, L. Jones, P. Holmans, D. G. Monckton, S. J. Tabrizi, TRACK-HD Investigators, OPTIMISTIC Consortium, MSH3 modifies somatic instability and disease severity in Huntington's and myotonic dystrophy type 1. *Brain* **142**, 1876-1886 (2019).
39. B.-H. Laabs, C. Klein, J. Pozojevic, A. Domingo, N. Brüggemann, K. Grütz, R. L. Rosales, R. D. Jamora, G. Saranza, C. C. E. Diesta, M. Wittig, S. Schaake, M. Dulovic-Mahlow, J. Quismundo, P. Otto, P. Acuna, C. Go, N. Sharma, T. Multhaupt-Buell, U. Müller, H. Hanssen, F. Kilpert, A. Franke, A. Rolfs, P. Bauer, V. Dobričić, K. Lohmann, L. J. Ozelius, F. J. Kaiser, I. R. König, A. Westenberger, Identifying genetic modifiers of age-associated penetrance in X-linked dystonia-parkinsonism. *Nature Communications* **12**, 3216 (2021).
40. B. McAllister, J. Donaldson, C. S. Binda, S. Powell, U. Chughtai, G. Edwards, J. Stone, S. Lobanov, L. Elliston, L. N. Schuhmacher, E. Rees, G. Menzies, M. Ciosi, A. Maxwell, M. J. Chao, E. P. Hong, D. Lucente, V. Wheeler, J. M. Lee, M. E. MacDonald, J. D. Long, E. H. Aylward, G. B. Landwehrmeyer, A. E. Rosser, J. S. Paulsen, N. M. Williams, J. F. Gusella, D. G. Monckton, N. D. Allen, P. Holmans, L. Jones, T. H. Massey, Exome sequencing of individuals with Huntington's disease implicates FAN1 nuclease activity in slowing CAG expansion and disease onset. *Nat Neurosci* **25**, 446-457 (2022).
41. H. w. e. s. project. (2022).
42. M. Lek, K. J. Karczewski, E. V. Minikel, K. E. Samocha, E. Banks, T. Fennell, A. H. O'Donnell-Luria, J. S. Ware, A. J. Hill, B. B. Cummings, T. Tukiainen, D. P. Birnbaum, J. A. Kosmicki, L. E. Duncan, K. Estrada, F. Zhao, J. Zou, E. Pierce-Hoffman, J. Berghout, D. N. Cooper, N. Deflaux, M. DePristo, R. Do, J. Flannick, M. Fromer, L. Gauthier, J. Goldstein, N. Gupta, D. Howrigan, A. Kiezun, M. I. Kurki, A. L. Moonshine, P. Natarajan, L. Orozco, G. M. Peloso, R. Poplin, M. A. Rivas, V. Ruano-Rubio, S. A. Rose, D. M. Ruderfer, K. Shakir, P. D. Stenson, C. Stevens, B. P. Thomas, G. Tiao, M. T. Tusie-Luna, B. Weisburd, H.-H. Won, D. Yu, D. M. Altshuler, D. Ardissino, M. Boehnke, J. Danesh, S. Donnelly, R. Elosua, J. C. Florez, S. B. Gabriel, G. Getz, S. J. Glatt, C. M. Hultman, S. Kathiresan, M. Laakso, S. McCarroll, M. I. McCarthy, D. McGovern, R. McPherson, B. M. Neale, A. Palotie, S. M. Purcell, D. Saleheen, J. M. Scharf, P. Sklar, P. F. Sullivan, J. Tuomilehto, M. T. Tsuang, H. C. Watkins, J. G. Wilson, M. J. Daly, D. G. MacArthur, C. Exome Aggregation, Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285-291 (2016).
43. K. J. Karczewski, B. Weisburd, B. Thomas, M. Solomonson, D. M. Ruderfer, D. Kavanagh, T. Hamamsy, M. Lek, K. E. Samocha, B. B. Cummings, D. Birnbaum, The Exome Aggregation Consortium, M. J. Daly, D. G. MacArthur, The ExAC browser: displaying reference data information from over 60 000 exomes. *Nucleic Acids Research* **45**, D840-D845 (2017).
44. R. R. Iyer, A. Pluciennik, DNA Mismatch Repair and its Role in Huntington's Disease. *Journal of Huntington's Disease* **10**, 75-94 (2021).
45. N. de Wind, M. Dekker, N. Claij, L. Jansen, Y. van Klink, M. Radman, G. Riggins, M. van der Valk, K. van't Wout, H. te Riele, HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. *Nat Genet* **23**, 359-362 (1999).

46. W. Edelmann, A. Umar, K. Yang, J. Heyer, M. Kucherlapati, M. Lia, B. Kneitz, E. Avdievich, K. Fan, E. Wong, G. Crouse, T. Kunkel, M. Lipkin, R. D. Kolodner, R. Kucherlapati, The DNA Mismatch Repair Genes Msh3 and Msh6 Cooperate in Intestinal Tumor Suppression1. *Cancer Research* **60**, 803-807 (2000).
47. P. Peltomäki, Update on Lynch syndrome genomics. *Fam Cancer* **15**, 385-393 (2016).
48. P. Peltomäki, Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J Clin Oncol* **21**, 1174-1179 (2003).
49. M. Abedalthagafi, Constitutional mismatch repair-deficiency: current problems and emerging therapeutic strategies. *Oncotarget* **9**, 35458-35469 (2018).
50. R. A. Barnetson, A. Tenesa, S. M. Farrington, I. D. Nicholl, R. Cetnarskyj, M. E. Porteous, H. Campbell, M. G. Dunlop, Identification and Survival of Carriers of Mutations in DNA Mismatch-Repair Genes in Colon Cancer. *New England Journal of Medicine* **354**, 2751-2763 (2006).
51. R. S. Hamad, M. E. Ibrahim, CMMRD caused by PMS1 mutation in a sudanese consanguineous family. *Hereditary Cancer in Clinical Practice* **20**, 16 (2022).
52. H. Hampel, W. L. Frankel, E. Martin, M. Arnold, K. Khanduja, P. Kuebler, H. Nakagawa, K. Sotamaa, T. W. Prior, J. Westman, J. Panescu, D. Fix, J. Lockman, I. Comeras, A. de la Chapelle, Screening for the Lynch Syndrome (Hereditary Nonpolyposis Colorectal Cancer). *New England Journal of Medicine* **352**, 1851-1860 (2005).
53. G. E. Palomaki, M. R. McClain, S. Melillo, H. L. Hampel, S. N. Thibodeau, EGAPP supplementary evidence review: DNA testing strategies aimed at reducing morbidity and mortality from Lynch syndrome. *Genetics in Medicine* **11**, 42-65 (2009).
54. R. H. Sijmons, R. M. W. Hofstra, Review: Clinical aspects of hereditary DNA Mismatch repair gene mutations. *DNA Repair (Amst)* **38**, 155-162 (2016).
55. K. Wimmer, J. Etzler, Constitutional mismatch repair-deficiency syndrome: have we so far seen only the tip of an iceberg? *Human Genetics* **124**, 105-122 (2008).
56. R. Adam, I. Spier, B. Zhao, M. Kloth, J. Marquez, I. Hinrichsen, J. Kirfel, A. Tafazzoli, S. Horpaopan, S. Uhlhaas, D. Stienen, N. Friedrichs, J. Altmüller, A. Laner, S. Holzapfel, S. Peters, K. Kayser, H. Thiele, E. Holinski-Feder, G. Marra, G. Kristiansen, M. M. Nöthen, R. Büttner, G. Möslein, R. C. Betz, A. Brieger, R. P. Lifton, S. Aretz, Exome Sequencing Identifies Biallelic MSH3 Germline Mutations as a Recessive Subtype of Colorectal Adenomatous Polyposis. *Am J Hum Genet* **99**, 337-351 (2016).
57. A. S. Aelvoet, D. R. Hoekman, B. J. W. Redeker, J. Weegenaar, E. Dekker, C. J. M. van Noesel, F. A. M. Duijkers, A large family with MSH3-related polyposis. *Fam Cancer* **22**, 49-54 (2023).
58. M.-C. Villy, J. Masliah-Planchon, A. Schnitzler, H. Delhomelle, B. Buecher, M. Filser, K. Merchadou, L. Golmard, S. Melaabi, S. Vacher, M. Blanluet, V. Suybeng, C. Corsini, M. Dhooge, N. Hamzaoui, S. Farelly, A. Ait Omar, R. Benamouzig, V. Caumette, M. Bahuau, J. Cucherousset, Y. Allory, D. Stoppa-Lyonnet, I. Bieche, C. Colas, MSH3 : a confirmed predisposing gene for adenomatous polyposis. *J Med Genet*, jmg-2023-109341 (2023).
59. A. Gavric, M. Krajc, L. Strnisa, A. U. Gavric, S. Plut, MSH3-related adenomatous polyposis in a patient with the negative family history of colorectal polyps. *Gastroenterología y Hepatología* **47**, 397-400 (2024).

60. M. Koi, B. H. Leach, S. McGee, S. S. Tseng-Rogenski, C. A. Burke, J. M. Carethers, Compound heterozygous MSH3 germline variants and associated tumor somatic DNA mismatch repair dysfunction. *npj Precision Oncology* **8**, 12 (2024).
61. N. Pecina-Slaus, A. Kafka, I. Salamon, A. Bukovac, Mismatch Repair Pathway, Genome Stability and Cancer. *Front Mol Biosci* **7**, 122 (2020).
62. G. X. Reyes, T. T. Schmidt, R. D. Kolodner, H. Hombauer, New insights into the mechanism of DNA mismatch repair. *Chromosoma* **124**, 443-462 (2015).
63. M. Zhao, Y. Liu, G. Ding, D. Qu, H. Qu, Online database for brain cancer-implicated genes: exploring the subtype-specific mechanisms of brain cancer. *BMC Genomics* **22**, 458 (2021).
64. C. F. Bennett, B. F. Baker, N. Pham, E. Swayze, R. S. Geary, Pharmacology of antisense drugs. *Annual review of pharmacology and toxicology* **57**, 81-105 (2017).
65. S. T. Crooke, J. L. Witztum, C. F. Bennett, B. F. Baker, RNA-targeted therapeutics. *Cell metabolism* **27**, 714-739 (2018).
66. C. F. Bennett, H. B. Kordasiewicz, D. W. Cleveland, Antisense Drugs Make Sense for Neurological Diseases. *Annual Review of Pharmacology and Toxicology* **61**, 831-852 (2021).
67. M. E. McCauley, C. F. Bennett, Antisense drugs for rare and ultra-rare genetic neurological diseases. *Neuron* **111**, 2465-2468 (2023).
68. C. Arber, S. V. Precious, S. Cambray, J. R. Risner-Janiczek, C. Kelly, Z. Noakes, M. Fjodorova, A. Heuer, M. A. Ungless, T. A. Rodriguez, A. E. Rosser, S. B. Dunnett, M. Li, Activin A directs striatal projection neuron differentiation of human pluripotent stem cells. *Development* **142**, 1375-1386 (2015).
69. R. Ferguson, R. Goold, L. Coupland, M. Flower, S. J. Tabrizi, Therapeutic validation of MMR-associated genetic modifiers in a human ex vivo model of Huntington disease. *The American Journal of Human Genetics* **111**, 1165-1183 (2024).
70. R. Goold, M. Flower, D. H. Moss, C. Medway, A. Wood-Kaczmar, R. Andre, P. Farshim, G. P. Bates, P. Holmans, L. Jones, S. J. Tabrizi, FAN1 modifies Huntington's disease progression by stabilizing the expanded HTT CAG repeat. *Human Molecular Genetics* **28**, 650-661 (2018).
71. R. Goold, J. Hamilton, T. Menneteau, M. Flower, E. L. Bunting, S. G. Aldous, A. Porro, J. R. Vicente, N. D. Allen, H. Wilkinson, G. P. Bates, A. A. Sartori, K. Thalassinos, G. Balmus, S. J. Tabrizi, FAN1 controls mismatch repair complex assembly via MLH1 retention to stabilize CAG repeat expansion in Huntington's disease. *Cell Rep* **36**, 109649 (2021).
72. F. Rigo, S. J. Chun, D. A. Norris, H. Gene, L. Sam, M. John, A. F. Robert, G. Hans, H. Yimin, S. G. John, R. K. Adrian, P. H. Scott, C. F. Bennett, Pharmacology of a Central Nervous System Delivered 2′-*O*-Methoxyethyl–Modified Survival of Motor Neuron Splicing Oligonucleotide in Mice and Nonhuman Primates. *Journal of Pharmacology and Experimental Therapeutics* **350**, 46 (2014).
73. R. Driscoll, L. Hampton, N. A. Abraham, J. D. Larigan, N. F. Joseph, J. C. Hernandez-Vega, S. Geisler, F.-C. Yang, M. Deninger, D. T. Tran, N. Khatri, B. M. D. C. Godinho, G. A. Kinberger, D. R. Montagna, W. D. Hirst, C. L. Guardado, K. E. Glajch, H. M. Arnold, C. L. Gallant-Behm, A. Weihofen, Dose-dependent reduction of somatic expansions but not Htt aggregates by di-valent siRNA-mediated silencing of MSH3 in HdhQ111 mice. *Scientific Reports* **14**, 2061 (2024).

74. V. Ezzatizadeh, R. M. Pinto, C. Sandi, M. Sandi, S. Al-Mahdawi, H. Te Riele, M. A. Pook, The mismatch repair system protects against intergenerational GAA repeat instability in a Friedreich ataxia mouse model. *Neurobiol Dis* **46**, 165-171 (2012).
75. M. M. Slean, G. B. Panigrahi, A. L. Castel, A. B. Pearson, A. E. Tomkinson, C. E. Pearson, Absence of MutSβ leads to the formation of slipped-DNA for CTG/CAG contractions at primate replication forks. *DNA Repair (Amst)* **42**, 107-118 (2016).
76. C. Savouret, E. Brisson, J. Essers, R. Kanaar, A. Pastink, H. te Riele, C. Junien, G. Gourdon, CTG repeat instability and size variation timing in DNA repair-deficient mice. *The EMBO Journal* **22**, 2264-2273 (2003).
77. Z. L. McLean, G. Dadi, C. Kevin, C. L. R. Jennie, S. Shota, N. F. Iris, V.-M. Zoe, R. Manasa, M. Elisabetta, R. Jayla, G. Tammy, L. Diane, P. K. Benjamin, L. Jong-Min, E. M. Marcy, C. W. Vanessa, P. Ricardo Mouro, F. G. James, *PMS1* as a target for splice modulation to prevent somatic CAG repeat expansion in Huntington's disease. *bioRxiv*, 2023.2007.2025.550489 (2023).
78. R. M. Plenge, E. M. Scolnick, D. Altshuler, Validating therapeutic targets through human genetics. *Nature Reviews Drug Discovery* **12**, 581-594 (2013).
79. S. J. Tabrizi, S. Schobel, E. C. Gantman, A. Mansbach, B. Borowsky, P. Konstantinova, T. A. Mestre, J. Panagoulias, C. A. Ross, M. Zauderer, A. P. Mullin, K. Romero, S. Sivakumaran, E. C. Turner, J. D. Long, C. Sampaio, A biological classification of Huntington's disease: the Integrated Staging System. *The Lancet Neurology* **21**, 632-644 (2022).
80. J.-M. Lee, J. Zhang, A. I. Su, J. R. Walker, T. Wiltshire, K. Kang, E. Dragileva, T. Gillis, E. T. Lopez, M.-J. Boily, M. Cyr, I. Kohane, J. F. Gusella, M. E. MacDonald, V. C. Wheeler, A novel approach to investigate tissue-specific trinucleotide repeat instability. *BMC Systems Biology* **4**, 29 (2010).
81. M. Gomes-Pereira, S. I. Bidichandani, D. G. Monckton, in *Trinucleotide Repeat Protocols,* Y. Kohwi, Ed. (Humana Press, Totowa, NJ, 2004), pp. 61-76.
82. D. G. Monckton, L.-J. C. Wong, T. Ashizawa, C. T. Caskey, Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. *Human Molecular Genetics* **4**, 1-8 (1995).
83. R. Bourgon, R. Gentleman, W. Huber, Independent filtering increases detection power for high-throughput experiments. *Proceedings of the National Academy of Sciences* **107**, 9546-9551 (2010).
84. P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
85. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).
86. S. X. Ge, D. Jung, R. Yao, ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628-2629 (2020).
87. B. Zhang, S. Horvath, A General Framework for Weighted Gene Co-Expression Network Analysis. **4**, (2005).
88. T. Wu, E. Hu, S. Xu, M. Chen, P. Guo, Z. Dai, T. Feng, L. Zhou, W. Tang, L. Zhan, X. Fu, S. Liu, X. Bo, G. Yu, clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**, 100141 (2021).
89. E. Z. Macosko, A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, E. M. Martersteck, J. J. Trombetta, D. A. Weitz, J. R. Sanes, A. K.

Shalek, A. Regev, S. A. McCarroll, Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202-1214 (2015).

Acknowledgments

We wish to thank Drs Tom Vogt, Brinda Prasad and Michael Finley for helpful discussions and input on the research. Thank you to Richard Newton for assistance with figures and tables. We thank Taconic Biosciences for their contribution to the generation of MSH3 knock-in mice. Thank you to Professor Peter Holmans for consultation on our statistical approach. We would also like to thank Dr. Branduff McAllister for performing the calculation based on published data ***(40, 41)*** regarding the delay in age at onset in *MSH3* heterozygous LoF HD individuals.

Funding

Medical Research Council (MRC) MR/N013867/1 (ELB)

CHDI Foundation (ELB, JD, MF, JO, SJT, DGM, SAC)

UK Dementia Research Institute that receives its funding from DRI Ltd., funded by the UK MRC, and the CHDI Foundation (MF, JO, SJT, GB)

Wellcome Trust, 223082/Z/21/Z (SJT)

Romanian Ministry of Research, Innovation and Digitization (no. PNRR-III-C9-2022-I8-66; contract 760114). (GB, MM)

Author contributions

Conceptualization: SJT

Methodology: ELB, JD, MF, HTZ, JO, EB, SAC, MM, WSL, MT, JB, JH, SL, BF, BH, MR, GB, SM, HBK, DGM, SJT

Investigation: ELB, JD, MF, HTZ, JO, EB, SAC, MM, WSL, MT, JB, JH, BH

Visualization: JD, MF, JH, ELB, HTZ, JO, EB, SAC, MM, WSL, MT, JB, BH, MR, GB, SM, HBK, DGM, SJT

Funding acquisition: ELB, SJT

Project administration: SJT

Supervision: SJT, DGM, JD, SAC, MF, GB

Writing – original draft: ELB

Writing – review & editing: JD, ELB, MF, JO, EB, HTZ, SAC, MM, WSL, MT, JB, JH, BH, MR, GB, SM, FB, HBK, DGM, SJT

Competing interests

SJT competing interests: In the previous 12 months Sarah J. Tabrizi has also received research grant funding from the UK Medical Research Council, the Wellcome Trust and the CHDI Foundation. In the previous 12 months, through the offices of UCL Consultants Ltd, a wholly owned subsidiary of University College London, Sarah J. Tabrizi has undertaken consultancy services for Alnylam Pharmaceuticals, Annexon, Ascidian Therapeutics, Arrowhead Pharmaceuticals, Atalanta Therapeutics, Design Therapeutics, F. Hoffman-La Roche, Iris Medicine, Latus Bio, LifeEdit, Novartis Pharma, Pfizer, Prilenia Neurotherapeutics, PTC Therapeutics, Rgenta Therapeutics, Takeda Pharmaceuticals, UniQure Biopharma, Vertex Pharmaceuticals. In the past 12 months, University College London Hospitals NHS Foundation Trust, Professor Tabrizi's host clinical institution, received funding to run clinical trials for F. Hoffman-La Roche, Novartis Pharma, PTC Therapeutics, and UniQure Biopharma. STJ is founder of Function RX Ltd.

DGM competing interests: Within the last five years Darren G. Monckton has been a scientific consultant and/or received an honoraria/grants from AMO Pharma, Dyne, F. Hoffman-La Roche, LoQus23, MOMA Therapeutics, Novartis, Ono Pharmaceuticals, Pfizer Pharmaceuticals, Rgenta Therpeutics, Sanofi, Sarepta Therapeutics Inc, Script Biosciences, Triplet Therapeutics, and Vertex Pharmaceuticals. Darren G. Monckton also had research contracts with AMO Pharma and Vertex Pharmaceuticals.

G.B. is founder and chief executive officer (part time) of Function RX Ltd.

HTZ, SL, BF and HBK are full-time employees and shareholders at Ionis Pharmaceuticals. All work conducted at Ionis was funded by Ionis.

ELB is now a full-time employee of F. Hoffman-La Roche.

Data and materials availability

All the data are included in the main text or in the Supplementary Materials. For gene-edited iPSC lines, please contact corresponding author at UCL. RNA sequencing data are deposited at SRA; bulk RNA-Seq accession number PRJNA1177884, snRNA-Seq accession number PRJNA1180543. Individual-level data and statistical analyses are provided in data files S6-S8. Data and analysis scripts for repeat sizing are available at https://github.com/mike-flower/MSH3_aso_juliet, and for other analyses at https://github.com/mike-flower/MSH3_aso_analysis.

Supplementary Materials

Fig. S1: *in vitro* screening of ASOs.

Fig. S2: Effect of MSH3 ASO treatment on cell viability.

Fig. S3: Fragment analysis to assess the effect of MSH3 lowering on somatic instability.

Fig. S4: Generation of *FAN1*$^{-/-}$ and *MSH3*$^{-/-}$ 125 CAG iPSC lines.

Fig. S5: Fragment analysis to assess the effect of MSH3 knock-out on somatic instability.

Fig. S6: Fragment analysis to assess the effects of MSH3 lowering on somatic instability in 125 CAG *FAN1*$^{-/-}$ MSN-enriched cultures.

Fig. S7: Fragment analysis to measure the effects of MSH3 lowering on somatic instability in 125 CAG unedited MSN-enriched cultures.

Fig. S8: Small-pool PCR to measure the effect of *MSH3* ASO treatment on large, rare expansion events.

Fig. S9: Bulk RNA-Seq analysis of the global transcriptome following treatment with ASOs vs Vehicle.

Fig. S10: Predicted off-targets of MSH3-targeting ASOs.

Fig. S11: Effect of *MSH3* ASO on *MSH3* expression across the snRNA-Seq clusters.

Fig. S12: Effect of ASO treatment on *MSH3* and calbindin in vivo.

Fig. S13: Overlap analysis between DNA repair genes and transcripts excluded during RNA-Seq quality control.

Fig. S14: RNA-Seq WGCNA analysis selection.

Fig. S15: Cluster dendrogram of gene modules in RNA-Seq WGCNA analysis.

Table S1: *MSH3* ASO sequence and chemistry.

Table S2: Samples used for WGCNA network construction.

Dataset S1: Expression data of all the 23869 ENSEMBL IDs used in our analyses. Lists of DEGs are presented in individual sheets for each timepoint and experimental conditions. Additionally, expression values of all the genes (*i.e.*, DEGs and non-DEGs) are provided.

Dataset S2: GO enrichment analyses.

Dataset S3: DNA repair and cancer gene sets from GO and KEGG.

Dataset S4: Enrichment analyses from WGCNA. Gene set enrichment analysis results for each gene module.

Dataset S5: WGCNA modules. The 'module_genes' tab lists the genes present in each module. The 'hub_genes' tab gives the hub gene for each module. The 'module_summary' tab provides detail about each module, including the number of genes they contain, correlation coefficients and p-values for association with each trait, and the top GO term.

Dataset S6: Fragment analysis data.

Dataset S7: Fragment analysis map.

Dataset S8: Individual-level data and statistical analyses.

Figures

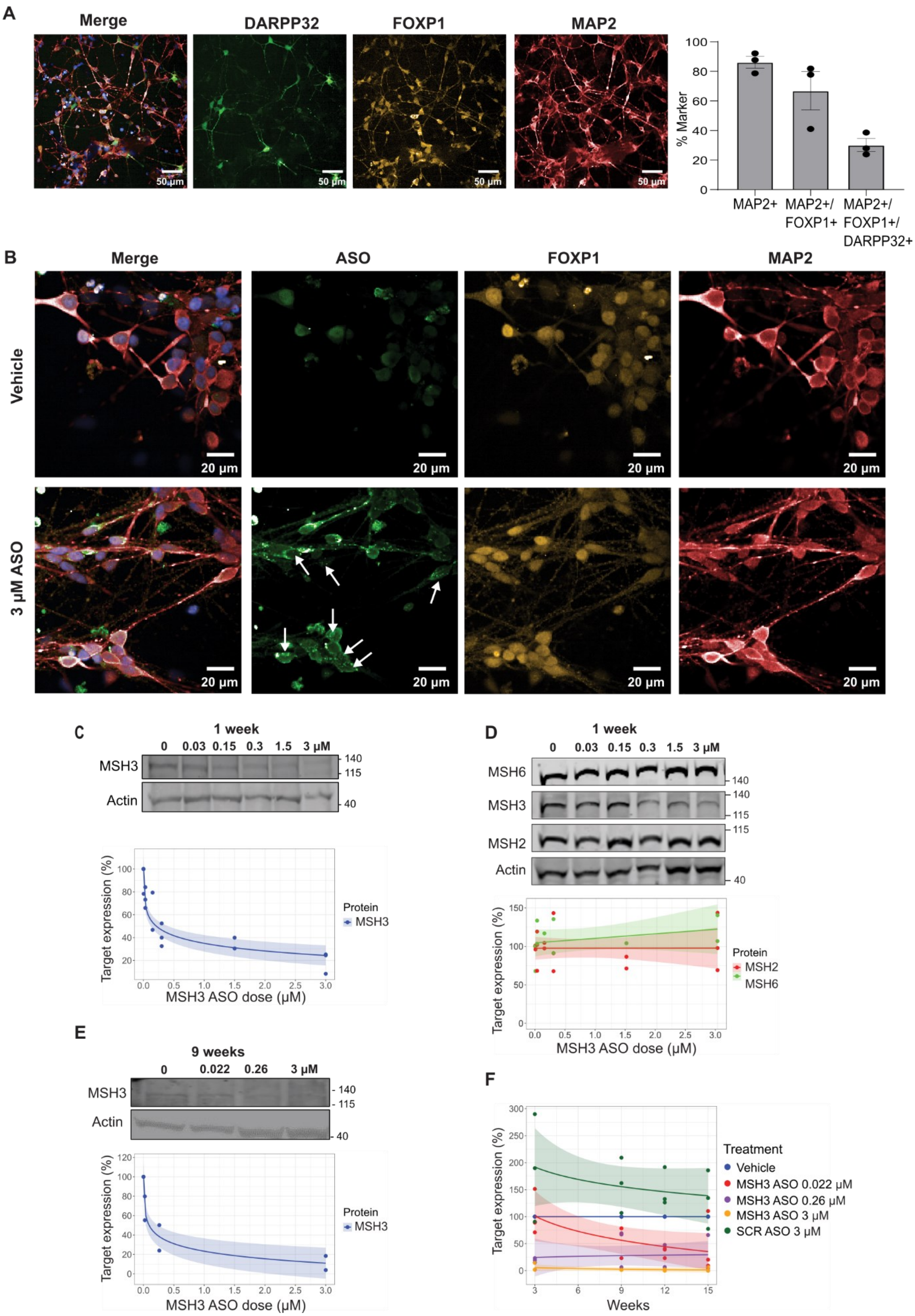
A
Merge
DARPP32
FOXP1
MAP2
50 µm
% Marker
0
20
40
60
80
100
MAP2+
MAP2+/ FOXP1+
MAP2+/ FOXP1+/ DARPP32+
B
Merge
ASO
FOXP1
MAP2
Vehicle
3 µM ASO
20 µm
C
1 week
0
0.03
0.15
0.3
1.5
3 µM
MSH3
Actin
140
115
40
Target expression (%)
MSH3 ASO dose (µM)
Protein
MSH3
D
1 week
0
0.03
0.15
0.3
1.5
3 µM
MSH6
MSH3
MSH2
Actin
140
115
40
Target expression (%)
MSH3 ASO dose (µM)
Protein
MSH2
MSH6
E
9 weeks
0
0.022
0.26
3 µM
MSH3
Actin
140
115
40
Target expression (%)
MSH3 ASO dose (µM)
Protein
MSH3
F
Target expression (%)
Weeks
Treatment
Vehicle
MSH3 ASO 0.022 µM
MSH3 ASO 0.26 µM
MSH3 ASO 3 µM
SCR ASO 3 µM

Fig. 1. Dose titration of MSH3 in MSN-enriched neuronal cultures using an *MSH3* ASO. (**A**) Immunofluorescent staining of 125 CAG iPSC-derived neuronal cultures at day 36 of differentiation protocol. The co-staining of DARPP32 (green), FOXP1 (orange), MAP2 (red) and Hoechst 33342 (blue) indicates successful production of MSNs. Quantification of marker positive cell frequency for MAP2, FOXP1 and DARPP32 as proportion of the total cells counterstained with Hoechst. (**B**) Immunofluorescent staining of the ASO (green, with white arrows indicating ASO foci), FOXP1 (orange), MAP2 (red) and Hoechst 33342 (blue), following 1 week of treatment. (**C-E**) (Upper) Representative western blot showing dose titration of (C, D, E) MSH3 and (D) MSH2 and MSH6 in response to (C, D) 1 week or (E) 9 weeks of *MSH3* ASO treatment. (Lower) Quantification of protein levels, normalised to actin and presented relative to the Vehicle treated control (0 µM) (N = 2-3, n = 3). (**F**) qRT-PCR assessing *MSH3* mRNA levels at every time-point throughout the long-term time-course. N = biological replicates *i.e.* differentiation of each clone; n = technical replicates *i.e.* number of samples taken from each differentiation at each timepoint.

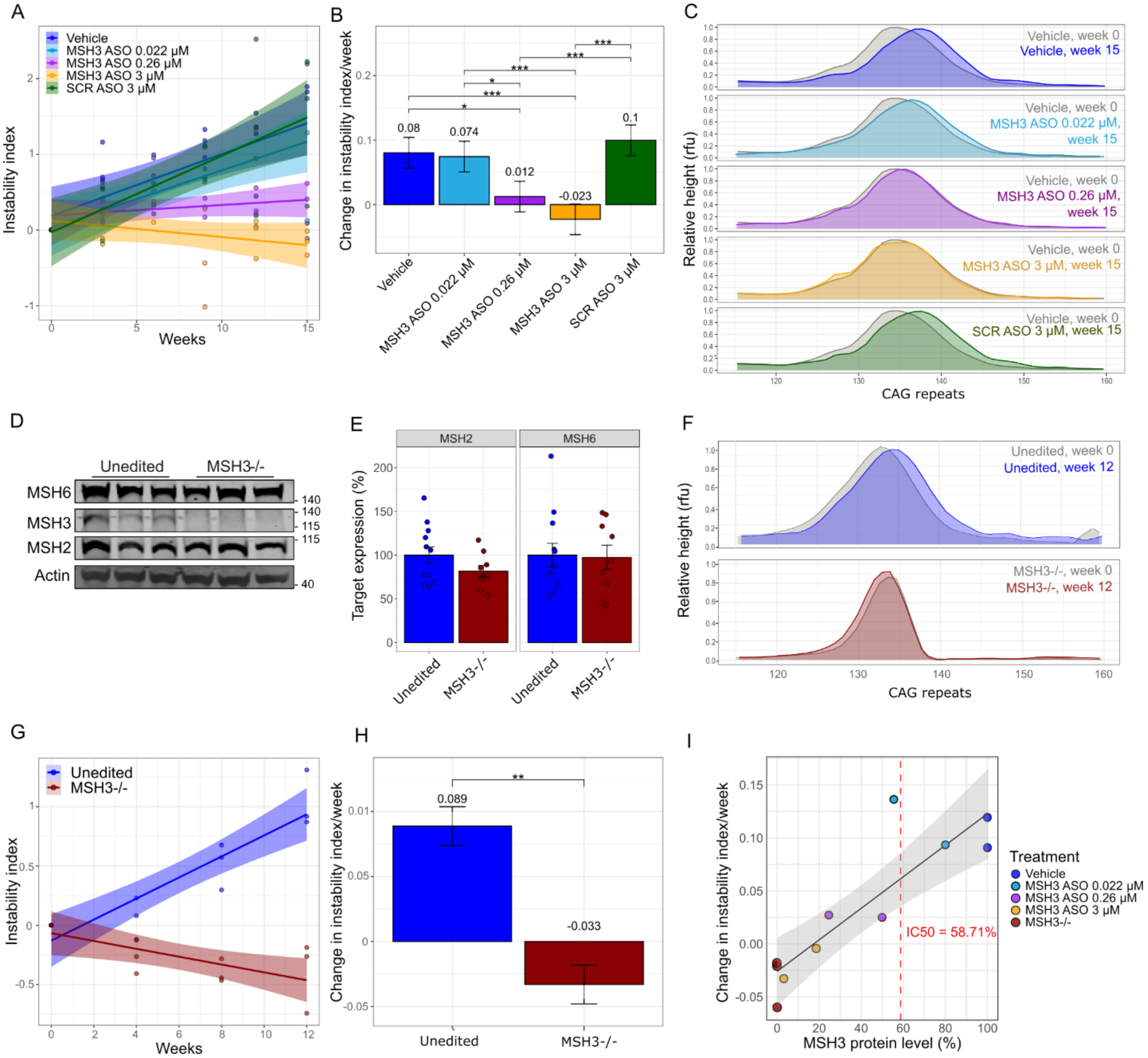

Fig. 2. Fragment analysis to assess the effect of MSH3 lowering on somatic instability. (A) Change in CAG length over time in 125 CAG MSN-enriched cultures treated with varying doses of *MSH3* ASO (sky blue = 0.022 µM, purple = 0.26 µM and orange = 3 µM) as well as Vehicle (PBS, blue) and 3 µM SCR ASO (green) controls. Each data point shows an independent biological replicate, calculated from the mean of three technical replicates (N = 4, n = 3). The shaded regions indicate the 95% confidence intervals. (**B**) The rate of change in CAG per week was determined from a linear mixed effects model (lmer(instability_metric ~ treatment * time + (1 + time | differentiation)) by the estimated marginal means method, with differences between

conditions adjusted for multiple testing using the Bonferroni method (N = 4, n = 3). (**C**) Representative fragment length traces (normalised to modal peak height): Baseline (grey), 15 weeks treatment with Vehicle (blue), 15 weeks treatment with 0.022 µM (sky blue), 0.26 µM (purple) or 3 µM *MSH3* ASO (orange) and 15 weeks treatment with 3 µM SCR ASO (green). (**D**) Representative western blot of MSH3, MSH2 and MSH6 following CRISPR-Cas9 editing of *MSH3* in 125 CAG iPSCs. Each lane shows a separate clone. (**E**) Quantification of MSH2 and MSH6 protein levels, normalised to actin and presented relative to the average unedited cells (N = 9–12, n = 1) (**F**) Representative traces (normalised to modal peak height): Baseline (grey), 125 CAG unedited at 12 weeks (blue) and *MSH3*$^{-/-}$ at 12 weeks (dark red). (**G**) Effect of *MSH3* knock-out on somatic instability. Change in CAG length over time in 125 CAG unedited and *MSH3*$^{-/-}$ MSN-enriched cultures. Each data point shows an independent biological replicate, calculated from the mean of three technical replicates (N = 4, n = 3). The shaded regions indicate the 95% confidence intervals. (**H**) The rate of change in CAG per week was determined from a linear mixed effects model (lmer(instability_metric ~ treatment * time + (1 + time | differentiation)) by the estimated marginal means method (N = 4, n = 3). (**I**) Regression analysis of the relationship between MSH3 protein abundance and CAG expansion rate in HD MSNs. MSH3 protein levels were determined by densitometry of western blots, and instability rates represent the change in instability index per week, relative to the baseline sample in each differentiation and genotype. Datapoints are coloured by treatment group and genotype; Vehicle and *MSH3* ASO dosing groups had an unedited genetic background. The black line represents the fitted regression. The shaded region indicates the 95% confidence intervals. The IC50 is depicted by a red dashed vertical line at an MSH3 protein abundance 58.7% of Vehicle. It was calculated from the regression model coefficients and represents the MSH3 protein abundance at which the CAG expansion rate is halved. *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$. N = biological replicates *i.e.* differentiation of each clone, n = technical replicates *i.e.* number of samples taken from each differentiation at each timepoint.

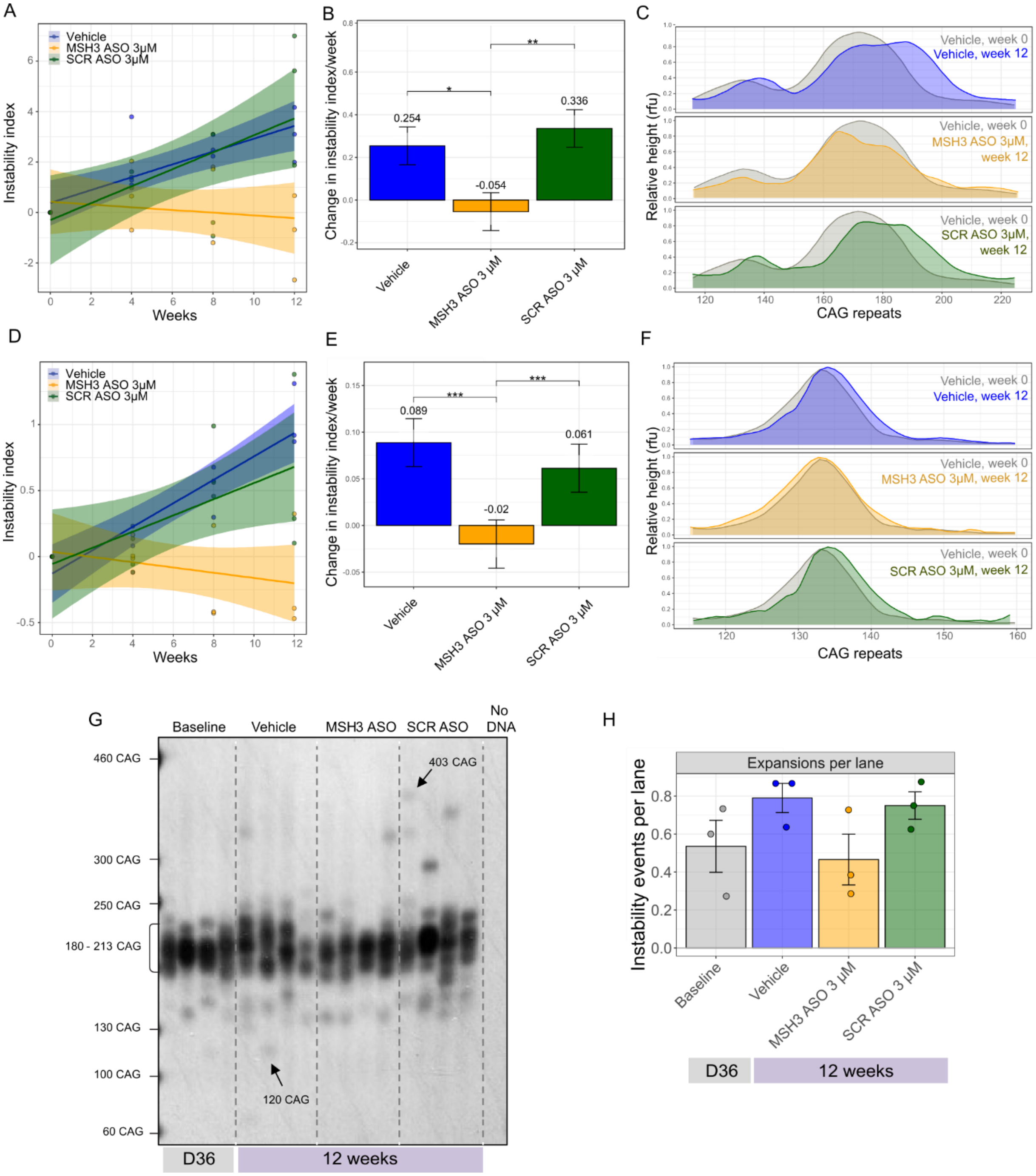

Fig. 3. The effects of MSH3 lowering on CAG repeat instability in 125 CAG *FAN1*$^{-/-}$ MSN-enriched cultures. (**A-C**) 125 CAG *FAN1* $^{-/-}$ MSN-enriched cultures (N = 4, n = 3). (**D-F**) 125 CAG unedited MSN-enriched cultures (N = 3, n = 3), derived from a 125 CAG iPSC clone which had undergone the CRISPR-editing process with no successful cutting. (A, D) Change in CAG length over time in cultures treated with Vehicle (PBS, blue), 3 µM *MSH3* ASO (orange) and 3

μM SCR ASO (green). Each data point shows an independent biological replicate, calculated from the mean of three technical replicates (N = 3-4, n = 3). The shaded regions indicate the 95% confidence intervals. (B, E) The rate of change in CAG per week was determined from a linear mixed effects model (lmer(instability_metric ~ treatment * time + (1 + time | differentiation)) by the estimated marginal means method, with differences between conditions adjusted for multiple testing using the Bonferroni method. (C, F) Representative traces: Baseline (day 36, grey), 12 weeks of treatment post-baseline with Vehicle (blue), 3 μM *MSH3* ASO (orange) and 3 μM SCR ASO (green). (**G**) Small-pool PCR to measure the effect of *MSH3* ASO treatment on large, rare expansion events in 125 CAG *FAN1*$^{-/-}$ MSN-enriched cultures at baseline (day 36) or following Vehicle, SCR ASO or MSH3 ASO treatment for 12 weeks post-baseline (N = 3, n = 3, with 16 separate PCRs assessed per sample). The length of the most extreme alleles was measured by GelAnalyzer. (**H**) Frequency of large expansions per lane, with large expansions determined as those falling outside the modal spread of alleles. Each data point shows an independent biological replicate. *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$. N = biological replicates *i.e.* differentiation of each clone, n = technical replicates *i.e.* number of samples taken from each differentiation at each timepoint.

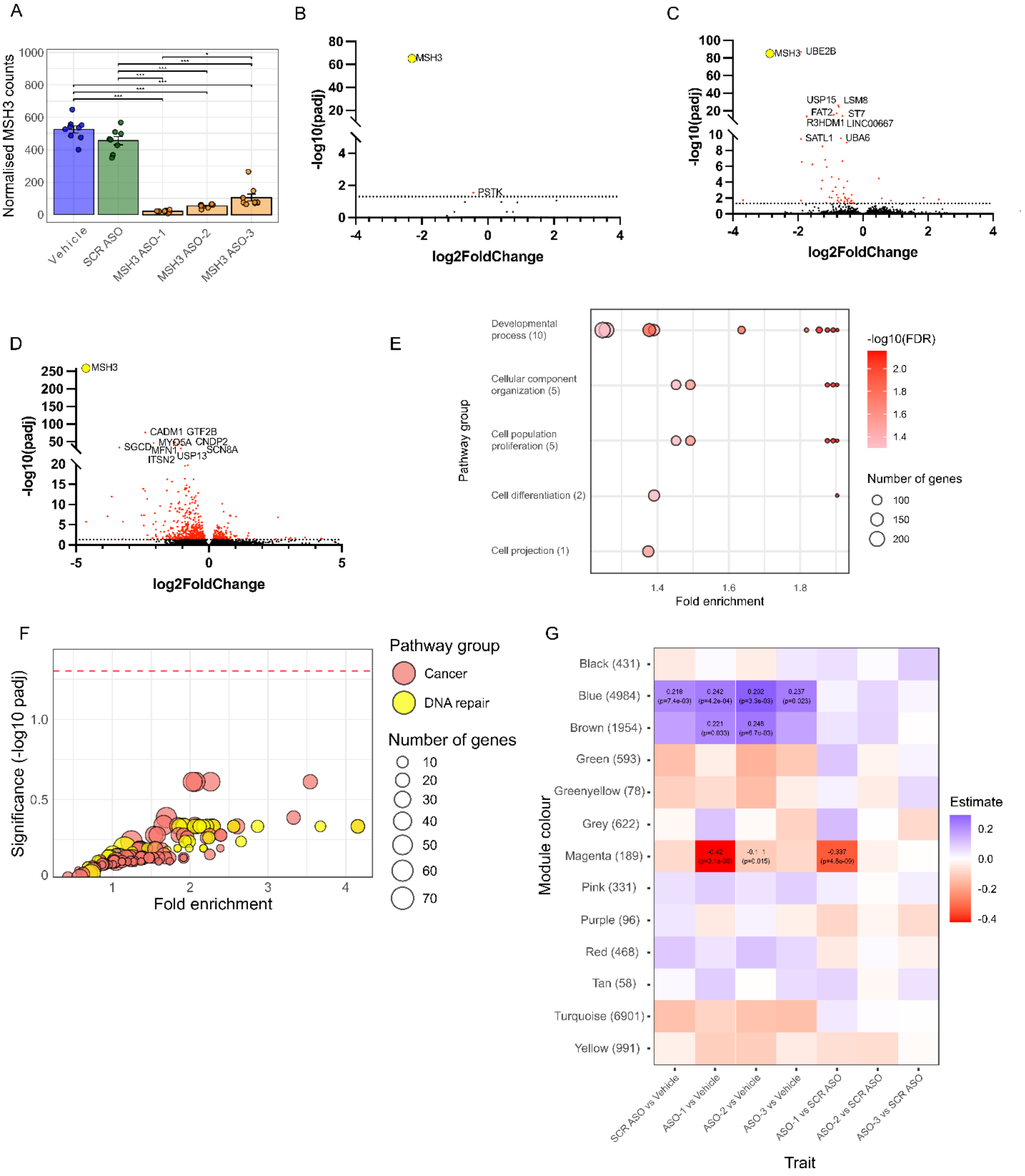

Fig. 4. Bulk RNA-Seq to assess the effect of MSH3 ASO treatment on the global transcriptome. (**A**) Normalised read counts of *MSH3* in Vehicle, SCR, *MSH3* ASO-1, *MSH3* ASO-2 and *MSH3* ASO-3-treated samples. (**B-D**) Volcano plots of DEGs (red) and non-DEGs

(black) for *MSH3* ASOs vs SCR ASO treated neurons: (B) *MSH3* ASO-3, (C) *MSH3* ASO-2, (D) *MSH3* ASO-1. The horizontal black dashed line represents padj = 0.05. The top 10 DEGs are labelled, with *MSH3* highlighted in yellow. The full list of DEGs is available in Data File S1. (**E**) Nodal pathway plots summarising pathways dysregulated in *MSH3* ASO-1 vs SCR ASO treated samples. The size of the bubble depends on the number of genes enriching the GO term. The full details for the GO enriched terms are available in Data File S2. (**F**) Gene set enrichment analysis of DEGs in *MSH3* ASO-1 comparisons using NCG and custom DNA repair and cancer pathways. Fold enrichment was plotted against enrichment significance (-log10p adjusted), with a red dashed horizontal line representing padj = 0.05 significance. *MSH3* ASO-2 and -3 are not included in the figure due to either insufficient DEGs to perform analyses (ASO-3) or no significant results (ASO-2). (**G**) Heatmap showing correlation of WGCNA gene expression modules with *MSH3* ASO or SCR ASO treatment compared to Vehicle in iPSC-derived MSN cultures. Modules on the y-axis are ranked by their significance, with arbitrary colours assigned as names for each module and numbers in brackets indicating gene counts. Colour of cells within heatmap reflect the correlation strength and direction between each module's gene signature and the treatment: red for negative and blue for positive. Correlation coefficients and p-values are displayed where significant. Full details are in Data File S3**.**

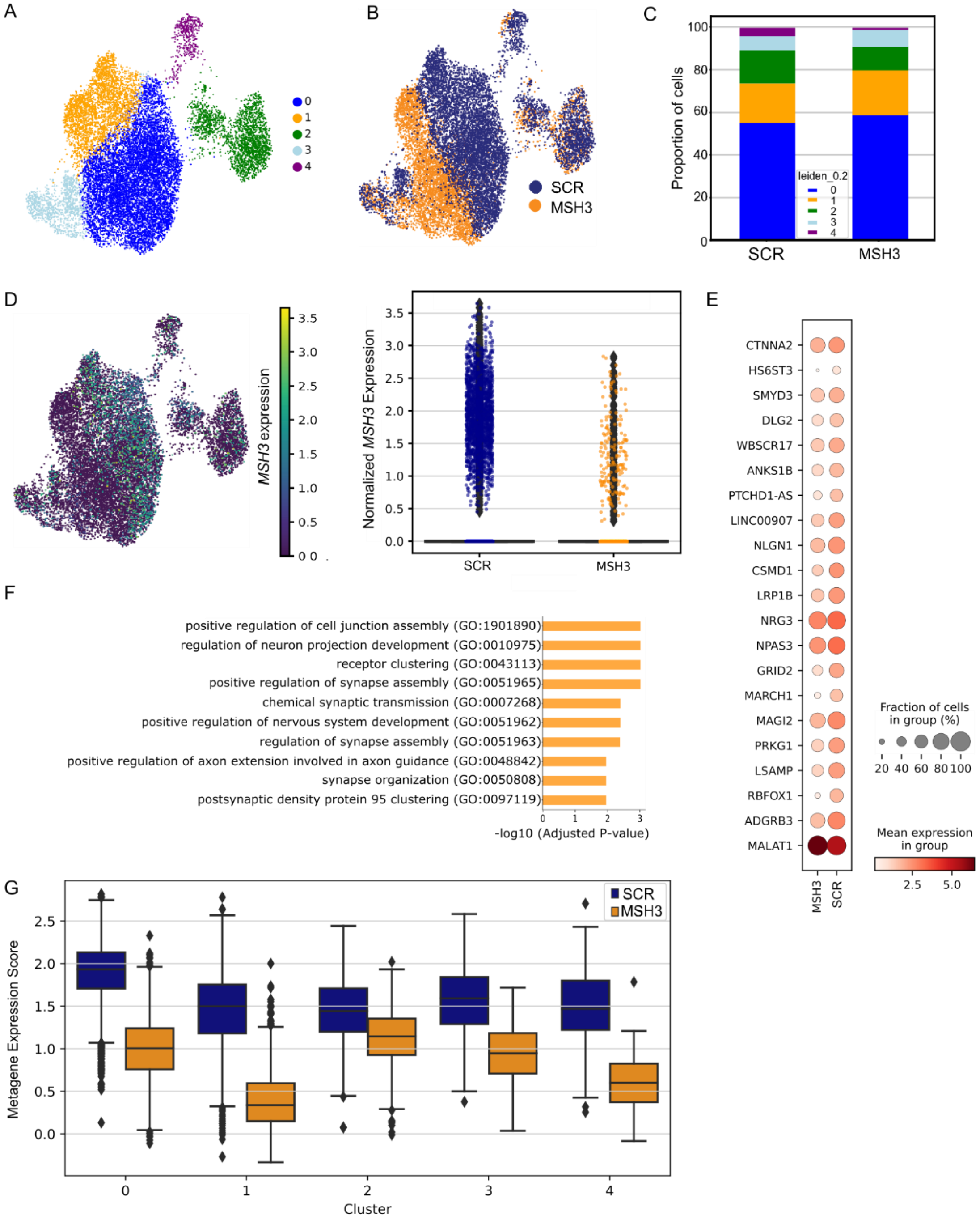

Fig. 5. snRNA-Seq reveals no cell-type specific effects of *MSH3* ASO treatment. (**A, B**) Uniform manifold approximation and projection of snRNA-Seq data from *MSH3* and SCR ASO treatment in neurons labelled by (A) Leiden unsupervised clustering, and (B) treatment. (**C**) Barplot showing proportion of cells from each treatment category in each cluster. (**D**) (Left)

UMAP labelled by *MSH3* expression. (Right), jitter plot displaying *MSH3* expression. In both cases, *MSH3* expression is normalised for library size and log1p transformed. (**E**) Dotplot of top differentially expressed genes between *MSH3* ASO and SCR ASO treated neurons. (**F**) Gene set enrichment for downregulated genes following exposure to *MSH3* ASO. (**G**) Distribution of per-cell differential expression effect of *MSH3* ASO treatment across Leiden clusters.

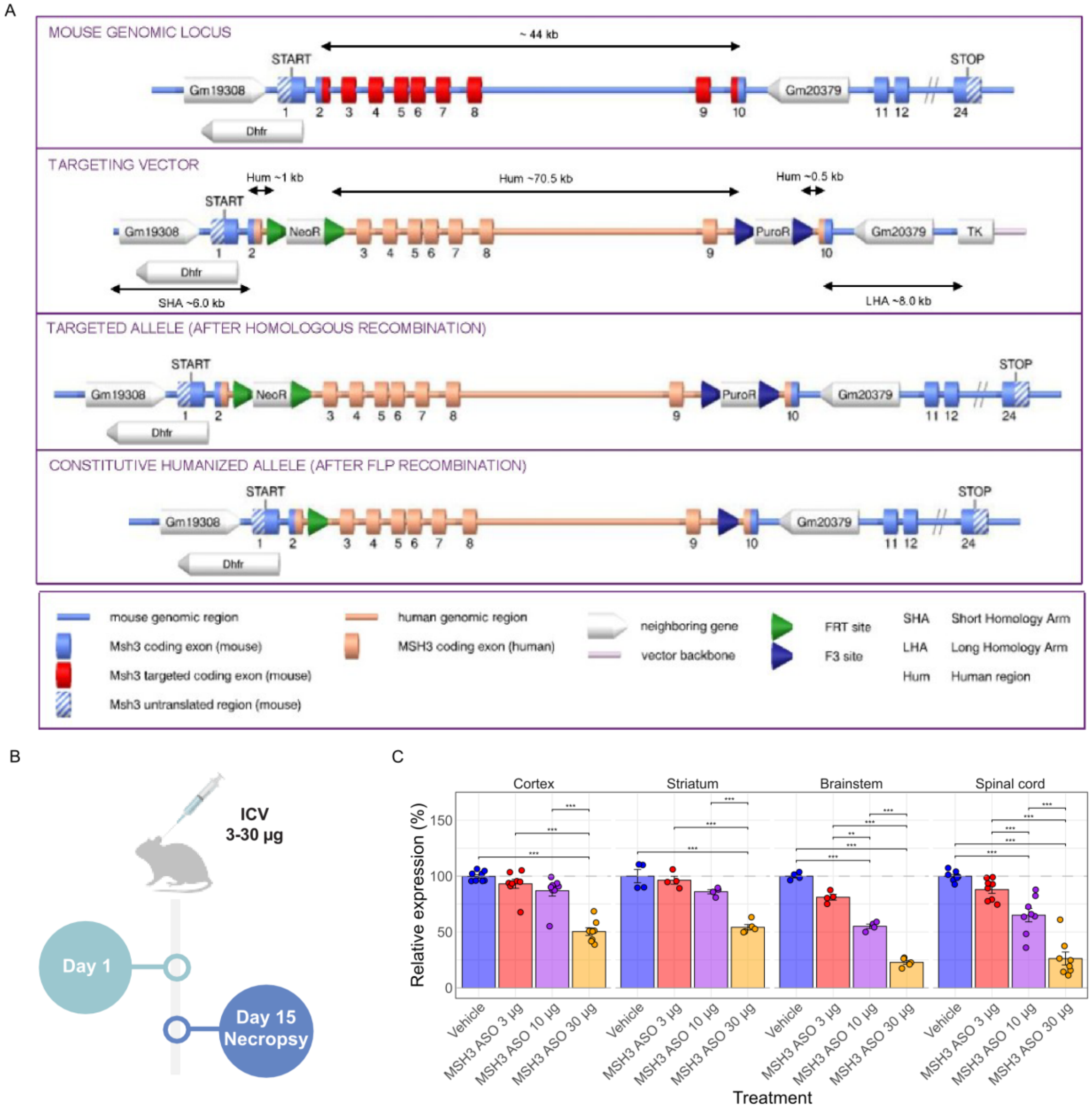

Fig. 6. *in vivo* target engagement of the *MSH3* ASO. (**A**) Schematic of humanisation of MSH3 exon 2-10 via homologous recombination and FLP combination. (**B**) Schematic describing treatment of *MSH3* knock-in mice with PBS (Vehicle, N = 8) or 3 µg (N = 8), 10 µg (N = 8) or 30 µg (N = 8) of the *MSH3* ASO via ICV, and tissue harvesting 14 days later. (**C**) qRT-PCR of *MSH3* in tissue samples harvested from mice 14 days following treatment with PBS (Vehicle, N = 8), or 3 µg (N = 8), 10 µg (N = 8) or 30 µg (N = 8) of the *MSH3* ASO via ICV. Data were analysed using a two-way ANOVA and comparisons were adjusted for multiple testing using emmeans. **** = $p<0.0001$, *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$.

Supplementary figures

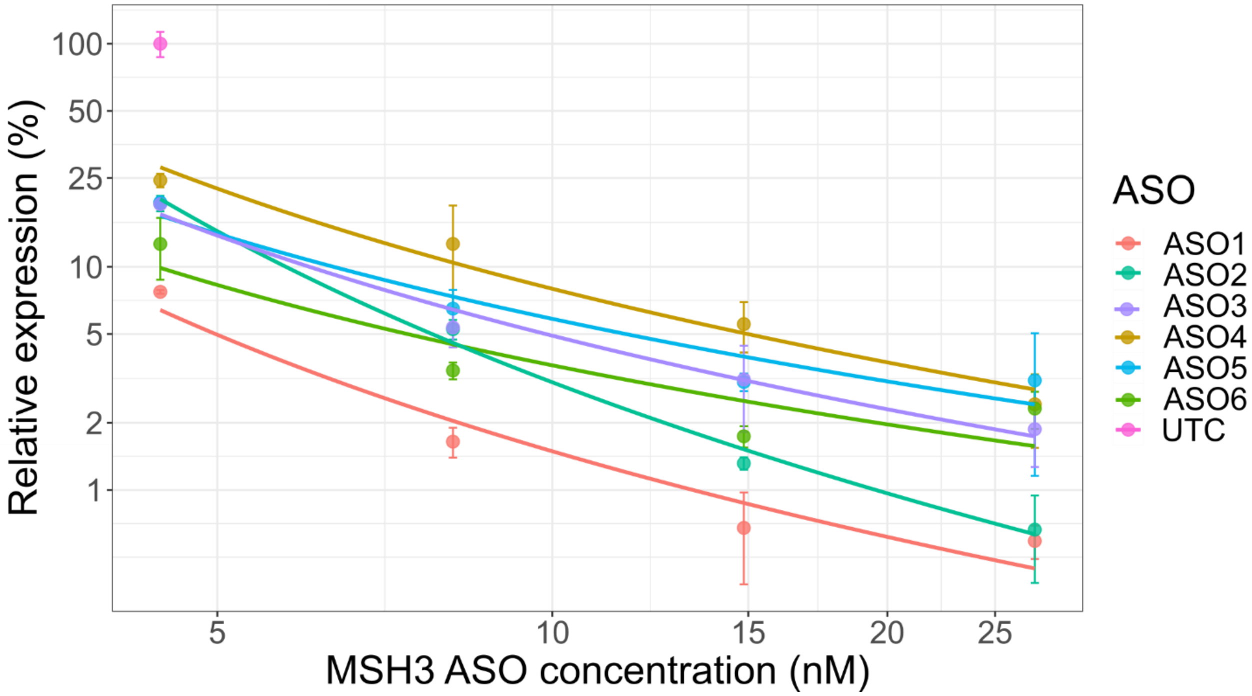

Fig. S1. *in vitro* screen of *MSH3* ASOs. qRT-PCR assessing *MSH3* mRNA levels in A-431 cells following 48 hours treatment with the top six most potent *MSH3* ASOs. ASO-1 (herein referred to as the *MSH3* ASO) was used for all further *in vitro* and *in vivo* characterization work. ASO-2 and ASO-3 were additionally included in the bulk RNA-sequencing experiments, to elucidate ASO-specific vs MSH3 lowering effects.

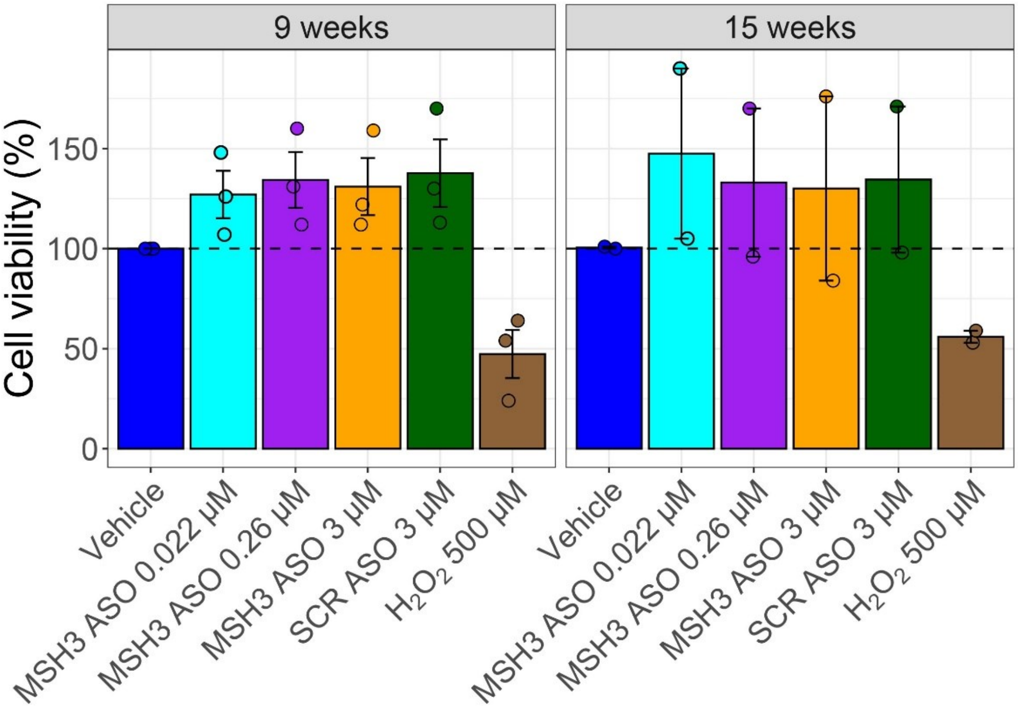

Fig. S2. Effect of *MSH3* ASO treatment on cell viability. Cell viability of 125 CAG MSN-enriched cultures treated with varying doses of *MSH3* ASO and maximal dose (3 µM) of SCR ASO, relative to Vehicle, as measured by AlamarBlue™ assay after 9 weeks and 15 weeks of continuous treatment. Hydrogen peroxide (H_2O_2) treatment at 500 µM was used as a positive control.

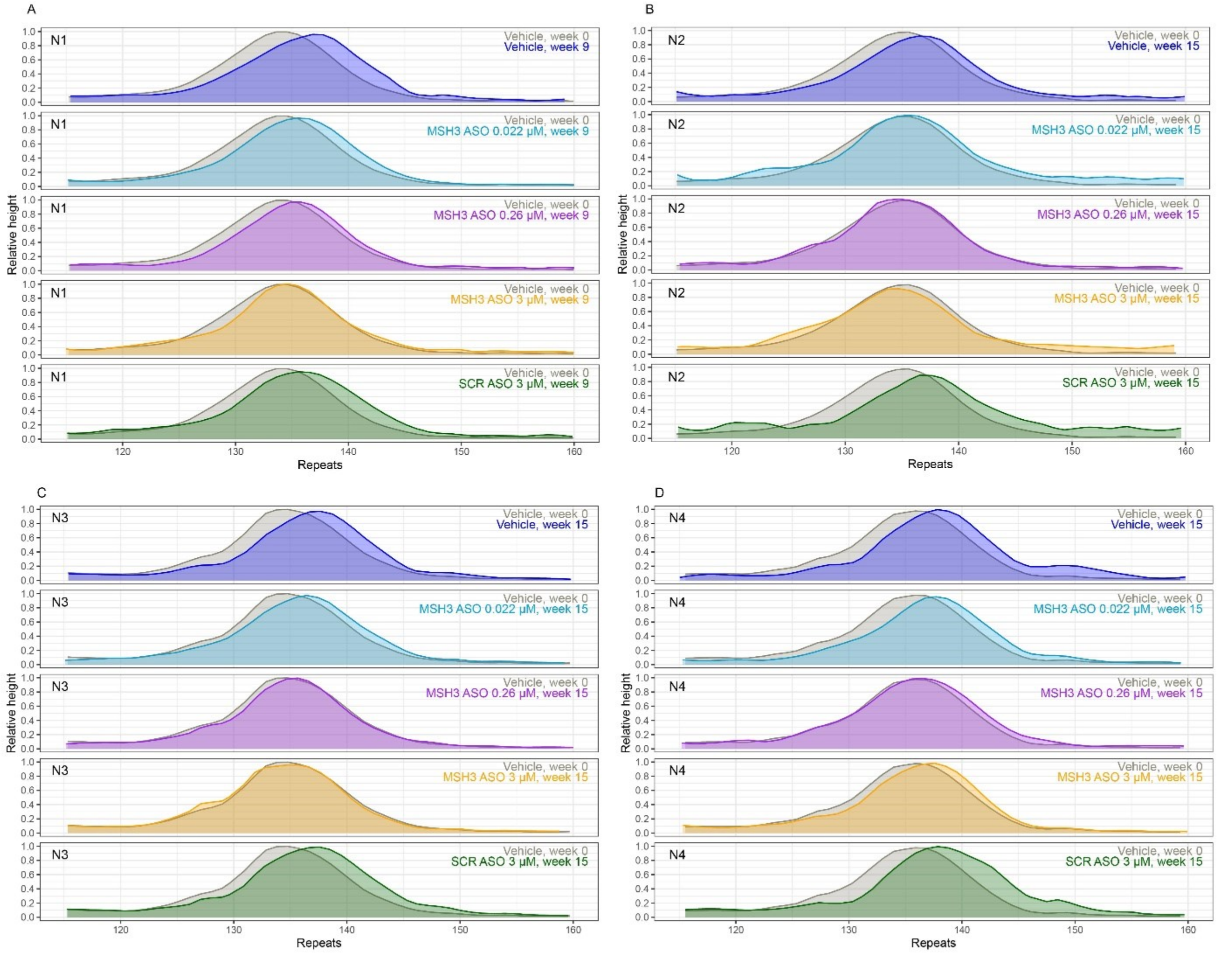

Fig. S3. Fragment analysis to assess the effect of MSH3 lowering on somatic instability. (**A-D**) Representative fragment length traces for each biological replicate (*i.e*. differentiations 1-4): Baseline (day 36, grey), and end-point (differentiation N1 = 9 weeks, N2, 3, and 4 = 15 weeks) after treatment with Vehicle (blue), 0.022 µM (red), 0.26 µM (purple) or 3 µM *MSH3* ASO (orange), or 3 µM SCR ASO (green). Relative height is normalised to modal peak height.

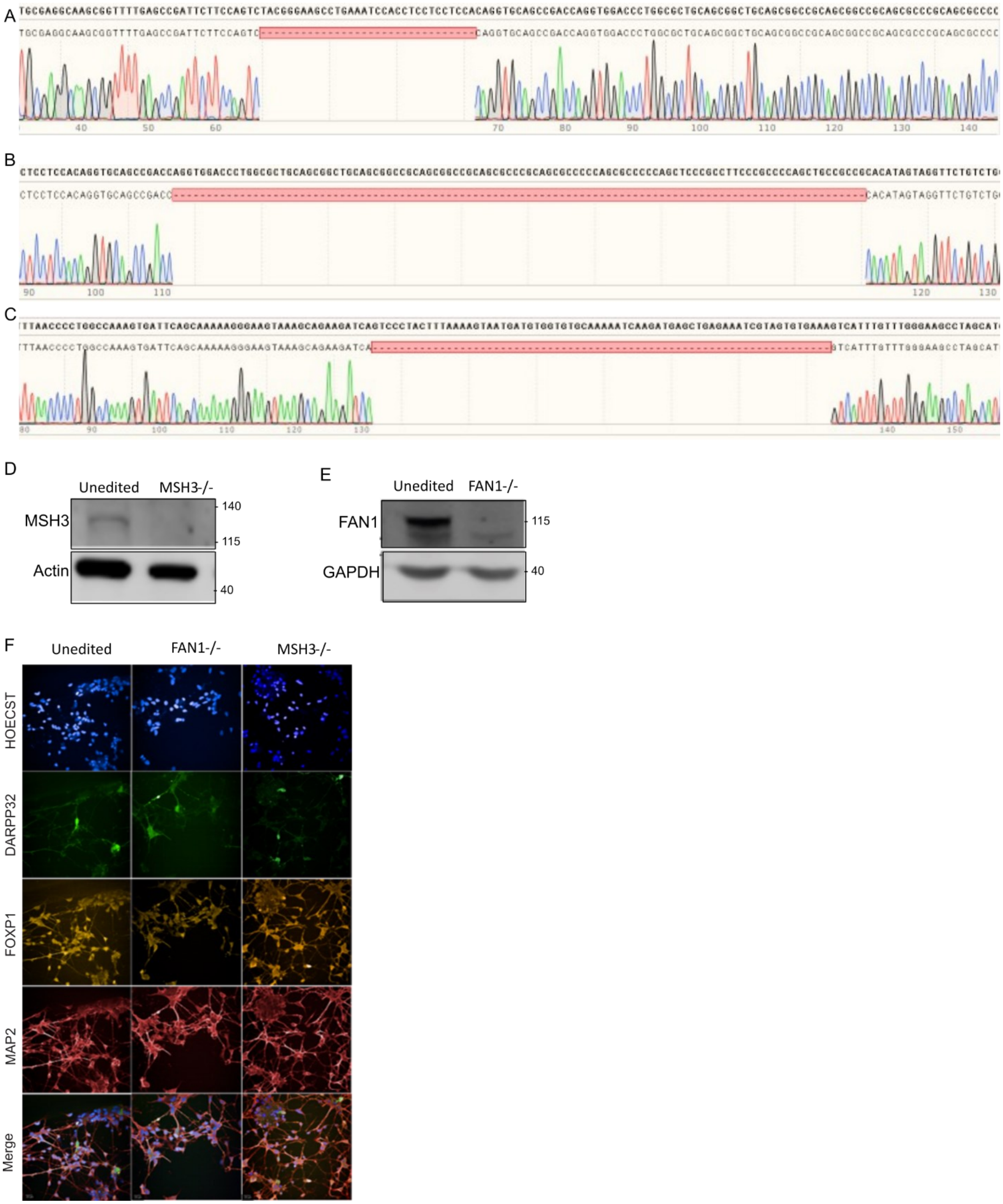

Fig. S4. Generation of *MSH3*$^{-/-}$ and *FAN1*$^{-/-}$ 125 CAG iPSC lines. (**A**, **B**) Sanger sequencing results of a successful 125 CAG *MSH3*$^{-/-}$ iPSC clone with (A) a 32 bp deletion or (B) a 104 bp deletion, generated using two different sets of guide RNAs. (**C**) Sanger sequencing results of a successful 125 CAG *FAN1*$^{-/-}$ iPSC clone with a 68 bp deletion. Dashed red line denotes the site of the deletion event. (**D**, **E**) Western blot confirms loss of (D) MSH3 or (E) FAN1 following CRISPR-editing. (**F**) Immunofluorescent staining of MSN-enriched cultures at day 36, generated from *MSH3*$^{-/-}$, *FAN1*$^{-/-}$ and unedited 125 CAG lines, stained with MAP2 (post-mitotic neuronal marker), FOXP1 (lateral ganglionic eminence marker), DARPP32+ (MSN-specific marker, when co-staining with FOXP1).

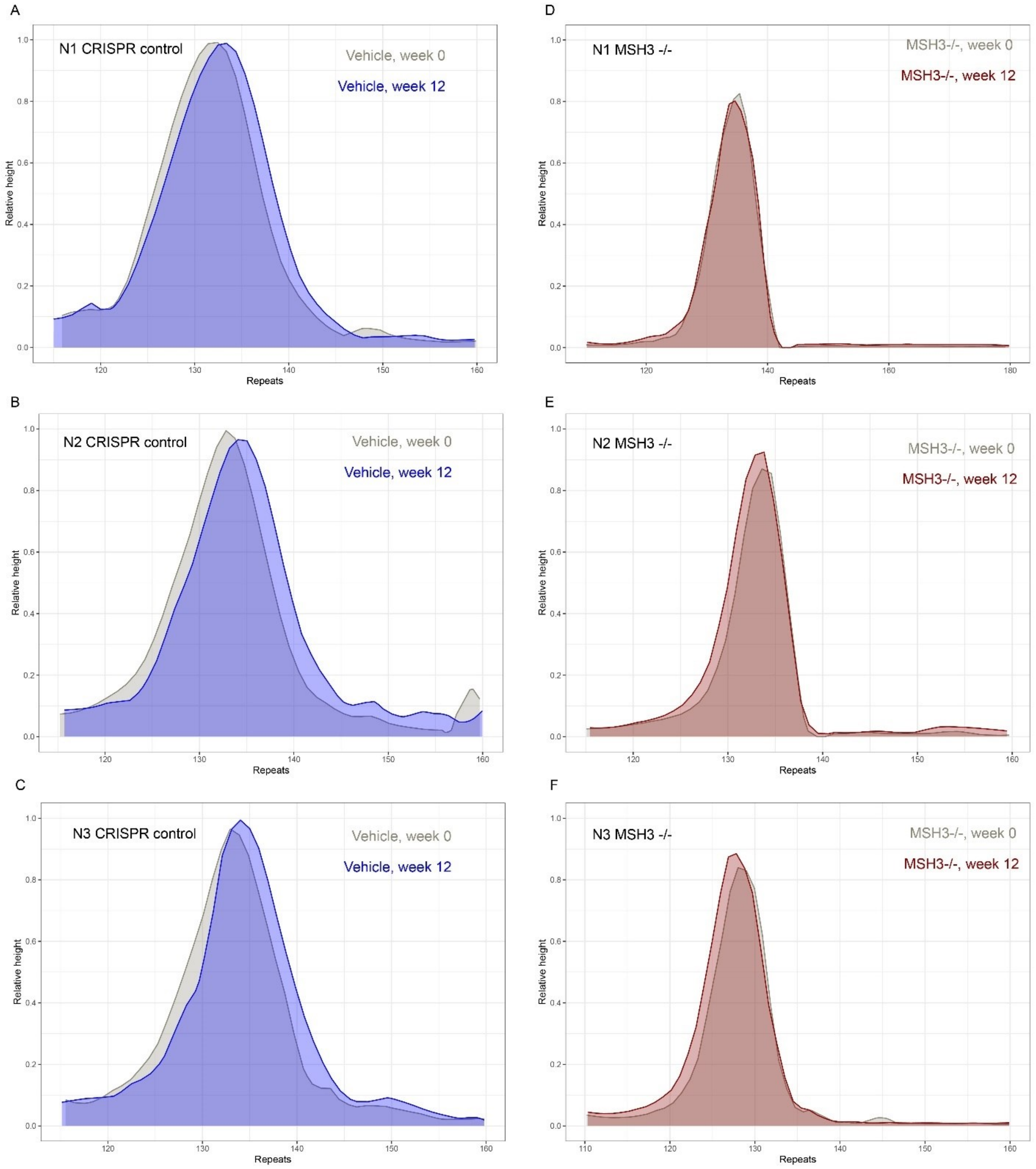

Fig. S5. Fragment analysis to assess the effect of MSH3 knock-out on somatic instability. (**A-C**) Representative fragment length traces for each biological replicate (*i.e.* differentiation 1-3) of 125Q CAG unedited at baseline (grey) and 12 weeks (blue). (**D-F**) Representative fragment length traces for each biological replicate (*i.e.* differentiation 1-3) of MSH3$^{-/-}$ at baseline (grey) and 12 weeks (dark red). Relative height is normalised to modal peak height.

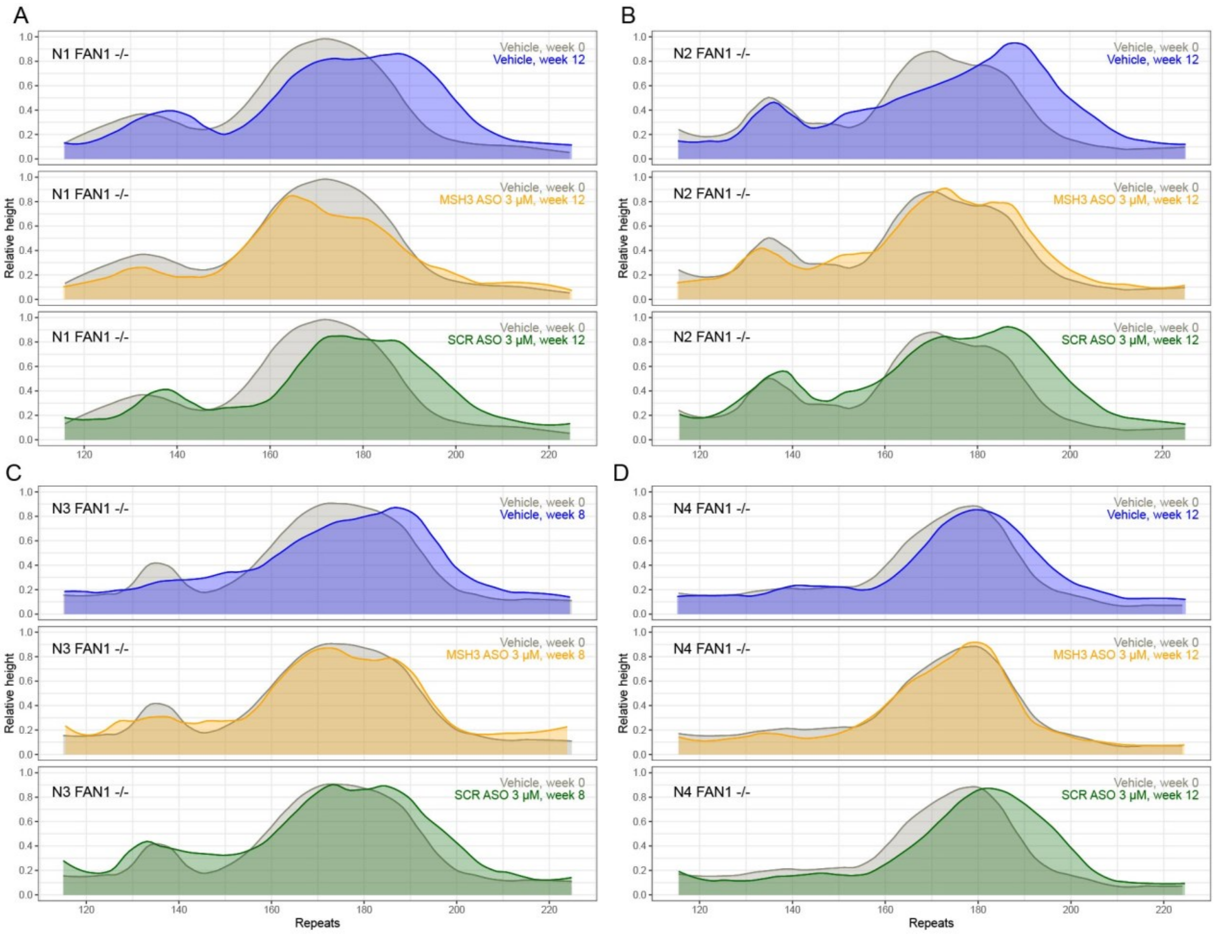

Fig. S6. Fragment analysis to assess the effects of MSH3 lowering on somatic instability in 125 CAG *FAN1*$^{-/-}$ MSN-enriched cultures. (**A-D**) Representative traces for each biological replicate (*i.e.* differentiation 1-4): Baseline (day 36, grey), 12 weeks of treatment post-baseline with Vehicle (blue), 3 µM *MSH3* ASO (orange) and 3 µM SCR ASO (green). Relative height is normalised to modal peak height.

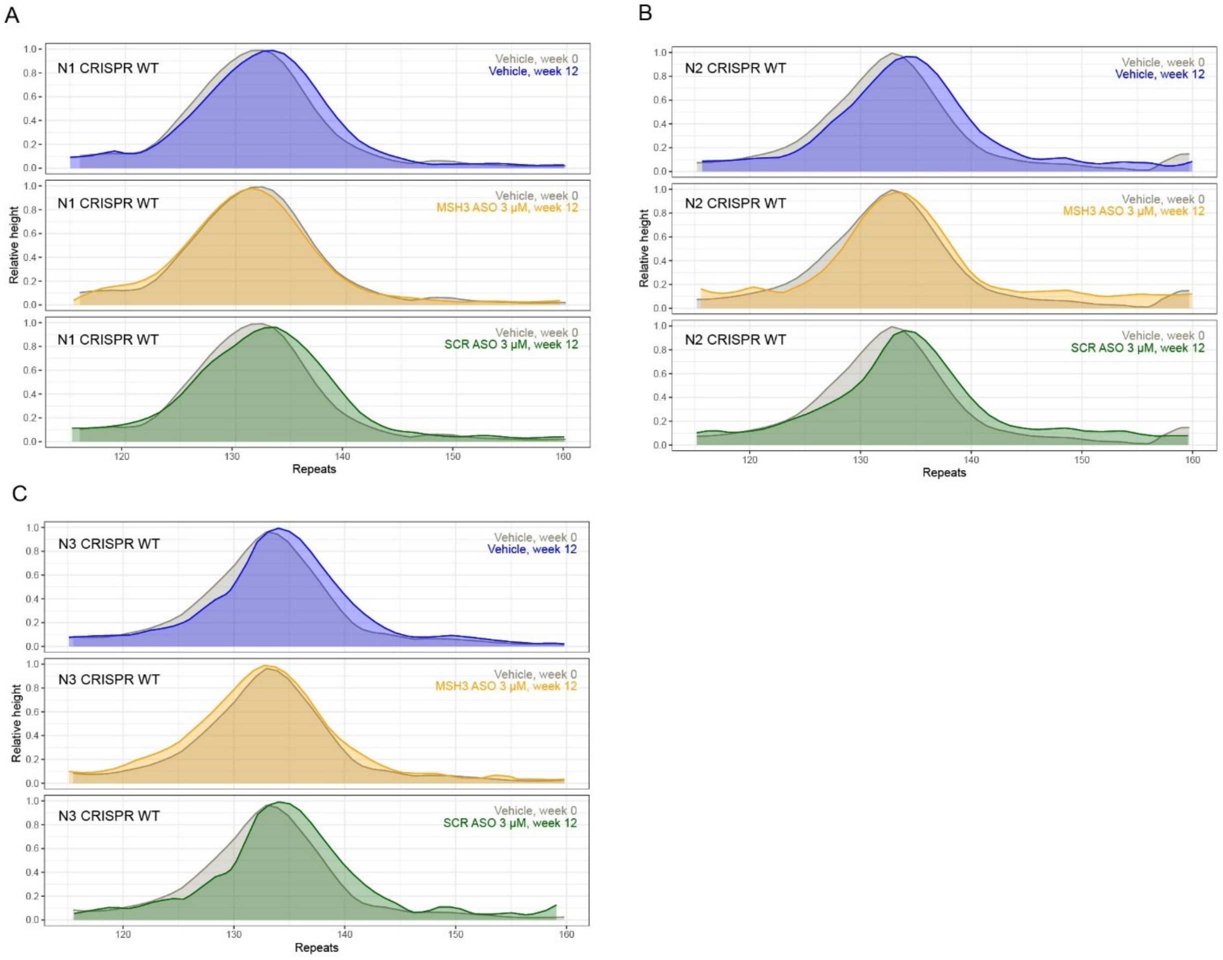

Fig. S7. Fragment analysis to measure the effects of MSH3 lowering on somatic instability in 125 CAG unedited MSN-enriched cultures. (**A-C**) Representative traces of biological replicates (*i.e.* differentiations 1-3): Baseline (day 36, grey), 12 weeks of treatment post-maturity with Vehicle (blue), 3 µM *MSH3* ASO (orange) and 3 µM SCR ASO (green). Relative height is normalised to modal peak height.

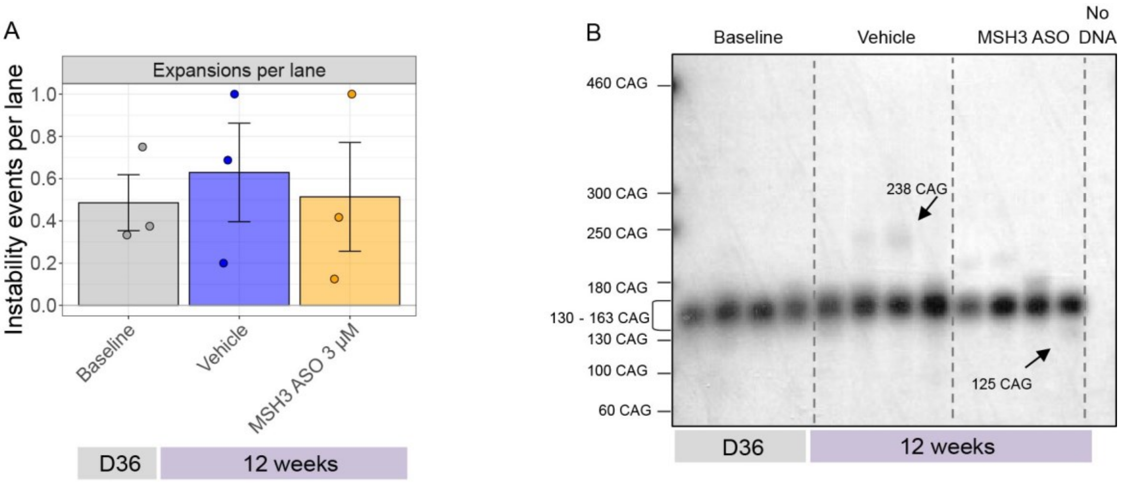

Fig. S8. Small-pool PCR to measure the effect of *MSH3* ASO treatment on large, rare expansion events. 125 CAG unedited MSN-enriched cultures (N = 1 - 3, n = 3, with 16 separate PCR reactions assessed per sample) at baseline (day 36) or following Vehicle, SCR ASO or *MSH3* ASO treatment for 12 weeks post-baseline. (**A**) Large expansions per lane (**B**) representative autoradiograph. The length of the most extreme alleles was measured by GelAnalyzer. Each data point shows the mean ± SEM of the three biological replicates. Differences between conditions were assessed by one-way ANOVA with no significant differences found. N = biological replicates *i.e.* differentiation of each clone, n = technical replicates *i.e.* number of samples taken from each differentiation at each timepoint.

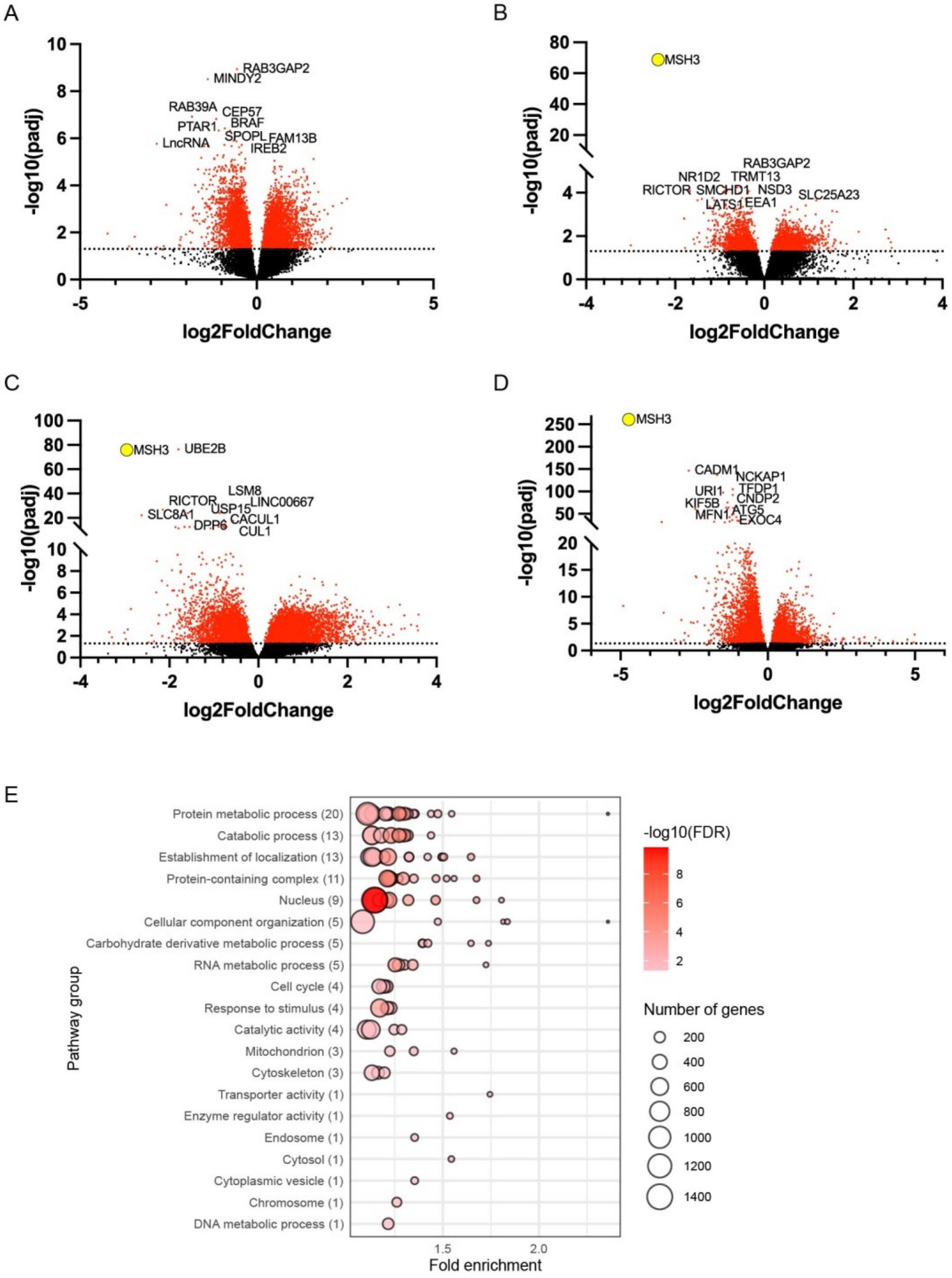

Fig. S9. Bulk RNA-Seq analysis of the global transcriptome following treatment with ASOs vs Vehicle for 2 weeks post-baseline. (**A-D**) Volcano plot of DEGs (red) and non-DEGs (black) for ASOs vs Vehicle treated neurons: (A) SCR ASO, (B) *MSH3* ASO-3, (C) *MSH3* ASO-2 (D) *MSH3* ASO-1. The horizontal black dashed line represents padj = 0.05. The top 10 DEGs are labelled, with *MSH3* highlighted in yellow. The full list of

DEGs is available in Supplementary Dataset 1. (**E**) Nodal pathway plot showing top 20 pathway groups dysregulated for SCR ASO vs Vehicle treated neurons. The size of the bubble depends on the number of genes enriching the GO term. The full details for the GO enriched terms are available in Supplementary Dataset 2.

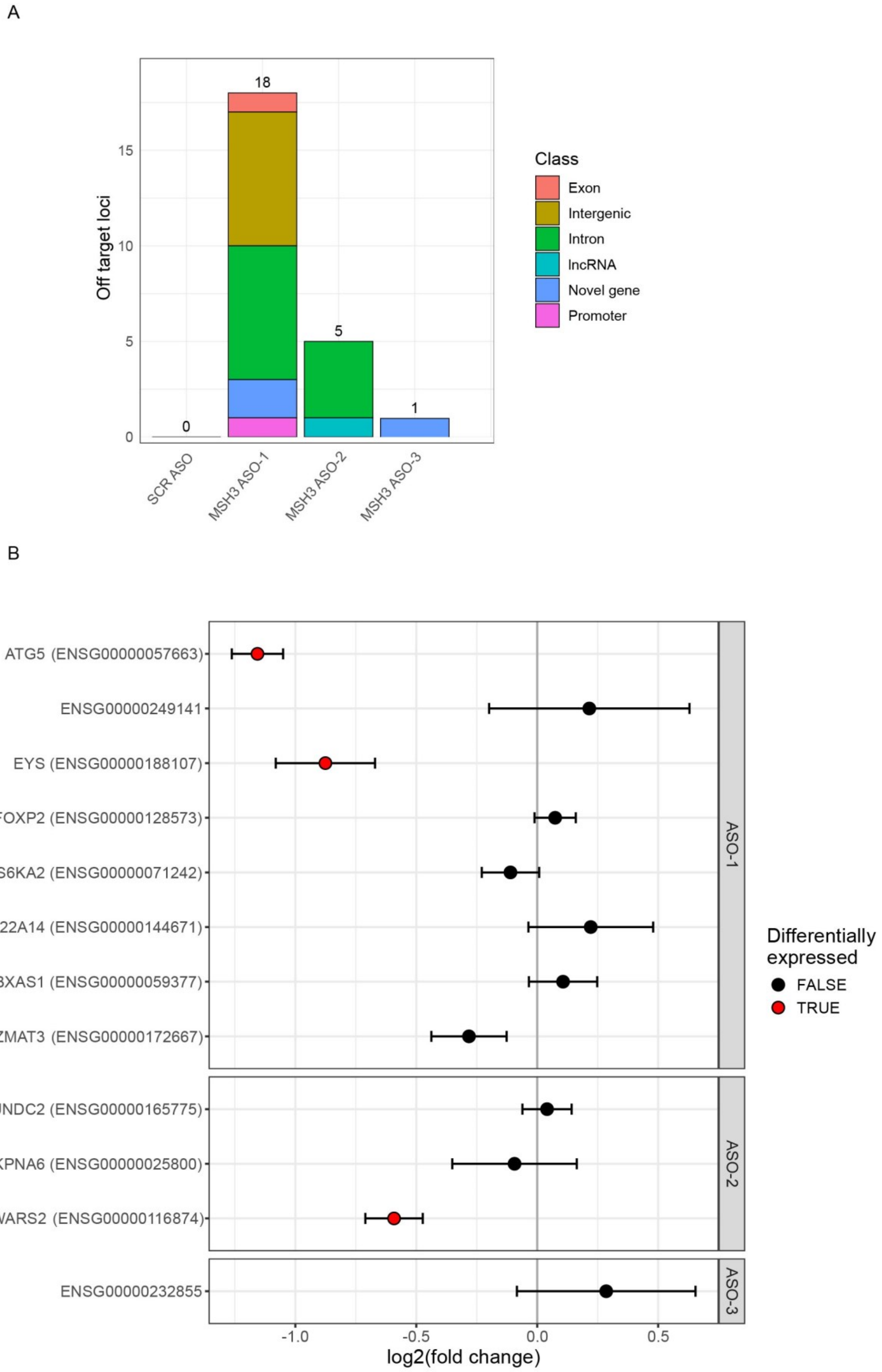
A
Off target loci
0
5
10
15
18
5
1
0
SCR ASO
MSH3 ASO-1
MSH3 ASO-2
MSH3 ASO-3
Class
Exon
Intergenic
Intron
lncRNA
Novel gene
Promoter
B
Predicted off target loci
ATG5 (ENSG00000057663)
ENSG00000249141
EYS (ENSG00000188107)
FOXP2 (ENSG00000128573)
RPS6KA2 (ENSG00000071242)
SLC22A14 (ENSG00000144671)
TBXAS1 (ENSG00000059377)
ZMAT3 (ENSG00000172667)
FUNDC2 (ENSG00000165775)
KPNA6 (ENSG00000025800)
WARS2 (ENSG00000116874)
ENSG00000232855
ASO-1
ASO-2
ASO-3
-1.0
-0.5
0.0
0.5
log2(fold change)
Differentially expressed
FALSE
TRUE

Fig. S10. Predicted off-targets of *MSH3*-targeting ASOs. (**A**) Predicted off-targets for the six *MSH3* ASOs listed in Supplementary Table 1. (**B**) Differential expression of predicted off target loci that were annotated in our dataset following treatment with *MSH3* ASOs. Off target loci for each ASO are plotted separately, with transcript ID and gene name on the y-axis, and log2 (fold change) (lfc) in expression on the x-axis. Points represent lfc, and error bars the standard error (SE). A differential expression analysis employed the DESeq2 package to estimate lfc and their SE for gene expression levels between treatment with the indicated *MSH3* ASO and SCR ASO. Red points indicate genes that were differentially expressed by treatment with the indicated MSH3 ASO, relative to SCR ASO.

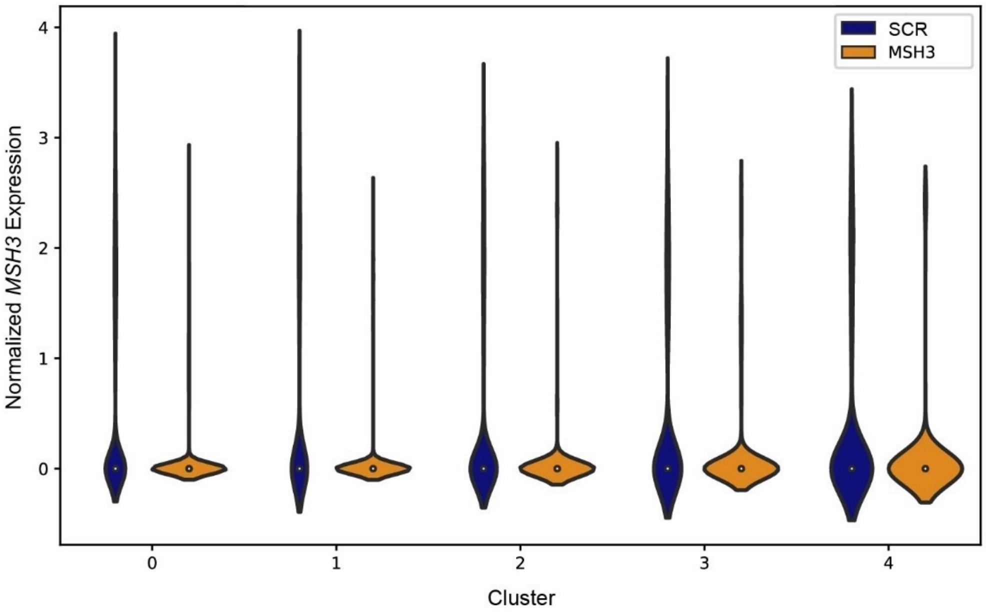

Fig. S11. Effect of *MSH3* ASO on *MSH3* expression across the snRNA-Seq clusters. Violin plot showing *MSH3* expression normalised for library size and log1p transformed for neuronal cultures treated with SCR ASO 3 μM (blue) or *MSH3* ASO 3 μM (orange) for 2 weeks post-baseline.

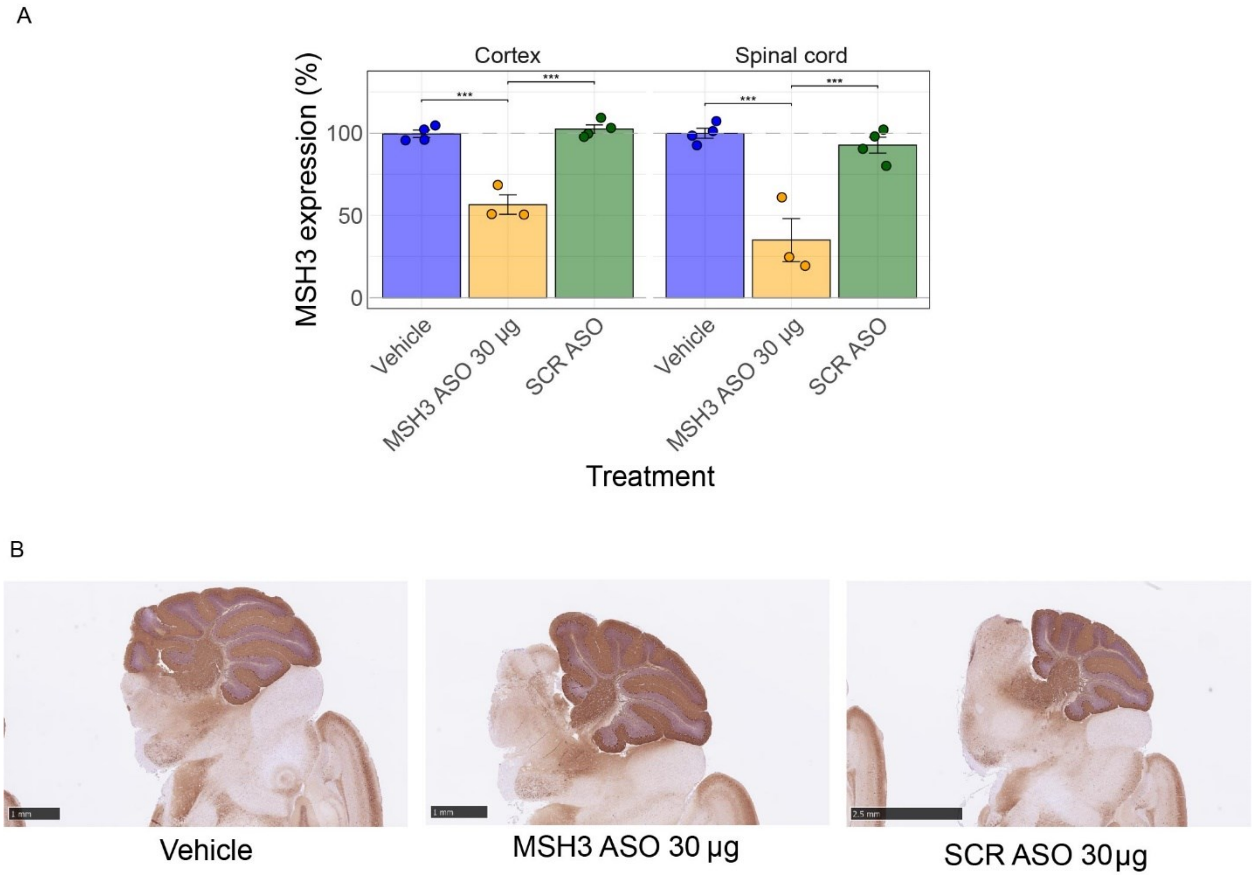

Fig. S12. Effect of ASO treatment on *MSH3* and calbindin *in vivo*. (**A**) qRT-PCR of *MSH3* in tissue samples harvested from mice 14 days after treatment with PBS (Vehicle, n = 4), 30 µg (n = 3) of the *MSH3* ASO or 30 µg of SCR ASO via ICV. Data were analysed using a two-way ANOVA and comparisons were adjusted for multiple testing using emmeans. **** = p<0.0001, *** = p<0.001, ** = p<0.01, * = p <0.05. (**B**) Representative image of Vehicle, SCR ASO (30 µg) or *MSH3* ASO (30 µg) treated humanised *MSH3* knock-in mouse brains stained with calbindin, a marker for neurodegeneration.

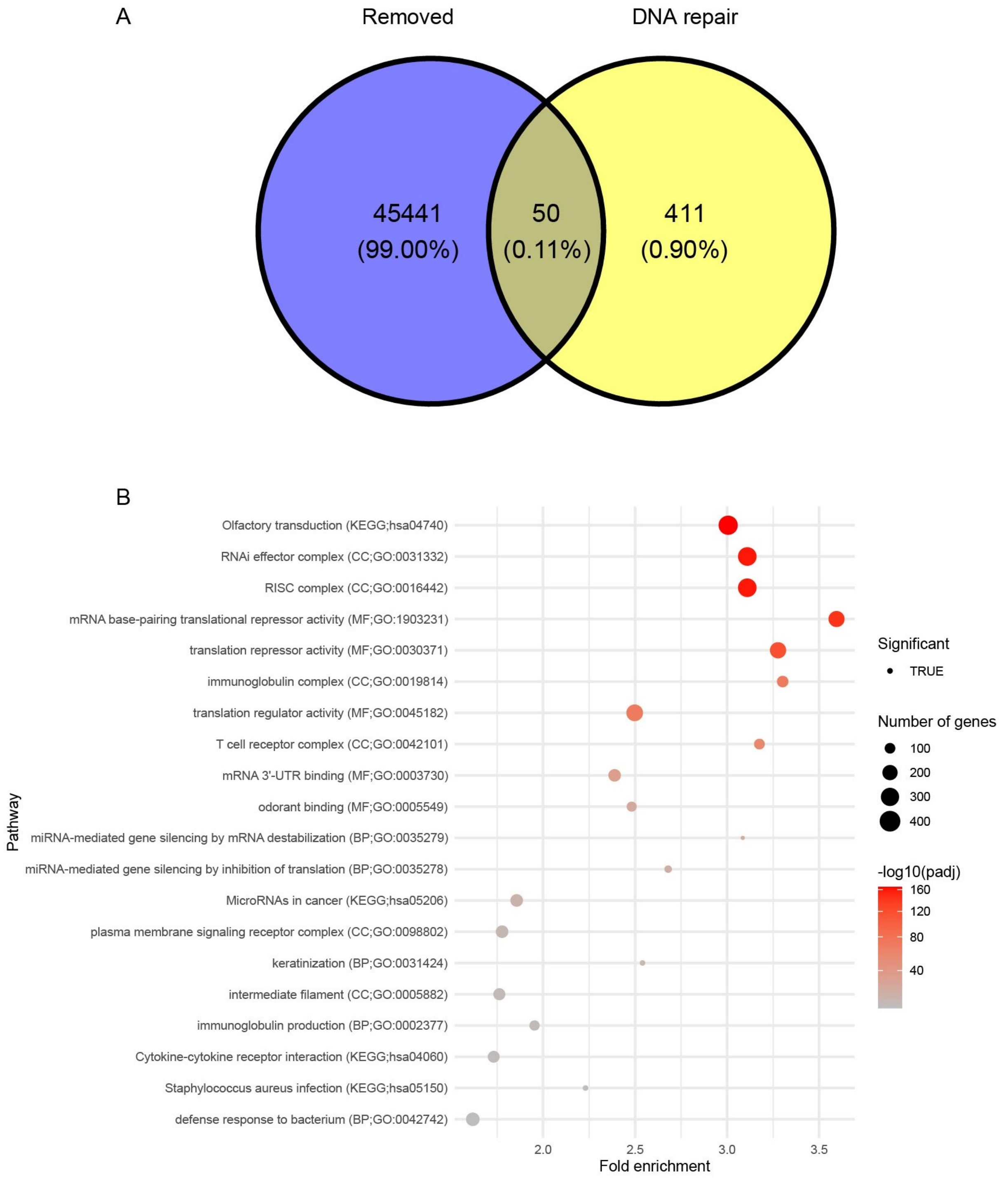

Fig. S13. Overlap analysis between DNA repair genes and transcripts excluded during RNA-Seq quality control. (**A**) Venn diagram illustrating the overlap between transcripts removed for low read count and DNA repair genes. (**B**) Pathway analysis of genes excluded for low read counts in WGCNA analysis.

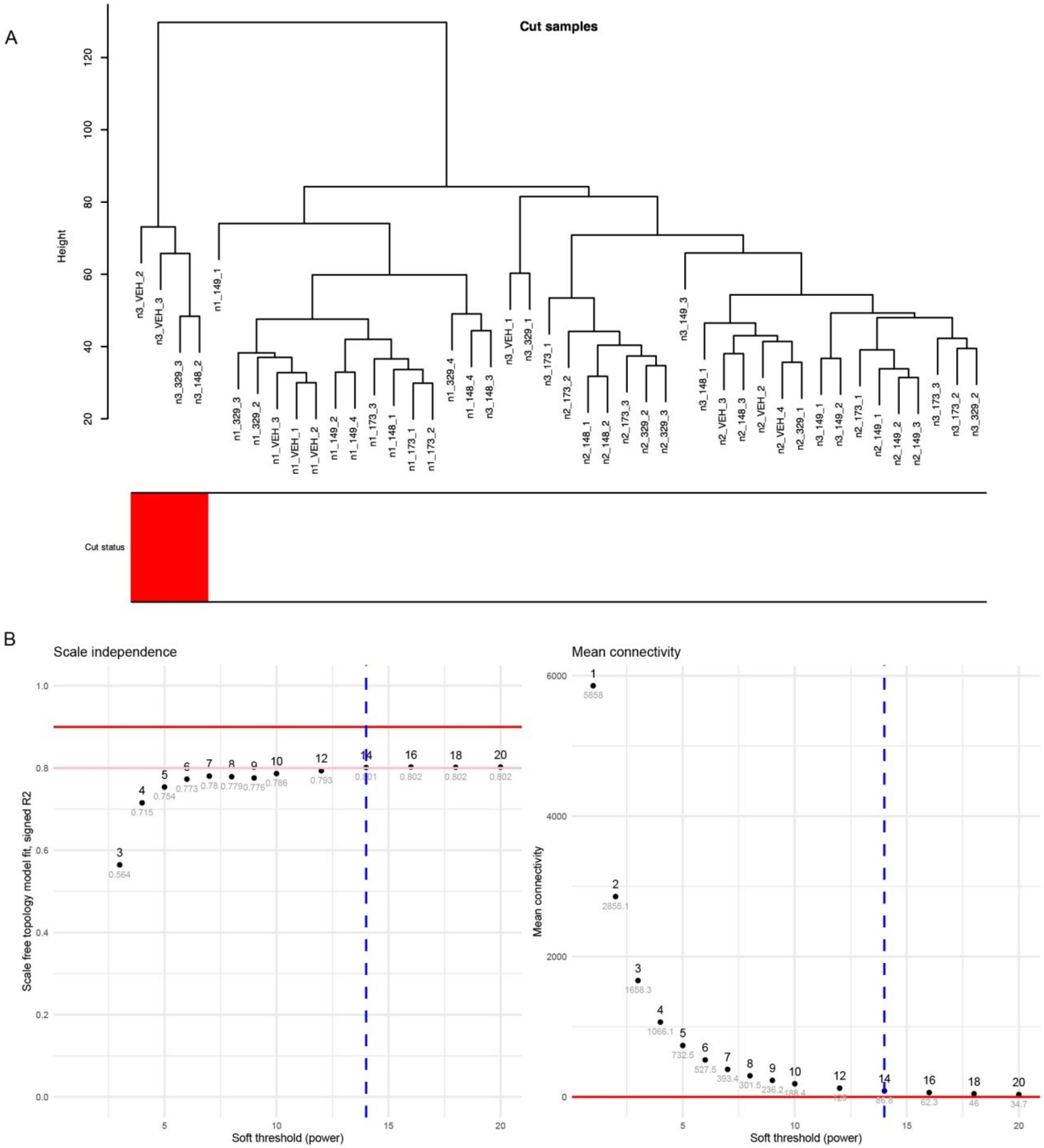

Fig. S14. RNA-Seq WGCNA analysis selection. (**A**) Sample clustering illustrating why five samples (left branch) were removed. (**B**) Network topology illustrating the selection of the soft thresholding power.

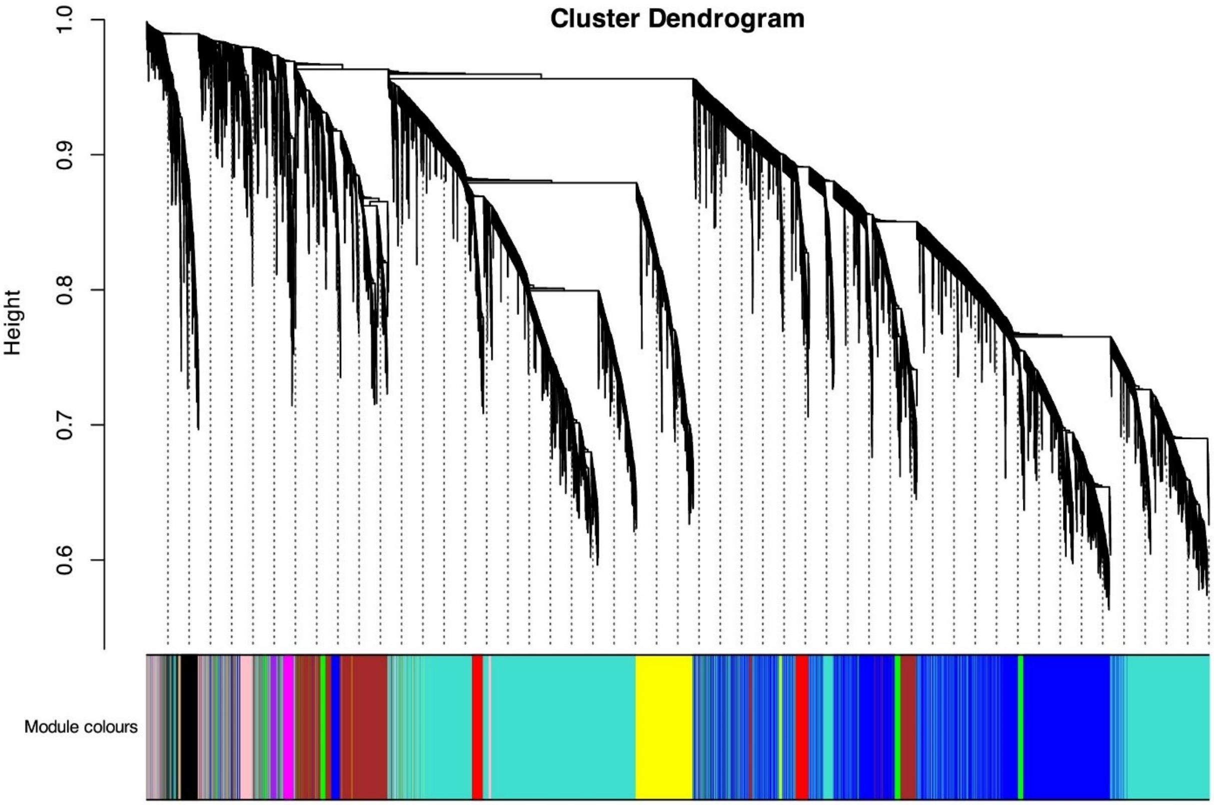

Fig. S15. Cluster dendrogram of gene modules in RNA-Seq WGCNA analysis. Dendrogram representing the hierarchical clustering of gene modules identified through WGCNA of RNA-Seq data. Each branch of the dendrogram corresponds to a cluster. Each branch of the dendrogram corresponds to a cluster of co-expressed genes, with the height indicating the degree of similarity between clusters. The colour coding (module colors) reflects the assignment of genes to different modules based on their co-expression patterns.

***MSH3* ASO**	**Sequence and chemistry**
ASO-1 (the *MSH3* ASO)	GCTTTTGGACCGTTTG
ASO-2	AGAAGATAGCTGGTAG
ASO-3	GTATATTGATGTAGGT
ASO-4	ATATTGAGCTCTCGGG
ASO-5	AAGTTAACAATGCCAG
ASO-6	AGGTATACTTGCTGTC
SCR ASO	GGCTACTACGCCGTCA
SCR ASO (used in *in vivo* studies)	CGCCGATAAGGTACAC

Table S1. *MSH3* ASO sequence and chemistry. Nucleotides with 2'-O-ethyl sugar modification are denoted in blue, DNA are in black, and molecules have full phosphorothioate backbone modifications. ASO-1 was the MSH3 ASO selected for further *in vitro* and *in vivo* characterization.

Treatment	**Number of samples**
MSH3 ASO -1	10
MSH3 ASO -2	9
MSH3 ASO -3	10
SCR ASO	10
Vehicle	10

Table S2. Samples used for WGCNA network construction.