bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

A CAG repeat threshold for therapeutics targeting somatic instability in Huntington's disease.

Sarah G. Aldous[1], Edward J. Smith[1], Christian Landles[1], Georgina F. Osborne[1], Maria Cañibano-Pico[1], Iulia M. Nita[1], Jemima Phillips[1], Yongwei Zhang[2], Bo Jin[2], Marissa B. Hirst[3], Caroline L. Benn[4], Brian C. Bond[5], Winfried Edelmann[2], Jonathan R. Greene[3], Gillian P. Bates[1]

[1]Huntington's Disease Centre, Department of Neurodegenerative Disease and UK Dementia Research Institute at UCL, Queen Square Institute of Neurology, University College London, London WC1N 3BG, UK

[2]Department of Cell Biology, Albert Einstein College of Medicine, NY 10461, USA

[3]Rancho BioSciences, San Diego, California 92127, USA

[4]LoQus23 Therapeutics, Riverside, Babraham Research Campus, Cambridge CB22 3AT, UK

[5]Prism Training and Consultancy Limited, St John's Innovation Centre, Cowley Road Cambridge, CB4 0WS, UK

Correspondence to: Gillian Bates

Huntington's Disease Centre, Department of Neurodegenerative Disease

Queen Square Institute of Neurology, UCL

Queen Square, London WC1N 3BG, UK

Email: gillian.bates@ucl.ac.uk

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

The Huntington's disease mutation is a CAG repeat expansion in the huntingtin gene that results in an expanded polyglutamine tract in the huntingtin protein. The CAG repeat is unstable, and expansions of hundreds of CAGs have been detected in Huntington's disease *post-mortem* brains. The age of disease onset can be predicted partially from the length of the CAG repeat as measured in blood. Onset age is also determined by genetic modifiers, which in six cases involve variation in DNA mismatch repair pathways genes. Knocking-out specific mismatch repair genes in mouse models of Huntington's disease prevents somatic CAG repeat expansion. Taken together, these results have led to the hypothesis that somatic CAG repeat expansion in Huntington's disease brains is required for pathogenesis. Therefore, the pathogenic repeat threshold in brain is longer than $(CAG)_{40}$, as measured in blood, and is currently unknown.

The mismatch repair gene *MSH3* has become a major focus for therapeutic development, as unlike other mismatch repair genes, nullizygosity for *MSH3* does not cause malignancies associated with mismatch repair deficiency. Potential treatments targeting *MSH3* currently under development include gene therapy, biologics and small molecules, which will be assessed for efficacy in mouse models of Huntington's disease. The zQ175 knock-in model carries a mutation of approximately $(CAG)_{185}$ and develops early molecular and pathological phenotypes that have been extensively characterised. Therefore, we crossed the mutant huntingtin allele onto heterozygous and homozygous *Msh3* knock-out backgrounds to determine the maximum benefit of targeting *Msh3* in this model. Ablation of *Msh3* prevented somatic expansion throughout the brain and periphery, and reduction of *Msh3* by 50% decreased the rate of expansion. This had no effect on the deposition of huntingtin aggregation in the nuclei of striatal neurons, nor on the dysregulated striatal transcriptional profile. This contrasts with ablating *Msh3* in knock-in models with shorter CAG repeat expansions. Therefore, further expansion of a $(CAG)_{185}$ repeat in striatal neurons does not accelerate the onset of molecular and neuropathological phenotypes. It is striking that highly expanded CAG repeats of a similar size in humans cause disease onset before 2 years of age, indicating that somatic CAG repeat expansion in the brain is not required for pathogenesis. Given that the trajectory for somatic CAG expansion in the brains of Huntington's disease mutation carriers is unknown, our study underlines the importance of administering treatments targeting somatic instability as early as possible.

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Keywords: Huntington's disease, somatic CAG repeat instability, MSH3, genetic modifiers, pathogenic CAG repeat length

Abbreviations: CRISPR = clustered regularly interspaced short palindromic repeats; DAB = 3.3′-diaminobenzidine; FELASA = Federation of European Laboratory Animal Science Associations; gRNA = guide RNA; GOBP = gene ontology biological process; GWAS = genome-wide association study; HTRF = homogeneous time-resolved fluorescence; HTTexon1 = exon 1 HTT protein; IQR = interquartile range; PBS = phosphate buffer saline; PCA = principal component analysis; polyA = polyadenylation; polyQ = polyglutamine; Q = glutamine; qPCR = real-time quantitative PCR; RNA-seq = RNA sequencing; SEM = standard error of the mean; SMRT = single molecule real-time sequencing; WT = wild-type.

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

Huntington's disease is a heritable neurodegenerative disorder, characterised by motor dysfunctions, psychiatric changes, and cognitive decline.[1] The disease is caused by an expanded CAG repeat in exon 1 of the huntingtin gene (*HTT*), which encodes an expanded polyglutamine (polyQ) tract in the huntingtin protein (HTT). Individuals with a repeat of $(CAG)_{40}$ or more will develop Huntington's disease in a normal life span, while repeats of around $(CAG)_{65}$ or more will lead to disease onset in childhood or adolescence.[2] The CAG repeat is unstable in both somatic and germ cells,[2,3] and large expansions up to approximately $(CAG)_{1000}$ have been observed in patient brains.[4] In the presence of an expanded CAG repeat, the *HTT* pre-mRNA can be alternatively processed to generate the *HTT1a* transcript, through the activation of two cryptic polyA sites within intron 1.[5,6] This encodes the highly aggregation prone and pathogenic HTTexon1 protein.[5,7,8] Huntington's disease pathology is characterised by the deposition of nuclear and cytoplasmic inclusions,[9,10] as well as synaptic and neuronal loss in the striatum and other regions of the brain.[11,12] Nuclear huntingtin aggregation is linked to transcriptional dysregulation, a molecular phenotype observed in mouse models and *post-mortem* brains.[13-15]

Although the age of disease onset is inversely correlated with CAG repeat length, genetic and environmental modifiers also contribute to the variation in age of onset and disease progression.[16] Genome-wide association studies (GWAS) have identified potential genetic modifiers, a subset of which have variants associated with genes involved in DNA damage repair pathways including *FAN1*, *MSH3*, *MLH1*, *LIG1*, *PMS1* and *PMS2*.[17-20] That the variants affected the function of these DNA repair genes was supported by previous experiments showing that the ablation of specific mis-match repair genes (*Msh2*, *Msh3*, *Mlh1* and *Mlh3*) reduced somatic CAG repeat instability in mouse models of Huntington's disease.[21-23] Mechanistically, it has since been shown that FAN1 nuclease activity is implicated in slowing CAG repeat expansion,[24] that the promotion of somatic CAG instability by FAN1 is blocked in the absence of MLH1[25] and that FAN1 controls mismatch repair assembly via MLH1 to stabilise the CAG repeat.[26,27]

It is now widely accepted that the somatic expansion of a mutant CAG repeat in brain cells is the first step in the molecular pathogenesis of Huntington's disease, followed by a second process that leads to cell dysfunction, once a CAG repeat threshold has been achieved. Therefore, it is the rate of somatic CAG repeat expansion that drives the age of onset and rate

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

of progression of the disease, and the genetic modifiers that impact this process are potential therapeutic targets.[28,29] Of these, MSH3, part of the MutSβ heterodimer that is thought to recognise the potential slip outs caused by the expanded CAG repeats,[30,31] is of particular interest. Previous work has established that *Msh3* is required for somatic instability in the *Hdh*Q111 mouse model of Huntington's disease,[32] and that variants of mouse *Msh3*[33] and levels of *MSH3* expression in Huntington's disease patients[34] affect somatic expansion. Importantly, for therapeutic potential, unlike other DNA mis-match repair genes, loss of function of MSH3 does not cause Lynch syndrome, which is associated with a high risk of colon cancer and a spectrum of other malignancies.[35,36]

The design of mouse trials for the target validation of MSH3 and for the preclinical testing of novel MSH3 targeting therapeutics is not straightforward. Ideally, the chosen model will develop robust phenotypes within a timescale amenable to the administration and durability of the potential therapeutic agents. However, such a model is likely to have been created with a highly expanded CAG repeat that would cause childhood onset of the disease in humans, and therefore, does not model the CAG repeat expansion process that occurs in the adult-onset disease. The zQ175 knock-in mouse of Huntington's disease is an attractive candidate as it is well characterised[7,37,38] and has been used for HTT lowering approaches.[39] However, the mutant allele in zQ175 generally has approximately $(CAG)_{190}$. To test whether the prevention of somatic CAG repeat expansion could result in measurable improvements in zQ175 mice, we crossed the mutant *Htt* allele onto heterozygous and homozygous *Msh3* knockout backgrounds. Although, in the absence of *Msh3*, the pronounced somatic CAG repeat instability in the striatum was prevented, there was no improvement in striatal transcription dysregulation, as measured by RNA-seq, nor was there a decrease in striatal HTT aggregation. These data point to a CAG repeat threshold above which the rate of disease pathogenesis does not increase, and at which the prevention of somatic CAG repeat expansion will have no beneficial consequences. The trajectory for CAG repeat-expansion in the human brain is not known and this work further underlines the importance of developing treatments that can be given to Huntington's disease mutation carriers as early as possible.

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Materials and methods

Ethics statement

All procedures were performed in accordance with the Animals (Scientific Procedures) Act, 1986, and approved by the University College London Ethical Review Process Committee.

Generation of *Msh3* knock out mice

Msh3 knockout mice were generated by CRISPR technology.[40] In brief, guide RNAs (gRNAs) targeting exon 3 of mouse *Msh3*: Msh3 gRNA 51/57（targeting sequence: TCTGCAACGTGCAAAGAATGCGG）and exon 8, Msh3 gRNA 89/69, (Targeting sequence: GTTGGCACAGACAGGTCGGAGGG) were designed by Benchling online software and generated by *in vitro* transcription. Fertilised eggs from C57BL/6J mice were injected with the two gRNAs (Msh3 gRNA 51/57 and Msh3 gRNA 89/69) and Cas9 protein (PNA BIO INC) and transferred to pseudopregnant CD1 female mice. *Msh3*$^{+/-}$ founders (F0) mice were backcrossed to C57BL/6J (Charles River) mice and maintained on C57BL/6J background (Charles River).

Animal colony breeding and maintenance

zQ175 mice on a C57BL/6J background were obtained from the CHDI Foundation colony at the Jackson Laboratory (Bar Harbour, Maine). zQ175 males were bred to *Msh3*$^{+/-}$ females to generate zQ175:*Msh3*$^{+/-}$ males that were then bred to *Msh3*$^{+/-}$ females to generate the six required genotypes as littermates and aged to 6 months. The mean and standard deviation of the CAG repeat sizes were zQ175 186 ± 2.3, zQ175:*Msh3*$^{+/-}$ 185 ± 2.4 and zQ175:*Msh3*$^{-/-}$ 182 ± 2.0. A second smaller generation was bred and aged to 3 months, with CAG sizes zQ175 185 ± 3.2, zQ175:*Msh3*$^{+/-}$ 183 ± 2.6 and zQ175:*Msh3*$^{-/-}$ 190 ±6.0. The CAG repeat size for the 2-month-old zQ175 mice for HTRF was 192 ± 6.4.

Mouse husbandry and health monitoring were carried out as previously described.[41] Mice were group-housed by gender, with mixed genotypes per cages. Animals were kept in individually ventilated cages with Aspen Chips 4 Premium bedding (Datesand) and environmental enrichment through access to chew sticks and a play tunnel (Datesand). All mice had constant

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

access to water and food (Teklad global 18% protein diet, Envigo, the Netherlands). Temperature was regulated at 21°C ± 1°C and mice were kept on a 12 h light/dark cycle. The facility was barrier-maintained and quarterly non-sacrificial FELASA (Federation of European Laboratory Animal Science Associations) screens found no evidence of pathogens. Tissues were harvested, snap frozen in liquid nitrogen and stored at -80°C. Mice were transcardially perfused with 4% paraformaldehyde as previously described.[7]

Real-time quantitative PCR for *Msh3* transcripts

RNA was isolated and real-time quantitative PCR (qPCR) was performed as previously described.[42] Taqman assays to the deleted *Msh3* region (Thermo, Mm00487756_m1) and outside the deleted region (Thermo, Mm01290051_m1) were used, and to housekeeping genes for *Atp5b* and *Eif4α2* as described previously.[42]

Genotyping and CAG repeat sizing

Genomic DNA isolation from ear biopsies or from frozen brain regions and peripheral tissues was carried out as previously described.[43] DNA from ear biopsies was diluted to 50 ng/µL. The DNA concentration from brain regions and peripheral tissues was variable but within the range of 100 – 300 ng/µL.

Genotyping PCRs for zQ175 mice were performed as previously described.[43] The *Msh3* PCR was carried out in 10 µL reactions using DreamTaq Buffer (Thermo). The *Msh*$^{-/-}$ knockout allele was amplified with primers Msh3 51 SF (5′-GAATTGTCCTAGTGAAGCCAAAA-3′) and Msh3 89SR (5′-CTTGGGAGGAAAAGCAGATG-3′) generating a product of 501 bp. The wild-type allele was amplified with primers Msh3 51 SF and Msh3 51 SR (5′-AACCAACTACTCTCCCCCAAA-3′) generating a 479 bp product. Primers were at 5 pmol/µL, and cycling conditions were 2.5 min at 95 °C, 32 x (20 s at 95 °C, 30 s at 56 °C, 1 min at 72 °C), 6 min at 72 °C.

CAG repeat sizing for ear biopsy DNA was carried out as described previously and for somatic instability analysis of DNA from tissue samples, reaction volumes were scaled to 25 µL.[43] The PCR product was denatured in HiDi Formamide (Thermo Fisher Scientific) and run in triplicate on ABI3730XL Genetic Analyser with MapMarker ROX 1000 (Bioventures) internal size standards and analysed using GeneMapper v6 software (Thermo Fisher Scientific). The tissue-

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

wide analysis of mode and instability index was analysed as described,[27,44] using a custom R script.[27] A more detailed statistical analysis of striatal samples was run using JMP®, v17 (SAS Institute Inc., Cary, NC).

Western blotting

Cortex was lysed on ice in KCL buffer (50 mM Tris pH 8.0, 10% Glycerol, 150 mM KCL, 5 mM EDTA, with 1 mM phenylmethylsulphonyl fluoride and cOmplete protease inhibitor cocktail tablets (Roche)). Tissue homogenisation and western blotting was performed with 50 µg of total protein as previously described.[15] Membranes were incubated overnight with gentle agitation at 4 °C with the MSH3 antibody (Merck MSH3MAB(1F6)) in 5% milk PSB-T (phosphate buffer saline (PBS), 0.1% Tween-20). Tubulin was detected by incubating for 1 h at room temperature in 5% Milk-PBST (Marvel) with an α-tubulin antibody (Sigma T9026) at a 1:40.000 dilution.

Tissue lysis and QuantiGene assays

The lysis of tissue samples and the QuantiGene multiplex assays were carried out as previously described,[42] with 1:2 dilutions for each brain region. All QuantiGene probe sets are available from Thermo Fisher Scientific and the individual accession numbers and probe set regions are summarised in Supplementary Figure 1 and Supplementary Tables 1 and 2. The sequences for the capture extenders, label extenders and blocking probes are available from Thermo Fisher Scientific on request. Data analysis was performed by normalising raw MFI signals obtained to the geometric mean of the housekeeping gene signals.

Homogeneous time-resolved fluorescence

Homogeneous time-resolved fluorescence (HTRF) was performed as previously described,[7] using n = 8/ genotype/ age with an equal ratio of males to females. The antibody pairings and concentrations for the soluble HTTexon1 assay were donor: 2B7-Tb, 1 ng and acceptor: MW8-d2, 40 ng. For the aggregated HTTexon1 assay the antibody pairings were donor: 4C9-Tb, 1 ng and acceptor: MW8-d2, 10 ng. The donor and acceptor antibodies were added in 5 µL HTRF detection buffer (50 mM NaH_2PO_4, 0.2 M KF, 0.1% bovine serum albumin, 0.05% Tween-20) to give the optimised antibody concentration when added to 10 µL of a 10 % lysate

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

concentration in a 384-well ProxiPlate (Greiner Bio-One). Following incubation for 3 h on an orbital shaker (250 rpm), plates were read using an EnVision (Perkin Elmer) plate reader as previously described.[43]

Immunohistochemistry

The immunostaining and imaging of coronal sections using S830 was carried out as previously outlined.[7] Sections from 3- and 6-month-old mice were stained and imaged separately, but the four genotypes for each age were processed together.

To perform an analysis of aggregation throughout the striatum, four images per striatum across five hemispheres/sections were captured (20 images in total). Images were obtained using a Zeiss AxioSkop2 Plus microscope fitted with a Zeiss AxioCam 208 colour camera and were recorded using Zeiss ZEN 3.2 (blue edition). For quantification and thresholding, images were converted to 8-bit in ImageJ and a signal intensity level was chosen to indicate staining above background. Signals were segmented based on pixel intensity and the number of positive pixels in contact with each other. A gate of 80 pixels was applied to all signals over 190 to identify S830 positive objects. Inclusions were identified by applying a gate of 2 pixels and a signal of 150, the vast majority of which would be nuclear inclusions. The average area of each object and the percentage of the image that contained S830 immunostaining was recorded.

RNA Sequencing

Tissues were homogenised in Qiazol lysis reagent (Qiagen) and RNA was extracted from the same set of samples as used for somatic instability. Total RNA isolation was performed as outlined[42] using the RNeasy mini kit (Qiagen). RNA concentration (Thermo, Q32855), integrity and quality (Thermo, Q33222) were measured using a Qubit Fluorometer (Themo Fisher Scientific Q33238). RNA sequencing and analysis was then performed as described previously.[45] RNA was converted into cDNA, indexed and libraries sequenced on the Illumina NovaSeq 6000 system to generate 30 million paired-end reads. Differential expression tests were performed in R using DESeq2[46] with independent filtering disabled, and a significance threshold of an adjusted $P < 0.05$ after multiple-test correction, and fold changes of at least 20% in either direction. Analysis of RNA-Seq data comparing zQ175 with wild-type to

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

generate the "HD signature" lists of transcriptionally dysregulation genes was also based on the same significance criteria. The posterior probabilities-based method[45] was used to quantify the probability that *Msh3* knockout out restored transcriptional dysregulation, either partially or fully, on a gene-by-gene basis. This was done by viewing the effect of *Msh3* genotype as a multiple of the HD model signature (in log2). If $\alpha < 0$, then the disease effect is reversed by *Msh3* knockout, if $\alpha > 0$, the disease effect is exacerbated by loss of *Msh3*. The thresholds for each type of reversal class were as previously calculated[45] and were as follows: super-reversal: $\alpha < -1.3$; full reversal: $-1.3 < \alpha < -0.7$; partial reversal: $-0.7 < \alpha < -0.3$; negligible reversal: $-0.3 < \alpha < 0.3$; and exacerbation: $\alpha > 0.3$. The genes were then scored for normalization based on the overall reversal probability. The 'HD signature' dysregulated gene list was tested for gene set overrepresentation against the GOBP (gene ontology biological process)[47] version 18.9.29 using the enricher function within the R clusterProfiler[48] package. A clusterProfiler $q < 0.05$ was deemed significant.

Statistical analysis

Data were screened for outliers at sample level using the ROUT test (Q = 10%; GraphPad Prism v9) and outliers were removed from the analysis. All data sets were tested for a normal gaussian distribution (Shapiro-wilk, Prism v9). Statistical analysis was by one-way ANOVA, two-way ANOVA or mixed-effects model (REML) with either Tukey's or Bonferroni *post hoc* tests. Non-parametric analysis was performed by the Kruskal-Wallis test with Dunn's *post hoc* analysis. Graphs were prepared using Prism v9 (GraphPad Software, California, USA). *P*-values <0.05 were considered statistically significant. PCA analysis of the summary statistics for the CAG repeat profiles of the striatal and ear tissue was performed. The Standard Least Squares model with Effect Leverage was conducted with ear trace data included as a covariate. Data for variance were log transformed and graphs prepared using JMP® v17 (SAS Institute Inc., Cary, NC).

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Results

An *Msh3* knockout mouse line was generated using CRISPR Cas9 editing,[40] targeting exons three to eight as shown in Fig. 1A. A colony was established on a C57BL/6J background and mice were bred to homozygosity. The level of *Msh3* mRNA in cortical lysates was shown to have decreased by 50% in heterozygous (*Msh3*+/-), and to have been ablated in homozygous (*Msh3*-/-), knockout mice by real-time quantitative PCR (qPCR) (Fig. 1B). MSH3 protein levels were determined in cortical lysates by western blotting and similarly showed that MSH3 had been reduced to at least 50% in *Msh3*+/- heterozygotes and could not be detected in the *Msh3*-/- homozygotes (Fig. 1C). To determine the effects of MSH3 reduction and MSH3 ablation in zQ175 mice, zQ175 males were bred to *Msh3*+/- females to generate zQ175 males that were heterozygous for *Msh3* knockout. These were then bred to *Msh3*+/- females to generate all six genotypes of interest as littermates in a single generation (Fig. 1D).

Loss of *Msh3* prevents somatic CAG repeat instability in zQ175 mice.

We first investigated the extent of somatic CAG repeat instability in ten CNS regions and several peripheral tissues from zQ175 mice at 6 months of age (N = 8). The GeneMapper CAG repeat profiles for the brain regions and tissues for each mouse were compared to that obtained from the ear sample taken at post-natal day 12 (P12) (Fig. 2). Within the CNS, instability was most pronounced in the striatum, brain stem, colliculus, thalamus, hypothalamus and spinal cord, consistent with that observed in other mouse models.[49] In the periphery, instability was extensive in the liver as expected,[23,33,49] but absent from the heart. Unexpectedly, CAG repeat contraction had led to a reduction in the CAG repeat length mode in the pituitary. Examination of the traces for zQ175:*Msh3*-/- mice showed that nullizygosity for *Msh3* prevented CAG repeat expansion and that the repeat contractions had resulted in a decrease in the mode in most brain regions and peripheral tissues (Fig. 2). Heterozygosity for *Msh3* in zQ175:*Msh3*+/- mice caused in a reduction in the extent of repeat expansion (Fig. 2).

The instability index quantifies the change in CAG repeat length profiles between samples.[44] This was calculated for each brain region and peripheral tissue for every mouse by comparing its CAG repeat trace with that of the P12 ear sample (N = 8/genotype). Plotting the change in the mean instability index for each tissue, demonstrated the highly significant,

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

stepwise reduction in CAG repeat instability with loss of MSH3 (Fig. 3A). In many of the brain regions and tissues, e.g. olfactory bulb, colliculus thalamus and retina, the level of instability in zQ175:*Msh3*+/- mice was decreased to approximately 50% of the reduction observed in the zQ175*Msh*3-/- mice. In some brain regions, e.g. striatum, the decrease in instability index between zQ175 and zQ175:*Msh3*+/- mice was less pronounced than between zQ175:*Msh3*+/- and zQ175:*Msh3*-/- mice.

The change in mode CAG repeat length between each brain region and tissue at 6 months of age as compared to the P12 ear sample was also calculated for each mouse (N = 8/genotype). The average mode CAG length changed little with age in most brain regions and tissues for zQ175 mice, exceptions being the cerebellum, brain stem, pituitary, and retina (Fig. 3B). However, there was a reduction in the average mode CAG length in the zQ175:*Msh3*-/- mice in most brain regions and tissues (Fig. 3B).

As the most somatically unstable region, we set out to analyse the striatal traces in more detail using summary statistics in a multivariate analysis. The Principal Component Analysis (PCA) investigated correlations between a set of summary statistics of CAG repeat profiles and found new independent trends (Fig. 4A**)**. The first two components explained 77.3 % of the total variation with the first component explaining 62.2 % alone. The first component in the scores plot separated the striatum for the zQ175:*Msh3*+/- and zQ175:*Msh3*-/- mice from the ear tissues of all three genotypes and the striatum of the zQ175 mice. To a lesser degree the second component separated the ear tissue from the striatum of the zQ175 mice. The loadings plot in Fig. 4B represents the summary statistics, and the relationship between them. The summary statistics related to location (mean/median/mode) and variation (log variance/log interquartile range (IQR) along with skewness, instability and expansion were all closely correlated in the first component, with contraction and kurtosis negatively correlated in the second component. The information from both the scores and loadings plots confirmed the difference seen visually with the traces in Fig 2**,** and that these can be attributed to specific summary statistic differences.

We also analysed specific variables of the striatal traces in more detail; instability index, expansion, and variation (Fig. 4C). The instability index as calculated in Fig 3. A is normalised to the ear as a control, however due to the genetic nature of the intervention, these ear samples were also affected by the genotype, as visible in the zQ175:*Msh3*-/- ear trace in Fig 2. We reasoned that instability normalised to the ear trace, whilst a good indication of expansion, may

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

not be the most accurate, as the resulting ratio could have an inflated variability. Therefore, we used a standard least squares model, and included the ear data of each variable, as a co-variate. This analysis of covariance (ANCOVA) approach enabled us to ensure that any difference detected indicated a true effect of the genotype on the striatum, over and above the non-specific effect. These results again demonstrated a clear effect of genotype on the instability index, expansion, and variation of the striatal traces, confirming the effect of *Msh3* genotype on instability, and other statistical variables that characterise the traces seen in Fig 2.

The absence of MSH3 had no effect on transcriptional dysregulation in the striatum of zQ175 mice

Transcriptional dysregulation is a molecular phenotype that has been well-characterised in Huntington's disease *post-mortem* brains and mouse models[14,50]. This is particularly the case for the striatal transcriptome of an allelic series of knock-in mice[13] for which a robust disease signature has been established.[51] Therefore, we performed RNA-seq on striata from wild-type, zQ175, *Msh3*-/- and zQ175:*Msh3*-/- mice to investigate the effect of *Msh3* ablation on the striatal transcriptome of zQ175 mice at 6 months of age.

We began by comparing the expression levels of striatal genes between zQ175 and wild-type mice. The extent of transcriptional dysregulation is illustrated in the volcano plot (Fig. 5A). This zQ175 signature correlated with those previously obtained for the striata from the zQ175, Q140 and *Hdh*Q111 knock-in mouse models at 6 and 10 months of age (Supplementary Fig. 2A).[13] Similarly, the gene set overrepresentation of the 2,486 dysregulated genes allocated to gene ontology biological process (GOBP) confirmed that the transcriptional dysregulation in the zQ715 mice was consistent with that previously reported (Supplementary Fig. 2B).[45]

The extent of transcriptional dysregulation in both zQ175 and zQ175:*Msh3*$^{-/-}$ mice as compared to wild-type was very similar (Fig. 5B). Comparison of the striatal gene expression profiles between zQ175 and zQ175:*Msh3*-/- identified only two changes, *Msh3* and *Shox2*, the most significant of which was *Msh3* (Fig. 5C). The ablation of MSH3 resulted in the expression of only 3% of the 2,486 genes dysregulated in the zQ175 striata being reversed toward wild-type levels whilst the dysregulated levels of 1% of these genes was exacerbated (Fig. 5E). Of the 65 genes for which the expression level was reversed, 14 were fully reversed and 51

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

partially reversed, numbers corresponding to below background (Fig. 5E). We also considered the 266 genes that comprise the 'striatal signature' of robustly dysregulated genes across multiple Huntington's disease models.[51] Of these, 258 were dysregulated in the zQ175 striatum, and a reversal of expression toward wild-type levels in the absence of MSH3 was not observed for any of these genes (Fig. 5E). The expression levels of a few key striatal genes in wild-type, Msh3-/-, zQ175 and zQ175:Msh3-/- striata are illustrated in Fig. 5F.

Ablation of MSH3 had no effect on the striatal transcriptome of wild-type mice; *Msh3* being the only transcript found to differ between the two genotypes (Fig. 5D). Consistent with this, ablation of MSH3 had no effect on the expression level of a panel of mis-match repair genes in either wild-type or zQ175 mice in the striatum (Fig. 5G and Supplementary Fig. 3) or other brain regions (Supplementary Fig. 3) as measured by a QuantiGene multiplex assay.

Striatal nuclear HTT aggregation levels were unchanged with loss of *Msh3* in zQ175 mice.

Several lines of evidence indicate that the retention of aggregated HTT in neuronal nuclei precedes transcriptional dysregulation.[7,13,15] Therefore, we used immunohistochemistry with the S830 anti-HTT antibody to quantify the level of HTT aggregation in striata. Using this approach, we have previously shown that aggregated HTT could first be detected in striatal nuclei from zQ175 mice before two months of age as a diffuse immunostain that filled the nuclei.[7] By six months of age, this had often resolved to a nuclear inclusion either with or without the diffusely distributed aggregated HTT.[7] Coronal sections from zQ175, zQ175:*Msh3*$^{+/-}$ and zQ175:*Msh3*-/- mice, together with wild-type controls, at three and six months of age were immunostained with S830. At three months of age, the S830 immunostain appeared diffusely nuclear as expected, with the initiation of inclusion bodies being detected in some cells (Fig. 6A). The application of thresholding and quantification of signal intensity indicated that there was no difference in the number of S830 positive objects, the average area of these objects and the percentage of the image stained between the three genotypes (Fig. 6B). Therefore, *Msh3* genotype did not influence the first appearance of HTT aggregation in striatal nuclei. At six months of age, in addition to the diffuse HTT aggregation, nuclear inclusions were more prominent (Fig. 6C). Thresholding was applied to quantify the number of S830 positive objects, their average size and the percentage of the image stained both for all objects (Fig. 6D) and for inclusions, the majority of which were nuclear (Fig. 6E). As at three months

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

of age, at six months, there was no difference in the level of HTT aggregation in striatal nuclei between genotypes.

We have previously established homogeneous time-resolved fluorescence (HTRF) assays to compare the levels of soluble and aggregated HTT isoforms between tissues.[43] These included an assay specific for the soluble HTTexon1 protein (2B7-MW8) and a HTT aggregation assay (4C9-MW8) that we have shown is specific for the aggregated HTTexon1 protein.[7] The epitope for 2B7 is within the first 17 amino acids of HTT, that for 4C9 is within the human-specific proline-rich region and MW8 acts as a C-terminal-neo-epitope antibody, and is specific for HTTexon1 in this HTRF assay (Fig. 7A). We applied these assays to the striatal, cortical and brain stem tissues of zQ175, zQ175:*Msh3*+/- and zQ175:*Msh3*-/- mice to investigate whether the prevention of somatic CAG repeat expansion had resulted in a decrease in HTT aggregation. Striatum, cortex, and brain stem from zQ175 and wild-type mice at two months of age were included in the analysis.

The 4C9-MW8 assay showed that HTTexon1 aggregation had increased from two to six months of age in all three brain regions (Fig. 7B). At six months of age, there was no difference in HTTexon1 aggregation between zQ175, zQ175:*Msh3*+/- and zQ175:*Msh3*-/- in the striatum, consistent with the immunohistochemical analysis (Fig. 7B). However, there was a small reduction HTTexon1 aggregation in the cortex of zQ175:*Msh3*+/- mice as compared to zQ175, and in zQ175:*Msh3*-/- mice as compared to zQ175:*Msh3*+/- (Fig. 7B). Similarly, the level of HTTexon1 aggregation in the brain stem of zQ175:*Msh3*-/- mice was decreased as compared to zQ175 and zQ175:*Msh3*+/- mice (Fig. 7B**)**. The 2B7-MW8 HTRF assay demonstrated that the soluble levels of HTTexon1 in zQ175 mice remain unchanged by *Msh3* genotype, in all three brain regions (Fig. 7C). Consistent with this, the level of the *Htt1a* transcript was not consistently changed in the striatum, cortex and brain stem between zQ175, zQ175:*Msh3*+/- and zQ175:*Msh3*-/- (Supplementary Fig. 4).

Discussion

Genetic variation in DNA mismatch repair pathway genes modifies the age of onset and rate of progression of Huntington's disease by slowing the rate of somatic CAG repeat expansion. Of these, MSH3 has become a major focus for therapeutic development because loss of *MSH3* appears to be well tolerated.[36] In preparation for the preclinical evaluation of these potential treatments, we genetically ablated *Msh3* in the zQ175 knock-in mouse model of Huntington's

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

disease to determine the maximum benefit of targeting *Msh3* in this model. Although loss of *Msh3* prevented somatic CAG repeat instability in striatal neurons, there was no effect on early molecular or pathological phenotypes. These data indicate that highly expanded CAG repeats can reach a threshold above which further expansion does not accelerate the onset of cellular pathology.

The zQ175 knock-in mouse model has been used extensively for mechanistic and preclinical studies, because its highly expanded CAG repeat results in molecular and pathological phenotypes at a younger age than in other knock-in mouse models of Huntington's disease.[37,38] However, somatic CAG repeat instability had not previously been characterised in these mice. We generated CAG-repeat traces for ten CNS regions and some peripheral tissues for zQ175 mice at six months of age, and compared these to traces from ear biopsies sampled when the mice were 12 days old. Pronounced instability was detected in the olfactory bulb, striatum, brain stem, colliculus, thalamus and spinal cord. This pattern of expansion was comparable to that previously published for the transgenic R6/1 mice with an repeat of $(CAG)_{115}$,[49] and so was not influenced by genomic context. In the periphery, there was considerable instability in the liver, consistent with that reported for the knock-in *Hdh*Q111 model.[23,33,49] That the CAG repeat contracted in the pituitary was an unexpected finding, and could provide mechanistic insights in a future study. We have consistently found the heart CAG repeat to be comparatively stable and propose this could be used as a reference tissue, should an ear or tail biopsy from young mice not be available.

Capillary electrophoresis of bulk-PCR products, as used here, will not have captured the full extent of CAG repeat expansions for any brain region or tissue, as shorter CAG tracts will be preferentially detected. This is particularly challenging for the zQ175 model, for which the baseline CAG repeat length is already very high. This issue has been circumvented with small-pool PCR approaches whereby samples were diluted to contain a few DNA molecules and alleles were detected by Southern blotting.[4,52] However, higher throughput analyses are required. A comparison of capillary electrophoresis, MiSeq or PacBio SMRT sequencing for bulk-tissue analysis demonstrated that PCR-based approaches cannot be used to estimate modal length or to quantify expansions greater than $(CAG)_{250}$,[53] and approaches that avoid an amplification step (no-amp) are under development.[54,55] Bulk-PCR analysis also obscures any cell-type specificity for CAG repeat instability. Laser dissection showed that large expansions were present in neurons rather than glia in mouse models and *post-mortem* brains.[56,57] More recently, CAG repeat expansions within a given cell population were determined by sorting

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

cell-specific nuclei from Huntington's disease *post-mortem* brains followed by high-throughput sequencing. This showed that, in the striatum, expansions had occurred in medium spiny neurons and CHAT+ interneurons but not in glial populations.[58] Interestingly, CAG repeat expansions were detected in cerebellar Purkinje cells, a rare cell type in this brain structure,[58] whereas CAG repeat instability in the cerebellum had not been uncovered by small-pool PCR.[4]

The ablation of MSH3 completely prevented somatic CAG repeat expansion in all brain regions and peripheral tissues and in many cases resulted in a reduction of the repeat length mode by up to five CAGs. Comparison of the instability index for zQ175, zQ175:*Msh3*+/- and zQ175:*Msh3*-/- demonstrated that for most brain regions and tissues, a 50% reduction in MSH3 resulted in a 50% decrease in CAG repeat expansion. In the case of the striatum and spinal cord, the reduction in expansion between zQ175 and zQ175:*Msh3*+/- mice was much less than 50%, possibly reflecting an underestimate of the extent of expansion in these brain regions, as discussed above. MSH3 levels only affected expansion, as neither a loss of MSH3, or a reduction in its levels reversed the CAG repeat contraction observed in the pituitary.

Our rationale for performing the intercross between zQ175 and *Msh3* knock-out mice was to determine whether zQ175 might be useful preclinically for assessing therapeutics targeting MSH3. At six months of age zQ175 mice have robust molecular and pathological phenotypes in the form of striatal transcriptional dysregulation[13] and the deposition of aggregated HTT throughout the brain,[7] and have been used in a zQ175 preclinical trial to assess HTT-lowering with zinc finger proteins.[39] Therefore, we performed RNA-seq on striatal RNA for wild-type, *Msh3*-/-, zQ175 and zQ175:*Msh3*-/- mice. Remarkably, the ablation of MSH3 had no impact on the dysregulated transcriptional profile of the zQ175 striatum; of 2,486 dysregulated genes, only 3% of dysregulated genes had an expression that was reversed toward wild-type levels, and in 1%, this was exacerbated; changes that would be below background levels. This is in stark contrast to the effect of ablating MSH3 in CAG140 knock-in mice (zQ175 mice are CAG140 mice with a longer CAG repeat) in which HTT aggregation in striatal nuclei was decreased and a considerable percentage of the dysregulated transcripts were reversed toward wild-type levels (X. William Yang, personal communication, manuscript submitted).

It is now widely accepted that it is the somatic CAG repeat expansion in the brain that drives the age of onset and rate of progression of Huntington's disease,[59] but the mechanism

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

through which this initiates pathogenesis has not been confirmed. The alternative processing of the *HTT* pre-mRNA to generate *HTT1a*, provides a candidate for this second step.[5,6] The production of *HTT1a* increases with increasing CAG repeat length and it encodes the HTTexon1 protein that is known to be highly aggregation-prone and pathogenic.[15] Therefore somatic CAG repeat expansion would be expected to increase the levels of *Htt1a* and of the soluble HTTexon1 protein at the single cell level. However, when comparing zQ175:*Msh3*-/- mice with zQ175, we did not identify a reduction in *Htt1a* or HTTexon1 in bulk tissue by QuantiGene or HTRF analyses, respectively. These techniques may not be sufficiently sensitive to detect these changes at the tissue-level, given that the CAG repeat in zQ175 mice is already highly expanded and that in most cells further expansions do not increase the CAGs repeat length to a dramatic degree.

Changes in HTT aggregation might be more amenable to detection, given the cumulative nature of the process and that only the cells in which aggregation occurs contribute to the analysis. HTT aggregation is a concentration-dependent process[60] and whether this occurs in a given subcellular location will depend on a number of factors e.g. the concentration of HTTexon1 and other HTT N-terminal fragments, the polyQ length (as the longer the polyQ tract the lower the concentration required), post-translational modifications, interacting proteins or RNAs, and the proteostasis network. In zQ175 mice, it is the soluble HTTexon1 protein that is depleted from brain regions during the aggregation process.[7] The N-terminus of HTT contains a potent nuclear export signal,[61,62] and HTT is retained in neuronal nuclei because it aggregates there, appearing first as a diffuse immunoreactive product that fills the nucleus.[15,61,63] The threshold concentration for the nucleation of aggregation will be cell-type and subcellular location specific.

Knocking out the mismatch repair genes *Msh2* or *Mlh1* delayed the accumulation of aggregated HTT in striatal nuclei of the *Hdh*Q111 knock-in model.[22,23] However, prevention of somatic CAG repeat expansion in zQ175 mice had no effect on the deposition of aggregated HTT in striatal nuclei. Given that HTT aggregation precedes transcriptional changes,[7,13,15] this was consistent with MSH3 ablation having no effect on transcriptional dysregulation in the zQ175 striatum. Therefore, in this subcellular location, $(CAG)_{185}$ in the *Htt* gene produced HTTexon1 at or above the threshold concentration needed to initiate aggregation, and further CAG expansion could not accelerate this process. It is striking that highly expanded CAG repeats (of greater than $(CAG)_{200}$) in humans cause onset of Huntington's disease before two years of age, whereas those of $(CAG)_{100–130}$ caused onset at around four years.[64,65] This suggests

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

that for CAG repeats approaching $(CAG)_{200}$, initiation of the pathogenic process does not require further expansion through somatic CAG repeat instability.

MSH3 has significant therapeutic potential for Huntington's disease and other trinucleotide repeat disorders,[34] and a wide range of MSH3-targeting modalities are under development.[29,66] Our data indicate that whilst ablation of MSH3 prevented somatic CAG repeat expansion in zQ175 mice, this model will not be useful for evaluating these MSH3-targeting treatments, and that knock-in mice with shorter CAG repeat mutations should be used. However, more importantly, we found that somatic expansions in zQ175 striatal neurons did not accelerate HTT aggregate pathology or transcriptional dysregulation, indicative of a pathogenic repeat threshold at, or close to, $(CAG)_{185}$. Early childhood onset cases of Huntington's disease with repeat expansions close to $(CAG)_{200}$, are consistent with a pathogenic threshold occurring at around this repeat length. Available data indicate that successful MSH3-targeting treatments will slow down the rate of somatic CAG repeat expansion in the brains of Huntington's disease mutation carriers. The prevention of somatic CAG repeat expansion is unlikely as that would require MSH3 ablation. The trajectory for repeat expansion in the Huntington's disease brain is not understood and these results indicate that the greatest benefit for MSH3-targeting therapeutics will be realised by treating as early as possible before cell-specific pathogenic thresholds have been achieved.

Acknowledgements

The authors wish to thank Jian Chen at the CHDI Foundation for advice on RNA sequencing and facilitating the data analysis. We also thank Kirupa Sathasivam and Pamela Farshim for help with sectioning brains and Casandra Gomez-Paredes for help with QuantiGene analysis.

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Funding

This work was supported by grants from the CHDI Foundation, UK Dementia Research Institute, which receives its funding from Dementia Research Institute Ltd, funded by the Medical Research Council, Alzheimer's Society and Alzheimer's Research UK and the Wellcome Trust (223082/Z/21/Z). Mice were generated with support from NIH grants R01CA248536 and 5P30CA13330.

Competing interests

JRG and MBH are employees of Rancho BioSciences and CLB is an employee of LoQus23 Therapeutics.

Data availability

RNA-seq data that support the finding of this study have been deposited in GEO with the accession no. GSE245510. The authors confirm that all other data supporting the findings in this study are available within the article and its Supplementary Material. Raw data will be shared by the corresponding author upon request.

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

1. Bates GP, Dorsey R, Gusella JF, et al. Huntington disease. *Nat Rev Dis Primers*. Apr 23 2015;1:15005. doi:10.1038/nrdp.2015.5
2. Duyao M, Ambrose C, Myers R, et al. Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nat Genet*. Aug 1993;4(4):387-92. doi:10.1038/ng0893-387
3. Telenius H, Kremer B, Goldberg YP, et al. Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. *Nat Genet*. Apr 1994;6(4):409-14. doi:10.1038/ng0494-409
4. Kennedy L, Evans E, Chen CM, et al. Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. *Hum Mol Genet*. Dec 15 2003;12(24):3359-67. doi:10.1093/hmg/ddg352
5. Sathasivam K, Neueder A, Gipson TA, et al. Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease. *Proc Natl Acad Sci U S A*. Feb 5 2013;110(6):2366-70. doi:10.1073/pnas.1221891110
6. Neueder A, Landles C, Ghosh R, et al. The pathogenic exon 1 HTT protein is produced by incomplete splicing in Huntington's disease patients. *Sci Rep*. May 2 2017;7(1):1307. doi:10.1038/s41598-017-01510-z
7. Smith EJ, Sathasivam K, Landles C, et al. Early detection of exon 1 huntingtin aggregation in zQ175 brains by molecular and histological approaches. *Brain Commun*. 2023;5(1):fcad010. doi:10.1093/braincomms/fcad010
8. Mangiarini L, Sathasivam K, Seller M, et al. Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell*. Nov 1 1996;87(3):493-506. doi:10.1016/s0092-8674(00)81369-0
9. DiFiglia M, Sapp E, Chase KO, et al. Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science*. Sep 26 1997;277(5334):1990-3. doi:10.1126/science.277.5334.1990
10. Hickman RA, Faust PL, Marder K, Yamamoto A, Vonsattel JP. The distribution and density of Huntingtin inclusions across the Huntington disease neocortex: regional correlations with Huntingtin repeat expansion independent of pathologic grade. *Acta Neuropathol Commun*. Apr 19 2022;10(1):55. doi:10.1186/s40478-022-01364-1
11. Waldvogel HJ, Kim EH, Tippett LJ, Vonsattel JP, Faull RL. The Neuropathology of Huntington's Disease. *Curr Top Behav Neurosci*. 2015;22:33-80. doi:10.1007/7854_2014_354
12. Rub U, Vonsattel JP, Heinsen H, Korf HW. The Neuropathology of Huntington s disease: classical findings, recent developments and correlation to functional neuroanatomy. *Adv Anat Embryol Cell Biol*. 2015;217:1-146.
13. Langfelder P, Cantle JP, Chatzopoulou D, et al. Integrated genomics and proteomics define huntingtin CAG length-dependent networks in mice. *Nat Neurosci*. Apr 2016;19(4):623-33. doi:10.1038/nn.4256
14. Hodges A, Strand AD, Aragaki AK, et al. Regional and cellular gene expression changes in human Huntington's disease brain. *Hum Mol Genet*. Mar 15 2006;15(6):965-77. doi:10.1093/hmg/ddl013
15. Landles C, Milton RE, Ali N, et al. Subcellular Localization And Formation Of Huntingtin Aggregates Correlates With Symptom Onset And Progression In A Huntington'S Disease Model. *Brain Commun*. 2020;2(2):fcaa066. doi:10.1093/braincomms/fcaa066
16. Wexler NS, Lorimer J, Porter J, et al. Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset. *Proc Natl Acad Sci U S A*. Mar 9 2004;101(10):3498-503. doi:10.1073/pnas.0308679101
17. Genetic Modifiers of Huntington's Disease C. Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. *Cell*. Jul 30 2015;162(3):516-26. doi:10.1016/j.cell.2015.07.003

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

18. Genetic Modifiers of Huntington's Disease Consortium. Electronic address ghmhe, Genetic Modifiers of Huntington's Disease C. CAG Repeat Not Polyglutamine Length Determines Timing of Huntington's Disease Onset. *Cell*. Aug 8 2019;178(4):887-900 e14. doi:10.1016/j.cell.2019.06.036
19. Lee JM, Chao MJ, Harold D, et al. A modifier of Huntington's disease onset at the MLH1 locus. *Hum Mol Genet*. Oct 1 2017;26(19):3859-3867. doi:10.1093/hmg/ddx286
20. Moss DJH, Pardinas AF, Langbehn D, et al. Identification of genetic variants associated with Huntington's disease progression: a genome-wide association study. *Lancet Neurol*. Sep 2017;16(9):701-711. doi:10.1016/S1474-4422(17)30161-8
21. Manley K, Shirley TL, Flaherty L, Messer A. Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. *Nat Genet*. Dec 1999;23(4):471-3. doi:10.1038/70598
22. Wheeler VC, Lebel LA, Vrbanac V, Teed A, te Riele H, MacDonald ME. Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum. *Hum Mol Genet*. Feb 1 2003;12(3):273-81. doi:10.1093/hmg/ddg056
23. Pinto RM, Dragileva E, Kirby A, et al. Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. *PLoS Genet*. Oct 2013;9(10):e1003930. doi:10.1371/journal.pgen.1003930
24. McAllister B, Donaldson J, Binda CS, et al. Exome sequencing of individuals with Huntington's disease implicates FAN1 nuclease activity in slowing CAG expansion and disease onset. *Nat Neurosci*. Apr 2022;25(4):446-457. doi:10.1038/s41593-022-01033-5
25. Loupe JM, Pinto RM, Kim KH, et al. Promotion of somatic CAG repeat expansion by Fan1 knock-out in Huntington's disease knock-in mice is blocked by Mlh1 knock-out. *Hum Mol Genet*. Nov 4 2020;29(18):3044-3053. doi:10.1093/hmg/ddaa196
26. Goold R, Flower M, Moss DH, et al. FAN1 modifies Huntington's disease progression by stabilizing the expanded HTT CAG repeat. *Hum Mol Genet*. Feb 15 2019;28(4):650-661. doi:10.1093/hmg/ddy375
27. Goold R, Hamilton J, Menneteau T, et al. FAN1 controls mismatch repair complex assembly via MLH1 retention to stabilize CAG repeat expansion in Huntington's disease. *Cell Rep*. Aug 31 2021;36(9):109649. doi:10.1016/j.celrep.2021.109649
28. Tabrizi SJ, Estevez-Fraga C, van Roon-Mom WMC, et al. Potential disease-modifying therapies for Huntington's disease: lessons learned and future opportunities. *Lancet Neurol*. Jul 2022;21(7):645-658. doi:10.1016/S1474-4422(22)00121-1
29. Benn CL, Gibson KR, Reynolds DS. Drugging DNA Damage Repair Pathways for Trinucleotide Repeat Expansion Diseases. *J Huntingtons Dis*. 2021;10(1):203-220. doi:10.3233/JHD-200421
30. Acharya S, Wilson T, Gradia S, et al. hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proc Natl Acad Sci U S A*. Nov 26 1996;93(24):13629-34. doi:10.1073/pnas.93.24.13629
31. Owen BA, Yang Z, Lai M, et al. (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. *Nat Struct Mol Biol*. Aug 2005;12(8):663-70. doi:10.1038/nsmb965
32. Dragileva E, Hendricks A, Teed A, et al. Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiol Dis*. Jan 2009;33(1):37-47. doi:10.1016/j.nbd.2008.09.014
33. Tome S, Manley K, Simard JP, et al. MSH3 polymorphisms and protein levels affect CAG repeat instability in Huntington's disease mice. *PLoS Genet*. 2013;9(2):e1003280. doi:10.1371/journal.pgen.1003280
34. Flower M, Lomeikaite V, Ciosi M, et al. MSH3 modifies somatic instability and disease severity in Huntington's and myotonic dystrophy type 1. *Brain*. Jun 19 2019;142(7):1876-86. doi:10.1093/brain/awz115
35. Iyer RR, Pluciennik A. DNA Mismatch Repair and its Role in Huntington's Disease. *J Huntingtons Dis*. 2021;10(1):75-94. doi:10.3233/JHD-200438

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

36. Valle L, de Voer RM, Goldberg Y, et al. Update on genetic predisposition to colorectal cancer and polyposis. *Mol Aspects Med*. Oct 2019;69:10-26. doi:10.1016/j.mam.2019.03.001
37. Heikkinen T, Lehtimaki K, Vartiainen N, et al. Characterization of neurophysiological and behavioral changes, MRI brain volumetry and 1H MRS in zQ175 knock-in mouse model of Huntington's disease. *PLoS One*. 2012;7(12):e50717. doi:10.1371/journal.pone.0050717
38. Menalled LB, Kudwa AE, Miller S, et al. Comprehensive behavioral and molecular characterization of a new knock-in mouse model of Huntington's disease: zQ175. *PLoS One*. 2012;7(12):e49838. doi:10.1371/journal.pone.0049838
39. Zeitler B, Froelich S, Marlen K, et al. Allele-selective transcriptional repression of mutant HTT for the treatment of Huntington's disease. *Nat Med*. Jul 2019;25(7):1131-1142. doi:10.1038/s41591-019-0478-3
40. Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*. Sep 12 2013;154(6):1370-9. doi:10.1016/j.cell.2013.08.022
41. Fienko S, Landles C, Sathasivam K, et al. Alternative processing of human HTT mRNA with implications for Huntington's disease therapeutics. *Brain*. Dec 19 2022;145(12):4409-4424. doi:10.1093/brain/awac241
42. Papadopoulou AS, Gomez-Paredes C, Mason MA, Taxy BA, Howland D, Bates GP. Extensive Expression Analysis of Htt Transcripts in Brain Regions from the zQ175 HD Mouse Model Using a QuantiGene Multiplex Assay. *Sci Rep*. Nov 6 2019;9(1):16137. doi:10.1038/s41598-019-52411-2
43. Landles C, Milton RE, Jean A, et al. Development of novel bioassays to detect soluble and aggregated Huntingtin proteins on three technology platforms. *Brain Commun*. 2021;3(1):fcaa231. doi:10.1093/braincomms/fcaa231
44. Lee JM, Zhang J, Su AI, et al. A novel approach to investigate tissue-specific trinucleotide repeat instability. *BMC Syst Biol*. Mar 19 2010;4:29. doi:10.1186/1752-0509-4-29
45. Marchionini DM, Liu JP, Ambesi-Impiombato A, et al. Benefits of global mutant huntingtin lowering diminish over time in a Huntington's disease mouse model. *JCI Insight*. Oct 24 2022;7(20)doi:10.1172/jci.insight.161769
46. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. doi:10.1186/s13059-014-0550-8
47. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. May 2000;25(1):25-29. doi:10.1038/75556
48. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. May 2012;16(5):284-287. doi:10.1089/omi.2011.0118
49. Mangiarini L, Sathasivam K, Mahal A, Mott R, Seller M, Bates GP. Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation. *Nat Genet*. 1997;15(2):197-200.
50. Kuhn A, Goldstein DR, Hodges A, et al. Mutant huntingtin's effects on striatal gene expression in mice recapitulate changes observed in human Huntington's disease brain and do not differ with mutant huntingtin length or wild-type huntingtin dosage. *Hum Mol Genet*. Aug 1 2007;16(15):1845-1861.
51. Obenauer J, Chen J, Andreeva V, et al. Expression analysis of Huntington's disease mouse models reveals robust striatum disease signatures. *bioRxiv*. 2022:479180. doi:10.1101/2022.02.04.479180
52. Kennedy L, Shelbourne PF. Dramatic mutation instability in HD mouse striatum: does polyglutamine load contribute to cell-specific vulnerability in Huntington's disease? *Hum Mol Genet*. 2000;9(17):2539-2544.
53. Ciosi M, Cumming SA, Chatzi A, et al. Approaches to Sequence the HTT CAG Repeat Expansion and Quantify Repeat Length Variation. *J Huntingtons Dis*. 2021;10(1):53-74. doi:10.3233/JHD-200433

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

54. Tsai Y-C, Greenberg D, Powell J, et al. Amplification-free, CRISPR-Cas9 targeted enrichment and SMRT sequencing of repeat-expansion disease causing genomic regions. *bioRxiv*. 2017:203919. doi:10.1101/203919
55. Hafford-Tear NJ, Tsai YC, Sadan AN, et al. CRISPR/Cas9-targeted enrichment and long-read sequencing of the Fuchs endothelial corneal dystrophy-associated TCF4 triplet repeat. *Genet Med*. Sep 2019;21(9):2092-2102. doi:10.1038/s41436-019-0453-x
56. Shelbourne PF, Keller-McGandy C, Bi WL, et al. Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. *Hum Mol Genet*. May 15 2007;16(10):1133-42. doi:10.1093/hmg/ddm054
57. Gonitel R, Moffitt H, Sathasivam K, et al. DNA instability in postmitotic neurons. *Proc Natl Acad Sci U S A*. Mar 4 2008;105(9):3467-3472. doi:0800048105 [pii]

10.1073/pnas.0800048105
58. Matlik K, Baffuto M, Kus L, et al. Cell type specific CAG repeat expansion and toxicity of mutant huntingtin in human striatum and cerebellum. *bioRxiv*. 2023:538082. doi:10.1101/2023.04.24.538082
59. Gusella JF, Lee JM, MacDonald ME. Huntington's disease: nearly four decades of human molecular genetics. *Hum Mol Genet*. Oct 1 2021;30(R2):R254-R263. doi:10.1093/hmg/ddab170
60. Scherzinger E, Sittler A, Schweiger K, et al. Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. *Proc Natl Acad Sci U S A*. Apr 13 1999;96(8):4604-9. doi:10.1073/pnas.96.8.4604
61. Benn CL, Landles C, Li H, et al. Contribution of nuclear and extranuclear polyQ to neurological phenotypes in mouse models of Huntington's disease. *Hum Mol Genet*. Oct 15 2005;14(20):3065-3078.
62. Cornett J, Cao F, Wang CE, et al. Polyglutamine expansion of huntingtin impairs its nuclear export. *Nat Genet*. Feb 2005;37(2):198-204.
63. Landles C, Sathasivam K, Weiss A, et al. Proteolysis of mutant huntingtin produces an exon 1 fragment that accumulates as an aggregated protein in neuronal nuclei in Huntington disease. *J Biol Chem*. Mar 19 2010;285(12):8808-8823. doi:M109.075028 [pii]

10.1074/jbc.M109.075028
64. Moeller AA, Felker MV, Brault JA, Duncan LC, Hamid R, Golomb MR. Patients With Extreme Early Onset Juvenile Huntington Disease Can Have Delays in Diagnosis: A Case Report and Literature Review. *Child Neurol Open*. Jan-Dec 2021;8:2329048X211036137. doi:10.1177/2329048X211036137
65. Gonzalez-Alegre P, Afifi AK. Clinical characteristics of childhood-onset (juvenile) Huntington disease: report of 12 patients and review of the literature. *J Child Neurol*. Mar 2006;21(3):223-229. doi:10.2310/7010.2006.00055
66. O'Reilly D, Belgrad J, Ferguson C, et al. Di-valent siRNA-mediated silencing of MSH3 blocks somatic repeat expansion in mouse models of Huntington's disease. *Mol Ther*. Jun 7 2023;31(6):1661-1674. doi:10.1016/j.ymthe.2023.05.006

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure legends

Figure 1 Characterisation of the *Msh3* knockout line. (A) *Msh3* knockout allele generation by CRISPR/Cas9 showing the position of guide RNAs (gRNA) targeting *Msh3* exons 3 and 8 and the position of the qPCR assays. **(B)** Expression of the *Msh3* mRNA was confirmed, using a qPCR assay with primers in the deleted region (*Msh3*_756), in the cortex of heterozygous *Msh3*+/- and homozygous *Msh3*-/- mice at 8 weeks of age (N = 5/genotype). **(C)** Western blotting with an MSH3 antibody showed that MSH3 levels were decreased in *Msh3*+/- heterozygotes and absent in *Msh3*-/- homozygotes. MSH3 levels were quantified by normalising to α-tubulin (N = 5/wild-type, N = 4/*Msh3*+/- and *Msh3*-/-). **(D)** Schematic of the genetic cross performed to generate the six experimental genotypes as littermates. One-way ANOVA with Tukey's post hoc test. Error bars = SEM. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. The test statistic, degrees of freedom and p values are summarised in Supplementary Table 3. Full-length blots are shown in Supplementary Fig. 5. WT= wild-type.

Figure 2 CAG repeat instability in zQ175 mice is dependent on *Msh3* expression. Representative GeneMapper traces of CAG repeat profiles for brain regions and peripheral tissues from mice at 6 months of age (N = 8/genotype). The modal CAG repeat length in ear samples at 12 days of age (P12) for zQ175 mice was 186 ± 2.3, for the zQ175:*Msh3*+/- mice was 185 ± 2.4 and for zQ175:*Msh3*-/- mice was 182 ± 2.0. The representative traces are from the same mouse for each genotype with the CAG repeat mode at P12 indicated by the dashed line. Somatic instability in 6-month-old zQ175 mice was most pronounced in the striatum, brainstem, spinal cord, and liver, and the repeat had contracted in the pituitary. Somatic CAG expansion was prevented in zQ175 mice in which *Msh3* was absent, and contractions had led to a reduction in the CAG repeat mode. Heterozygosity for *Msh3* led to a decrease in the extent of repeat expansion.

Figure 3 Tissue specific changes in the instability index and CAG repeat mode of the of the CAG repeat profiles in zQ175, zQ175:*Msh3*$^{+/-}$ and zQ175:*Msh3*$^{-/-}$ mice. (A) Instability index quantified for the GeneMapper traces for each brain region and tissue normalised to P12 ear sample for each genotype. A stepwise reduction in the instability index was observed for each brain region / tissue with the reduction and loss in *Msh3* expression. **(B)** Quantification of the change in mode from the P12 ear sample. Modal CAG was reduced with *Msh3* genotype, in all regions apart from striatum, heart and liver. N = 8/genotype, 2-3 technical replicates for each mouse. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. Box and whisker plots represent median and

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

25th and 75th percentiles with bars drawn from minimum to maximum. Statistical analysis was mixed-effects model (REML), with Tukey's correction for multiple comparisons. The test statistic, degrees of freedom and *p* values are summarised in Supplementary Table 4.

Figure 4 Statistical analysis of the effect of heterozygosity and homozygosity for *Msh3* knock-out on somatic CAG repeat instability in the striatum. (A-B) Scores and Loading plots from the PCA of all summary statistics using ear and striatal data. The plots indicate the statistical variables contributing to the differences seen between genotypes: contraction, mode, median, skew, mean, expansion, II = instability index, log(IQR) = log interquartile range, log(var) = log variance, kurtosis and N = sum of peak heights. **(C)** Least Squares Means plots of instability index, expansion, and variance summary statistics showing means (geometric means for variance) adjusted for ear and 95% confidence intervals for striatum across genotype. These data again confirm the significant effect of *Msh3* genotype on instability, expansion, and variation. *P*-value of effect test (genotype) shown.

Figure 5 Transcriptional dysregulation in the striatum of zQ175 mice is not alleviated by loss of *Msh3* expression. (A-D) Volcano plots illustrating gene expression changes between **(A)** zQ175 and wild-type mice **(B)** zQ175:*Msh3*-/- and wild-type mice, **(C)** zQ175 and zQ175:*Msh3*-/- mice and **(D)** *Msh3*-/- and wild-type mice, at 6 months of age. **(E)** Of the 2486 genes dysregulated in zQ175 mice as compared to WT mice at 6 months of age, 2408 showed negligible reversal, with a total of 15 genes showing full reversal, and 51 partial reversal, indicating that only 3% of genes were rescued, a level considered to be background. Of the known 266 genes dysregulated in striatal tissues across models, 258 were detected as dysregulated, with none reversed and 1 exacerbated. **(F-H)** Expression levels for **(F)** striatal genes of interest: *Cnr1, Drd1, Pde10a, Penk,* and *Ppp1r1b* and **(G)** the DNA mismatch repair genes *Msh3, Msh2, Msh6* plotted as variance stabilizing transformed counts derived using DESeq2. Box and whisker plots represent median and 25th and 75th percentiles with bars drawn from minimum to maximum. They show there was no change in expression due to loss of *Msh3*. *N* = 8/genotype, 4 males/4 females for each genotype except zQ175*:Msh3*-/- at 5 males/3 females. WT= wild-type.

Figure 6 Striatal extranuclear, but not nuclear, HTT aggregation is reduced with loss of *Msh3.* Coronal striatal sections from WT, zQ175, zQ175:*Msh3*+/- and zQ175:*Msh3*-/- at 3 and 6 months of age were immunohistochemically stained with the S830 antibody. A nuclear counterstain was not applied to these images, as this would mask the nuclear signal. **(A)** HTT

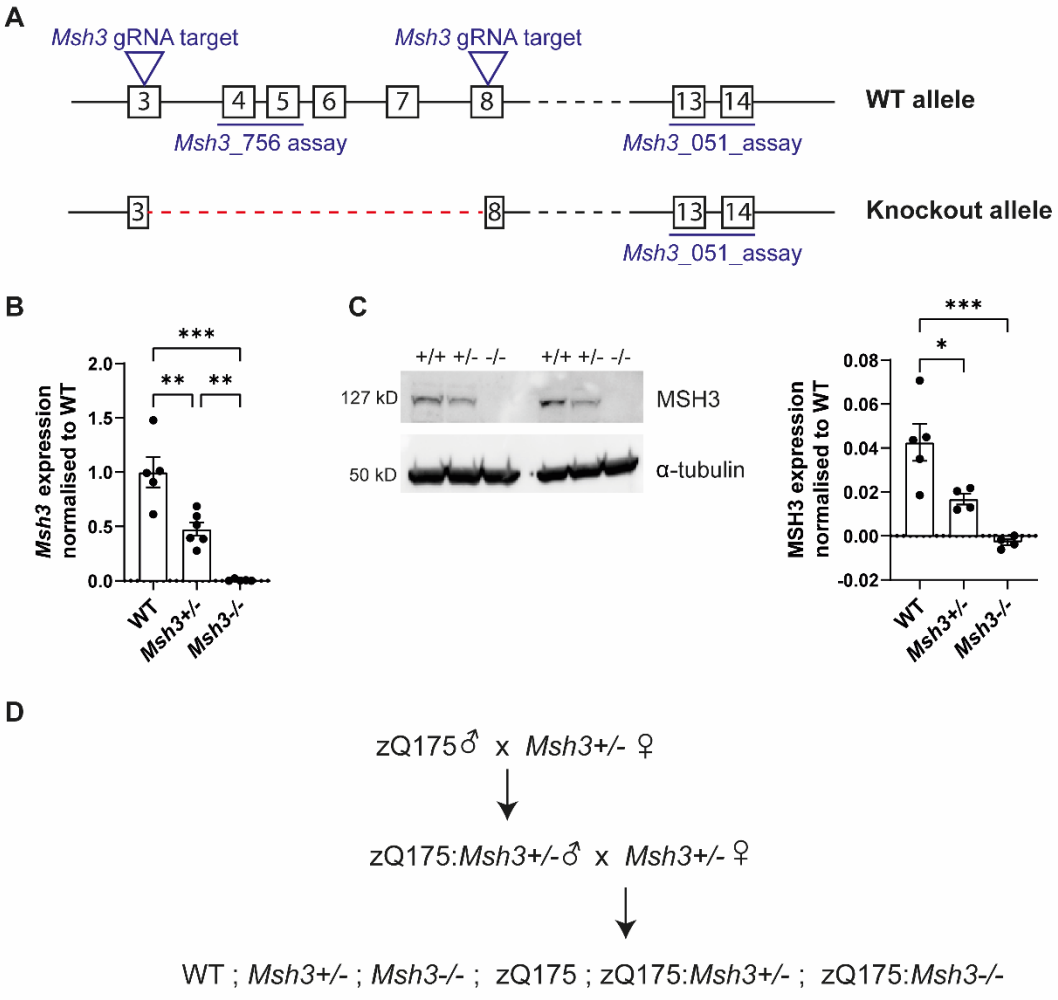
bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

aggregation was detectable by 3 months of age in the zQ175 mice and S830 immunostaining levels were comparable irrespective of *Msh3* genotype. **(B)** Thresholding was applied to images to quantify levels of staining of S830 positive objects, average area of object, and the percentage of image covered. The *Msh3* genotype had no impact on these three measures of S830 staining. **(C)** At 6 months of age, in addition to diffuse nuclear aggregation, nuclear inclusions were also prominent in many nuclei. **(D-E)** Thresholding was applied to quantify the number of **(D)** S830 positive objects and **(E)** inclusions. In each case, the number of objects, the average area of the objects, and the percentage of an image covered by the S830 signal was determined. One-way ANOVA with Bonferroni correction. Error bars = SEM. *N* = 5/zQ175 genotype, *N* = 3 WT. Scale bar in large image = 10 μm, scale bar in insert = 5 μm. M = months, WT = wild-type.

Figure 7 Striatal HTTexon1 aggregation is not affected by loss of *Msh3* expression. (A) Schematic of epitope location of antibodies used in HTTexon1 HTRF assays**. (B)** Levels of aggregated HTTexon1 as measured by the 4C9-MW8 assay indicated that aggregation increased from 2 to 6 months of age in zQ175 mice. There was no change in the levels of HTTexon1 aggregation detected between zQ175, zQ175:*Msh3*+/- and zQ175:*Msh3*+/- in the striatum, but small and significant differences were detected in the cortex and brainstem at 6 months of age. **(C)** High levels of soluble HTTexon1, as detected by the 2B7-MW8 assay, were present by 2 months of age in the zQ175 striatum, cortex and brain stem and had decreased at 6 months in all regions. There was no difference in soluble HTTexon1 at 6 months of age between the zQ175, zQ175:*Msh3*+/- or zQ175:*Msh3*-/- genotypes. 6-month samples: One-way ANOVA with Bonferroni correction; 2-month and 6-month zQ175 and WT: two-way ANOVA with Bonferroni correction. Error bars = SEM. *$P \leq 0.05$, ** $P \leq 0.01$, ***$P \leq 0.001$. *N* = 10/genotype. The test statistic, degrees of freedom and *p* values are summarised in Supplementary Table 5. AU= arbitrary units. WT= wild-type.

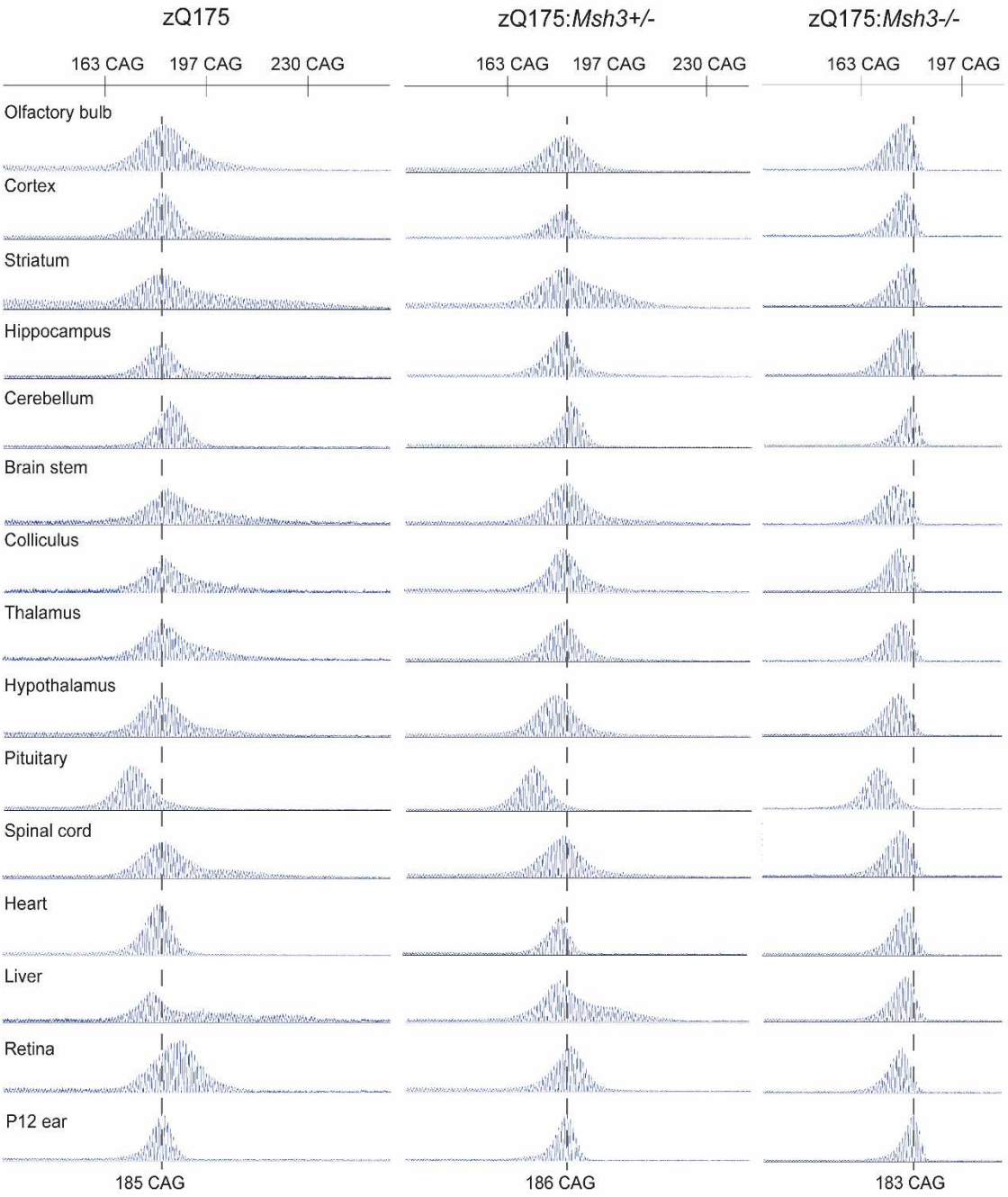
bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 1



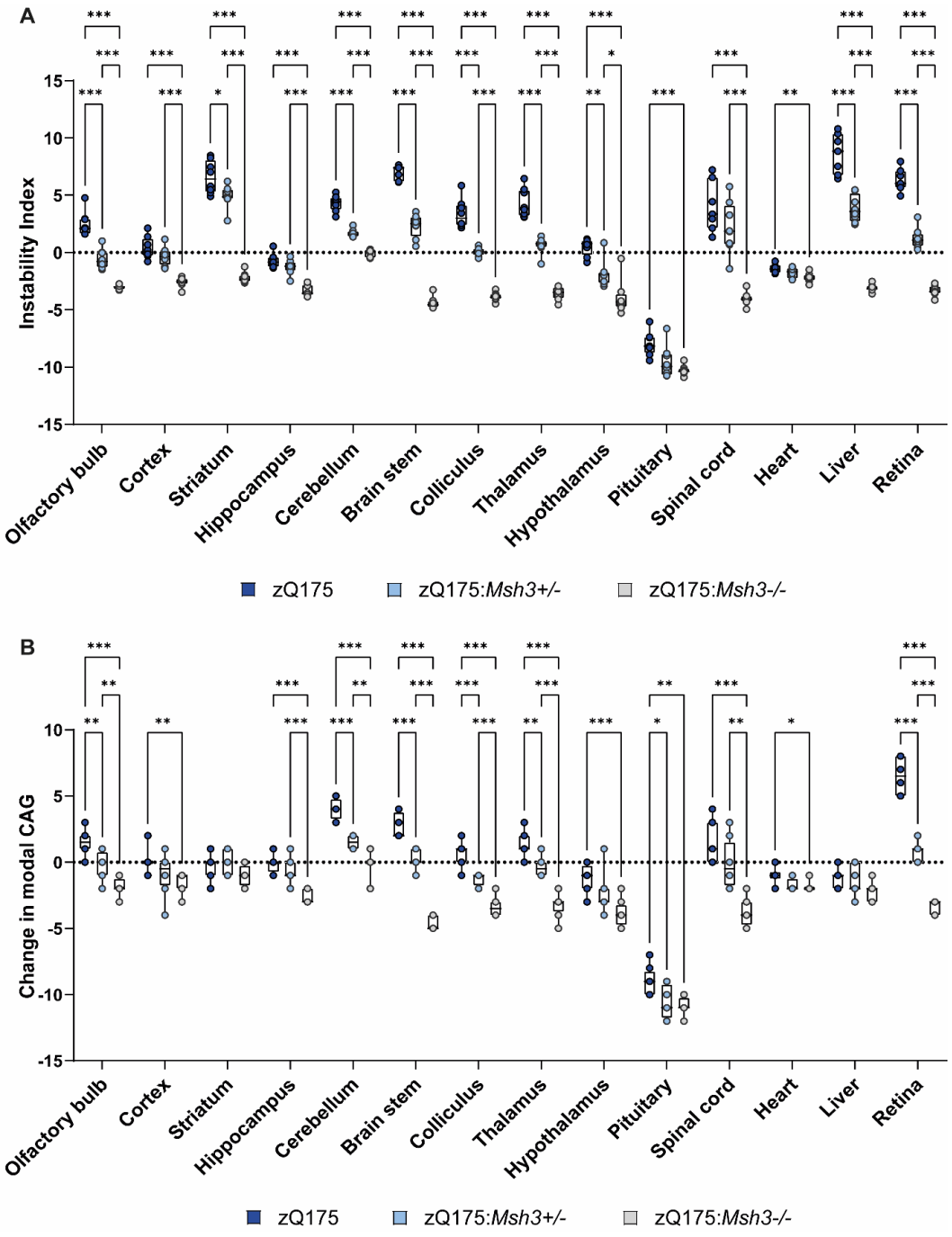
bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 2



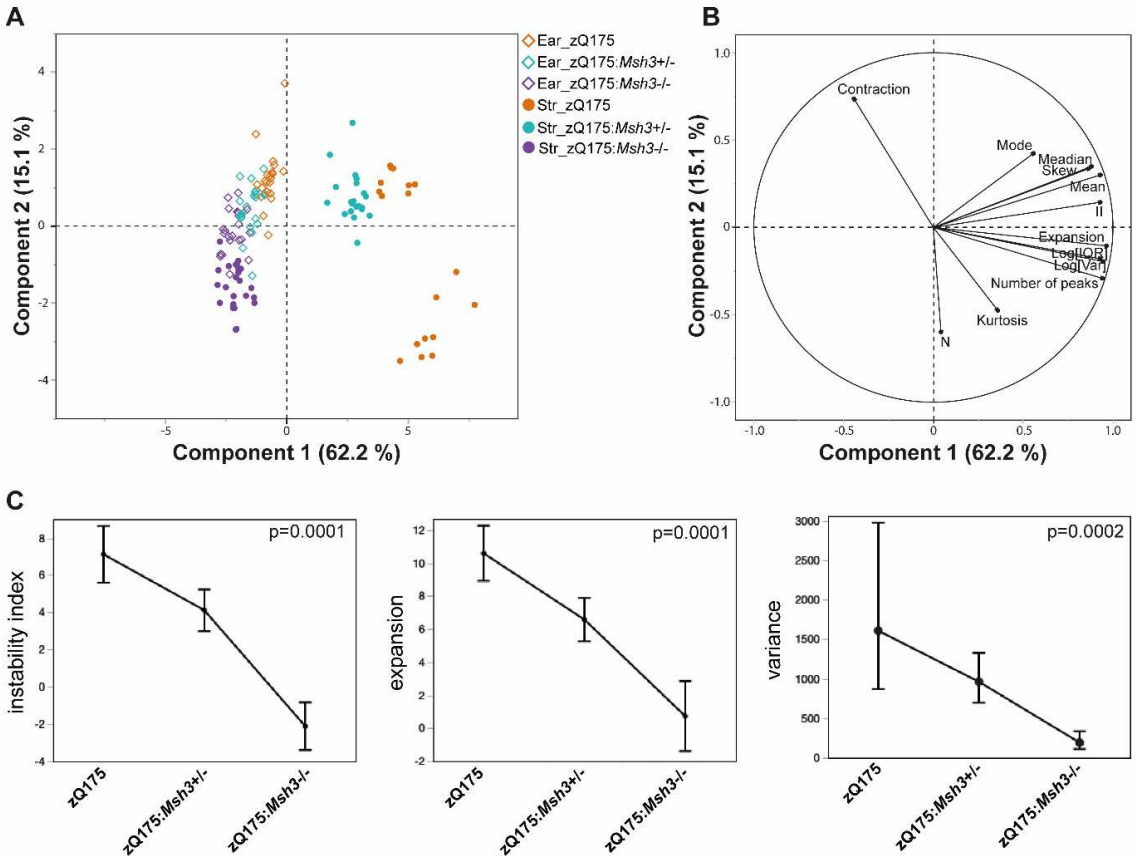
bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 3



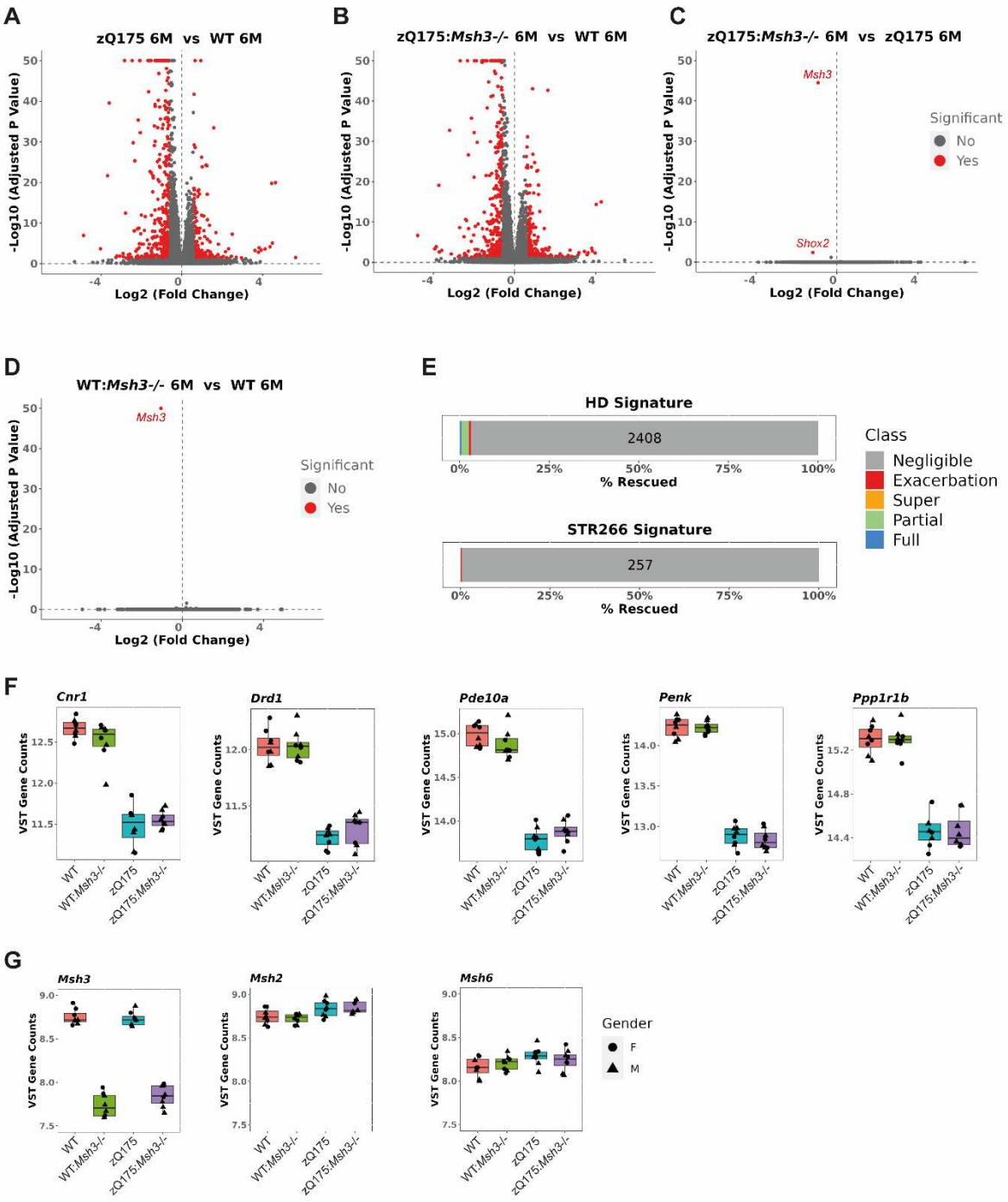
bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 4



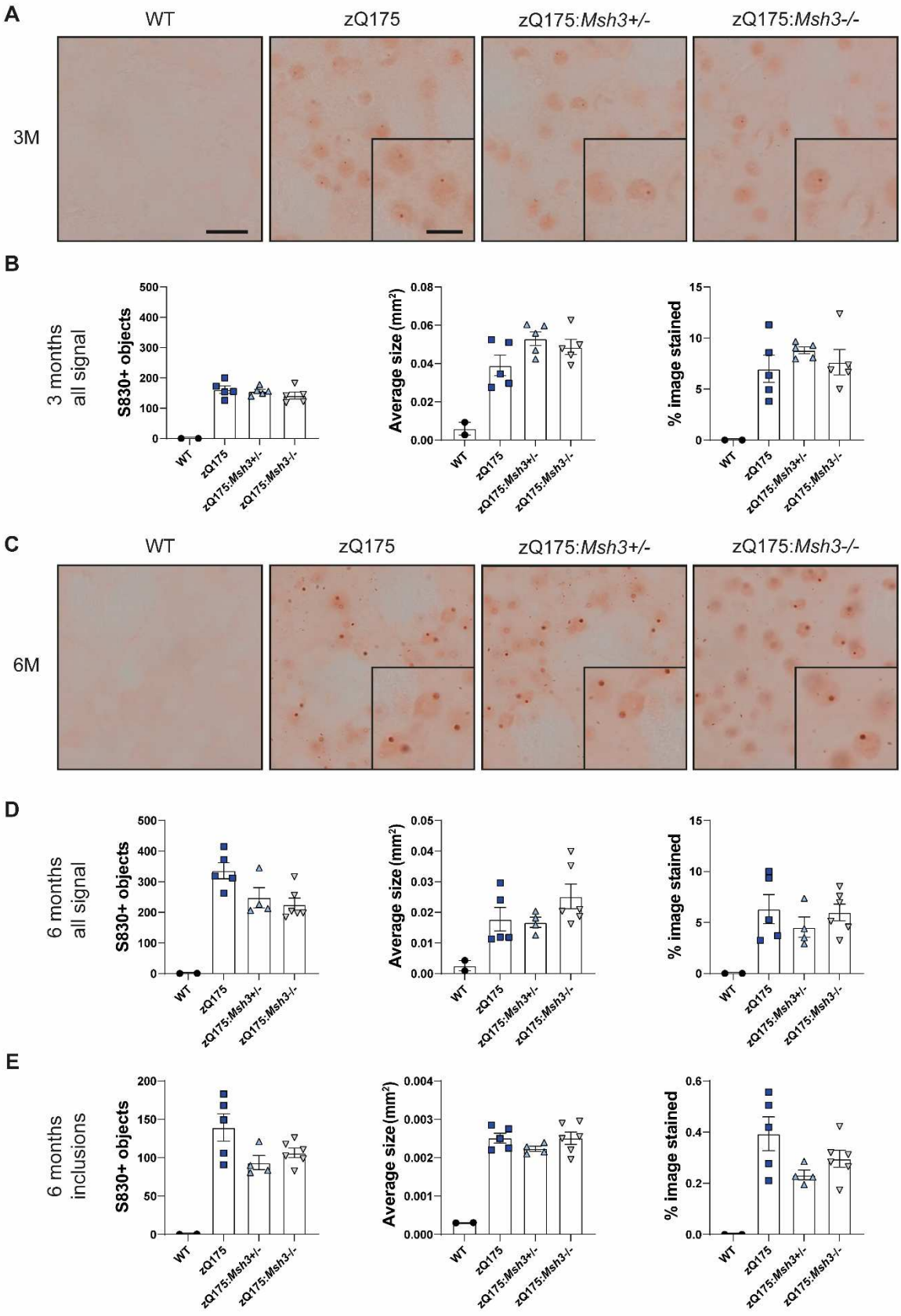
bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 5



bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 6



bioRxiv preprint doi: https://doi.org/10.1101/2023.12.15.571088; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

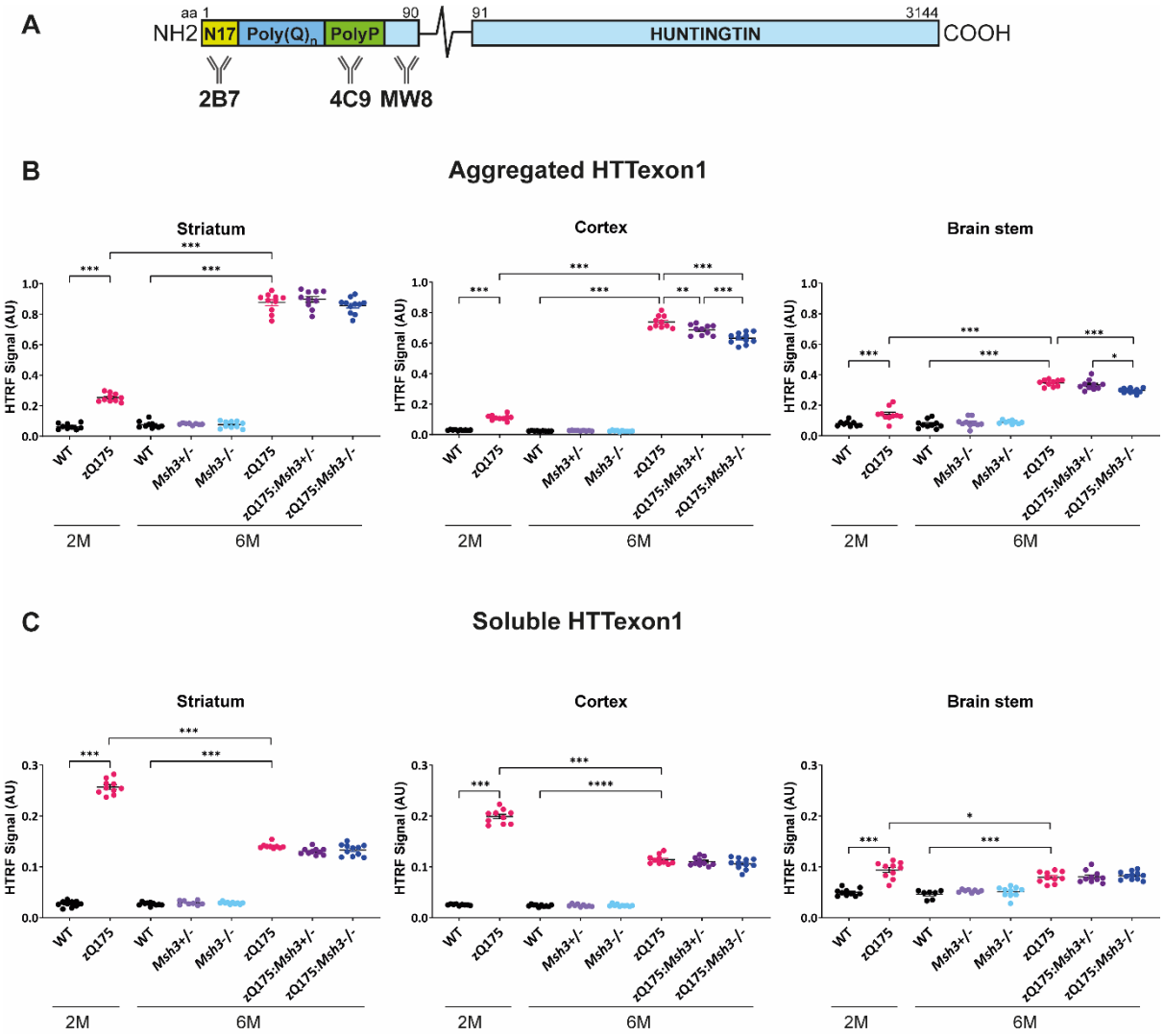



Figure 7