GAA•TTC repeat expansion in human cells is mediated by mismatch repair complex MutLγ and depends upon the endonuclease domain in MLH3 isoform one

Anasheh Halabi[1,†], Kayla T. B. Fuselier[2,†] and Ed Grabczyk[2,*]

[1]Division of Neurology, Department of Neurosciences, University of California, San Diego, CA 92103, USA and
[2]Department of Genetics, Louisiana State University Health Sciences Center, New Orleans, LA 70112, USA

ABSTRACT

DNA repeat expansion underlies dozens of progressive neurodegenerative disorders. While the mechanisms driving repeat expansion are not fully understood, increasing evidence suggests a central role for DNA mismatch repair. The mismatch repair recognition complex MutSβ (MSH2-MSH3) that binds mismatched bases and/or insertion/deletion loops has previously been implicated in GAA•TTC, CAG•CTG and CGG•CCG repeat expansion, suggesting a shared mechanism. MutSβ has been studied in a number of models, but the contribution of subsequent steps mediated by the MutL endonuclease in this pathway is less clear. Here we show that MutLγ (MLH1-MLH3) is the MutL complex responsible for GAA•TTC repeat expansion. Lentiviral expression of shRNA targeting MutL nuclease components MLH1, PMS2, and MLH3 revealed that reduced expression of MLH1 or MLH3 reduced the repeat expansion rate in a human Friedreich ataxia cell model, while targeting PMS2 did not. Using splice-switching oligonucleotides we show that MLH3 isoform 1 is active in GAA•TTC repeat expansion while the nuclease-deficient MLH3 isoform 2 is not. MLH3 isoform switching slowed repeat expansion in both model cells and FRDA patient fibroblasts. Our work indicates a specific and active role for MutLγ in the expansion process and reveals plausible targets for disease-modifying therapies.

INTRODUCTION

Genomic instability is a central tenet for a variety of diseases including malignancy, aging, and neurodegenerative disease. While DNA mismatch repair (MMR) is essential for maintaining genomic integrity, it has been shown that in the case of repetitive DNA sequences, MMR actually contributes to genome instability. Given the increasingly recognized pervasiveness of repeat expansion in neurodegenerative disease, understanding the mechanism by which the process is mediated is critical to the development of meaningful therapeutic interventions. MMR has been implicated in the repeat expansion underlying CAG•CTG expansion disorders such as Huntington's Disease (HD), myotonic dystrophy (DM) (1–3) and more recently, the GAA•TTC repeat expanded in Friedreich ataxia (4–6).

In the MMR pathway, there is a division of labor: a MutS recognition complex identifies a lesion and a MutL endonuclease makes a cut. MSH2 (MutS Homologue 2) is a component of both MutS heterodimer recognition complexes, which are responsible for identifying and binding mismatched bases and/or insertion/deletion loops (7). Upon mismatch recognition by a MutS complex, a MutL heterodimer is recruited to make an incision near the lesion recognition site (8). Under physiologic conditions, binding of MutL initiates recruitment of the necessary machinery that will excise the lesion and synthesize the DNA patch. Analogous to the role of MSH2 in MutS complexes, MLH1 (MutL Homologue 1) is the core subunit of known MutL complexes. MLH1 combines with one of three partners: PMS2 (post-meiotic segregation increased 2), PMS1, or MLH3 to form MutLα, MutLβ and MutLγ, respectively. It is estimated that about 90% of MLH1 in most human cells is bound to PMS2 (MutLα) (9,10); further, PMS1 and PMS2 are estimated to be present in 10-fold and 60-fold molar excess of MLH3 (10).

Friedreich ataxia (FRDA) is a progressive neurodegenerative disease caused by GAA•TTC repeat expansion in the frataxin (FXN) gene (11). The GAA•TTC expansion represses frataxin expression in a length dependent manner to cause FRDA. In prior work we have shown that GAA•TTC

---

expansion rate is associated with transcription rate within the repeat (12) and specifically requires the action of MMR complex MutSβ (MSH2-MSH3) but not MutSα (MSH2-MSH6) (5). Similarly, MutSβ and/or the MSH3 subunit in particular is also rate limiting for CAG•CTG expansion seen in Huntington's disease (HD), myotonic dystrophy (DM) and the CGG•CCG expansion seen in Fragile X syndrome (2,3,13,14) suggesting a shared MSH3 dependent expansion mechanism despite the sequence differences in the repeats.

Because nucleolytic cleavage by a MutL complex is the obligate next step after MutS recognition in MMR, we investigated the role of MutL nuclease complexes in GAA•TTC repeat expansion. We report here that shRNA mediated knockdown of MLH1 or MLH3 slows the rate of GAA•TTC expansion in model cell lines, while knockdown of PMS2 does not. Furthermore, repeat expansion rate is more sensitive to knockdown of MLH3 than it is to knockdown of MLH1. Finally, using splice-switching oligonucleotides (SSOs) we demonstrate a specific and active role for MutLγ in the GAA•TTC expansion process by showing that MLH3 isoform 1 is active in GAA•TTC repeat expansion while the nuclease deficient MLH3 isoform 2 is not. The SSO mediated switch from isoform 1 to isoform 2 slowed repeat expansion both in model cells and in FRDA patient derived fibroblasts.

MATERIALS AND METHODS

Plasmid construction

Knockdown constructs used the pLKO.1 vector system (Open Biosystems), which confers puromycin resistance and drives shRNA expression from a human U6 promoter. The shRNA sequences were chosen from The RNAi Consortium (TRC) TRC-Hs1.0 (human) shRNA library (Broad Institute). Four different sequences for each target were as follows: MLH1-1 (TRCN0000010380), MLH1-2 (V2HS_76497), MLH1-3 (V2HS_261056), MLH1-4 (TRCN0000040054), MLH3-1 (TRCN0000038737), MLH3-2 (TRCN0000038738), MLH3-3 (TRCN0000038736), MLH3-4 (TRCN0000038735), PMS2-1 (V2HS_93544), PMS2-2 (TRCN0000078549), PMS2-3 (TRCN0000078548) and PMS2-4 (V2HS_232932). Oligodeoxyribonucleotides were annealed to assemble the shRNA templates, which were cloned adjacent to the human U6 promoter between restriction sites AgeI and EcoRI in pLKO.1. Construction of the MSH3 expressing lentiviral construct PNL-MSH3-IRES2EGFP, which expresses both MSH3 and EGFP, has been detailed previously (5).

Cell culture

HEK 293 (15,16) cells, part of the Flp-In T-REx HEK 293 system from Invitrogen had a single copy construct bearing defined GAA•TTC repeats integrated via Flp recombinase as described (12). Long uninterrupted GAA•TTC repeats cannot be propagated in bacteria. Therefore we build repeat arrays for our constructs using an in vitro ligation strategy we devised (17) circumventing bacterial growth for long repeats and enabling the use of defined, uninterrupted repeats. Thus, non-FRDA patient derived repeats described in this work were made by serial ligation *in vitro*, circularized *in vitro* then directly co-transfected into human cells with a Flp recombinase expressing plasmid (12,18). Cell lines were derived independently. Cell lines were maintained in Dulbecco's modified Eagle's medium (DMEM) high glucose (Invitrogen) with 5% fetal bovine serum (Sigma) at 37°C in an atmosphere containing 5% CO2. Over time in culture the repeats expand to produce a population of sizes producing a characteristic smeared appearance in PCR sizing products (12). FRDA patient-derived fibroblasts GM04078 were from the Coriell Institute for Medical Research; FA4111 and FA6196 were obtained from the FRDA fibroblast repository (19). Fibroblasts were maintained in Dulbecco's modified Eagle's medium (DMEM) high glucose (Invitrogen) with 10% fetal bovine serum (Sigma) at 37°C in an atmosphere containing 5% CO_2.

Lentiviral production and cell transductions

Lentiviral particles were produced from vectors derived from the pLKO.1 vector (Open Biosystems). Viral particles were made via a three-plasmid expression system in HEK 293T cells as described (20) except that transfections used Lipofectamine 2000™ (Invitrogen) per the manufacturer's recommendations. Transductions were performed with HEK293 cells containing (GAA·TTC)$_{176}$ tandem reporter constructs or primary FRDA patient fibroblasts in the presence of 8 μg/ml of Polybrene (Sigma). Knockdowns were selected for with 1μg/ml of puromycin. Transduction with the MSH3 expressing lentiviral construct PNL-MSH3-IRES2EGFP was verified by EGFP fluorescence in the majority of fibroblasts.

Morpholino oligonucleotides and treatment regime

To effect splice switching in MLH3, oligonucleotide analogues were designed to mask the acceptor and donor splice sites surrounding MLH3 mRNA exon 7 to induce exon skipping of exon 7 and the consequent production of MLH3 iso2. Conjugated 'Vivo-morpholino' oligonucleotides (Gene Tools, LLC, Philomath, OR, USA) targeted the splice acceptor from exon 6 to exon 7 (MLH3acr6: 5′-TCCCACCTAGATGAGCAAGGATTGT-3′) and the splice donor from exon 7 to exon 8 (MLH3dnr8: 5′-TCTGGCTGCAAACAGATCCTTACCA-3′). FRDA model repeat expansion cells were treated with 1:1 mixtures of MLH3acr6 and MLH3dnr8 with total concentrations of 100nM, 500nM, 750nM in Opti-MEM Reduced-Serum Medium (Gibco) with 1% fetal bovine serum (Sigma) at 37°C in an atmosphere containing 5% CO_2. Low-serum media had the dual benefit of slowing cell division and not interfering with the morpholino activity. Controls were maintained in the same conditions. Treatments were given twice weekly for 4 weeks in the HEK 293 derived cells. For FRDA patient derived fibroblasts, SSO treatment was given once a week to confluent cell cultures, previously transduced with MSH3 expressing lentiviral construct PNL-MSH3-IRES2EGFP to induce repeat expansion (5). The SSOs were applied in Opti-MEM Reduced-Serum Medium (Gibco) with no serum added for 6 h to penetrate the cells. After 6 h additional Media was added with

supplemental FBS to bring the total FBS to 10%. Twenty four hours after SSO addition, the media was replaced with normal fibroblast maintenance media (DMEM high glucose with 10% FBS).

Reverse-transcription-PCR analysis of MLH3 isoforms

RNA was isolated from cells using TRIzol following the manufacturers protocol (Ambion). RNA was then used as a template to make cDNA with the High Capacity cDNA Reverse Transcription Kit following the manufacturers protocol (Applied Biosystems). PCR reactions for MLH3 isoforms consisted of 50 μl reactions with 2 μl of cDNA template, 200 nM primers, 250 μM each dNTP (Stratagene) and 2.5 U Paq5000 DNA polymerase (Stratagene). PCR was performed with 32 cycles of denaturation at 94°C for 30 s, primer annealing at 60°C for 15 s, and extension at 68°C for 60 s. Primers flanking MLH3 exon 7 were MLH3 × 7L3324 (5′-TCCTTTCCTTCCGAGAGCTC-3′) and MLH3 × 7R3757 (5′-TTTTCCGACCAGAGCCTTGT-3′) which produced 434 and 362 bp products from MLH3 isoform 1 and isoform 2, respectively. Products were resolved by electrophoresis on 1.4% agarose gels with 1KB Plus DNA Ladder (Invitrogen) as a marker. DNA was visualized with ethidium bromide and images were obtained using a Kodak Gel Logic 440 Imaging System.

PCR analysis of GAA•TTC repeats

Genomic DNA was isolated from cells using the PureLink Genomic DNA Minikit (Invitrogen) following the manufacturers protocol. Typically, 50 μl PCR reactions were performed with 100 ng of template, 200 nM primers, 250 μM each dNTP (Stratagene) and 2.5 U Paq5000 DNA polymerase (Stratagene). Primer pairs for model GAA•TTC expansions in HEK293 cells were TAN2767 (5′-GAGGACGCTGTCTGAAGTCC-3′) and MGR3537 (5′-TGAGCAACTGACTGAAATGCCTCAA-3′) annealed at 64°C for 15 s and extended at 68°C for 2 min for 32 cycles. These produced products that contained 667 base pairs of DNA flanking the repeat. PCR with FRDA patient derived fibroblasts was performed with primers 517F (5′-GGCTTGAACTTCCCACACGTGTT-3′) and 629R (5′-AGGACCATCATGGCCACACTT-3′) annealed at 64°C for 15 s and extended at 68°C for 2 min for 32 cycles. These primers produced PCR products from the endogenous FXN locus that added 498 base pairs of DNA flanking the repeat. Amplified products were resolved by electrophoresis on 1% agarose gels with the 1 Kb Plus DNA Ladder (Invitrogen) as a marker. DNA was visualized by staining with ethidium bromide and images were acquired with a Kodak Gel Logic 440 Imaging System. Images were analyzed with Carestream Molecular Imaging Software (5.0.2.26 for Mac OS). Bar graphs were created using KaleidaGraph 4.1 for Mac OS. Student's *t*-test for unpaired data with unequal variance was used for statistical analyses of GAA•TTC expansion.

Western blot

Analysis was performed with whole cell extracts as previously described (5). Briefly, cells were lysed in 4× Laemmli

Buffer (20% glycerol, 4% SDS, 100 mM Tris (pH 6.8), fresh 1 mM DTT) and boiled at 100°C for 10 min. Aliquots were separated on pre-cast 4–20% Tris-glycine gels (Bio-Rad Mini-PROTEAN TGX System) at 20 mA constant current. Protein was transferred to Immobilon-P (Millipore) with a Bio-Rad semi-dry transfer apparatus. Membranes were blocked for 45 min using 10% evaporated milk (Carnation) and 90% phosphate-buffered saline with 0.1% Tween (PBST). The membranes were incubated with primary antibodies MLH1 (BD Transduction Labs), PMS2 (BD Transduction Labs) and beta-actin (Sigma) in 10% milk/90% PBST for 90 min at room temperature or overnight at 4°C. After several washes with PBST the membranes were incubated for 45 min with horseradish peroxidase-conjugated goat anti-mouse (or goat anti-rabbit) secondary antibody (Pierce) in 10% milk / 90% PBST. After several PBST washes, the ECL Advance chemiluminescence kit (Amersham) was used to produce a signal captured by a Kodak Gel Logic 440 Imaging System. Image analysis was done with Carestream Molecular Imaging Software version 5.0.2.26 for Mac OS. Statistical analysis was done using the Student's t-test for unpaired data with unequal variance.

RESULTS

MutLγ is key to GAA•TTC repeat expansion in human cells

We used lentiviral-mediated shRNA knockdown of MLH1, PMS2 and MLH3 to probe the contribution of MutL to GAA•TTC repeat expansion. Lentiviral mediated shRNA knockdowns were carried out in four independent clones of HEK293 cells carrying a single copy of the tandem reporter vector bearing 176 GAA•TTC repeats (12). Each knockdown reflects a pool of four shRNA-expressing lentivirus. DNA and protein extracts were prepared after four weeks in culture as previously described (5). A representative PCR sizing gel is shown in Figure 1A. The original 176 triplets had expanded with time in culture for each clone, producing a population of sizes specific to each line at the start of experiments. Consequently, triplet gain rather than repeat length was compared. The number of triplet repeats gained in 4 weeks was calculated for all four cell lines (see (5)) and presented in graphical form in Figure 1B. The data reveal that DNA from MLH1 knockdown cells (lane: MLH1sh) have reduced expansion compared to the empty vector control cells (pLKO) at week 4, and that DNA from MLH3 knockdown cells exhibit the least repeat expansion (MLH3sh) in the same period of time. Both the MLH1 knockdown and MLH3 knockdown each show a statistically significant reduction in expansion rate. By contrast, PMS2 knockdown samples reveal a trend towards greater expansion (PMS2sh) that did not reach statistical significance. Thus, shRNA knockdown of MMR proteins MLH1 and MLH3, but not PMS2 slows GAA•TTC expansion in human cells, indicating a role for the MutLγ complex, a heterodimer of MLH1 and MLH3. Given data from PMS2 knockdown, MutLα (MLH1-PMS2) does not appear to have a role in promoting GAA•TTC repeat expansion.

Similar to MutS, MutL partners are more stable as heterodimers (9), and in the MutL system, MLH1 is the central common partner. Western blot analyses of protein extracts show the expected reduction of both PMS2 and MLH1 in

A

B
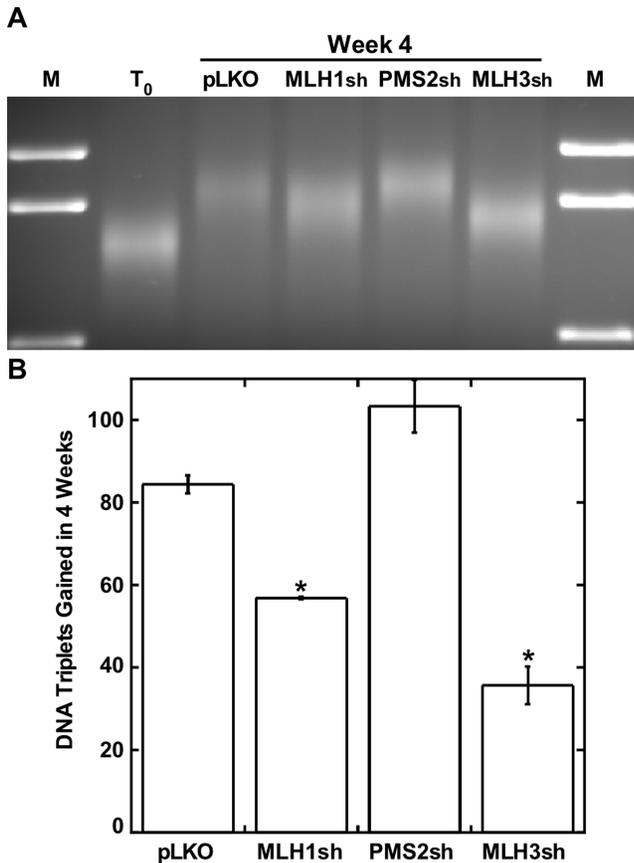
Figure 1. Knockdown of MLH1 or MLH3 significantly reduces GAA•TTC expansion rate in human cells. Four independent FRDA model lines were transduced with shRNA knockdown lentiviral pools targeting the indicated MutL subunits and an empty vector control virus (pLKO). (**A**) Representative gel image of PCR products measuring GAA•TTC expansion. T_0 is the day of transduction. PCR product sizes reflect 3X(number of repeats) plus 669 bp of flanking sequence. Lane M is the 1 Kb Plus DNA ladder showing 2, 1.6 and 1 Kb bands. (**B**) Mean repeats gained over four weeks. Knockdown of MLH1 and MLH3 reduced expansion compared to the empty vector control virus (pLKO). MLH1sh was significantly different ($P = 0.0009$) as was MLH3sh ($P = 0.00045$) whereas PMS2sh did not reach significance ($P = 0.053$). Error bars show SEM.
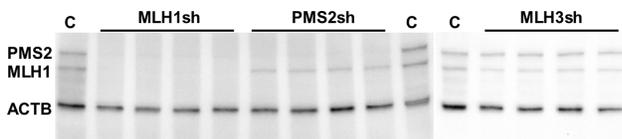
Figure 2. MLH3 knockdown does not appreciably alter MutLα expression. Western blot probed for PMS2, MLH1 and ACTB (β-actin) shows long-term lentiviral-mediated shRNA knockdown of MLH1, PMS2 and MLH3 protein in the cells used for DNA in the previous figure. Knockdown of MLH1 concurrently depletes PMS2 (lanes MLH1sh). Knockdown of PMS2 roughly halves MLH1 levels (lanes PMS2sh). Knockdown of MLH3 has little effect on MLH1 protein levels, and no effect on PMS2 levels (lanes MLH3sh). None of the MLH3 antibodies we tried were effective for western blots (not shown).

the MLH1 knockdown cells as compared to controls (Figure 2 compare lanes MLH1sh and control lanes 'C'). When PMS2 is knocked down, the protein level of MLH1 decreases but remains present (Figure 2 lanes PMS2sh). Im-

portantly, the MLH3 knockdown does not appreciably alter MLH1 or PMS2, suggesting that MutLα levels are not greatly affected by the loss of MLH3 in these cells. Taken together, these data indicate a central role for MutLγ in GAA•TTC expansion. It has been reported that endogenous expression of PMS2 is 60 times greater than MLH3 and that without enrichment, detection of MLH3 with commercially available antibodies is not feasible (10). We too are unable to detect MLH3 in western blots with the commercial antibodies that are currently available. In the absence of a useful antibody to demonstrate MLH3 knockdown at the protein level we sought another avenue to probe the role of MLH3 in repeat expansion.

MLH3 exon skipping provides an effective alternative to MLH3 knockdown

MLH3 is expressed in humans as two major isoforms differing by only 24 amino acids (Figure 3A). MLH3 isoform 1 includes exon 7, which contains a highly conserved endonuclease domain (21), while MLH3 isoform 2 lacks this 72 base exon but otherwise maintains the same translational frame (Figure 3A). Single amino acid mutations (asterisks in Figure 3B) in this conserved endonuclease domain have been shown to render mlh3 defective for both MMR and meiotic crossing over in yeast (21). Therefore we hypothesized that if the endonuclease activity of MLH3 isoform 1 is critical to repeat expansion, then preferential production of MLH3 isoform 2 by exclusion of exon 7 should approximate a functional knock out. We then sought to experimentally exclude MLH3 exon 7 in our model FRDA expansion cell lines as a means to functionally knockdown MLH3.

To skip MLH3 exon 7 we used a Splice Switching Oligonucleotide (SSO) strategy. With this, the SSO binds and masks the splice donor or acceptor signal flanking exon 7 in the unspliced pre-mRNA to cause exon skipping (Figure 4A). The SSO reagents used in this study are morpholino oligonucleotide analogues. Because of an uncharged backbone of morpholino subunits, morpholino oligonucleotides bind complementary target DNA or RNA very tightly and are resistant to nucleases and proteases (22). To enhance cellular uptake, we use a variant of these called 'vivo morpholinos' that have an octaguanidine moiety conjugated to the morpholino oligomer (23). We then designed DNA oligonucleotides to assay the ratio of MLH3 isoform 1 and MLH3 isoform 2 mRNA via reverse transcription PCR (RTPCR) to assess the efficacy of our SSOs. These primers produced PCR products that spanned MLH3 exon 7 corresponding to splice variants shown on the right of Figure 4A which differ by 72 bases and can be resolved in an agarose gel as seen in Figure 4B.

We tried different concentrations of acceptor SSO, donor SSO and combinations of the two in order to arrive at an effective dose. An example of these experiments is shown in Figure 4B. Masking the splice acceptor from exon 6 to exon 7 (acceptor lanes) or the splice donor from exon 7 to exon 8 (donor lanes) both resulted in varying degrees of exon 7 exclusion in the several hundred nanomolar range (Figure 4B). However, efficacy of splice switching to isoform 2 was enhanced when both donor and acceptor regions were simultaneously targeted despite halving the concentration of
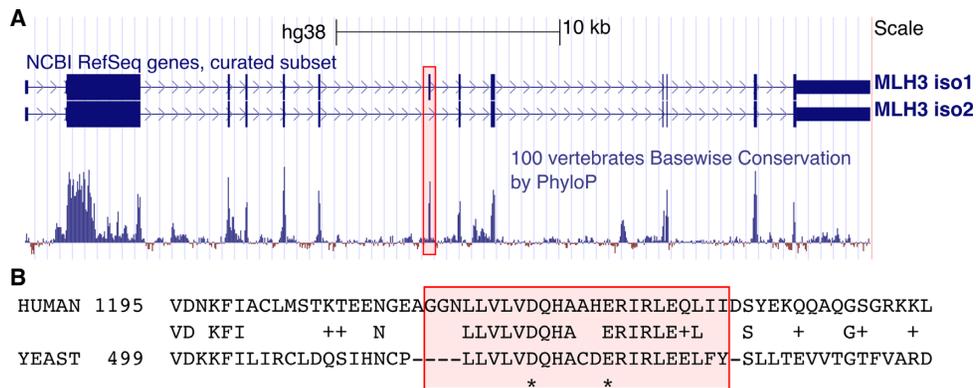
Figure 3. Human MLH3, a component of MutLγ, is expressed as two isoforms. (**A**) Image from UCSC Genome Browser (68) shows the two major human MLH3 isoforms. MLH3 isoform 1 includes exon 7; isoform 2 lacks this highly conserved 72 base exon (highlighted). Below the RefSeq track is the 100 vertebrates conservation track. The height of the bars indicates the degree of conservation, and exon 7 is among the most conserved. (**B**) BLASTP alignment (69) shows striking conservation between yeast and human MLH3 within the 24 amino acids coded by the human exon 7. This corresponds to a conserved endonuclease domain and mutations in it such as *mlh3-D523N* or *mlh3-E529K* (asterisks at bottom) have been shown to be defective for both MMR and meiotic crossing over in yeast (21).
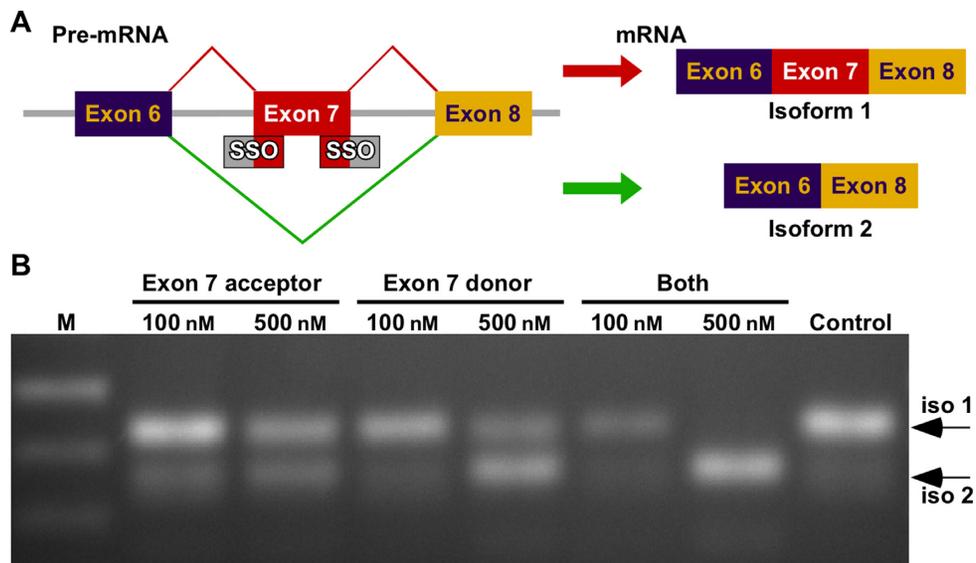
Figure 4. Splice switching oligonucleotides (SSOs) mask the acceptor or donor region of MLH3 exon 7 in pre-mRNA to promote exon skipping. (**A**) Design strategy. Morpholino antisense oligonucleotides were designed to bind the intron–exon junction in MLH3 pre-mRNA at either the 'acceptor' region (exon 6 side) or the 'donor' region (exon 8 side) of MLH3 exon 7. These SSOs were used to induce skipping of exon 7 and preferential production of MLH3 splice isoform 2 (iso2). (**B**) Gel image shows an example of SSO mediated splice switching as revealed by RT-PCR 48 h after treatment. RT-PCR products spanning exon 6 to exon 8 produced a 434 or 362 bp product from MLH3 isoform 1 or isoform 2, respectively (arrows at right). FRDA rapid expansion model cells in culture were treated with acceptor, donor or an equimolar mixture of both SSOs at the indicated total SSO concentrations. Lane M is the 1 Kb Plus DNA ladder showing 500, 400 and 300 bp.

the individual morpholino species (see Figure 4B 500 nM lanes). Consequently, we moved forward with experimental use of an equimolar combination of the two SSO reagents.

Preferential expression of MLH3 isoform 2 leads to slowed GAA•TTC repeat expansion

To test whether SSO mediated exon skipping would change repeat expansion rate we performed experiments in which FRDA model cell lines harboring GAA•TTC repeats were dosed repeatedly with SSOs at various concentrations over the course of four weeks. Typically, cells were fed fresh media containing SSOs twice a week and split as necessary. The

model cells are derivatives of HEK 293, and under normal growth conditions divide rapidly. For these experiments, we used media containing 1% serum that had the dual benefit of slowing division of the cells and allowing for better uptake of the vivo-morpholinos, which can be inhibited by serum.

We found that SSO treatment applied over four weeks slowed GAA•TTC repeat expansion in the model cells in a dose-dependent way (Figure 5). A representative gel is shown in Figure 5A. Application of equimolar donor and acceptor SSOs at the indicated total SSO concentration resulted in a dose-dependent slowing of expansion over the four-week treatment time. The averaged response to

A

B
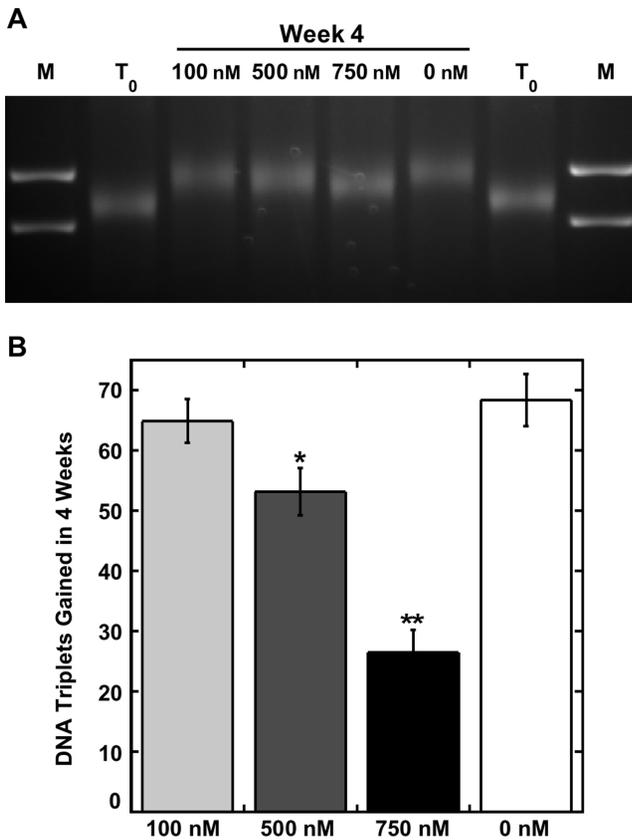
Figure 5. Dose-dependent reduction of GAA•TTC expansion rate with SSO treatment. (**A**) Representative gel image shows PCR sizing of GAA•TTC repeats at time zero (T_0) and at week 4 with paired SSO treatments at indicated concentrations. PCR product size equals 669 bp flanking sequence + 3x(repeat #). Lane M is the 1 Kb Plus DNA ladder showing 2 Kb and 1.6 Kb bands. (**B**) Mean repeats gained over four weeks. Samples treated with 750 nM of SSOs gained significantly less repeats than all other treatment groups (**P < 0.001), while the 500 nM treatment group was significantly less than the 0 nM control (*P = 0.038). Error bars equal SEM.

the treatment regimen by four different isolates bearing GAA•TTC repeats was calculated for all four cell lines (5) and is presented in graphical form as the number of triplet repeats gained in 4 weeks (Figure 5B). Cells treated with 750 nM SSO showed a more robust decline in repeat expansion than those treated with 500 nM SSO. In general, the SSO concentrations that effectively switch MLH3 to isoform 2 also slowed GAA•TTC repeat expansion (compare Figures 4 and 5). However, while the effect on isoform switching was fairly robust for mixed SSO concentrations at and above 500 nM at 48 h, during the course of experiments we found that a dose above 500 nM was superior for slowing repeat expansion. One difference between the two measures is that Figure 4B provides a snap shot of SSO activity on MLH3 isoforms 48 h after treatment while the SSO effects upon repeat expansion were exerted over the course of four weeks. It is possible that SSO concentrations were subject to some fluctuations in the long-term experiments.

The model cells we developed show repeat expansion in an experimentally useful time frame, however, the GAA•TTC repeat is not in the endogenous FXN locus.

Therefore, we tested the effect of altering MLH3 isoform ratios on repeat expansion in FRDA patient derived cells. GAA•TTC repeats are fairly stable in primary fibroblasts, which lack detectable expression of MSH3, but we have shown that ectopic expression of MSH3 will cause the repeats to expand (5). Primary FRDA patient-derived fibroblasts GM04078, FA4111 and FA6196 were transduced with lentiviral vectors expressing MSH3 to accelerate repeat expansion (5) and allowed to reach confluency, which stopped cell division by contact inhibition. The confluent cells were then treated once a week with 750 nM SSOs for six weeks. These treatments were effective by two measures. First, SSO treated cells showed a shift to MLH3 isoform 2 as measured by RT-PCR, while the untreated cells exhibited varying ratios of isoforms (Figure 6A). More importantly, the SSO treatment slowed expansion at the FXN locus in the alleles that expanded in each cell line over the course of six weeks (Figure 6B). Repeat expansion varied between individuals. For GM04078, both alleles expanded at a similar rate, while for FA4111 and FA6196 expansion of the small allele was slower than expansion of the large allele. Suppression of GAA•TTC repeat expansion by SSO treatment in these cells was a reproducible trend, but did not reach statistical significance due to the variability of repeat expansion rate both between individual patient derived lines and the variability of expansion rate between the alleles within individual lines.

In sum, we found that knocking down MLH3 via shRNA caused a robust decline in GAA•TTC repeat expansion rate with little impact on the constituent protein subunits of MutLα, thereby implicating MutLγ as the active agent in repeat expansion. Further, we showed that it is the endonuclease activity contributed by MLH3 that is responsible for repeat expansion by excluding only the nuclease domain of MLH3 via morpholino mediated splice switching.

DISCUSSION

We have previously found that transcription through a GAA•TTC repeat and the MutSβ recognition complex are both needed to drive GAA•TTC expansion (5,12). In that previous work and in this work we used a unique human cellular DNA repeat expansion model that facilitates such studies. We adapted the Invitrogen Flp-In™ T-REx 293 system to make cells with a single GAA•TTC repeat tract in a consistent chromosomal location and orientation. In these cell lines GAA•TTC expansion is not a rare event; it occurs nearly synchronously in the majority of cells (5,12). This model has a key advantage over authentic FXN repeat expansion: the repeats are not linked to an essential gene. Thus, we are able to avoid the well-known problems associated with selection against frataxin knockdown cells in culture (24). Hence there is no counter selection against expansion or 'brakes' on the rate of expansion. However, because the model does not have the repeat array in the endogenous FXN gene, select repeat expansion experiments can be validated in patient-derived cells; the downside being that it requires several months to complete such experiments. Other advantages of the model over primary cells include a single copy genomic location, the ability to control transcription into the repeat, and a cell environment permissive for expan-
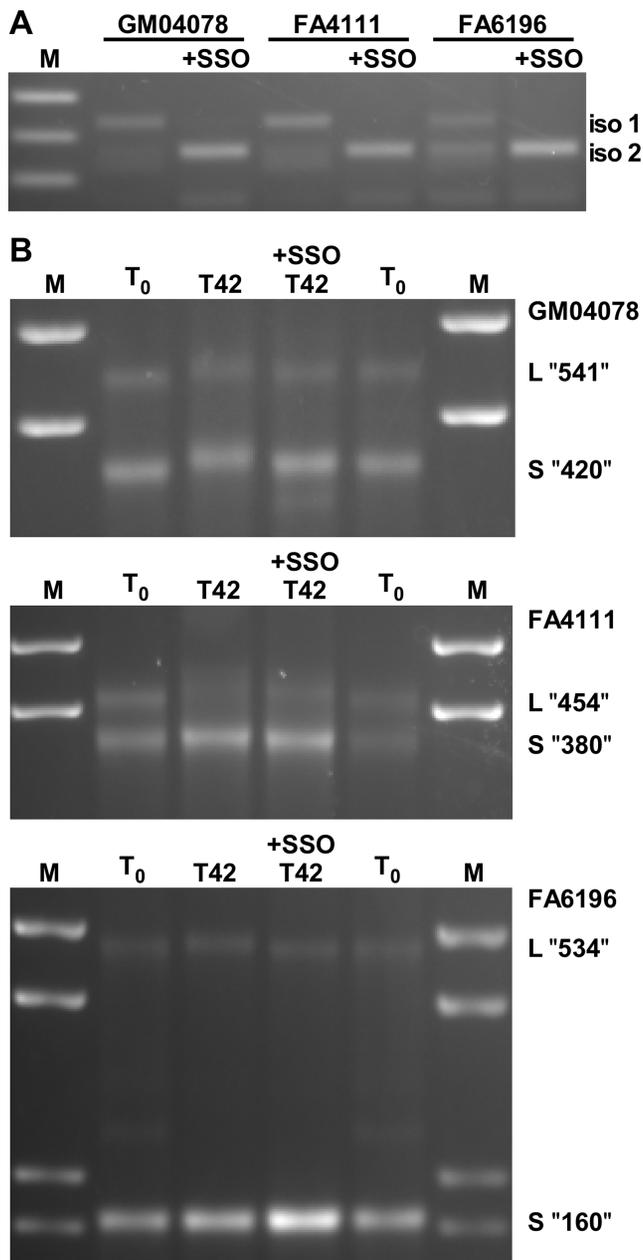
Figure 6. Weekly treatment with SSOs slows GAA•TTC expansion in non-dividing FRDA patient-derived fibroblasts. (**A**) Gel image shows an example of SSO mediated splice switching as revealed by RT-PCR 48 h after a treatment for FRDA patient derived fibroblasts. RT-PCR products spanning exon 6 to exon 8 produced a 434 or 362 bp product from MLH3 isoform 1 or isoform 2, respectively. Primary FRDA patient-derived fibroblasts GM04078, FA4111 and FA6196 were transduced with lentiviral vectors expressing MSH3 to accelerate expansion (5) and allowed to reach confluency, which stopped cell division by contact inhibition. The confluent cells were then treated once a week with 750 nM SSOs for 6 weeks. Lane M is the 1 Kb Plus DNA ladder showing 500, 400 and 300 bp. (**B**) Gel images show PCR sizing of GAA•TTC repeats for three different FRDA patient derived fibroblasts after six weeks in culture with and without SSO treatment. One set of cells was taken up at the time of the first treatment in the experiment (T_0) to provide a baseline repeat size. The T_0 samples were loaded to flank the control and SSO treated 6-week samples. Alongside the gel images 'L' represents the longer and 'S' the shorter allele for each patient with number indicating the nominal size of the alleles originally reported for these cells. The primers used add 498 bp to the size of the GAA•TTC repeat. M is a 1 Kb Plus DNA ladder showing 2 Kb and 1.6 Kb bands as well as 1 Kb and 850 bp for FA6196.

sion. Although HEK stands for human embryonic kidney, HEK 293 cells are neuronal in nature and express a large number of proteins typically expressed exclusively, or preferentially, in neurons (15,16). Our ability to study expansion with an accelerated time course is possibly aided by the neuronal nature of HEK293.

Using a combination of model cells and FRDA primary fibroblasts we have now determined that the next necessary step in the process leading to GAA•TTC repeat expansion is likely a nucleolytic incision by MutLγ (MLH1-MLH3). This new data is integrated in our working model for somatic repeat expansion shown in Figure 7. To move from a stable duplex in Figure 7A to an incrementally expanded duplex in 7F the DNA must first be unpaired to allow structure formation. Most somatic expansion pertinent to disease occurs in post mitotic tissues (3,25–27) thus un-pairing of repetitive DNA likely results from transcription as opposed to replication. Transcription into a DNA repeat can lead to formation of a structure. In Figure 7B, one such potential structure featuring an RNA•DNA hybrid, or R-loop, is shown. We have shown that very stable RNA•DNA hybrids form during transcription in GAA•TTC repeats (28) and in other purine•pyrimidine repetitive sequences (29). Other groups have subsequently verified the potential for RNA•DNA hybrid formation in several other repeats in addition to GAA•TTC (30–32). Out of register re-annealing by the repeats during resolution of the RNA•DNA hybrid can lead to loop-outs as shown in Figure 7C. These loop-outs resemble the slipped strand structures formed by tandem repeats re-annealed *in vitro* (33,34), and may attract MMR proteins such as MutSβ (35).

MutSβ (MSH2–MSH3), which has a preference for 3–12 base loops (36–38) is shown binding a 6 base loop in Figure 7D. Work supporting a role for MSH2 in GAA•TTC expansion (4–6,39) as well as the role of MSH3 (5) have helped unify repeat expansion mechanisms across different repeat sequences. In work focused on CAG•CTG repeats, MSH2 (1,40) and MSH3 (2,13,41) were implicated as necessary for repeat expansion in mouse models. Region specific CAG•CTG expansion in the mouse model has also been connected to MSH3 expression (42,43). CGG•CCG expansion in mouse models is more complicated, but MSH3 is responsible for the bulk of expansions (14,44). Thus, MutSβ (MSH2-MSH3) has a central, and well-supported, role in DNA repeat expansion regardless of repeat sequence.

Our data indicate that the step following MutSβ recognition, and actually leading to repeat expansion, is action by MutLγ (Figure 7E) with an intact endonuclease domain in the MLH3 subunit. We surmise that the cut made by MutLγ is then repaired in a manner that leads to an incremental increase in the number of repeats as reflected in Figure 7F. Data indicate that while *S. cerevisiae* MutLγ does not have an intrinsic affinity for mismatched DNA (45), it may be recruited and stimulated by MutSβ (46). Interactions have also been shown between MLH3, MLH1, and MSH3 (47–49) suggesting a means of recruitment. Further, MutLγ has been shown as necessary for all somatic expansions in an HD mouse model (50), substantiating a role for MutLγ that can be adapted more generally in DNA repeat expansion.
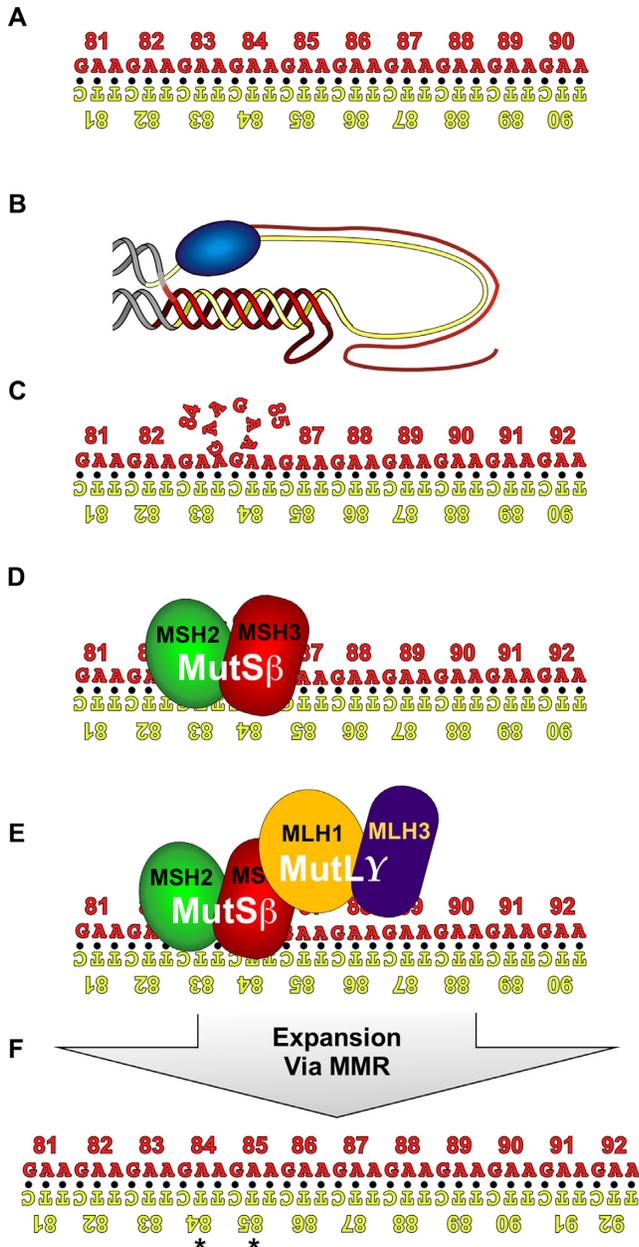
Figure 7. Working model for transcription initiated DNA repeat expansion via mismatch repair. (**A**) Part of a GAA•TTC repeat is depicted with the purine (or **R**) strand in red, and the pyrimidine (or **Y**) strand in yellow. The numbered bases show alignment in register. (**B**) During transcription the two strands are separated and a variety of structures may form. One likely example (28), an RNA•DNA hybrid, is shown. (**C**) Resolution of a structure can lead to an out-of-register re-annealing within the repeat. (**D**) The small loop formed becomes a target for binding by mismatch repair complex, MutSβ (MSH2-MSH3). (**E**) MutSβ in turn recruits MutLγ (MLH1-MLH3), and endonuclease activity contributed by MLH3 isoform 1 is required to initiate subsequent steps. Recruitment of the more prevalent MutLα, which cuts near the recognition complex (8,56), does not lead to expansion, suggesting that MutLγ cuts differently, possibly at a distance (57). (**F**) Repeat expansion has occurred with the addition of two trinucleotides (*) after repair synthesis targeted by MutSβ recognition and initiated by MutLγ endonuclease activity.

Contrary to MutLγ, we found that MutLα (MLH1-PMS2) does not contribute to repeat expansion. In fact, knockdown of PMS2 shows a trend to increased expansion rate. This result agrees with findings by other groups that have looked at the effect of PMS2 on repeat expansion in transgenic mouse models of FRDA (4,51) that indicated depletion of PMS2 causes an acceleration of repeat expansion. These results are at odds, however, with earlier findings indicating PMS2 depletion caused a reduction in the rate of CAG•CTG expansion (52).

PMS1, PMS2, and MLH3 have all been reported to compete for the same binding site on the C-terminal end of MLH1 (47); as with the MutS homologues, it is possible that the abundance of these proteins in relation to one another regulates their ability to compete for MLH1 and hence, their stability. Expression of MLH1 and PMS2 correlate strongly as reflected in our protein studies. We believe that the slightly increased repeat expansion rate seen with PMS2 knockdown may reflect more opportunity for MLH3 to form complexes with MLH1 due to reduced competition with PMS2. In addition to its already low abundance, MLH3 expression may not be tethered to the expression of other MutL proteins (10). We cannot address the potential changes in MLH3 protein directly given the lack of a commercially available antibody for MLH3 effective for western blots. Despite this, we were able to determine that knockdown of MLH3, while slowing repeat expansion has little impact on MLH1 and PMS2 levels. This is particularly relevant as predominantly expressed subunits of the MutL pathway are known to have a strong association with malignancies, particularly MLH1 (53–55).

Based on our data, we suspect that MutLγ is likely to be the active cause of repeat expansion, while MutSβ is the recruiting signal. Given the low abundance of MLH3, it is likely that MutLα (MLH1-PMS2) is the complex that is most frequently acting downstream of MutSβ, yet the MutSβ to MutLα pathway does not cause repeat expansion. This conclusion is consistent with our data showing reduced GAA•TTC expansion when MLH3 is knocked down despite MutLα expression remaining in relative abundance. Furthermore, this theory is consistent with the complete lack of CAG•CTG expansion seen in the MLH3 knockout mouse (50). This also raises the critical question of the differences between how MutLα and MutLγ process DNA downstream of MutSβ to yield such disparate outcomes in DNA repeats. It is possible that there are differences in accessory proteins that associate with the two complexes or differences in the way that the two complexes cut DNA.

Available evidence supports the idea that there is a difference in the location of the nick introduced by each complex. MutLα has been shown to introduce a break near the lesion recognition site (8,56) whereas exciting recent work by Manhart et al. indicates that MutLγ works as a polymer with a specific polarity, cutting only after sufficient polymerization has occurred to activate cutting (57). Some subunits act as a scaffold while the actual cut site may be at some distance from the initial DNA binding site; cutting may even occur *in trans* on another DNA template (57). It is possible that incision at some distance from the MutSβ recognition site results in repair synthesis in DNA repeats that leads to repeat expansion. Such differential cutting may explain the differ-

ent outcomes seen between MMR mediated by MutLα and MutLγ on repetitive templates. Furthermore, such an action at a distance by MutLγ may also contribute to determining the size threshold for repeat expansion that is seen in all the DNA repeats (58) if, for example, a cut at a particular distance away from the loop recognition site is required to initiate an expansion event.

Human MLH3 is expressed as two major isoforms differing by only 24 amino acids encoded by MLH3 exon seven. This small exon contains a conserved endonuclease domain. Skipping exon seven via SSOs parallels shRNA knockdown of MLH3 and slows GAA•TTC repeat expansion. Previous studies in yeast have shown that point mutations in this domain abrogate endonuclease function approximating mlh3Δ (21). This is in agreement with our findings that switching MLH3 from predominantly isoform 1 to predominantly isoform 2 evokes a MLH3 knockdown. The existence of two functionally distinct isoforms of MLH3 implies the existence of two distinct species of MutLγ. Manhart *et al.* found that in experiments mixing the nuclease deficient Mlh1-mlh3D523N with wild type Mlh1-Mlh3, the mutant MutLγ subunits can still take part in the polymer and stimulate nicking as long as wild type MutLγ is also incorporated in the polymer (57). We suggest that a mixture of cutting proficient and cutting deficient MutLγ subunits in a polymer could allow the polymer to grow to a larger size before cutting occurs, leading to a cut at a greater distance from the recognition site than might occur with a purely isoform 1 polymer. Whether that is the case, and any possible benefit to meiosis or mismatch repair, awaits future work.

Research in alternate DNA repair pathways and their contribution to repeat expansion has indicated that oxidative damage and repair by the base excision repair pathway (BER), particularly OGG1 (59), is implicated in CAG•CTG expansion in Huntington Disease models. While this may seem at odds with a MutSβ to MutLγ-mediated expansion process, it has been shown that MutSβ can be loaded onto a site after action by base excision repair machinery, linking BER to MMR in repeat expansion (60). Suppression of the oxidative damage leading into this pathway has been shown to reduce the rate of CAG•CTG expansion and delay the onset of pathology in an HD mouse model (61). Budworth et al. inhibited oxidative damage to limit entry into the BER to MMR pathway to slow CAG•CTG expansion (61). We suggest targeting MLH3, and specifically, switching MLH3 into isoform 2 as an additional way to limit repeat expansion in the shared downstream component of this repeat expansion pathway.

MLH3 has the potential to provide a much better therapeutic target than does MLH1. As with MSH2, MLH1 depletion is strongly associated with hereditary nonpolyposis colorectal cancer (HNPCC) and sporadic gastric and endometrial carcinomas (53–55). Currently, there is little evidence to support the loss of MLH3 as a contributing factor in development of cancer (62–64), perhaps in part due to its low abundance relative to the other MutL subunits (10,65). Furthermore, knockdown of MLH3 in our experiments did not affect MutLα levels suggesting MLH3 has a relatively safer potential as a therapeutic target to limit DNA repeat expansion than other MMR components. Currently there is no cure for Friedreich ataxia or any other neurodegen-

erative disease driven by DNA repeat expansion. For progressive repeat expansion diseases such as Friedreich ataxia or Huntington's disease slowed somatic expansion may delay or even prevent the onset of symptoms if treatment is started early. SSO mediated skipping of MLH3 exon seven to slow repeat expansion may provide a basis for the development of novel therapeutic agents, particularly as more data suggests that MLH3 is part of a central mechanism shared by repeat expansion disorders. The successful application of FDA approved splice-switching oligonucleotides as a useful therapeutic in spinal muscular atrophy (SMA) patients (66,67) is an invaluable step forward for using SSO-mediated exon skipping in neurodegenerative disease.

REFERENCES

1. Manley,K., Shirley,T.L., Flaherty,L. and Messer,A. (1999) Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. *Nat. Genet.*, **23**, 471–473.
2. Foiry,L., Dong,L., Savouret,C., Hubert,L., te Riele,H., Junien,C. and Gourdon,G. (2006) Msh3 is a limiting factor in the formation of intergenerational CTG expansions in DM1 transgenic mice. *Hum. Genet.*, **119**, 520–526.
3. Gonitel,R., Moffitt,H., Sathasivam,K., Woodman,B., Detloff,P.J., Faull,R.L. and Bates,G.P. (2008) DNA instability in postmitotic neurons. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 3467–3472.
4. Ezzatizadeh,V., Pinto,R.M., Sandi,C., Sandi,M., Al-Mahdawi,S., Te Riele,H. and Pook,M.A. (2012) The mismatch repair system protects against intergenerational GAA repeat instability in a Friedreich ataxia mouse model. *Neurobiol. Dis.*, **46**, 165–171.
5. Halabi,A., Ditch,S., Wang,J. and Grabczyk,E. (2012) DNA mismatch repair complex MutSbeta promotes GAA•TTC repeat expansion in human cells. *J. Biol. Chem.*, **287**, 29958–29967.
6. Du,J., Campau,E., Soragni,E., Ku,S., Puckett,J.W., Dervan,P.B. and Gottesfeld,J.M. (2012) Role of mismatch repair enzymes in GAA.TTC triplet-repeat expansion in Friedreich ataxia induced pluripotent stem cells. *J. Biol. Chem.*, **287**, 29861–29872.
7. Kolodner,R.D. and Marsischky,G.T. (1999) Eukaryotic DNA mismatch repair. *Curr. Opin. Genet. Dev.*, **9**, 89–96.
8. Kadyrov,F.A., Dzantiev,L., Constantin,N. and Modrich,P. (2006) Endonucleolytic function of MutLalpha in human mismatch repair. *Cell*, **126**, 297–308.
9. Raschle,M., Marra,G., Nystrom-Lahti,M., Schar,P. and Jiricny,J. (1999) Identification of hMutLbeta, a heterodimer of hMLH1 and hPMS1. *J. Biol. Chem.*, **274**, 32368–32375.
10. Cannavo,E., Marra,G., Sabates-Bellver,J., Menigatti,M., Lipkin,S.M., Fischer,F., Cejka,P. and Jiricny,J. (2005) Expression of the MutL homologue hMLH3 in human cells and its role in DNA mismatch repair. *Cancer Res.*, **65**, 10759–10766.

11. Campuzano,V., Montermini,L., Molto,M.D., Pianese,L., Cossee,M., Cavalcanti,F., Monros,E., Rodius,F., Duclos,F., Monticelli,A. *et al.* (1996) Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science*, **271**, 1423–1427.

12. Ditch,S., Sammarco,M.C., Banerjee,A. and Grabczyk,E. (2009) Progressive GAA•TTC repeat expansion in human cell lines. *PLoS Genet.*, **5**, e1000704.

13. Dragileva,E., Hendricks,A., Teed,A., Gillis,T., Lopez,E.T., Friedberg,E.C., Kucherlapati,R., Edelmann,W., Lunetta,K.L., MacDonald,M.E. *et al.* (2009) Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiol. Dis.*, **33**, 37–47.

14. Zhao,X.N., Kumari,D., Gupta,S., Wu,D., Evanitsky,M., Yang,W. and Usdin,K. (2015) Mutsbeta generates both expansions and contractions in a mouse model of the Fragile X-associated disorders. *Hum. Mol. Genet.*, **24**, 7087–7096.

15. Graham,F.L., Smiley,J., Russell,W.C. and Nairn,R. (1977) Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.*, **36**, 59–74.

16. Shaw,G., Morse,S., Ararat,M. and Graham,F.L. (2002) Preferential transformation of human neuronal cells by human adenoviruses and the origin of HEK 293 cells. *FASEB J.*, **16**, 869–871.

17. Grabczyk,E. and Usdin,K. (1999) Generation of microgram quantities of trinucleotide repeat tracts of defined length, interspersion pattern, and orientation. *Analyt. Biochem.*, **267**, 241–243.

18. Banerjee,A., Sammarco,M.C., Ditch,S., Wang,J. and Grabczyk,E. (2009) A novel tandem reporter quantifies RNA polymerase II termination in mammalian cells. *PLoS One*, **4**, e6193.

19. Li,Y., Polak,U., Clark,A.D., Bhalla,A.D., Chen,Y.Y., Li,J., Farmer,J., Seyer,L., Lynch,D., Butler,J.S. *et al.* (2016) Establishment and maintenance of primary fibroblast repositories for rare diseases—Friedreich's ataxia example. *Biopreserv. Biobank*, **14**, 324–329.

20. Marino,M.P., Luce,M.J. and Reiser,J. (2003) Small- to large-scale production of lentivirus vectors. *Methods Mol. Biol.*, **229**, 43–55.

21. Nishant,K.T., Plys,A.J. and Alani,E. (2008) A mutation in the putative MLH3 endonuclease domain confers a defect in both mismatch repair and meiosis in Saccharomyces cerevisiae. *Genetics*, **179**, 747–755.

22. Summerton,J. and Weller,D. (1997) Morpholino antisense oligomers: design, preparation, and properties. *Antisense Nucleic Acid Drug Dev.*, **7**, 187–195.

23. Morcos,P.A., Li,Y. and Jiang,S. (2008) Vivo-Morpholinos: a non-peptide transporter delivers Morpholinos into a wide array of mouse tissues. *BioTechniques*, **45**, 613–614.

24. Calmels,N., Seznec,H., Villa,P., Reutenauer,L., Hibert,M., Haiech,J., Rustin,P., Koenig,M. and Puccio,H. (2009) Limitations in a frataxin knockdown cell model for Friedreich ataxia in a high-throughput drug screen. *BMC Neurol.*, **9**, 46.

25. De Biase,I., Rasmussen,A., Endres,D., Al-Mahdawi,S., Monticelli,A., Cocozza,S., Pook,M. and Bidichandani,S.I. (2007) Progressive GAA expansions in dorsal root ganglia of Friedreich's ataxia patients. *Ann. Neurol.*, **61**, 55–60.

26. De Biase,I., Rasmussen,A., Monticelli,A., Al-Mahdawi,S., Pook,M., Cocozza,S. and Bidichandani,S.I. (2007) Somatic instability of the expanded GAA triplet-repeat sequence in Friedreich ataxia progresses throughout life. *Genomics*, **90**, 1–5.

27. Shelbourne,P.F., Keller-McGandy,C., Bi,W.L., Yoon,S.R., Dubeau,L., Veitch,N.J., Vonsattel,J.P., Wexler,N.S., Group,U.S.-V.C.R., Arnheim,N. *et al.* (2007) Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. *Hum. Mol. Genet.*, **16**, 1133–1142.

28. Grabczyk,E., Mancuso,M. and Sammarco,M.C. (2007) A persistent RNA•DNA hybrid formed by transcription of the Friedreich ataxia triplet repeat in live bacteria, and by T7 RNAP in vitro. *Nucleic Acids Res.*, **35**, 5351–5359.

29. Grabczyk,E. and Fishman,M.C. (1995) A long purine-pyrimidine homopolymer acts as a transcriptional diode. *J. Biol. Chem.*, **270**, 1791–1797.

30. Reddy,K., Tam,M., Bowater,R.P., Barber,M., Tomlinson,M., Nichol Edamura,K., Wang,Y.H. and Pearson,C.E. (2011) Determinants of R-loop formation at convergent bidirectionally transcribed trinucleotide repeats. *Nucleic Acids Res.*, **39**, 1749–1762.

31. Loomis,E.W., Sanz,L.A., Chedin,F. and Hagerman,P.J. (2014) Transcription-associated R-loop formation across the human FMR1 CGG-repeat region. *PLoS Genet.*, **10**, e1004294.

32. Groh,M., Lufino,M.M., Wade-Martins,R. and Gromak,N. (2014) R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. *PLoS Genet.*, **10**, e1004318.

33. Coggins,L.W. and O'Prey,M. (1989) DNA tertiary structures formed in vitro by misaligned hybridization of multiple tandem repeat sequences. *Nucleic Acids Res.*, **17**, 7417–7426.

34. Pearson,C.E., Wang,Y.H., Griffith,J.D. and Sinden,R.R. (1998) Structural analysis of slipped-strand DNA (S-DNA) formed in (CTG)n. (CAG)n repeats from the myotonic dystrophy locus. *Nucleic Acids Res.*, **26**, 816–823.

35. Owen,B.A., Yang,Z., Lai,M., Gajec,M., Badger,J.D. 2nd, Hayes,J.J., Edelmann,W., Kucherlapati,R., Wilson,T.M. and McMurray,C.T. (2005) (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. *Nat. Struct. Mol. Biol.*, **12**, 663–670.

36. Palombo,F., Iaccarino,I., Nakajima,E., Ikejima,M., Shimada,T. and Jiricny,J. (1996) hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Curr. Biol.: CB*, **6**, 1181–1184.

37. Dowen,J.M., Putnam,C.D. and Kolodner,R.D. (2010) Functional studies and homology modeling of Msh2-Msh3 predict that mispair recognition involves DNA bending and strand separation. *Mol. Cell. Biol.*, **30**, 3321–3328.

38. Gupta,S., Gellert,M. and Yang,W. (2012) Mechanism of mismatch recognition revealed by human MutSbeta bound to unpaired DNA loops. *Nat. Struct. Mol. Biol.*, **19**, 72–78.

39. Ku,S., Soragni,E., Campau,E., Thomas,E.A., Altun,G., Laurent,L.C., Loring,J.F., Napierala,M. and Gottesfeld,J.M. (2010) Friedreich's ataxia induced pluripotent stem cells model intergenerational GAATTC triplet repeat instability. *Cell Stem Cell*, **7**, 631–637.

40. Kovtun,I.V. and McMurray,C.T. (2001) Trinucleotide expansion in haploid germ cells by gap repair. *Nat. Genet.*, **27**, 407–411.

41. van den Broek,W.J., Nelen,M.R., Wansink,D.G., Coerwinkel,M.M., te Riele,H., Groenen,P.J. and Wieringa,B. (2002) Somatic expansion behaviour of the (CTG)n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. *Hum. Mol. Genet.*, **11**, 191–198.

42. Kovalenko,M., Dragileva,E., St Claire,J., Gillis,T., Guide,J.R., New,J., Dong,H., Kucherlapati,R., Kucherlapati,M.H., Ehrlich,M.E. *et al.* (2012) Msh2 acts in medium-spiny striatal neurons as an enhancer of CAG instability and mutant huntingtin phenotypes in Huntington's disease knock-in mice. *PLoS One*, **7**, e44273.

43. Tome,S., Simard,J.P., Slean,M.M., Holt,I., Morris,G.E., Wojciechowicz,K., te Riele,H. and Pearson,C.E. (2013) Tissue-specific mismatch repair protein expression: MSH3 is higher than MSH6 in multiple mouse tissues. *DNA Repair*, **12**, 46–52.

44. Lokanga,R.A., Zhao,X.N. and Usdin,K. (2014) The mismatch repair protein MSH2 is rate limiting for repeat expansion in a fragile X premutation mouse model. *Hum. Mut.*, **35**, 129–136.

45. Ranjha,L., Anand,R. and Cejka,P. (2014) The Saccharomyces cerevisiae Mlh1-Mlh3 heterodimer is an endonuclease that preferentially binds to Holliday junctions. *J. Biol. Chem.*, **289**, 5674–5686.

46. Rogacheva,M.V., Manhart,C.M., Chen,C., Guarne,A., Surtees,J. and Alani,E. (2014) Mlh1-Mlh3, a meiotic crossover and DNA mismatch repair factor, is a Msh2-Msh3-stimulated endonuclease. *J. Biol. Chem.*, **289**, 5664–5673.

47. Kondo,E., Horii,A. and Fukushige,S. (2001) The interacting domains of three MutL heterodimers in man: hMLH1 interacts with 36 homologous amino acid residues within hMLH3, hPMS1 and hPMS2. *Nucleic Acids Res.*, **29**, 1695–1702.

48. Flores-Rozas,H. and Kolodner,R.D. (1998) The Saccharomyces cerevisiae MLH3 gene functions in MSH3-dependent suppression of frameshift mutations. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 12404–12409.

49. Charbonneau,N., Amunugama,R., Schmutte,C., Yoder,K. and Fishel,R. (2009) Evidence that hMLH3 functions primarily in meiosis and in hMSH2-hMSH3 mismatch repair. *Cancer Biol. Ther.*, **8**, 1411–1420.

50. Pinto,R.M., Dragileva,E., Kirby,A., Lloret,A., Lopez,E., St Claire,J., Panigrahi,G.B., Hou,C., Holloway,K., Gillis,T. *et al.* (2013) Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. *PLoS Genet.*, **9**, e1003930.

51. Bourn,R.L., De Biase,I., Pinto,R.M., Sandi,C., Al-Mahdawi,S., Pook,M.A. and Bidichandani,S.I. (2012) Pms2 suppresses large expansions of the (GAA.TTC)n sequence in neuronal tissues. *PLoS One*, **7**, e47085.

52. Gomes-Pereira,M., Fortune,M.T., Ingram,L., McAbney,J.P. and Monckton,D.G. (2004) Pms2 is a genetic enhancer of trinucleotide CAG.CTG repeat somatic mosaicism: implications for the mechanism of triplet repeat expansion. *Hum. Mol. Genet.*, **13**, 1815–1825.

53. Bronner,C.E., Baker,S.M., Morrison,P.T., Warren,G., Smith,L.G., Lescoe,M.K., Kane,M., Earabino,C., Lipford,J., Lindblom,A. *et al.* (1994) Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature*, **368**, 258–261.

54. Risinger,J.I., Berchuck,A., Kohler,M.F., Watson,P., Lynch,H.T. and Boyd,J. (1993) Genetic instability of microsatellites in endometrial carcinoma. *Cancer Res.*, **53**, 5100–5103.

55. Nystrom-Lahti,M., Parsons,R., Sistonen,P., Pylkkanen,L., Aaltonen,L.A., Leach,F.S., Hamilton,S.R., Watson,P., Bronson,E., Fusaro,R. *et al.* (1994) Mismatch repair genes on chromosomes 2p and 3p account for a major share of hereditary nonpolyposis colorectal cancer families evaluable by linkage. *Am. J. Hum. Genet.*, **55**, 659–665.

56. Kadyrov,F.A., Holmes,S.F., Arana,M.E., Lukianova,O.A., O'Donnell,M., Kunkel,T.A. and Modrich,P. (2007) Saccharomyces cerevisiae MutLalpha is a mismatch repair endonuclease. *J. Biol. Chem.*, **282**, 37181–37190.

57. Manhart,C.M., Ni,X., White,M.A., Ortega,J., Surtees,J.A. and Alani,E. (2017) The mismatch repair and meiotic recombination endonuclease Mlh1-Mlh3 is activated by polymer formation and can cleave DNA substrates in trans. *PLoS Biol.*, **15**, e2001164.

58. Lee,D.Y. and McMurray,C.T. (2014) Trinucleotide expansion in disease: why is there a length threshold? *Curr. Opin. Genet. Dev.*, **26**, 131–140.

59. Kovtun,I.V., Liu,Y., Bjoras,M., Klungland,A., Wilson,S.H. and McMurray,C.T. (2007) OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature*, **447**, 447–452.

60. Lai,Y., Budworth,H., Beaver,J.M., Chan,N.L., Zhang,Z., McMurray,C.T. and Liu,Y. (2016) Crosstalk between MSH2-MSH3 and polbeta promotes trinucleotide repeat expansion during base excision repair. *Nat. Commun.*, **7**, 12465.

61. Budworth,H., Harris,F.R., Williams,P., Lee,D.Y., Holt,A., Pahnke,J., Szczesny,B., Acevedo-Torres,K., Ayala-Pena,S. and McMurray,C.T. (2015) Suppression of somatic expansion delays the onset of pathophysiology in a mouse model of Huntington's disease. *PLoS Genet.*, **11**, e1005267.

62. Wu,Y., Berends,M.J., Sijmons,R.H., Mensink,R.G., Verlind,E., Kooi,K.A., van der Sluis,T., Kempinga,C., van dDer Zee,A.G., Hollema,H. *et al.* (2001) A role for MLH3 in hereditary nonpolyposis colorectal cancer. *Nat. Genet.*, **29**, 137–138.

63. Hienonen,T., Laiho,P., Salovaara,R., Mecklin,J.P., Jarvinen,H., Sistonen,P., Peltomaki,P., Lehtonen,R., Nupponen,N.N., Launonen,V. *et al.* (2003) Little evidence for involvement of MLH3 in colorectal cancer predisposition. *Int. J. Cancer*, **106**, 292–296.

64. Liu,H.X., Zhou,X.L., Liu,T., Werelius,B., Lindmark,G., Dahl,N. and Lindblom,A. (2003) The role of hMLH3 in familial colorectal cancer. *Cancer Res.*, **63**, 1894–1899.

65. Korhonen,M.K., Raevaara,T.E., Lohi,H. and Nystrom,M. (2007) Conditional nuclear localization of hMLH3 suggests a minor activity in mismatch repair and supports its role as a low-risk gene in HNPCC. *Oncol Rep.*, **17**, 351–354.

66. Finkel,R.S., Chiriboga,C.A., Vajsar,J., Day,J.W., Montes,J., De Vivo,D.C., Yamashita,M., Rigo,F., Hung,G., Schneider,E. *et al.* (2016) Treatment of infantile-onset spinal muscular atrophy with nusinersen: a phase 2, open-label, dose-escalation study. *Lancet*, **388**, 3017–3026.

67. Meijboom,K.E., Wood,M.J.A. and McClorey,G. (2017) Splice-switching therapy for spinal muscular atrophy. *Genes (Basel)*, **8**, E161.

68. Kent,W.J., Sugnet,C.W., Furey,T.S., Roskin,K.M., Pringle,T.H., Zahler,A.M. and Haussler,D. (2002) The human genome browser at UCSC. *Genome Res.*, **12**, 996–1006.

69. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.