Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2012/07/13/M112.391961.DC1.html

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 287, NO. 35, pp. 29861–29872, August 24, 2012
© 2012 by The American Society for Biochemistry and Molecular Biology, Inc. Published in the U.S.A.

Role of Mismatch Repair Enzymes in GAA·TTC Triplet-repeat Expansion in Friedreich Ataxia Induced Pluripotent Stem Cells*[S]

Received for publication, June 14, 2012, and in revised form, July 12, 2012 Published, JBC Papers in Press, July 13, 2012, DOI 10.1074/jbc.M112.391961

Jintang Du[‡], Erica Campau[‡], Elisabetta Soragni[‡], Sherman Ku[‡], James W. Puckett[§], Peter B. Dervan[§], and Joel M. Gottesfeld[‡1]

From the [‡]Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037 and the [§]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Background: Friedreich ataxia is caused by a GAA·TTC triplet-repeat expansion in the first intron of the *FXN* gene.
Results: Expansion of the repeats is observed in induced pluripotent stem cells (iPSCs) and can be blocked with either shRNAs to mismatch repair enzymes or small molecules targeting the repeats.
Conclusion: MutSα and MutSβ are involved in repeat expansion.
Significance: iPSCs provide a model system for studying triplet-repeat expansion.

The genetic mutation in Friedreich ataxia (FRDA) is a hyperexpansion of the triplet-repeat sequence GAA·TTC within the first intron of the *FXN* gene. Although yeast and reporter construct models for GAA·TTC triplet-repeat expansion have been reported, studies on FRDA pathogenesis and therapeutic development are limited by the availability of an appropriate cell model in which to study the mechanism of instability of the GAA·TTC triplet repeats in the human genome. Herein, induced pluripotent stem cells (iPSCs) were generated from FRDA patient fibroblasts after transduction with the four transcription factors Oct4, Sox2, Klf4, and c-Myc. These cells were differentiated into neurospheres and neuronal precursors *in vitro*, providing a valuable cell model for FRDA. During propagation of the iPSCs, GAA·TTC triplet repeats expanded at a rate of about two GAA·TTC triplet repeats/replication. However, GAA·TTC triplet repeats were stable in FRDA fibroblasts and neuronal stem cells. The mismatch repair enzymes MSH2, MSH3, and MSH6, implicated in repeat instability in other triplet-repeat diseases, were highly expressed in pluripotent stem cells compared with fibroblasts and neuronal stem cells and occupied *FXN* intron 1. In addition, shRNA silencing of *MSH2* and *MSH6* impeded GAA·TTC triplet-repeat expansion. A specific pyrrole-imidazole polyamide targeting GAA·TTC triplet-repeat DNA partially blocked repeat expansion by displacing MSH2 from *FXN* intron 1 in FRDA iPSCs. These studies suggest that in FRDA, GAA·TTC triplet-repeat instability occurs in embryonic cells and involves the highly active mismatch repair system.

Nearly 30 hereditary diseases in humans are caused by expansion of a triplet-repeat sequence in genomic DNA (1, 2). These expanded sequences are unstable and frequently change in length during intergenerational transmission and within somatic cells. However, the mechanisms of expanded triplet-repeat sequence generation remain enigmatic (3).

Friedreich ataxia (FRDA)[2] is caused by heterochromatin-mediated silencing of the *FXN* gene, encoding the essential mitochondrial protein frataxin (4). The genetic mutation in FRDA is a GAA·TTC triplet-repeat expansion in the first intron of *FXN*, with unaffected alleles having 6–34 repeats in contrast to 66–1700 repeats in patient alleles. Longer repeats are associated with more severe gene repression, lower frataxin protein levels, and earlier onset and increased disease severity (5, 6). Since the GAA·TTC triplet-repeat expansion mutation was identified (5), several models have been established to explain the mechanism of repeat expansion. Using *Escherichia coli* (7) and yeast (8, 9) models, GAA·TTC triplet-repeat expansion has been observed during cell culture. Mammalian cell models using a plasmid construct with GAA·TTC triplet repeats have shown that the repeat expansions are transcription-, replication-, and position-dependent (10–12). In addition, mouse models with expanded GAA·TTC triplet repeats show somatic instability in different tissues (13–15). These systems differ substantially from humans and from each other in terms of DNA replication rate, the cell type in which GAA·TTC triplet-repeat expansion occurs, and chromatin structure. Therefore, studies on FRDA pathogenesis and therapeutic development are still limited by the availability of an appropriate cell model in which to study the mechanism of GAA·TTC triplet-repeat generation in the human genome.

In our previous work (16), we showed that GAA·TTC triplet repeats are highly unstable in FRDA induced pluripotent stem cells (iPSCs). Liu *et al.* (17) also reported that GAA·TTC triplet repeats are unstable during the generation of iPSCs from

* This work was supported, in whole or in part, by National Institutes Health Grant NS062856 from NINDS (to J. M. G.) and Grant GM275681 from NIGMS (to P. B. D.). This work was also supported by a postdoctoral fellowship from the Friedreich's Ataxia Research Alliance (to J. D.) and by a grant from the California Institute for Regenerative Medicine (CIRM).
[S] This article contains supplemental Figs. S1–S3.
[1] To whom correspondence should be addressed: Dept. of Molecular Biology, The Scripps Research Institute, MB-27, 10550 N. Torrey Pines Rd., La Jolla, CA 92037. Tel.: 858-784-8914; Fax: 858-784-8965; E-mail: joelg@scripps.edu.

[2] The abbreviations used are: FRDA, Friedreich ataxia; iPSC, induced pluripotent stem cell; MMR, mismatch repair; ESC, embryonic stem cell; SP-PCR, small-pool PCR; qRT-PCR, quantitative RT-PCR.

Downloaded from www.jbc.org at CALIFORNIA INSTITUTE OF TECHNOLOGY, on September 27, 2012

patient fibroblasts. Importantly, asymptomatic heterozygous carriers show GAA·TTC triplet-repeat expansion on the pathogenic allele, but not the normal allele. The mismatch repair (MMR) enzyme MSH2, implicated in repeat instability in other triplet-repeat diseases, is highly expressed in pluripotent stem cells and occupies *FXN* intron 1. In addition, shRNA silencing of *MSH2* impedes GAA·TTC triplet-repeat expansion (16). Our findings provide the first human cell model to study the mechanisms of GAA·TTC triplet-repeat expansion in the context of the endogenous cellular *FXN* gene in human cells. Herein, we found that GAA·TTC triplet-repeat expansion occurred only in the iPSCs, but not in FRDA fibroblasts or differentiated neurospheres. Repeat expansion involved the MMR complexes MutSα and MutSβ. These studies suggest that the generation and expansion of GAA·TTC triplet repeats in FRDA most likely occur in embryonic cells and involve the highly active MMR system.

EXPERIMENTAL PROCEDURES

Cell Culture—Fibroblasts, embryonic stem cells (ESCs)/iPSCs, neurospheres, neurons, and HEK293T cells were grown at 37 °C and 5% CO_2. Fibroblasts were cultured with 10% FBS in minimal essential medium, 2 mM glutamine, 1% nonessential amino acids, 20 mM HEPES, and 1% antibiotic/antimycotic (Invitrogen). ESCs/iPSCs were grown on γ-irradiated mouse embryonic fibroblasts (GlobalStem, Rockville, MD) in DMEM/nutrient mixture F-12 with 20% KnockOut serum replacement, 1 mM glutamine, 1% nonessential amino acids, 1% antibiotic/antimycotic, 0.1 mM β-mercaptoethanol (Invitrogen), and 20 ng/ml basic FGF (Stemgent, San Diego, CA) and passaged manually every 7 or 8 days. Neurospheres were grown in Neurobasal-A medium with 2% B-27 supplement, 1% ITS-A supplement, 1% N-2 supplement, 2 mM glutamine, 1% antibiotic/antimycotic, 10 mM HEPES, 20 ng/ml basic FGF, and 20 ng/ml EGF (R&D Systems). Neuronal cells were grown on laminin in Neurobasal-A medium with 2% B-27 supplement, 1% ITS-A supplement, 1% N-2 supplement, 2 mM glutamine, 1% antibiotic/antimycotic, and 10 mM HEPES. HEK293T cells were grown with 10% FBS in DMEM, 2 mM glutamine, 20 mM HEPES, 1% nonessential amino acids, and 1% antibiotic/antimycotic.

Derivation of iPSCs—Unaffected and FRDA iPSC derivation followed previous methods with minor deviations (18). These normal and FRDA iPSCs have been characterized by standard methods (16).

Derivation of Neurospheres and Neural Differentiation—In vitro differentiation of FRDA iPSCs to neurospheres and neurons was performed as described previously (19–22). Briefly, iPSC colonies were passaged onto high-density γ-irradiated mouse embryonic fibroblasts (at 60×10^3 cells/cm^2) with ROCK inhibitor (10 μM; Stemgent). Induction was performed with either noggin (0.5 μg/ml; PeproTech Inc., Rocky Hill, NJ) or dorsomorphin (5 μM; P5499, Sigma-Aldrich) and SB431542 (10 μM; 040010, Stemgent) in ESC medium. During induction, FRDA iPSCs were maintained for 2 weeks without passage. Subsequently, induced colonies were dissected and transferred to suspension culture as neurospheres in Neurobasal-A medium with EGF (20 ng/ml) and basic FGF (20 ng/ml). Neurospheres were maintained as a suspension culture and passaged manually every 4–5 days. Neural differentiation was performed by dissociating neurospheres with Accutase (Invitrogen) and replating onto polylysine/laminin-coated plates at 50×10^3 cells/cm^2. Neurons were maintained in Neurobasal-A medium without EGF or FGF for 7 or 8 days after replating.

Flow Cytometry and Immunocytochemistry—Neurospheres and neurons were dissociated with Accutase. Single cells were fixed using a BD Cytofix/Cytoperm fixation/permeabilization kit (BD Biosciences) following the manufacturer's protocol at 4 °C for 20 min. Primary antibodies were incubated for 30 min at room temperature: anti-nestin (ab5968, Abcam, Cambridge, MA) at 1:5000 dilution for neurospheres, anti-Tuj1 (MMS-435P, Covance, Emeryville, CA) at 1:2500 dilution for neuronal cells, and Alexa Fluor 488 anti-MAP2B (560399, BD Biosciences) at 1:20 dilution for neuronal cells. Secondary antibodies (Alexa Fluor 488 goat anti-rabbit IgG (H + L; A11034, Invitrogen) and Alexa Fluor 488 goat anti-mouse IgG (H + L; A11001, Invitrogen)) were incubated at 1:1000 dilution for 30 min at room temperature. Cells were analyzed with a FACSCalibur HTS instrument (BD Biosciences).

For immunocytochemistry, neuronal cells were fixed in 4% paraformaldehyde for 10 min and permeabilized with 0.5% Triton X-100 in PBS for 10 min at room temperature. Anti-MAP2 (AB5622, Millipore) and anti-Tuj1 primary antibodies in 10% goat serum in 0.5% Triton X-100 in PBS were incubated overnight at 4 °C, and Alexa Fluor 488 goat anti-rabbit IgG (H + L), Alexa Fluor 488 goat anti-mouse IgG (H + L), and Alexa Fluor 594 goat anti-rabbit IgG (H + L; A11012, Invitrogen) secondary antibodies were incubated at 1:1000 dilution for 1 h at room temperature, followed by nuclear staining with DAPI.

Nucleic Acid Purification—Total RNA was purified with an RNeasy Plus minikit (Qiagen) according to the manufacturer. Genomic DNA was purified by isopropyl alcohol precipitation of cell lysate prepared in total cell lysis buffer (100 mM Tris-HCl (pH 8.5), 5 mM EDTA, 0.2% SDS, 0.2 M NaCl, and 200 μg/ml proteinase K (Roche Applied Science)) (23). Generally, cells were collected and washed once with PBS. Cell pellets were incubated in 0.2 ml of lysis buffer overnight at room temperature. To the lysate was added 1.0 μl of 20 mg/ml glycogen (Roche Applied Science) and 0.2 ml of isopropyl alcohol. After incubation for 2 h, the precipitated genomic DNA was collected by centrifugation and washed twice with 0.5 ml of 70% isopropyl alcohol in TE buffer (10 mM Tris-HCl (pH 7.5) and 1 mM EDTA). The genomic DNA pellet was heated at 55 °C for 10 min and then resuspended in 40 μl of TE buffer overnight at 55 °C.

Conventional PCR, Small-pool PCR (SP-PCR) , and Quantitative RT-PCR (qRT-PCR)—For GAA·TTC triplet-repeat length conventional PCRs, Phusion polymerase (New England Biolabs, Ipswich, MA) was used according to the manufacturer. 20 ng of genomic DNA (~6600 genomic equivalents) and 0.1 μM primers GAA-104F and GAA-629R were used (12) in 20-μl reactions cycled through the following conditions: denaturation at 98 °C for 5 s, annealing at 70 °C for 15 s, and extension at 72 °C for 90 s for 40 cycles with a 5-min initial denaturation and a 5-min final extension. Quantitation of PCR band size was performed using an inverse power function directly correlating gel migration of a molecular weight ladder to its known sizes

Downloaded from www.jbc.org at CALIFORNIA INSTITUTE OF TECHNOLOGY, on September 27, 2012

(ImageJ software) (24). PCR products from the *FXN* locus contain 499 bp of non-repeat sequences, so GAA·TTC triplet-repeat number estimations were adjusted accordingly.

SP-PCR was performed using a modification of a previously published protocol (25). Serial dilutions of genomic DNA ranging from 75 to 750 pg (25–250 genomic equivalents) were prepared in microcentrifuge tubes with 2 μM primers GAA-104F and GAA-629R. Other PCR conditions were according to previous conventional PCR with 35 cycles. PCR products were resolved by electrophoresis on 0.8% agarose gels. The products were analyzed by Southern blotting using an end-labeled $(TTC)_9$ oligonucleotide probe that specifically hybridizes to the GAA triplet-repeat sequence.

qRT-PCR analysis was done with a qScript One-Step SYBR Green qRT-PCR kit (Bio-Rad) according to the manufacturer. Primers for pluripotent markers were as described (18). *MSH2*, *MSH3*, and *GAPDH* primers were as described (16). Other primers used for qRT-PCR were as follows: *MSH6*_RT02F (5′-AGGAGCAGGAAAATGGCAA-3′), *MSH6*_RT07R (5′-CATCTGGGCCATTACAGCTA-3′), *MAP2*_F2 (5′-TTTGGAGAGCATGGGTCAC-3′), *MAP2*_R2 (5′-TGAACTATCCTTGCAGACACC-3′), *nestin*_F2 (5′-AAGACTTCCCTCAGCTTTCAG-3′), and *nestin*_R2 (5′-AGCAAAGATCCAAGACGCC-3′). Analysis of relative qRT-PCR data was performed via the $\Delta\Delta C_t$ method (26). All qRT-PCR signals were normalized to *GAPDH*. Reverse transcription was performed at 50 °C for 20 min, followed by PCR cycling at 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s for 40 cycles.

Western Analysis—Whole cell extracts in 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 10% glycerol, 0.5% Triton X-100, and protease inhibitor (Roche Applied Science) were electrophoresed on polyacrylamide gels and transferred onto nitrocellulose membranes. Primary antibodies were incubated overnight, and secondary antibodies were incubated at room temperature for 1 h. Signals were detected using an ODYSSEY CLx infrared imaging system (LI-COR Bioscience, Lincoln, NE). Blocking was performed in 5% milk in TBS (137 mM NaCl, 2.7 mM KCl, and 25 mM Tris-HCl (pH 7.4)) at room temperature for 1 h. Antibody incubation was performed in 3% BSA in TBS containing 0.1% Tween 20. Anti-MSH2 (ab52266) and anti-MSH6 (ab50604) antibodies were obtained from Abcam and used at 1:4000 and 1:1000 dilutions, respectively. Anti-MSH3 (sc-11441) and anti-actin (sc-8432) antibodies were obtained from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA) and used at 1:200 and 1:400 dilutions, respectively. IRDye 680LT-conjugated goat anti-mouse IgG (H + L; 926-68020) and IRDye 800CW-conjugated goat anti-rabbit IgG (H + L; 926-32211) secondary antibodies were obtained from LI-COR Bioscience and used at a dilution of 1:5000.

Lentiviral shRNA Transduction—Lentiviral shRNA for targeting *MSH2* has been described (16). Lentiviral shRNA constructs for targeting *MSH6* were generated by cloning an MSH6-targeted oligonucleotide (5′-CCGGGCCAGAAGAATACGAGTTGAACTCGAGTTCAACTCGTATTCTTCTGGCTTTTTG-3′) into the vector pLKO.1-puro (Sigma-Aldrich) at *AgeI* and *EcoRI* sites downstream of the U6 promoter. A vector containing a scrambled shRNA (plasmid 1864) was obtained from Addgene (Cambridge, MA). Lentiviral particles were then packaged by cotransfecting human HEK293T cells using FuGENE 6 reagent (Roche Applied Science) with shRNA constructs along with psPAX2 (plasmid 12260) and pMD2.G (plasmid 12259) helper plasmids (Addgene) in HEK293T medium. Transfected cells were grown at 37 °C for a total of 12 h, and the cell medium was changed to human ESC medium. Lentivirus-containing supernatants were collected 24 and 48 h later. Lentivirus-containing supernatant was filtered through a 0.45-μm cellulose acetate mesh to remove any contaminating cells or cell debris. FRDA iPSCs were subjected to two lentiviral transductions with lentivirus for 12 h at 37 °C with 5 μg/ml Polybrene. Cells were then expanded and subjected to 6 days of puromycin selection (0.4 μg/ml) on drug-resistant DR4 mouse embryonic fibroblasts (GlobalStem).

Pyrrole-Imidazole Polyamide Treatment—FRDA iPSCs were incubated with polyamide FA1 (5 μM) (27) or HIV-1 (5 μM) (28, 29) or in the absence of polyamide for 6 weeks, and genomic DNA was purified after each passage. For chromatin immunoprecipitation, FRDA iPSCs were incubated with polyamide FA1 (5 μM) or in the absence of polyamide for 72 h before iPSCs were collected. Polyamides were synthesized as described previously (27, 28) and were verified for purity and identity by analytical HPLC and MALDI-TOF-MS. Characterization of the DNA-binding properties of the molecules by quantitative DNase I footprinting has been described (27, 28). A BODIPY FL-polyamide conjugate was also used for nuclear uptake studies (27).

Chromatin Immunoprecipitation—Cells were cross-linked first with 1.5 mM dithiobis(succinimidyl propionate), followed by 1% formaldehyde. Subsequent ChIP procedures were as described (4, 16) with anti-MSH2 antibody (sc-494, Santa Cruz Biotechnology, Inc.). Analysis by qPCR with primers for the *FXN* promoter, the region upstream of the GAA·TTC triplet repeats, and the region downstream of the repeats was as described (4, 16).

RESULTS

Generation of Expanded GAA·TTC Triplet Repeats in FRDA iPSCs—It is known that expanded GAA·TTC triplet repeats in FRDA-associated alleles are genetically highly unstable and undergo changes in repeat length in both the germ line and soma (30). Our previous work has shown that GAA·TTC triplet repeats in iPSCs in *FXN* exhibit repeat instability similar to patient families, where the repeats expand with discrete changes in length between generations (16). Liu *et al.* (17) also reported that reprogramming of FRDA fibroblasts to iPSCs results in both expansions and contractions of the GAA·TTC triplet repeats within the *FXN* locus. In these studies, conventional PCR (typically using >20 ng of genomic DNA, which represents 6600 genomic equivalents) was used to assess GAA·TTC triplet-repeat length variability in the FRDA fibroblasts and iPSCs. This method cannot detect rarer molecules comprising <10% of the total population.

It has been reported that SP-PCR can detect these rarer mutant molecules. This method constitutes PCR amplification of the GAA·TTC repeat sequence in multiple small pools of input genomic DNA, containing on the order of 0.5–200 genomic equivalents (31). To differentiate whether the GAA·TTC triplet-repeat instability observed in FRDA iPSCs is

Downloaded from www.jbc.org at CALIFORNIA INSTITUTE OF TECHNOLOGY, on September 27, 2012

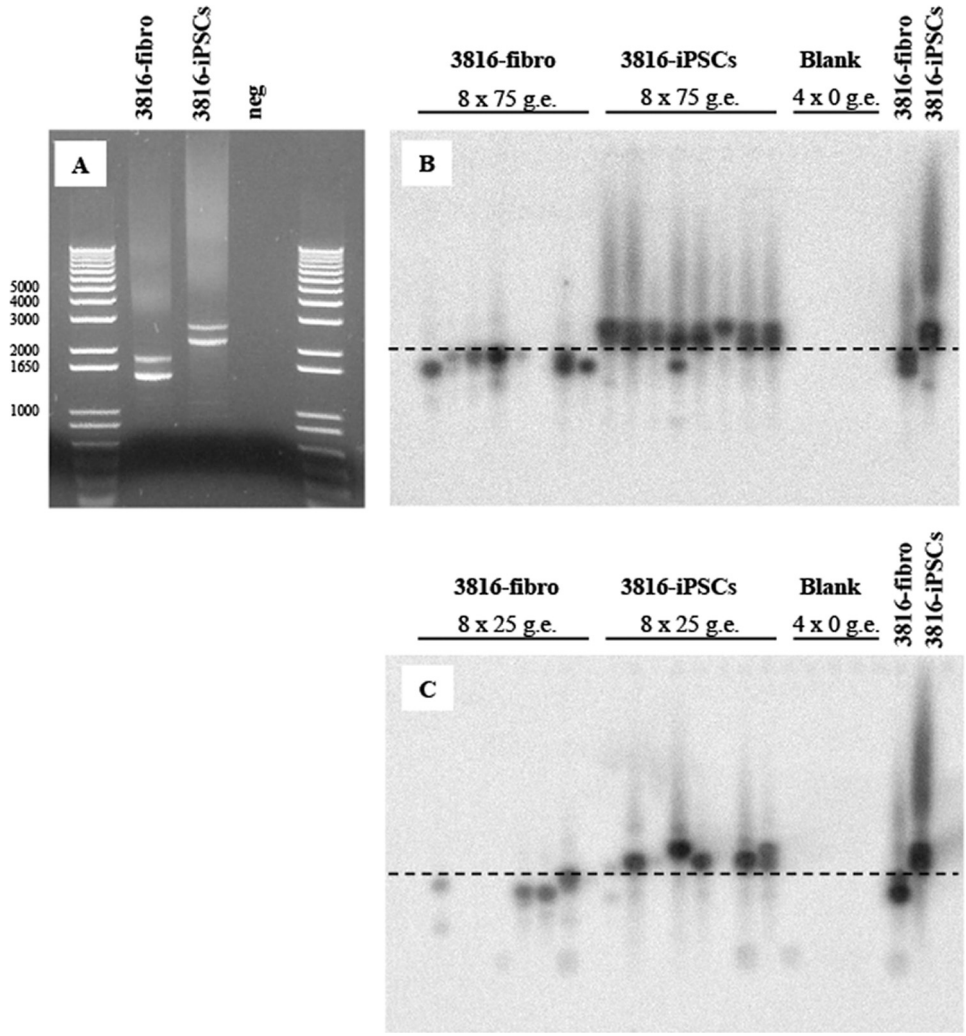
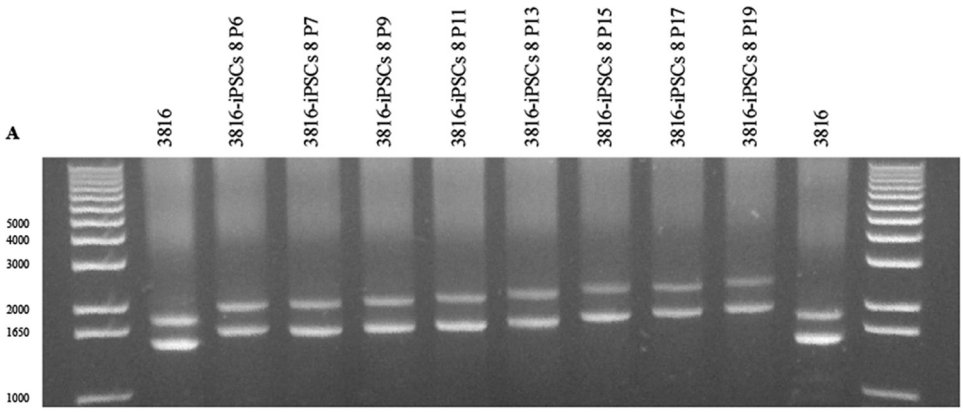


FIGURE 1. **Conventional PCR and SP-PCR of GAA·TTC triplet repeats from FRDA fibroblasts and iPSCs.** *A*, conventional PCR amplification (~6600 genomic equivalent molecules/reaction) showing the GAA·TTC triplet repeats from FRDA fibroblasts and iPSCs (passage 22). *neg*, negative. *B* and *C*, SP-PCR analysis of the GAA·TTC triplet repeats from FRDA fibroblasts and iPSCs (passage 22) using 75 and 25 genomic equivalent (*g.e.*) molecules/reaction in *B* and *C*, respectively. *3816-fibro* and *3816-iPSCs* (passage 22) are controls for SP-PCR using 250 genomic equivalent molecules/reaction. The *dashed lines* are set between the PCR bands from fibroblasts and iPSCs.

similar to somatic or intergenerational instability, SP-PCR was used to detect rarer molecules with discrete lengths of GAA·TTC triplet repeats from FRDA fibroblasts (line GM03816 from the NIGMS Coriell Depository, homozygous for expanded GAA·TTC triplet alleles) and iPSC line 23-3 (passage 22, derived from GM03816 cells) (16). Conventional PCR (20 ng, ~6600 genomic equivalent molecules/reaction) showed that fibroblasts and iPSCs carried a distinguished spectrum of expanded alleles (Fig. 1*A*). The expanded GAA·TTC triplet alleles, seen as double bands by conventional PCR, were thus resolved into similar bands from the respective constitutional GAA·TTC allele sizes (Fig. 1*B*) in SP-PCR with 75 genomic equivalent molecules/reaction. Lower DNA template concentrations (25 genomic equivalent molecules/reaction) frequently showed single-allele discrimination, in which only one of two homologous expanded alleles was amplified (Fig. 1*C*). We believe that these bands reflect the allele length in individual FRDA genes (haploid genomes). The repeat instability we observed in FRDA iPSCs resembles the intergenerational repeat expansions seen in patient families rather than the repeat instability seen in the somatic cells that are affected in the disease (6). In this latter case, heterogeneous expansions and contractions of the repeats are observed, rather than the discrete changes in repeat length observed on maternal transmission of pathogenic alleles. Our findings are therefore more similar to this intergenerational repeat instability; however, this conclusion must be tempered by the fact that derivation of iPSCs from fibroblasts does not involve meiosis, as does passage through the maternal germ line.

GAA·TTC Triplet-repeat Expansion Rate in FRDA iPSCs—Upon propagation of the FRDA iPSCs, we found that the GAA·TTC triplet repeats continued to expand (Fig. 2*A*). This makes FRDA iPSCs an ideal system in which to study the accumulation rate of the GAA·TTC triplet repeats. We found that both alleles expanded on passage in culture (from passages 6 to 19). PCR band sizes were quantified by an inverse power function directly correlating gel migration of a molecular weight ladder to its component lengths. PCR products from the *FXN* locus contain 499 bp of non-repeat sequences, so GAA·TTC triplet-repeat numbers were estimated accordingly (Fig. 2*B*).

Downloaded from www.jbc.org at CALIFORNIA INSTITUTE OF TECHNOLOGY, on September 27, 2012



B

	allele 1 (# GAA)	allele 2 (# GAA)
3816	433	323
3816-iPSCs P6	516	377
3816-iPSCs P7	527	379
3816-iPSCs P9	548	392
3816-iPSCs P11	569	403
3816-iPSCs P13	592	422
3816-iPSCs P15	632	447
3816-iPSCs P17	645	468
3816-iPSCs P19	688	493

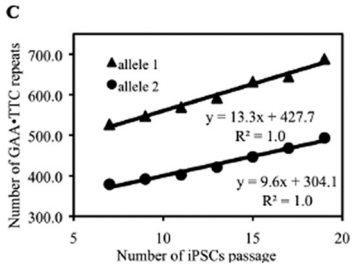


FIGURE 2. **GAA·TTC triplet-repeat expansion over time.** *A*, gel analysis (from *left* to *right*) of PCR products from DNA obtained from FRDA fibroblasts (GM03816) and FRDA iPSCs (clone 8) at different passages (*P*). *B*, analysis of the GAA·TTC triplet-repeat length using ImageJ software. Repeat numbers are shown for each allele. *C*, GAA·TTC triplet-repeat expansion rate (over passage).

Because FRDA is a recessive disease, both alleles are affected, and both alleles expand during iPSC propagation. The rate of GAA·TTC triplet-repeat expansion in the longer expanded allele was greater than in the shorter expanded allele (Fig. 2*C*). At each iPSC passage, the longer expanded allele accumulated 13.3 GAA·TTC triplet repeats, whereas the shorter expanded allele accumulated 9.6 GAA·TTC triplet repeats.

Potential links between GAA·TTC triplet-repeat expansion and DNA replication have been studied in a mammalian cell model using a plasmid with GAA·TTC triplet repeats. In this plasmid model, GAA·TTC triplet-repeat expansions have been found to be replication-dependent (10). In the iPSCs, the accumulation of repeats in each replication could be estimated. Generally, iPSCs were passaged about every 7 days. It has been reported that iPSC doubling time is ~29 h (32). Therefore, it can be calculated that GAA·TTC triplet repeats accumulated at a rate of ~2.3 (longer expanded allele) and ~1.7 (shorter expanded allele) repeats/replication cycle.

GAA·TTC Triplet-repeat Expansion Occurs in FRDA iPSCs, but Not in Differentiated Cells—iPSCs are a type of pluripotent stem cell artificially derived from a non-pluripotent patient fibroblast cell by inducing the "forced" expression of Oct4, Sox2, Klf4, and c-Myc. The generated iPSCs are remarkably similar to naturally isolated pluripotent stem cells (16). We were curious to determine whether GAA·TTC triplet-repeat expansion also occurs in FRDA fibroblasts and neurospheres (neural stem cells), which can be differentiated from FRDA iPSCs.

In vitro differentiation of FRDA iPSCs to neurospheres was performed as described previously (19–22). Neurospheres showed higher expression of the genes *MAP2* and *nestin* compared with the original fibroblasts and iPSCs (Fig. 3*A*). Flow cytometry showed that neurosphere populations were >90% positive for the marker protein nestin (supplemental Fig. S1). In addition, these neurospheres could be further differentiated to neuronal precursor cells (17). The FRDA neuronal precursor cells expressed neural marker proteins Tuj1 and MAP2, with a cell population of 93% (Tuj1) and 84% (MAP2), respectively (supplemental Fig. S1, *C–E*, and Fig. S2).

FRDA fibroblasts (GM03816 cells) were cultured and passaged every week and collected every 3 weeks for genomic DNA (passages 11–23). GAA·TTC triplet repeats were stable in FRDA fibroblasts during passage for 12 weeks (Fig. 3*B*). GAA·TTC triplet repeats began to expand in FRDA iPSCs (Fig. 3*C*) and stop expansion once differentiated into neurospheres, showing stability at weeks 0, 2, and 4 post-neurosphere formation (Fig. 3*C*). Interestingly, GAA·TTC triplet repeats continued expanding during the 2 weeks of noggin induction (passages 9–11*). Moreover, we also observed GAA·TTC triplet-repeat expansion between passage 11* and week 0 prior to complete neurosphere formation (2 weeks). One possibility is that intermediate cells may also express some of the pluripotency genes, which maintain GAA·TTC triplet-repeat expansion. GAA·TTC triplet repeats were also stable in neurospheres during a 10-week observation period (Fig. 3*D*).

Recently, several studies have reported triplet-repeat expansion and contraction during iPSC derivation (17, 33). To confirm whether GAA·TTC triplet-repeat expansion occurred during the reprogramming, FRDA iPSCs were collected after reprogramming (first collection from the reprogramming dishes, passage 0). Two independent clones both showed a clear expansion at passage 0 and at later passages (Fig. 3*E*). One explanation for contractions observed in other studies (17) is that these iPSC clones came from different fibroblasts,

Downloaded from www.jbc.org at CALIFORNIA INSTITUTE OF TECHNOLOGY, on September 27, 2012

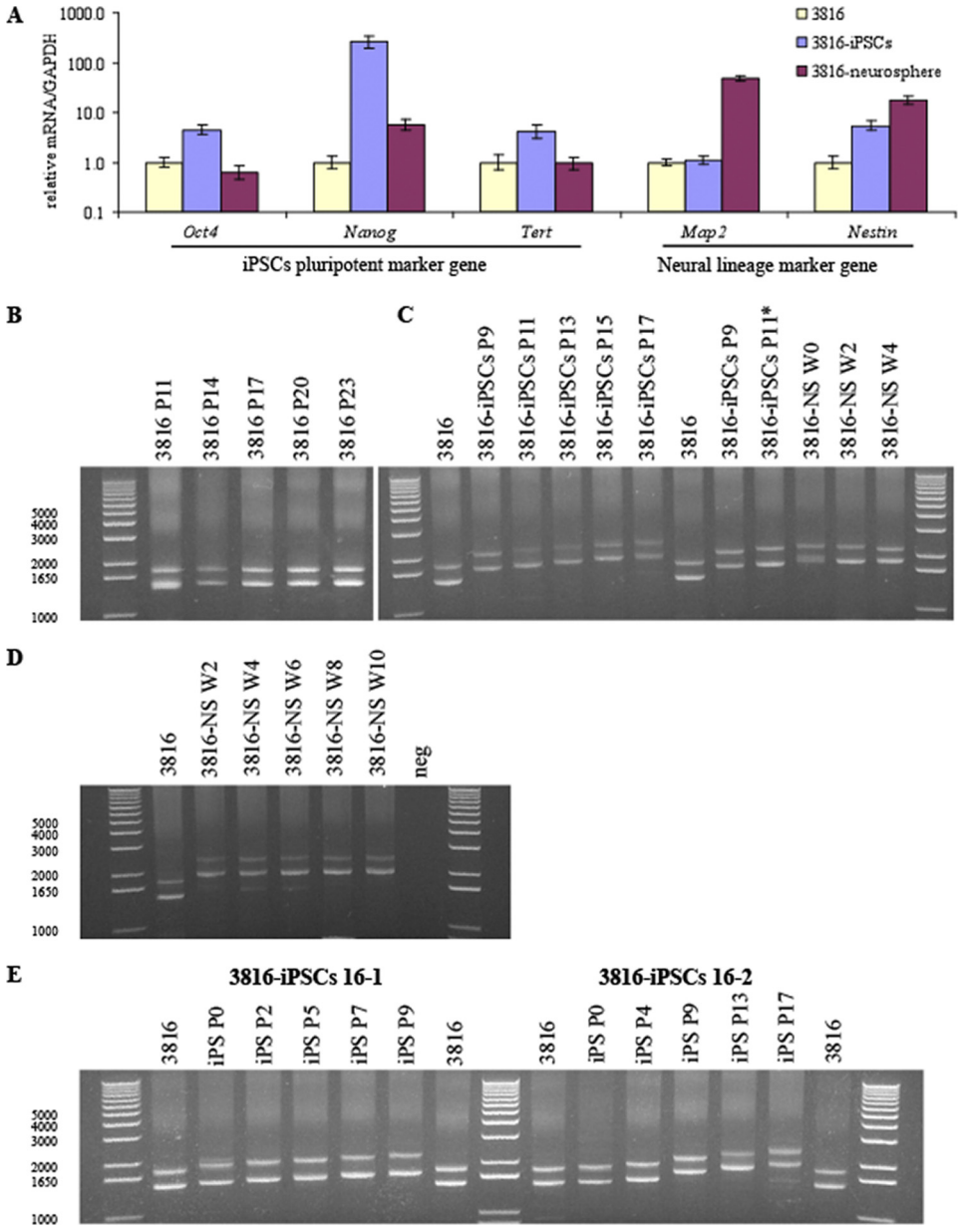


FIGURE 3. **GAA·TTC triplet-repeat expansion occurs in FRDA iPSCs, but not in fibroblasts or neurospheres.** *A*, quantitative RT-PCR analysis of mRNA markers of pluripotency (*Oct4*, *Nanog*, and *Tert*) and neuronal stem cells (*MAP2* and *nestin*). *B*, gel analysis of PCR products from genomic DNA obtained from FRDA fibroblasts (GM03816) at different passages. *C*, gel analysis of PCR products from genomic DNA obtained from FRDA iPSCs and FRDA neurosphere (*NS*) at different passages (*P*) or weeks (*W*). *P11** means that iPSCs were induced with noggin. *D*, GAA·TTC triplet repeats were stable in FRDA neurospheres over time. Gel analysis (from *left* to *right*) of PCR products from DNA obtained from FRDA fibroblasts (GM03816) and FRDA neurospheres at different collection times (weeks 2, 4, 6, 8, and 10). *neg*, negative. *E*, GAA·TTC triplet-repeat expansion over time. Shown are the results from gel analysis of PCR products from DNA obtained from FRDA fibroblasts (GM03816) and FRDA iPSCs (clones 16-1 and 16-2) at different passages.

each with different GAA·TTC triplet-repeat lengths. Another explanation is that expansion was not due to the cell reprogramming but was a consequence of iPSC formation after reprogramming.

GAA·TTC triplet repeats accumulate during propagation of the FRDA iPSCs in a manner analogous to the expansion observed in FRDA patient families (intergenerational repeat instability). Importantly, asymptomatic heterozygous carriers show GAA·TTC triplet-repeat expansion on the pathogenic allele, but not the normal allele (16). These observations open a new avenue for the investigation of the molecular mechanisms that underlie GAA·TTC triplet-repeat expansion.

MMR Proteins Are Highly Expressed in iPSCs—It has been reported that the MMR system is involved in triplet-repeat instability. In DM1 (myotonic dystrophy 1) transgenic mice, MSH3 is a limiting factor in the formation of intergenerational CTG·CAG triplet-repeat expansions, whereas MSH6 is not (34). Our previous work also showed that the MMR enzyme MSH2 is highly expressed in pluripotent cells and occupies *FXN* intron 1, and shRNA silencing of *MSH2* impedes GAA·TTC triplet-repeat expansion in the FRDA iPSCs (16).

In eukaryotes, MutS homologs MSH2, MSH3, and MSH6 have been identified and participate in MMR in the form of heterodimers: MutSα, MSH2/MSH6, and MutSβ, MSH2/

Downloaded from www.jbc.org at CALIFORNIA INSTITUTE OF TECHNOLOGY, on September 27, 2012

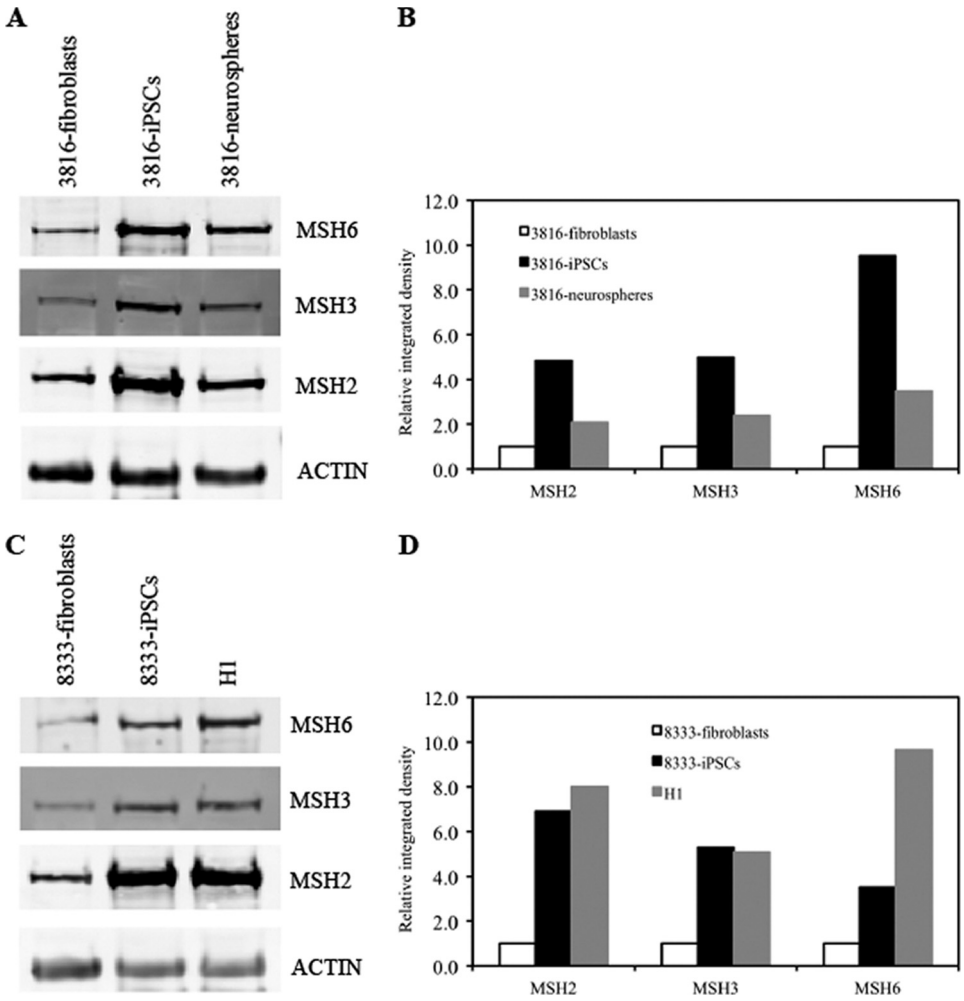


FIGURE 4. **MMR protein expression in fibroblasts, iPSCs, neurospheres, and H1 human ESCs.** *A*, Western blot results for MSH2, MSH3, and MSH6 in GM03816 fibroblasts, iPSCs, and neurospheres. *B*, densitometric analysis of MSH2, MSH3, and MSH6 protein normalized to actin in *A*. *C*, Western blot results for MSH2, MSH3, and MSH6 in unaffected fibroblasts (GM08333), unaffected iPSCs (GM08333), and H1 human ESCs. *D*, densitometric analysis of MSH2, MSH3, and MSH6 protein normalized to actin in *C*. Actin levels were used as a loading control.

MSH3 (35, 36). Additionally, results in DM1-derived human ESCs show significantly higher expression levels of MMR genes in undifferentiated cells compared with differentiated stages (37).

Because the protein levels of MSH2, MSH3, and MSH6 do not always reflect the level of their respective mRNA transcripts but are dependent upon Mutα and Mutβ complex stability (38), we monitored the levels of MSH2, MSH3, and MSH6 by Western blotting using extracts from the FRDA fibroblasts, iPSCs, and neurospheres. Indeed, expression of MSH2, MSH3, and MSH6 proteins was up-regulated in FRDA iPSCs compared with the parent fibroblasts (Fig. 4, *A* and *B*). These increased protein levels are not related to a FRDA phenotype, as unaffected GM08333 iPSCs showed similar increases in MSH2, MSH3, and MSH6 protein levels compared with fibroblasts and similar to the expression in H1 ESCs (Fig. 4, *C* and *D*). Once iPSCs were differentiated into neurospheres, MSH2, MSH3, and MSH6 protein levels were reduced (Fig. 4, *A* and *B*).

MMR Enzymes MSH2 and MSH6 Are Involved in GAA·TTC Triplet-repeat Expansion—It has been reported that MutSα (MSH2/MSH6) preferentially recognizes base-base mismatches and small insertion/deletion mispaired nucleotides, whereas MutSβ (MSH2/MSH3) preferentially recognizes larger insertion/deletion mispairs (35). Several studies have pointed to the involvement of MutSβ in triplet-repeat instability (8, 39), and MutSα (MSH2/MSH6) may also be involved in the formation of CTG·CAG triplet expansions during maternal transmission (34). Here, we investigated the role of MutSα and MutSβ in GAA·TTC triplet-repeat expansion in the FRDA iPSCs.

It has been shown that shRNA silencing of *MSH2* impedes GAA·TTC triplet-repeat expansion in the FRDA iPSCs (16); however, that study did not determine the effect of *MSH2* knockdown on the rate of repeat expansion or the role of the other MutSα and MutSβ subunits (MSH3 and MSH6) in repeat expansion. We therefore performed shRNA-mediated knockdown of the mRNAs encoding each of these proteins in FRDA iPSCs and monitored the effects of MMR knockdown on repeat expansion. shRNA knockdown of *MSH2* reduced the level of MSH3 protein, leaving MSH6 protein unaffected (Fig. 5*A*). Due to the sensitivity of our PCR assay, where amplification of the shorter allele is more efficient that that of the longer allele and where changes in repeat length are easier to quantify, we meas-

Downloaded from www.jbc.org at CALIFORNIA INSTITUTE OF TECHNOLOGY, on September 27, 2012

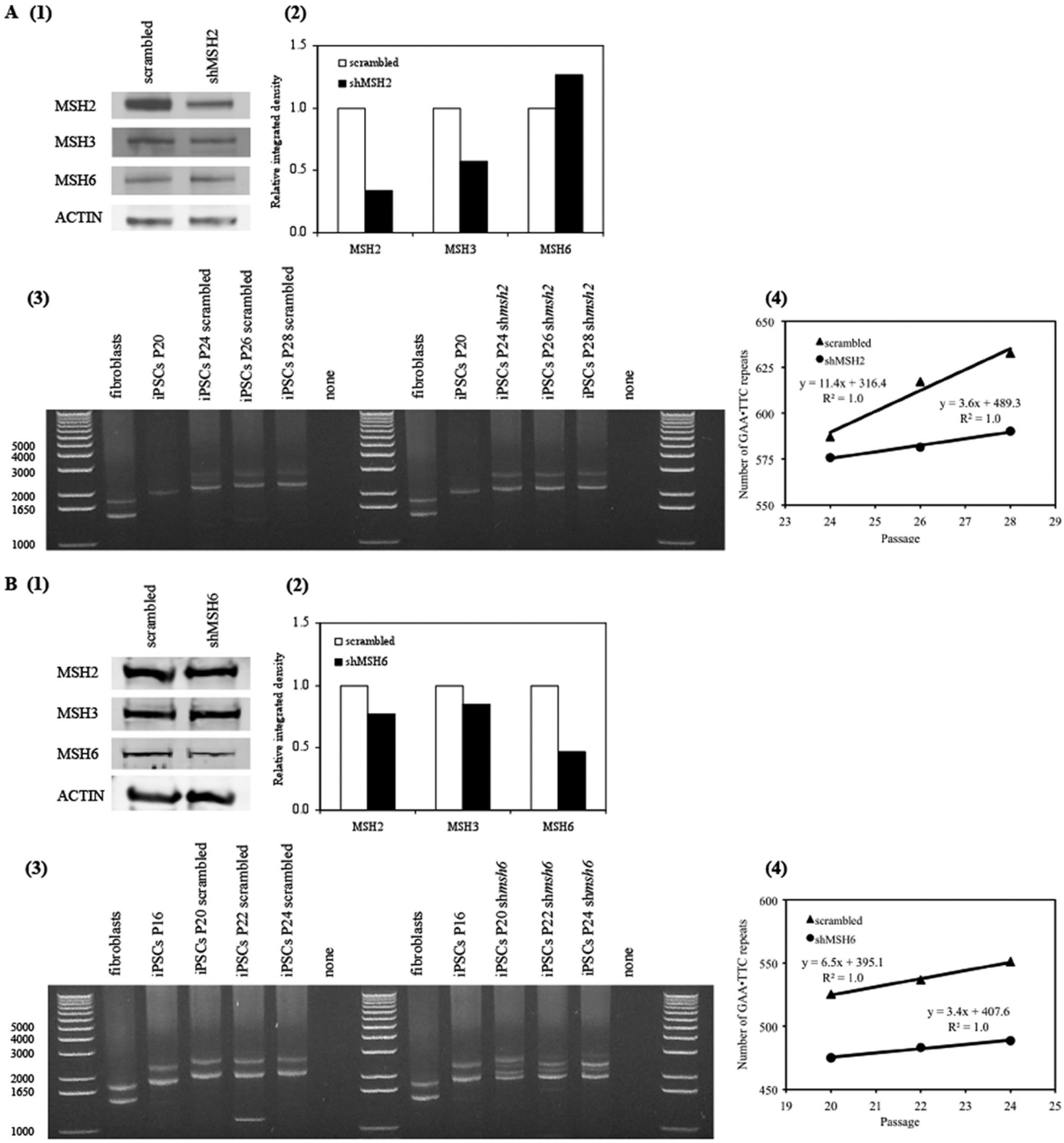


FIGURE 5. **GAA·TTC triplet-repeat expansion involves MMR enzymes MSH2 and MSH6.** *A*, *panel 1*, Western blot results for MSH2, MSH3, and MSH6 protein in FRDA iPSCs with scrambled shRNA and *MSH2*-targeted shRNA (*shMSH2*). *Panel 2*, densitometric analysis of MSH2, MSH3, and MSH6 protein normalized to actin in *panel 1*. *Panel 3*, PCR of GAA·TTC triplet repeats in fibroblasts, iPSCs before silencing (*P20*) and at four passages (*P24*), six passages (*P26*), and eight passages (*P28*) after transduction with scrambled and *MSH2* shRNAs. *Panel 4*, GAA·TTC triplet-repeat expansion rate (over passage) in iPSCs with scrambled and *MSH2* shRNAs. Here, the rate of GAA·TTC triplet-repeat expansion in the shorter expanded allele was analyzed. *B*, *panel 1*, Western blot results for MSH2, MSH3, and MSH6 protein in FRDA iPSCs with scrambled and *MSH6*-targeted shRNAs. *Panel 2*, densitometric analysis of MSH2, MSH3, and MSH6 protein normalized to actin in *panel 1*. *Panel 3*, PCR of GAA·TTC triplet-repeat length in fibroblasts before silencing (starting iPSCs; *P16*) and at four passages (*P20*), six passages (*P22*), and eight passages (*P24*) after transduction with scrambled and *MSH6* shRNAs. *Panel 4*, GAA·TTC triplet-repeat expansion rate (over passage) in iPSCs with scrambled and MSH6 shRNAs. Here, the rate of GAA·TTC triplet-repeat expansion in the shorter expanded allele was analyzed.

ured changes in repeat length on the shorter allele. The rate of GAA·TTC triplet-repeat expansion in the shorter expanded allele was analyzed after *MSH2* gene knockdown (Fig. 5*A*). iPSC genomic DNA was collected after four, six, and eight passages of *MSH2* knockdown. PCR analysis showed that the shorter expanded alleles accumulated 11.4 GAA·TTC triplet repeats at every passage in scrambled control iPSCs, but only 3.6 GAA·TTC triplet repeats in *MSH2* knockdown iPSCs (Fig. 5*A*). These data further implicate the involvement of MSH2 in GAA·TTC triplet-repeat instability.

Downloaded from www.jbc.org at CALIFORNIA INSTITUTE OF TECHNOLOGY, on September 27, 2012

MSH6 gene knockdown was also achieved (Fig. 5*B*), whereas *MSH3* gene knockdown was unsuccessful (data not shown). *MSH6* gene knockdown was analyzed at the protein level for all three MMR subunits, showing that MSH6 protein knockdown did not affect the MSH2 or MSH3 protein levels significantly (Fig. 5*B*). PCR analysis for repeat lengths showed that the MMR enzyme MutSα (MSH2/MSH6) is involved in GAA·TTC triplet-repeat expansion. The shorter expanded allele accumulated 6.5 GAA·TTC triplet repeats at every passage in scrambled control iPSCs, but only 3.4 GAA·TTC repeats in *MSH6* knockdown iPSCs (Fig. 5*B*). The difference between the scrambled control iPSCs in Fig. 5 (*A* and *B*) may come from a passage number difference between lines during the shRNA gene knockdown.

GAA·TTC Triplet Repeat-targeted Pyrrole-Imidazole Polyamide Impedes Repeat Expansion—Biochemical studies have documented that expanded GAA·TTC triplet repeats adopt unusual non-B DNA structures *in vitro*, such as triplexes, containing two purine GAA strands along with one pyrimidine TTC strand, flanking a single-stranded pyrimidine region (40), as well as intramolecular "sticky" DNA (41). Structural studies indicate that β-alanine-linked polyamides bind the minor groove of canonical B DNA (42). Given the high affinity of β-alanine-linked polyamides for purine tracts (43), these molecules might act as a thermodynamic "sink" and lock GAA·TTC triplet repeats into double-stranded B DNA. Such an event would disfavor duplex unpairing, which is necessary for formation of FRDA triplexes and sticky DNA, which may lead to GAA·TTC triplet-repeat expansion in FRDA iPSCs. It has been shown that β-alanine-linked pyrrole-imidazole polyamide FA1 (ImPy-β-ImPy-β-Im-β-Dp, where Py is pyrrole, Im is imidazole, β is β-alanine, and Dp is dimethylaminopropylamine) binds GAA·TTC triplet-repeat DNA in its B-type conformation, shifts the equilibrium from sticky DNA back to B-type DNA, and relieves transcriptional repression of the *FXN* gene in FRDA cells (Fig. 6*A*) (27).

Quantitative DNase I footprinting experiments showed that polyamide FA1 binds the 9-bp site 5′-AAGAAGAAG-3′ (as in GAA·TTC triplet-repeat DNA) with an apparent dissociation constant (K_d) of 0.1 nM (27). Quantitative footprint titration experiments also revealed that polyamide HIV-1 (ImPy-β-ImPy-γ-ImPy-β-ImPy-β-Dp, where γ is γ-butyric acid) binds to sites adjacent to and overlapping the HIV-1 TATA element (5′-(A/T)GC(A/T)GC(A/T)-3′) with a K_d of 0.05 nM (28), and subsequent studies showed that this same molecule targets CAG·CTG triplet-repeat DNA with high affinity (Fig. 6*B*) (29). Both molecules have been found to be cell-permeable, and fluorescent conjugates localize in the cell nucleus (supplemental Fig. S3) (27, 44). Thus, polyamide HIV-1 was used as a control in the following experiments. We asked whether these polyamides would have an effect on GAA·TTC triplet-repeat expansion in FRDA iPSCs. At 5 μM, FA1 impeded GAA·TTC triplet-repeat expansion in FRDA iPSCs, whereas the control polyamide HIV-1 did not (Fig. 6, *C* and *D*).

We postulate that polyamide FA1 reverses a non-B DNA conformation, causing displacement of MMR enzymes and thereby preventing repeat expansion. ChIP assays showed that FA1 removed the MMR enzyme MSH2 from the region downstream of the repeats (Fig. 6*E*). This result was consistent with the *MSH2* shRNA knockdown that removed the MMR enzyme MSH2 from the region downstream of the repeats (Fig. 6*F*).

DISCUSSION

Accumulating data from *E. coli*, yeast, mammalian cell, and mouse models reveal that somatic gene mosaicism and triplet-repeat expansion occur as the result of multiple small changes in length-dependent, transcription-dependent, and DNA MMR-dependent processes. The most parsimonious model that can account for these dynamics involves the DNA MMR machinery (45).

Generation of expanded GAA·TTC triplet repeats in FRDA iPSCs can be explained using a MMR model. In this model, possibly due to transcription and RNA·DNA hybrid formation (46), the GAA·TTC triplet repeats adopt an unusual non-B DNA structure. The loop outs are recognized by MutSβ (MSH2/MSH3) and MutSα (MSH2/MSH6), which initiate a DNA MMR reaction by recruiting a MutL heterodimer. If the complementary loop outs are far enough apart, they will be repaired as independent events, and if repair is directed to incorporate the loop out on the opposite stand, a net gain of repeats will result that is equivalent to the size of the original misalignment. Multiple rounds of misalignment and MMR could lead to the accumulation of large GAA·TTC triplet-length changes.

Whether the mechanism of GAA·TTC triplet-repeat generation is really as simple as inappropriate DNA MMR remains to be elucidated. However, progress achieved in understanding the important factors mediating repeat instability will be extended through the use of an impressive array of model systems now established to test competing models. FRDA iPSCs would be a valuable model to study repeat instability in the human genome.

Is there any relationship between triplet-repeat instability and reprogramming? Recently, several studies have reported triplet-repeat expansions and contractions during iPSC derivation (17, 33). There are three ways to explain these repeat expansions and contractions. First, iPSC clones come from a fibroblast cell with shorter or longer repeats. Second, repeat expansion occurs during the reprogramming stage. The repeat size difference between fibroblasts and iPSCs comes from the propagation of iPSCs. The very early iPSCs (passage 0) showing expansion in this study resulted after 3 weeks of propagation, in which transduced fibroblasts grew and formed iPSC clones. The repeat expansion rate between fibroblasts and the very early iPSCs (passage 0) is similar to expansion during later iPSC propagation (Fig. 3*E*). Third, there is a true expansion and contraction during cell reprogramming. However, this possibility needs more experimental support. For example, one could differentiate iPSCs back to fibroblasts and check the repeat lengths. These new fibroblasts will have less somatic instability because the iPSCs come from one fibroblast cell. Then, reprogram the newly generated fibroblasts back into iPSCs and check the length of the repeats. Although this process will not fully answer the question, it can further confirm expansion and contraction during the reprogramming.

Is MSH6 involved in GAA·TTC triplet-repeat expansion? Several studies have pointed to the involvement of MutSβ

Downloaded from www.jbc.org at CALIFORNIA INSTITUTE OF TECHNOLOGY, on September 27, 2012

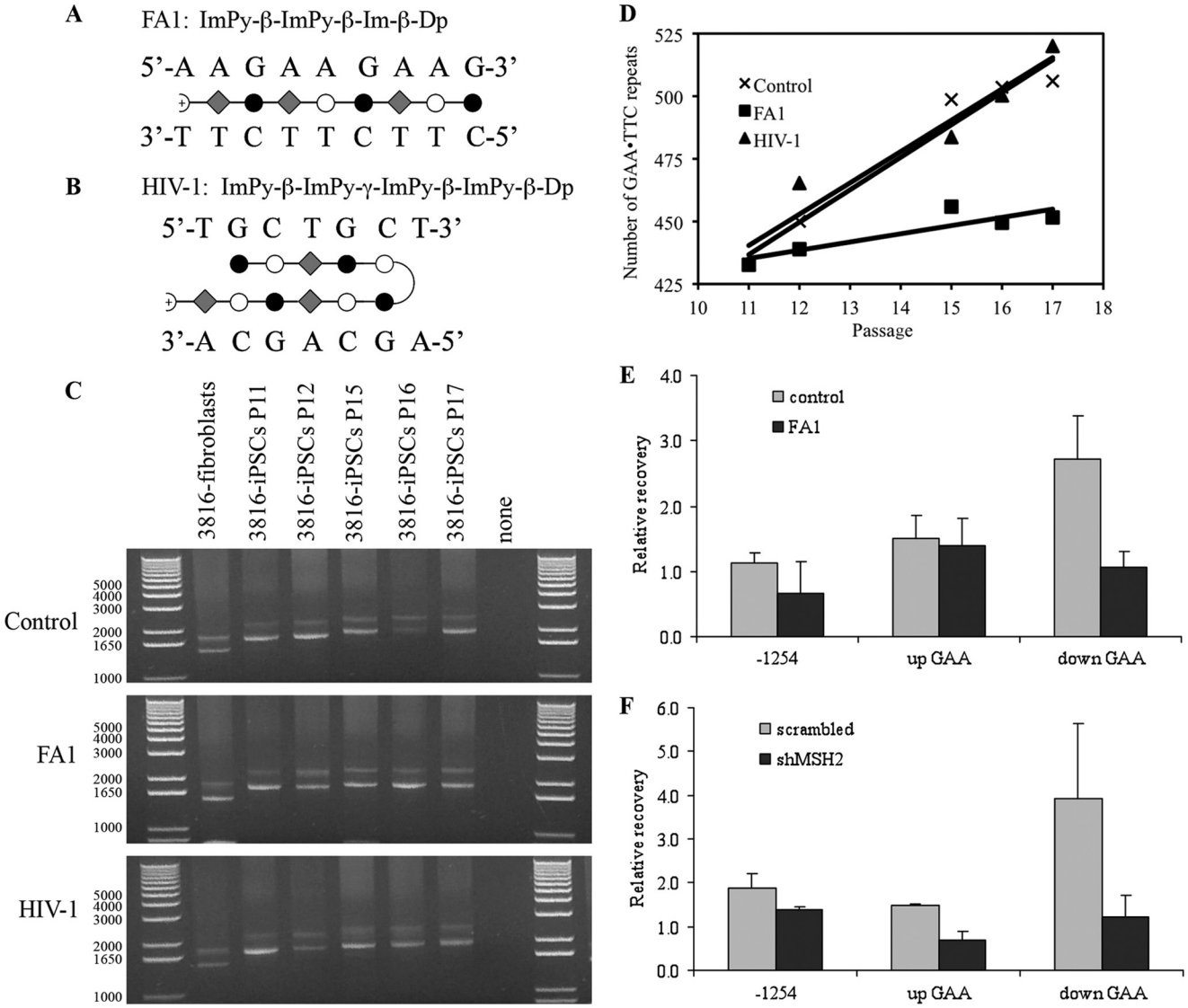


FIGURE 6. **Polyamide FA1, which targets GAA·TTC triplet-repeat DNA, partially blocks repeat expansion in FRDA iPSCs.** *A*, polyamide FA1 sequence and binding model. *B*, polyamide HIV-1 sequence and binding model. *Filled* and *open circles* are imidazole (*Im*) and pyrrole (*Py*) rings, respectively; *diamonds* are *β*-alanine (*β*); the *curved line* is γ-aminobutyric acid (*γ*); and the *semicircle* with a *plus sign* is dimethylaminopropylamine (*Dp*). *C* and *D*, PCR of GAA·TTC triplet repeats and GAA·TTC triplet-repeat expansion rate (over passage (*P*)), respectively, in untreated, FA1-treated and HIV-1-treated iPSCs. The HIV-1 polyamide targets CTG·CAG triplet-repeat DNA. *E*, polyamide FA1 targeting GAA·TTC triplet-repeat DNA displaces MSH2 from *FXN* intron 1 in FRDA iPSCs. *F*, *MSH2*-targeted shRNA (*shMSH2*) displaces MSH2 from *FXN* intron 1 in FRDA iPSCs. ChIP experiments were performed with anti-MSH2 antibody in FRDA iPSCs using PCR primers for a region on the *FXN* gene 1254 bp upstream of the transcriptional start site (*−1254*) and upstream (*up GAA*) and downstream (*down GAA*) of the GAA·TTC triplet-repeat expansion. Immunoprecipitation recovery is relative to intron 2.

(MSH2/MSH3) in triplet-repeat instability in transgenic mice (8, 15, 39). Our previous ChIP assays (16) also show that at a resolution of ~1 kb, there is an increased occupancy of MSH2 and MSH3 downstream of the GAA·TTC triplet repeats in FRDA iPSCs compared with an unaffected iPSC line, but, in contrast, not 1254 bp upstream of the *FXN* transcriptional start site or directly upstream of the GAA·TTC triplet repeats. In addition, no differences in MSH6 occupancy were found at any of the regions probed. However, these results cannot rule out the involvement of MSH6 in GAA·TTC triplet-repeat expansion. One possibility is that MSH6 may have an indirect effect on GAA·TTC triplet-repeat expansion in iPSCs. Because the mechanism of MMR in mammalian cells is not well understood, future biochemical studies could investigate the function of MSH6 in MMR. iPSCs clearly provide a mammalian cell model to study the mechanisms of MMR.

Are other factors involved in GAA·TTC triplet-repeat expansion? Higher expression of mismatch enzymes in iPSCs could be one key factor that promotes the GAA·TTC triplet-repeat expansion. However, two results suggest that there may be other factors contributing to GAA·TTC triplet-repeat expansion in iPSCs. First, asymptomatic heterozygous carriers show GAA·TTC triplet-repeat expansion only on the pathogenic allele, but not on the normal allele in FRDA iPSCs (16). This result strongly supports models in which expanded alleles adopt non-B DNA structures, which in turn promote further expansion by recruitment of the MMR enzymes. However, the exact mechanisms involved in MMR recruitment remain to be

Downloaded from www.jbc.org at CALIFORNIA INSTITUTE OF TECHNOLOGY, on September 27, 2012

identified. Second, in FRDA patients, somatic GAA·TTC triplet-repeat instability is tissue-specific. This finding could implicate other, as yet unidentified factors in repeat expansion. Because we cannot target *FXN* transcription, DNA replication, or active MMR enzymes to block GAA·TTC triplet-repeat instability, factors that are specifically related to this instability could represent a potential drug target to treat FRDA.

Specific β-alanine-linked pyrrole-imidazole polyamides could be one way to impede the GAA·TTC triplet-repeat expansion. Although the existence of GAA·TTC and other triplet-repeat sequences has been known for some time, there is still no effective method or compound to prevent repeat instability *in vivo*. In this study, we have shown that a specific polyamide (FA1) can impede the GAA·TTC triplet-repeat expansion in FRDA iPSCs, likely by binding and displacing MMR enzymes from the repeats. This observation opens a new avenue for investigating and designing new compounds that have the ability to target GAA·TTC triplet repeats, as well as other repeats, to block repeat instability. The ability to impede expansion raises the hope that the severity of pathophysiology might be reduced or its onset delayed, thereby widening the therapeutic window for these deadly triplet-repeat diseases (3).

Acknowledgment—We thank Dr. Chunping Xu for HPLC and MS analyses of the polyamides.

REFERENCES

1. Mirkin, S. M. (2007) Expandable DNA repeats and human disease. *Nature* **447,** 932–940
2. Kovtun, I. V., and McMurray, C. T. (2008) Features of trinucleotide repeat instability *in vivo*. *Cell Res.* **18,** 198–213
3. McMurray, C. T. (2010) Mechanisms of trinucleotide repeat instability during human development. *Nat. Rev. Genet.* **11,** 786–799
4. Herman, D., Jenssen, K., Burnett, R., Soragni, E., Perlman, S. L., and Gottesfeld, J. M. (2006) Histone deacetylase inhibitors reverse gene silencing in Friedreich ataxia. *Nat. Chem. Biol.* **2,** 551–558
5. Campuzano, V., Montermini, L., Moltò, M. D., Pianese, L., Cossée, M., Cavalcanti, F., Monros, E., Rodius, F., Duclos, F., Monticelli, A., Zara, F., Cañizares, J., Koutnikova, H., Bidichandani, S. I., Gellera, C., Brice, A., Trouillas, P., De Michele, G., Filla, A., De Frutos, R., Palau, F., Patel, P. I., Di Donato, S., Mandel, J. L., Cocozza, S., Koenig, M., and Pandolfo, M. (1996) Friedreich ataxia: autosomal recessive disease caused by an intronic GAA triplet-repeat expansion. *Science* **271,** 1423–1427
6. Bidichandani, S. I., Ashizawa, T., and Patel, P. I. (1998) The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. *Am. J. Hum. Genet.* **62,** 111–121
7. Pollard, L. M., Bourn, R. L., and Bidichandani, S. I. (2008) Repair of DNA double-strand breaks within the $(GAA{\cdot}TTC)_n$ sequence results in frequent deletion of the triplet-repeat sequence. *Nucleic Acids Res.* **36,** 489–500
8. Kim, H. M., Narayanan, V., Mieczkowski, P. A., Petes, T. D., Krasilnikova, M. M., Mirkin, S. M., and Lobachev, K. S. (2008) Chromosome fragility at GAA tracts in yeast depends on repeat orientation and requires mismatch repair. *EMBO J.* **27,** 2896–2906
9. Shishkin, A. A., Voineagu, I., Matera, R., Cherng, N., Chernet, B. T., Krasilnikova, M. M., Narayanan, V., Lobachev, K. S., and Mirkin, S. M. (2009) Large-scale expansions of Friedreich ataxia GAA repeats in yeast. *Mol. Cell* **35,** 82–92
10. Rindler, P. M., and Bidichandani, S. I. (2011) Role of transcript and interplay between transcription and replication in triplet-repeat instability in mammalian cells. *Nucleic Acids Res.* **39,** 526–535
11. Ditch, S., Sammarco, M. C., Banerjee, A., and Grabczyk, E. (2009) Progressive GAA·TTC repeat expansion in human cell lines. *PLoS Genet.* **5,** e1000704
12. Rindler, P. M., Clark, R. M., Pollard, L. M., De Biase, I., and Bidichandani, S. I. (2006) Replication in mammalian cells recapitulates the locus-specific differences in somatic instability of genomic GAA triplet repeats. *Nucleic Acids Res.* **34,** 6352–6361
13. Clark, R. M., De Biase, I., Malykhina, A. P., Al-Mahdawi, S., Pook, M., and Bidichandani, S. I. (2007) The GAA triplet repeat is unstable in the context of the human *FXN* locus and displays age-dependent expansions in cerebellum and DRG in a transgenic mouse model. *Hum. Genet.* **120,** 633–640
14. Al-Mahdawi, S., Pinto, R. M., Ruddle, P., Carroll, C., Webster, Z., and Pook, M. (2004) GAA repeat instability in Friedreich ataxia YAC transgenic mice. *Genomics* **84,** 301–310
15. 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
16. 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 ataxia induced pluripotent stem cells model intergenerational GAA·TTC triplet-repeat instability. *Cell Stem Cell* **7,** 631–637
17. Liu, J., Verma, P. J., Evans-Galea, M. V., Delatycki, M. B., Michalska, A., Leung, J., Crombie, D., Sarsero, J. P., Williamson, R., Dottori, M., and Pébay, A. (2011) Generation of induced pluripotent stem cell lines from Friedreich ataxia patients. *Stem Cell Rev.* **7,** 703–713
18. Park, I. H., Zhao, R., West, J. A., Yabuuchi, A., Huo, H., Ince, T. A., Lerou, P. H., Lensch, M. W., and Daley, G. Q. (2008) Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451,** 141–146
19. Dottori, M., and Pera, M. F. (2008) Neural differentiation of human embryonic stem cells. *Methods Mol. Biol.* **438,** 19–30
20. Zhou, J., Su, P., Li, D., Tsang, S., Duan, E., and Wang, F. (2010) High-efficiency induction of neural conversion in human ESCs and human induced pluripotent stem cells with a single chemical inhibitor of transforming growth factor β superfamily receptors. *Stem Cells* **28,** 1741–1750
21. Chambers, S. M., Fasano, C. A., Papapetrou, E. P., Tomishima, M., Sadelain, M., and Studer, L. (2009) Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat. Biotechnol.* **27,** 275–280
22. Kim, D. S., Lee, J. S., Leem, J. W., Huh, Y. J., Kim, J. Y., Kim, H. S., Park, I. H., Daley, G. Q., Hwang, D. Y., and Kim, D. W. (2010) Robust enhancement of neural differentiation from human ES and iPS cells regardless of their innate difference in differentiation propensity. *Stem Cell Rev.* **6,** 270–281
23. Laird, P. W., Zijderveld, A., Linders, K., Rudnicki, M. A., Jaenisch, R., and Berns, A. (1991) Simplified mammalian DNA isolation procedure. *Nucleic Acids Res.* **19,** 4293
24. Abramoff, M. D., Magalhaes, P. J., and Ram, S. J. (2004) Image Processing with ImageJ. *Biophotonics Int.* **11,** 36–42
25. Monckton, D. G., Wong, L. J., Ashizawa, T., and Caskey, C. T. (1995) Somatic mosaicism, germ-line expansions, germ-line reversions, and intergenerational reductions in myotonic dystrophy males: small-pool PCR analyses. *Hum. Mol. Genet.* **4,** 1–8
26. Livak, K. J., and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25,** 402–408
27. Burnett, R., Melander, C., Puckett, J. W., Son, L. S., Wells, R. D., Dervan, P. B., and Gottesfeld, J. M. (2006) DNA sequence-specific polyamides alleviate transcription inhibition associated with long GAA·TTC repeats in Friedreich ataxia. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 11497–11502
28. Dickinson, L. A., Gulizia, R. J., Trauger, J. W., Baird, E. E., Mosier, D. E., Gottesfeld, J. M., and Dervan, P. B. (1998) Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proc. Natl. Acad. Sci. U.S.A.* **95,** 12890–12895
29. Fujimoto, J., Bando, T., Minoshima, M., Uchida, S., Iwasaki, M., Shinohara, K., and Sugiyama, H. (2008) Detection of triplet repeat sequences in the double-stranded DNA using pyrene-functionalized pyrrole-imidazole polyamides with rigid linkers. *Bioorg. Med. Chem.* **16,** 5899–5907
30. Sharma, R., Bhatti, S., Gomez, M., Clark, R. M., Murray, C., Ashizawa, T.,

Downloaded from www.jbc.org at CALIFORNIA INSTITUTE OF TECHNOLOGY, on September 27, 2012

and Bidichandani, S. I. (2002) The GAA triplet-repeat sequence in Friedreich ataxia shows a high level of somatic instability *in vivo*, with a significant predilection for large contractions. *Hum. Mol. Genet.* **11,** 2175–2187
31. Gomes-Pereira, M., Bidichandani, S. I., and Monckton, D. G. (2004) Analysis of unstable triplet repeats using small-pool polymerase chain reaction. *Methods Mol. Biol.* **277,** 61–76
32. Ding, V. M., Ling, L., Natarajan, S., Yap, M. G., Cool, S. M., and Choo, A. B. (2010) FGF-2 modulates Wnt signaling in undifferentiated hESC and iPS cells through activated PI3K/GSK3β signaling. *J. Cell. Physiol.* **225,** 417–428
33. Sheridan, S. D., Theriault, K. M., Reis, S. A., Zhou, F., Madison, J. M., Daheron, L., Loring, J. F., and Haggarty, S. J. (2011) Epigenetic characterization of the *FMR1* gene and aberrant neurodevelopment in human induced pluripotent stem cell models of fragile X syndrome. *PLoS ONE* **6,** e26203
34. 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
35. Li, G. M. (2008) Mechanisms and functions of DNA mismatch repair. *Cell Res.* **18,** 85–98
36. Jiricny, J. (2006) The multifaceted mismatch repair system. *Nat. Rev. Mol. Cell Biol.* **7,** 335–346
37. Seriola, A., Spits, C., Simard, J. P., Hilven, P., Haentjens, P., Pearson, C. E., and Sermon, K. (2011) Huntington and myotonic dystrophy hESCs: down-regulated trinucleotide repeat instability and mismatch repair machinery expression upon differentiation. *Hum. Mol. Genet.* **20,** 176–185
38. Marra, G., Iaccarino, I., Lettieri, T., Roscilli, G., Delmastro, P., and Jiricny, J. (1998) Mismatch repair deficiency associated with overexpression of the *MSH3* gene. *Proc. Natl. Acad. Sci. U.S.A.* **95,** 8568–8573
39. Dragileva, E., Hendricks, A., Teed, A., Gillis, T., Lopez, E. T., Friedberg, E. C., Kucherlapati, R., Edelmann, W., Lunetta, K. L., MacDonald, M. E., and Wheeler, V. C. (2009) Intergenerational and striatal CAG repeat instability in Huntington disease knock-in mice involves different DNA repair genes. *Neurobiol. Dis.* **33,** 37–47
40. Sakamoto, N., Chastain, P. D., Parniewski, P., Ohshima, K., Pandolfo, M., Griffith, J. D., and Wells, R. D. (1999) Sticky DNA: self-association properties of long GAA·TTC repeats in R·R·Y triplex structures from Friedreich ataxia. *Mol. Cell* **3,** 465–475
41. Napierala, M., Dere, R., Vetcher, A., and Wells, R. D. (2004) Structure-dependent recombination hot spot activity of GAA·TTC sequences from intron 1 of the Friedreich ataxia gene. *J. Biol. Chem.* **279,** 6444–6454
42. Urbach, A. R., Love, J. J., Ross, S. A., and Dervan, P. B. (2002) Structure of a β-alanine-linked polyamide bound to a full helical turn of purine tract DNA in the 1:1 motif. *J. Mol. Biol.* **320,** 55–71
43. Urbach, A. R., and Dervan, P. B. (2001) Toward rules for 1:1 polyamide: DNA recognition. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 4343–4348
44. Dudouet, B., Burnett, R., Dickinson, L. A., Wood, M. R., Melander, C., Belitsky, J. M., Edelson, B., Wurtz, N., Briehn, C., Dervan, P. B., and Gottesfeld, J. M. (2003) Accessibility of nuclear chromatin by DNA-binding polyamides. *Chem. Biol.* **10,** 859–867
45. 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
46. Lin, Y., Dent, S. Y., Wilson, J. H., Wells, R. D., and Napierala, M. (2010) R loops stimulate genetic instability of CTG·CAG repeats. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 692–697

Downloaded from www.jbc.org at CALIFORNIA INSTITUTE OF TECHNOLOGY, on September 27, 2012