PLOS ONE

Friedreich Ataxia Patient Tissues Exhibit Increased 5-Hydroxymethylcytosine Modification and Decreased CTCF Binding at the *FXN* Locus

Sahar Al-Mahdawi[1], Chiranjeevi Sandi[1], Ricardo Mouro Pinto[2], Mark A. Pook[1*]

1 Ataxia Research Group, Division of Biosciences, School of Health Sciences and Social Care, Brunel University, Uxbridge, Middlesex, United Kingdom,
2 Molecular Neurogenetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, United States of America

Abstract

Background: Friedreich ataxia (FRDA) is caused by a homozygous GAA repeat expansion mutation within intron 1 of the *FXN* gene, which induces epigenetic changes and *FXN* gene silencing. Bisulfite sequencing studies have identified 5-methylcytosine (5mC) DNA methylation as one of the epigenetic changes that may be involved in this process. However, analysis of samples by bisulfite sequencing is a time-consuming procedure. In addition, it has recently been shown that 5-hydroxymethylcytosine (5hmC) is also present in mammalian DNA, and bisulfite sequencing cannot distinguish between 5hmC and 5mC.

Methodology/Principal Findings: We have developed specific MethylScreen restriction enzyme digestion and qPCR-based protocols to more rapidly quantify DNA methylation at four CpG sites in the *FXN* upstream GAA region. Increased DNA methylation was confirmed at all four CpG sites in both FRDA cerebellum and heart tissues. We have also analysed the DNA methylation status in FRDA cerebellum and heart tissues using an approach that enables distinction between 5hmC and 5mC. Our analysis reveals that the majority of DNA methylation in both FRDA and unaffected tissues actually comprises 5hmC rather than 5mC. We have also identified decreased occupancy of the chromatin insulator protein CTCF (CCCTC-binding factor) at the *FXN* 5' UTR region in the same FRDA cerebellum tissues.

Conclusions/Significance: Increased DNA methylation at the *FXN* upstream GAA region, primarily 5hmC rather than 5mC, and decreased CTCF occupancy at the *FXN* 5' UTR are associated with FRDA disease-relevant human tissues. The role of such molecular mechanisms in FRDA pathogenesis has now to be determined.

Introduction

FRDA is an autosomal recessive neurodegenerative mitochondrial disorder caused primarily by a homozygous GAA repeat expansion mutation within intron 1 of the *FXN* gene [1]. Unaffected individuals have up to 43 GAA repeats, while affected individuals have 44 to 1700 GAA repeats, most commonly between 600–900 GAA repeats [2,3]. The length of the smaller GAA repeat correlates with FRDA disease severity and inversely correlates with the age of onset [4,5]. The effect of the GAA repeat expansion is to decrease expression of the essential and ubiquitously expressed mitochondrial frataxin protein, with levels in FRDA patients ranging from 4% to 29% that of normal [6]. Reduced levels of frataxin in FRDA patients

are associated with defects of iron–sulphur (Fe–S) cluster biosynthesis [7], mitochondrial iron accumulation in the heart, spinal cord and dentate nucleus of the cerebellum [8] and increased susceptibility to oxidative stress [9].

Two main hypotheses have been proposed to link *FXN* GAA repeat expansions with decreased frataxin expression. Firstly, evidence from *in vitro* and cell transfection studies suggests that GAA repeat expansions may adopt abnormal non-B DNA structures (triplexes or "sticky DNA") or DNA•RNA hybrid structures (R-loops), which impede the process of RNA polymerase II and thus reduce *FXN* gene transcription [10,11]. Secondly, there is evidence that GAA repeat expansions can induce heterochromatin-mediated gene silencing effects [12]. Consistent with the latter hypothesis, several FRDA-related

epigenetic changes have been identified in the immediate vicinity of the expanded GAA repeats of the *FXN* gene [13]. An initial investigation of DNA methylation of the *FXN* gene by bisulfite sequencing revealed hypermethylation of the cytosine residue of specific CpG sites upstream of the GAA repeat sequence in FRDA patient lymphoblastoid cells compared to cells derived from unaffected individuals [14]. We have subsequently used bisulfite sequencing to identify increased DNA methylation at the upstream GAA repeat region in FRDA patient cerebellum and heart autopsy tissues, which are clinically relevant tissues in FRDA [15]. Interestingly, we also identified reduced levels of DNA methylation in the downstream GAA repeat region in FRDA patient tissues compared with controls. These findings have now been confirmed by bisulfite-based EpiTYPER MassARRAY analysis of blood and buccal cell samples from a large cohort of FRDA patients, where a significant inverse correlation was also detected between the level of DNA methylation in the upstream GAA region and the level of *FXN* expression [16]. Yet another study has shown that the degree of DNA methylation in the upstream GAA repeat region in FRDA patients correlates with the length of the GAA repeats and inversely correlates with the age of disease onset [17]. Therefore, there is good evidence that DNA methylation may have some role to play in the molecular mechanism of GAA repeat induced FRDA disease. However, bisulfite sequencing is a somewhat time-consuming procedure. Therefore, we now report a more rapid MethylScreen restriction enzyme digestion and qPCR procedure, which we have validated by confirming increased DNA methylation at the upstream GAA repeat region in FRDA patient cerebellum and heart autopsy tissues.

Methylation of mammalian DNA primarily occurs as 5mC modification of CpG dinucleotides. However, recent studies have revealed the existence of an alternative modification, 5hmC, which is formed by oxidation of 5mC by ten-eleven translocation (TET) enzymes [18,19]. The overall levels of 5hmC in the mammalian genome are approximately 10% of 5mC levels [20], although higher levels have been detected in tissues of the central nervous system [21]. For example, 5hmC is approximately 40% as abundant as 5mC in the DNA of Purkinje cells of the cerebellum [18]. There are two main theories regarding the function of 5hmC. Firstly, 5hmC may be an intermediate in the removal of 5mC by an active demethylation process that involves the base excision repair pathway [22]. Secondly, 5hmC may be an epigenetic modification in its own right, regulating chromatin or transcriptional factors involved in processes such as neurodevelopment [23]. Therefore, it is possible that 5hmC may also be involved in epigenetic-based disease mechanisms, such as those proposed for FRDA. However, useful information cannot be obtained from previous bisulfite sequence DNA methylation data, because this method does not distinguish between 5mC and 5hmC [24]. Therefore, we have analysed the DNA methylation status of one of the *FXN* upstream GAA CpG sites in FRDA cerebellum and heart tissues using a procedure that does allow distinction between 5hmC and 5mC and we report that the majority of the methylated DNA at this CpG residue comprises 5hmC rather than 5mC.

In addition, another factor that has been proposed to be involved in silencing of the *FXN* gene in FRDA is the chromatin insulator protein CTCF, because decreased occupancy of CTCF has been identified at the *FXN* 5' UTR region of FRDA fibroblast cells [25]. Therefore, we have decided to further investigate CTCF occupancy in FRDA cerebellum tissue and we report decreased occupancy of CTCF at the *FXN* 5' UTR region. Overall we show that increased 5hmC at the *FXN* upstream GAA region and decreased CTCF occupancy at the *FXN* 5' UTR are both associated with FRDA disease-relevant human tissues.

Materials and Methods

DNA preparation from human tissues

Human cerebellum and heart tissue samples were obtained from autopsies of four FRDA patients (aged 17, 24, 36 and 47, each homozygous for GAA repeat expansions within the range of approximately 550-750 units) and four unaffected individuals (aged 69-82, obtained from NDRI). Research ethical approval was obtained from the SHSSC Research Ethics Committee, Brunel University, and samples were obtained with written consent and were stored in accordance with UK Human Tissue Authority ethical guidelines. For the FRDA patient aged 17, written informed consent was obtained from the next of kin under the organisation of the Friedreich's Ataxia Group tissue donor scheme. Genomic DNA was isolated from frozen tissues by standard phenol/chloroform extraction and ethanol precipitation.

MethylScreen analysis

Four CpG sites in the *FXN* upstream GAA repeat region, designated 3, 6, 11 and 13 (numbered according to [14], Figure 1), were investigated using adaptations of the previously described MethylScreen assay [26]. 1µg of genomic DNA was digested with: (1) a methylation-sensitive restriction enzyme (MSRE), (2) a methylation-dependent restriction enzyme (MDRE), (3) both MSRE and MDRE (double digest, DD), and (4) neither MSRE or MDRE (mock control). The MSREs used for CpGs 3, 6, 11 and 13 were *Aci*I (Fermentas), *Hpy*188III (New England Biolabs), *Aji*I (Fermentas) and *Eco*72I (Fermentas), respectively. The MDRE used for all four CpGs was *Mcr*Bc (Fermentas). A 50ng aliquot of digested DNA was then amplified by quantitative PCR using SYBR® Green (Applied Biosystems) and an ABI7900HT Fast Real-Time PCR System with the following primers: CpG3 F 5'-GAGACGTGGCTTTGTTTTCTG-3' and R 5'-GTTTCCTCCTTTCAAGCCGTG-3'; CpG6 F 5'-GAAGATGCCAAGGAAGTGGTAG-3' and R 5'-GAGCAACACAAATATGGCTTGG-3'; CpG11 and CpG13 F-met 5'-GAACCGTCTGGGCAAAGGCCAG-3' and R-met 5'-ATCCCAAAGTTTCTTCAAACACAATG-3'. PCR quantification was carried out using the ΔCt method (values were calculated as $2^{\Delta Ct \, (mock - digest)}$ with the mock value set at 100%) and RQ Manager software (Applied Biosystems). Each qPCR reaction for each sample was performed in triplicate in a 96-well plate.
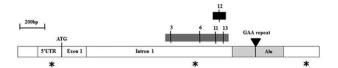
Figure 1. *FXN* region of analysis. Schematic diagram of 2.5kb of the *FXN* exon 1 and intron 1 region showing the GAA repeat within an Alu sequence (light grey box), together with the regions of 5mC/5hmC MethylScreen (dark grey box) and 5hmC EpiMarkTM (black box) analysis. The numbers represent the CpG sites, which are numbered according to [13]. Asterisks indicate the positions of the three potential CTCF binding sites that were analysed by EMSA.
doi: 10.1371/journal.pone.0074956.g001

MethylScreen DNA methylation values were then calculated as follows: Densely methylated (DM) = (MSRE-DD)/(100-DD) x 100; unmethylated (UM) = (MDRE-DD)/(100-DD) x 100; intermediately methylated (IM) = 100 - (DM+UM).

5hmC analysis

Levels of 5mC and 5hmC were determined at CpG site 12 (numbered according to [14]) of the *FXN* upstream GAA repeat region (Figure 1) using the EpiMark™ Analysis Kit (New England Biolabs). 1μg genomic DNA was glucosylated with T4 B-glucosyltransferase (T4-BGT), and then both glucosylated (G) and non-glucosylated (NG) DNA samples were digested with the methylation sensitive restriction enzyme, *Hpa*II, or the methylation insensitive restriction enzyme, *Msp*I. Digested DNA, together with non-digested (mock control) DNA samples were then amplified by quantitative PCR using SYBR® Green (Applied Biosystems) on an ABI7900HT Fast Real-Time PCR System with F-met and R-met PCR primers (the same as used for MethylScreen analysis). PCR quantification was carried out using the ΔCt method (values were calculated as $2^{\Delta Ct\,(mock+G\,-\,digest)}$ with the mock+G value set at 100%) and RQ Manager software (Applied Biosystems). Each qPCR reaction for each sample was performed in duplicate in a 96-well plate.

CTCF electrophoretic mobility shift assay (EMSA)

EMSA was performed with recombinant CTCF protein and three γ-^{32}P end-labelled *FXN* probes, together with CTCF-positive (DM1 (1)) and CTCF-negative (DM1 (3)) probes [27]. The probes were prepared by PCR amplification of human genomic DNA using the following primer pairs: *FXN* 5' UTR, F 5'-AAGCAGGCTCTCCATTTTTG-3', R 5'-CGAGAGTCCACATGCTGCT-3'; *FXN* GAA Upstream, F 5'-GAAACCCAAAGAATGGCTGTG-3', R 5'-TTCCCTCCTCGTGAAACACC-3'; *FXN* GAA Downstream, F 5'-TGGGTTGTCAGCAGAGTTGT-3', R 5'-CCGATAATCCCAGCTACTCG-3'; DM1 (1), F 5'-GCCTGCCAGTTCACAACC-3', R 5'-AGCAGCATTCCCGGCTAC-3'; DM1 (3), F 5'-AGCTTTCTTGTGCATGACG-3', R 5'-GGTTGTTGGGGGTCCTGTAG-3'. PCR products were gel purified using Geneclean III (BIO 101) and end-labelled with

γ-^{32}P ATP using T4 polynucleotide kinase. Labelled PCR products were put through a Microspin S-200 HR column (GE HealthCare) and then incubated with 0.2μg of CTCF full-length recombinant protein (Abnova) in PBS buffer containing 5mM MgCl$_2$, 0.1mM ZnSO$_4$, 1mM DTT, 0.1% Nonidet P-40, 10% glycerol and 1μg of double-strand competitor DNA poly(dI-dC) (Sigma). After incubation for 30 minutes at room temperature, samples were run on a 5% nondenaturing polyacrylamide gel in 0.5X TBE buffer at 100V for 1-2 hours. Gels were then vacuum dried and exposed to ECL Hyperfilm (GE HealthCare), followed by film development.

CTCF chromatin immunoprecipitation (ChIP)

Human cerebellum tissue samples (30mg from each of three FRDA and three unaffected controls, the same as used for DNA preparation) were homogenised, followed by cross-linking of DNA and protein by formaldehyde treatment. Cell nuclei were isolated by centrifugation and then lysed in the presence of protease inhibitors [15]. DNA was sheared by sonication, followed by immunoprecipitation with anti-CTCF antibody (Upstate 07-729). After reversal of cross-linking, qPCR amplification of the resultant co-immunoprecipitated DNA was carried out with SYBR® Green in an ABI7400 sequencer (Applied Biosystems) using *FXN* 5' UTR primers and *GAPDH* (negative for CTCF) control primers as previously described [25]. PCR quantification was carried out using the ΔΔCt method and RQ Manager software (Applied Biosystems). Each value of immunoprecipitated DNA was normalised to both input and *GAPDH* values. Each tissue sample was subjected to four independent ChIP procedures, followed by triplicate qPCR analysis.

Statistical analysis

Comparison of data from FRDA patients and unaffected control samples was carried out using the student's t test with a significance value set at $P<0.05$.

Results

MethylScreen analysis is a rapid method to quantify DNA methylation in the *FXN* upstream GAA repeat region

We, and others, have previously used bisulfite sequencing to quantify DNA methylation changes at the *FXN* upstream GAA repeat region in FRDA cells and tissues [14–17]. However, bisulfite sequencing is a somewhat time-consuming procedure, which can be subject to errors due to incomplete bisulfite treatment. Therefore, we have now adapted the previously described MethylScreen assay [26] to rapidly quantify the DNA methylation status at four CpG sites at the *FXN* upstream GAA repeat region, designated 3, 6, 11 and 13 according to [14] (Figure 1). The MethylScreen method does not require bisulfite treatment of DNA, but involves digestion of genomic DNA with a MSRE, a MDRE, and both a MSRE and a MDRE, followed by qPCR amplification of the residual non-digested DNA using an amplicon that spans the CpG site of interrogation. The output data is expressed as three levels of DNA methylation status:

DNA methylation - Cerebellum

Figure 2. DNA methylation levels in cerebellum. MethylScreen analysis of four CpG sites in the FXN upstream GAA repeat region of DNA from FRDA and control cerebellum tissues. UM = unmethylated, IM = intermediately methylated, DM = densely methylated. Error bars = s.e.m. n=4.

doi: 10.1371/journal.pone.0074956.g002
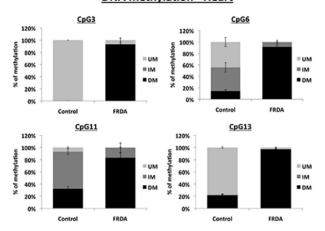
DNA methylation - Heart

Figure 3. DNA methylation levels in heart. MethylScreen analysis of four CpG sites in the FXN upstream GAA repeat region of DNA from FRDA and control heart tissues. UM = unmethylated, IM = intermediately methylated, DM = densely methylated. Error bars = s.e.m. n=4.

doi: 10.1371/journal.pone.0074956.g003

unmethylated (UM), intermediately methylated (IM) and densely methylated (DM). We validated the MethylScreen assay by investigating the levels of DNA methylation at CpG11 and CpG 13 in cerebellum and heart tissue from FRDA patients and unaffected controls, which allowed a direct comparison with our previously reported bisulfite sequence analysis [15]. In addition, we investigated CpG3 and CpG6 in cerebellum and heart tissue, since these sites together with CpG13 have previously been identified as the three sites of most significant DNA methylation increase in FRDA versus control lymphoblastoid cells [14]. Our results for cerebellum tissue reveal an increase in DNA methylation at all four CpG sites in FRDA tissue compared with unaffected control tissue, in agreement with all previous bisulfite sequencing data (Figure 2). However, the magnitude of this DNA methylation increase was found to vary from a non-statistically significant small increase in DM of 6% to 20% at CpG3, through an intermediate level of increase in DM of 4% to 68% at CpG6 ($P<0.001$), to high levels of increase in DM of 14% to 96% at both CpG11 and CpG13 (both $P<0.001$). The results for heart tissue also reveal significant increases in DNA methylation at all four CpG sites in FRDA tissue compared with unaffected control tissue, with DM values increasing from 0% to 93% at CpG3 ($P<0.001$), 15% to 92% at CpG6 ($P<0.001$), 32% to 83% at CpG11 ($P<0.01$) and 21% to 97% at CpG13 ($P<0.001$), again in agreement with all previous bisulfite sequencing data (Figure 3). Therefore, the specific MethylScreen protocols that we have developed can be considered as both rapid and robust procedures to investigate DNA methylation status at the FXN locus.

5hmC is the predominant constituent of DNA methylation in the *FXN* upstream GAA repeat region and is increased in FRDA cerebellum and heart tissues

Levels of 5mC and 5hmC were determined at CpG site 12 of the *FXN* upstream GAA repeat region (Figure 1) using the EpiMark™ Analysis Kit (New England Biolabs). This CpG site was chosen because it resides within a CCGG motif, which is required by the EpiMark™ procedure for differential restriction enzyme digestion by *Hpa*II and *Msp*I. The principle of this technique depends on the addition of glucose to the hydroxyl group of 5hmC via an enzymatic reaction utilising T4-BGT. Genomic DNA is glucosylated, then digested with a methylation sensitive restriction enzyme, *Hpa*II, and a methylation insensitive restriction enzyme, *Msp*I. If 5hmC is present in the context of CCGG, this modification converts a normally cleavable *Msp*I site to a non-cleavable one.
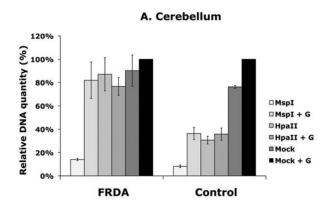
Initial determination of DNA methylation levels (combined 5mC and 5hmC) by observing *Hpa*II + G and *Hpa*II qPCR values at CpG12 in cerebellum tissue reveals a significant increase in DNA methylation in FRDA tissue compared with unaffected control tissue (Figure 4A), at levels similar to those previously identified by bisulfite sequencing. Thus, DNA methylation levels increase from either 36% to 77% (*Hpa*II + G, $P<0.01$) or 30% to 87% (*Hpa*II, $P<0.05$) compared with previously reported bisulfite sequence values of 40% to 95% [15]. However, more detailed analysis reveals that there is no difference in the *Msp*I + G and *Hpa*II + G qPCR results from unaffected cerebellum tissues, and similarly there is no difference in the *Msp*I + G and *Hpa*II + G qPCR results from FRDA cerebellum tissues (Figure 4A). This indicates that all methylated DNA in both the unaffected and the FRDA cerebellum tissue comprises 5hmC rather than 5mC. Therefore, we can now conclude that a characteristic feature of FRDA cerebellum tissue is an increase in 5hmC-modified DNA

in the *FXN* upstream GAA repeat region, at least at CpG site 12. Determination of DNA methylation levels (combined 5mC and 5hmC) in heart tissue also reveals a significant increase in DNA methylation in FRDA tissue compared with unaffected control tissue (Figure 4B), at levels similar to those previously identified by bisulfite sequencing. Thus, DNA methylation levels increase from either 44% to 100% (*Hpa*II + G, *P*<0.01) or 37% to 89% (*Hpa*II, *P*<0.001) compared with previously reported bisulfite sequence values of 65% to 95% [15]. However, more detailed analysis of the *Msp*I + G compared with *Hpa*II + G qPCR results indicates that the methylated DNA comprises approximately 64% 5hmC in the unaffected heart tissues and only 37% 5hmC in the FRDA heart tissues, although there is still an overall increase in 5hmC levels (*Msp*I + G values) in FRDA compared with controls (Figure 4B). This suggests that, while there is increased DNA methylation in both cerebellum and heart tissues, there are also tissue specific differences: DNA methylation is entirely due to 5hmC in the cerebellum, but only partially due to 5hmC in the heart.

CTCF occupancy at the *FXN* 5' UTR region is decreased in FRDA cerebellum tissue

We firstly performed *in silico* screening of a 40kb region spanning the *FXN* gene for potential CTCF binding sites using the bioinformatics web tool designed by Klenova and colleagues (http://www.essex.ac.uk/bs/molonc/spa.htm) and we identified three regions of interest at the start of the *FXN* gene. We then investigated the three potential CTCF binding regions, which were located within the 5' UTR, GAA upstream and GAA downstream regions of the *FXN* locus (Fig 1), by EMSA using recombinant full length CTCF protein. Only the 5' UTR probe proved capable of binding CTCF protein to cause a gel shift (Figure 5A). This result is in agreement with a previous report that identified CTCF binding in the *FXN* 5' UTR region using fibroblast cell lysates [25]. We then performed ChIP analysis to assess the level of CTCF binding at the 5' UTR site in three FRDA cerebellum tissue lysates compared with three unaffected controls, using the same qPCR amplicons as previously described [25]. A statistically significant decrease (*P*<0.05) in CTCF binding was detected in FRDA cerebellum, with levels falling to only 20% of the levels detected in unaffected cerebellum controls (Figure 5B).

Discussion

We have investigated two molecular mechanisms that may be involved in the disease process in FRDA patient cerebellum and heart tissues. Firstly, we have developed a rapid and robust MethylScreen method to determine the general DNA methylation status (5mC and 5hmC) at the *FXN* upstream GAA repeat region. The level of DNA methylation in FRDA cerebellum tissue showed a gradation across the four CpG sites investigated, progressively increasing as you move closer to the GAA repeat locus. This suggests that the GAA repeat expansion mutation may act as focal point to induce only a localised effect of DNA methylation, spreading upstream from this point. However, this gradation effect was not seen in FRDA heart tissue, which demonstrated high levels of DNA
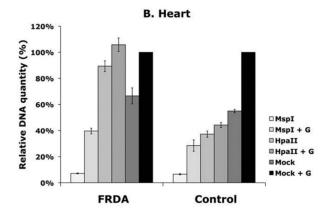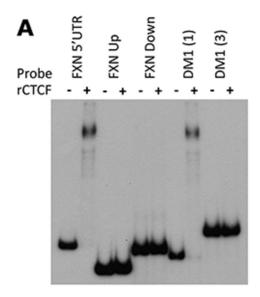

Figure 4. 5hmC levels in cerebellum and heart. EpiMark™ 5hmC and 5mC analysis of CpG site 12 in the FXN upstream GAA repeat region of DNA from FRDA and control cerebellum (**A**) and heart (**B**) tissues. G = glucosylated DNA. MspI values represent complete digestion control levels of DNA; MspI + G represents 5hmC levels, HpaII and HpaII + G represent combined 5mC and 5hmC levels; Mock and Mock +G represent undigested control levels of DNA. Error bars = s.e.m. n=4.

doi: 10.1371/journal.pone.0074956.g004

methylation at all 4 CpG sites, indicating a degree of variability in the DNA methylation process in different FRDA tissues. Such variability may have some bearing on the previously identified differences in GAA repeat somatic instability, with cerebellum showing high levels of instability, while heart is comparatively stable [28].

Subsequently, we have shown that the methylated DNA at one CpG site in the *FXN* upstream GAA region comprises entirely 5hmC in FRDA cerebellum, but only partially 5hmC in FRDA heart. Our results are in agreement with the recent report that 5hmC is particularly associated with human cerebellum throughout development [29]. However, we can discount the possibility that the increased 5hmC that we have observed in FRDA tissues is a non-specific age-related effect in neuronal tissues, as previously reported for mouse hippocampus and cerebellum [23,30], because our FRDA patients are actually younger than the unaffected controls.
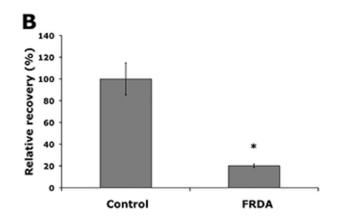
Figure 5. CTCF analysis. (**A**) EMSA showing a mobility shift for the *FXN* 5' UTR probe and CTCF-positive DM1 (1) probe [35], while there is no such shift detected for the FXN upstream GAA (Up) probe, *FXN* downstream GAA (Down) probe or CTCF-negative DM1 (3) probe [35]. (**B**) ChIP analysis showing the relative CTCF occupancy in the *FXN* 5' UTR of DNA from FRDA and control cerebellum tissues. Error bars = s.e.m. n=3. * = P<0.05.

doi: 10.1371/journal.pone.0074956.g005

Further studies will be required to determine if our finding is a general increased 5hmC effect for CpG sites within the *FXN* upstream GAA region, or indeed within other regions of the genome. The high levels of 5hmC that we have detected at the *FXN* upstream GAA region could be explained by either a passive or an active process. Thus, 5hmC could merely be a non-functioning intermediate in a demethylation process that is attempting to reverse the GAA repeat expansion-induced 5mC DNA methylation that is associated with *FXN* gene silencing [17]. In this case, the differences that we have detected between 5hmC levels in cerebellum and heart tissues could perhaps be due to tissue-selective differences in the efficiency

of the TET enzymes that convert 5mC to 5hmC [18,19] or the DNA repair systems (e.g. base excision repair) that are required to convert 5hmC to cytosine [22]. Alternatively, 5hmC may have an as yet unidentified role in regulation of gene expression. For example, it has shown *in vitro* that 5hmC at the promoter but not in the gene body negatively regulates gene expression [31]. Furthermore, it has been proposed that 5hmC may be involved in the regulation of genes that are silent but transcriptionally poised for activation, where there is loss of 5mC, recruitment of Polycomb repressive complex 2 (PRC2) and, therefore, methylation of histone 3 at lysine 27 (H3K27me3) [20]. Indeed, 5hmC and H3K27me3 have been shown to have high correlation across a variety of somatic tissues [32]. Therefore, it is interesting to note that the *FXN* gene demonstrates RNA polymerase II pausing and FRDA lymphoblastoid cells have shown increased levels of H3K27me3 in the *FXN* upstream GAA repeat region [33,34]. In addition, it has been proposed that 5hmC may inhibit DNMT1 binding and may therefore function as a negative regulator of 5mC levels at protein-DNA interaction sites, allowing the binding of enhancer proteins or transcription factors [35]. Furthermore, 5hmC has been reported to bind MECP2, particularly in the CNS, although the functional consequences of such an interaction are not yet known [36]. Although the close positioning of the GAA repeat expansion mutation and the increased 5hmC modification of DNA in the *FXN* upstream GAA repeat region in FRDA would suggest a likely causal relationship, there may also be other more general factors involved in the formation of 5hmC. For example, the TET enzymes that produce 5hmC are known to be dependent on Fe(II) and α-ketoglutarate. Therefore, disruptions of iron metabolism, oxidative stress and the TCA cycle, processes which are highly relevant to FRDA, may all potentially affect 5hmC levels [37].

The second molecular change that we have identified in FRDA cerebellum is reduced CTCF occupancy at the *FXN* 5' UTR region, confirming previous findings in FRDA fibroblast cells [25]. It has been hypothesized that decreased binding of CTCF in FRDA cells may be an epigenetic switch that allows increased frataxin antisense (*FAST-1*) transcription, spread of heterochromatin and resultant *FXN* gene silencing. DNA methylation has not been thought to play an active role in this process, because increased DNA methylation has not been detected at the CTCF binding site in the *FXN* 5' UTR region of FRDA cells or tissues [15,16,25], and the increased DNA methylation at the *FXN* upstream GAA repeat region has been considered to be an unconnected marker of the heterochromatin state. However, if there is a demethylation process occurring at the *FXN* upstream GAA repeat region, as indicated by the high levels of 5hmC that we have now identified, then this may indeed be involved in increasing *FAST-1* expression, CTCF depletion and hence *FXN* gene silencing. In conclusion, we can propose a model for FRDA that combines both the increased 5hmC modification of DNA and decreased CTCF binding at the *FXN* locus. Further studies will now be required to determine the exact relationship between these molecular mechanisms and FRDA pathogenesis.

Acknowledgements

Author Contributions

Conceived and designed the experiments: SA RMP MAP. Performed the experiments: SA CS RMP MAP. Analyzed the data: SA CS RMP MAP. Wrote the manuscript: SA CS RMP MAP.

References

1. Campuzano V, Montermini L, Moltò MD, Pianese L, Cossée M et al. (1996) Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. Science 271: 1423-1427. doi: 10.1126/science.271.5254.1423. PubMed: 8596916.
2. Pandolfo M (2009) Friedreich ataxia: the clinical picture. J Neurol 256 Suppl 1: 3-8. doi:10.1007/s00415-009-0097-x. PubMed: 19283344.
3. Sharma R, De Biase I, Gómez M, Delatycki MB, Ashizawa T et al. (2004) Friedreich ataxia in carriers of unstable borderline GAA triplet-repeat alleles. Ann Neurol 56: 898-901. doi:10.1002/ana.20333. PubMed: 15562408.
4. Dürr A, Cossee M, Agid Y, Campuzano V, Mignard C et al. (1996) Clinical and genetic abnormalities in patients with Friedreich's ataxia. N Engl J Med 335: 1169-1175. doi:10.1056/NEJM199610173351601. PubMed: 8815938.
5. Filla A, De Michele G, Cavalcanti F, Pianese L, Monticelli A et al. (1996) The relationship between trinucleotide (GAA) repeat length and clinical features in Friedreich ataxia. Am J Hum Genet 59: 554-560. PubMed: 8751856.
6. Campuzano V, Montermini L, Lutz Y, Cova L, Hindelang C et al. (1997) Frataxin is reduced in Friedreich ataxia patients and is associated with mitochondrial membranes. Hum Mol Genet 6: 1771-1780. doi: 10.1093/hmg/6.11.1771. PubMed: 9302253.
7. Bradley JL, Blake JC, Chamberlain S, Thomas PK, Cooper JM et al. (2000) Clinical, biochemical and molecular genetic correlations in Friedreich's ataxia. Hum Mol Genet 9: 275-282. doi:10.1093/hmg/9.2.275. PubMed: 10607838.
8. Koeppen AH, Michael SC, Knutson MD, Haile DJ, Qian J et al. (2007) The dentate nucleus in Friedreich's ataxia: the role of iron-responsive proteins. Acta Neuropathol 114: 163-173. doi:10.1007/s00401-007-0220-y. PubMed: 17443334.
9. Wong A, Yang J, Cavadini P, Gellera C, Lonnerdal B et al. (1999) The Friedreich's ataxia mutation confers cellular sensitivity to oxidant stress which is rescued by chelators of iron and calcium and inhibitors of apoptosis. Hum Mol Genet 8: 425-430. doi:10.1093/hmg/8.3.425. PubMed: 9949201.
10. Grabczyk E, Mancuso M, Sammarco MC (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. doi:10.1093/nar/gkm589. PubMed: 17693431.
11. Wells RD (2008) DNA triplexes and Friedreich ataxia. FASEB J 22: 1625-1634. doi:10.1096/fj.07-097857. PubMed: 18211957.
12. Saveliev A, Everett C, Sharpe T, Webster Z, Festenstein R (2003) DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing. Nature 422: 909-913. doi:10.1038/nature01596. PubMed: 12712207.
13. Sandi C, Al-Mahdawi S, Pook MA (2013) Epigenetics in Friedreich's Ataxia: Challenges and Opportunities for Therapy. Genet Res Int, 2013: 852080. PubMed: 23533785
14. Greene E, Mahishi L, Entezam A, Kumari D, Usdin K (2007) Repeat-induced epigenetic changes in intron 1 of the frataxin gene and its consequences in Friedreich ataxia. Nucleic Acids Res 35: 3383-3390. doi:10.1093/nar/gkm271. PubMed: 17478498.
15. Al-Mahdawi S, Mouro Pinto R, Ismail O, Varshney D, Lymperi S et al. (2008) The Friedreich ataxia GAA repeat expansion mutation induces comparable epigenetic changes in human and transgenic mouse brain and heart tissues. Hum Mol Genet 17: 735-746. PubMed: 18045775.
16. Evans-Galea MV, Carrodus N, Rowley SM, Corben LA, Tai G et al. (2012) FXN methylation predicts expression and clinical outcome in Friedreich ataxia. Ann Neurol 71: 487-497. doi:10.1002/ana.22671. PubMed: 22522441.
17. Castaldo I, Pinelli M, Monticelli A, Acquaviva F, Giacchetti M et al. (2008) DNA methylation in intron 1 of the frataxin gene is related to GAA repeat length and age of onset in Friedreich's ataxia patients. J Med Genet 45: 808-812. doi:10.1136/jmg.2008.058594. PubMed: 18697824.
18. Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324: 929-930. doi:10.1126/science.1169786. PubMed: 19372393.
19. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H et al. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324: 930-935. doi: 10.1126/science.1170116. PubMed: 19372391.
20. Branco MR, Ficz G, Reik W (2011) Uncovering the role of 5-hydroxymethylcytosine in the epigenome. Nat Rev Genet 13: 7-13. doi: 10.1038/ni.2192. PubMed: 22083101.
21. Globisch D, Münzel M, Müller M, Michalakis S, Wagner M et al. (2010) Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. PLOS ONE 5: e15367. doi:10.1371/journal.pone.0015367. PubMed: 21203455.
22. Guo JU, Su Y, Zhong C, Ming GL, Song H (2011) Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell 145: 423-434. doi:10.1016/j.cell.2011.03.022. PubMed: 21496894.
23. Szulwach KE, Li X, Li Y, Song CX, Wu H et al. (2011) 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. Nat Neurosci 14: 1607-1616. doi:10.1038/nn.2959. PubMed: 22037496.
24. Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR et al. (2010) The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. PLOS ONE 5: e8888. doi:10.1371/journal.pone.0008888. PubMed: 20126651.
25. De Biase I, Chutake YK, Rindler PM, Bidichandani SI (2009) Epigenetic silencing in Friedreich ataxia is associated with depletion of CTCF (CCCTC-binding factor) and antisense transcription. PLOS ONE 4: e7914. doi:10.1371/journal.pone.0007914. PubMed: 19956589.
26. Holemon H, Korshunova Y, Ordway JM, Bedell JA, Citek RW et al. (2007) MethylScreen: DNA methylation density monitoring using quantitative PCR. BioTechniques 43: 683-693. doi:10.2144/000112597. PubMed: 18072598.
27. Filippova GN, Thienes CP, Penn BH, Cho DH, Hu YJ et al. (2001) CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. Nat Genet 28: 335-343. doi: 10.1038/ng570 PubMed: 11479593
28. De Biase I, Rasmussen A, Endres D, Al-Mahdawi S, Monticelli A et al. (2007) Progressive GAA expansions in dorsal root ganglia of Friedreich's ataxia patients. Ann Neurol 61: 55-60. doi:10.1002/ana.21052. PubMed: 17262846.
29. Wang T, Pan Q, Lin L, Szulwach KE, Song CX et al. (2012) Genome-wide DNA hydroxymethylation changes are associated with neurodevelopmental genes in the developing human cerebellum. Hum Mol Genet 21: 5500-5510. doi:10.1093/hmg/dds394. PubMed: 23042784.
30. Chen H, Dzitoyeva S, Manev H (2012) Effect of aging on 5-hydroxymethylcytosine in the mouse hippocampus. Restor Neurol Neurosci 30: 237-245. doi:10.3233/RNN-2012-110223. PubMed: 22426040.
31. Robertson J, Robertson AB, Klungland A (2011) The presence of 5-hydroxymethylcytosine at the gene promoter and not in the gene body negatively regulates gene expression. Biochem Biophys Res Commun 411: 40-43. doi:10.1016/j.bbrc.2011.06.077. PubMed: 21703242.
32. Haffner MC, Pellakuru LG, Ghosh S, Lotan TL, Nelson WG et al. (2013) Tight correlation of 5-hydroxymethylcytosine and Polycomb marks in health and disease. Cell Cycle 12: 1835-1841. doi:10.4161/cc.25010. PubMed: 23676216.
33. Chan PK, Torres R, Yandim C, Law PP, Khadayate S et al. (2013) Heterochromatinization induced by GAA-repeat hyperexpansion in Friedreich's ataxia can be reduced upon HDAC inhibition by vitamin B3. Hum Mol Genet 22: 2662-2675. doi:10.1093/hmg/ddt115. PubMed: 23474817.
34. Kim E, Napierala M, Dent SY (2011) Hyperexpansion of GAA repeats affects post-initiation steps of FXN transcription in Friedreich's ataxia. Nucleic Acids Res 39: 8366-8377. doi:10.1093/nar/gkr542. PubMed: 21745819.
35. Stroud H, Feng S, Morey Kinney S, Pradhan S, Jacobsen SE (2011) 5-Hydroxymethylcytosine is associated with enhancers and gene bodies

in human embryonic stem cells. Genome Biol 12: R54. doi:10.1186/gb-2011-12-6-r54. PubMed: 21689397.

36. Mellén M, Ayata P, Dewell S, Kriaucionis S, Heintz N (2012) MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell 151: 1417-1430. doi:10.1016/j.cell.2012.11.022. PubMed: 23260135.

37. Chia N, Wang L, Lu X, Senut MC, Brenner C et al. (2011) Hypothesis: environmental regulation of 5-hydroxymethylcytosine by oxidative stress. Epigenetics 6: 853-856. doi:10.4161/epi.6.7.16461. PubMed: 21617369.