

The Role of AGG Interruptions in the Transcription of *FMR1* Premutation Alleles

Carolyn M. Yrigollen¹, Federica Tassone¹, Blythe Durbin-Johnson², Flora Tassone^{1,3*}

1 Department of Biochemistry and Molecular Medicine, School of Medicine, University of California Davis, Davis, California, United States of America, **2** Department of Public Health Sciences School of Medicine, University of California Davis, Davis, California, United States of America, **3** M.I.N.D. Institute, University of California Davis Medical Center, Davis, California, United States of America

Abstract

Fragile X associated disorders are caused by a premutation allele in the fragile X mental retardation 1 gene (*FMR1*) and are hypothesized to result from the toxic effect of elevated levels of expanded *FMR1* transcripts. Increased levels of *FMR1* mRNA have indeed been reported in premutation carriers; however the mechanism by which expanded alleles lead to elevated levels of *FMR1* mRNA in premutation carriers is unknown. Within the CGG repeat tract AGG interruptions are found, generally 1–3 present in normal/intermediate alleles (6–54 CGG repeats) and usually 0–1 in premutation alleles (55–200 CGG repeats). They are present at specific locations, generally occurring after 9 or 10 uninterrupted CGG repeats [(CGG)₉AGG(CGG)₉AGG(CGG)_n]. We evaluated both the number of AGG interruptions and the resulting length of the uninterrupted 3' CGG repeat pure tract in premutation alleles derived from two large cohorts of male and female carriers to determine whether the presence of AGG interruptions or the length of a pure stretch of CGG repeats influence the levels of *FMR1* mRNA in blood. Our findings indicate that neither the number of AGG interruptions, nor their position along the CGG tract have a significant affect on mRNA levels in premutation carriers. We also, as expected based on previous findings, observed a highly significant correlation between CGG repeat number (as both total length and length of pure CGG stretch) and *FMR1* mRNA expression levels, in both males and females. Importantly, we did not observe any significant difference in *FMR1* mRNA levels in premutation carriers based on age.

Citation: Yrigollen CM, Tassone F, Durbin-Johnson B, Tassone F (2011) The Role of AGG Interruptions in the Transcription of *FMR1* Premutation Alleles. PLoS ONE 6(7): e21728. doi:10.1371/journal.pone.0021728

Editor: Michael Nicholas Weedon, Peninsula College of Medicine and Dentistry, University of Exeter, United Kingdom

Received: April 21, 2011; **Accepted:** June 7, 2011; **Published:** July 19, 2011

Copyright: © 2011 Yrigollen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the National Institutes of Health through individual research awards [HD02274, HD04066, HD36071]. Statistical support for this publication was made possible by the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), and NIH Roadmap for Medical Research [UL1 RR024146]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Flora Tassone is a non-paid collaborator with Asuragen, Inc. She has a patent application for a method for the detection of *FMR1* allele size and category. This does not alter the authors' adherence to all the PLoS ONE policies on sharing data and materials.

* E-mail: *ftassone@ucdavis.edu

Introduction

Expansion of the CGG trinucleotide repeat in the 5' untranslated (UTR) region of the *FMR1* gene is causative for a number of disorders including fragile X syndrome (FXS), fragile X associated tremor ataxia syndrome (FXTAS) and *FMR1* related primary ovarian insufficiency (FXPOI) [1]. A CGG expansion in the *FMR1* gene has also been associated with an increased incidence of several other clinical involvements ranging from anxiety to autism spectrum disorder (ASD) [2]. The length of the repeat falls into several categories, defined by The American College of Medical Genetics [3], which are useful as molecular indicators and based on predicting the risk of expansion in the next generations. Individuals with an allele within the normal range (5–44 CGG repeats) are considered unaffected, while full mutation alleles (>200 CGG repeats) result in transcriptional silencing of the *FMR1* gene, absence of fragile X mental retardation 1 protein (FMRP) and fragile X syndrome. Intermediate (45–54 CGG repeats) and premutation alleles (55–200 CGG repeats) were originally characterized as having an increased risk of expansion in future generations, but having no phenotypic effect on the carrier of these alleles. Within the last decade, biological effects have been

identified and studied [4,5,6], but more needs to be learned about the mechanism that leads to the clinical involvement in these transitional ranges.

Increased levels of *FMR1* mRNA correlating with the length of the repeat tract, have been reported in both male and female carriers of a premutation allele [7,8,9]. Although the elevated levels result from an increase in transcriptional activity [10] a concomitant decreased protein expression, due to a deficit in translational efficiency especially in longer expanded alleles, was also reported [7,8,9,11,12]. Recently, increased FMRP expression levels have been reported in premutation carriers with an allele within the 80–89 CGG repeats range compared to those in the lower and in the higher premutation range [12].

Within the CGG repeat tract are interspersions of AGG segments that occur in normal, intermediate, and premutation range alleles and commonly occur after 9 or 10 uninterrupted CGG repeats [(CGG)₉AGG(CGG)₉AGG(CGG)_n]. These AGG interruptions can be stably inherited, such that the number of AGG interruptions and their position is likely to correspond in parent and offspring; however like the length of CGG repeat tracts in *FMR1*, there is a level of instability, such as the loss of an interruption during transmission [13].

Moreover, AGG interruptions have been predicted to increase the stability of the repeat tract during maternal transmission by decreasing the size of pure CGG repeats in the tract [14,15]. AGG interruptions have been shown to affect both DNA and RNA structure at the *FMRI* locus. The DNA at a pure CGG repeat tract of *FMRI* has been shown to adopt a non- β conformation, which is increasingly unstable in the presence of AGG interruptions and leads to the formation of bulges and larger loops [16]. The CGG repeat region is known to form secondary structures in DNA, such as hairpins and tetraplexes, and in RNA has also been shown to form hairpins, though tetraplexes were not observed [16].

In recent studies, the presence of AGG interruptions within the *FMRI* trinucleotide repeat has been investigated to determine if transcription or translation is affected [10,12,17]. Tassone et al. [10] looked at the effect of AGG interruptions on *FMRI* transcription levels and reported a lack of association between *FMRI* mRNA levels and the number of AGG interruptions in 147 subjects including 76 subjects with normal range alleles, 10 subjects with an intermediate allele and 59 subjects with a premutation allele. More recently, Peprah et al. [12] looked at the transcriptional and translational activity of *FMRI* and AGG interruptions in 74 premutation carriers and found no correlations, confirming previous findings from an *in vitro* system [17].

Here, we describe our results from an in-depth investigation of the influence of the number of AGG interruptions, their location within the CGG tract and the length of the 3' CGG repeat pure tract on the level of *FMRI* mRNA in a cohort of 770 premutation carriers, including 288 males and 482 females. This study adds to the body of knowledge with an increased sample size for male participants and confirms the same findings in a cohort of female participants, which have not been previously studied. Additionally, using a new triple primer PCR, we have been able to precisely map AGG interruptions along the CGG repeat tract and to correlate their location with the *FMRI* mRNA transcriptional level. This new approach allowed us to analyze both the total length of the repeat region and the length of uninterrupted CGG repeats (a 3' CGG repeat pure stretch). Previous analysis of the role AGG interruptions have on transcription may have been unable to detect a significant correlation because the analysis was based on the presence or the absence of AGG interruptions, but not on the number of uninterrupted CGG repeats.

Results

DNA samples from 770 unrelated individuals comprising 288 males and 482 females were selected for a detailed molecular analysis including CGG repeat number, length of pure CGG stretch, activation ratio in females, AGG number and distribution, and *FMRI* mRNA expression levels. Males and females with the premutation were selected from fragile X pedigrees because they were known to carry the premutation (and confirmed by *FMRI* DNA testing). None were selected on the basis of clinical symptoms.

Of the 770 samples, 265 male and 435 female samples had measured mRNA transcript levels and 231 male and 397 female samples had information about the age of the participant at the time the DNA and RNA samples were collected in males and females respectively.

The mean age in male subjects at the time of sample collection was 45.8 years (± 26.3), and the mean age of female subjects was 42.6 years (± 16.8). Most of the female premutation carriers were identified consequent of having a child with FXS and male

premutation carriers were part of projects focused on fragile X associated disorders (distributions included in Figure S1a,b).

The mean *FMRI* mRNA level in males (3.1 ± 1.2) was significantly different from the mean in the females (2.4 ± 0.8) (t test $P < 0.001$) due to the contribution of the normal allele on the other X chromosome (Table 1; distributions are presented in Supplementary Material (Figure S2a,b). CGG repeat allele sizes were comparable between males (mean 93.1 ± 25.8) and females (89.2 ± 20.7) (t test $P = 0.035$) (Figure S3a,b). The distribution of Activation ratio measured in 465 females had a mean of 0.54 and followed a normal distribution (Figure 1). Detailed molecular measures are presented in Table 1.

Of the 288 male samples, 23 had missing data on *FMRI* mRNA levels, 6 were missing the number of AGG interruptions and length of the longest pure CGG stretch, and 57 samples were missing participant age at time of sample collection. Missing observations were excluded from analyses as needed. The *FMRI* mRNA level from one additional sample was excluded from analyses due to an excessively large standard error of replicate measurements.

Of the 482 female subjects with the *FMRI* premutation allele included in the study, 47 samples had no data on *FMRI* mRNA levels, 17 samples did not have the activation ratio, and 85 samples were missing data on the participant age at time of sample collection. Missing observations were excluded from analyses as needed. The *FMRI* mRNA levels in females are determined by two alleles. In this study one allele was within the premutation range and the other was within the normal range (with the exception of 2 subjects, one with 2 premutation alleles and one with an intermediate and a premutation allele). Adjusted *FMRI* mRNA levels (estimates of $mRNA_{hp}$, the contribution to the observed mRNA level from expanded alleles) were derived from the observed *FMRI* mRNA level and the activation ratios using the method described in Tassone et al. [9].

FMRI mRNA levels correlate with both total CGG repeat length and the length of 3' pure CGG stretch in male and female carriers

In order to assess the relationship between *FMRI* mRNA levels and total CGG length, a linear regression model, with *FMRI* mRNA level as the response variable and total CGG length (as a

Table 1. A comparison of age, CGG repeat size, and *FMRI* mRNA levels in male and female premutation carriers.

	Male	Female
Samples (N)	288	482
Mean age	48.8	42.6
Age range	1–89	1–95
Mean Total CGG length (standard deviation)	93.1 (25.8)	89.2 (20.7)
Total CGG size range	55–220	55–190
Mean 3' pure CGG stretch	84.9	82.6
3' Pure CGG range	34–220	35–190
Number of AGG interruptions range	0–2	0–3
Mean <i>FMRI</i> mRNA levels	3.1	2.4
<i>FMRI</i> mRNA levels	1.1–9.0	0.8–6.4
Mean activation ratio		0.54
Activation ratio range		0.08–0.96

doi:10.1371/journal.pone.0021728.t001

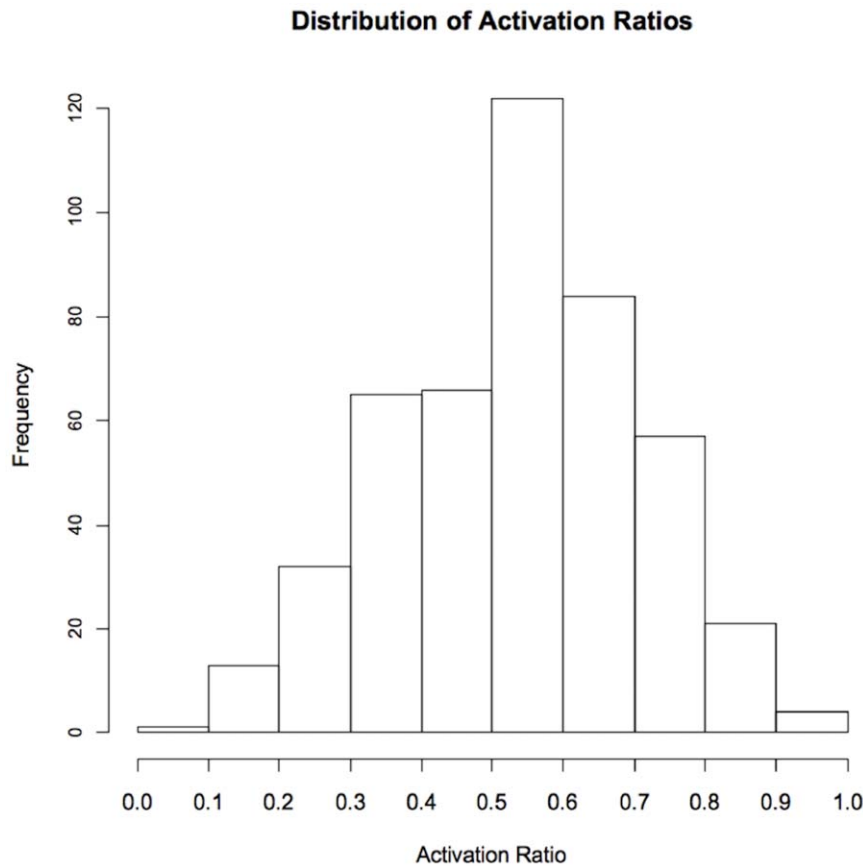


Figure 1. Distribution of activation ratio. Activation ratio was determined by Southern blot analysis. The distribution approximates a normal distribution with a mean of 0.54.
doi:10.1371/journal.pone.0021728.g001

continuous variable) as the predictor, was used. This and subsequent analysis was conducted separately on data from males and females. Based on the linear regression model fit using total CGG length as a continuous variable, *FMRI* mRNA levels appear to be strongly related to total CGG length ($P < 0.001$) in both males (Figure 2a) and females (Figure 2b) confirming previous reports [7,8,9,10,12].

In order to assess the relationship between *FMRI* mRNA levels and length of 3' pure CGG stretch the same analyses were conducted. Using a linear regression model, with *FMRI* mRNA level as the response variable and length of pure CGG stretch (as a continuous variable) as the predictor, *FMRI* mRNA levels appear to be strongly related to the length of 3' pure CGG stretch ($P < 0.001$) in both males (Figure 3a) and females (Figure 3b). *FMRI* mRNA levels were adjusted for activation ratio. Results were confirmed using an ANOVA model, with *FMRI* mRNA level as the response variable and length of pure CGG stretch as the predictor.

The number of AGG Interruptions correlate with total CGG Length

The relationship between total CGG length and number of AGG interruptions was assessed using an ANOVA model, with total CGG length as the response variable and number of AGG interruptions as the predictor. The test for an AGG effect was highly significant ($P < 0.001$) in both genders, rejecting the null hypothesis that mean total CGG length is equal in each AGG

group. Tukey HSD post-hoc tests were conducted in order to test for pairwise differences between each group. In males, the total CGG length differed significantly between subjects with 0 interruptions and those with 1 or 2 interruptions ($P < 0.001$ in both cases). However, the data did not suggest a difference in total CGG length between subjects with 1 AGG interruption and those with 2 AGG interruptions ($P = 0.940$). A box plot of total CGG length for each number of AGG interruptions is shown in Figure 4a. In females, the total CGG length differed significantly between subjects with 0 and 1 interruption ($P = 0.001$) and between those with 0 and 2 interruptions ($p < 0.001$). However, the data did not suggest a difference in total CGG length between the remaining pairs of groups (note that this lack of significance should be interpreted cautiously for comparisons between patients with 3 interruptions and those with other numbers of interruptions due to the small sample size). A box plot of total CGG length for each number of AGG interruptions is shown in Figure 4b. Finally, no difference was observed in the interspersion pattern of AGG interruptions within the premutation alleles as already reported [13].

FMRI mRNA Levels do not correlate with the number of AGG interruptions

The relationship between *FMRI* mRNA levels and number of AGG interruptions was assessed using an ANOVA model, with *FMRI* mRNA level as the response variable and number of AGG interruptions as the predictor. The global test for an AGG effect

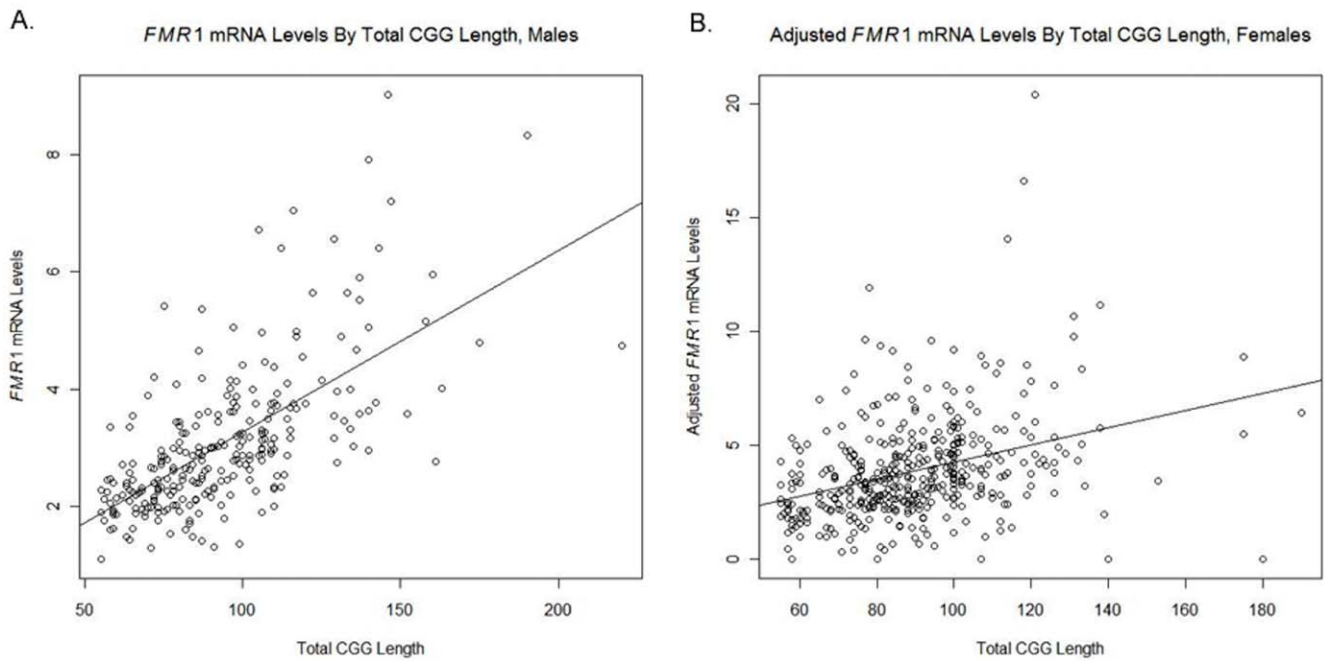


Figure 2. Scatterplot of *FMR1* mRNA levels by total CGG length. A linear regression model with model fit using CGG length as a continuous variable a) in premutation males and b) in premutation females. *FMR1* mRNA levels show strong ($P < 0.001$) correlation to total CGG length. doi:10.1371/journal.pone.0021728.g002

did not suggest differences in *FMR1* mRNA levels between any of the AGG groups ($P = 0.060$ in males and $P = 0.066$ in females) (Figure 5a,b). These findings confirmed previous reports [10].

Given that the total CGG length, which is highly predictive of *FMR1* mRNA levels, differs significantly between subjects with different number of AGG interruptions, further analyses were

conducted in which a separate linear regression of *FMR1* mRNA levels on total CGG length was fitted for each number of AGG interruptions. Differences in the intercepts of these lines would suggest differences in *FMR1* mRNA levels between subjects with different numbers of AGG interruptions but comparable total CGG length, whereas differences in the slopes of these lines would

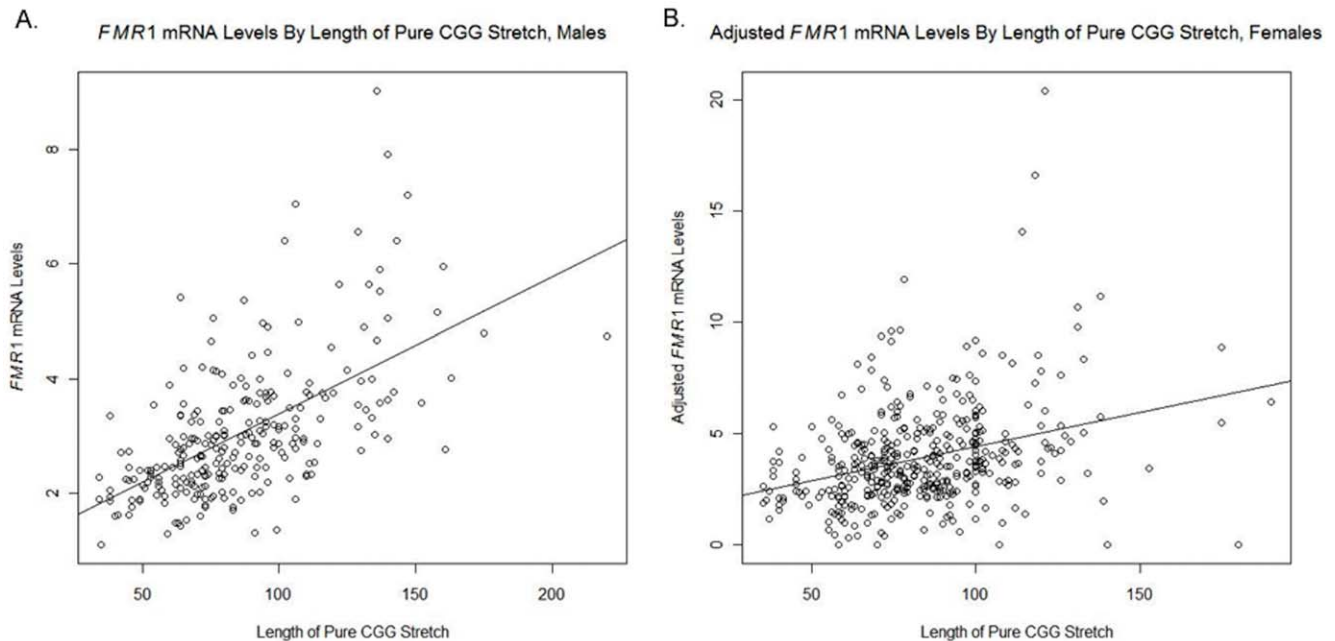


Figure 3. Scatterplot of *FMR1* mRNA levels by length of CGG pure stretch. A linear regression model with model fit using CGG pure stretch as a continuous variable a) in premutation males and b) in premutation females. *FMR1* mRNA levels appear strongly related to pure CGG stretch ($P < 0.001$). doi:10.1371/journal.pone.0021728.g003

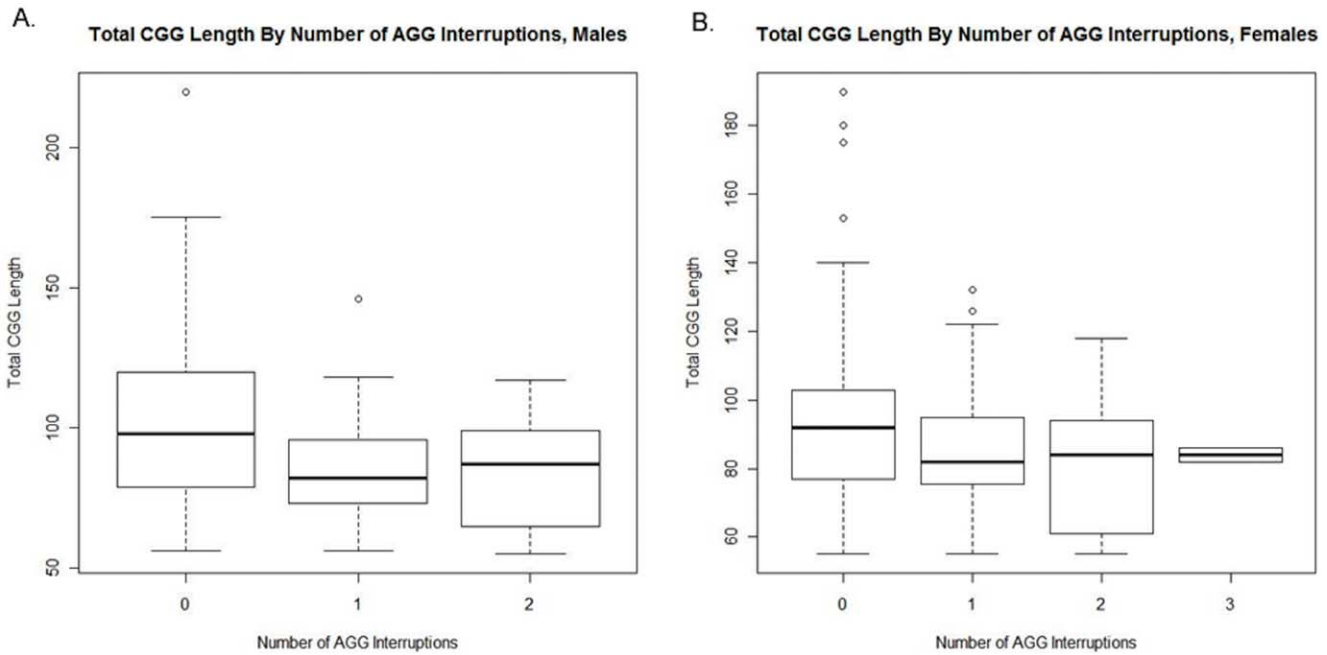


Figure 4. A box plot of total CGG length for each number of AGG interruptions. A) In premutation male data, significant difference in total CGG length of subjects with no AGG interruptions (130 subjects) and those with 1 or 2 interruptions (96 and 55 subjects respectively) was found ($P < 0.001$ in both cases). No significance was found in total CGG length of subjects with 1 AGG interruption compared to 2 AGG interruptions ($P = 0.935$). B) Premutation female data showed significant difference in total CGG length of subjects with no AGG interruptions (259 subjects) and those with 1 or 2 interruptions (143 and 78 subjects respectively) ($P < 0.001$) but no difference with remaining pairs. doi:10.1371/journal.pone.0021728.g004

suggest that the impact of total CGG length on *FMRI* mRNA levels differs based on the number of AGG interruptions. In male subjects, the slope of the regression line fitted to data from subjects with 0 AGG interruptions showed a marginally significant

difference from the slope of the regression line fitted to data from subjects with 1 interruption ($P = 0.046$); no other significant differences in slopes or intercepts of regression lines were seen for male or female subjects. (Analyses on data from female subjects

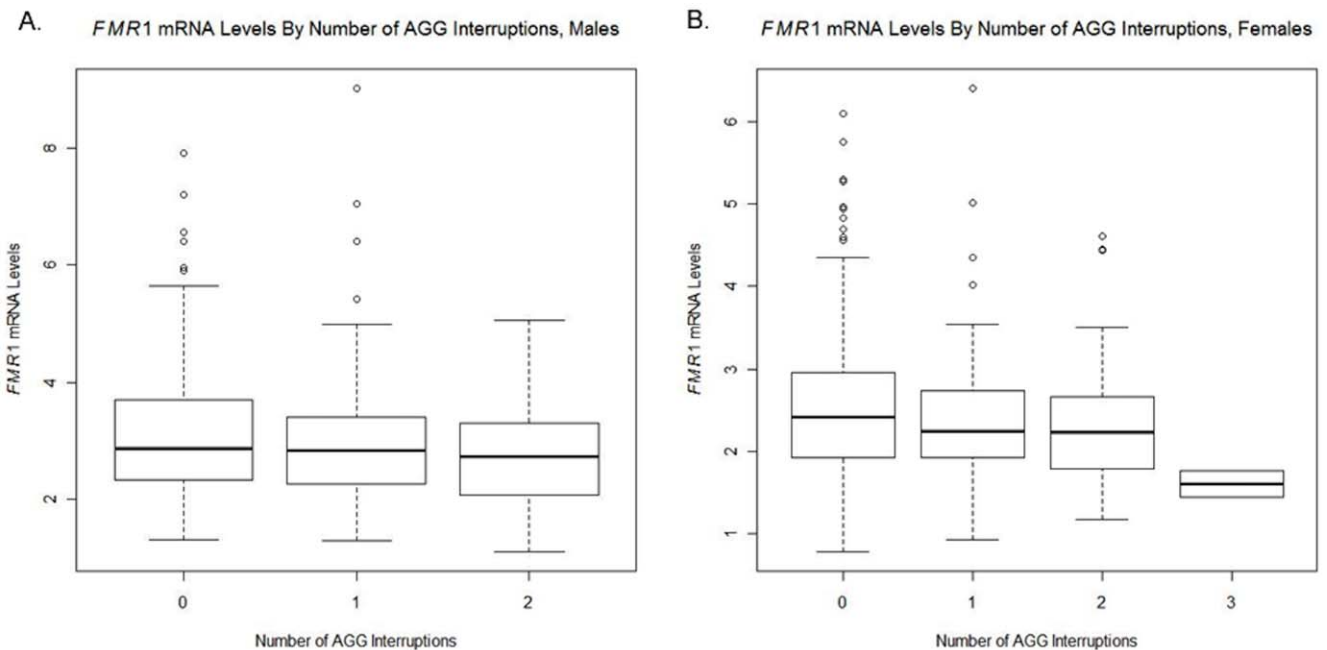


Figure 5. A box plot of *FMRI* mRNA levels for each number of AGG interruptions. No difference was found in *FMRI* mRNA levels between any of the AGG interruption groups a) in males ($P = 0.060$) or b) females ($P = 0.066$). doi:10.1371/journal.pone.0021728.g005

used adjusted mRNA levels and combined 2 and 3 AGG interruptions into a single category.) When controlling for total CGG length, the number of AGG interruptions does not appear to have a meaningful impact on *FMRI* mRNA levels for both genders. (Figure 6a,b; and Table 2)

FMRI mRNA levels are not affected by age

The relationship between *FMRI* mRNA and age, controlling for CGG, was assessed using a linear regression model, with *FMRI* mRNA level as the response variable and total CGG length (continuous) and age as predictors. Although the total CGG length remains strongly predictive of *FMRI* mRNA levels when age is included in the model ($P < 0.001$) for both genders, our data do not suggest a statistically significant relationship between *FMRI* mRNA levels and age in either males or females when CGG is controlled ($P = 0.100$ in males and $P = 0.297$ in females). The scatter plots of *FMRI* mRNA levels by age are shown in Figure 7a,b.

Discussion

Premutation alleles (55–200 CGG repeats) are relatively common in the general population with a prevalence of approximately 1 in 113–259 females and 1 in 251–813 males [13,18,19,20,21]. Although individuals with the premutation were previously thought to be clinically unaffected, it has become clear in the past decade that premutation carriers present with a number of clinical involvement including the fragile X-associated primary ovarian insufficiency (FXPOI) in approximately 20% of the females [22,23,24] and the fragile X-associated tremor ataxia syndrome (FXTAS) [25,26,27,28] which occurs in approximately 1/3 of the male carriers and less frequently in female carriers. In addition, a variety of medical and developmental problems including ADHD and autism spectrum disorders (ASD) [29,30]

and psychiatric and emotional problems [31,32,33,34], most notably depression and anxiety, are also commonly seen in premutation carriers. Psychological problems have been demonstrated to be significantly associated with an increased level of the *FMRI* mRNA [32] commonly observed in premutation carriers [7,8,9]. The increased levels of *FMRI* mRNA has led to the RNA gain-of-function model in which the *FMRI* mRNA itself is causative of the clinical problems observed in carriers, including FXTAS. Consistent with the mRNA toxicity model and the neurodegeneration associated with FXTAS, murine hippocampal neurons from CGG expanded neurons show reduced viability and display deficit in dendritic complexity including shorter dendritic length and fewer branches between 7 and 21 day compared to wild type age matched controls [35]. This provides evidence that disorders associated with the premutation genotype are the result of a global increase in *FMRI* mRNA levels that starts early in life. Our data showing no significant age-dependence on *FMRI* transcriptional levels is in agreement with these findings.

Although it has been demonstrated that the *FMRI* transcription rate, not the stability of the message, is increased in premutation alleles [9,10] the mechanism by which this phenomena occurs has not been clarified yet. Thus, we have studied the potential role of AGG interruption and length of the 3' CGG repeat pure stretch in a large cohort of male and female carriers of the *FMRI* premutation. Our findings indicate that *FMRI* transcription levels, which are elevated in premutation carriers, are not significantly altered by the presence or by the position of AGG interruptions. This comprehensive analysis compared the number of AGG interruptions within a premutation allele and *FMRI* mRNA levels and found significant difference between groups of AGG interruptions when the total length of the CGG tract was factored in. Genders were analyzed separately, adjustments were made for activation ratio in females, each gender comparison resulted in no correlation. These results are consistent with previous studies that

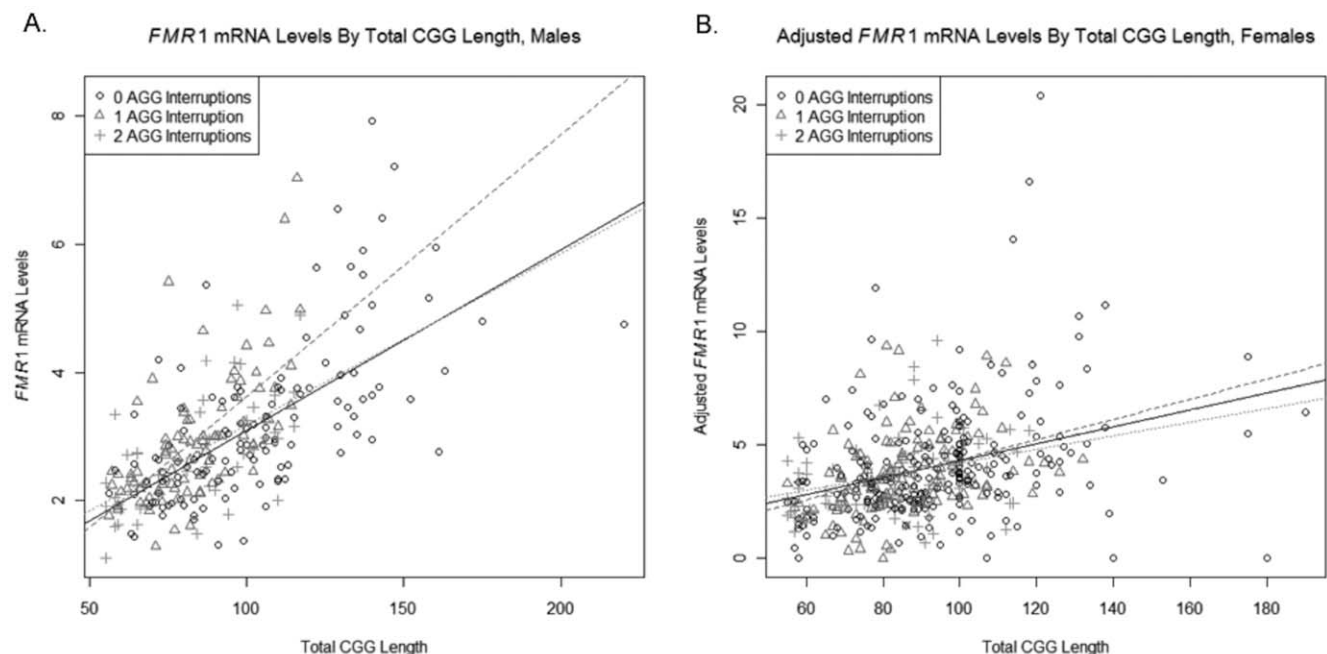


Figure 6. Scatterplot of *FMRI* mRNA levels by total CGG length. A linear regression model showed no significant difference in the slopes for 0, 1, or 2 AGG interruptions a) in males ($P > 0.046$) or b) females ($P > 0.448$). See also Table 2. doi:10.1371/journal.pone.0021728.g006

Table 2. Comparison of linear regression fits for *FMR1* mRNA levels by total CGG length.

Comparison	P-Value, Males	P-Value, Females
0 AGG vs. 1 AGG, Slope	0.046	0.637
0 AGG vs. 2 AGG, Slope	0.840	0.617
1 AGG vs. 2 AGG, Slope	0.101	0.449
0 AGG vs. 1 AGG, Intercept	0.191	0.626
0 AGG vs. 2 AGG, Intercept	0.708	0.642
1 AGG vs. 2 AGG, Intercept	0.185	0.448

doi:10.1371/journal.pone.0021728.t002

compared *FMR1* mRNA levels by the number of AGG interruptions present in a given allele (Tassone et al., 2007). However, given that the total CGG length, which is highly predictive of *FMR1* mRNA levels, differs significantly between subjects with different number of AGG interruptions, further analyses were conducted in which a separate linear regression of *FMR1* mRNA levels on total CGG length was fitted for each number of AGG interruptions. When controlling for total CGG length, the number of AGG interruptions does not appear to have a meaningful impact on *FMR1* mRNA levels.

Finally, we have analyzed the correlation of *FMR1* mRNA levels and age in premutation carriers. This analysis was performed on both genders separately, and results were consistent between genders, as no significant association was made between age and mRNA level. Thus, the increase levels of *FMR1* mRNA in premutation carriers do not appear to change with time (either increasing or decreasing). However, *FMR1* transcript levels were determined in total RNA isolated from peripheral blood

leukocytes. Thus, it is possible that the levels may change with time in other tissues (i.e. brain).

In conclusion, we have demonstrated that the elevated *FMR1* mRNA levels observed in premutation carriers, which are strongly correlated to the CGG repeat number, are not affected by the presence of the AGG interruption. Importantly we have shown that the *FMR1* mRNA levels in premutation carriers are not influenced by the age of the subject. This is of relevance as elevated levels of *FMR1* mRNA have been demonstrated to be causative or related not only to FXTAS, but also to the reported behavioral problems, cognitive impairment and seizure observed in early development in premutation carriers [27,29,30,35,36,37] to autoimmune disorders in female premutation carriers [38,39,40] and perhaps to the primary ovarian insufficiency (FXPOI), observed in ~20% female carriers of a premutation [23,24,41].

Thus, the elevated levels of *FMR1* mRNA, which appear to not change with time, are likely to be causative for both neurodevelopmental and neurodegenerative phenotypes in premutation carriers throughout their life.

Materials and Methods

Subjects

Individuals were recruited through the M.I.N.D. Institute Clinic and provided informed consent for participation. Individuals were selected from fragile X pedigrees based on molecular testing identifying a premutation size allele. None were selected on the basis of clinical symptoms.

Blood was collected in EDTA (for DNA isolation) and Tempus vacutainers (for total RNA isolation) under protocols approved by the Institutional Review Board. Samples were chosen for inclusion in this study based on the CGG repeat allele within the size range of a premutation. In total, 770 DNA samples, comprising 288

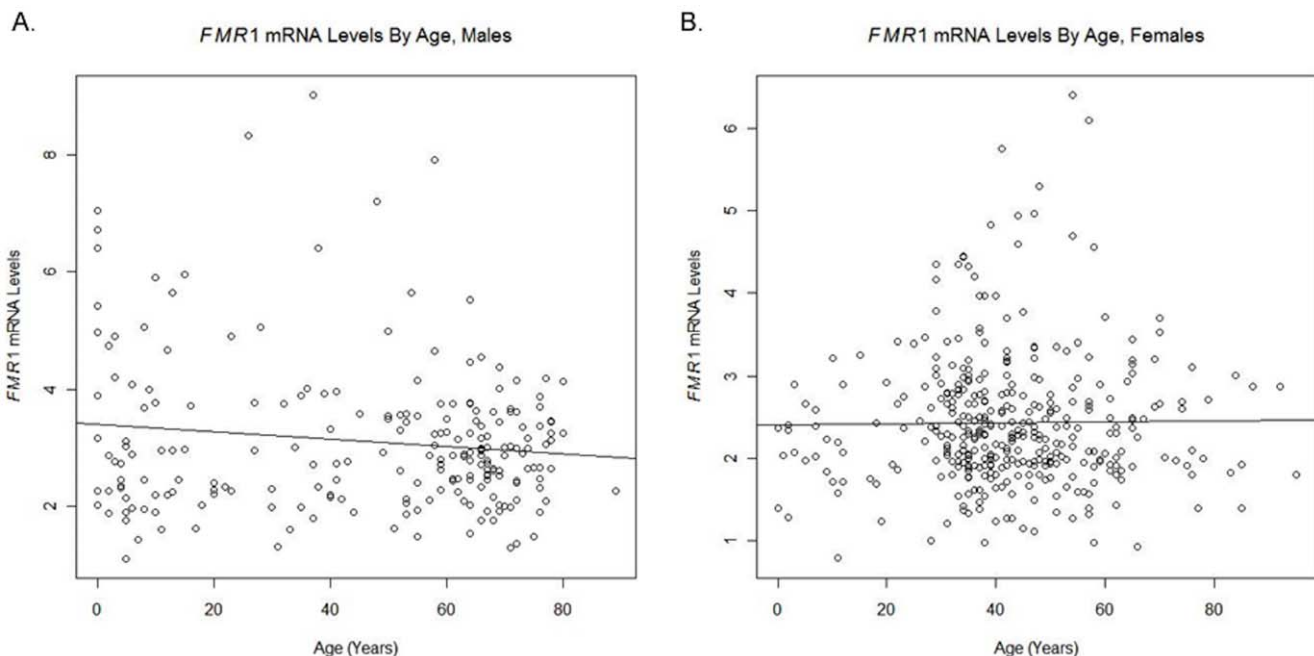


Figure 7. Scatterplot of *FMR1* mRNA levels by age. A linear regression model showed no significant relationship between *FMR1* mRNA levels and age, controlling for CGG length, in a) males ($P=0.093$) or b) females ($P=0.297$).

doi:10.1371/journal.pone.0021728.g007

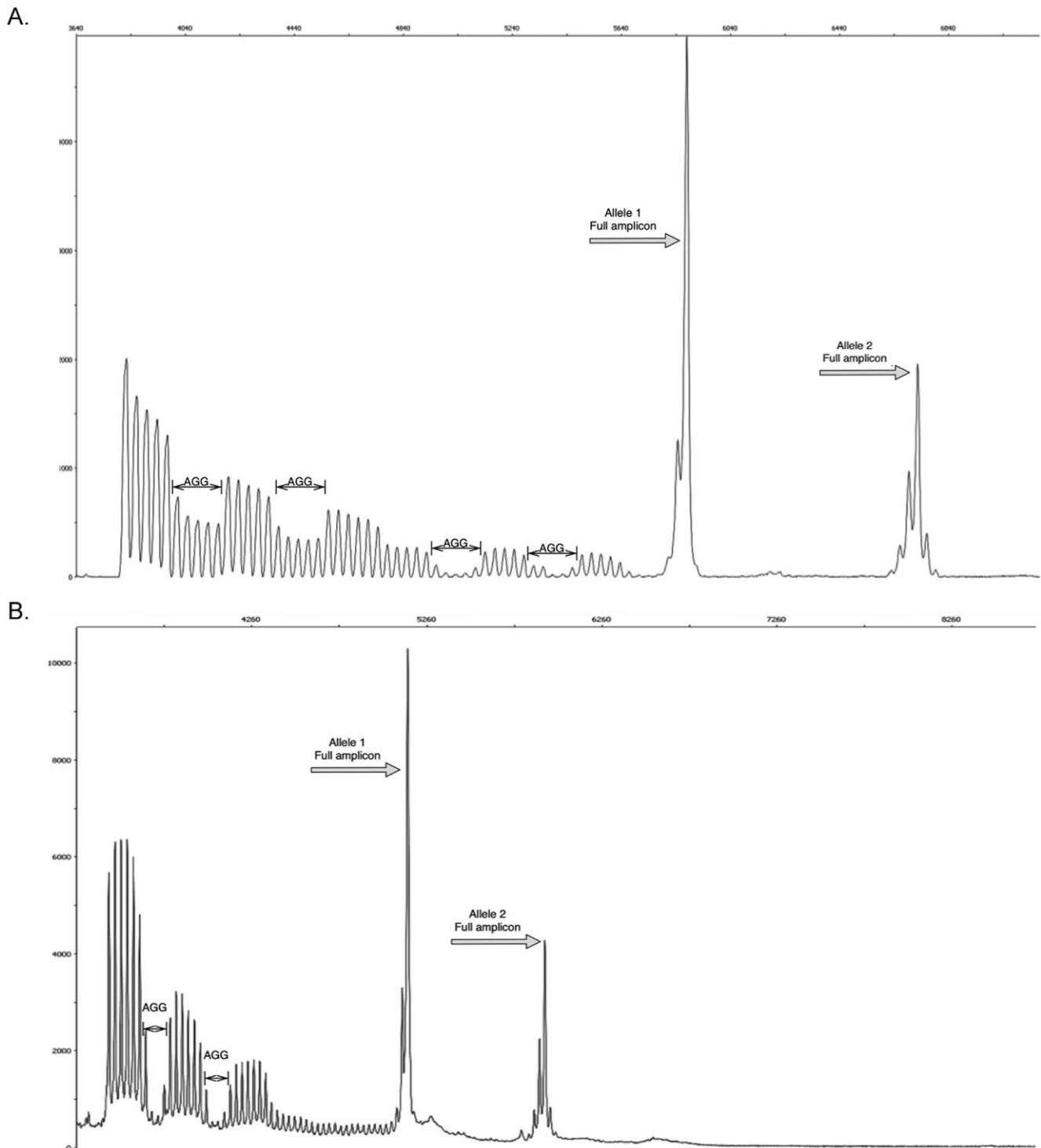


Figure 8. Electropherogram of FMR1 CGG linker PCR. a) PCR product from a premutation female with 1 normal allele (Allele 1) with 30 CGG repeats and 1 premutation allele (Allele 2) with 55 CGG repeats. Using a CCG chimeric primer, AGG interruptions can be identified as loss of signal intensity spanning 5 CGG peaks, followed by restoration of signal intensity. AGG interruptions occurring in the normal allele have a loss of approximately half the CGG peak intensity due to the presence of CGG peaks from the premutation allele. Allele 1: (CGG)₁₀-AGG-(CGG)₉-AGG-(CGG)₉; Allele 2: (CGG)₉-AGG-(CGG)₉-AGG-(CGG)₃₅. b) PCR product using a GGC chimeric primer and the same DNA sample as Fig. 8a. AGG interruptions result in loss of signal in 4 consecutive CGG peaks, followed by restoration in signal. Overlapping AGG interruptions results in peak intensity falling to baseline levels.

doi:10.1371/journal.pone.0021728.g008

males and 482 females, were selected to be molecularly characterized (including CGG repeat number, length of 3' pure CGG stretch, activation ratio, *FMRI* mRNA expression levels and presence of AGG interruptions).

Molecular measures

CGG repeat size. Genomic DNA (gDNA) was isolated from 5 mL of peripheral blood leukocytes using Gentra Puregene Blood Kit (Qiagen). gDNA was used to determine the size of the CGG repeat allele, methylation status and the *FMRI* activation ratio in females by Southern Blot analysis and PCR amplification, using established diagnostic protocols [42].

***FMRI* mRNA levels.** Total RNA was isolated from whole blood collected in Tempus vacutainers (Applied Biosystems). Reverse transcription followed by real time PCRs (qRT-PCR) to measure the *FMRI* transcript levels were performed on total RNA samples, as previously described [9].

The relative abundance of *FMRI* mRNA was assessed using the β -glucuronidase (*GUS*) mRNA as a reference gene. For each sample, RT reactions were run starting from 3 different total RNA concentrations (500–250 and 125 ng) and quantitative-PCR reactions were performed for both *FMRI* and *GUS* amplicons for a total of 12 reactions. Control reactions (*FMRI* and *GUS*) were run in parallel with total RNA isolated from a standard control lymphoblastoid used as a calibrator (RPM 7666; ATCC CCL114). Details of the methodology and analysis are as described in [9].

AGG interruptions. A PCR approach including a CGG repeat primer, as previously described [35,42] was used to determine the number and the distribution of AGG interruptions within the *FMRI* gene. PCR was performed using 100 ng of gDNA following manufacturer's recommendations (Asuragen) and as described in Chen et al., [35]. The presence and the distribution of AGG interruptions was determined by the analysis of an electropherogram as illustrated in Figure 8a and in Figure S4. Briefly, a three primer assay is designed to amplify the 239 bp flanking region of the trinucleotide repeat (using c and f primers) [43], resulting in a PCR fragment corresponding to the CGG repeat size. The third primer, a CGG chimeric primer (CCG)₅ which anneals to CGG tracts, amplifies intermediate fragments that are variable in length. The chimeric primer is unable to anneal to a CGG tract that is interrupted with an AGG, this allows the location of interruptions to be mapped by identification of a corresponding absence of PCR product.

gDNA was amplified in a 15 μ L total reaction volume using ABI GeneAmp 9700 (Applied Biosystems) under the following conditions: 95°C for 5 minutes; followed by 10 cycles of 97°C for 35 seconds, 62°C for 35 seconds, 68°C for 4 minutes; 20 cycles of 97°C for 35 seconds, 62°C for 35 seconds, 68°C for 4 minutes and an additional 20 seconds per cycle; 72°C extension for 10 minutes and finally incubation at 4°C. Two microliters of PCR product was denatured with 2 μ L of ROX 1000 size marker (Asuragen) in Hi-Di Formamide (Applied Biosystems) to a total volume of 15 μ L for 2 minutes at 95°C followed by 3 minutes at 4°C. The single stranded amplified product was visualized by capillary electrophoresis using the ABI 3100 Genetic Analyzer (Applied Biosystems). Fragments were separated through a 36-cm capillary array using POP4 polymer (Applied Biosystems) following manufacturers recommendations with the following changes to the default run parameters: injection for 20 seconds at 1250 volts, and a run time of 50 minutes. AGG positions within the CGG element were determined from raw Genescan files imported into Peakscanner software (v1.0) (Applied Biosystems). FAM-6 labeled PCR product is displayed as an electrophero-

gram. The full fragment visualized as a distinct band(s) amplified by the c and f primers is first identified and sized using the size standard marker (Asuragen). The CGG linker primer generates PCR fragments that are smaller than the sizing amplicons and are varying in size. They are identified on Peakscanner as peaks with lower signal intensity than sizing peaks of normal range alleles and are smaller fragment lengths than sizing peak that correspond to the same allele. The presence of an AGG interruption can be visualized on the electropherogram as a loss of signal intensity spanning 5 CGG peaks, followed by the restoration of signal intensity (Figure 8a,b). These 5 peaks correspond to the positions of the AGG interruption and 4 CGG segments up and downstream of the AGG interruption. Similarly, the first CGG peak on the electropherogram is the 5th CGG repeat of the amplicon, the first 5 repeats being the earliest position on the DNA template for the CGG linker primer to anneal. To characterize the CGG tract the peaks and dips are ordered in reverse of their position on the electropherogram, given the directionality of the f and CGG linker primers. AGG interruptions are expected at the far end of the electropherogram, where signal intensity naturally decreases from longer PCR amplification being less thermodynamically favorable than shorter fragments.

A sample set including 47 DNAs underwent an additional PCR step to confirm the position of the AGG interruption using a (GGC)₄ chimeric primer as described in Tassone et al., [42] (Figure 8b). In this case the linker anneals to the CGG tract on the reverse strand, and only requires 4 uninterrupted repeats. This primer amplifies in conjunction with the c primer, creating the same AGG interruption dips that were previously seen at the tail end of the electropherogram but now visualized at the beginning of the electropherogram with the highest signal intensity. It is to be noted that potential ambiguity of AGG position may exist in female samples in cases where the location of the AGG interruption could be attributed to the same position of both alleles (normal and premutation). There is a small risk for AGG interruptions of the premutation allele to be incorrectly interpreted as being present on the normal allele in females. The ambiguity of one AGG interruption in a location where two CGG repeat sequences overlap can typically be resolved in premutation carriers using the CGG linker because interruptions are expected to occur at the 3' end of the CGG repeat. The CGG linker assay amplifies the 5' end of the CGG repeat first which provides a length of space between the AGG interruptions on a premutation allele with the AGG interruptions of a normal allele. The cases where this may not be true are when the premutation allele is a low premutation and the normal is a high normal, also many AGG interruptions will decrease the space between premutation and normal AGG interruption ranges, and very rare AGG interruption patterns that are located more than 20 repeats from the 3' end of the CGG repeat. In order to resolve these ambiguous female premutation carrier samples the typical CGG linker assay was used and the presence or absence of AGG interruptions was determined by digest with EciI as previously described [10].

Supporting Information

Figure S1 Distribution of age in participants. The distribution of ages in a) males and b) females. (TIFF)

Figure S2 Distribution of *FMRI* mRNA levels. Distributions of *FMRI* mRNA levels in a) males and b) females. (TIFF)

Figure S3 Distribution of CGG repeats in the *FMR1* gene. Distributions of *FMR1* CGG repeats are similar in a) males and b) females. (TIFF)

Figure S4 Electropherogram of *FMR1* CGG linker PCR. PCR product from a premutation female with 1 normal allele (Allele 1) with 2 AGG interruptions in 30 CGG repeats and 1 premutation allele (Allele 2) with 2 AGG interruptions in 87 CGG repeats. Determination of the number and location of the AGG interruptions within each allele was performed as described in the method section. (TIFF)

References

- Oostra BA, Willemsen R (2009) FMR1: a gene with three faces. *Biochim Biophys Acta* 1790: 467–477.
- Chonchaiya W, Schneider A, Hagerman RJ (2009) Fragile X: a family of disorders. *Adv Pediatr* 56: 165–186.
- Maddalena A, Richards CS, McGinniss MJ, Brothman A, Desnick RJ, et al. (2001) Technical standards and guidelines for fragile X: the first of a series of disease-specific supplements to the Standards and Guidelines for Clinical Genetics Laboratories of the American College of Medical Genetics. *Genetics in Medicine* 3: 200–205.
- Hagerman PJ, Hagerman RJ (2004) The fragile-X premutation: a maturing perspective. *Am J Hum Genet* 74: 805–816.
- Amiri K, Hagerman RJ, Hagerman PJ (2008) Fragile X-associated tremor/ataxia syndrome: an aging face of the fragile X gene. *Arch Neurol* 65: 19–25.
- Boyle L, Kaufmann WE (2010) The behavioral phenotype of FMR1 mutations. *Am J Med Genet C Semin Med Genet* 154C: 469–476.
- Allen EG, He W, Yadav-Shah M, Sherman SL (2004) A study of the distributional characteristics of FMR1 transcript levels in 238 individuals. *Hum Genet* 114: 439–447.
- Kenneson A, Zhang F, Hagedorn CH, Warren ST (2001) Reduced FMRP and increased FMR1 transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. *Hum Mol Genet* 10: 1449–1454.
- Tassone F, Hagerman RJ, Taylor AK, Gane LW, Godfrey TE, et al. (2000) Elevated levels of FMR1 mRNA in carrier males: a new mechanism of involvement in the fragile-X syndrome. *Am J Hum Genet* 66: 6–15.
- Tassone F, Beilina A, Carosi C, Albertosi S, Bagni C, et al. (2007) Elevated FMR1 mRNA in premutation carriers is due to increased transcription. *RNA* 13: 555–562.
- Primerano B, Tassone F, Hagerman RJ, Hagerman PJ, Amaldi F, et al. (2002) Reduced FMR1 mRNA translation efficiency in Fragile X patients with premutations. *RNA* 8: 1482–1488.
- Peprah E, He W, Allen E, Oliver T, Boyne A, et al. (2010) Examination of FMR1 transcript and protein levels among 74 premutation carriers. *J Hum Genet* 55: 66–68.
- Dombrowski C, Levesque ML, Morel ML, Rouillard P, Morgan K, et al. (2002) Premutation and intermediate-size FMR1 alleles in 10 572 males from the general population: loss of an AGG interruption is a late event in the generation of fragile X syndrome alleles. *Hum Mol Genet* 11: 371–378.
- Eichler EE, Kunst CB, Lugenbeel KA, Ryder OA, Davison D, et al. (1995) Evolution of the cryptic FMR1 CGG repeat. *Nat Genet* 11: 301–308.
- Eichler EE, Nelson DL (1996) Genetic variation and evolutionary stability of the FMR1 CGG repeat in six closed human populations. *Am J Med Genet* 64: 220–225.
- Jarem DA, Huckaby LV, Delaney S (2010) AGG interruptions in (CGG)_n DNA repeat tracts modulate the structure and thermodynamics of non-B conformations in vitro. *Biochemistry* 49: 6826–6837.
- Ludwig AL, Raske C, Tassone F, Garcia-Arocena D, Hershey JW, et al. (2009) Translation of the FMR1 mRNA is not influenced by AGG interruptions. *Nucleic Acids Res* 37: 6896–6904.
- Fernandez-Carvajal I, Walichiewicz P, Xiaosen X, Pan R, Hagerman PJ, et al. (2009) Screening for expanded alleles of the FMR1 gene in blood spots from newborn males in a Spanish population. *J Mol Diagn* 11: 324–329.
- Rousseau F, Rouillard P, Morel ML, Khandjian EW, Morgan K (1995) Prevalence of carriers of premutation-size alleles of the FMR1 gene—and implications for the population genetics of the fragile X syndrome. *Am J Hum Genet* 57: 1006–1018.
- Toledano-Alhadeif H, Basel-Vanagaite L, Magal N, Davidov B, Ehrlich S, et al. (2001) Fragile-X Carrier Screening and the Prevalence of Premutation and Full-Mutation Carriers in Israel. *Am J Hum Genet* 69: 351–360.
- Hagerman PJ (2008) The fragile X prevalence paradox. *J Med Genet* 45: 498–499.
- Cronister A, Schreiner R, Wittenberger M, Amiri K, Harris K, et al. (1991) Heterozygous fragile X female: historical, physical, cognitive, and cytogenetic features. *Am J Med Genet* 38: 269–274.
- Sullivan AK, Marcus M, Epstein MP, Allen EG, Anido AE, et al. (2005) Association of FMR1 repeat size with ovarian dysfunction. *Hum Reprod* 20: 402–412.
- Wittenberger MD, Hagerman RJ, Sherman SL, McConkie-Rosell A, Welt CK, et al. (2007) The FMR1 premutation and reproduction. *Fertil Steril* 87: 456–465.
- Hagerman RJ, Leehey M, Heinrichs W, Tassone F, Wilson R, et al. (2001) Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. *Neurology* 57: 127–130.
- Jacquemont S, Hagerman RJ, Leehey M, Grigsby J, Zhang L, et al. (2003) Fragile X premutation tremor/ataxia syndrome: molecular, clinical, and neuroimaging correlates. *Am J Hum Genet* 72: 869–878.
- Garcia-Arocena D, Hagerman PJ (2010) Advances in understanding the molecular basis of FXTAS. *Hum Mol Genet* 19: R83–89.
- Leehey MA (2009) Fragile X-associated tremor/ataxia syndrome: clinical phenotype, diagnosis, and treatment. *J Investig Med* 57: 830–836.
- Aziz M, Stathopulu E, Callias M, Taylor C, Turk J, et al. (2003) Clinical features of boys with fragile X premutations and intermediate alleles. *Am J Med Genet* 121B: 119–127.
- Farzin F, Perry H, Hessl D, Loesch D, Cohen J, et al. (2006) Autism spectrum disorders and attention-deficit/hyperactivity disorder in boys with the fragile X premutation. *J Dev Behav Pediatr* 27: S137–144.
- Bourgeois J, Coffey S, Rivera S, Hessl D, Gane L, et al. (2009) Fragile X Premutation Disorders - Expanding the Psychiatric Perspective. in press.
- Hessl D, Tassone F, Loesch DZ, Berry-Kravis E, Leehey MA, et al. (2005) Abnormal elevation of FMR1 mRNA is associated with psychological symptoms in individuals with the fragile X premutation. *Am J Med Genet B Neuropsychiatr Genet* 139: 115–121.
- Roberts JE, Bailey DB, Jr., Mankowski J, Ford A, Sideris J, et al. (2009) Mood and anxiety disorders in females with the FMR1 premutation. *Am J Med Genet B Neuropsychiatr Genet* 150B: 130–139.
- Rodríguez-Revenga L, Madrigal I, Alegret M, Santos M, Mila M (2008) Evidence of depressive symptoms in fragile-X syndrome premutated females. *Psychiatr Genet* 18: 153–155.
- Chen L, Hadd A, Sah S, Filipovic-Sadic S, Krosting J, et al. (2010) An information-rich CGG repeat primed PCR that detects the full range of fragile X expanded alleles and minimizes the need for southern blot analysis. *J Mol Diagn* 12: 589–600.
- Bailey Jr. DBRM, Olmsted M, Holiday DB (2008) Co-Occurring Conditions Associated With FMR1 Gene Variations: Findings From a National Parent Survey. *Am J Med Genet* 146A: 2060–2069.
- Chen Y, Tassone F, Berman RF, Hagerman PJ, Hagerman RJ, et al. (2010) Murine hippocampal neurons expressing Fmr1 gene premutations show early developmental deficits and late degeneration. *Hum Mol Genet* 19: 196–208.
- Greco CM, Tassone F, Garcia-Arocena D, Tartaglia N, Coffey SM, et al. (2008) Clinical and neuropathologic findings in a woman with the FMR1 premutation and multiple sclerosis. *Arch Neurol* 65: 1114–1116.
- Coffey SM, Cook K, Tartaglia N, Tassone F, Nguyen DV, et al. (2008) Expanded clinical phenotype of women with the FMR1 premutation. *Am J Med Genet A* 146: 1009–1016.
- Zhang L, Coffey S, Lua LL, Greco CM, Schafer JA, et al. (2009) FMR1 premutation in females diagnosed with multiple sclerosis. *J Neurol Neurosurg Psychiatry* 80: 812–814.
- Sherman SL (2000) Premature ovarian failure in the fragile X syndrome. *Am J Med Genet* 97: 189–194.
- Tassone F, Pan R, Amiri K, Taylor AK, Hagerman PJ (2008) A rapid polymerase chain reaction-based screening method for identification of all expanded alleles of the fragile X (FMR1) gene in newborn and high-risk populations. *J Mol Diagn* 10: 43–49.
- Yu S, Pritchard M, Kremer E, Lynch M, Nancarrow J, et al. (1991) Fragile X genotype characterized by an unstable region of DNA. *Science* 252: 1179–1181.

Acknowledgments

We thank all of the participants and their families for their contributions. We also thank Ka Pou Iong and Tzuhan Tong for their technical assistance, and Andrew Hadd for his help with AGG interruption profiles. This work is dedicated to the memory of Matteo.

Author Contributions

Conceived and designed the experiments: CMY Flora Tassone. Performed the experiments: CMY Federica Tassone. Analyzed the data: CMY Federica Tassone BDJ Flora Tassone. Contributed reagents/materials/analysis tools: Flora Tassone BDJ. Wrote the paper: CMY Flora Tassone BDJ. Designed the new methodology to determine the number of AGG interruptions and distribution: Flora Tassone CMY.