



# Large Interruptions of GAA Repeat Expansion Mutations in Friedreich Ataxia Are Very Rare

Sahar Al-Mahdawi<sup>1,2</sup>, Heather Ging<sup>3</sup>, Aurelien Bayot<sup>3†</sup>, Francesca Cavalcanti<sup>4</sup>, Valentina La Cognata<sup>5</sup>, Sebastiano Cavallaro<sup>5</sup>, Paola Giunti<sup>3</sup> and Mark A. Pook<sup>1,2\*</sup>

<sup>1</sup> Ataxia Research Group, Division of Biosciences, Department of Life Sciences, College of Health and Life Sciences, Brunel University London, Uxbridge, United Kingdom, <sup>2</sup> Synthetic Biology Theme, Institute of Environment, Health and Societies, Brunel University London, Uxbridge, United Kingdom, <sup>3</sup> Ataxia Centre, Department of Molecular Neuroscience, Institute of Neurology, University College London, London, United Kingdom, <sup>4</sup> Institute of Neurological Sciences, National Research Council, Mangone, Italy, <sup>5</sup> Institute of Neurological Sciences, National Research Council, Catania, Italy

## OPEN ACCESS

### Edited by:

Egidio D'Angelo,  
University of Pavia, Italy

### Reviewed by:

Renata Santos,  
INSERM U894 Centre de Psychiatrie  
et Neurosciences, France  
Michael Huang,  
The University of Sydney, Australia

### \*Correspondence:

Mark A. Pook  
Mark.Pook@brunel.ac.uk

### †Present address:

Aurelien Bayot,  
Mitochondrial Biology Group, CNRS  
UMR 3691, Department of Cell  
Biology and Infection, Institut Pasteur,  
Paris, France

**Received:** 26 July 2018

**Accepted:** 05 November 2018

**Published:** 21 November 2018

### Citation:

Al-Mahdawi S, Ging H, Bayot A,  
Cavalcanti F, La Cognata V,  
Cavallaro S, Giunti P and Pook MA  
(2018) Large Interruptions of GAA  
Repeat Expansion Mutations  
in Friedreich Ataxia Are Very Rare.  
*Front. Cell. Neurosci.* 12:443.  
doi: 10.3389/fncel.2018.00443

Friedreich ataxia is a multi-system autosomal recessive inherited disorder primarily caused by homozygous GAA repeat expansion mutations within intron 1 of the frataxin gene. The resulting deficiency of frataxin protein leads to progressive mitochondrial dysfunction, oxidative stress, and cell death, with the main affected sites being the large sensory neurons of the dorsal root ganglia and the dentate nucleus of the cerebellum. The GAA repeat expansions may be pure (GAA)<sub>n</sub> in sequence or may be interrupted with regions of non-GAA sequence. To our knowledge, there has been no large-scale study of FRDA patient DNA samples to determine the frequency of large interruptions in GAA repeat expansions. Therefore, we have investigated a panel of 245 Friedreich ataxia patient and carrier DNA samples using GAA repeat PCR amplification and *MbolI* restriction enzyme digestion. We demonstrate that the vast majority (97.8%) of Friedreich ataxia GAA repeat expansion samples do not contain significant sequence changes that would result in abnormal *MbolI* digestion profiles, indicating that they are primarily pure GAA repeats. These results show for the first time that large interruptions in the GAA repeats are very rare.

**Keywords:** Friedreich ataxia, FRDA, frataxin, GAA repeat mutation, GAA repeat interruptions, trinucleotide repeat expansion disease

## INTRODUCTION

Friedreich ataxia (FRDA) is a multi-system autosomal recessive inherited disorder characterized by neurological features of ataxia, dysarthria, weakness, ocular fixation instability, deep sensory loss, and visual and hearing impairment, together with non-neurological features such as hypertrophic cardiomyopathy, diabetes mellitus, kyphoscoliosis, and foot deformities (Reetz et al., 2015). The mean age of onset of FRDA is 15 years, with most cases developing by age 25, although rare cases of late onset FRDA (26–39 years), or very late onset FRDA (40 years or over) have been reported (Durr et al., 1996; Filla et al., 1996). FRDA is primarily caused by homozygous GAA trinucleotide repeat expansion mutations within intron 1 of the frataxin (*FXN*) gene, leading to deficit of the essential mitochondrial protein frataxin. The resulting deficiency of frataxin protein leads to progressive mitochondrial dysfunction, oxidative stress, and cell death, with the main affected sites being the large sensory neurons of the dorsal root ganglia and the dentate nucleus

of the cerebellum (Campuzano et al., 1996, 1997). The number of GAA repeats in unaffected individuals generally ranges from 6 to 27 repeats, although rare cases of 33 to 130 repeats have been identified (Cossee et al., 1997; Montermini et al., 1997; Ohshima et al., 1999). In contrast, approximately 96% of FRDA patients have homozygous GAA repeat expansions ranging from 44 to 1,700 repeats, with 600–900 GAA repeats being the most common. The remaining FRDA patients are compound heterozygous for a GAA repeat expansion and a second *FXN* mutation. The age of onset is inversely correlated with the GAA repeat number, particularly for the shorter allele (GAA1), with a prediction of 2.3 years earlier onset for every 100 GAA repeats added to GAA1 (Reetz et al., 2015). Thus, the GAA1 allele is considered to give the closest genotype–phenotype relationship. However, the GAA repeat size accounts for only approximately 36% to 56% of the variation in age of onset (Durr et al., 1996; Filla et al., 1996; Reetz et al., 2015). This suggests that there are other contributory mechanisms such as somatic mosaicism, interruptions in the GAA repeat sequence, and other modifying genes or environmental factors, which influence age of onset (Filla et al., 1996; Pandolfo, 2009; Reetz et al., 2015). Earlier age of onset increases the frequency and severity of neurological and non-neurological symptoms. In contrast, patients with late-onset disease have a milder phenotype and slower disease progression (Durr et al., 1996).

In the majority of cases of FRDA, the GAA repeat expansion mutations have been characterized only in terms of the overall repeat size rather than sequence content, which can be either pure GAA or interrupted GAA. Where full sequencing of *FXN* GAA repeats has been possible due to comparatively short repeats (up to approximately 130 repeats), there are reports of interrupted GAA repeat expansion sequences, such as (GAGGAA)<sub>5–9</sub> or (GAAGGA)<sub>65</sub>, and these are associated with either absence of FRDA disease phenotype (Cossee et al., 1997; Montermini et al., 1997; Ohshima et al., 1999) or atypical mild late-onset or very late-onset FRDA disease phenotype (Cossee et al., 1997; Epplen et al., 1997; Moseley et al., 1998; McDaniel et al., 2001; Sharma et al., 2004; Stolle et al., 2008). *In vitro* studies have shown that interrupted GAA repeats inhibit triplex formation, alleviate transcription inhibition, and reduce repeat instability (Ohshima et al., 1999; Sakamoto et al., 2001). Therefore, it is thought that interruptions in the GAA repeat sequence can result in maintenance of *FXN* expression levels and reduced somatic instability of GAA repeats when compared to similar sized pure GAA repeats, thereby impacting upon FRDA disease progression. Further studies have shown that pure GAA repeats are stable up to a threshold of 44 repeats, after which they become unstable (Sharma et al., 2004). In a study of FRDA carriers, 107 pure GAA repeats were unstable, whereas 114 interrupted GAA repeats were stable by small-pool PCR of blood samples (Pollard et al., 2004). Also a (GAA)<sub>90</sub>(GAAAGAA)<sub>9</sub>(GAA)<sub>20</sub> interrupted 112 GAA repeat has been shown to stably transmit through two generations (Cossee et al., 1997). Finally, somatic instability of the GAA1 allele has been identified in two FRDA patients who have one large GAA2 allele and a small GAA1 allele of either 44 or 66 repeats, while their sibling with one large GAA2 allele and one somatically stable GAA1 allele of 37 repeats was clinically

normal (Sharma et al., 2004). So, it would appear that somatic instability may have a significant role to play in FRDA disease progression and interrupted GAA repeats are likely to confer some protection in the form of somatic stability and maintenance of *FXN* expression levels.

Analysis of the much more frequent large GAA2 alleles (>130 repeats) for the presence or absence of interruptions has not previously been performed to any great extent, primarily because of the technical difficulty in obtaining accurate sequence of an entire long GAA repeat expansion. Sakamoto et al. (2001) sequenced 11 expanded FRDA alleles for as long as was technically possible, up to approximately 200 repeats, and they identified interruptions in 5 of the 11 samples, which were clustered within the last 10–15 repeats at the 3' end. They also used *EarI* (GAAGAG recognition sequence) and *MnII* (GAGG recognition sequence) restriction enzyme digestion to identify specified interruptions within 4 out of 22 FRDA patient samples (Sakamoto et al., 2001). Subsequently, Holloway et al. (2011) reported the development of a useful method to determine non-specified GAA repeat interruptions comprising long-range GAA PCR followed by *MboII* (GAAGA recognition sequence) restriction enzyme digestion. Only four FRDA patients were investigated in this study, but one of the four was found to contain an extensive region of GAA repeat interruptions in their larger 1140 GAA repeat allele, which was partially sequenced to be (GAA)<sub>21</sub>(GGAGAA)<sub>5</sub>(GGAGGAGAA)<sub>70</sub>(GAA)<sub>n</sub>. To expand upon these previous small-number studies, we have now addressed the question, “Do most FRDA patients have pure GAA repeat expansions?” by investigating a large panel of 245 FRDA patient and carrier DNA samples using long-range GAA repeat PCR amplification and *MboII* restriction enzyme digestion. We find that the vast majority (97.8%) of FRDA patient DNA samples do not contain substantial sequence changes that would result in abnormal *MboII* digestion profiles, indicating that they are primarily pure GAA repeats. We discuss the implications of this finding with regard to GAA repeat instability and FRDA disease progression.

## MATERIALS AND METHODS

### *MboII* Digestion Analysis

We obtained 245 peripheral blood genomic DNA samples from FRDA patients (238 samples) and carriers (7 samples) that had previously undergone GAA repeat expansion size determination (**Supplementary Table S1**) (ethics approval granted within the European Union Seventh Framework Programme [FP7/2007–2013] under grant agreement number 242193/EFACTS). We also obtained genomic DNA samples from cerebellum autopsy tissues from three FRDA patients (tissues were registered with the HTA under Brunel University Licensing number 12543) and five ear biopsies from previously described GAA repeat expansion-based Y47R and YG8sR FRDA mouse models (Anjomani Virmouni et al., 2015) (animal procedures were carried out in accordance with the UK Home Office “Animals (Scientific Procedures) Act 1986” and with approval from the Brunel University London Animal Welfare and Ethical Review Board). We then performed

long-range PCR of the samples (approximately 100 ng input DNA) using either the Expand High Fidelity PCR System, dNTPack (Roche), or the Long Range PCR Kit (Qiagen) together with GAA-B-F (5'-AATGGATTTCCTGGCAGGACGC-3') and GAA-B-R (5'-GCATTGGGCGATCTTGGCTTAA-3') primers as previously described (Holloway et al., 2011). The thermocycling conditions used were (i) Roche Kit: 94°C for 2 min; 10 cycles of 94°C for 10 s, 60°C for 30 s, 68°C for 45 s; 20 cycles of 94°C for 10 s, 60°C for 30 s, 68°C for 1 min with 20 s increments; and a final cycle of 68°C for 10 mins, or (ii) Qiagen Kit: 93°C for 3 min; 35 cycles of 93°C for 15 s, 62°C for 30 s, 68°C for 5 min, and a final cycle of 68°C for 10 min. The amplified PCR products contained the GAA trinucleotide repeat expansion with flanking sequences of 170bp at the 5' end and 120bp at the 3' end. A 5 µl sample of each PCR product was checked by running on 1% agarose gels. The positive samples were then digested with *MboII*, which has a cleavage sequence of 5'-GAAGA(8/7)-3'. PCR sample (17 µl) was digested in a total reaction volume of 20 µl at 37°C for 1 h. Digested DNA fragments were subjected to heating at 95°C for 10 min followed by slow cooling to room temperature to prevent potential heteroduplex formation, then separated by running on 2% agarose gels [1% Nusieve (Seakem Agarose GTG) and 1% Metaphor agarose]. As the *MboII* restriction enzyme cuts at the recognition site 5'-GAAGA (8/7)-3', pure GAA repeats are completely cut by *MboII* leaving only two fragments from the uncut flanking sequences, 171/170bp upstream (designated hereafter as "170bp") and 117/118bp downstream (designated hereafter as "120bp"), which do not contain *MboII* sites. On the other hand, if the GAA repeat expansion contains interrupted GAA sequences that are not cut by *MboII*, a different pattern of bands would be obtained upon agarose gel electrophoresis: either two bands with sizes that differ from the expected 170 and 120bp or else more than two bands.

## DNA Sequencing

GAA PCR products were purified from agarose gels using GeneClean (MP Biomedicals), then cloned into pCR4.0 using the TOPO TA cloning kit (Invitrogen). For each PCR product, plasmid DNA from two independent colonies was Sanger sequenced by Genewiz using T3 and T7 sequencing primers.

## Statistical Analysis

Correlation and regression analysis of GAA repeat sizes versus age of onset and subsequent ANOVA statistics were performed using Microsoft Excel data analysis tools (GAA sizes for samples 89–245 were obtained from the EFACTS database).

## RESULTS

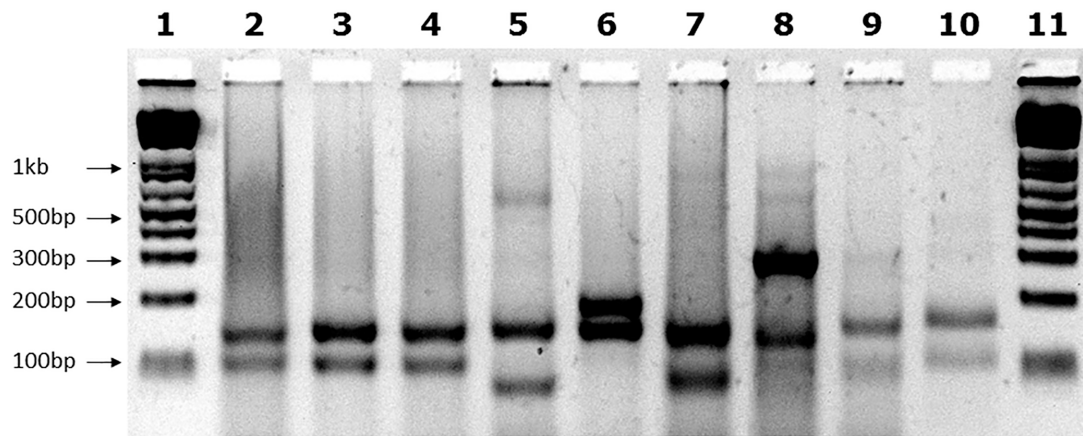
### The Majority of FRDA GAA Repeats Do Not Contain Interruptions Detected by *MboII* Analysis

We examined the GAA repeat status in 245 FRDA patient and carrier peripheral blood DNA samples by long-range GAA PCR followed by *MboII* restriction enzyme digestion

and agarose gel electrophoresis. This method can detect significant GAA interruptions in either of the two alleles at a resolution of approximately 20bp added to either of the two expected *MboII* bands of 170 and 120bp or approximately 50bp of internal interruptions (Holloway et al., 2011). As a positive control for known interrupted FXN GAA repeat sequence, we used DNA from the "NEP" BAC transgenic mouse that contains approximately 500 triplet repeats with the previously determined interrupted sequence of (GAA)<sub>21</sub>(GGAGAA)<sub>5</sub>(GGAGGAGAA)<sub>70</sub>(GAA)<sub>n</sub> (Holloway et al., 2011). Eighteen of our FRDA DNA samples failed to generate sufficient PCR product for the assay, leaving 227 samples for further analysis; 220 FRDA patients and 7 carriers (Table 1). 222 out of these 227 samples (97.8%) produced only the expected two *MboII* bands of 170 and 120bp, indicating the likely presence of pure GAA repeats in both alleles in the vast majority of cases (Figure 1, lanes 2–4). However, the remaining five samples (2.2%) showed abnormal *MboII* band profiles, indicating disruption of at least one of the two alleles. One sample produced *MboII* bands of 170 and 200bp, indicating an insertion of approximately 80bp in the 3' GAA flanking sequence (Figure 1, lane 6). Sequencing of the cloned GAA PCR products from this sample confirmed that there was an 80bp duplication in the 3' GAA flanking region of the smaller allele, duplicating the sequence 24–104bp downstream of the final GAA repeat, that did not involve any interruption of the GAA repeat itself. Another sample produced *MboII* bands of 170 and 100bp, indicating a deletion of approximately 20bp in the 3' GAA flanking sequence (Figure 1, lane 7). Sequencing of the cloned GAA PCR products from this sample confirmed that there was a 19bp deletion of the smaller allele immediately after the GAA repeat sequence, but not involving any interruption of the GAA repeat itself. The remaining three abnormal *MboII* samples produced bands of approximately 100, 180, and 300bp (Figure 1, lanes 8–10), but sequencing did not identify any substantial insertions or deletions in the 5' or 3' flanking regions or in the GAA repeats close to these flanking regions, indicating the likely presence of significant GAA repeat interruptions within the body of the GAA repeat expansion alleles and beyond the capability of our sequencing technology. However, sequencing of up to 56 repeats at the 5' and 3' ends of the nine FRDA samples shown in Figure 1 did reveal small GAA repeat sequence interruptions in three of the nine samples (i.e., 33%) (Figure 1, lane 7 sample: ... (GAA)<sub>23</sub>A(GAA)<sub>5</sub>AGAA, lane

TABLE 1 | Summary of *MboII* digestion results.

<i>MboII</i> digestion results	Total numbers	Patient numbers	Carrier numbers
Total no. of samples	245	238	7
Total no. of <i>MboII</i> digestions	227	220	7
170 and 120bp <i>MboII</i> bands only	222	215	7
Other sizes of <i>MboII</i> bands	5	5	0
Confirmed insertion in 3' flanking region	1	1	0
Confirmed deletion of 3' flanking region	1	1	0
Unknown change	3	3	0

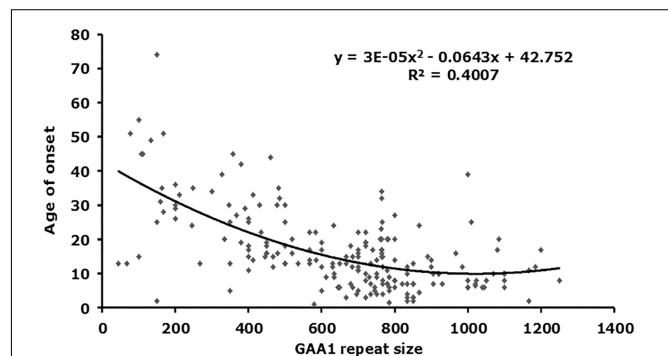


**FIGURE 1 |** *MboII* digest results. Agarose gel showing *MboII* digests of GAA PCR products of FRDA samples. The expected 170bp (5') and 120bp (3') undigested GAA-flanking fragments from normal pure GAA repeat expansion FRDA samples are shown in lanes 2, 3, and 4. These band sizes can be seen in between the 200 and 100bp fragments of the 1 Kb+ DNA ladder markers, which are loaded into lanes 1 and 11 of the gel. Lane 5 shows a large *MboII* band of approximately 600bp that was obtained from the positive interrupted GAA repeat sequence from the "NEP" BAC transgenic mouse that contains approximately 500 triplet repeats with the previously determined interrupted sequence of  $(GAA)_{21}(GGAGAA)_5(GGAGGAGAA)_{70}(GAA)_n$  (Holloway et al., 2011). In addition for this positive sample, we also identified the expected 5' flanking band of 170bp, together with a smaller band of less than 100bp that we sequenced and we showed to contain a 27bp deletion in the 3' flanking region. Lane 6 shows an abnormal band of 200bp representing the 80bp duplication in the 3' GAA flanking region. Lane 7 shows an abnormal band of approximately 100bp representing the 19bp deletion in the 3' GAA flanking region. Lanes 8, 9, and 10 contain abnormal bands of approximately 300, 100, and 180bp, respectively, that are likely to contain a region of interrupted GAA repeat sequence within the body of one or other of the large FRDA GAA repeat expansions.

eight sample:  $\dots(GAA)_{26}A(GAA)_4G(GAA)_2$  and lane 10 sample:  $\dots(GAA)_4GAG(GAA)_5$ . These interruptions were all within the last seven repeats at the 3' end of the GAA repeat sequence, in a similar location and at a similar frequency to the small 3' GAA repeat interruptions that were detected in 5 out of 11 (i.e., 45%) FRDA samples as previously reported by Sakamoto et al. (2001).

### Confirmation of the Inverse Correlation Between GAA Repeat Size and Age of Disease Onset

From our starting cohort of 245 samples, we identified 199 homozygous GAA repeat expansion FRDA patient samples for which data were available on age of onset in addition to both GAA repeat allele sizes (Supplementary Table S1). In agreement with previous studies (Durr et al., 1996; Filla et al., 1996; Reetz et al., 2015), our analysis of all 199 samples revealed a strong inverse correlation between the size of the GAA1 allele and age of onset, with a Pearson's correlation coefficient,  $R = -0.58$ , and  $R^2 = 0.34$ , using a linear model, and an even stronger correlation,  $R = -0.63$ , and  $R^2 = 0.40$ , using a quadratic model (Figure 2 and Table 2). This would suggest that up to 40% of the variation in age of onset in our sample collection is determined by GAA1 repeat expansion length. The correlations for either the larger GAA2 allele repeat size or the average GAA repeat size with age of onset were also significant, but to a lesser degree ( $R^2 = 0.30$  and 0.12, respectively, using a quadratic model, Table 2). The five samples that showed abnormal *MboII* band profiles came from FRDA patients with ages of onset ranging from 5 to 45 years and, therefore, did not reveal any unusual age of onset subgrouping.



**FIGURE 2 |** Correlation analysis of GAA repeat size with age of onset. A graph representing GAA1 repeat size v age of onset is shown ( $n = 199$ ). The best fit was for a quadratic model, for which the equations and  $R^2$ -values are shown (ANOVA  $P$ -values are listed in Table 2).

**TABLE 2 |** Summary of GAA repeat size versus age of onset correlation and regression analysis.

Samples	$R$ (linear)	$R^2$ (linear)	$R$ (quadratic)	$R^2$ (quadratic)	$P$ -value	$n$
All GAA1	-0.58	0.34	-0.63	0.40	<0.0001	199
All GAA2	-0.28	0.08	-0.35	0.12	<0.001	199
All average GAA	-0.51	0.26	-0.55	0.30	<0.0001	199

### Analysis of GAA Repeat Expansion Dynamics

GAA repeat expansions are known to be dynamic entities, exhibiting both intergenerational and somatic instability of repeat

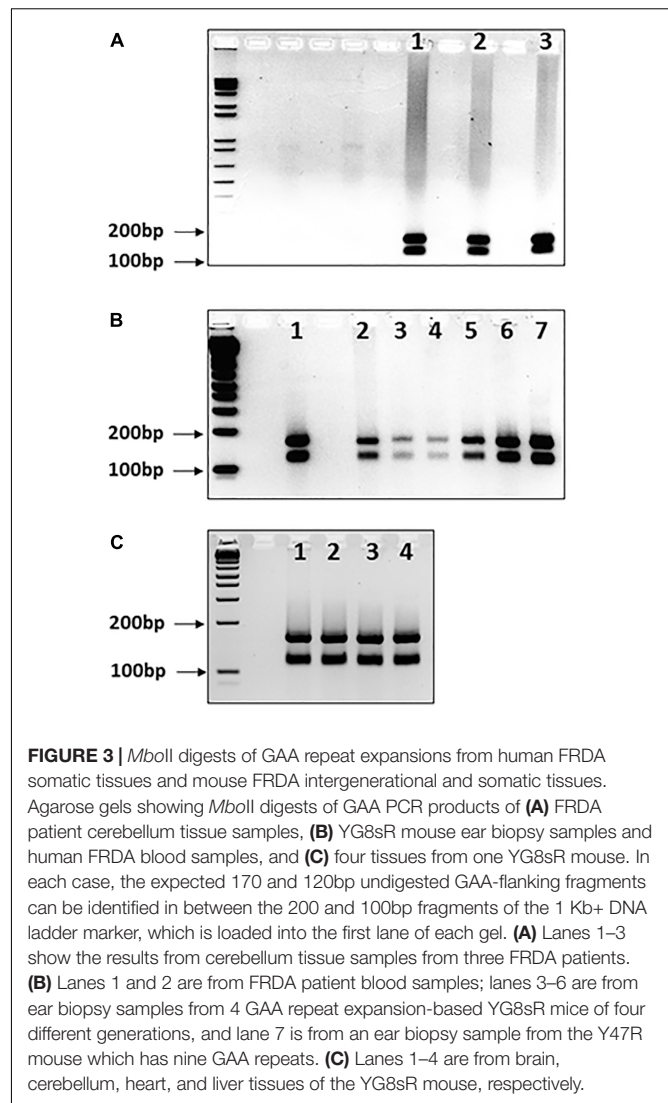
lengths. However, it is possible that GAA repeat expansions may also vary at either intergenerational or somatic levels according to sequence content (pure GAA or interrupted repeat sequence). To address potential intergenerational repeat changes, we identified four examples of carrier parents and FRDA-affected offspring within our collection of FRDA patient and carrier DNA samples (**Supplementary Table S1**) and examined these samples by *MboII* digestion analysis. We found that for all four examples only the expected two *MboII* bands of 170 and 120bp were produced in both parents and offspring, indicating primarily pure GAA repeats. This indicates that no gross GAA repeat interruptions detectable by *MboII* digestion occurred across generations.

Another possible occurrence is that pure GAA sequences may develop interruptions in different somatic tissues. We addressed this issue by determining the *MboII* digestion profiles from cerebellum autopsy tissue samples from three FRDA patients, since cerebellum tissue is known exhibit a high degree of GAA repeat instability in FRDA (Al-Mahdawi et al., 2004). Pure GAA repeat expansions were identified for each sample, indicating that, at least in these cases, there was no evidence of tissue-specific differences of large GAA repeat interruptions that could be detected by *MboII* analysis (**Figure 3A**, lanes 1–3).

Our data obtained from human samples is supported by similar *MboII* digestion studies of GAA repeat expansion-based FRDA YAC transgenic mice. We analyzed DNA samples from four generations of YG8sR FRDA mice, which have previously been shown to contain 120 pure GAA repeats (Anjomani Virmouni et al., 2015). The expected two *MboII* bands of 170 and 120bp were obtained for all samples, indicating intergenerational transmission of pure GAA repeats as detected by *MboII* analysis (**Figure 3B**, lanes 3–6). Furthermore, the expected two *MboII* bands of 170 and 120bp were obtained for samples of brain, cerebellum, heart, and liver tissues from the same 1-year-old YG8sR FRDA mouse, indicating somatic stability of pure GAA repeats as detected by *MboII* analysis (**Figure 3C**, lanes 1–4).

## DISCUSSION

To our knowledge, there has been no large-scale study of FRDA patient DNA samples to determine the frequency of large interruptions in GAA repeat expansions. Therefore, we aimed to address the question, “Do most FRDA patients have mainly pure GAA repeat expansions?” by performing long-range GAA repeat PCR amplification and *MboII* restriction enzyme digestion upon a large panel of 245 FRDA patient and carrier DNA samples. We found that the majority of GAA repeats from FRDA patient and carrier samples do not show significant extensive sequence changes that would result in abnormal *MboII* profiles, indicative of primarily pure GAA repeats throughout most of the repeat length. So the answer to our question is, yes, most FRDA patients have mainly pure GAA repeat expansions. However, we also confirmed the previous finding by Sakamoto et al. (2001) that a significant number of FRDA samples contain small sequence interruptions localized to the 3' end of the GAA repeats. These small localized repeat interruptions would not be detected by the *MboII* enzymatic digestion method and they would not be



expected to have any major impact upon GAA repeat expansion dynamics in the way that a large *MboII*-detected interruption may have.

We recognize that there are some limitations to the *MboII* digestion method of analysis. Firstly, *MboII* digestion does not identify sequence at the base pair level. Sequencing of long homozygous GAA repeats, which can be up to 1,700 repeats in size, is technically still too difficult to perform. However, sequencing of small GAA repeat alleles is possible after separation from larger GAA repeat alleles by gel purification techniques. In future, technologies such as GAA TP-PCR and single molecule sequencing may prove useful for more detailed analysis of GAA repeat interruptions in FRDA. TP-PCR studies may allow the detection of non-GAA interruptions within approximately the first 100 repeats at the 5' or 3' ends of GAA repeat expansions (Ciotti et al., 2004), although again TP-PCR will not identify sequence at the base pair level. Alternatively, single molecule sequencing has been used successfully for the identification of AGG interruptions in CGG repeats of fragile X syndrome

premutation individuals (Ardui et al., 2017). Secondly, the *MboII* method will only detect approximately 20bp added to the 170 or 120bp flanking regions or >50bp of internal interruptions in the GAA repeat expansion, but will not detect smaller interruptions or (GAAGA)<sub>n</sub> interruptions because these will be cut by *MboII*. Furthermore, care must be taken to exclude possible heteroduplex formation, which may cause difficulties in interpretation of *MboII* digests by producing an apparent third *MboII* band that is not real.

Overall, our findings indicate that, since most FRDA patients appear to carry primarily pure GAA repeat expansions throughout most of the length of the repeat, they are likely to exhibit a classical FRDA phenotype, showing well established decreases in frataxin gene expression due to epigenetic changes, R loop and heterochromatin formation, with consequential mitochondrial dysfunction, oxidative stress, and cell death, especially of large sensory neurons (Campuzano et al., 1996; Saveliev et al., 2003; Groh et al., 2014; Abeti et al., 2015, 2016, 2018a,b). Pure GAA repeat expansions are also likely to be unstable, increasing in length in aging somatic tissues, particularly in expansion-susceptible tissues such as the dorsal root ganglia and cerebellum (De Biase et al., 2007), and thereby increasing the severity of FRDA disease progression. In contrast, the lesser number of FRDA patients who may be shown to have *MboII*-detected GAA repeat interruptions may have a less progressive frataxin deficit within their cells and thus a less severe FRDA disease progression, which may suggest a prognosis of milder disease. Due to the unstable nature of pure GAA repeat expansions, individuals with pure GAA repeat expansions are likely to transmit even larger GAA repeats to any offspring. However, this may not be the case for individuals with significant interruptions in their GAA repeat expansions, since a (GAAAGAA)<sub>n</sub> interrupted 112 GAA repeat has been reported to be stably transmitted through two generations (Cossee et al., 1997). Understanding whether potential FRDA carriers in the population have either pure GAA repeats or large interrupted GAA repeats may also be of significance, since large interrupted GAA repeats may block the appearance of pathologically sized GAA repeat expansions.

Our findings concerning the paucity of large GAA repeat interruptions in FRDA also add to the understanding of trinucleotide repeat disorders in general. It is known that interruptions within trinucleotide repeat sequences, such as the (CAG)<sub>n</sub> *ATXN1* repeat, (CTG)<sub>n</sub> *DM1* repeat, and (CGG)<sub>n</sub> *FMR1* repeat, can make them less prone to expansion. Studies in yeast have shown that such interruptions in (CTG)<sub>n</sub> repeats reduce hairpin formation during DNA replication and thus inhibit expansion rates (Rolfmeier and Lahue, 2000). Other *in vitro* studies of (CAG)<sub>n</sub> and (CGG)<sub>n</sub> repeats have demonstrated inhibitory effects of interruptions on the formation of slipped strand DNA structures, thereby reducing their instability (Pearson et al., 1998). Whatever the exact mechanisms, the interplay between trinucleotide repeat interruptions and stability may have important consequences on related genetic disease phenotype. Thus, the importance of interruptions in (CAG)<sub>n</sub> repeat expansions with regard to disease pathology has previously been reported for spinocerebellar ataxia type 1 (SCA1)

(Menon et al., 2013). Interruptions in the (ATTCT)<sub>n</sub> repeat have also been shown to modify the SCA10 disease phenotype (Matsuura et al., 2006). To this list, we can now add our findings in FRDA, where large interruptions in the (GAA)<sub>n</sub> repeat are very rare and therefore are unlikely to significantly modify the age of onset FRDA disease phenotype on an overall cohort basis, but may have significant effects for individual FRDA cases.

Further insights into the relevance of interruptions of trinucleotide repeats and associated genetic diseases may be gained by studying suitable mouse models. For example, the comparison of FRDA and Huntington disease (HD) transgenic and knock-in mice also sheds some light on our understanding of the importance of trinucleotide repeat sequence content, somatic instability, and disease phenotype. Both CAG transgenic and CAG knock-in mice can give rise to HD-like phenotypes. However, both of the HD transgenic mice, BACHD and YAC128, contain interrupted repeats (Pouladi et al., 2012), which in both cases leads to complete somatic stability, whereas the three HD knock-in mice (Shelbourne et al., 1999; Wheeler et al., 1999; Ishiguro et al., 2001) have pure CAG repeats and extensive somatic instability. Further studies of HD transgenic mice with pure CAG repeats and knock-in mice with interrupted repeats should help to sort out the underlying mechanisms here. The same is true for investigations of the GAA repeats of FRDA mouse models. YG8R and KIKO mice have pure GAA repeats and extensive somatic and intergenerational GAA repeat instability with a bias toward expansion, whereas the BAC transgenic mice have extensive interruptions in the GAA repeat sequence and show no somatic instability (Al-Mahdawi et al., 2004). However, this does not translate directly into a significant difference in disease phenotype since all three FRDA mouse models exhibit similarly mild FRDA-like phenotypes (Miranda et al., 2002; Al-Mahdawi et al., 2006). Our reported finding now that YG8sR FRDA YAC transgenic mice contain pure GAA repeats that remain uninterrupted throughout intergenerational transmission and across different somatic tissues means that such mice will be liable to further expansion of the GAA repeat size across generations and within tissues throughout age. It will be interesting to see how increasing GAA repeat sizes correlate with FRDA-like disease severity in these mice.

## CONCLUSION

We have shown that major interruptions of GAA repeat expansions are very rare, but when they do occur, they may be of importance to FRDA disease progression, in a similar manner to interruptions in other trinucleotide repeat sequences and their associated genetic diseases. In contrast, we have confirmed that small GAA repeat sequence interruptions occur more frequently at the 3' end of the GAA repeat sequence, but such small interruptions are not expected to have such a major impact on GAA repeat expansion dynamics. Ultimately, future studies that enable GAA repeat expansion sequence verification at the base pair resolution are needed to allow further in depth conclusions to be made regarding the importance of GAA repeat interruptions and FRDA disease.

## AUTHOR CONTRIBUTIONS

SA-M, HG, AB, FC, VLC, SC, and MP performed the experiments. SA-M, HG, AB, PG, and MP conceived and designed the study. SA-M, HG, PG, and MP wrote the manuscript. All authors read and approved the manuscript.

## FUNDING

This work was supported by funding from the European Union Seventh Framework Programme [FP7/2007–2013] under grant agreement number 242193/EFACTS.

## REFERENCES

- Abeti, R., Baccaro, A., Esteras, N., and Giunti, P. (2018a). Novel Nrf2-inducer prevents mitochondrial defects and oxidative stress in Friedreich's ataxia models. *Front. Cell. Neurosci.* 12:188. doi: 10.3389/fncel.2018.00188
- Abeti, R., Brown, A. F., Maiolino, M., Patel, S., and Giunti, P. (2018b). Calcium deregulation: novel insights to understand Friedreich's ataxia pathophysiology. *Front. Cell. Neurosci.* 12:264. doi: 10.3389/fncel.2018.00264
- Abeti, R., Parkinson, M. H., Hargreaves, I. P., Angelova, P. R., Sandi, C., Pook, M. A., et al. (2016). 'Mitochondrial energy imbalance and lipid peroxidation cause cell death in Friedreich's ataxia'. *Cell Death Dis.* 7:e2237. doi: 10.1038/cddis.2016.111
- Abeti, R., Uzun, E., Renganathan, I., Honda, T., Pook, M. A., and Giunti, P. (2015). Targeting lipid peroxidation and mitochondrial imbalance in Friedreich's ataxia. *Pharmacol. Res.* 99, 344–350. doi: 10.1016/j.phrs.2015.05.015
- 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. doi: 10.1016/j.ygeno.2004.04.003
- Al-Mahdawi, S., Pinto, R. M., Varshney, D., Lawrence, L., Lowrie, M. B., Hughes, S., et al. (2006). GAA repeat expansion mutation mouse models of Friedreich ataxia exhibit oxidative stress leading to progressive neuronal and cardiac pathology. *Genomics* 88, 580–590. doi: 10.1016/j.ygeno.2006.06.015
- Anjomani Virmouni, S., Ezzatizadeh, V., Sandi, C., Sandi, M., Al-Mahdawi, S., Chutake, Y., et al. (2015). A novel GAA-repeat-expansion-based mouse model of Friedreich's ataxia. *Dis. Model Mech.* 8, 225–235. doi: 10.1242/dmm.018952
- Ardui, S., Race, V., Zablotskaya, A., Hestand, M. S., Van Esch, H., Devriendt, K., et al. (2017). Detecting AGG interruptions in male and female FMR1 premutation carriers by single-molecule sequencing. *Hum. Mutat.* 38, 324–331. doi: 10.1002/humu.23150
- Campuzano, V., Montermini, L., Lutz, Y., Cova, L., Hindelang, C., Jiralerspong, S., 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
- Campuzano, V., Montermini, L., Molto, M. D., Pianese, L., Cossee, M., Cavalcanti, F., 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
- Ciotti, P., Di Maria, E., Bellone, E., Ajmar, F., and Mandich, P. (2004). Triplet repeat primed PCR (TP PCR) in molecular diagnostic testing for Friedreich ataxia. *J. Mol. Diagn.* 6, 285–289. doi: 10.1016/S1525-1578(10)60523-5
- Cossee, M., Schmitt, M., Campuzano, V., Reutenauer, L., Moutou, C., Mandel, J. L., et al. (1997). Evolution of the Friedreich's ataxia trinucleotide repeat expansion: founder effect and premutations. *Proc. Natl. Acad. Sci. U.S.A.* 94, 7452–7457. doi: 10.1073/pnas.94.14.7452
- De Biase, I., Rasmussen, A., Endres, D., Al-Mahdawi, S., Monticelli, A., Coccozza, S., 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

## ACKNOWLEDGMENTS

We would like to thank Massimo Pandolfo and Myriam Rai (Université Libre de Bruxelles, Hôpital Erasme Laboratory of Experimental Neurology) for EFACTS DNA sample data, and Joseph Sarsero for “NEP” (GAA) 500 FXN BAC transgenic mouse DNA sample data.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncel.2018.00443/full#supplementary-material>

- Durr, A., Cossee, M., Agid, Y., Campuzano, V., Mignard, C., Penet, 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
- Epplen, C., Epplen, J. T., Frank, G., Mitterski, B., Santos, E. J., and Schols, L. (1997). Differential stability of the (GAA)<sub>n</sub> tract in the Friedreich ataxia (STM7) gene. *Hum. Genet.* 99, 834–836. doi: 10.1007/s004390050458
- Fillia, A., De Michele, G., Cavalcanti, F., Pianese, L., Monticelli, A., Campanella, G., et al. (1996). The relationship between trinucleotide (GAA) repeat length and clinical features in Friedreich ataxia. *Am. J. Hum. Genet.* 59, 554–560.
- Groh, M., Lufino, M. M., Wade-Martins, R., and Gromak, N. (2014). R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. *PLoS Genet.* 10:e1004318. doi: 10.1371/journal.pgen.1004318
- Holloway, T. P., Rowley, S. M., Delatycki, M. B., and Sarsero, J. P. (2011). Detection of interruptions in the GAA trinucleotide repeat expansion in the FXN gene of Friedreich ataxia. *Biotechniques* 50, 182–186. doi: 10.2144/000113615
- Ishiguro, H., Yamada, K., Sawada, H., Nishii, K., Ichino, N., Sawada, M., et al. (2001). Age-dependent and tissue-specific CAG repeat instability occurs in mouse knock-in for a mutant Huntington's disease gene. *J. Neurosci. Res.* 65, 289–297. doi: 10.1002/jnr.1153
- Matsuura, T., Fang, P., Pearson, C. E., Jayakar, P., Ashizawa, T., Roa, B. B., et al. (2006). Interruptions in the expanded ATTCT repeat of spinocerebellar ataxia type 10: repeat purity as a disease modifier? *Am. J. Hum. Genet.* 78, 125–129. doi: 10.1086/498654
- McDaniel, D. O., Keats, B., Vedanarayanan, V. V., and Subramony, S. H. (2001). Sequence variation in GAA repeat expansions may cause differential phenotype display in Friedreich's ataxia. *Mov. Disord.* 16, 1153–1158. doi: 10.1002/mds.1210
- Menon, R. P., Nethisinghe, S., Faggiano, S., Vannocci, T., Rezaei, H., Pemble, S., et al. (2013). The role of interruptions in polyQ in the pathology of SCA1. *PLoS Genet.* 9:e1003648. doi: 10.1371/journal.pgen.1003648
- Miranda, C. J., Santos, M. M., Ohshima, K., Smith, J., Li, L., Bunting, M., et al. (2002). Frataxin knockin mouse. *FEBS Lett.* 512, 291–297. doi: 10.1016/S0014-5793(02)02251-2
- Montermini, L., Andermann, E., Labuda, M., Richter, A., Pandolfo, M., Cavalcanti, F., et al. (1997). The Friedreich ataxia GAA triplet repeat: premutation and normal alleles. *Hum. Mol. Genet.* 6, 1261–1266. doi: 10.1093/hmg/6.8.1261
- Moseley, M. L., Benzow, K. A., Schut, L. J., Bird, T. D., Gomez, C. M., Barkhaus, P. E., et al. (1998). Incidence of dominant spinocerebellar and Friedreich triplet repeats among 361 ataxia families. *Neurology* 51, 1666–1671. doi: 10.1212/WNL.51.6.1666
- Ohshima, K., Sakamoto, N., Labuda, M., Poirier, J., Moseley, M. L., Montermini, L., et al. (1999). A nonpathogenic GAAGGA repeat in the Friedreich gene: implications for pathogenesis. *Neurology* 53, 1854–1857. doi: 10.1212/WNL.53.8.1854
- Pandolfo, M. (2009). Friedreich ataxia: the clinical picture. *J. Neurol.* 256(Suppl. 1), 3–8. doi: 10.1007/s00415-009-1002-3

- Pearson, C. E., Eichler, E. E., Lorenzetti, D., Kramer, S. F., Zoghbi, H. Y., Nelson, D. L., et al. (1998). Interruptions in the triplet repeats of SCA1 and FRAXA reduce the propensity and complexity of slipped strand DNA (S-DNA) formation. *Biochemistry* 37, 2701–2708. doi: 10.1021/bi972546c
- Pollard, L. M., Sharma, R., Gomez, M., Shah, S., Delatycki, M. B., Pianese, L., et al. (2004). Replication-mediated instability of the GAA triplet repeat mutation in Friedreich ataxia. *Nucleic Acids Res.* 32, 5962–5971. doi: 10.1093/nar/gkh933
- Pouladi, M. A., Stanek, L. M., Xie, Y., Franciosi, S., Southwell, A. L., Deng, Y., et al. (2012). Marked differences in neurochemistry and aggregates despite similar behavioural and neuropathological features of Huntington disease in the full-length BACHD and YAC128 mice. *Hum. Mol. Genet.* 21, 2219–2232. doi: 10.1093/hmg/dds037
- Reetz, K., Dogan, I., Costa, A. S., Dafotakis, M., Fedosov, K., Giunti, P., et al. (2015). Biological and clinical characteristics of the European Friedreich's ataxia consortium for translational studies (EFACTS) cohort: a cross-sectional analysis of baseline data. *Lancet Neurol.* 14, 174–182. doi: 10.1016/S1474-4422(14)70321-7
- Rolfmeier, M. L., and Lahue, R. S. (2000). Stabilizing effects of interruptions on trinucleotide repeat expansions in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 20, 173–180. doi: 10.1128/MCB.20.1.173-180.2000
- Sakamoto, N., Larson, J. E., Iyer, R. R., Montermini, L., Pandolfo, M., and Wells, R. D. (2001). GGA\*TCC-interrupted triplets in long GAA\*TTC repeats inhibit the formation of triplex and sticky DNA structures, alleviate transcription inhibition, and reduce genetic instabilities. *J. Biol. Chem.* 276, 27178–27187. doi: 10.1074/jbc.M101852200
- Saveliev, A., Everett, C., Sharpe, T., Webster, Z., and Festenstein, R. (2003). DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing. *Nature* 422, 909–913. doi: 10.1038/nature01596
- Sharma, R., De Biase, I., Gomez, M., Delatycki, M. B., Ashizawa, T., and Bidichandani, S. I. (2004). Friedreich ataxia in carriers of unstable borderline GAA triplet-repeat alleles. *Ann. Neurol.* 56, 898–901. doi: 10.1002/ana.20333
- Shelbourne, P. F., Killeen, N., Hevner, R. F., Johnston, H. M., Tecott, L., Lewandoski, M., et al. (1999). A Huntington's disease CAG expansion at the murine Hdh locus is unstable and associated with behavioural abnormalities in mice. *Hum. Mol. Genet.* 8, 763–774. doi: 10.1093/hmg/8.5.763
- Stolle, C. A., Frackelton, E. C., McCallum, J., Farmer, J. M., Tsou, A., Wilson, R. B., et al. (2008). Novel, complex interruptions of the GAA repeat in small, expanded alleles of two affected siblings with late-onset Friedreich ataxia. *Mov. Disord.* 23, 1303–1306. doi: 10.1002/mds.22012
- Wheeler, V. C., Auerbach, W., White, J. K., Srinidhi, J., Auerbach, A., Ryan, A., et al. (1999). Length-dependent gametic CAG repeat instability in the Huntington's disease knock-in mouse. *Hum. Mol. Genet.* 8, 115–122. doi: 10.1093/hmg/8.1.115

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Al-Mahdawi, Ging, Bayot, Cavalcanti, La Cognata, Cavallaro, Giunti and Pook. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.