



Repeat Interruptions Modify Age at Onset in Myotonic Dystrophy Type 1 by Stabilizing *DMPK* Expansions in Somatic Cells

Jovan Pešović^{1*}, Stojan Perić^{2,3}, Miloš Brkušanić¹, Goran Brajušković¹, Vidosava Rakočević-Stojanović^{2,3} and Dušanka Savić-Pavićević^{1*}

¹ Center for Human Molecular Genetics, Faculty of Biology, University of Belgrade, Belgrade, Serbia, ² School of Medicine, University of Belgrade, Belgrade, Serbia, ³ Neurology Clinic, Clinical Center of Serbia, Belgrade, Serbia

OPEN ACCESS

Edited by:

Jordi Pérez-Tur,
Instituto de Biomedicina de Valencia
(IBV), Spain

Reviewed by:

Rosanna Cardani,
Policlinico San Donato (IRCCS), Italy
Giovanni Meola,
Università degli Studi di Milano, Italy

*Correspondence:

Jovan Pešović
jovan.pesovic@bio.bg.ac.rs
orcid.org/0000-0002-8304-2067
Dušanka Savić-Pavićević
duska@bio.bg.ac.rs
orcid.org/0000-0002-2079-4077

Specialty section:

This article was submitted to
Genetic Disorders,
a section of the journal
Frontiers in Genetics

Received: 31 August 2018

Accepted: 15 November 2018

Published: 27 November 2018

Citation:

Pešović J, Perić S, Brkušanić M, Brajušković G, Rakočević-Stojanović V and Savić-Pavićević D (2018) Repeat Interruptions Modify Age at Onset in Myotonic Dystrophy Type 1 by Stabilizing *DMPK* Expansions in Somatic Cells. *Front. Genet.* 9:601. doi: 10.3389/fgene.2018.00601

CTG expansions in *DMPK* gene, causing myotonic dystrophy type 1 (DM1), are characterized by pronounced somatic instability. A large proportion of variability of somatic instability is explained by expansion size and patient's age at sampling, while individual-specific differences are attributed to additional factors. The age at onset is extremely variable in DM1, and inversely correlates with the expansion size and individual-specific differences in somatic instability. Three to five percent of DM1 patients carry repeat interruptions and some appear with later age at onset than expected for corresponding expansion size. Herein, we characterized somatic instability of interrupted *DMPK* expansions and the effect on age at onset in our previously described patients. Repeat-primed PCR showed stable structures of different types and patterns of repeat interruptions in blood cells over time and buccal cells. Single-molecule small-pool PCR quantification of somatic instability and mathematical modeling showed that interrupted expansions were characterized by lower level of somatic instability accompanied by slower progression over time. Mathematical modeling demonstrated that individual-specific differences in somatic instability had greater influence on age at onset in patients with interrupted expansions. Therefore, repeat interruptions have clinical importance for disease course in DM1 patients due to stabilizing effect on *DMPK* expansions in somatic cells.

Keywords: age at onset, CTG expansion, *DMPK*, myotonic dystrophy 1, repeat interruptions, somatic instability

INTRODUCTION

Myotonic dystrophy type 1 (DM1, MIM #160900) is a multisystem disease that affects skeletal muscles, heart, brain, eyes and endocrine system, and has an extremely variable age of disease onset (Harper, 2001). It is an autosomal dominant disorder caused by an expansion of 50 to a few thousands repeats in (CTG)_n tract located at the 3' untranslated region of the *DMPK* gene (MIM #605377) (Brook et al., 1992; Fu et al., 1992; Mahadevan et al., 1992). The number of CTG repeats broadly correlates with the disease severity (Harley et al., 1993; De Antonio et al., 2016).

Although CTG repeat tract at the *DMPK* locus was considered to be uninterrupted in both wild-type and expanded alleles, less than a decade ago expanded alleles containing repeat interruptions (also known as variant repeats) were identified (Musova et al., 2009; Braida et al., 2010). Since then, CCG, GGC, CTC and CAG interruptions have been described within 3' and 5' ends of *DMPK* expansions, with various patterns of CCG repeats being

the most commonly detected (Musova et al., 2009; Braida et al., 2010; Santoro et al., 2013; Botta et al., 2017; Pešović et al., 2017; Cumming et al., 2018; Tomé et al., 2018). The frequency of repeat interruptions has been estimated at 3–5% in DM1 patients of different origin (Musova et al., 2009; Braida et al., 2010; Santoro et al., 2013; Botta et al., 2017; Pešović et al., 2017; Cumming et al., 2018). Even though the majority of clinical signs in those patients were associated with DM1 phenotype, some patients appeared with milder or atypical symptoms (Musova et al., 2009; Pešović et al., 2017; Cumming et al., 2018), even atypical neurological phenotype (Braida et al., 2010), and without described congenital form. A noticeable feature of some DM1 patients with repeat interruptions described by us (Pešović et al., 2017) and others (Musova et al., 2009; Botta et al., 2017) was a later age at onset than expected for the corresponding number of CTG repeats. These findings have made repeat interruptions a potential genetic modifier of DM1 phenotype, likewise described in spinocerebellar ataxia type 10 (Matsuura et al., 2006).

Repeat interruptions are known *cis*-acting modifiers with a stabilizing effect either on wild type (Cleary and Pearson, 2003) or expanded (Moseley et al., 2000; Matsuura et al., 2006; Gao et al., 2008; Hu et al., 2017) repeat arrays associated with numerous unstable repeat expansion diseases. Therefore, one of the mechanisms by which repeat interruptions can influence the effect of DM1 causing mutation is genetic stabilization of CTG repeat array.

Accumulated data have indicated that interrupted *DMPK* expansions were prone to relatively stable intergenerational transmissions or even contractions in almost all described cases (Musova et al., 2009; Braida et al., 2010; Botta et al., 2017; Pešović et al., 2017; Tomé et al., 2018). Contrary to transmissions of uninterrupted *DMPK* expansions (Lavedan et al., 1993; Rakocević-Stojanović et al., 2005), these transmissions were independent on the sex of the transmitting parent (Pešović et al., 2017) and likely explain the absence of congenital cases (Braida et al., 2010; Tomé et al., 2018).

A pronounced instability of uninterrupted *DMPK* expansions in somatic cells leads to somatic mosaicism that is tissue-specific, age-dependent, biased toward further expansions throughout the patient's life (Monckton et al., 1995; Wong et al., 1995; Martorell et al., 1998) and contributes to progressive nature of the symptoms (Morales et al., 2012). In a small number of DM1 patients studied so far, a reduced instability of interrupted *DMPK* expansions in blood cells was described (Braida et al., 2010; Cumming et al., 2018; Tomé et al., 2018). In addition, it has been suggested that a low level of somatic instability may partially explain the mild symptoms seen in some of these patients (Cumming et al., 2018).

Quantification of somatic instability by small-pool PCR in a large group of DM1 patients with uninterrupted expansions and linear regression modeling have shown that a large proportion of variability of somatic instability is explained by the progenitor allele length and the patient's age at sampling (Morales et al., 2012). The unexplained, inter-individual variability is a heritable quantitative trait influenced by individual-specific genetic and environmental factors (Morales et al., 2012). The same study has also shown that age at onset inversely correlates with the

progenitor allele length and inter-individual variability of somatic instability. Therefore, identification of factors that contribute to individual-specific differences in the level of somatic instability may help to gain a better understanding of factors influencing variability of age at onset in DM1 patients.

Given that some of our previously genetically and clinically characterized DM1 patients with repeat interruptions appeared to have later age at onset than expected for the corresponding expansion size (Pešović et al., 2017), herein we characterized mutational dynamics of interrupted *DMPK* expansions in somatic cells and examined, for the first time, whether repeat interruptions modified age at onset by influencing the level of somatic instability.

MATERIALS AND METHODS

Subjects

The study included seven DM1 patients with interrupted *DMPK* expansions from Serbia, who gave informed consent. Six patients (DF1-1, DF1-2, DF1-3, DF2-1, DF3-1, and DF5-2) from four families have been previously genetically characterized (Pešović et al., 2017) (**Table 1**). Patient DF3-2 has been recently identified with the pattern of repeat interruptions identical to her mother (DF3-1) (**Table 1**) and an intergenerational transmission accompanied by an apparent contraction. Patients underwent a detailed clinical examination at the Neurology Clinic, Clinical Center of Serbia, Belgrade at two different time points (except DF3-2) in the period between 2013 and 2017. The first clinical examination was carried out upon the initial detection of interrupted *DMPK* expansions, and the second one approximately two and half to four years later. Clinical characteristics of patients at the first examination have been described (Pešović et al., 2017), and no significant progression of the disease was observed at the second examination. Blood samples were collected at both time points, while the buccal swab samples were obtained from five patients at their second examination (DF1-1, DF1-2, DF1-3, DF5-2, and DF5-3). Four DM1 patients with uninterrupted *DMPK* expansions served as controls (herein referred to as the control group) (**Table 1**). One of them was a sister (DF5-3) of the patient with a single, *de novo* CTC repeat interruption (DF5-2). The remaining three were unrelated to each other and to patients carrying repeat interruptions, and their blood DNA samples were available from two different time points ranging the time interval of 8–11 years. These DNA samples belong to the DNA bank that accompanies the Serbian patient registry for myotonic dystrophies. The study was approved by the Ethical Committee of the University of Belgrade, School of Medicine (number 339/5). It was performed in accordance with the ethical standards of the 1964 Declaration of Helsinki and its later amendments.

Molecular Genetic Analyses

Genomic DNA was extracted from blood and buccal swab samples using QIAamp® DNA Blood Mini Kit (Qiagen, Hilden, Germany). To determine the patterns of repeat interruptions in buccal cells and to investigate them in blood

TABLE 1 | Main features of DM1 patients analyzed in this study.

Patient ID	Sex	Interruption pattern	AO	Onset symptoms	PAL	AS t1	10th t1	AS t2	10th t2
DF1-1	F	(CTG) _n (CCGCTG) ₃ (CTG) ₄ (CCGCTG) ₂ CTGCCG(CTG) ₁₇	39.5	Hand grip myotonia	520	61.5	459	65.5	443
DF1-2	M		30.5	Hand grip myotonia	350	34.5	348	38.5	384
DF1-3	M		15.5	Hand grip myotonia	450	35.5	441	39.5	469
DF2-1	M	(CTG) _n (CCG) ₃₆ (CTG) _n CCG(CTG) ₇ CCG(CTG) ₁₂	40.5	Proximal arm and leg weakness	320	45.5	371	48	350
DF3-1	F	(CTG) _n (CCG) ₃ (CTG) ₆ (CCG) ₃ (CTG) ₇ CCG(CTG) ₈ CCG(CTG) ₈	45.5	Proximal leg weakness	240	50	266.5	52.5	289
DF3-2	F		31.5	Hand grip myotonia, leg pain	NA	NA	NA	31.5	187
DF5-2	F	(CTG) _n CTC(CTG) ₂₆	22.5	Masticatory muscles myotonia	250	26.5	235	30.5	262
DF5-3*	F	pure CTG expansion	21.5	Hand grip myotonia	300	23.5	314	27.5	422
MD70*	F		35.5	Hand grip myotonia	NA	50.5	106	61.5	146
MD179*	F		13.5	Hand grip myotonia	NA	14.5	326	23	415
MD180*	F		14.5	Hand grip myotonia	NA	22.5	400	30.5	462

AO, age at onset, self-reported by the patients in years; PAL, the progenitor allele length in the number of repeats; AS, age at sampling in years; 10th, the 10th percentile allele length in the number of repeats; t1 and t2, the first and the second sampling time point, respectively; *, analyzed patients with uninterrupted *DMPK* expansions; NA, not available. Data for interruption pattern and PAL are from our previous study (Pešović et al., 2017). PAL was estimated at t1 as a sharp lower boundary of expansion range (Monckton et al., 1995) using small-pool PCR with ~300 pg of input DNA. In this study, patients were analyzed by single-molecule small pool PCR, and we used the 10th percentile allele length as an unbiased measurement (Higham, 2013) of the progenitor allele length.

cells over time, repeat-primed PCR was performed as previously described (Pešović et al., 2017). Additional primer P5CCG (5'-AGCGTCTACTGTCTCGGCACTTGCCCGCCGCCCG-3') was used in repeat-primed PCR to determine the number of repeats in a large CCG block in DF2-1 patient.

Quantification of somatic instability of *DMPK* expanded alleles was performed by single-molecule small-pool PCR containing ~30–60 pg of DNA (5–10 genomic equivalents). PCR protocol was as previously described (Savić et al., 2002), with the exception of DF2-1, for whom a protocol developed for the detection of *FMRI* CGG expansions (Saluto et al., 2005) was used with primers 101 and 102 (Brook et al., 1992). Briefly, upon transfer to membrane, products were hybridized with (CAG)₁₂ oligonucleotide labeled at the 3' end with DIGddUTP, and visualized by chemiluminescence detection using anti-DIG alkaline phosphatase and CDP-star (Roche Life Science, Mannheim, Germany). Single-molecule small-pool PCR products were sized using Core Laboratory Image Quantification Software – CLIQS (TotalLab Ltd., Newcastle upon Tyne, United Kingdom). At least 200 alleles per sample were amplified and sized, and used to obtain allele frequency distribution. The 10th percentile of allele frequency distribution was used as an estimate for the progenitor allele length (Higham, 2013).

Statistical Analyses

To examine tissue specificity of somatic mosaicism, the allele frequency distributions in blood and buccal cells were compared in each patient by Wilcoxon–Mann–Whitney test, after testing data normality using Shapiro–Wilk test.

To characterize mutational dynamics of *DMPK* expansions in blood cells, the observed and expected levels of somatic instability were compared in each group of patients using Wilcoxon signed-rank test. The difference in the number of repeats at the 10th and the 90th percentile of the allele frequency distribution was used

as a measure of observed level of somatic instability (Morales et al., 2012; Higham, 2013). To estimate the expected level of somatic instability, the linear regression models that correlate somatic instability with the progenitor allele length and age at sampling were considered (Morales et al., 2012). The linear regression models were retested using the 10th percentile allele length of 136 patients carrying uninterrupted *DMPK* expansions (herein referred to as the reference group) (Morales et al., 2012) and patients analyzed in this study. Retesting was performed separately for the first and the second sampling time point (Supplementary Table 1). Model 8* explained 90% of variability of somatic instability at both time points, and was used to estimate the expected level of somatic instability.

To examine inter-individual variability of somatic instability according to model 8*, residual variance was considered. This residual variance reflects individual-specific differences in the level of somatic instability that is not accounted for by the progenitor allele length and age at sampling. It is attributed to individual-specific factors (genetic and environmental). Residuals were standardized to a mean of zero and a standard deviation (SD) of 1. Standardized residuals were compared between patients analyzed in this study and the reference group by Wilcoxon–Mann–Whitney test, after testing data normality using the Shapiro–Wilk test.

To characterize mutational dynamics of somatic instability in blood cells over time, the allele frequency distributions from two time points were compared in each patient by Wilcoxon–Mann–Whitney test. In addition, the observed and expected expansion size increments over time were compared in each group of patients using Wilcoxon signed-rank test. The expected expansion size increment over time was estimated using the linear regression model explaining ~56% of variability by the modal expansion size and time interval between samplings as independent variables (Martorell et al., 1998). As a large number

of alleles sized by single-molecule small-pool PCR resulted in non-normal allele frequency distributions, the mode allele length was assumed as corresponding measure of the modal expansion size. In the original study, the modal expansion size was determined at the point of highest band intensity on the autoradiograph of genomic Southern blot or at the center of the smear for very diffuse bands (Martorell et al., 1998). The observed expansion size increment over time was calculated as difference between the mode allele length at the first and the second time point.

To characterize mutational dynamics of *DMPK* expansions without interrupted part (herein referred to as the 5' end) in blood cells over time, we used the same approach as above described for the whole run of repeats. The size of the 5' ends was obtained by subtracting the corresponding interrupted part at the 3' end from all detected expanded alleles in each sample. Retested model 8** (Morales et al., 2012) for both sampling time points explained 90% of variability of somatic instability (Supplementary Table 2).

To analyze age at onset, the observed and expected age at onset were compared in each group of patients by Wilcoxon signed-rank test. The observed age at onset was self-reported by the patients (Table 1). To estimate the expected age at onset, the following linear regression models were considered (Morales et al., 2012): models 9 and 10, correlating age at onset with the progenitor allele length, and model 11, correlating age at onset with the progenitor allele length and residual variance of somatic instability obtained from model 8*. The models were retested using the 10th percentile allele length of 121 symptomatic patients from the reference group (Morales et al., 2012) and patients examined in this study. Retesting was again performed separately for the first and the second sampling time point. Retested models explained ~43% (model 9*), ~49% (model 10*) and ~55% (model 11*) of variability of age at onset (Supplementary Table 3). Models 10* and 11* were used to assess the expected age at onset.

To examine inter-individual variability of age at onset according to models 10* and 11*, standardized residuals were compared between patients analyzed in this study and the reference group using Wilcoxon–Mann–Whitney test, after testing data normality using the Shapiro–Wilk test. Residual variance in model 10* reflects individual-specific differences in age at onset that are not explained by the progenitor allele length. Residual variance in model 11* reflects individual-specific differences in age at onset that are not explained by the progenitor allele length and by inter-individual variability of somatic instability according to model 8*. To assess contribution of inter-individual variability of somatic instability, standardized residuals from model 11* were compared with standardized residuals from model 10* in each group of patients using Wilcoxon signed-rank test.

Statistical analyses were performed in the R environment for statistical computing and visualization ver 3.3.3 (R Core Team, 2016). Statistical analyses were performed for the first and the second sampling time point, where appropriate. The significance level was set at 0.05 (two-tailed *p*-value) in all tests.

RESULTS

Interrupted *DMPK* Expansions Show Tissue-Specific Somatic Instability Prone to Further Expansions

A total number of 5,701 alleles were amplified and sized by single-molecule small-pool PCR, with at least 200 alleles per sample (the dataset is given in Supplementary Table 4). 4,572 alleles were amplified from blood sampled at two time points from six DM1 patients carrying interrupted expansions and four patients with uninterrupted expansions, and at one time point from DF3-2, carrying an interrupted expansion. Additional 1,129 alleles were amplified from buccal swab samples, available from four patients with repeat interruptions and one with pure CTG repeats. A representative blot of a single-molecule small-pool PCR is shown in Figure 1.

Allele frequency distribution in all samples was skewed toward larger alleles. Comparison of the allele frequency distribution between blood and buccal cells showed statistical difference in DF1 family members carrying runs of CCGCTG hexamer (mother DF1-1, $W = 19,379$, $p = 0.004$; son DF1-2, $W = 32,252$, $p = 0.01$; son DF1-3, $W = 19,936$, $p = 0.002$) (Figure 2). The progenitor allele length was higher (50–150 repeats more) and the level of somatic instability was lower in buccal than in blood cells (Figure 2 and Supplementary Table 5). Patient DF5-2, carrying a single *de novo* CTC interruption, showed no statistical difference in the allele frequency distribution between examined tissues ($W = 27,716$, $p = 0.68$) (Figure 2 and Supplementary Table 5). The similar result was obtained for her sister (DF5-3) carrying an uninterrupted *DMPK* expansion ($W = 22,488$, $p = 0.38$) (Figure 2 and Supplementary Table 5). Therefore, *DMPK* expansions with repeat interruptions were characterized by allele length heterogeneity prone to further expansions in somatic cells. Tissue specificity of somatic instability was likely influenced by types and structure of repeat interruptions, and additionally, by factors clustering within examined families.

Repeat Interruptions Are Individual Specific Factors Contributing to Lower Level of Somatic Instability of *DMPK* Expansions

All patients with interrupted *DMPK* expansions had lower level of observed somatic instability than expected based on the progenitor allele length and age at sampling (Figure 3 and Supplementary Table 5) and the difference was statistically significant ($V = 0$, $p = 0.03$ for the first time point; $V = 0$, $p = 0.016$ for the second time point). In line with no statistical difference between the observed and expected levels of somatic instability in the reference group ($V = 5132$, $p = 0.18$), the control group also showed no difference ($V = 7$, $p = 0.625$ for both time points) (Figure 3 and Supplementary Table 5). It is of note that the patient with a single *de novo* CTC interruption (DF5-2) was characterized by a lower level of somatic instability at both time points in comparison to her sister (DF5-3) with similar expansion size and age at sampling, but without the interruption

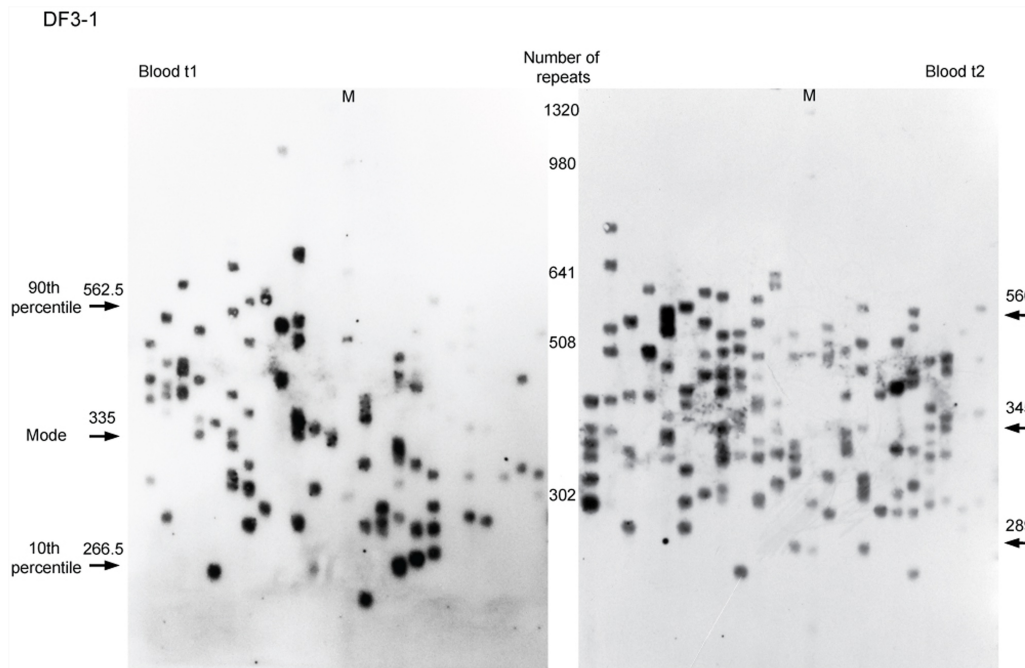


FIGURE 1 | Single-molecule small-pool PCR analysis of *DMPK* expanded alleles with repeat interruptions. Analysis of the blood sample for DF3-1 when she was 50 years old (the first time point, blood t1) and 52.5 years old (the second time point, blood t2), showing allele size heterogeneity. Each lane of agarose gel (24 cm in length) contained PCR products amplified from ~30–60 pg of DNA (5–10 genomic equivalents). Arrows with number indicate the position and size of the mode, the 10th and the 90th percentile allele length in number of trinucleotide repeats, inferred from more than 200 alleles sized per sample. M, DNA Molecular Weight Marker X (Roche Life Science, Mannheim, Germany) shown in number of CTG repeats.

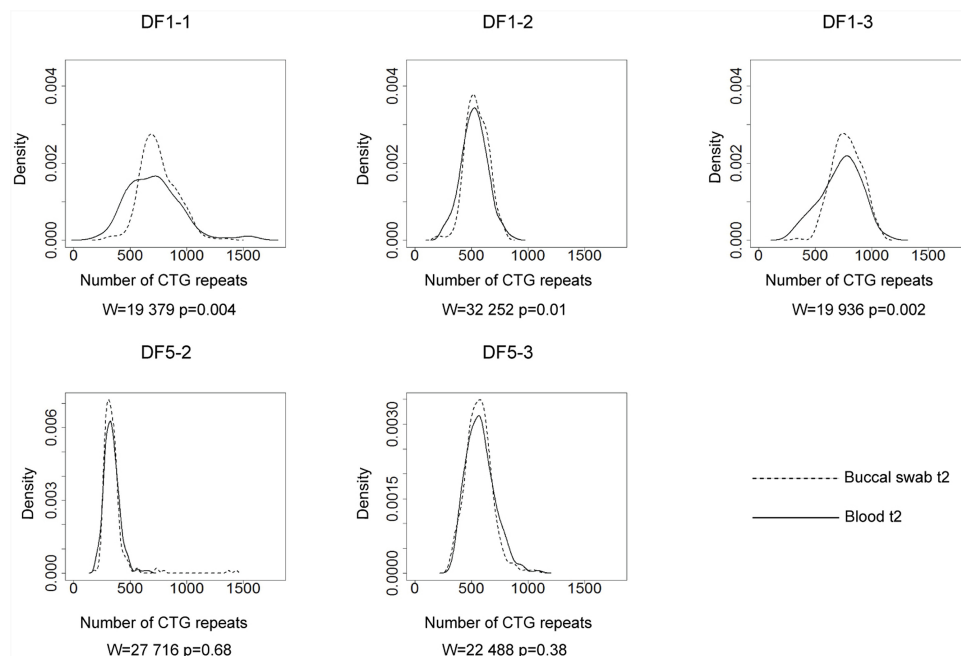


FIGURE 2 | Tissue specific somatic instability of *DMPK* expansions. Density plots show the allele frequency distributions skewed toward larger expansions in blood (solid lines) and buccal cells (dashed lines), sampled at the same time point (t2). All allele frequency distributions were derived from sizing at least 200 alleles and their comparisons were performed by Wilcoxon–Mann–Whitney test (*W* and *p*-values are shown). DF1 family members (**top graphs**), carrying identical pattern of CGG interruptions, showed difference in the allele frequency distribution between examined tissues. DF5 family members (**bottom graphs**) showed no difference in the allele frequency distribution between examined tissues. DF5-2 has a single *de novo* CTC interruption, while her sister DF5-3 has an uninterrupted *DMPK* expansion.

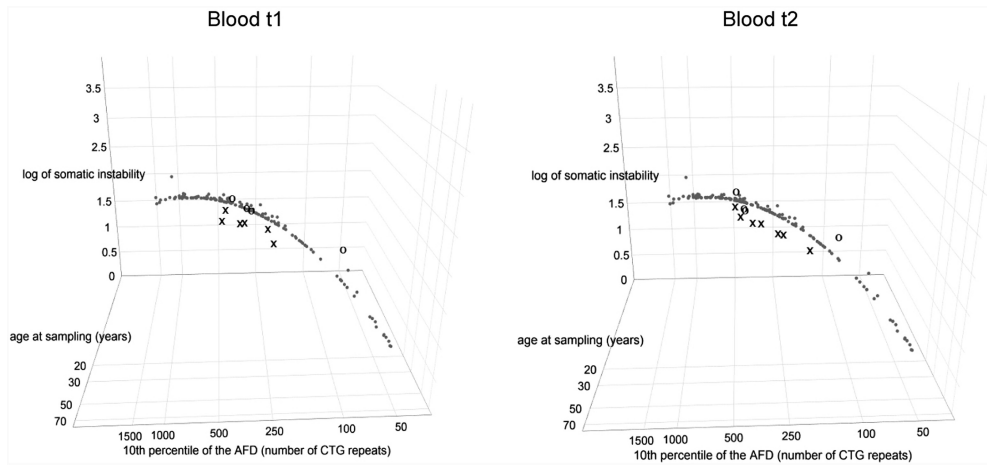


FIGURE 3 | DM1 patients with repeat interruptions show a lower level of somatic instability. The graphs represent the correlation of the level of somatic instability (SI), the 10th percentile (10^{th}) allele length (as an estimate of the progenitor allele length (Higham, 2013)), and age at sampling (AS) according to model 8*. The analyzed group of DM1 patients included our patients with interrupted expansions, our patients with pure expansions (the control group), and 136 patients from the reference group (Morales et al., 2012). Observed level of SI in blood was estimated as the range between the 10th and the 90th percentile of the allele frequency distribution (Morales et al., 2012). Patients with interrupted expansions are shown in X and the control group is shown in O, at the first time point (blood t1) and the second time point (blood t2) (2.5–4 years later for patients with interrupted expansions and 4–11 years later for the control group). Patients with interrupted *DMPK* expansions showed a large residual variance, and the median of their standardized residuals of SI was significantly lower relative to the reference group. The control group showed no difference. AFD, allele frequency distribution.

(Figure 4 and Supplementary Table 5). These results indicated that interrupted *DMPK* expansions were more stable relative to expansions with pure CTG repeats in blood cells and this feature was maintained over time.

Patients with interrupted *DMPK* expansions showed a noticeably larger residual variance of somatic instability according to model 8*, and median of standardized residuals was statistically lower in patients with interrupted *DMPK* expansions relative to the reference group ($W = 727$, $p = 1.2e^{-3}$ for the first time point and $W = 864$, $p = 2.8e^{-4}$ for the second time point) (Figure 3). On the contrary, median of standardized residuals of somatic instability in the control group showed no statistical difference relative to the reference group ($W = 221$, $p = 0.53$ for the first time point and $W = 253$, $p = 0.82$ for the second time point) (Figure 3). Since the repeat interruptions were a common factor for patients showing larger individual-specific differences in the level of somatic instability (with negative standardized residuals), these results indicated that repeat interruptions could be individual-specific factors significantly contributing to reduced somatic instability in DM1 patients.

Interrupted *DMPK* Expansions Show Slower Progression of Somatic Instability Over Time

The comparison of the allele frequency distributions from two time points during the interval of two and a half to four years in DF1-1, DF2-1 and DF3-1 (Figure 4) showed no statistical change (Table 2), indicating a relative stability of interrupted *DMPK* expansions over time. On the other hand, the observed shift in the distribution toward larger allele lengths in DF1-2, DF1-3 and DF5-2 (Figure 4) was statistically significant (Table 2), indicating

that the level of somatic instability was still enough to cause a shift of allele frequency distribution toward larger expansions. As expected, the control group had statistically significant shift in the distributions toward larger allele lengths during time interval of 4–11 years (Figure 4 and Table 2).

The observed expansion size increment over time was statistically lower than expected based on the mode allele length and time interval between samplings only in patients with interrupted *DMPK* expansions ($V = 0$, $p = 0.03$), but not in the control group ($V = 0$, $p = 0.13$) (Table 2). Both sisters from the DF5 family were characterized by a shift in the distribution toward larger allele lengths during the same time interval of 4 years. However, DF5-2 with a single *de novo* CTC interruption, showed the smaller expansion size increment over time relative to DF5-3 with similar expansion size, but without interruptions (Figure 4). These results indicated that patients with interrupted *DMPK* expansions were characterized by a slower progression of somatic expansions in comparison to patients with uninterrupted expansions in blood cells.

Repeat Interruptions Are Stable in Somatic Cells and *in Cis* Stabilize the Remaining Part of *DMPK* Expansions

To examine the influence of repeat interruptions on the demonstrated lower level and slower progression of somatic instability at the *DMPK* locus, we separately characterized mutational dynamics of the 3' end, containing repeat interruptions, and the 5' end, the remaining part of expansions.

For each examined patient, repeat-primed PCR profiles of the 3' end were identical in buccal and blood cells, and in blood cells sampled at two time points (Figure 5). Blood cells at two time

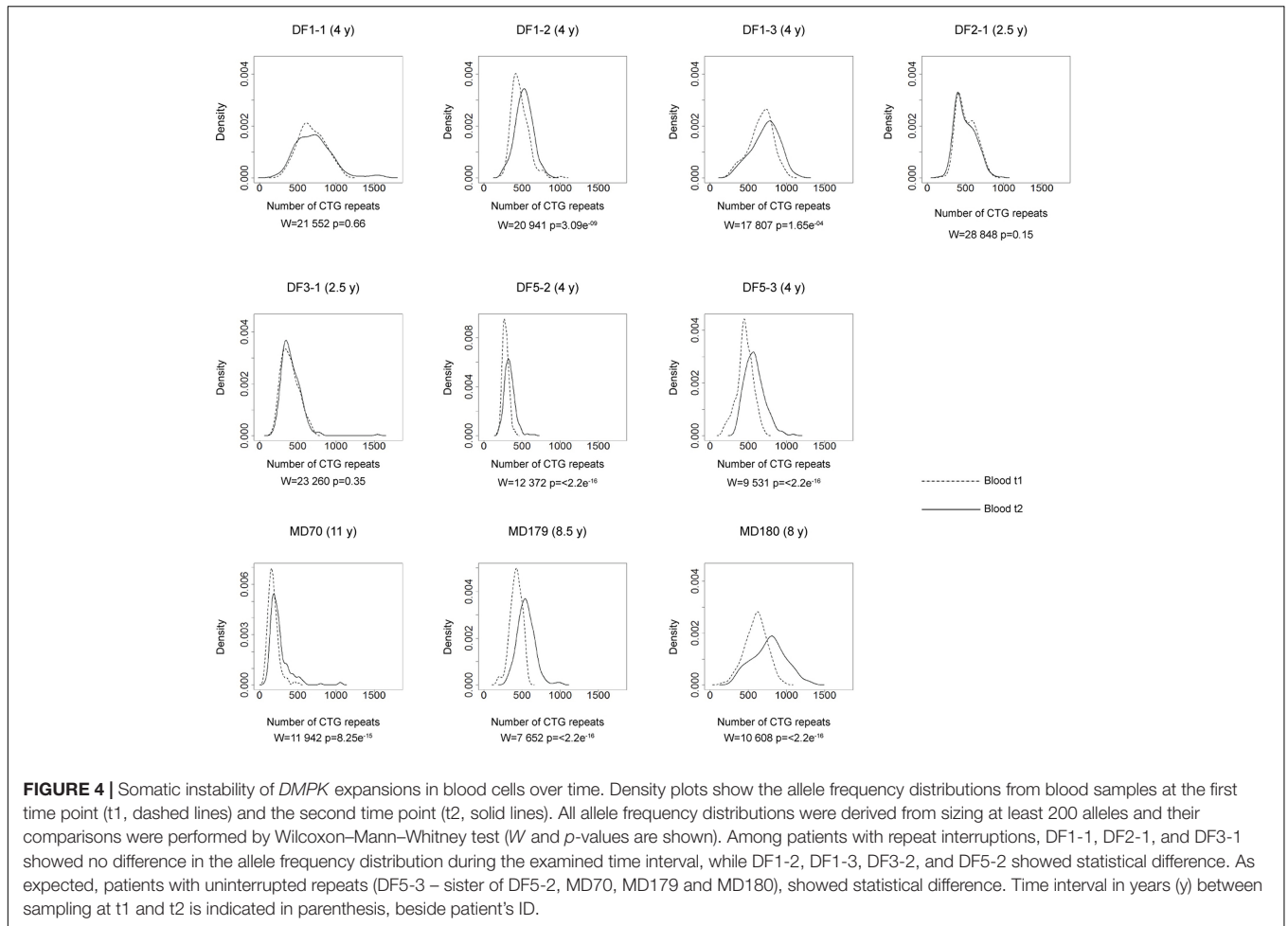


TABLE 2 | Comparisons of the allele frequency distributions of patients with interrupted and uninterrupted *DMPK* expansions obtained from blood samples at two time points, and the observed and expected expansion size increments over time.

Patient ID	Mode t1	Mode t2	Mode (5' end) t1	Mode (5' end) t2	Wilcoxon–Mann–Whitney test		Time between two samplings (years)	Observed increment	Expected increment	Expected increment (5' end)	Wilcoxon signed-rank test	
					<i>W</i>	<i>P</i> -value					<i>V</i>	<i>P</i> value
DF1-1	612	720	577	685	21,552	0.66	4	108	147	142	0	0.02
DF1-2	417	527	382	492	20,941	3.09e⁻⁰⁹	4	110	119	114		
DF1-3	732	779	697	744	17,807	1.65e⁻⁰⁴	4	47	165	160		
DF2-1	410	401	325	316	28,848	0.15	2.5	-9	62	50		
DF3-1	335	345	296	306	23,260	0.35	2.5	10	51	46		
DF5-2	267	324	238	295	12,372	<2.2e⁻¹⁶	4	57	97	93		
DF5-3*	445	565	445	565	9,531	<2.2e⁻¹⁶	4	120	123	123	0	0.125
MD70*	158	186	158	186	11,492	8.25e⁻¹⁵	11	28	340	340		
MD179*	427	540	427	540	7,652	<2.2e⁻¹⁶	8.5	113	287	287		
MD180*	622	806	622	806	10,608	<2.2e⁻¹⁶	8	184	297	297		

Mode, the mode allele length in the number of repeats; *t1* and *t2*, the first and the second time point, respectively; *W*, statistics of the Wilcoxon–Mann–Whitney test used for comparison of allele frequency distribution at two time points in each patient; The observed expansion size increment over time was calculated as a difference between the mode allele length at the first and the second time point both for the entire interrupted *DMPK* expansions and their 5' ends and had the same values. The expected expansion size increment over time was calculated using the equation of a linear regression model for uninterrupted *DMPK* expansions that shows a correlation between expansion size increment, initial expansion in the number of repeats (IE) and time in years between sampling (T): $increment = -90 + 37 \times T + 0.146 \times IE$ (Martorell et al., 1998). As a measure of IE we used the mode allele length inferred from more than 200 sized alleles in each sample. *V*, statistics of the Wilcoxon signed-rank test used for comparison of the observed and expected expansion size increment over time, as well as the observed and expected expansion size increment at the 5' end over time in each group of patients (the same values were obtained). *Analyzed patients with uninterrupted *DMPK* expansions.

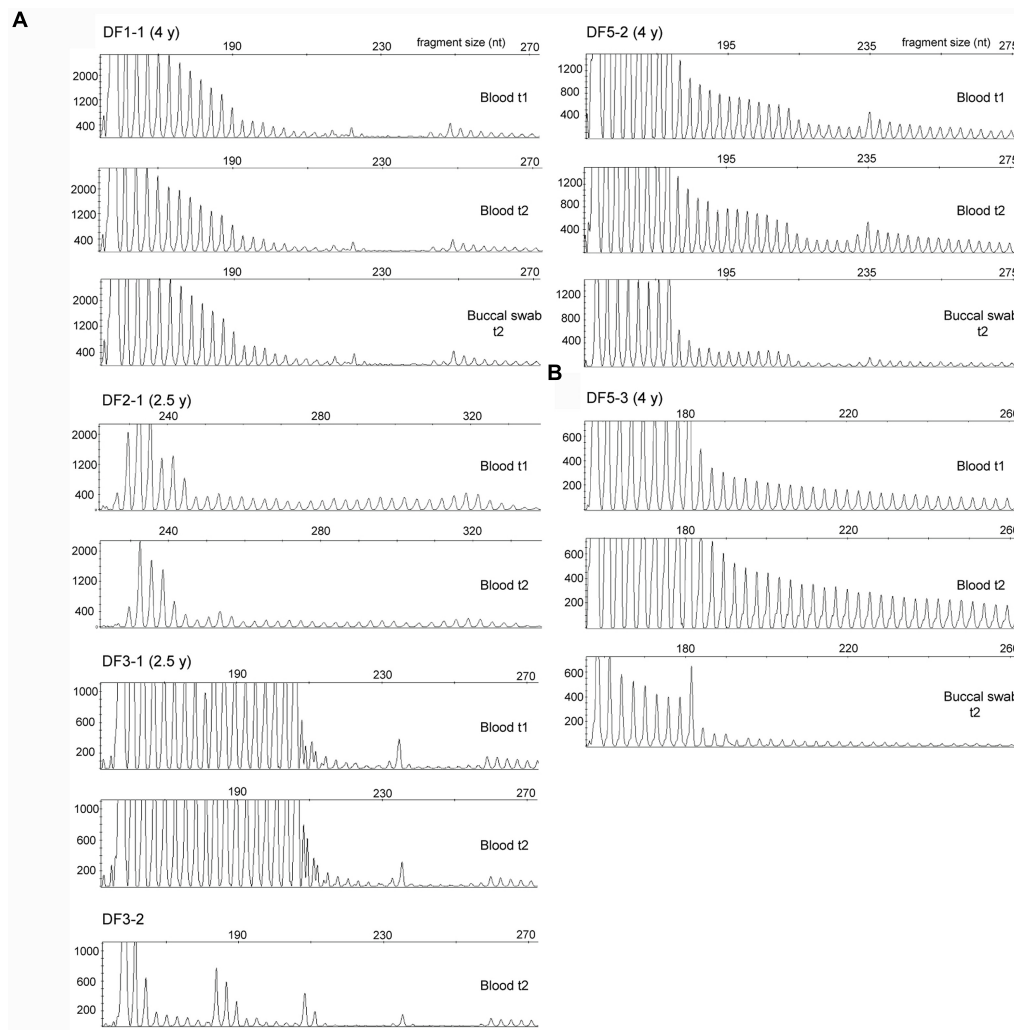


FIGURE 5 | Somatic stability of repeat interruptions at the 3' end of *DMPK* expansions between tissues and over time. **(A)** Identical repeat-primed PCR profiles in blood sample at the first (t1) and the second time point (t2) and in buccal swab sample (only for patients DF1-1 and DF5-2) in patients with interrupted *DMPK* expansions. All profiles were obtained using repeat-binding primer P4CTG-for (Pešović et al., 2017). The only exception was DF2-1, whose profiles were obtained using P5CCG primer, binding to a large CCG block containing 36 repeats. DF1-2 and DF1-3 are not shown, as they had the identical repeat-primed PCR profiles from all analyzed samples as their mother DF1-1 due to identical wild-type allele of 5 CTG repeats and the same pattern of repeat interruptions. **(B)** Repeat-primed PCR profiles in blood sample at the first time point (t1) and the second time point (t2) and in buccal swab sample in patient DF5-3 with an uninterrupted *DMPK* expansion. The profiles were obtained using repeat-binding primer P4CTG-for. Time interval in years (y) between sampling at t1 and t2 is indicated in parenthesis, beside patient's ID.

points and buccal cells were characterized by identical structure of CCGCTG hexamer runs in DF1 family, and by a single CTC repeat in DF5-2. Identical structures of small CCG blocks were identified in blood cells of DF3-1 over time. Additional characterization of repeat interruptions in DF2-1 showed the presence of previously assumed large CCG block containing 36 repeats, similar to his daughter (DF2-2) who was characterized by a block of 40 CCG repeats (Pešović et al., 2017). This interruption also showed the identical structure in blood cells over time. The obtained results demonstrated a relative somatic stability of different types and patterns of repeat interruptions at the 3' end of *DMPK* expansions in two examined tissues and throughout time in blood cells.

The 5' ends showed lower level of the observed somatic instability than expected for the corresponding progenitor allele length and age at sampling ($V = 0$, $p = 0.03$ for the first time point; $V = 0$, $p = 0.04$ for the second time point) (Supplementary Table 6). No statistical difference was observed neither in the control group, nor the reference group according to the retested model (data not shown). In line with the results for the whole run of repeats in interrupted *DMPK* expansions, the comparison of the 5' end frequency distributions at two time points showed no statistical change in DF1-1, DF2-1, and DF3-1, and a statistically significant shift toward larger allele lengths in DF1-2, DF1-3, and DF5-2 (Table 2). The 5' ends also showed statistically smaller observed increment than expected based on the mode allele

length and time interval between samplings ($V = 0$, $p = 0.04$, **Table 2**). A lower level and a slower progression of somatic instability at the 5' end of interrupted *DMPK* expansions were suggestive of a stabilizing effect of repeat interruptions not only at the 3' end, but also on the remaining part of *DMPK* expansions. Collectively, these results established repeat interruptions as *cis*-acting factors with a stabilizing effect at the *DMPK* locus in blood cells.

Repeat Interruptions Are Positive Genetic Modifiers of Age at Onset in DM1 Patients

According to model 10* (correlating age at onset with the progenitor allele length), the observed age at onset was 11–23 years higher than expected in four out of seven patients with repeat interruptions (DF1-1, DF1-2, DF2-1, and DF3-1) (**Supplementary Table 7**). Model 11* (correlating age at onset with the progenitor allele length and residual variance of somatic instability obtained from model 8*) gave a somewhat better prediction, but the observed age at onset was still higher than expected (7–20 years) in the above mentioned patients (**Supplementary Table 7**). Both models closely predicted age at onset in the control group (**Supplementary Table 7**). However, there was no statistical difference between the observed and expected age at onset neither in the group of patients with repeat interruptions, nor in the control group according to models 10* and 11* (**Supplementary Table 7**).

We considered residual variance in models 10* and 11* to examine inter-individual variability of age at onset. According to model 10*, patients with interrupted *DMPK* expansions showed a noticeably larger residual variance, and median of their standardized residuals was statistically higher than in the reference group ($W = 147$, $p = 0.01$ for the first time point and $W = 171$, $p = 0.008$ for the second time point) (**Figure 6**). The control group showed no statistical difference ($W = 289$, $p = 0.51$ for the first time point and $W = 218$, $p = 0.74$ for the second time point) (**Figure 6**). This result indicated the existence of factors with a significant contribution to individual-specific differences in age at onset in patients carrying repeat interruptions. Notably, inclusion of residual variance of somatic instability (according to model 8*) as a variable in model 11*, resulted in no statistical difference in medians of standardized residuals between patients with interrupted expansions and the reference group ($W = 207$, $p = 0.08$ for the first time point and $W = 252$, $p = 0.07$ for the second time point) (**Figure 6**). The control group again showed no statistical difference ($W = 249$, $p = 0.93$ for the first time point and $W = 204$, $p = 0.6$ for the second time point) (**Figure 6**). This result indicated that individual-specific differences in the level of somatic instability could explain individual-specific differences in age at onset, observed in patients carrying repeat interruptions according to model 10*. In addition, median of standardized residuals in age at onset in model 11* was statistically lower than in model 10* in patients with interrupted *DMPK* expansions ($V = 0$, $p = 0.03$ for the first time point and $V = 0$, $p = 0.02$ for the second time point), but not in the control group ($V = 6$, $p = 0.88$ for the first time point and $V = 6$, $p = 0.88$ for the second time

point), nor the reference group ($V = 3854$, $p = 0.67$). According to this result, individual specific differences in the level of somatic instability had a greater impact on age at onset in patients with interrupted expansions. Given that repeat interruptions significantly contributed to individual-specific differences in the level of somatic instability by stabilization of *DMPK* expansions in somatic cells, these results implied that repeat interruptions could be considered as factors with a positive modifying effect on age at onset in DM1 patients.

DISCUSSION

This study reports on the characterization of mutational dynamics of *DMPK* expanded alleles in somatic cells in DM1 patients carrying repeat interruptions and how this dynamics influences the age at onset. The novelty of our study is the finding that repeat interruptions contribute to later appearance of disease symptoms and the mechanism by which they operate is a genetic stabilization of *DMPK* expansions in somatic cells. In addition, we showed that mutational dynamics of interrupted *DMPK* expansions was tissue-specific and the structures of repeat interruptions were stable in somatic tissues over time.

We used single-molecule small-pool PCR to quantify somatic instability at the *DMPK* locus in our previously characterized DM1 patients with repeat interruptions (Pešović et al., 2017). In that study, the progenitor allele length was estimated as a sharp lower boundary of expansion range (Monckton et al., 1995), since we used small-pool PCR with ~300 pg of input DNA. This kind of estimation of the progenitor allele length was not clear and could have been biased when our patients were analyzed by single-molecule small pool PCR. Instead, we used the 10th percentile allele length as an unbiased measurement (Higham, 2013) and obtained approximately the same length of the progenitor alleles. All retested linear regression models using the 10th percentile allele length explained almost the same percent of variability of somatic instability and age at onset as the original models (Morales et al., 2012), additionally demonstrating that the 10th percentile allele length was a precise measurement of the progenitor allele length. This enabled us to use the originally studied large group of patients (Morales et al., 2012) as a valuable reference group for comparisons of somatic instability and age at onset. Since the limited number of DM1 patients was available for our study, quantification of somatic instability at two sampling time points, followed by the same results in statistical analyses, supported the reliability of our results. Testing of the control DM1 patients, who lacked repeat interruptions and were not matched by expansion size and age at sampling, showed a proper fit to the applied linear regression models (Martorell et al., 1998; Morales et al., 2012). Analysis of the control group additionally confirmed the reliability of the obtained results and showed the absence of possible population-specific differences relative to the reference group. Therefore, our approach to study somatic instability and age at onset in DM1 patients with repeat interruptions does not require a group of patients with uninterrupted expansions matched by expansion size and age at sampling. As such, it may offer an advantage for

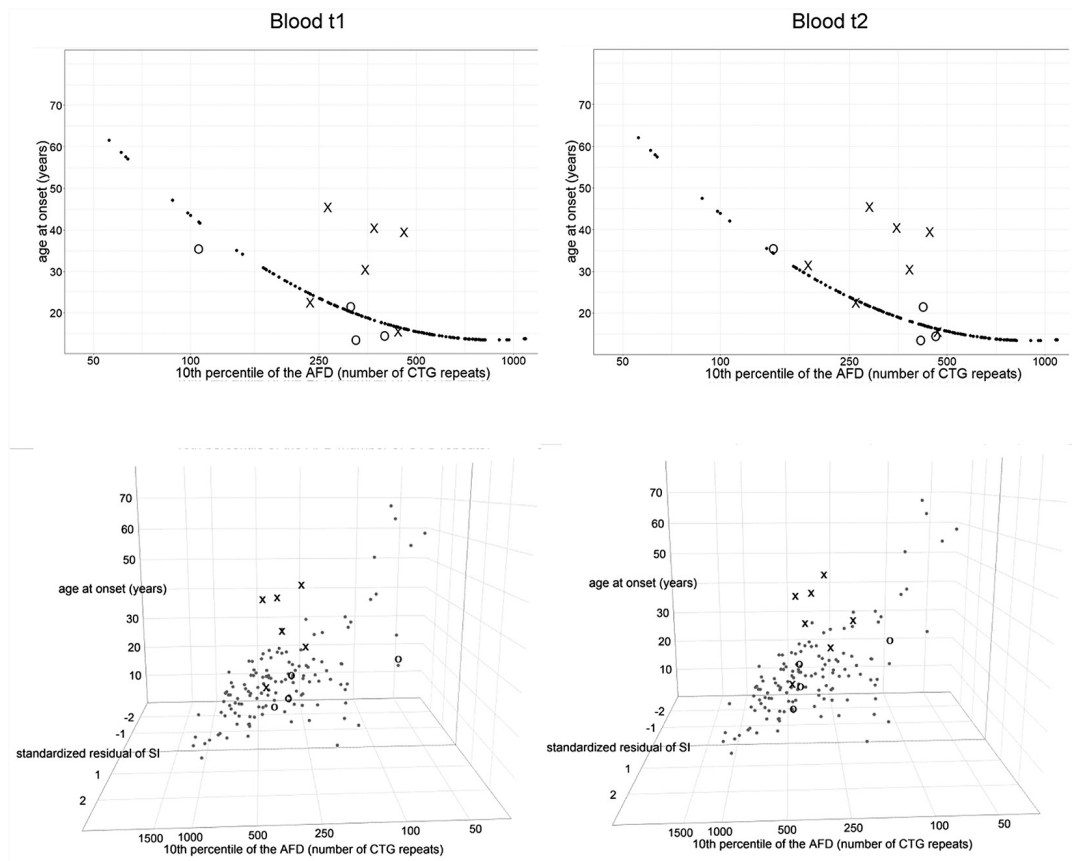


FIGURE 6 | Repeat interruptions are positive genetic modifiers of age at onset in DM1 patients. The graphs represent the correlation of age at onset (AO) with the 10th percentile (10^{th}) allele length (as an estimate of the progenitor allele length, Higham, 2013) according to model 10* (**top graphs**), and additionally with standardized residual of somatic instability (SI) according to model 11* (**bottom graphs**). The analyzed group of DM1 patients included our patients with interrupted expansions, our patients with pure expansions (the control group), and 121 symptomatic patients from the reference group (Morales et al., 2012). Observed values of AO (self-reported by patients) are shown in X in patients with interrupted expansions and in O in the control group, at the first time point (blood t1) and the second time point (blood t2) (2.5–4 years later for patients with interrupted expansions and 4–11 years later for the control group). Patients with interrupted expansions showed a large residual variance when AO was correlated with the 10th allele length (model 10*) and median of their standardized residuals was statistically higher than in the reference group (**top graphs**). The inclusion of residual variance of SI as a variable (model 11*) resulted in no statistical difference in medians of standardized residuals between patients with interrupted expansions and the reference group (**bottom graphs**). The control group showed no statistical difference according to both models. AFD, allele frequency distribution.

future studies that would enroll a larger number of patients from different countries (Wood et al., 2018).

The similar approach to study somatic instability of interrupted expansions in DM1 patients was used by Braida et al. (2010) with only difference being the estimate of the progenitor allele size as originally proposed (Monckton et al., 1995). Two recently published studies compared small-pool PCR profiles (Tomé et al., 2018) and a degree of somatic instability, measured as a difference between the progenitor and modal allele length (Cumming et al., 2018), in patients with interruptions relative to a few age matched patients with uninterrupted expansions of similar size. However, comparisons to a large group of DM1 patients with uninterrupted expansions better address variability in somatic instability seen in DM1.

Observed allele frequency distributions that are skewed toward larger allele lengths and characterized by differences between two time points in blood cells, as well as between

blood and buccal cells in each patient, indicate that somatic mosaicism of interrupted *DMPK* expansions has characteristics similar to expansions consisting of pure $(CTG)_n$ tract: a bias toward further expansions throughout the patient's life, an age related dependency and a tissue specificity (Monckton et al., 1995; Wong et al., 1995; Martorell et al., 1998). However, there is a principal difference in their mutational dynamics, because the interrupted *DMPK* expansions are more stable. In particular, we demonstrated that patients with interrupted *DMPK* expansions had different mutational dynamics in blood and buccal cells that was likely influenced by types and structure of repeat interruptions, and additionally by factors clustering within examined families. Patients with interrupted *DMPK* expansions were characterized by a lower level of allele size heterogeneity in blood cells than expected for the progenitor allele length and age at sampling. They also showed larger individual-specific differences in the level of somatic instability,

which additionally confirmed reduced level of somatic instability. These individual-specific differences can be attributed to repeat interruptions due to their stabilizing effect on *DMPK* expansions (as discussed below). The similar result was reported for members of the family with *DMPK* expansions with multiple GGC and CCG interruptions at the 3' end relative to the same reference group (Braidà et al., 2010), although the effect of other factors clustering within this family and potentially influencing variability of somatic instability (Morales et al., 2012) cannot be excluded. Factors contributing to individual-specific differences in the level of somatic instability are mostly unknown. Our results demonstrate that repeat interruptions are the individual-specific, genetic modifier of somatic instability. Among other genetic factors, the single nucleotide variant in *MSH3* mismatch repair gene is the only known *trans*-acting modifier of somatic instability in DM1 patients (Morales et al., 2016). We further showed that patients with interrupted *DMPK* expansions were characterized by the lower expansion size increment in blood cells over time than expected for the corresponding allele length and time interval between two samplings. Three examined patients were even characterized by a relative stability of interrupted *DMPK* expansions over time. Patient with a single *de novo* CTC interruption (DF5-2) showed the smaller expansion size increment over the same time interval relative to her sister with similar expansion size, but without interruptions (DF5-3). The smaller net expansion gain over time in our DM1 patients with repeat interruptions at the 3' end is in line with the finding in two DM1 patients with repeat interruptions at the 5' end (Tomé et al., 2018).

Our findings extend the previous knowledge of mutational dynamics of interrupted *DMPK* expansions, and collectively with previous studies (Braidà et al., 2010; Cumming et al., 2018; Tomé et al., 2018), indicate that different types and structures of repeat interruptions at both ends of expansions, including only a single interruption and *de novo* arisen interruptions, confer genetic stability at the *DMPK* locus in somatic cells. Repeat interruptions were also identified in expanded alleles associated with SCA8 (Moseley et al., 2000; Hu et al., 2017), SCA10 (Matsuura et al., 2006; Landrian et al., 2017) and SCA17 (Gao et al., 2008), but their influence on somatic instability was studied only in SCA17, showing a greater stability of *TBP* expansions interrupted by CAA repeats (Gao et al., 2008).

The greater stability of interrupted *DMPK* expansions in somatic cells is a consequence of *in cis* stabilizing effect of repeat interruptions, both locally and beyond the part of expansion harboring interruptions. We found the identical structures of repeat interruptions at the 3' end of *DMPK* expansions in blood and buccal cells, and in blood cells over time. This feature of somatic tissues is in contrast to germline tissues, since intergenerational changes in the structures of repeat interruptions were reported (our DF2 family and two families reported by Musova et al., 2009). Although repeat interruptions do not completely abolish instability at the *DMPK* locus in somatic cells, we showed that the remaining part (5' end) of interrupted expansions was more stable than expected for the corresponding allele size and age at sampling. This was previously described for *DMPK* expansions with multiple GGC and CCG

interruptions (Braidà et al., 2010), and according to our findings it seems to be a general characteristic of *DMPK* expansions with various types and patterns of repeat interruptions, at least at the 3' end. In addition, we demonstrated that this stability is maintained over time in blood cells, since the 5' end of interrupted expansions was characterized by lower expansion size increment over time than expected based on the allele length and time interval between two samplings.

The ability of repeat interruptions to interfere with the adoption of mutagenic intermediates by CTG/CAG repeats (slipped DNAs and R-loop) (Pearson et al., 2002; Axford et al., 2013; Reddy et al., 2014) is a likely mechanism that maintains the stable local structure of the part of *DMPK* expansion harboring interruptions. In addition, the presence of repeat interruptions in hairpins may facilitate their correct repair upon recognizing the mismatches that arise from the misalignment of repeat interruptions, at least in yeast cells (Rolfmeier et al., 2000). The mechanisms that contribute to increased stability of the remaining part of expansions are not known and may involve epigenetic mechanisms. It can be speculated that repeat interruptions may alter nucleosome positioning, since CTG repeats at the *DMPK* locus act as strong nucleosome positioning elements (Wang et al., 1994), and CAT repeat interruptions in SCA1 expanded repeats significantly reduce their ability to assemble into nucleosomes (Mulvihill et al., 2005). In addition, CCG/GGC repeats, the most commonly observed interruptions at *DMPK* expanded alleles, preferentially exclude nucleosome formation (Wang et al., 1996; Mulvihill et al., 2005). Altered nucleosome positioning may modulate the fidelity of DNA metabolic processes involved in repeat instability (Cleary and Pearson, 2003), or may interfere with the action of *cis* elements that are placed in the flanking sequences and confer instability at the *DMPK* locus (Gourdon et al., 1997; Brock et al., 1999). Polarized hypermethylation of CpG islands downstream of the CTG repeats, very rarely observed in juvenile and adult DM1 patients (Barbé et al., 2017), was described in patients with CCG interruptions at the 3' end of *DMPK* expansion (Santoro et al., 2015). Interestingly, in *E. coli*-based *in vivo* methylation system, CpG methylation of flanking sequences in DM1 patients-derived clones has shown a mild stabilizing effect on expanded CTG repeats (Nichol and Pearson, 2002). However, hypermethylation of flanking CpG islands has not been observed in two DM1 families carrying repeat interruption at the 5' end (Tomé et al., 2018), suggesting that DNA methylation could not be the main mechanism, or it can merely be a consequence of the presence of CCG interruptions at the 3' end. Furthermore, it is noteworthy that bioinformatic predictions showed similar secondary structures of RNA formed by various patterns of CCG and CTG repeats without effect on the affinity for MBNL binding (Braidà et al., 2010). This finding was supported by the study from Santoro et al. (2013), which demonstrated the presence of nuclear foci in muscle biopsies from DM1 patients with variant repeats and their co-localization with MBNL proteins. Additionally, they confirmed the aberrant splicing events that had been postulated as main pathomechanism underlying DM1. Nevertheless, the interruption patterns described in our patients are unique and additional studies are required to examine the possible effect

on formation of nuclear foci and the consequence for RNA metabolism.

It has been assumed that the presentation of DM1 symptoms appears when a sufficient proportion of cells acquire sufficiently long expansions that cause dysfunction, seen at the tissue level (Morales et al., 2012). Similar assumption, known as the somatic threshold model, has been proposed for Huntington's disease (HD) (Budworth et al., 2015). Indeed, patient cohort studies have shown that individual-specific differences in the level of somatic instability of *DMPK* and *HTT* expansions contribute to variability in age at onset (Swami et al., 2009; Morales et al., 2012). Herein, we demonstrated that individual-specific differences in the level of somatic instability had a greater contribution to age at onset in DM1 patients with repeat interruptions relative to patients with uninterrupted expansions. Since repeat interruptions significantly contribute to individual-specific differences in the level of somatic instability and stabilize *DMPK* expansions in somatic cells, our results indicate the role of repeat interruptions as factors with a positive modifying effect on age at onset in DM1 patients. In general, DM1 patients carrying repeat interruptions have predisposition to develop the disease later than average due to a slower progression of expansions toward larger size than average. Therefore, the stabilizing effect of repeat interruptions at the *DMPK* locus in somatic cells can explain an apparently later age at onset in four out of seven patients in our study. The association of repeat interruptions with later age at onset in DM1 patients has also been reported in other studies (Musova et al., 2009; Botta et al., 2017), as well as for SCA8 patients with interrupted *ATXN8OS* expansions (Hu et al., 2017).

The lack of difference between the observed and expected age at onset in the whole group of analyzed DM1 patients with repeat interruptions likely reflects the influence of additional factors that contribute to variability of age at onset, since the applied linear regression models (Morales et al., 2012) explained up to ~55% of variability. Contribution of additional factors that influence the individual-specific differences in the age at onset may also explain why intergenerational instability, biased toward contractions in our DF1, DF2, and DF3 families, and in two families described by Botta et al. (2017), was still accompanied by the genetic anticipation (documented as the decrease in age at onset).

Our evidence that repeat interruptions can positively modulate age at onset in DM1 patients, through stabilizing effect on *DMPK* expansions in somatic cells, is in line with the results showing that suppression of somatic instability of CAG expansions substantially delayed the onset of symptoms in HD mouse models with deletion of certain repair genes (*Ogg1*, *Msh2* and *Msh3*) (Wheeler et al., 2003; Dragileva et al., 2009; Budworth et al., 2015). Additionally, in HD and other patients with polyglutamine diseases, single nucleotide variants in DNA repair genes have been identified as genetic factors that modify

age at onset, probably by influencing somatic instability (Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium, 2015; Bettencourt et al., 2016). Importantly, a study on the HD mouse model has shown that pharmacological treatment that suppresses somatic expansions during life rescues motor decline in these animals (Budworth et al., 2015). Our results on age at onset in DM1 patients with repeat interruptions and the reduced level of somatic instability point to the significance of somatic instability for the disease course and support studies on pharmacological agents that suppress somatic expansions with a potential to delay disease progression.

By detailed characterization of mutational dynamics of *DMPK* expansions in somatic cells, our study has provided firm evidence that various types and patterns of repeat interruptions, including only a single *de novo* interruption, *in cis* confer stability to *DMPK* expansions in somatic cells. Due to stabilizing effect on the *DMPK* locus in somatic cells, repeat interruptions predispose DM1 patients to develop disease later than average, supporting the assumption that they can act as genetic modifier of DM1 phenotype. Our results underlie the clinical significance of repeat interruptions and somatic instability of *DMPK* expansions for progression of the disease. Studies on larger groups of DM1 patients are needed to further elucidate phenotypic consequences of repeat interruptions. They would be important for appropriate genetic counseling, improved stratification of patients in clinical trials and new potential therapeutic approaches.

AUTHOR CONTRIBUTIONS

JP performed genetic and statistical analysis and drafted the manuscript. SP examined patients, collected the samples, and provided clinical data. MB verified statistical analysis. GB supported genetic analyses. VR-S examined patients and reviewed the manuscript. DS-P designed the study and reviewed the manuscript. All authors read and approved the final manuscript.

FUNDING

This study was supported by the Ministry of Education, Science and Technological Development, Republic of Serbia (Grant No. 173016).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2018.00601/full#supplementary-material>

REFERENCES

Axford, M. M., Wang, Y. H., Nakamori, M., Zannis-Hadjopoulos, M., Thornton, C. A., and Pearson, C. E. (2013). Detection of slipped-DNAs at the trinucleotide repeats of the myotonic dystrophy type I disease locus in patient tissues. *PLoS Genet.* 9:e1003866. doi: 10.1371/journal.pgen.1003866

Barbé, L., Lanni, S., López-Castel, A., Franck, S., Spits, C., Keymolen, K., et al. (2017). CpG methylation, a parent-of-origin effect for maternal-biased transmission of congenital myotonic dystrophy. *Am. J. Hum. Genet.* 100, 488–505. doi: 10.1016/j.ajhg.2017.01.033

Bettencourt, C., Hensman-Moss, D., Flower, M., Wiethoff, S., Brice, A., Goizet, C., et al. (2016). DNA repair pathways underlie a common genetic mechanism

- modulating onset in polyglutamine diseases. *Ann. Neurol.* 79, 983–990. doi: 10.1002/ana.24656
- Botta, A., Rossi, G., Marcarelio, M., Fontana, L., D'Apice, M. R., Brancati, F., et al. (2017). Identification and characterization of 5' CCG interruptions in complex DMPK expanded alleles. *Eur. J. Hum. Genet.* 25, 257–261. doi: 10.1038/ejhg.2016.148
- Braida, C., Stefanatos, R. K., Adam, B., Mahajan, N., Smeets, H. J., Niel, F., et al. (2010). Variant CCG and GGC repeats within the CTG expansion dramatically modify mutational dynamics and likely contribute toward unusual symptoms in some myotonic dystrophy type 1 patients. *Hum. Mol. Genet.* 19, 1399–1412. doi: 10.1093/hmg/ddq015
- Brock, G. J., Anderson, N. H., and Monckton, D. G. (1999). Cis-acting modifiers of expanded CAG/CTG triplet repeat expandability: associations with flanking GC content and proximity to CpG islands. *Hum. Mol. Genet.* 8, 1061–1067. doi: 10.1093/hmg/8.6.1061
- Brook, J. D., McCurrach, M. E., Harley, H. G., Buckler, A. J., Church, D., Aburatani, H., et al. (1992). Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* 68, 799–808. doi: 10.1016/0092-8674(92)90154-5
- Budworth, H., Harris, F. R., Williams, P., Lee, D. Y., Holt, A., Pahnke, J., et al. (2015). Suppression of somatic expansion delays the onset of pathophysiology in a mouse model of huntington's disease. *PLoS Genet.* 11:e1005267. doi: 10.1371/journal.pgen.1005267
- Cleary, J. D., and Pearson, C. E. (2003). The contribution of cis-elements to disease-associated repeat instability: clinical and experimental evidence. *Cytogenet. Genome Res.* 100, 25–55. doi: 10.1159/000072837
- R Core Team (2016). *R: A Language And Environment For Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Cumming, S. A., Hamilton, M. J., Robb, Y., Gregory, H., McWilliam, C., Cooper, A., et al. (2018). De novo repeat interruptions are associated with reduced somatic instability and mild or absent clinical features in myotonic dystrophy type 1. *Eur. J. Hum. Genet.* 26, 1635–1647. doi: 10.1038/s41431-018-0156-9
- De Antonio, M., Dogan, C., Hamroun, D., Mati, M., Zerrouki, S., Eymard, B., et al. (2016). Unravelling the myotonic dystrophy type 1 clinical spectrum: a systematic registry-based study with implications for disease classification. *Rev. Neurol.* 172, 572–580. doi: 10.1016/j.neurol.2016.08.003
- Dragileva, E., Hendricks, A., Teed, A., Gillis, T., Lopez, E. T., Friedberg, E. C., et al. (2009). Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiol. Dis.* 33, 37–47. doi: 10.1016/j.nbd.2008.09.014
- Fu, Y. H., Pizzuti, A., Fenwick, R. G. Jr., King, J., Rajnarayan, S., Dunne, P. W., et al. (1992). An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* 255, 1256–1258. doi: 10.1126/science.1546326
- Gao, R., Matsuura, T., Coolbaugh, M., Zühlke, C., Nakamura, K., Rasmussen, A., et al. (2008). Instability of expanded CAG/CAA repeats in spinocerebellar ataxia type 17. *Eur. J. Hum. Genet.* 16, 215–222. doi: 10.1038/sj.ejhg.5201954
- Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium (2015). Identification of genetic factors that modify clinical onset of Huntington's disease. *Cell* 162, 516–526. doi: 10.1016/j.cell.2015.07.003
- Gourdon, G., Radvanyi, F., Lia, A. S., Duros, C., Blanche, M., Abitbol, M., et al. (1997). Moderate intergenerational and somatic instability of a 55-CTG repeat in transgenic mice. *Nat. Genet.* 15, 190–192. doi: 10.1038/ng0297-190
- Harley, H. G., Rundle, S. A., MacMillan, J. C., Myring, J., Brook, J. D., Crow, S., et al. (1993). Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission in myotonic dystrophy. *Am. J. Hum. Genet.* 52, 1164–1174.
- Harper, P. S. (2001). *Myotonic Dystrophy*. London: WB Saunders.
- Higham, C. F. (2013). *Dynamic Dna and Human Disease: Mathematical Modelling and Statistical Inference for Myotonic Dystrophy Type 1 and Huntington Disease*. Ph.D. thesis, University of Glasgow, Glasgow.
- Hu, Y., Hashimoto, Y., Ishii, T., Rayle, M., Soga, K., Sato, N., et al. (2017). Sequence configuration of spinocerebellar ataxia type 8 repeat expansions in a Japanese cohort of 797 ataxia subjects. *J. Neurol. Sci.* 382, 87–90. doi: 10.1016/j.jns.2017.08.3256
- Landrian, I., McFarland, K. N., Liu, J., Mulligan, C. J., Rasmussen, A., and Ashizawa, T. (2017). Inheritance patterns of ATCCCT repeat interruptions in spinocerebellar ataxia type 10 (SCA10) expansions. *PLoS One* 12:e0175958. doi: 10.1371/journal.pone.0175958
- Lavedan, C., Hofmann-Radvanyi, H., Shelbourne, P., Rabes, J. P., Duros, C., Savoy, D., et al. (1993). Myotonic dystrophy: size- and sex-dependent dynamics of CTG meiotic instability, and somatic mosaicism. *Am. J. Hum. Genet.* 52, 875–883.
- Mahadevan, M., Tsilifidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., et al. (1992). Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* 255, 1253–1255. doi: 10.1126/science.1546325
- Martorell, L., Monckton, D. G., Gamez, J., Johnson, K. J., Gich, I., Lopez de Munain, A., et al. (1998). Progression of somatic CTG repeat length heterogeneity in the blood cells of myotonic dystrophy patients. *Hum. Mol. Genet.* 7, 307–312. doi: 10.1093/hmg/7.2.307
- 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
- Monckton, D. G., Wong, L. J., Ashizawa, T., and Caskey, C. T. (1995). Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. *Hum. Mol. Genet.* 4, 1–8. doi: 10.1093/hmg/4.1.1
- Morales, F., Couto, J. M., Higham, C. F., Hogg, G., Cuenca, P., Braida, C., et al. (2012). Somatic instability of the expanded CTG triplet repeat in myotonic dystrophy type 1 is a heritable quantitative trait and modifier of disease severity. *Hum. Mol. Genet.* 21, 3558–3567. doi: 10.1093/hmg/ddsl85
- Morales, F., Vásquez, M., Santamaría, C., Cuenca, P., Corrales, E., and Monckton, D. G. (2016). A polymorphism in the MSH3 mismatch repair gene is associated with the levels of somatic instability of the expanded CTG repeat in the blood DNA of myotonic dystrophy type 1 patients. *DNA Repair* 40, 57–66. doi: 10.1016/j.dnarep.2016.01.001
- Moseley, M. L., Schut, L. J., Bird, T. D., Koob, M. D., Day, J. W., and Ranum, L. P. (2000). SCA8 CTG repeat: en masse contractions in sperm and intergenerational sequence changes may play a role in reduced penetrance. *Hum. Mol. Genet.* 9, 2125–2130. doi: 10.1093/hmg/9.14.2125
- Mulvihill, D. J., Edamura, K. N., Hagerman, K. A., Pearson, C. E., and Wang, Y. H. (2005). Effect of CAT or AGG interruptions and CpG methylation on nucleosome assembly upon trinucleotide repeats on spinocerebellar ataxia, type 1 and fragile X syndrome. *J. Biol. Chem.* 280, 4498–4503. doi: 10.1074/jbc.M413239200
- Musova, Z., Mazanec, R., Krepelova, A., Ehler, E., Vales, J., Jaklova, R., et al. (2009). Highly unstable sequence interruptions of the CTG repeat in the myotonic dystrophy gene. *Am. J. Med. Genet. A* 149A, 1365–1374. doi: 10.1002/ajmg.a.32987
- Nichol, K., and Pearson, C. E. (2002). CpG methylation modifies the genetic stability of cloned repeat sequences. *Genome Res.* 12, 1246–1256. doi: 10.1101/gr.74502
- Pearson, C. E., Tam, M., Wang, Y. H., Montgomery, S. E., Dar, A. C., Cleary, J. D., et al. (2002). Slipped-strand DNAs formed by long (CAG)ⁿ(CTG) repeats: slipped-out repeats and slip-out junctions. *Nucleic Acids Res.* 30, 4534–4547. doi: 10.1093/nar/gkf572
- Pešović, J., Perić, S., Brkušanić, M., Brajušković, G., Rakočević-Stojanović, V., and Savić-Pavićević, D. (2017). Molecular genetic and clinical characterization of myotonic dystrophy type 1 patients carrying variant repeats within DMPK expansions. *Neurogenetics* 18, 207–218. doi: 10.1007/s10048-017-0523-7
- Rakočević-Stojanović, V., Savić, D., Pavlović, S., Lavrnjić, D., Stević, Z., Basta, I., et al. (2005). Intergenerational changes of CTG repeat depending on the sex of the transmitting parent in myotonic dystrophy type 1. *Eur. J. Neurol.* 12, 236–237. doi: 10.1111/j.1468-1331.2004.01075.x
- Reddy, K., Schmidt, M. H., Geist, J. M., Thakkar, N. P., Panigrahi, G. B., Wang, Y. H., et al. (2014). Processing of double-R-loops in (CAG)_n(CTG)_n and C9orf72 (GGGGCC)_n(GGCCCC)_n repeats causes instability. *Nucleic Acids Res.* 42, 10473–10487. doi: 10.1093/nar/gku658
- Rolfmeier, M. L., Dixon, M. J., and Lahue, R. S. (2000). Mismatch repair blocks expansions of interrupted trinucleotide repeats in yeast. *Mol. Cell* 6, 1501–1507. doi: 10.1016/S1097-2765(00)00146-5
- Saluto, A., Brussino, A., Tassone, F., Arduino, C., Cagnoli, C., Pappi, P., et al. (2005). An enhanced polymerase chain reaction assay to detect pre-and full mutation alleles of the fragile X mental retardation 1 gene. *J. Mol. Diagn.* 7, 605–612. doi: 10.1016/S1525-1578(10)60594-6

- Santoro, M., Fontana, L., Masciullo, M., Bianchi, M. L. E., Rossi, S., Leoncini, E., et al. (2015). Expansion size and presence of CCG/CTC/CGG sequence interruptions in the expanded CTG array are independently associated to hypermethylation at the DMPK locus in myotonic dystrophy type 1 (DM1). *Biochim. Biophys. Acta* 1852, 2645–2652. doi: 10.1016/j.bbdis.2015.09.007
- Santoro, M., Masciullo, M., Pietrobono, R., Conte, G., Modoni, A., Bianchi, M. L., et al. (2013). Molecular, clinical, and muscle studies in myotonic dystrophy type 1 (DM1) associated with novel variant CCG expansions. *J. Neurol.* 260, 1245–1257. doi: 10.1007/s00415-012-6779-9
- Savić, D., Rakocvic-Stojanovic, V., Keckarevic, D., Culjkovic, B., Stojkovic, O., Mladenovic, J., et al. (2002). 250 CTG repeats in DMPK is a threshold for correlation of expansion size and age at onset of juvenile-adult DM1. *Hum. Mutat.* 19, 131–139. doi: 10.1002/humu.10027
- Swami, M., Hendricks, A. E., Gillis, T., Massood, T., Mysore, J., Myers, R. H., et al. (2009). Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. *Hum. Mol. Genet.* 18, 3039–3047. doi: 10.1093/hmg/ddp242
- Tomé, S., Dandelot, E., Dogan, C., Bertrand, A., Geneviève, D., Péréon, Y., et al. (2018). Unusual association of a unique CAG interruption in 5' of DM1 CTG repeats with intergenerational contractions and low somatic mosaicism. *Hum. Mutat.* 39, 970–982. doi: 10.1002/humu.23531
- Wang, Y. H., Amirhaeri, S., Kang, S., Wells, R. D., and Griffith, J. D. (1994). Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science* 265, 669–671. doi: 10.1126/science.8036515
- Wang, Y. H., Gellibolian, R., Shimizu, M., Wells, R. D., and Griffith, J. (1996). Long CCG triplet repeat blocks exclude nucleosomes: a possible mechanism for the nature of fragile sites in chromosomes. *J. Mol. Biol.* 263, 511–516. doi: 10.1006/jmbi.1996.0593
- Wheeler, V. C., Lebel, L. A., Vrbanac, V., Teed, A., TeRiele, H., and MacDonald, M. E. (2003). Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum. *Hum. Mol. Genet.* 12, 273–281. doi: 10.1093/hmg/ddg056
- Wong, L. J., Ashizawa, T., Monckton, D. G., Caskey, C. T., and Richards, C. S. (1995). Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. *Am. J. Hum. Genet.* 56, 114–122.
- Wood, L., Bassez, G., Bleyenheuft, C., Campbell, C., Cossette, L., Jimenez-Moreno, A. C., et al. (2018). Eight years after an international workshop on myotonic dystrophy patient registries: case study of a global collaboration for a rare disease. *Orphanet J. Rare Dis.* 13:155. doi: 10.1186/s13023-018-0889-0

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 Pešović, Perić, Brkušanić, Brajušković, Rakočević-Stojanović and Savić-Pavičević. 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.