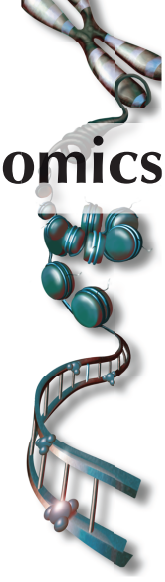
Short Communication

For reprint orders, please contact: reprints@futuremedicine.com

Epigenomics



DNA methylation at the *DMPK* gene locus is associated with cognitive functions in myotonic dystrophy type 1

Édith Breton[‡,1,2], Cécilia Légaré[‡,1,2], Gayle Overend[3], Simon-Pierre Guay[1,4], Darren Monckton[3], Jean Mathieu[2,5], Cynthia Gagnon[2,5], Louis Richer[2,6], Benjamin Gallais[2,5,7] & Luigi Bouchard*[,1,2,8]

[1]Department of Biochemistry & Functional Genomics, Université de Sherbrooke, Sherbrooke, Québec J1E 4K8, Canada
[2]Groupe de recherche interdisciplinaire sur les maladies neuromusculaires (GRIMN), Centre intégré universitaire de santé et de services sociaux (CIUSSS) du Saguenay–Lac-St-Jean - Hôpital de Jonquière, Saguenay, Québec G7X 7X2, Canada
[3]Institute of Molecular, Cell & Systems Biology, University of Glasgow, Glasgow G12 8QQ, United Kingdom
[4]Department of Specialized Medicine, Division of Medical Genetics, McGill University Health Centre, Montreal, QC H4A 3J1, Canada
[5]Centre de recherche Charles-Le-Moyne-Saguenay-Lac-Saint-Jean sur les innovations en santé (CR-CSIS), Université de Sherbrooke, Saguenay, Québec G7H 5H6, Canada
[6]Department of Health Sciences, Université du Québec à Chicoutimi (UQAC), Saguenay, Québec G7H 2B1, Canada
[7]ÉCOBES - Recherche et transfert, Cégep de Jonquière, Saguenay, Québec G7X 7W2, Canada
[8]Department of Medical Biology, Centre intégré universitaire de santé et de services sociaux (CIUSSS) du Saguenay–Lac-St-Jean – Hôpital de Chicoutimi, Saguenay, Québec G7H 5H6, Canada
*Author for correspondence: Tel.: +1 (418) 541-1234 # 2475; luigi.bouchard@usherbrooke.ca
[‡]These authors contributed equally to this work

Aim: Myotonic dystrophy type 1 (DM1) is caused by an unstable trinucleotide (CTG) expansion at the *DMPK* gene locus. Cognitive dysfunctions are often observed in the condition. We investigated the association between *DMPK* blood DNA methylation (DNAm) and cognitive functions in DM1, considering expansion length and variant repeats (VRs). **Method:** Data were obtained from 115 adult-onset DM1 patients. Molecular analyses consisted of pyrosequencing, small pool PCR and Southern blot hybridization. Cognitive functions were assessed by validated neuropsychological tests. **Results:** For patients without VRs (n = 103), blood DNAm at baseline independently contributed to predict cognitive functions 9 years later. Patients with VRs (n = 12) had different DNAm and cognitive profiles. **Conclusion:** DNAm allows to better understand DM1-related cognitive dysfunction etiology.

Lay Abstract: Aim: Myotonic dystrophy type 1 (DM1) is a rare neuromuscular genetic disease caused by an abnormal expansion of a small DNA sequence (CTG) within the *DMPK* gene locus. Cognitive dysfunctions are often observed in DM1. DNA methylation (DNAm) is an epigenetic process regulating gene expression with potential phenotypic impacts. Our study aimed to investigate whether DNAm levels measured in blood at the *DMPK* gene locus are associated with cognitive functions in DM1, in addition to the length of the CTG expansion. **Method:** Data were obtained from 115 adult-onset DM1 patients followed up over a 9-year period. Cognitive functions were assessed by validated neuropsychological tests. **Results:** For patients without CTG repeat interruptions (n = 103), blood DNAm at baseline independently predicted cognitive functions 9 years later. Patients with CTG repeat interruptions (n = 12) had different DNAm and cognitive profiles. **Conclusion:** DNAm might allow us to better understand DM1-related cognitive dysfunction.

First draft submitted: 17 August 2020; Accepted for publication: 18 October 2020; Published online: 10 December 2020

Keywords: copy number variant • epigenetics • genetics • neuromuscular disease • neuropsychology

Myotonic dystrophy type 1 (DM1) is a multisystemic condition causing serious morbidity [1] for the patients themselves and their family [2,3], with muscular impairment as the cardinal sign. A variety of CNS impairments have been described, although they are highly variable between individuals [4,5]. Indeed, various neuropsychological

Future Medicine

studies have shown that DM1 patients present at least subtle changes in their brain functions [2,4–7]. These alterations are more obvious in the severe forms of the disease (i.e., congenital and juvenile DM1), although the adult and late-onset forms have been associated with greater and faster cognitive decline than normally observed with aging [6,8–11]. Such deficits are thought to involve many cognitive functions, like executive function, attention, visuospatial perception, memory, social cognition and processing speed [2,4,7,12].

CNS alterations are supported by neuro-imaging studies, which have identified brain structural changes that might be part of the pathophysiological processes underlying cognitive impairments in DM1 [6,7,9,10,13]. Overall, results from previous studies reveal that cognitive functions are highly variable between individuals, but the biological mechanisms behind this remain mostly unknown [5,14].

DM1 is an autosomal dominant genetic condition [15] caused by an unstable cytosine–thymine–guanine (CTG) trinucleotide repeat expansion within the 3′-untranslated region (UTR) of the *DMPK* gene [16–18]. The CTG expansion becomes pathological over 50 repeats but can reach more than 1000 repeats [16] in the more severe forms. Usually, longer CTG expansions are associated with greater severity of symptoms, including neuropsychological profiles [19], although results have been inconsistent across studies [8,12]. Symptoms generally worsen in subsequent generations as the CTG expansion is unstable and tend to expand during germline transmission [15]. In addition, a small subset of DM1 patients (∼5%) also carries interruptions in their CTG repeat, often called variant repeats (VRs) [20,21]. VRs, which include CCG, GGC or CTC motifs, have been associated with milder muscular symptoms, as they seem to reduce the somatic instability of the CTG expansion, which is also present in DM1 patients over time [20,21]. To date, the association between VRs and cognitive profile in DM1 in not well understood. In addition, genetic variation alone does not explain all DM1 phenotypic variability [11,22], suggesting that other factors, such as epigenetic modifications [23–34], should be taken into account.

DNA methylation (DNAm) is a widely studied epigenetic modification, which regulates gene expression [35,36]. DNAm is environment-sensitive, reversible and can be transmitted through generations [36]. Interestingly, a previous study conducted by our group identified associations between DNAm levels at the *DMPK* gene locus and muscular strength and respiratory functions in adult-onset DM1 patients [23]. Another study found associations between DNAm levels at the *DMPK* gene locus and CTG length, disease severity as well as age of symptoms onset [26]. DNAm levels at the *DMPK* gene locus have also been associated with its own transcriptional activity (e.g., mRNA levels) in various tissues [25]. Interestingly, *DMPK* DNAm levels have also been correlated with the expression of other nearby genes that may also be associated with DM1 pathology, such as *SIX5* [25]. This is of clear interest as deregulation of *SIX5* expression has been suggested as a potential mechanism leading to neuropsychological alterations in DM1 patients [37].

To the best of our knowledge, the impacts of both VRs and DNAm on cognitive functions in DM1 have never been reported. In this study, we have thus investigated the associations between blood DNAm levels at the *DMPK* gene locus and cognitive profiles over a 9-year follow-up period, taking into account both the CTG repeat length and the presence of VRs. We hypothesized that DNAm levels at the *DMPK* gene locus contribute to interindividual variability in cognitive functions of adult-onset phenotypes of DM1, independently from the CTG repeat length and the presence of VRs. Moreover, we expected DNAm levels at the *DMPK* gene locus to be associated with *DMPK* and *SIX5* expression (mRNA levels), which might also be associated with patient's cognitive functions. This would further support the functional role of *DMPK* gene expression regulation (partly thought DNAm) at this locus in DM1 cognitive symptoms.

Participants & methods

This study was part of a larger longitudinal study involving a 3-day evaluation [38,39]. Data were collected at two time points; the first between 2002 and 2004 (T1), and then 9 years later from 2011 to 2013 (T2). In total, 115 DM1 patients presented at both T1 and T2. Details about the recruitment and follow-up have been previously published [38,39].

Participants

DM1 patients were recruited at the Neuromuscular clinic of the *Centre intégré universitaire de santé et de services sociaux du Saguenay-Lac-Saint-Jean* (CIUSSS-SLSJ), Québec, Canada. Adult and late-onset phenotypes were included, and all participants had a molecular diagnostic through genetic testing. Participants were diagnosed with late-onset phenotype if they presented with two of the three following criteria: number of CTG repeats under 200, a score of 1 (no muscle impairment) or 2 (minimal signs) on the Muscular Impairment Rating Scale, and age at

onset of symptoms >40 years [8]. All other participants were classified as having the adult phenotype, according to their molecular DM1 diagnostic. In our cohort, participants self-reported their age at symptom onset. Patient information such as age, sex and educational level were collected through a socio-demographic questionnaire. All participants gave their informed consent and the project was approved by the Research Ethic Committee of the CIUSSS-SLSJ.

Cognitive functions

Among the neuropsychological battery administered to participants, three tests were selected as they showed the highest score variations between T1 and T2 in a previous study by our group [8]. The digit symbol coding subscale of the Wechsler Adult Intelligence Scale-Revised (WAIS-R) is a well-used neuropsychological tool permitting to assess processing speed and that has been shown to be sensitive to the presence of cognitive dysfunction as well as to change in cognitive function [40]. It is used clinically to evaluate patient cognition changes with time and its score is influenced by a wide range of cognitive skills, such as attention, visuospatial perception, learning and working memory [40]. The California Verbal Learning Test (CVLT) measures learning ability and verbal memory (short and long term). The CVLT is also a widely used test in research as well as to clinically monitor episodic verbal learning and memory [41,42]. The CVLT A1-A5 subscore specifically assess the total number of recalled right words and reflects verbal semantic learning abilities. The Ruff 2&7 evaluates sustained and selective attention [43]. Participants completed the French validated version of these tests at both time points [8]. The tests were administered by two well-trained research assistants using standardized procedures, including evaluation across two separate half days to minimize fatigue [8]. All three tests used normative tables allowing us to compute each participant's results into z-scores that were normalized for age (i.e., based on community samples). This means that no change in z-score would be expected over time in a healthy population. For each test, a negative z-score means that the patient performed worse than would be expected for his age. In addition, the CVLT was also normalized for patients' sex, while the Ruff 2&7 was corrected for sex and education [8].

Blood sampling & nucleic acids extraction

DNA was extracted from blood samples at T1 and T2, which were collected in vacutainer tubes containing EDTA. Buffy coats were stored at -80°C until DNA extraction was performed using a standardized protocol from the Gentra Puregene Blood Kit (Qiagen, MD, USA). Double stranded DNA was quantified using the Quant-iT™ PicoGreen™ dsDNA Assay Kit (Thermo Fisher Scientific, USA).

At T2, blood samples were also collected in PAXgene Blood RNA Tubes (Qiagen) and stored at -80°C (as recommended by the manufacturer) until RNA extraction, which was performed using the PAXgene Blood RNA Kit following standardized procedures (Qiagen). RNA integrity (mean RNA Integrity Number (RIN) = 8.7 ± 0.60) was assessed on an Agilent 2100 Bioanalyser, using the Agilent RNA 6000 Nano Kit (Agilent Technologies, CA, USA). Only RNA samples with RIN ≥ 6 were included in mRNA quantification analyses (n = 107, only at T2).

DMPK DNA methylation levels at T1 & T2

DNAm levels at *DMPK* gene locus at T1 and T2 were quantified on a PyroMark Q24 pyrosequencer (Qiagen). Pyrosequencing is a real time and quantitative sequencing method quantifying site-specific DNAm levels as a percentage (%) for each CpG site of a given sequence [44]. Before sequencing, DNA samples were treated with sodium bisulfite using the EpiTect Bisulfite Kit (Qiagen), followed by PCR amplification with the PyroMark PCR Kit (Qiagen), according to the manufacturer's standard protocol. Based on a previous study from our group, we selected CpG sites with the largest variability in DNAm levels between participants [23]. PCR and sequencing primers were designed with Pyromark assay design 2.0 (Qiagen), as described before [23]. A schema of the selected region (i.e., L6), which is located in the 3′-UTR (downstream) of *DMPK* (GRCh38: Chromosome 19: 45770041-45769996), is presented in Figure 1. 14 samples (four at T1 and ten at T2), failed to pass the pyrosequencing quality controls and were thus excluded from the statistical analyses.

DMPK & *SIX5* mRNA levels at T2

DMPK and *SIX5* mRNA levels were quantified using quantitative real-time PCR (Applied Biosystems 7500 Real Time PCR system; ThermoFisher Scientific, MA, USA) relative to *B2M* and *GAPDH* endogenous controls. For each sample (n = 107), 200 ng of RNA was reverse transcribed into cDNA using the High Capacity cDNA RT kit

10.2217/epi-2020-0328

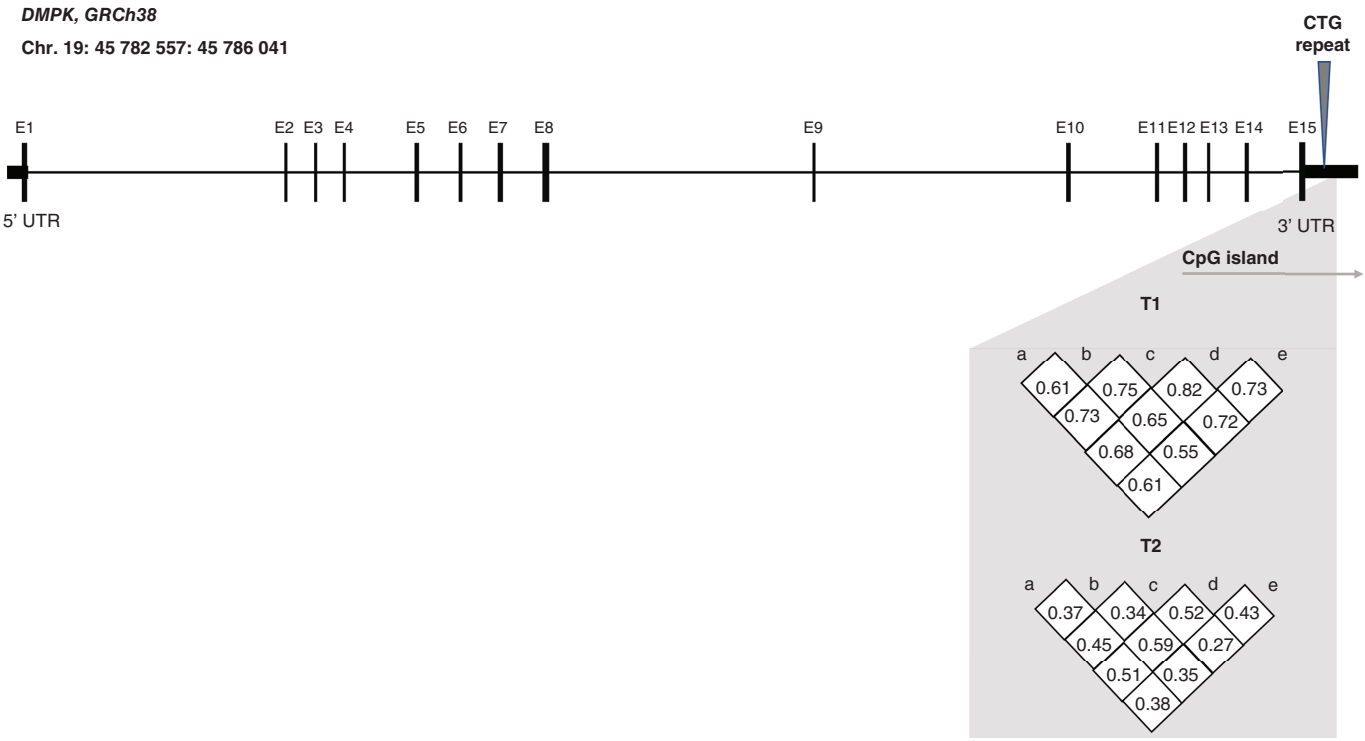


Figure 1. Schema of the *DMPK* gene locus. The region studied was located downstream of the CTG expands, in a large CpG island. Spearman correlation coefficients between CpG sites DNAm levels are shown for T1 and T2.
DNAm: DNA methylation; UTR: Untranslated region.

(ThermoFisher Scientific), and amplified, in duplicate, using the Taqman Fast Universal PCR Master Mix (ThermoFisher Scientific), as well as primers and Taqman probes (ThermoFisher Scientific; *DMPK*: Hs00189385_m1, *SIX5*: Hs01650774_m1, *B2M*: Hs00984230_m1, *GAPDH*: Hs02758991_g1). *B2M* and *GAPDH* mean mRNA levels were shown to be stable in the blood samples of participants (18.3 ± 0.6) and were measured in parallel to *DMPK* and *SIX5* as reference genes, according to the manufacturer's recommendations. Mean mRNA level ratios (reference genes/*DMPK*; reference genes/*SIX5*) were computed and used for the statistical analyses. *SIX5* mRNA duplicates values were too different for ten samples (i.e., more than 0.5 cycle threshold differences), thus they were excluded from the analyses.

CTG repeat length & variant repeats analyses

Modal CTG repeat length at each timepoint was measured using small pool PCR and Southern blot hybridization. As previously described [23,45], the flanking primers DM-C and DM-DR were used to amplify the *DMPK* CTG repeat by PCR. The PCR buffer Custom PCR Master Mix-No Taq (Thermo Scientific, Renfrew, UK) was supplemented with 69 mM of 2–mercaptoethanol, and Taq polymerase (Sigma-Aldrich, Gillingham, UK) was used at 1 unit per 10 µl. When amplification of the expanded allele failed due to the suspected presence of variant repeats, 10% DMSO (Sigma-Aldrich) was added to the reactions, along with a decrease of the annealing temperature from 68 to 63.5°C. According to the manufacturer's protocol, PCR products were digested with AciI (New England Biolabs, Hitchin, UK) for detection of VRs. PCR products were resolved on agarose gels (1%) in 0.5X TBE, blotted and hybridized. Labeling of DNA probes complementary to the CTG repeat and the molecular weight marker Invitrogen 1 kb+ ladder (Thermo Fisher Scientific) and the was performed with α-32P-dCTP (Perkin Elmer, Beaconsfield, UK) and the Invitrogen Random Primers DNA labeling system, respectively (Thermo Fisher Scientific). Afterward, autoradiographs were scanned, and allele lengths were estimated in comparison to the 1 kb+ ladder using CLIQS 1D gel analysis software (TotalLabs, St Helens, UK), as previously described [45]. CTG repeat length was unavailable for eight participants at T1 and six participants at T2.

Statistical analyses

Because differences between VRs groups were observed for DNAm levels, all the following analyses were conducted in both groups separately (if not otherwise specified). IBM SPSS 25.0 software (IBM, NY, USA) was used for all the statistical analyses. Normality of the variables were assessed with Shapiro-Wilk test and histogram inspection. As most variables did not present a normal distribution, non-parametric tests were employed when needed. We used Mann-Whitney U-test to assess differences between VRs groups (i.e., with versus without VRs) for CTG repeat length, DNAm levels and cognitive functions as well as for age and age at onset. Chi Square tests were applied to assess group differences in DM1 phenotype and sex. We also applied a Kruskal-Wallis test to investigate the presence of differences in DNAm levels according to the parental inheritance of the mutant allele (maternal/paternal/unknown), at the five CpGs of interest in our study.

Wilcoxon signed-rank tests were used to study the presence of change (or decline) in cognitive functions over the 9 year follow-up period. Delta z-score were also calculated (i.e., cognitive test z-scores at T2 minus cognitive test z-scores at T1) as indicators of change over time (i.e., negative delta z-score mean the patient's performance according to his age declined from T1 to T2). We used this score of change to compare the two groups (i.e., with vs without VRs) using Mann-Whitney U-test.

To explore how cognitive test scores at T2 could be predicted by the patient's characteristics and DNAm levels at T1, we first computed Spearman correlation coefficients to identify potential CpG sites to include in our models (correlation p-value < 0.15 for association between DNAm levels at T1 and the cognitive test scores at T2). We then used stepwise linear regression models for each neuropsychological test. Patients with AciI-sensitive VRs (n = 12) were excluded, as the sample size of that group was too small to apply stepwise analyses. As we aimed to predict cognitive functions at T2 using the patient's profile at T1, the following variables were included in the models, based on available data from the original study: DNAm levels (T1, only CpGs correlated with the outcome), modal CTG repeat length (T1), DM1 phenotype, age (T1), sex and education levels. As stated above, the z-score for the CVLT already considered patient's sex; thus, we did not add sex to the CVLT regression model. Using variance inflation factors, we tested each model to make sure there was no multicollinearity of variables. DNAm levels at T1 were not associated with Ruff 2&7 z-score at T2; therefore, we did not pursue the analyses for this test.

Finally, we assessed the correlations between *DMPK* and *SIX5* mRNA levels with DNAm levels at the *DMPK* gene locus and the cognitive test scores (only at T2) using Spearman coefficients.

Results

Participant's characteristics

Table 1 presents the participant's characteristics at T1 and T2. Most had the adult phenotype (79.1%). Mean age at T1 was 43.6 ± 10.3 years, whereas the age at onset was 21.2 ± 9.5 years, on average. Patients with AciI-sensitive VRs had lower CVLT A1-A5 z-score (p = 0.002, Mann–Whitney U-test) and tend to have higher Ruff 2&7 total speed scores (p = 0.07, Mann–Whitney U-test) at T1 as compared with those without the VRs (Table 1). No other differences between groups were found (Table 1).

Participant's cognitive functions change over time

Patients without VRs presented a significant decline in their cognitive performance for the three tests: CVLT A1-A5 ($p = 4.9 \times 10^{-9}$, Wilcoxon signed-rank test), Ruff 2&7 ($p = 1.1 \times 10^{-12}$, Wilcoxon signed-rank test) and WAIS-R digit symbol coding ($p = 4.6 \times 10^{-9}$, Wilcoxon signed-rank test). On the opposite, only WAIS-R digit symbol coding significantly declined from T1 to T2 for patients with VRs (p = 0.01, Wilcoxon signed-rank test).

Based on delta z-scores (i.e., z-score T2 – z-score T1), we also found that decline in CVLT A1-A5 performance tend to be greater in patients without VRs as compared with those with the VRs (VRs: -0.10 ± 0.97; No VRs: -0.78 ± 1.2; p = 0.08) (Figure 2).

DMPK DNA methylation levels at T1 & T2

As shown in Figure 1, the five CpGs tended to be only moderately correlated with each other ($0.27 \leq r_s \leq 0.82$) and were thus considered independently in the analyses. At T1, all CpGs showed lower DNAm levels in the absence of AciI-sensitive VRs, and these differences were similar at T2 (Table 2 and Supplementary Table 1; adjusted for age and sex). There were differences in DNAm levels between T1 and T2 for patients without VRs at CpGa (p = 0.02), CpGb (p = 0.04), CpGc (p = 0.001), CpGd (p = 0.03) and CpGe ($p < 0.001$) (data not shown). For patients with VRs, there was only a statistically significant difference between T1 and T2 at CpGd (p = 0.02). DNAm levels

10.2217/epi-2020-0328

Table 1. Patient's characteristics at first and second follow-up.

	Patients without variant repeats			Patients with variant repeats			p-value
	n	Mean ± SD OR n (%)	Range (min – max)	n	Mean ± SD OR n (%)	Range (min – max)	
First follow-up (T1)							
Age	103	43.3 ± 10.5	20–77	12	46.8 ± 8.4	39–69	0.11
Patient self-reported age at symptom onset	75	20.7 ± 9.2	4–64	7	27.0 ± 11.7	12–43	0.18
Sex (F/M)	103	65/38 (63.1% F)	NA	12	7/5 (58.3% F)	NA	0.75
Education level (high school not completed/ high school completed/ higher education)	103	45/37/21 (43.7%/ 35.9%/ 20.4%)	NA	12	5/3/4 (41.7%/ 25%/ 33.3%)	NA	0.06
Phenotype (Adult/Late-onset)	103	82/21 (79.6% adult)	NA	12	9/3 (75.0% adult)	NA	0.71
Modal CTG repeat length	96	510 ± 318	68–1321	11	443 ± 243	150–859	0.74
WAIS-R digit symbol coding (z-score)†	103	-1.01 ± 0.62	-2.33–1.00	11	-0.91 ± 0.84	-2.00–0.33	0.77
CVLT A1-A5 (z-score)‡	103	0.31 ± 1.10	-2.70–2.30	12	-0.81 ± 1.04	-2.60–0.30	0.002
Ruff 2&7 total speed (z-score)§	99	-0.73 ± 0.92	-2.90–1.90	11	-0.23 ± 1.24	-2.80–2.30	0.07
Second follow-up (T2)							
Age	103	51.9 ± 10.49	29–85	12	55.33 ± 8.38	47–77	0.13
CTG repeat length	99	649 ± 431	63–1545	10	558 ± 296	202–1071	0.73
WAIS-R digit symbol coding (z-score)†	102	-1.33 ± 0.68	-3.00–0.67	12	-1.58 ± 1.00	-3.00–0	0.25
CVLT A1-A5 (z-score)‡	103	-0.48 ± 1.16	-4.50–2.30	11	-0.96 ± 1.60	-3.5–1.00	0.45
Ruff 2&7 total speed (z-score)§	98	-1.45 ± 1.00	-4.00–1.3	11	-0.78 ± 1.62	-2.9–2.8	0.14

†z-score corrected for patient age.
‡z-score corrected for patients age and sex.
§z-score corrected for patient age and education.
CVLT: California Verbal Learning Test; OR: Odds ratio; SD: Standard deviation; WAIS-R: Wechsler Adult Intelligence Scale-Revised.

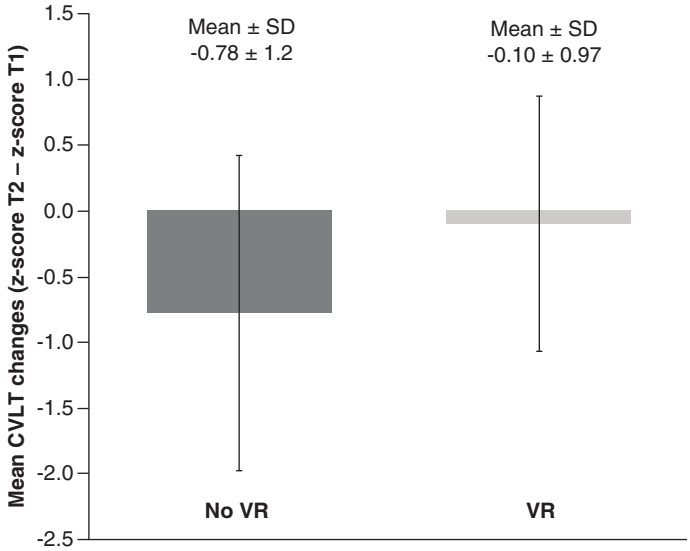


Figure 2. Cognitive function changes from T1 to T2 (patients with or without variant repeats) for CVLT A1-A5, compared with Mann–Whitney U-test. A negative delta z-score mean the patient's performance according to his age declined more than would be expected in a community sample from T1 to T2.
CVLT: California Verbal Learning Test; SD: Standard deviation; VR: Variant repeats.

were only moderately correlated with the CTG repeat length at both T1 and T2 (Supplementary Table 2). Finally, we did not find DNAm differences between participants who inherited their CTG expansion maternally (n = 40) or paternally (n = 53; unknown n = 18; $p > 0.12$; Supplementary Figure 1).

DMPK DNAm levels at T1 predict cognitive functions at T2 for patients without VRs

Spearman correlations were applied to select the CpG sites to include in the stepwise linear regression models. We found that DNAm levels at CpGb (r_s = -0.15; p = 0.14) were correlated with the WAIS-R digit-symbol coding model, whereas CpGb (r_s = -0.17; p = 0.09), CpGc (r_s = -0.23; p = 0.02), CpGd (r_s = -0.32; p = 0.001) and CpGe (r_s = -0.18; p = 0.08) were correlated with the CVLT A1-A5. We did not observe any correlation between DNAm

Table 2. Differences in DNA methylation levels at the *DMPK* gene locus between patients with or without AciI-sensitive variant repeat.

	DNA methylation levels (%) at the *DMPK* gene locus				p-value[†]
	Patients without variant repeats		Patients with variant repeats		
	Mean ± SD	Range (min – max)	Mean ± SD	Range (min – max)	
Baseline (T1)	**(n = 100)**		**(n = 11)**		
CpGa	2.78 ± 1.69	1.72–16.06	8.55 ± 8.84	1.85–26.59	0.002
CpGb	1.61 ± 0.91	0.70–7.49	7.01 ± 7.45	0.97–21.41	4.3×10^{-4}
CpGc	4.40 ± 2.08	2.46–19.15	12.33 ± 10.91	3.37–32.89	3.7×10^{-4}
CpGd	2.26 ± 1.84	1.01–16.32	10.24 ± 10.89	1.03–31.28	1.9×10^{-4}
CpGe	5.91 ± 1.70	3.53–16.61	12.64 ± 9.90	4.55–32.98	0.005
9-year follow-up (T2)	**(n = 94)**		**(n = 11)**		
CpGa	2.90 ± 0.88	1.42–7.02	7.91 ± 6.85	2.36–19.08	0.003
CpGb	1.76 ± 0.93	0–7.36	5.96 ± 6.71	0–19.27	0.06
CpGc	4.84 ± 1.88	2.69–17.15	11.76 ± 8.45	3.98–25.72	5.9×10^{-5}
CpGd	2.42 ± 1.55	0–14.25	7.71 ± 8.01	1.55–21.42	0.02
CpGe	8.35 ± 2.35	4.32–18.51	12.83 ± 7.45	5.75–27.54	0.10

[†]Corrected for patients age and sex.
SD: Standard deviation.

Table 3. Stepwise linear regression model aiming to predict the patients' performance at the WAIS-R digit symbol coding (z-score) test at T2 based on their characteristics at T1.

	WAIS-R digit symbol coding at T2 (β ± SE; p-value)
Age	$\beta = -0.57$; $p = 1.0 \times 10^{-10}$
Sex	$\beta = -0.32$; $p = 6.0 \times 10^{-5}$
Education	$\beta = 0.24$; $p = 0.003$
DM1 phenotype	$\beta = -0.30$; $p = 0.001$
DNAm levels CpGb	$\beta = -0.21$; $p = 0.01$
Model R square	0.495
Model p-value	3.9×10^{-12}

Variables tested are DNAm levels at CpGb (T1), CTG repeat length (T1), age (T1), DM1 phenotype, sex and education.
DNAm: DNA methylation; SE: Standard error; WAIS-R: Wechsler Adult Intelligence Scale-Revised.

Table 4. Step-wise linear regression aiming to predict the patients' performance at the CVLT A1-A5 (z-score) test at T2 based on their characteristics at T1.

	CVLT A1-A5 z-score at T2 (β ± SE; p-value)
DM1 phenotype	$\beta = 0.32$; $p = 0.001$
Education	$\beta = 0.23$; $p = 0.02$
DNAm levels CpGd	$\beta = -0.18$; $p = 0.06$
Model R square	0.166
Model p-value	0.001

Variables tested are DNAm levels (CpGb-CpGe) (T1), CTG repeat length (T1), age (T1), DM1 phenotype and education.
CVLT: California Verbal Learning Test; DNAm: DNA methylation;SE: Standard error.

levels T1 and Ruff 2&7 total speed at T2. We also found that a longer number of CTG repeats at baseline was correlated with worst outcomes at the Ruff 2&7 ($r_s = -0.45$; $p = 6.0 \times 10^{-6}$) and the WAIS-R digit-symbol coding ($r_s = -0.26$; $p = 0.01$) at follow-up in patients without VRs.

Then applying stepwise regression models, we found that age at T1, sex, education level, DM1 phenotype and *DMPK* DNAm levels (T1) at CpGb explained 49.5% of WAIS-R digit symbol coding test score variability at T2 (model $p = 3.9 \times 10^{-12}$) (Table 3). Using a similar approach, we found that patients' DM1 phenotype, education level and *DMPK* DNAm levels (T1) at CpGd contributed independently to CVLT A1-A5 test score (Table 4).

10.2217/epi-2020-0328

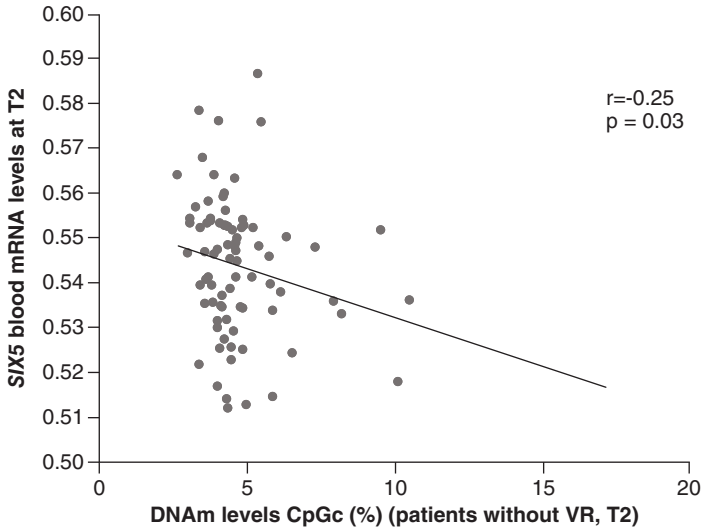


Figure 3. Correlation between CpGc DNA methylation levels and *SIX5* mRNA levels at T2.
VR: Variant repeat.

Together, these explained 16.6% of the CVLT A1-A5 test score variance at T2 (model p = 0.001) (Table 4). Interestingly, the CTG length at T1 was not retained in the step-wise models when including DNAm levels.

DMPK DNAm levels at T1 are associated with cognitive functions at T2 for patients with AciI-sensitive VRs

The limited number of patients with AciI-sensitive VRs prevented to test regression models as above. Nevertheless, we investigated the association between *DMPK* DNAm levels and cognitive test scores in this patients' subgroup. Using Spearman correlation, we found coefficients of interesting effect size between *DMPK* DNAm levels at CpGb-CpGe and WAIS-R digit symbol coding test scores (r_s range = -0.16 to -0.27), as well as between DNAm levels at CpGb-CpGe and CVLT A1-A5 test scores (r_s range = -0.24 to -0.35). Although none reached statistical significance (Supplementary Table 3), these effect sizes are of similar magnitude and in the same direction as those observed for patients without the VRs. As for CTG length, a statistically significant correlation was only observed with the WAIS-R digit-symbol coding (r_s = -0.61; p = 0.04) at follow-up.

DMPK & *SIX5* mRNA quantification analyses

First, *DMPK* and *SIX5* mRNA levels quantified in blood cells (RNA samples only available at T2) were similarly positively and highly correlated with each other in both patients without (r_s = 0.78, p = 2.2×10^{-18}) and with VRs (r_s = 0.81, p = 0.003).

We further tested the association between *DMPK* DNAm and *DMPK* and *SIX5* mRNA levels. We found a negative correlation between CpGc DNAm at T2 and *SIX5* mRNA levels (r_s = -0.25, p = 0.03) in patients without VRs (Figure 3). This correlation did not reach statistical significance in patients with the VRs, although it was of higher effect size (same direction) as that reported for patients without VRs (r_s = -0.44, p = 0.20). There was no other evidence of an association between DNAm levels and *DMPK* or *SIX5* mRNA levels in our cohort (data not shown). Of note, the CTG repeat length at T2 was marginally correlated with *DMPK* (r_s = 0.19, p = 0.07) and *SIX5* (r_s = -0.18, p = 0.10) mRNA levels in patients without VRs.

As for the correlations between gene expression and phenotype, *DMPK* mRNA levels were positively and marginally correlated with the CVLT A1-A5 at T2 in patients without VRs (r_s = 0.18, p = 0.07). This correlation did not reach the significance threshold in patients with VRs, and the effect size was on the opposite direction (r_s = -0.14, p = 0.66). No other evidence of correlation was found between mRNA levels and cognitive profile (data not shown).

Discussion

There is an urgent clinical need to better understand the mechanisms behind DM1 multi-systemic involvement, and the high interindividual variability of the condition makes it a great challenge [1,10,46,47]. Since cognitive impairments greatly impede on DM1 patients' social participation, it is important to better understand their

origins [3,47,48]. To the best of our knowledge, we are the first to report that DNAm levels at *DMPK* gene locus could independently predict patients' cognitive profile almost a decade later, in adult-onset DM1.

Currently, the *DMPK* CTG repeat length is recognized as the best predictor available to estimate the patient's multisystemic impairments in DM1, and some studies suggested that it might also apply to patients' cognition, but results are inconsistent between studies [8,12,19,49]. With this knowledge gap in mind, we have tested associations and identified CpG sites at the *DMPK* gene locus where DNAm levels contributed to predict patients' WAIS-R digit symbol coding and the CVLT A1-A5 test scores, two neuropsychological tests performed 9 years after baseline blood collection. These two tests are both highly involved in age-related cognitive diseases, such as dementia as well as in the DM1 age and disease duration-related cognitive decline [8,50]. Importantly, performance on the WAIS-R digit symbol coding correlates with real-world functional outcomes (e.g., the ability to accomplish everyday tasks). This makes the current findings even more significant, since a decrease in abilities to accomplish such everyday activities in DM1 patients is not only due to diminished motor capacities but also relate to cognitive functions.

Our results thus emphasize the promise of DNAm as a potential biomarker of disease progression and as a potential prognostic tool for the cognitive phenotypes. Consequently, quantifying DNAm modifications in DM1 might lead to key findings leading toward a better understanding of DM1 pathophysiology and accordingly, to improve personalized interventions. As an example, studies carried on an elderly population also found an association between blood DNAm levels and abnormal cognitive decline during normal brain aging [51]. DM1 is often compared with a hastened aging process, and as such, finding in DM1 may also improve our understanding of mechanisms of abnormal aging [8,52].

Interestingly, we found that *DMPK* DNAm levels at T1 explain a fraction of the variance in WAIS-R digit symbol coding and of the CVLT A1-A5 tests scores at T2 independently of the CTG repeat length, which was not retained in stepwise models (even if there was a significant Spearman correlation between CTG length and WAIS-R digit symbol coding) underlying the potential significance of DNAm in DM1 cognitive functions. These results need to be interpreted with caution, as we only studied a few CpG sites and three neuropsychological tests – including one (i.e., the Ruff 2&7) for which DNAm levels were not correlated with the test's z-score, whereas CTG repeat length was. This is partly in line with a recent study suggesting the importance of the CTG length in DM1 cognitive functions [49]. Of note, methodological differences between this previous study and ours might partly explain some result discrepancies, as we used different neuropsychological tests, as well as different study designs (i.e., transversal versus longitudinal). DNAm levels and CTG repeat length are correlated with each other and might well interact in DM1 pathophysiology. More study including both measures of CTG length and DNAm levels (at more CpG sites) will be needed to clearly elucidate their respective and shared roles in DM1 cognitive functions. Moreover and as underlined in a recent study, it might also be important to consider DM1 parental inheritance of the mutant allele when studying DNAm at *DMPK* gene locus (increased DNAm levels when the mutant allele is maternally inherited) [53]. However, we did not replicate these findings at any of the five CpG sites we investigated; nevertheless, parental inheritance is a very interesting concept that warrants further studies. Overall, our findings open the door for further studies that could help understand cognitive dysfunction's progression in DM1 patients through a combination of biological markers.

Previous findings support that the degree of instability of the CTG expansion is associated with age at onset (i.e., patients with more somatic instability develop DM1 symptoms earlier) [22,49]. Interestingly, the VRs have been associated with an increased stability of the CTG repeat, as well as a later age at onset compared with patients with similar CTG repeat length, thus leading to a milder phenotype than what is normally observed in DM1 patients without VRs [11,20,21,49,54]. Therefore, our study does not entirely corroborate these previous findings (CVLT A1-A5 performance was poorer in patients with VRs at baseline when compared with those without the VRs). Further studies are thus needed to clarify whether the VRs are associated with milder cognitive symptoms and/or decline over time. Interestingly, DNAm has also been suggested to increase CTG repeat stability in DM1 [25,26,55] and the VRs might well modulate this association [56]. In the current study, we observed significantly higher *DMPK* DNAm levels at CpG sites located downstream of the CTG repeat in patients with the VRs compared with those without VRs. Patients with the VRs (and higher DNAm levels) presented lower verbal learning memory at T1 (measured by the CVLT A1–A5), but their decline over time was milder than for patients without the VRs. However, we have to take into account that scores at T1 were relatively low in patients with the VRs, suggesting a potential threshold effect (i.e., patients who already had low scores at T1 cannot decline as much as patients who had better scores). Overall, our results suggest that patients with VRs would probably benefit from being studied separately, as they possibly have a different DNAm profile than those without it, which might impact their disease profile as well.

10.2217/epi-2020-0328

In support of the findings above, we found that *DMPK* DNAm levels are correlated with *SIX5* gene expression in blood, both in patients with and without VRs, as in other studies [24,25,27]. We also found a correlation between *DMPK* DNAm levels at CpGc and *SIX5* mRNA levels. It is possible that DNAm levels at these sites have a tissue-specific effect on these two gene's expressions. Interestingly, *SIX5* expression regulation could play a role in cognitive alterations in DM1 [37] and has been suggested as a regulator of *DMPK* gene expression. *SIX5* knock out in mice generated some phenotypes that may be reflective of DM1 symptoms, supporting the idea that DM1 pathology is mediated by multiple genes in the region [57,58].

Strengths & limitations

The current study has been carried on a large cohort of patients with DM1 (n = 115) for which rich longitudinal data were available over a 9-year period follow-up. This is a remarkable achievement in the context of a rare genetic disease. Moreover, the use of robust and widely used neuropsychological tests for the assessment of patients' cognitive profile is another clear advantage of our study. This is also the first assessment of the relationship between the presence of VRs and DNAm levels at the *DMPK* gene locus and patient's cognitive decline in adult-onset DM1, which also considered the contribution of the CTG repeat length.

Yet, our study also has some limitations. The small sample size of patients with the VRs (n = 12) limited our capacity to assess the potential predictive properties of DNAm at the *DMPK* locus on patients' cognitive functions over 9 years in this group. This specific genetic variation is rare among patients with DM1 although the prevalence in our cohort is about twice higher (10.4%) than expected [20,21]. Also, our results are, for now, only applicable to adult and late-onset DM1 phenotypes. Moreover, we exclusively studied DNAm levels of a relatively small region of the *DMPK* gene locus proximal to the CTG expansion; thus, it would be interesting to extend the analyses to the entire *DMPK* gene locus as well as to *SIX5*. The low DNAm levels, especially in the non-VRs group, might also be considered as a limitation of this study. However, the presence of DNAm variability at one timepoint between groups, but also the DNAm levels variability over time at these CpG sites in patients without VRs counterpart this limitation by supporting the plausible role of these regions in the difference between patients with and without VRs. Moreover, the clear DNAm level distinction between the groups emphasizes our findings supporting the need to consider patients with VR separately. Importantly, DNAm levels, CTG expansion and the gene's expression were measured in blood, which might not always accurately represent their values in the brain, as previous studies support that DNAm at *DMPK* gene locus correlates with mRNA levels in a tissue-specific manner [25]. Nevertheless, blood remains an accessible and common tissue in the quest for reliable biomarkers, including in DM1. Finally, we did not compare our results with unaffected controls, which limit our capacity to determine whether our findings are limited to DM1 or not.

Conclusion

Our longitudinal study established for the first time an association between DNAm levels at the *DMPK* gene locus and patient's cognitive profile over 9 years of follow-up in adult-onset DM1. Our results suggest the importance of considering the contribution of DNAm to DM1 pathophysiology and phenotypic variability, at least when studying cognitive impairments associated with this condition. They also provide evidence that presence of a VR in the CTG expansion might influence or be influenced by DNAm levels in DM1, affecting patient's cognitive profile. In brief, our results highlight the complexity of DM1 physiopathology and suggest that the condition involves several genes and regulatory elements that must be further investigated.

Future perspective

Although CTG repeat length has been used to estimate DM1 disease severity, this study and other studies showed that other factors such as the presence of a variant repeat or DNAm at the *DMPK* gene locus may also be implicated. Thus, in addition to CTG repeat length, information about the presence of a variant repeat and DNA methylation may also be collected to be taken into account in future studies as they also affect disease severity. Further studies are also needed to better understand how DNA methylation and the presence of a variant repeat affect DM1 pathogenesis.

Summary points

- DM1 is a rare neuromuscular disease caused by an unstable cytosine–thymine–guanine (CTG) expansion at the *DMPK* gene locus. There are cognitive dysfunctions in the adult and late-onset forms of the condition, but their origin is not well understood.
- The high inter-individual variability in DM1 symptoms is complex, possibly involving epigenetic, as well as other genetic processes such as the presence of variant repeats (VRs) in the CTG expansion.
- Using a large longitudinal cohort of DM1 patients, this study investigated for the first time the association between DNA methylation and cognitive functions in DM1, considering both CTG length and VRs.
- We found that patients with VRs have different DNA methylation and cognitive profiles than patients without VR.
- By combining DNA methylation and clinical characteristics measured at T1, our regression models explain 49.5% of Wechsler Adult Intelligence Scale-Revised digit symbol coding ($p = 3.9 \times 10^{-12}$) and 16.6% of California Verbal Learning Test A1-A5 ($p = 0.001$) tests score variance in patients without VRs at T2, independently of the CTG length.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/epi-2020-0328

Acknowledgements

The authors acknowledge all research participants, and the staff of the Groupe the recherche interdisciplinaire sur les maladies neuromusculaires (GRIMN) who participated to the follow-up and data collection. We also thank V Desgagné (Research Assistant), who carefully revised the manuscript.

Financial & competing interests disclosure

During this study, E Breton was recipient of a financial support from the Faculté de médecine et des Sciences de la Santé from the Université de Sherbrooke, and from the Fondation du Grand Défi Pierre Lavoie. E Breton and C Légaré are also supported by a doctoral training award from the Fonds de recherche en Santé du Québec (FRQS) and C Légaré received a funding from the Fondation du Grand Défi Pierre Lavoie and Corporation de recherche et d'action sur les maladies héréditaires (CORAMH). During this research, SP Guay was the recipient of a doctoral research award from the Canadian Institutes for Health and Research (CIHR). CG and LB are senior research scholars from the FRQS. L Bouchard is a member of the CR-CHUS, a FRQS-funded research center. C Gagnon and B Gallais are members of the Centre de recherche Charles-Le-Moyne-Saguenay-Lac-Saint-Jean, which is affiliated to the Université de Sherbrooke. G Overend and DG Monckton were supported by an award from Muscular Dystrophy UK and donations from the Myotonic Dystrophy Support Group.

Within the last 3 years DG Monckton has been a scientific consultant and/or received an honoraria/stock options from AMO Pharma, Charles River, Vertex Pharmaceuticals, Triplet Therapeutics, LoQus23 and Small Molecule RNA, research contracts with AMO Pharma and Vertex Pharmaceuticals. DG Monckton is on the Scientific Advisory Board of Myotonic (formerly the Myotonic Dystrophy Foundation) and EuroDyMA (European Dystrophia Myotonica Association), is a scientific advisor to the Myotonic Dystrophy Support Group and is a vice president of Muscular Dystrophy UK. He has received research funding from Muscular Dystrophy UK and the Myotonic Dystrophy Support Group. C Gagnon received an honorarium from Biogen Idec and a funding from Bioblast Pharma inc. B Gallais has received an honorarium from Expansion Therapeutics, Inc and from Biogen Idec.

The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. De Antonio M, Dogan C, Hamroun D *et al.* Unravelling the myotonic dystrophy type 1 clinical spectrum: a systematic registry-based study with implications for disease classification. *Rev. Neurol. (Paris)* 172(10), 572–580 (2016).
2. Fujino H, Shingaki H, Suwazono S *et al.* Cognitive impairment and quality of life in patients with myotonic dystrophy type 1. *Muscle Nerve* 57(5), 742–748 (2018).

3. Van Heugten C, Meuleman S, Hellebrekers D, Kruitwagen-van Reenen E, Visser-Meily J. Participation and the role of neuropsychological functioning in myotonic dystrophy type 1. *J. Neuromuscul. Dis.* 5(2), 205–214 (2018).
4. Okkersen K, Buskes M, Groenewoud J *et al.* The cognitive profile of myotonic dystrophy type 1: a systematic review and meta-analysis. *Cortex* 95, 143–155 (2017).
5. Peric S, Rakocevic Stojanovic V, Mandic Stojmenovic G *et al.* Clusters of cognitive impairment among different phenotypes of myotonic dystrophy type 1 and type 2. *Neurol. Sci. Milano* 38(3), 415–423 (2017).
6. Wozniak JR, Mueller BA, Bell CJ, Muetzel RL, Lim KO, Day JW. Diffusion tensor imaging reveals widespread white matter abnormalities in children and adolescents with myotonic dystrophy type 1. *J. Neurol.* 260(4), 1122–1131 (2013).
7. van Dorst M, Okkersen K, Kessels RPC *et al.* Structural white matter networks in myotonic dystrophy type 1. *NeuroImage Clin.* 21, 101615 (2019).
8. Gallais B, Gagnon C, Mathieu J, Richer L. Cognitive decline over time in adults with myotonic dystrophy type 1: a 9-year longitudinal study. *Neuromuscul. Disord.* 27(1), 61–72 (2017).

•• **Shows the cognitive decline over 9 years in dystrophy type 1 (DM1) patients and which cognitive function decline the most.**

9. Caso F, Agosta F, Peric S *et al.* Cognitive impairment in myotonic dystrophy type 1 is associated with white matter damage. *PLoS One* 9(8), e104697 (2014).
10. Ashizawa T. Myotonic dystrophy as a brain disorder. *Arch. Neurol.* 55(3), 291–293 (1998).
11. Thornton CA. Myotonic dystrophy. *Neurol. Clin.* 32(3), 705–719 (2014).

• **This is a review of both DM1 and myotonic dystrophy type 2, which describes its prevalence, its genetics as well as the symptoms present in the various phenotype of the disease.**

12. Winblad S, Lindberg C, Hansen S. Cognitive deficits and CTG repeat expansion size in classical myotonic dystrophy type 1 (DM1). *Behav. Brain Funct.* 2(1), 16 (2006).
13. Baldanzi S, Cecchi P, Fabbri S *et al.* Relationship between neuropsychological impairment and grey and white matter changes in adult-onset myotonic dystrophy type 1. *NeuroImage Clin.* 12, 190–197 (2016).
14. Hamilton MJ, McLean J, Cumming S *et al.* Outcome measures for central nervous system evaluation in myotonic dystrophy type 1 may be confounded by deficits in motor function or insight. *Front. Neurol.* 9, 780 (2018).
15. Harley HG, Rundle SA, MacMillan JC *et al.* Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission in myotonic dystrophy. *Am. J. Hum. Genet.* 52(6), 1164–1174 (1993).
16. Brook JD, McCurrach ME, Harley HG *et al.* 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(4), 799–808 (1992).
17. Fu YH, Pizzuti A, Fenwick RG *et al.* An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* 255(5049), 1256–1258 (1992).
18. Mahadevan M, Tsilfidis C, Sabourin L *et al.* Myotonic dystrophy mutation: an unstable CTG repeat in the 3′ untranslated region of the gene. *Science* 255(5049), 1253–1255 (1992).
19. Sistiaga A, Urreta I, Jodar M *et al.* Cognitive/personality pattern and triplet expansion size in adult myotonic dystrophy type 1 (DM1): CTG repeats, cognition and personality in DM1. *Psychol. Med.* 40(3), 487–495 (2010).
20. Overend G, Légaré C, Mathieu J, Bouchard L, Gagnon C, Monckton DG. Allele length of the DMPK CTG repeat is a predictor of progressive myotonic dystrophy type 1 phenotypes. *Hum. Mol. Genet.* 28(13), 2245–2254 (2019).

•• **Shows that both cytosine–thymine–guanine (CTG) repeat length and the presence of a variant repeat affects disease severity particularly age at symptom onset.**

21. Cumming SA, Hamilton MJ, Robb Y *et al. 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. EJHG* 26(11), 1635–1647 (2018).
22. Morales F, Couto JM, Higham CF *et al.* 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(16), 3558–3567 (2012).
23. Légaré C, Overend G, Guay S-P *et al.* DMPK gene DNA methylation levels are associated with muscular and respiratory profiles in DM1. *Neurol. Genet.* 5(3), e338 (2019).

•• **Shows that DNAm at the *DMPK* gene locus as well as CTG repeat length explained variability of muscular and respiratory profile in DM1 patients.**

24. He F, Todd PK. Epigenetic mechanisms in repeat expansion disorders. *Semin. Neurol.* 31(5), 470–483 (2011).
25. Buckley L, Lacey M, Ehrlich M. Epigenetics of the myotonic dystrophy-associated DMPK gene neighborhood. *Epigenomics* 8(1), 13–31 (2016).
26. Steinbach P, Gläser D, Vogel W, Wolf M, Schwemmle S. The DMPK gene of severely affected myotonic dystrophy patients is hypermethylated proximal to the largely expanded CTG repeat. *Am. J. Hum. Genet.* 62(2), 278–285 (1998).
27. Filippova GN, Thienes CP, Penn BH *et al.* CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat. Genet.* 28(4), 335–343 (2001).

28. Shaw DJ, Chaudhary S, Rundle SA *et al.* A study of DNA methylation in myotonic dystrophy. *J. Med. Genet.* 30(3), 189–192 (1993).
29. Barbé L, Lanni S, López-Castel A *et al.* CpG methylation, a parent-of-origin effect for maternal-biased transmission of congenital myotonic dystrophy. *Am. J. Hum. Genet.* 100(3), 488–505 (2017).
30. Brouwer JR, Huguet A, Nicole A, Munnich A, Gourdon G. Transcriptionally repressive chromatin remodelling and CpG methylation in the presence of expanded CTG-repeats at the DM1 Locus. *J. Nucleic Acids* 2013, 567435 (2013).
31. López Castel A, Nakamori M, Tomé S *et al.* Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues. *Hum. Mol. Genet.* 20(1), 1–15 (2011).
32. Santoro M, Fontana L, Masciullo M *et al.* 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(12), 2645–2652 (2015).

- **Shows that DM1 patients with variant repeat are hypermethylated compared with those without them.**

33. Spits C, Seneca S, Hilven P, Liebaers I, Sermon K. Methylation of the CpG sites in the myotonic dystrophy locus does not correlate with CTG expansion size or with the congenital form of the disease. *J. Med. Genet.* 47(10), 700–703 (2010).
34. Yanovsky-Dagan S, Avitzour M, Altarescu G *et al.* Uncovering the role of hypermethylation by CTG expansion in myotonic dystrophy type 1 using mutant human embryonic stem cells. *Stem Cell Rep.* 5(2), 221–231 (2015).
35. Moore LD, Le T, Fan G. DNA methylation and its basic function. *Neuropsychopharmacology* 38(1), 23–38 (2013).
36. Portela A, Esteller M. Epigenetic modifications and human disease. *Nat. Biotechnol.* 28(10), 1057–1068 (2010).
37. Sato S, Nakamura M, Cho DH, Tapscott SJ, Ozaki H, Kawakami K. Identification of transcriptional targets for Six5: implication for the pathogenesis of myotonic dystrophy type 1. *Hum. Mol. Genet.* 11(9), 1045–1058 (2002).
38. Gagnon C, Petitclerc É, Kierkegaard M, Mathieu J, Duchesne É, Hébert LJ. A 9-year follow-up study of quantitative muscle strength changes in myotonic dystrophy type 1. *J. Neurol.* 265(7), 1698–1705 (2018).
39. Raymond K, Levasseur M, Mathieu J, Gagnon C. Progressive decline in daily and social activities: a 9-year longitudinal study of participation in myotonic dystrophy type 1. *Arch. Phys. Med. Rehabil.* 100(9), 1629–1639 (2019).
40. Jaeger J. Digit symbol substitution test. *J. Clin. Psychopharmacol.* 38(5), 513–519 (2018).
41. Woods SP, Delis DC, Scott JC, Kramer JH, Holdnack JA. The California Verbal Learning Test – Second Edition: test-retest reliability, practice effects, and reliable change indices for the standard and alternate forms. *Arch. Clin. Neuropsychol.* 21(5), 413–420 (2006).
42. Delis DC, Freeland J, Kramer JH, Kaplan E. Integrating clinical assessment with cognitive neuroscience: construct validation of the california verbal learning test. *J. Consult. Clin. Psychol.* 56(1), 123–130 (1988).
43. Ruff RM, Niemann H, Allen CC, Farrow CE, Wylie T. The Ruff 2 and 7 selective attention test: a neuropsychological application. *Percept. Mot. Skills* 75(3 Pt 2), 1311–1319 (1992).
44. Marsh S. Pyrosequencing applications. *Methods Mol. Biol. Clifton NJ* 373, 15–24 (2007).
45. Monckton DG, Wong LJ, Ashizawa T, Caskey CT. Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. *Hum. Mol. Genet.* 4(1), 1–8 (1995).
46. Smith CA, Gutmann L. Myotonic dystrophy type 1 management and therapeutics. *Curr. Treat. Options Neurol.* 18(12), 52 (2016).
47. Bosco G, Diamanti S, Meola G. Workshop report: consensus on biomarkers of cerebral involvement in myotonic dystrophy, 2–3 December 2014, Milan, Italy. *Neuromuscul. Disord.* 25(10), 813–823 (2015).
48. Gallais B, Montreuil M, Gargiulo M, Eymard B, Gagnon C, Laberge L. Prevalence and correlates of apathy in myotonic dystrophy type 1. *BMC Neurol.* 15, 148 (2015).
49. Cumming SA, Jimenez-Moreno C, Okkersen K *et al.* Genetic determinants of disease severity in the myotonic dystrophy type 1 OPTIMISTIC cohort. *Neurology* 93(10), e995–e1009 (2019).

- **Shows that the presence of variant repeat decreases disease severity.**

50. Winblad S, Samuelsson L, Lindberg C, Meola G. Cognition in myotonic dystrophy type 1: a 5-year follow-up study. *Eur. J. Neurol.* 23(9), 1471–1476 (2016).
51. Chouliaras L, Pishva E, Haapakoski R *et al.* Peripheral DNA methylation, cognitive decline and brain aging: pilot findings from the Whitehall II imaging study. *Epigenomics* 10(5), 585–595 (2018).
52. Mateos-Aierdi AJ, Goicoechea M, Aiastui A *et al.* Muscle wasting in myotonic dystrophies: a model of premature aging. *Front. Aging Neurosci.* 7, 125 (2015).
53. Hildonen M, Knak KL, Dunø M, Vissing J, Tümer Z. Stable longitudinal methylation levels at the CpG sites flanking the CTG repeat of DMPK in patients with myotonic dystrophy type 1. *Genes* 11(8), 936 (2020).
54. Ballester-Lopez A, Koehorst E, Almendrote M *et al.* A DM1 family with interruptions associated with atypical symptoms and late onset but not with a milder phenotype. *Hum. Mutat.* 41(2), 420–431 (2019).
55. K N, Ce P. CpG methylation modifies the genetic stability of cloned repeat sequences. *Genome Res.* 12(8), 1246–1256 (2002).

56. Santoro M, Masciullo M, Silvestri G, Novelli G, Botta A. Myotonic dystrophy type 1: role of CCG, CTC and CGG interruptions within DMPK alleles in the pathogenesis and molecular diagnosis. *Clin. Genet.* 92(4), 355–364 (2017).

57. Kawakami K, Sato S, Ozaki H, Ikeda K. Six family genes—structure and function as transcription factors and their roles in development. *BioEssays* 22(7), 616–626 (2000).

58. Larkin K, Fardaei M. Myotonic dystrophy – a multigene disorder. *Brain Res. Bull.* 56(3), 389–395 (2001).