ARTICLE

CpG Methylation, a Parent-of-Origin Effect for Maternal-Biased Transmission of Congenital Myotonic Dystrophy

Lise Barbé,[1] Stella Lanni,[2] Arturo López-Castel,[2,12] Silvie Franck,[1] Claudia Spits,[1] Kathelijn Keymolen,[3] Sara Seneca,[3] Stephanie Tomé,[2,13] Ioana Miron,[4] Julie Letourneau,[5] Minggao Liang,[2,6] Sanaa Choufani,[2] Rosanna Weksberg,[2,6,7,8,9] Michael D. Wilson,[2,6] Zdenek Sedlacek,[10] Cynthia Gagnon,[5] Zuzana Musova,[10] David Chitayat,[4,7] Patrick Shannon,[11] Jean Mathieu,[5] Karen Sermon,[1,*] and Christopher E. Pearson[2,6,*]

CTG repeat expansions in *DMPK* cause myotonic dystrophy (DM1) with a continuum of severity and ages of onset. Congenital DM1 (CDM1), the most severe form, presents distinct clinical features, large expansions, and almost exclusive maternal transmission. The correlation between CDM1 and expansion size is not absolute, suggesting contributions of other factors. We determined CpG methylation flanking the CTG repeat in 79 blood samples from 20 CDM1-affected individuals; 21, 27, and 11 individuals with DM1 but not CDM1 (henceforth non-CDM1) with maternal, paternal, and unknown inheritance; and collections of maternally and paternally derived chorionic villus samples (7 CVSs) and human embryonic stem cells (4 hESCs). All but two CDM1-affected individuals showed high levels of methylation upstream and downstream of the repeat, greater than non-CDM1 individuals ($p = 7.04958 \times 10^{-12}$). Most non-CDM1 individuals were devoid of methylation, where one in six showed downstream methylation. Only two non-CDM1 individuals showed upstream methylation, and these were maternally derived childhood onset, suggesting a continuum of methylation with age of onset. Only maternally derived hESCs and CVSs showed upstream methylation. In contrast, paternally derived samples (27 blood samples, 3 CVSs, and 2 hESCs) never showed upstream methylation. CTG tract length did not strictly correlate with CDM1 or methylation. Thus, methylation patterns flanking the CTG repeat are stronger indicators of CDM1 than repeat size. Spermatogonia with upstream methylation may not survive due to methylation-induced reduced expression of the adjacent *SIX5*, thereby protecting DM1-affected fathers from having CDM1-affected children. Thus, *DMPK* methylation may account for the maternal bias for CDM1 transmission, larger maternal CTG expansions, age of onset, and clinical continuum, and may serve as a diagnostic indicator.

Introduction

Currently, more than 40 diseases are known to be caused by expanded microsatellite repeats at specific locations in the genome. Unaffected individuals have genetically stable repeat tract lengths below a certain threshold while the expanded repeats display an unstable character. Of the 43 repeat diseases, 16 are caused by expansions of CTG/CAG repeats, including myotonic dystrophy.[1]

Myotonic dystrophy type 1 (DM1 [MIM: 160900]) is a multisystemic autosomal-dominant disease, showing a continuum of disease severity and ages of onset. DM1 mainly affects neuronal and muscular systems, where myotonia, muscle wasting, cardiac conduction defects, respiratory problems, and cataracts are among the most common symptoms. The profound genetic anticipation observed in DM1-affected families has been explained to arise by the inheritance of larger CTG expansions in subsequent generations.[2,3] Postnatal non-congenital DM1 (referred collectively here as classical DM1) shows disease onset in children and juveniles or in adults or older adults and is often reported to have CTG expansions ranges of 600 to 1,000 or 50 to 600, where the first two and last two age categories are grouped. However, there is considerable overlap and these ranges are not always the rule. The most severe form is congenital myotonic dystrophy (CDM1), which presents prenatal symptoms during pregnancy (polyhydramnios, reduced fetal movements, and preterm delivery). At birth, CDM1-affected individuals present difficulties of breathing and feeding, marked generalized hypotonia, hyporeflexia, and high perinatal mortality. A striking difference between CDM1 and DM1 is the degree of intellectual dysfunction clearly evident in CDM1-affected individuals, which generally decreases from childhood- to juvenile- and adult-onset DM1-affected individuals.[4–6] DM1 is genetically characterized by an

[1]Department for Reproduction and Genetics, Vrije Universiteit Brussel, Brussels 1090, Belgium; [2]Genetics & Genome Biology, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada; [3]Centre for Medical Genetics, Universitair Ziekenhuis Brussel, Brussels 1090, Belgium; [4]The Prenatal Diagnosis and Medical Genetics Program, Department of Obstetrics and Gynecology, University of Toronto, Toronto, ON M5G 1E2, Canada; [5]Groupe de Recherche Interdisciplinaire sur les Maladies Neuromusculaires, Saguenay, QC 75204, Canada; [6]The Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada; [7]Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, ON M5G 1X8, Canada; [8]Institute of Medical Science, School of Graduate Studies, University of Toronto, Toronto, ON M5S 1A8, Canada; [9]Department of Pediatrics, University of Toronto, Toronto, ON M5S 1A1, Canada; [10]Department of Biology and Medical Genetics, Charles University Prague, Prague 128 00, Czech Republic; [11]Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, University of Toronto, Toronto, ON M5G 1X5, Canada
[12]Present address: Genera Biotech, Valencia 46980, Spain
[13]Present address: INSERM, Laboratory of CTG Repeat Instability and Myotonic Dystrophy, Paris 75654, France
*Correspondence: cepearson.sickkids@gmail.com (C.E.P.), karen.sermon@uzbrussel.be (K.S.)
http://dx.doi.org/10.1016/j.ajhg.2017.01.033.

expanded CTG repeat in the 3′ untranslated region of the dystrophia myotonica-protein kinase (*DMPK* [MIM: 605377]) gene on chromosome 19.[7,8] CDM1 is almost exclusively associated with maternal transmission and it has been suggested that it is linked to large repeat size (>1,000 repeats),[9–11] but this link is not true for all CDM1-affected individuals. Only a handful of rare paternally transmitted CDM1-affected case subjects are known.[12–17] However, many CDM1-affected individuals inherit shorter CTG tracts than some classical DM1-affected individuals and many individuals with classical DM1 have expansions considerably larger than 1,500 repeats.[18–20] For example, numerous individuals with CDM1 have repeats in the classical DM1 range, some with as few as 550 repeats, indicating that other unknown factors must contribute to CDM1.[10,20–26] Moreover, prenatal tissues (amniocentesis or chorionic villus sampling) from pregnancies that led to the birth of CDM1-affected children can have repeat lengths considerably shorter than 1,000 repeats, even fewer than the transmitting mother—complicating a definite prenatal diagnosis based only upon repeat length.[18,24,26–30] Similarly, some individuals with CTG expansions >1,000 repeats present with very mild symptoms with late onset, one case as late as 44 years old.[18–20] Ongoing somatic CTG repeat expansions can hamper correlations of repeat length to disease state.[31] Correction for somatic instability by estimating the inherited progenitor allele can improve genotype-phenotype relationships.[31] While such assessment and interpretations of repeat length might improve genotype-phenotype correlations, the existence of CDM1-affected individuals having <1,000 CTG repeats[10,20–26] and non-congenital DM1-affected individuals with expansions considerably larger than 1,500 repeats[18–20] argues against repeat length as the sole determinant of either the maternal bias or disease etiology of CDM1. Together these findings suggest that some maternal factors other than repeat size may be linked to the manifestation of CDM1.

The DM1 CTG repeat is located in a 3.5 kb CpG island, with two putative CCCTC-binding factor (CTCF) sites flanking the CTG repeat (Figure 1). Binding of CTCF to the CTCF binding sites, together with the DM1 CTG repeat, was suggested to establish an insulator element between the *DMPK* promoter and the six homeobox 5 (*SIX5* [MIM: 600963]) enhancer.[32] Limited data from a single cell line, derived from a fetal termination (without clinical assignment), suggested that hypermethylation of the region upstream of the expanded CTG tract is correlated with suppressive histone marks.[32–34]

López Castel et al.[35] reported CpG methylation upstream but not downstream of the CTG repeat in classical DM1 tissues from adults and fetuses as well as a loss of methylation levels with aging. Both López Castel et al.[35] and Brouwer et al.[36] assessed methylation patterns in DM1 mice and the latter study found that the increase in repeat size is correlated with an increase in up- and downstream methylation in DM1 transgenic mice. More recently, Yanovsky-Dagan et al.[37] investigated the effect of CpG methylation on *SIX5* expression in human DM1 embryonic stem cell lines (hESCs) that carry CTG expansions. That study revealed the importance of the DM1 locus as a regulatory element that can become dysfunctional due to epigenetic changes linked to an enlarged CTG repeat.[37]

To gain insight into the relationship between methylation, expansion size, and clinical presentation, specifically CDM1, we determined the CpG methylation state upstream and downstream of the *DMPK* CTG repeat in a large cohort of clinically characterized DM1-affected individuals and samples. Peripheral blood DNA from a total of 79 individuals, of which 20 had the congenital form, was investigated, as well as seven chorionic villus samples (CVSs), one fetal sample (cultured fibroblast from a skin biopsy), and one sperm sample, representative of different developmental stages from gamete to adult. We were able to analyze several affected family members over two or three generations, leading to a detailed view of the mechanisms of inheritance of CDM1. As a proxy for an earlier developmental stage, we also analyzed four DM1-affected hESC lines and blood from their affected donors, included in our analysis of classic DM1, as well as three control non-DM1-affected hESC lines. For all samples, we analyzed 25 CpG sites upstream of the CTG repeat (including 5 sites in the CTCF1 binding site) and 11 CpG sites downstream of the CTG repeat (including the 3 sites in the CTCF2 binding site) to identify the methylation status in hundreds of individual alleles. Interestingly, a close to 100% correlation was found between methylation upstream of the CTG repeat and CDM1 (19 of 20 CMD1-affected individuals were methylated). Moreover, a significantly lower number of classical DM1-affected individuals (2 out of 59 samples) showed methylation upstream ($p = 7.04958 \times 10^{-12}$). This could indicate that CpG methylation is a contributing factor to the development of CDM1. Moreover, our results argue against CTG repeat size as the sole determining factor for CDM1, since in our cohort there is a significant overlap in CTG repeat size between classical DM1- and CDM1-affected individuals.

Material and Methods

DM1-Affected Individual Samples

92 pre- and postnatal samples were analyzed for repeat size and CpG methylation up- and downstream of the CTG repeat. DM1-affected individual classifications conformed to prior published criteria.[4–6,24,27,38–46] Congenital DM1 (CDM1) showed prenatal, perinatal, and neonatal symptoms; childhood/infantile DM1 showed onset after the first year during the first decade; juvenile DM1 showed onset between 10 and 20 years of age; adult DM1 showed onset from 20 to 40 years of age; and late DM1 showed onset at >40 years of age (Tables 1 and S1). Non-congenital DM1 is collectively referred to here as classical DM1. 20 individuals had CDM1, 59 had classical DM1, and for 7 CVSs (samples B-29, B-30, B-33, B-34, B-36, B-38, and B-40) and one fetal sample

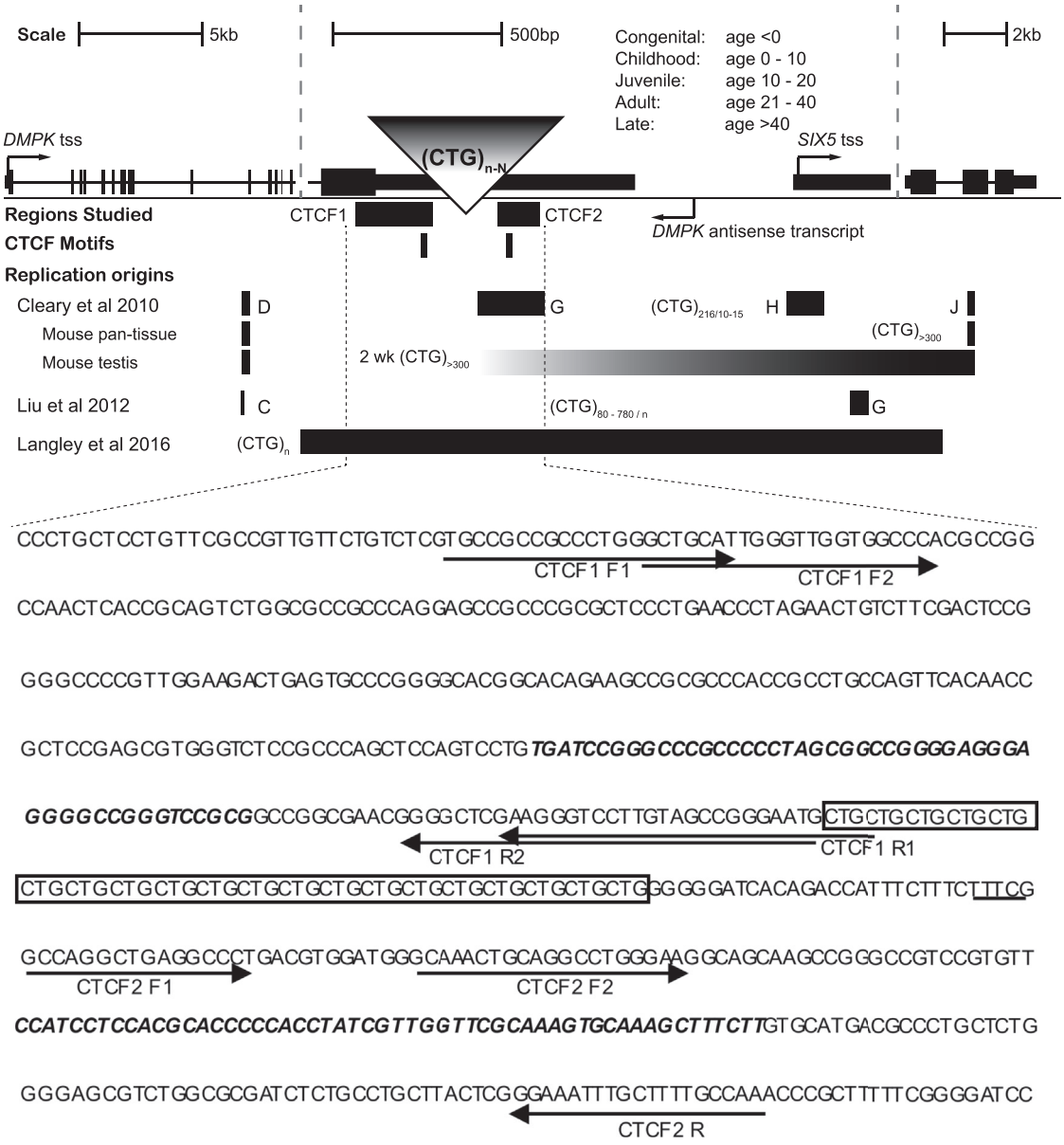

Figure 1. The DM1 Locus
The DM1 locus, associated genes, and mapped functional regions are schematically shown. The CTG repeat is located in the 3′ UTR and *SIX* promoter, of which part of the DNA sequence is shown. The initiation sites for DNA replication in control,[92,110,111] DM1 individuals cells,[92,110] and DM1 transgenic mouse tissues[92] of various ages are indicated. Primers used for PCR amplification of the CTCF binding sites are listed in Table S1 and indicated in the sequence in this figure (arrows). The CTCF binding sites up- and downstream of the repeat are in bold and italics.

(B-31), the disease type could not be identified. Except for the CVSs, one fetal sample, and one sperm sample (T-6S, adult DM1), all samples were DNA extracted from peripheral blood. The sperm sample was collected using a method that contains a step with SDS treatment prior to lysing the sperm heads, which eliminates any contaminating non-germline cells.[47] 41 DM1-affected individuals were members of the 19 families analyzed in this report. All individuals signed informed consent forms allowing their DNA to be used for research purposes and all procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation. Table S1 shows the description of the individuals by family, sex, repeat size, inheritance, disease type and subtype, age at sampling, and methylation status. Sample numbers starting with a "B" were collected at the UZ Brussel, Belgium; samples starting with a "T" were collected in Saguenay, Quebec, Canada in collaboration with The Hospital for Sick Children, Toronto, Canada; and samples starting with a "P" were collected at the Motol University Hospital of Prague, Czech Republic.

hESC Culture

Human embryonic stem cell lines used in this study are VUB01, VUB04_CF, VUB13_FXS, VUB03_DM1, VUB19_DM1, VUB24_DM1, and VAL6M (Table S1). All lines were derived with individuals' informed consent and after obtaining permission

Table 1. Summarized Samples and Methylation Data

	No. of Samples	Inheritance	No. of Samples	Disease Form	Methylation Upstream	Expansion Range (CTG)n-N	Methylation Downstream
Controls[a]	10	N/A	10		0%	non-affected	0%
DM1	59	paternal	27	late-onset	(0/1) 0%	1,500	(0/1) 0%
				adult	(0/22) 0%	57–2,800	(1/22) 4.8%
				juvenile	(0/3) 0%	500–1,600	(0/3) 0%
				childhood	(0/1) 0%	650	(0/1) 0%
		maternal	21	late-onset	(0/1) 0%	140	(0/1) 0%
				adult	(0/12) 0%	90–1,400	(1/12) 8.3%
				juvenile	(0/2) 0%	500	(2/2) 100%
				childhood	(2/6) 33.3%	500–2,000	(4/6) 66.6%
		unknown	11		0%	54–2,000	0%
CDM1	20	maternal	20		(19/20) 95%	1,100–4,700	(19/20) 95%
Sperm	1	paternal	1		0%	230	0%
hESCs	4	paternal	2		0%	140–500	50%
		maternal	2		100%	1,800–2,100	100%
CVS	7	paternal	3		0%	250–1,100	100%
		maternal	4		100%	<300–2,200	100%
Fibroblasts	1	paternal	1		0%	200	0%

The table shows grouped and summarized data for all DM1 samples analyzed in this study. For each sample, estimated repeat size, inheritance, disease form, and methylation up- and downstream of the CTG repeat are listed. DM1-affected individual classifications were according to previously published criteria.[4,5,24,27,38,39,41–46] Congenital DM1 (CDM1) showed prenatal, perinatal, and neonatal symptoms; childhood/infantile DM1 showed onset after the first year during the first decade; juvenile DM1 showed onset between 10 and 20 years of age; adult DM1 showed onset between 20 and 40 years of age; late DM1 showed onset at >40 years of age.
[a]Controls included a total of >60 non-affected individuals for the threshold (see Material and Methods) and 13 sequenced samples, shown here. For details, please see complete data outlined in Table S1.

from the Commission for Medical Ethics of the UZ Brussel and from the Federal Commission for Research on Embryos. VAL6M was obtained from the Centro de Investigación Príncipe Felipe, Valencia, Spain. All lines are registered with the EU hESC registry (see Web Resources). The female donors of the cell lines VUB03_DM1 and VUB24_DM1 were affected, and for VUB19_DM1 and VAL6M the male donors were affected. VUB01, VUB04_CF, and VUB13_FXS were used as control non-DM1 lines. The hESC lines were analyzed at following passages (P): VUB03_DM1 at P45, VUB19_DM1 at P71, VUB24_DM1 at P41, VAL6M at P32, VUB01 at P306, VUB04_CF at P66, and VUB13_FXS at P50. The control non-DM1 lines have two non-expanded DM1 alleles: VUB01 has 5 and 11 CTG repeats, VUB04_CF has 5 and 13 repeats, and VUB13_FXS is homozygous with 5 repeats.

All hESC lines were confirmed to be chromosomally balanced using array-comparative genomic hybridization, carried out according to the manufacturer's instructions with minor modifications (Human Genome CGH Microarray 4x44K, Agilent Technologies), as previously described.[48]

Human ESCs were cultured on inactivated CF1 mouse embryonic fibroblasts (MEFs; 2–4 × 10^4 cells/cm^2) at 37°C, 5% CO_2, and atmospheric O_2 conditions in hESC medium consisting of knockout D-MEM (Invitrogen) supplemented with 20% Serum Replacement (SR) (Invitrogen), 2 mM L-glutamin (Invitrogen), 0.01 mM non-essential amino acids (NEAA) (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma Aldrich), 4 ng/mL human recombinant basic fibroblast growth factor (bFGF) (Invitrogen), and 100 U/mL penicillin and streptomycin (Pen Strep) (Invitrogen). Medium was changed daily and mechanical passaging was performed every 6 days.[49]

CTG Repeat Length Analysis

Repeat length of non-expanded alleles were estimated by PCR amplification across the repeat. 10 ng of DNA was amplified by PCR using Amplitaq polymerase (Thermo Fisher). The total reaction volume was 25 μL, containing 1× PCR Buffer (Thermo Fisher), 2 mM $MgCl_2$ (Thermo Fisher), 0.2 mM dNTPs (Illustra DNA polymerization mix, GE Healthcare), 0.4 μM DM101 and DM102 primers (Integrated DNA Technologies (IDT),[8] and 1.25 U Amplitaq polymerase. Primer sequences are listed in Table S2. DM101 primers were FAM-labeled in order to detect the length of the repeat by fragment analysis. Fluorescent PCR products were analyzed by capillary electrophoresis using the 3130xl or 3730 Genetic Analyzer (Applied Biosystems). Fragment analysis was performed with Gene Mapper 4.0 Software (Applied Biosystems).

For analysis of the expanded CTG, a specific PCR for long fragments followed by Southern blot was performed in DM1 samples as described previously.[50,51] 10 to 100 pg of DNA was amplified using the LongAmp Taq PCR Kit (New England Biolabs) with 2.5 units LongAmp Taq DNA polymerase, 1x LongAmp buffer (New

England Biolabs), 0.2 mM dNTPs, and 0.4 μM of primers DM101 and DM102[8] (IDT) (sequences in Table S2) in a total reaction volume of 25 μL on a Veriti thermal cycler (Life Technologies). PCR conditions were as follows: 4 min of initial denaturation at 94°C, 35 cycles of 30 s denaturation at 94°C, 8 min annealing and extension at 65°C, and a final extension step at 65°C for 10 min. After denaturing 20 PCR products per sample by boiling, they were separated overnight on a denaturing alkaline agarose gel.[52] DNA fragments were transferred to a positively charged nylon membrane (Roche) by alkaline transfer. Hybridization with a digoxigenin labeled, DM1-repeat specific probe, 5′-/5DigN/CAGCAGCAGCAGCAG-3′ (IDT), was performed overnight and detected by chemoluminescence using the anti-Dig-CSPD system (Anti-Digoxigenin-AP Fab Fragments and CSPD, Roche). Chemoluminescent signals were visualized after film exposure (Carestream). For some samples from Quebec, CTG sizing was performed by Southern blotting of 3–5 μg of blood DNA, digested with EcoRI using the 2.2 kb BamHI/EcoRI subclone of probe pGB2.66, as previously described.[53] For some samples from the Czech Republic, CTG sizing was by Southern blotting using 10 μg of genomic DNA digested with EcoRI, and a PCR-probe, as previously described.[54] For ease of presentation and comparison, all CTG expansions were rounded to the nearest factor of ten (Table S1).

Ongoing somatic expansion-biased instability of the CTG tract throughout the lives of DM1- and CDM1-affected individuals may lead to over- and under-estimates in CTG lengths in more mildly affected individuals and CMD1-affected individuals, where the former are often sampled at older ages and the latter typically sampled at birth.[35,55–63] While somatic expansion variations for the various ages of blood DNA samplings used herein (Table S1) may hamper a direct comparison between all individuals, it is noteworthy that the degree of somatic instability in DM1- and CDM1-affected individuals is considerably lower in the blood and greater in muscle/heart.[35,59–63] A more direct comparison would be possible if the progenitor alleles of each sample were known, essentially accounting for a confounding effect of somatic instability and age at sampling. While methods to estimate the CTG size of the progenitor allele in a sample have been developed,[31,58] this technically demanding estimation is beyond the scope and capacity of our study and DNA availability. Our study and all other published studies, excluding a few,[31,58] cannot account for these variables. Considering our address to variables of age-stability of methylation, the absence or low correlation of age at sampling with up- or downstream methylation, the broad range of age at sampling for each disease class, the range of developmental stages, an absence of a correlation of repeat size with methylation status, our sample size, and our current statistical power, we do not feel that the strength of our correlations depends upon accounting for the age at sampling.

Bisulfite Treatment and Sanger Sequencing

Methylation around the CTG repeat was first screened by bisulfite treatment and Sanger sequencing to determine whether methylation hallmarks existed in the different sample types.

200 ng of DNA was bisulfite treated using the Imprint DNA Modification Kit (Sigma Aldrich) using the one-step modification procedure and DNA was eluted in 20 μL elution buffer. Bisulfite-treated DNA was amplified by nested and hemi-nested PCR for the upstream (CTCF1) and downstream (CTCF2) region, respectively (Table S2), using the Jumpstart Taq DNA polymerase kit (Sigma Aldrich) on a Veriti thermal cycler (Life Technologies). For the first PCR, 50 ng of bisulfite-treated DNA was added to a reaction mix consisting of 1.25 units Jumpstart Taq DNA polymerase, 1× reaction buffer with $MgCl_2$ (Sigma Aldrich), 0.2 mM dNTPs, and 0.4 μM primers (IDT) in a total reaction volume of 25 μL. PCR conditions for the first PCR were as follows: 5 min of initial denaturation at 94°C, 40 cycles of 30 s denaturation at 94°C, 30 s annealing at 55°C, and 30 s of extension at 72°C. A final extension step of 5 min at 72°C was applied. For the second PCR, 3 μL of PCR product from the first round was amplified using the same PCR reaction mix, with second round primers, and the same thermocycler conditions. Primer sequences are listed in Table S2; the number 1 or 2 at the end of the name indicates the PCR round. For the downstream region (CTCF2), the same reverse primer is used for both PCR rounds.

Amplicons were purified using the High Pure PCR Product Purification kit (Roche) and cycle-sequenced using BigDye chemistry 3.1 according to the manufacturer's instructions (Life Technologies). Afterward, amplicons were run on an ABI 3130XL automatic sequencer (Applied Biosystems). Results were analyzed using Applied Biosystems Sequence Scanner v.1.0 and aligned to the reference sequence using BiQ Analyzer.[64] Samples were considered to show methylation and were processed further for Massive Parallel Sequencing (MPS) when Sanger sequencing showed two peaks at one or more of the CpG sites, one for a C (methylated) and one for a T (unmethylated), including those where the two peaks signals were unequal in height.

Bisulfite Treatment and Massive Parallel Sequencing

All samples that showed some degree of methylation after Sanger sequencing, as well as the appropriate negative controls, were subjected to MPS for a detailed examination of the methylation status at every CpG site for a large number of alleles simultaneously. For samples with limited amounts of DNA, MPS was performed immediately, skipping Sanger sequencing analysis. For each sample, MPS yielded about 4,000 reads of which 100 sequences for both the upstream and downstream region were randomly selected and compared to the reference sequence. Bisulfite treatment as well as the first round of PCR was performed as described above, using only a different reverse primer for the upstream region (MPS CTCF1 R; Table S2). The second round of PCR used primers with linker sequences containing the next generation first read (forward) or second read (reverse) primers attached to an adaptor sequence. This adaptor sequence was used for library preparation by attaching indices and P5 or P7 flow cell binding sites. The PCR reaction mix consisted of 1.25 units Jumpstart Taq DNA polymerase, 1× reaction buffer with $MgCl_2$, 0.2 mM dNTPs, and 0.4 μM primers in a total reaction volume of 25 μL. PCR conditions included 5 min of initial denaturation at 95°C, 40 cycles of 30 s denaturation at 95°C, 30 s annealing at 65°C, and 30 s of extension at 72°C. A final extension step of 7 min at 72°C was applied. Primer sequences are listed in Table S2 with the suffix "Miseq." Library preparation was carried out by PCR using the NEBNext High-Fidelity 2X PCR Master Mix (Bioké) according to manufacturer's instructions, with Illumina Nextera indexed primers (Illumina), and with the following PCR program: 30 s at 98°C, 15 cycles of 10 s at 98°C, 30 s at 65°C, and 30 s at 72°C, followed by the final extension step of 5 min at 72°C. The products were purified using magnetic bead purification and size selection using a 0.7

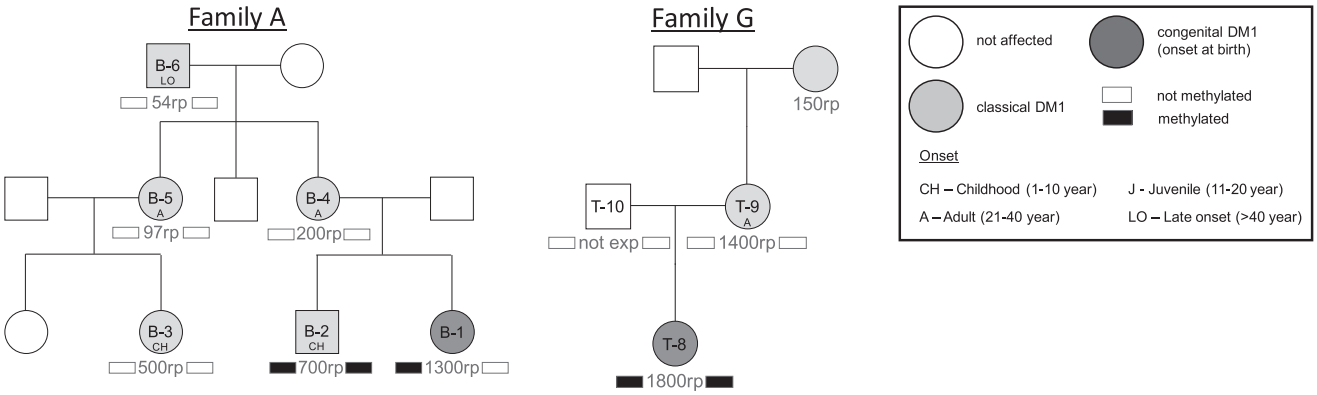

Figure 2. Pedigrees of Families A and G
Pedigree symbols indicate non-affected individuals as hollow, non-CDM1 individuals are light gray, and congenital DM1-affected individuals are filled. Disease subtype is represented as Ch for childhood onset, A for adult onset, and Lo for late onset. For each family member, the number of expanded CTG repeats and methylation up- and downstream of the CTG repeat is shown. Repeat sizes in non-affected individuals are not expanded and thus denoted as "not exp." Methylated CpG sites upstream and downstream of the CTG repeat are schematized in bold; non-methylated sites are hollow. All DNA analyzed was extracted from blood. For all pedigrees assessed in this paper, please see Figure S2.

beads-to-sample ratio (Analis). The libraries were spiked in with a 20% PhiX sequence (Illumina) to compensate for low complexity. The final libraries were loaded on the MiSeq Reagent Nano Kit v2 (500 cycles) according to manufacturer's instructions and sequenced at 2x250 bp (Illumina).

Sequencer output files were processed using a homemade script. Sequence lengths less than 150 bp were discarded and sequences of the upstream region were selected out from sequences of the downstream region by reference sequence similarity. For each sample, 100 sequences were randomly selected for both regions for analysis. CpG methylation analysis was performed using BiQ Analyzer, a free program that aligns the reference sequence to the 100 randomly selected sequencing files using Clustal W alignment.[64] Sequences with less than 90% bisulfite conversion rate and/or less than 80% sequence identity with the reference sequence were discarded.

Methylation was quantified per allele for each sample, by calculating the percentage of methylated CpG sites of the total number of potentially methylated sites in the upstream and downstream regions, and each allele was binned into categories according to their methylation percentage: no methylation (under threshold), low methylation (threshold to 35%), medium methylation (35%–65%), and high methylation (>65%). Methylation thresholds were set based on sequencing results of DNA from the control subjects, non-DM1-affected individuals (>60) for the DM1-affected individuals, and non-DM1 hESCs (VUB01, VUB04_CF, and VUB13_FXS) for the DM1 hESCs. Threshold selection: for the non-DM1-affected individuals, 50 samples used for whole-genome methylation analysis previously described,[65] ten non-affected relatives from DM1-affected families (Table S1), and three non-related control subjects were analyzed. An allele was considered to be methylated if at least two or more CpG sites were methylated. To avoid misinterpretation of the data due to outliers, a 10% interval was used. This means that a sample was considered as methylated only if more than 10% of the alleles showed methylation (with ≥2 CpG site being methylated). All the non-DM1-affected individuals and non-DM1 hESC lines had at least 90% of their alleles showing no methylation, being none or maximum 1 CpG site that showed methylation.

Statistical Analysis

DNA methylation for each sample was established in hundreds of individual alleles by DNA bisulfite conversion and massive parallel sequencing. Linear regression analysis was performed for the region upstream (CTCF1) and downstream (CTCF2) of the repeat separately on the samples for which the repeat size, inheritance, age at sampling, and disease form were known (Table S4). Analyses were performed where CpG methylation is the dependent variable and the independent variables are as follows: gender (male versus female), repeat size of the expanded allele, disease (classical DM1 versus CDM1), and age of onset (Tables S4A and S4B); repeat size of the expanded allele as a continuous variable (Tables S4C and S4D); inheritance (paternal versus maternal) (Tables S4F and S4G); maternal inheritance (maternally transmitted non-congenital DM1 versus CDM1) (Tables S4H and S4I); and DM1 disease form (childhood and juvenile versus adult and late onset) (Tables S4J and S4K). Further analyses were performed where repeat size of the expanded allele is compared to disease status (non-congenital DM1 versus CDM1) (Table S4E).

Results

DM1 Samples, CTG Lengths, and Methylation Study Parameters

We collected 92 DM1 samples, including a total of 79 blood samples (59 classical and 20 congenital myotonic dystrophy-affected individuals), 7 CVSs, 1 fetal sample, 1 sperm sample, as well as 4 hESC lines carrying the DM1 mutation and blood DNAs of their affected donors (included in the classical DM1-affected individuals) (Tables 1 and S1). 41 DM1 samples were part of 19 families for which also 10 non-affected relatives were analyzed (pedigrees are shown in Figures 2, S1A, and S2). Repeat lengths were determined (Figures S1B and S1C) and are indicated in the pedigrees (Figures 2 and S2) and Table S1.

Methylation was determined and quantified. DNA methylation was independently determined for regions 300 bp upstream (hg19: 46,277,287–46,277,059) of the CTG repeat and 229 bp downstream (hg19: 46,276,890–46,276,767; Figure 1) by bisulfite conversion and sequencing (Figures S1D–S1F).

In non-affected individuals with non-expanded CTG tracts, the DM1 region we studied herein has been epigenetically characterized.[66] Buckley et al. found the DM1 region to exist as a long, constitutively unmethylated region[66] in a 3.5 kb CpG island,[67] a pattern that was consistent among 16 non-affected tissues (blood, muscle, heart, brain, liver, etc.).[66] The methylation-free zone spanned from 800 bp upstream to 2.2 kb downstream of the CTG tract.[66] Our methylation analyses of control blood DNA from 13 non-affected individuals revealed only low background levels of methylation (Table 1, Table S1). Based upon the above findings, we set our threshold of methylation at more than one CpG site that is methylated and this in more than 10% of the alleles per sample to exclude outliers showing abnormal methylation (Material and Methods, section Bisulfite Treatment and Massive Parallel Sequencing).

Methylation Status in Classical DM1-Affected Individuals

We determined CpG methylation status of the region surrounding the DM1 CTG repeat tract in our collection of 60 classical DM1-affected individuals (59 blood samples and 1 sperm), of which 24 were part of 19 families (Figure S2). Overall, only 2 out of 21 classical DM1-affected individuals with maternal inheritance of the expansion showed methylation both upstream and downstream of the CTG tract (Figures 3 and S4, Table S3). All other DM1-affected individuals with maternal inheritance, as well as those with paternal inheritance, showed methylation only downstream of the repeats or absence of methylation in both regions analyzed (Figure 3). Results are detailed below per family, group of families, and/or DNA sample origins.

Classical DM1-Affected Family A

Family A is a three-generation DM1-affected family with five classical DM1-affected individuals and one CDM1-affected individual. Family A is representative of the methylation results obtained in all 19 families (Figure S1, pedigrees in Figures 2 and S2). Repeat sizes increased from generation to generation, going from 54 repeats (B-6) to 97 (B-5) and 200 (B-4) repeats in the second generation, and to 500 (B-3), 700 (B-2), and 1,300 (B-1) repeats in the third generation. There were no signs of CpG methylation upstream or downstream of the DM1 repeat (Figure 3) in the DNA samples of the second-generation daughters (B-4 and B-5) that inherited the disease paternally (see blue bubbles). The third-generation offspring B-3, with maternally inherited expansion, did not show any signs of CpG methylation upstream or downstream of the DM1 repeat. The offspring of the DM1-affected individual B-4—a classical DM1-affected individual (B-2) and a CDM1-affected individual (B-1)—showed some methylation at both regions analyzed in individual B-2 (41% of methylated alleles upstream as well as 41% downstream) and large increases in repeat size and methylation upstream of the repeats in the CDM1-affected individual (B-1) (95% of methylated allele in the upstream region; 50% of alleles with high methylation) (Figures 3 and S4). Repeat sizes are listed in Table S1 and in the pedigree (Figures 2 and S1A), methylation results can be found in Figures 3 and S4, and detailed percentages of methylation in Tables S3A and S3B.

Classical DM1-Affected Family B

Family B consists of a classically affected mother (B-9) transmitting classical DM1 to her two daughters (B-7 and B-8, pedigree in Figure S2, Table S1). Methylation appeared in the two daughters downstream of the CTG repeat (32% and 27% of methylated alleles for B-7 and B-8, respectively). There was no methylation present upstream of the CTG repeat. The mother did not show methylation at either site (Figures 3 and S4, Tables S3A and S3B).

Classical DM1-Affected Families C, D, and E

Families C, D, and E consist of classical DM1-affected individuals (B-10, T-1, and T-3) with 900 (T-1) and 1,000 (B-10 and T-3) repeats who inherited DM1 maternally (pedigrees shown in Figure S2). For family C, DNA of the affected parent (B-11, 160 repeats) could be analyzed. For family D, DNA only from the unaffected parent (T-2) was available, while for family E, DNA from both the affected parent (T-4, 650 repeats) and the unaffected parent (T-5) was available. For all three families, no methylation was found in the parents. B-10 showed methylation both upstream and downstream of the CTG repeat, while T-1 and T-3 only showed methylation downstream of the repeat (Figures 3 and S4, Tables S3A and S3B).

Classical DM1-Affected Family F, Including Sperm Sample

Family F consists of a classically affected father (T-7), with 57 repeats, who transmitted classical DM1 to his son (T-6), with 100 repeats (blood) (pedigree in Figure S2). DNA from T-6S sperm, with 230 repeats, and from T-7 blood was analyzed for methylation up- and downstream of the CTG repeat. No methylation at either sites was found in the father or son.

Classical DM1-Affected Individual Samples without DNA from Relatives

The blood DNA of an additional 36 classical DM1-affected individuals was analyzed for methylation at the DM1 locus. Of these, 9 had inherited the disease maternally and 17 paternally, while for 10 samples the inheritance was not known. The methylation analysis of these individuals is summarized in Figure 3. Overall, 2 out of 21 classical

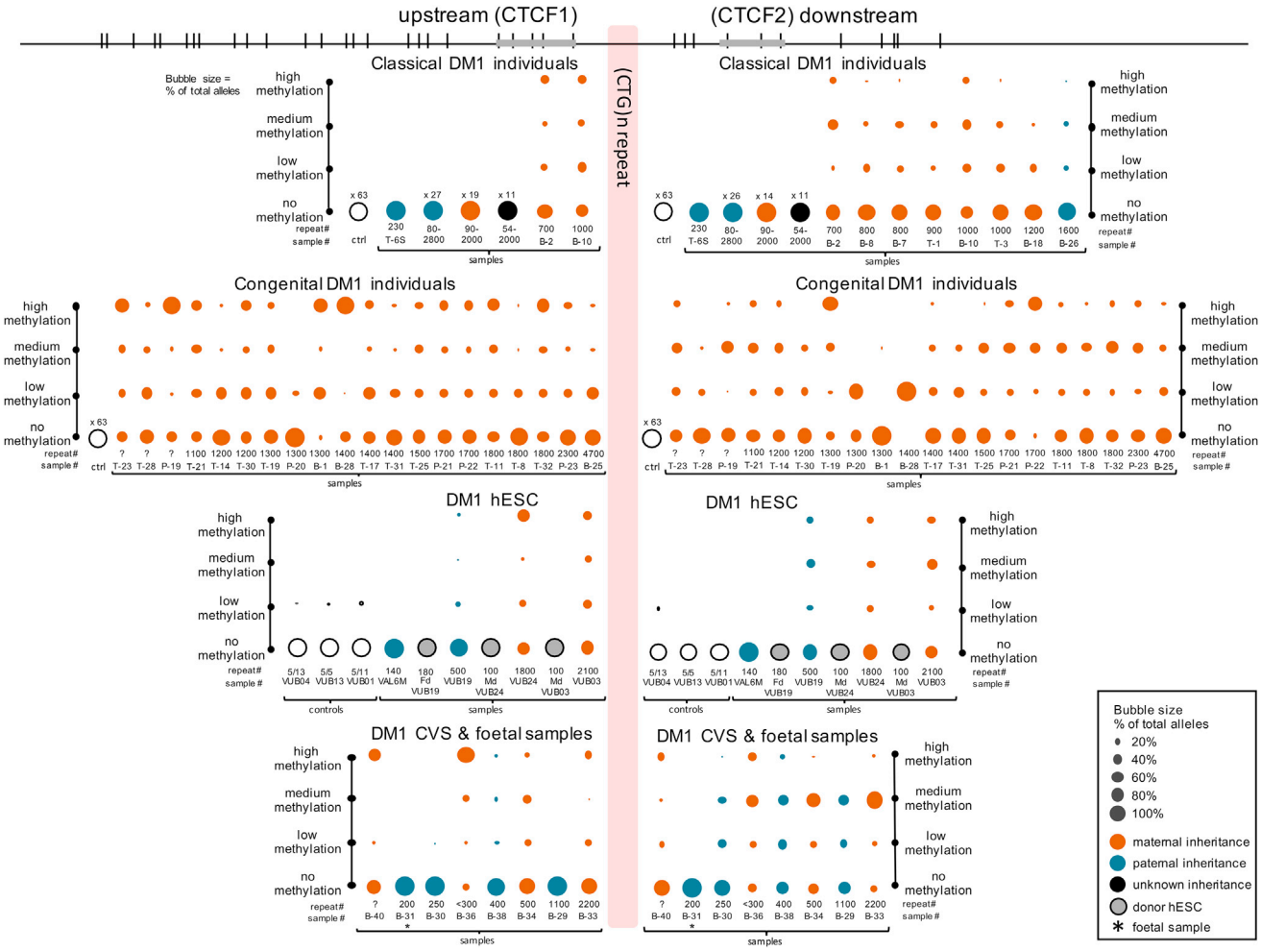

Figure 3. Methylation Up- and Downstream of the CTG Repeat in Classical and Congenital DM1-Affected Individuals, hESC Lines, and CVSs and Fetal Samples

CpG methylation up- and downstream of the CTG repeat was analyzed in all samples. For the upstream site we analyzed 25 CpG sites (indicated by the vertical lines) of which five were part of a CTCF binding site (CTCF1; indicated by the gray box). For the downstream site we analyzed 11 CpG sites of which three were part of a CTCF binding site (CTCF2; indicated by the gray box).

Methylation is summarized in plots for regions upstream and downstream of the repeats. One hundred epi-alleles were randomly selected for each sample. The methylation percentage per allele for the 25 upstream CpG sites and the 11 downstream CpG sites was calculated, and these were binned into four categories according to methylation levels (no methylation [under threshold], low methylation [threshold to 35%], medium methylation [>35%–65%], and high methylation [>65%]). Categories were established based on MPBS of more than 60 non-DM1 samples. In all control samples, more than 90% of the alleles had no methylation as defined in the Material and Methods section. The y axis plots the percentage of CpG site methylation per allele upstream or downstream of the repeats, while the different samples (including the repeat size of the largest allele of each sample) are plotted on the x axis. The bubble size reflects the number of alleles in a given sample that show this percentage of methylation. Maternally inherited DM1 alleles are colored in orange and paternally inherited DM1 alleles are colored in blue. DM1 samples with unknown inheritance are colored in black. Germline donors of the hESC lines are shown in gray.

For the classical DM1-affected individuals, many samples showed no methylation. These samples were grouped according to inheritance (maternal, paternal, and unknown) and the number of affected individuals without methylation for each group is indicated above the bubble, while the CTG repeat size range is indicated below. CVS stands for chorionic villus sample. In this bubble plot, the fetal sample is also included and indicated with an asterisk (*).

For complete methylation details, please see lollipop diagrams in Figure S4 and Table S3.

individuals with maternal inheritance showed methylation upstream and 7 out of 21 showed methylation downstream of the repeat. This is a slightly higher methylation level compared to the individuals with paternal inheritance of DM1 showing no methylation upstream and only 1 out of 27 individuals showed methylation downstream of the repeat. No methylation at either CTCF site was found for all samples for which the inheritance was not known. Detailed results for the methylated samples are shown in Figure S4 and Tables S3A and S3B.

Congenital DM1-Affected Individuals

20 individuals with CDM1 were analyzed. All CDM1-affected individuals inherited the expanded allele from

the mother. DNA from one or both parents was available for ten of the congenital individuals. For two congenital individuals, a classical DM1-affected (B-2) and an unaffected (T-27) brother (families A and N, respectively) were also available. For all families where parental DNA was available, the repeat size increased upon transmission to the CDM1-affected offspring, with increases of 400 up to 1,050 repeats in a single generation (Table S1, Figure S2).

Congenital individuals showed methylation above the set threshold (two or more CpG sites methylated per CTCF site in >10% of the alleles) at both CTCF binding sites (95% of the affected individuals are methylated upstream and 95% of the affected individuals are methylated downstream). In all of the CDM1-affected individuals, around half of the alleles showed no methylation either up- or downstream of the repeat. These alleles most likely represent the non-mutant DM1 allele. Most of the expanded alleles showed methylation (Figure 3), of which a few alleles showed high methylation with more than 65% of methylated CpGs. Exact percentages per group are listed in Tables S3C and S3D and lollipop diagrams are shown in Figure S4. In all families, the affected parent showed no methylation either up- or downstream of the repeat. CDM1 transmission concurred with an expansion of the repeat. Although there were large differences in repeat sizes between CDM1-affected individuals, ranging from 1,100 to 4,700 repeats (Figure S2, Table S1), there was a considerable overlap in CTG repeat size ranges with classical DM1-affected individuals, whom in our study had repeat sizes from 54 to 2,800 repeats. However, there was a significant difference in methylation between our CDM1-affected individuals and the classical DM1-affected individuals ($p = 7.04958 \times 10^{-12}$ upstream, $p = 9.55873 \times 10^{-5}$ downstream) (Tables S4A and S4B).

Methylation Status in DM1 hESC Lines

Human embryonic stem cell lines are derived from the inner cell mass of 6-day-old preimplantation embryos and therefore can be used as a proxy for this developmental stage. We used hESCs derived from embryos shown to carry the expanded allele after preimplantation genetic diagnosis for DM1; for VUB03_DM1 and VUB24_DM1, the affected germline donor was female, and for VUB19_DM1 and VAL6M, the affected germline donor was male. We also used three non-DM1 hESC lines (VUB01, VUB04_CF, and VUB13_FXS) as control lines. For the three control lines, Sanger sequencing and next generation sequencing after bisulfite treatment showed no CpG methylation either up- or downstream of the CTG repeat (Figure 3).

VUB19_DM1 and VAL6M derived from a male affected germline donor showed no methylation at the upstream site, while maternally derived VUB03_DM1 and VUB24_DM1 showed methylation in about half of the alleles (Figures 3 and S4, Tables S3E and S3F).

All DM1 hESC lines except for VAL6M, of paternal origin, showed methylation at the downstream site. In particular, half of the alleles were partially methylated, as illustrated in Figure 3. Exact percentages are listed in Table S3F and lollipop diagrams are in Figure S4. About half of the alleles were not methylated and were presumed to be the non-expanded allele: VUB03_DM1 showed 47% unmethylated alleles, VUB19_DM1 showed 54%, and VUB24_DM1 showed 56% (Figure 3 and Table S3F).

Blood DNA of the affected germline donors of the embryos from which the DM1 hESC lines were derived was also analyzed (Table S1). None of the germline donors' blood DNA samples showed methylation either up- or downstream of the CTG repeat (Figure 3 and Tables S3E and S3F).

Chorionic Villus Samples

We collected CVSs from four families—P, Q, R, and S (Figure S2). The four DM1 CVSs (B-33, B-34, B-36, B-40) from affected mothers showed methylation both up- and downstream of the repeat, while the three CVSs with paternally derived expansions (B-29, B-30, and B-38) showed methylation only downstream of the CTG repeat (Figures 3 and S4, Tables S3G and S3H). None of the classical DM1-affected parents showed methylation in their blood DNA. For family P, for which two CVSs (B-29 and B-30) were available, we also collected a fetal skin biopsy after termination of pregnancy, which was put in culture to obtain a fibroblast cell line (B-31). No methylation at either side of the CTG repeat was found for the fetal sample (Figure 3). Methylation levels in family P at the downstream region were lower than in families Q and R (Figure 3, Tables S3G and S3H) with CVSs B-29 and B-30 showing less than 50% methylation for most of the alleles. Remarkably, DM1 was paternally transmitted in both families P and S with low methylation, whereas it was maternally transmitted in families Q and R as well as in sample B-40, showing high methylation levels. The repeat sizes did not correlate with methylation levels. For example, sample B-34 has only 500 repeats and is methylated at both the upstream and the downstream site, while B-29 has 1,100 repeats and is methylated only at the downstream site.

Significance of Methylation

Statistical analysis (linear regression) as a function of various biological markers (gender, repeat size, disease, age at sampling) revealed that the congenital disease form is most significantly methylated at the repeat ($p = 7.04958 \times 10^{-12}$ upstream, $p = 9.55873 \times 10^{-5}$ downstream) (Tables S4A and S4B). This showed a strong correlation ($R^2 = 0.78$ upstream and $R^2 = 0.66$ downstream). In this multivariate analysis, repeat size was not significantly correlated with methylation at both CTCF binding sites (Tables S4A and S4B). There is no correlation of age at sampling with methylation upstream ($p = 0.26$) and a mild correlation downstream ($p = 0.03$) (Tables S4A and S4B). In another comparison, when methylation is assessed relative to only repeat length as a continuous variable (Tables S4C and S4D), the comparison was significant

($p = 7.11321 \times 10^{-6}$, upstream, and $p = 2.3349 \times 10^{-5}$, downstream). Similarly, when repeat length was compared to disease status as a dichotomous trait (with or without CDM1), expansion size is significantly associated with CDM1 (Table S4E, $p = 1.17093 \times 10^{-6}$). However, in either of these comparisons the correlations were not strong ($r^2 = 0.24$, upstream, and $r^2 = 0.22$, downstream, and $r^2 = 0.27$, respectively). There is a strong correlation between maternal inheritance and methylation ($p = 2.89731 \times 10^{-9}$ upstream, $p = 3.87005 \times 10^{-7}$ downstream) (Tables S4F and S4G). Taken together, our results suggest that while repeat length is important, CpG methylation is a stronger correlation/indicator of CDM1 than repeat length. Importantly, looking only at maternal inheritance, there is a significant correlation between CDM1 and methylation compared to maternal transmission of non-congenital DM1, providing strength for a maternal effect with 18 CDM1 ($p = 1.14853 \times 10^{-12}$ upstream, $p = 6.63241 \times 10^{-6}$ downstream) (Tables S4H and S4I). Comparing the different forms of non-congenital DM1, we found a stronger correlation between childhood- and juvenile-onset forms and CpG methylation versus adult- and late-onset forms ($p = 0.01$ upstream, $p = 0.00034$ downstream) (Tables S4J and S4K). This supports a continuum of methylation with age of onset.

Discussion

This study reports the largest dataset to date on DNA methylation up- and downstream of the CTG repeat in *DMPK*. Both the large number of clinically diagnosed classical and congenital DM1-affected individuals, as well as the large numbers of alleles assessed for each sample by MPS, gives strength to our findings and the implications they may have. In particular, aberrant methylation at the DM locus may account for both the maternal bias for CDM1 transmission, larger maternal versus paternal expansions, and the distinct clinical features of CDM1, as well as serve as a diagnostic indicator.

Repeat expansion size has long been suggested to be linked to CDM1.[10,11] A grandpaternal predominance for CDM1 transmissions was observed.[20,25,41,68–70] This predominance is probably explained by the paternal expansion bias of pre- and proto-mutation lengths (<80 repeats) compared to the relative stability of these lengths during maternal transmissions. In contrast, full mutation lengths (>79 repeats) show an expansion bias in maternal transmissions and a tendency to contract in paternal transmissions. The mechanism of these parent-of-origin mutation effects is poorly understood.[71] The maternal bias for transmitting very large expansions has been argued as a source of the maternal predominance of CDM1. Ongoing somatic expansions of the repeat in many tissues can give larger CTG expansions than the actual inherited allele. Since most CDM1-affected individuals are sampled shortly after birth when levels of somatic mosaicism in blood DNA are low, the lengths are probably close to the inherited length (however, this is not true of our sample set, see below). In contrast, non-congenital DM1-affected individuals frequently have their repeat sized later in life when somatic expansions can be considerable. The confounding variable of somatic expansions can hamper correlations of repeat length to disease state.[31] Correction for somatic instability by estimating the inherited progenitor allele can improve genotype-phenotype correlations.[31] For example, a poor association of parental repeat length with transmission of CDM1 has been found. Redman et al.[72] found that mothers with as few as 75 or as many as 2,500 repeats in their blood could transmit CDM1 and estimated that in DM1-affected mothers with >100 repeats, 62% of affected offspring would be affected by CDM1. Similarly, Rudnik-Schöneborn et al.[73] found that DM1-affected mothers with all ranges of CTG expansions in blood (<500 to >1,500) could have CDM1-affected offspring, suggesting an absence of an absolute correlation of maternal repeat length with CDM1 transmission. Recent data estimating for progenitor alleles suggest that age-dependent expansions occur in the germline of both DM1-affected parents, yielding an estimate that mothers with >164 repeats have a 64% risk of transmitting CDM1.[31] However, in that study age at onset in the child depended on the length of the allele transmitted, but not on the sex of the parent.[31] Moreover, since the age effect of expansion occurs in both male and female germlines, its contribution to the maternal bias for CDM1 transmissions is probably complex. Either way, while expansion size is important, it cannot explain the 25% of CDM1-affected individuals with relatively short expansions[10,20–26] or the absence of CDM1 in the presence of very large expansions.[18,19]

It is not possible to exclude a contribution of CTG length to CDM1 and our data do not argue against it's possible contribution. As noted above and in the Material and Methods, age-dependent somatic CTG expansions can weaken the presence of or the absence of phenotype correlations to repeat length.[31] We have not accounted for this confounding variable. However, our analyses are predominantly from blood samples, where the levels of somatic expansions are limited,[62,63] compared to affected tissues such as heart, muscle, and brain.[35] It is not likely that the limited somatic expansions in the blood, detectable in some but not all DM1-affected individuals, would exceed the overlapping CTG length differences observed between non-congenital DM1- and CDM1-affected individuals (our study and others). Notably, we observe large differences in repeat sizes between CDM1-affected individuals, ranging from 1,100 to 4,700 repeats, with considerable overlap in CTG repeat size ranges with non-congenital DM1-affected individuals, with tract sizes of 54 to 2,800 repeats. Furthermore, our CDM1-affected individuals were sampled at many ages, many as old as 33–50 years old, not predominantly newborns. A similar age range of sampling was evident in the non-congenital samples. Thus, although possible, it is unlikely that the somatic expansions in the

blood will significantly confound the strength of our findings. Considering the age stability of the methylation we observe, the absence or low correlation of age-at-sampling with up- or downstream methylation, the broad range of age-at-sampling for each disease class, the range of developmental stages, an absence of a correlation of repeat size with methylation status, our sample size, and our current statistical power, we do not feel that the strength of our correlations depend upon accounting for the age at sampling. We propose that the strong correlation of upstream methylation we observe in blood DNA be considered as a critical marker of CDM1, in addition to repeat length of the mother,[31] the age of the mother,[31] and the clinical state of the mother (see more below).[27,38]

What factors other than repeat length might contribute to the maternal specificity of CDM1? A maternal circulating (intrauterine) or environmental factor had been proposed.[74] The clinical status of the mother during pregnancy was also suggested,[27,38] as were the number of previous affected births,[38] a maternal mitochondrial genome contribution,[75] an effect of the non-expanded maternal allele effect,[20] or an epigenetic mark on the mutant maternal allele.[76] Direct analyses argue against a mitochondrial factor.[75] While the contribution of either an intrauterine/environmental factor or the clinical status of the mothers at the time of pregnancy and delivery to having congenitally affected offspring[38] was not directly tested and could not be ruled out, the absence of evidence supporting these was recently used as an argument against the involvement of intrauterine factors.[31] The clinical state of the mother may contribute to intrauterine state, as previously suggested.[27,38] While the number of families with comprehensive data is limited in our collection, it is interesting to note that, where known, all mothers of CDM1-affected children showed adult or earlier onset (Table S1) and were symptomatic prior to and during pregnancy. This is consistent with the observation that of 63 CDM1-affected children, all were born from DM1-affected mothers that showed obvious symptoms with adult or earlier onsets, and none were born from mothers with late-onset or "at risk" of disease.[38] Our results further support the reports that affected mothers are more likely to have CDM1-affected children, and their clinical state may contribute to having CDM1-affected children.[27,38,77] Obstetric complications of DM1-affected mothers during pregnancy and labor can be subdivided to those common to most pregnancies, those that are unique to fetuses with CTG expansions, and those unique to CDM1-affected fetuses.[73] Intrauterine exposure to environmental, chemical, genetic, and/or physiological stressors can induce epigenetic modifications, in tissue- and locus-specific manners.[78] Here we revisited the possibility that a maternal-specific epigenetic mark may contribute to CDM1. Based upon the presence of aberrant methylation in hESCs and CVSs, we suggest that methylation reprogramming during germline formation may specifically affect the methylation pattern and levels at the DM1 locus.[79–84] Methylation may both drive maternal transmission of CDM1 and protect against paternal transmissions of CDM1. Our evidence argues that an epigenetic mark that is frequently transmitted by mothers, and rarely transmitted by fathers, contributes to CDM1.

Our results reveal a strong correlation of DNA methylation patterns flanking the DM1 CTG repeat with maternal transmission of the expanded repeat. Our findings, as with the vast majority of cases, find CDM1 transmission exclusively through maternal transmission. CpG methylation was found upstream of the CTG repeat in all but 1 congenital individual (95%), in only 2 of the 21 classical DM1-affected individuals of maternal origin (9%), in 2/2 maternally derived hESC lines, and in 4/4 maternally transmitted DM1 CVSs. In striking contrast, we never found methylation upstream of the DM1 repeat after paternal transmission in either blood, CVSs, or hESCs. Moreover, the levels of upstream methylation per allele was greater in CDM1 samples relative to maternally transmitted classical DM1 samples ($p = 1.14853 \times 10^{-12}$, Table S4H), and certainly greater than the absence of upstream methylation in paternally derived classical DM1-affected individuals. It would be interesting to know whether the few paternally derived CDM1-affected case subjects may have also inherited aberrant methylation of the upstream region.[12–17] Interestingly, upstream methylation was evident in only two non-congenital individuals, and both were maternally transmitted cases of childhood DM1, with onsets prior to 4 years of age (B2 and B10, with 700 and 1,000 repeats) (Tables 1 and S1). We found a stronger correlation between childhood- and juvenile-onset forms and CpG methylation versus adult- and late-onset forms ($p = 0.01$ upstream, $p = 0.00034$ downstream) (Tables S4J and S4K). This might suggest that a continuum of age of onset and disease severity correlates with the levels of upstream CpG methylation (Figure 4).

One other study investigated the methylation pattern in a large collection of DM1 hESC lines in a broad region upstream (but not downstream) of the DM1 CTG repeat.[37] Yanovsky-Dagan et al. did not comment upon a methylation difference between maternally and paternally derived DM1 hESCs. They found high levels of upstream methylation with only the largest repeats (>300 repeats).[37] Although we did not observe upstream methylation for VUB19_DM1 with a paternally derived 500 repeats, this difference may be due to the different region analyzed between the two studies; Yanovsky-Dagan et al. assessed a region further upstream that does not overlap with the region we analyzed, which is closer to the CTG repeat (Figure 1). Taking all these elements together, we find indications to suggest that in hESCs, there is a differential methylation status according to parent-of-origin of the expanded CTG tract.

What is the source of the maternal bias for CDM1 transmission? The maternal-specific transmission of CDM1 and the extreme rarity of paternal transmission may be due to methylation patterns specific to the germ cells. We propose

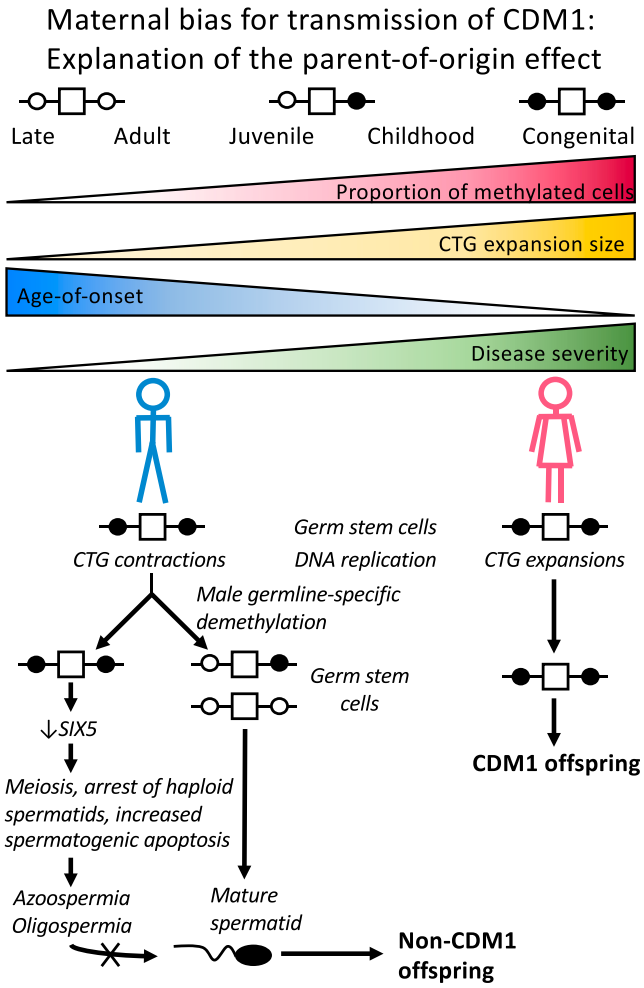

Figure 4. CpG Methylation, Parent-of-Origin Effect for Maternal Biased Transmission of CDM1

The region analyzed is schematized by a rectangle for the CTG repeat, a hollow dot for unmethylated sites, and a filled dot for methylated sites. The findings reported herein reveal that CpG methylation upstream of the expanded *DMPK* CTG repeat can account for the maternal-biased transmission of CDM1. While large CTG expansions are often associated with CDM1, this link is not absolute. Specific replication patterns affected by CpG methylation may allow for CTG contractions or expansions in male and female germline stem cells, thereby protecting affected fathers from transmitting CDM1 to their offspring. Furthermore, methylation levels and location may be selected against in the haploid spermatozoa by the reduced levels of SIX5 expression caused by methylation in its promoter (upstream of the CTG expansion). For details see Discussion.

three possible paths that might account for the maternal transmission bias of CDM1, including aberrant methylation of *DMPK* in the female germline from which a male-germline-specific CpG methylation erasure occurs, differential instability between male and female germlines, and selective growth disadvantage for the male germline with upstream methylation (Figure 4).

It is noteworthy that the methylation-free zone present in 16 different tissues of unaffected individuals in which the DM1 CTG repeat lies is further extended only in the sperm, to encompass the full 3.5 kb CpG island and a downstream CpG island (CAG repeat: chr19: 46,273,463–46,273,524; DMR tissue: chr19: 46,271,277–46,274,292 = 2,186 bp downstream, 768 bp upstream [relative to coding direction]; DMR sperm chr19: 42,670,117–46,275,169 = 3,346 kb downstream, 1,675 upstream [relative to coding direction]).[66] We propose that, in some instances, in oocytes of DM1-affected mothers, an aberrant CpG methylation pattern of the DM1 CTCF1 site is set up and transmitted to CDM1-affected offspring. Waves of methylation/demethylation specific to oogonia and spermatogonia occur with strict regulation[81–84] (Figure 4). This regulation, and the pattern of methylation, can be perturbed by proximal repeat mutations, as occurs in fragile X.[85–88] Although we did not observe a significant correlation of DM1 CTG tract length with methylation, a threshold length may be associated with aberrant methylation, but may not be evident in our sample set. A threshold CGG length linked to aberrant methylation was recently detected in FRAXA cells.[89] Sperm of DM1-affected fathers may escape aberrant CpG methylation due to a differential de-methylation program, which forces an absence of methylation further up- and downstream from the CTG tract.[66]

Selective differential proliferation of oogonia or spermatogonia with a particular methylation pattern may ultimately be the basis of the preponderance of the maternal bias for CDM1, as occurs in fragile X.[85–88] The process of this selection may arise from the dependence of spermatogenic survival upon SIX5 protein levels,[90] which can reflect the prominent progressive testicular atrophy, oligo-, and azoospermia in DM1-affected males.[91] Haploid spermatids with upstream methylation may not survive to mature spermatozoa due to the reduced SIX5 levels caused by *DMPK* methylation,[37] and hence protect DM1-affected fathers from having CDM1-affected children (Figure 4). While the methylation pattern in oocytes or a larger sample set of sperm of DM1-affected parents is unknown, the presence of the aberrant methylation pattern in hESCs and CVSs derived from germ cells of affected mothers but not affected fathers, observed herein, supports our suggestion of a maternal oocyte-specific aberrant epigenetic mark for CDM1.

Although large CTG expansions alone may not be sufficient to cause CDM1, such large expansions may be associated with the aberrant methylation patterns. Within each family studied herein, an increase in overall methylation and CTG repeat size could be found from one generation to the next (Figure S2). Where known, only 6 of the 9 three-generation families studied herein (families A, H, I, J, K and L) showed grandpaternal transmission leading to CDM1-affected grandchildren (Figure S2), previously noted for their larger expansions.[20,25,41,68–70] Upstream CpG methylation is associated with increased CTG expansions, which in turn may be linked with DNA replication patterns[92] (Figure 1, top). We suggest that for DM1-affected parents with larger CTG tracts, the transmission to larger expansions by DM1-affected mothers, relative to DM1-affected fathers, may be due to methylation-enhanced expansions and contractions, respectively (Figure 4). This

would be consistent with the suggestion of Brunner et al.[70] who proposed that repeat contractions from fathers protects their offspring from CDM1. Our suggestion is also consistent with the direct analysis of CTG lengths in sperm that suggested an increase of germline contraction frequencies for fathers with larger alleles.[63,93–96] We previously demonstrated that DM1 mouse tissues with lowest adjacent methylation showed the most asymmetric DM1 replication fork profiles, that may alter repeat instability.[92] In spermatocytes, an absence of methylation would permit binding of the CTCF paralog CTCFL, which may stabilize or induce contractions of large CTG tracts, just as CTCF binding appears to promote repeat stability in somatic tissues.[92] Thus, for DM1-affected parents with larger CTG tracts, methylation may enhance the transmission to larger maternal expansions and shorter (possibly contracted) CTG sizes by fathers and thereby account, in part for the maternal bias of CDM1 transmission (Figure 4). A similar expansion and contraction bias occurs in the female and male gametes in fragile X.[86,87] Thus, CpG methylation adjacent to the expanded DM1 allele in the oocytes may enhance large expansions.

Age-dependent and tissue-specific increases or decreases of methylation can arise, so it is important to know the stability of the methylation state we observe. An extensive analysis of numerous tissues from multiple DM1 fetal terminations and DM1-affected adult post-mortems found tissue-specific levels of CpG methylation.[35] Methylation levels in some tissues decreased between fetuses and adults, suggesting a loss in methylation with age.[35] Arguing against a loss or gain of CpG methylation with age, we have assessed upstream methylation and consistently detected its presence in maternally derived hESCs and CVSs and in 19 CDM1-affected individuals, many sampled as newborns and as late as at 33–50 years old (Table S1). Thus, it is unlikely that age is reducing the levels of *DMPK* methylation in blood. Also, we have observed an absence of upstream methylation in paternally derived hESCs, CVSs, and most non-CDM1 individuals (regardless of inheritance) with ages of sampling spanning 16–63 years (Table S1), which argues against an age dependency on the induction of upstream methylation. Thus, we conclude that the upstream *DMPK* methylation state is stable in blood. Epigenetic status between tissues, particularly those with varying clinical affectation, may relate to disease pathology.

The various methylation patterns may determine disease pathogenesis by altered gene expression. Model system data suggest that methylation-sensitive binding of CTCF at the DM1 locus may alter insulator activity or sense/anti-sense regulation (for extensive discussion see Filippova et al.,[32] Cho et al.,[33] and López Castel et al.[35]). However, the "CTCF1 insulator" model for myotonic dystrophy has recently been questioned.[37] Alternatively, it has been proposed that methylation of the DM1 locus can lead to altered chromatin packaging which may alter adjacent gene expression levels.[37] Such mechanisms may vary dramatically between tissues showing different methylation patterns.[35] Numerous studies have shown that CTG expansion in *DMPK* in DM1 individual cells downregulates expression of the adjacent *SIX5*/*DMAHP*.[37,97–100] *SIX5* expression was shown to be further decreased with CpG methylation adjacent to the *DMPK* CTG expansion.[37] Contradictory claims of up- and downregulation of *DMPK* expression have been reported in individual cells or tissues.[99,101–105] These varying results may be due to the extraction-resistant nuclear foci of retained toxic-CUG *DMPK* RNA and/or varying tissues or clinical states.[106,107] Recent analysis of two DM1 fetal terminations (12–33 weeks gestation) with CTG expansions (1,100 and 2,500 repeats), presumed to be CDM1 but without clinical assignment (it is not possible to diagnose CDM1 prenatally), revealed both sense and anti-sense *DMPK* transcripts in heart and brain.[108] Abundant RNA foci for both sense and antisense transcripts were observed in heart, skeletal muscle, and brain as early as 12 weeks of gestation.[108] Interestingly, expression of both strands was greater in heart than brain. It is possible that tissue-specific CpG methylation contributed to expression differences, but methylation status of the samples was not characterized. Our finding that differential methylation patterns exist in some CDM1 tissues[35] opens new questions of how this may lead to the clinical state. For example, many DNA-binding transcription factors that are known to be sensitive to CpG methylation state can bind proximal to the CTG tract (Table S5). Methylation-mediated misregulation of any one of these factors could contribute to the CDM1 phenotype.

The tight association of upstream methylation levels with CDM1 may provide some guidance to the difficulties hampering a definitive molecular prenatal determination of CDM1. Prenatal diagnosis of CDM1 is hampered by the poor association of repeat size to clinical presentation.[4,5,23–25,27–29,43,109] Remarkably, for each of the three CVSs with paternally derived CTG expansions, methylation was present only downstream of the repeat, while for all four maternally transmitted CVSs, methylation was present both upstream and downstream of the repeat (Tables 1 and S1, Figure 3). This strong methylation pattern was also observed in the four hESC lines studied, with methylation upstream in all the maternally inherited lines, and no difference in methylation, based on inheritance, downstream of the CTG repeat (Tables 1 and S1, Figure 3). While none of these CVSs or hESCs studied herein led to a live birth, disallowing a clinical diagnosis of either DM1 or CDM1, it is interesting to speculate that the presence of methylation upstream of the CTG tract might serve as a molecular marker of CDM1.

In conclusion, our study shows a positive correlation between methylation, particularly upstream of the CTG repeat, and maternal inheritance in DM1-affected individuals. This correlation of upstream methylation and maternal inheritance of CDM1 is nearly absolute, while only 2 out of 21 classical DM1-affected individuals with maternal inheritance showed methylation in the same upstream region. This could indicate that methylation is an

early step for development of CDM1 onset but that additional elements are necessary to develop the full-blown clinical picture. Our results argue against the CTG repeat size as the only and essential element, since in the literature and in our cohort there was a significant overlap in CTG repeat size between classical DM1- and CDM1-affected individuals. These results are corroborated in our DM1 hESC and CVSs of maternal origin, suggesting that methylation of an allele not methylated in the blood of the previous generation may occur quite early, possibly in the oocyte or embryo. Analysis of more hESC lines would reinforce our conclusions. Although the functional and tissue-specific consequences of methylation in DM1 are still unclear, this study firmly establishes that methylation upstream of the expanded *DMPK* CTG repeat occurs exclusively with maternal transmission and that it is somehow linked to the development of CDM1 ($p = 2.89731 \times 10^{-9}$) (Tables S4F and S4G). Analysis of this large and comprehensive collection highlights CpG methylation as a potential prenatal indicator of CDM1 that is stronger than the repeat size itself. This may guide families faced with an affected pregnancy to reach a reproductive decision based on a more accurate prediction of the risk for CDM1 in their child.

Supplemental Data

Supplemental Data include four figures and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2017.01.033.

Acknowledgments

This work was partially supported by the Canadian Institutes of Health Research (FRN94966 and FRN291634) (C.E.P.), the Muscular Dystrophy Canada (C.E.P.), Tribute Communities (C.E.P.), the Kazman Family Fund (C.E.P.), The Marigold Foundation (C.E.P.), the Petroff Family (C.E.P.), the Fonds voor Wetenschappelijk Onderzoek-Vlaanderen (K.S.), the Leerstoel Mireille Aerens (K.S.), the Methusalem grant (K.S.), and Association Belge contre les Maladies neuro-Musculaires fund (Projet 2013/10) (K.S.). Saguenay-Lac-St-Jean DM1 Biobank was supported by The Marigold Foundation (C.E.P. and C.G.). M.D.W. was supported by a Canada Research Chair. M.L. was supported by Hospital for Sick Children Foundation Student Scholarship Program and NSERC grant 436194-2013. S.L. was supported by Tribute Communities, the Kazman Family Fund, and The Marigold Foundation. L.B. was funded by a PhD grant from Leerstoel Mireille Aerens. L.B.'s exchange visit to the Pearson Lab was covered in part by Association Belge contre les Maladies neuro-Musculaires fund ("Prijs Denys Bourgeois"). We acknowledge the technical support by Dr. Lisa Yu, Mila Mirceta, Neha Thakkar, Kimberly Dee, Pierre Hilven, and the team of the BRIGHTCore for the Miseq Analyses. We acknowledge Pieter-Jan Van de Maele for his informatics expertise in writing the script used for analysis of the MPBS data. We acknowledge Olivia Shea for support in producing figures.

Received: August 30, 2016
Accepted: January 26, 2017
Published: March 2, 2017

Web Resources

BiQ Analyzer, http://biq-analyzer.bioinf.mpi-inf.mpg.de
hPSC Registry, https://hpscreg.eu/
OMIM, http://www.omim.org/

References

1. Schmidt, M.H., and Pearson, C.E. (2016). Disease-associated repeat instability and mismatch repair. DNA Repair (Amst.) *38*, 117–126.
2. Ashizawa, T., Dunne, C.J., Dubel, J.R., Perryman, M.B., Epstein, H.F., Boerwinkle, E., and Hejtmancik, J.F. (1992). Anticipation in myotonic dystrophy. I. Statistical verification based on clinical and haplotype findings. Neurology *42*, 1871–1877.
3. Ashizawa, T., Dubel, J.R., Dunne, P.W., Dunne, C.J., Fu, Y.H., Pizzuti, A., Caskey, C.T., Boerwinkle, E., Perryman, M.B., Epstein, H.F., et al. (1992). Anticipation in myotonic dystrophy. II. Complex relationships between clinical findings and structure of the GCT repeat. Neurology *42*, 1877–1883.
4. Echenne, B., and Bassez, G. (2013). Congenital and infantile myotonic dystrophy. Handb. Clin. Neurol. *113*, 1387–1393.
5. Ekström, A.B., Hakenäs-Plate, L., Samuelsson, L., Tulinius, M., and Wentz, E. (2008). Autism spectrum conditions in myotonic dystrophy type 1: a study on 57 individuals with congenital and childhood forms. Am. J. Med. Genet. B. Neuropsychiatr. Genet. *147B*, 918–926.
6. Turner, C., and Hilton-Jones, D. (2010). The myotonic dystrophies: diagnosis and management. J. Neurol. Neurosurg. Psychiatry *81*, 358–367.
7. Aslanidis, C., Jansen, G., Amemiya, C., Shutler, G., Mahadevan, M., Tsilfidis, C., Chen, C., Alleman, J., Wormskamp, N.G., Vooijs, M., et al. (1992). Cloning of the essential myotonic dystrophy region and mapping of the putative defect. Nature *355*, 548–551.
8. Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J.P., Hudson, T., 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 *69*, 385.
9. Salehi, L.B., Bonifazi, E., Stasio, E.D., Gennarelli, M., Botta, A., Vallo, L., Iraci, R., Massa, R., Antonini, G., Angelini, C., and Novelli, G. (2007). Risk prediction for clinical phenotype in myotonic dystrophy type 1: data from 2,650 patients. Genet. Test. *11*, 84–90.
10. Tsilfidis, C., MacKenzie, A.E., Mettler, G., Barceló, J., and Korneluk, R.G. (1992). Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy. Nat. Genet. *1*, 192–195.
11. Myring, J., Meredith, A.L., Harley, H.G., Kohn, G., Norbury, G., Harper, P.S., and Shaw, D.J. (1992). Specific molecular prenatal diagnosis for the CTG mutation in myotonic dystrophy. J. Med. Genet. *29*, 785–788.
12. Di Costanzo, A., de Cristofaro, M., Di Iorio, G., Daniele, A., Bonavita, S., and Tedeschi, G. (2009). Paternally inherited case of congenital DM1: brain MRI and review of literature. Brain Dev. *31*, 79–82.
13. Zeesman, S., Carson, N., and Whelan, D.T. (2002). Paternal transmission of the congenital form of myotonic dystrophy

type 1: a new case and review of the literature. Am. J. Med. Genet. *107*, 222–226.
14. de Die-Smulders, C.E., Smeets, H.J., Loots, W., Anten, H.B., Mirandolle, J.F., Geraedts, J.P., and Höweler, C.J. (1997). Paternal transmission of congenital myotonic dystrophy. J. Med. Genet. *34*, 930–933.
15. Tanaka, Y., Suzuki, Y., Shimozawa, N., Nanba, E., and Kondo, N. (2000). Congenital myotonic dystrophy: report of paternal transmission. Brain Dev. *22*, 132–134.
16. Ohya, K., Tachi, N., Chiba, S., Sato, T., Kon, S., Kikuchi, K., Imamura, S., Yamagata, H., and Miki, T. (1994). Congenital myotonic dystrophy transmitted from an asymptomatic father with a DM-specific gene. Neurology *44*, 1958–1960.
17. Bergoffen, J., Kant, J., Sladky, J., McDonald-McGinn, D., Zackai, E.H., and Fischbeck, K.H. (1994). Paternal transmission of congenital myotonic dystrophy. J. Med. Genet. *31*, 518–520.
18. Cobo, A.M., Baiget, M., López de Munain, A., Poza, J.J., Emparanza, J.I., and Johnson, K. (1993). Sex-related difference in intergenerational expansion of myotonic dystrophy gene. Lancet *341*, 1159–1160.
19. Clark, C., Petty, R.K., and Strong, A.M. (1998). Late presentation of myotonic dystrophy. Clin. Exp. Dermatol. *23*, 47–48.
20. Barceló, J.M., Pluscauskas, M., MacKenzie, A.E., Tsilfidis, C., Narang, M., and Korneluk, R.G. (1994). Additive influence of maternal and offspring DM-kinase gene CTG repeat lengths in the genesis of congenital myotonic dystrophy. Am. J. Hum. Genet. *54*, 1124–1125.
21. Hilbert, J.E., Johnson, N.E., and Moxley, R.T., 3rd. (2013). New insights about the incidence, multisystem manifestations, and care of patients with congenital myotonic dystrophy. J. Pediatr. *163*, 12–14.
22. Novelli, G., Gennarelli, M., Menegazzo, E., Angelini, C., and Dallapiccola, B. (1995). Discordant clinical outcome in myotonic dystrophy relatives showing (CTG)n > 700 repeats. Neuromuscul. Disord. *5*, 157–159.
23. Spranger, M., Janssen, B., Rating, D., and Spranger, S. (1999). [Disease picture of myotonic muscular dystrophy in patients with large CTG triplet expansion]. Nervenarzt *70*, 131–135.
24. Geifman-Holtzman, O., and Fay, K. (1998). Prenatal diagnosis of congenital myotonic dystrophy and counseling of the pregnant mother: case report and literature review. Am. J. Med. Genet. *78*, 250–253.
25. Lavedan, C., Hofmann-Radvanyi, H., Shelbourne, P., Rabes, J.P., Duros, C., Savoy, D., Dehaupas, I., Luce, S., Johnson, K., and Junien, C. (1993). Myotonic dystrophy: size- and sex-dependent dynamics of CTG meiotic instability, and somatic mosaicism. Am. J. Hum. Genet. *52*, 875–883.
26. Campbell, C., Levin, S., Siu, V.M., Venance, S., and Jacob, P. (2013). Congenital myotonic dystrophy: Canadian population-based surveillance study. J. Pediatr. *163*, 120–5.e1, 3.
27. Verrijn Stuart, A.A., Huisman, M., van Straaten, H.L., Bakker, J.C., and Arabin, B. (2000). "Shake hands"; diagnosing a floppy infant–myotonic dystrophy and the congenital subtype: a difficult perinatal diagnosis. J. Perinat. Med. *28*, 497–501.
28. Dalphin, M.L., Noir, A., Monnier, G., and Menget, A. (1992). [Congenital myotonic dystrophy. Diagnostic difficulties]. Pediatrie *47*, 677–680.
29. DiRocco, M., Gennarelli, M., Veneselli, E., Bado, M., Romanengo, M., Celle, M.E., Cordone, G., and Borrone, C. (1996). Diagnostic problems in congenital myotonic dystrophy. Eur. J. Pediatr. *155*, 995.
30. Hojo, K., Yamagata, H., Moji, H., Fujita, T., Miki, T., Fujimura, M., and Kidoguchi, K. (1995). Congenital myotonic dystrophy: molecular diagnosis and clinical study. Am. J. Perinatol. *12*, 195–200.
31. Morales, F., Vásquez, M., Cuenca, P., Campos, D., Santamaría, C., Del Valle, G., Brian, R., Sittenfeld, M., and Monckton, D.G. (2015). Parental age effects, but no evidence for an intrauterine effect in the transmission of myotonic dystrophy type 1. Eur. J. Hum. Genet. *23*, 646–653.
32. Filippova, G.N., Thienes, C.P., Penn, B.H., Cho, D.H., Hu, Y.J., Moore, J.M., Klesert, T.R., Lobanenkov, V.V., and Tapscott, S.J. (2001). CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. Nat. Genet. *28*, 335–343.
33. Cho, D.H., Thienes, C.P., Mahoney, S.E., Analau, E., Filippova, G.N., and Tapscott, S.J. (2005). Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. Mol. Cell *20*, 483–489.
34. Steinbach, P., Gläser, D., Vogel, W., Wolf, M., and Schwemmle, S. (1998). The DMPK gene of severely affected myotonic dystrophy patients is hypermethylated proximal to the largely expanded CTG repeat. Am. J. Hum. Genet. *62*, 278–285.
35. López Castel, A., Nakamori, M., Tomé, S., Chitayat, D., Gourdon, G., Thornton, C.A., and Pearson, C.E. (2011). Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues. Hum. Mol. Genet. *20*, 1–15.
36. Brouwer, J.R., Huguet, A., Nicole, A., Munnich, A., and Gourdon, G. (2013). Transcriptionally repressive chromatin remodelling and CpG methylation in the presence of expanded CTG-repeats at the DM1 locus. J. Nucleic Acids *2013*, 567435.
37. Yanovsky-Dagan, S., Avitzour, M., Altarescu, G., Renbaum, P., Eldar-Geva, T., Schonberger, O., Mitrani-Rosenbaum, S., Levy-Lahad, E., Birnbaum, R.Y., Gepstein, L., et al. (2015). Uncovering the role of hypermethylation by CTG expansion in myotonic dystrophy type 1 using mutant human embryonic stem cells. Stem Cell Reports *5*, 221–231.
38. Koch, M.C., Grimm, T., Harley, H.G., and Harper, P.S. (1991). Genetic risks for children of women with myotonic dystrophy. Am. J. Hum. Genet. *48*, 1084–1091.
39. Afifi, A.M., Bhatia, A.R., and Eyal, F. (1992). Hydrops fetalis associated with congenital myotonic dystrophy. Am. J. Obstet. Gynecol. *166*, 929–930.
40. Hsu, C.D., Feng, T.I., Crawford, T.O., and Johnson, T.R. (1993). Unusual fetal movement in congenital myotonic dystrophy. Fetal Diagn. Ther. *8*, 200–202.
41. Harley, H.G., Rundle, S.A., MacMillan, J.C., Myring, J., Brook, J.D., Crow, S., Reardon, W., Fenton, I., Shaw, D.J., and Harper, P.S. (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.
42. Steyaert, J., Umans, S., Willekens, D., Legius, E., Pijkels, E., de Die-Smulders, C., Van den Berghe, H., and Fryns, J.P. (1997). A study of the cognitive and psychological profile in 16 children with congenital or juvenile myotonic dystrophy. Clin. Genet. *52*, 135–141.
43. Zaki, M., Boyd, P.A., Impey, L., Roberts, A., and Chamberlain, P. (2007). Congenital myotonic dystrophy: prenatal

ultrasound findings and pregnancy outcome. Ultrasound Obstet. Gynecol. *29*, 284–288.
44. De Antonio, M., Dogan, C., Hamroun, D., Mati, M., Zerrouki, S., Eymard, B., Katsahian, S., Bassez, G.; and French Myotonic Dystrophy Clinical Network (2016). Unravelling the myotonic dystrophy type 1 clinical spectrum: A systematic registry-based study with implications for disease classification. Rev. Neurol. (Paris) *172*, 572–580.
45. Johnson, N.E., Ekstrom, A.B., Campbell, C., Hung, M., Adams, H.R., Chen, W., Luebbe, E., Hilbert, J., Moxley, R.T., 3rd, and Heatwole, C.R. (2016). Parent-reported multi-national study of the impact of congenital and childhood onset myotonic dystrophy. Dev. Med. Child Neurol. *58*, 698–705.
46. Winblad, S., Samuelsson, L., Lindberg, C., and Meola, G. (2016). Cognition in myotonic dystrophy type 1: a 5-year follow-up study. Eur. J. Neurol. *23*, 1471–1476.
47. Jeffreys, A.J., Tamaki, K., MacLeod, A., Monckton, D.G., Neil, D.L., and Armour, J.A. (1994). Complex gene conversion events in germline mutation at human minisatellites. Nat. Genet. *6*, 136–145.
48. Jacobs, K., Mertzanidou, A., Geens, M., Nguyen, H.T., Staessen, C., and Spits, C. (2014). Low-grade chromosomal mosaicism in human somatic and embryonic stem cell populations. Nat. Commun. *5*, 4227.
49. Mateizel, I., De Becker, A., Van de Velde, H., De Rycke, M., Van Steirteghem, A., Cornelissen, R., Van der Elst, J., Liebaers, I., Van Riet, I., and Sermon, K. (2008). Efficient differentiation of human embryonic stem cells into a homogeneous population of osteoprogenitor-like cells. Reprod. Biomed. Online *16*, 741–753.
50. De Temmerman, N., Seneca, S., Van Steirteghem, A., Haentjens, P., Van der Elst, J., Liebaers, I., and Sermon, K.D. (2008). CTG repeat instability in a human embryonic stem cell line carrying the myotonic dystrophy type 1 mutation. Mol. Hum. Reprod. *14*, 405–412.
51. Seriola, A., Spits, C., Simard, J.P., Hilven, P., Haentjens, P., Pearson, C.E., and Sermon, K. (2011). Huntington's and myotonic dystrophy hESCs: down-regulated trinucleotide repeat instability and mismatch repair machinery expression upon differentiation. Hum. Mol. Genet. *20*, 176–185.
52. Sambrook, J., and Russell, D.W. (2006). Purification of PCR products in preparation for cloning. CSH protocols 2006.
53. Pratte, A., Prévost, C., Puymirat, J., and Mathieu, J. (2015). Anticipation in myotonic dystrophy type 1 parents with small CTG expansions. Am. J. Med. Genet. A. *167A*, 708–714.
54. Musova, Z., Mazanec, R., Krepelova, A., Ehler, E., Vales, J., Jaklova, R., Prochazka, T., Koukal, P., Marikova, T., Kraus, J., et al. (2009). Highly unstable sequence interruptions of the CTG repeat in the myotonic dystrophy gene. Am. J. Med. Genet. A. *149A*, 1365–1374.
55. Wong, L.J., and Ashizawa, T. (1997). Instability of the (CTG)n repeat in congenital myotonic dystrophy. Am. J. Hum. Genet. *61*, 1445–1448.
56. Tachi, N., Ohya, K., Chiba, S., Sato, T., and Kikuchi, K. (1995). Minimal somatic instability of CTG repeat in congenital myotonic dystrophy. Pediatr. Neurol. *12*, 81–83.
57. Joseph, J.T., Richards, C.S., Anthony, D.C., Upton, M., Perez-Atayde, A.R., and Greenstein, P. (1997). Congenital myotonic dystrophy pathology and somatic mosaicism. Neurology *49*, 1457–1460.
58. Morales, F., Couto, J.M., Higham, C.F., Hogg, G., Cuenca, P., Braida, C., Wilson, R.H., Adam, B., del Valle, G., Brian, R., 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.
59. Thornton, C.A., Johnson, K., and Moxley, R.T., 3rd. (1994). Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. Ann. Neurol. *35*, 104–107.
60. Ohya, K., Tachi, N., Kon, S., Kikuchi, K., and Chiba, S. (1995). Somatic cell heterogeneity between DNA extracted from lymphocytes and skeletal muscle in congenital myotonic dystrophy. Jpn. J. Hum. Genet. *40*, 319–326.
61. Zatz, M., Passos-Bueno, M.R., Cerqueira, A., Marie, S.K., Vainzof, M., and Pavanello, R.C. (1995). Analysis of the CTG repeat in skeletal muscle of young and adult myotonic dystrophy patients: when does the expansion occur? Hum. Mol. Genet. *4*, 401–406.
62. Martorell, L., Monckton, D.G., Gamez, J., Johnson, K.J., Gich, I., Lopez de Munain, A., and Baiget, M. (1998). Progression of somatic CTG repeat length heterogeneity in the blood cells of myotonic dystrophy patients. Hum. Mol. Genet. *7*, 307–312.
63. Martorell, L., Gámez, J., Cayuela, M.L., Gould, F.K., McAbney, J.P., Ashizawa, T., Monckton, D.G., and Baiget, M. (2004). Germline mutational dynamics in myotonic dystrophy type 1 males: allele length and age effects. Neurology *62*, 269–274.
64. Bock, C., Reither, S., Mikeska, T., Paulsen, M., Walter, J., and Lengauer, T. (2005). BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. Bioinformatics *21*, 4067–4068.
65. Roifman, M., Choufani, S., Turinsky, A.L., Drewlo, S., Keating, S., Brudno, M., Kingdom, J., and Weksberg, R. (2016). Genome-wide placental DNA methylation analysis of severely growth-discordant monochorionic twins reveals novel epigenetic targets for intrauterine growth restriction. Clin. Epigenetics *8*, 70.
66. Buckley, L., Lacey, M., and Ehrlich, M. (2016). Epigenetics of the myotonic dystrophy-associated DMPK gene neighborhood. Epigenomics *8*, 13–31.
67. Boucher, C.A., King, S.K., Carey, N., Krahe, R., Winchester, C.L., Rahman, S., Creavin, T., Meghji, P., Bailey, M.E., Chartier, F.L., et al. (1995). A novel homeodomain-encoding gene is associated with a large CpG island interrupted by the myotonic dystrophy unstable (CTG)n repeat. Hum. Mol. Genet. *4*, 1919–1925.
68. Ashizawa, T., Anvret, M., Baiget, M., Barceló, J.M., Brunner, H., Cobo, A.M., Dallapiccola, B., Fenwick, R.G., Jr., Grandell, U., Harley, H., et al. (1994). Characteristics of intergenerational contractions of the CTG repeat in myotonic dystrophy. Am. J. Hum. Genet. *54*, 414–423.
69. López de Munain, A., Cobo, A.M., Poza, J.J., Navarrete, D., Martorell, L., Palau, F., Emparanza, J.I., and Baiget, M. (1995). Influence of the sex of the transmitting grandparent in congenital myotonic dystrophy. J. Med. Genet. *32*, 689–691.
70. Brunner, H.G., Brüggenwirth, H.T., Nillesen, W., Jansen, G., Hamel, B.C., Hoppe, R.L., de Die, C.E., Höweler, C.J., van Oost, B.A., Wieringa, B., et al. (1993). Influence of sex of the transmitting parent as well as of parental allele size on the CTG expansion in myotonic dystrophy (DM). Am. J. Hum. Genet. *53*, 1016–1023.

71. Pearson, C.E. (2003). Slipping while sleeping? Trinucleotide repeat expansions in germ cells. Trends Mol. Med. *9*, 490–495.
72. Redman, J.B., Fenwick, R.G., Jr., Fu, Y.H., Pizzuti, A., and Caskey, C.T. (1993). Relationship between parental trinucleotide GCT repeat length and severity of myotonic dystrophy in offspring. JAMA *269*, 1960–1965.
73. Rudnik-Schöneborn, S., Nicholson, G.A., Morgan, G., Röhrig, D., and Zerres, K. (1998). Different patterns of obstetric complications in myotonic dystrophy in relation to the disease status of the fetus. Am. J. Med. Genet. *80*, 314–321.
74. Harper, P.S., and Dyken, P.R. (1972). Early-onset dystrophia myotonica. Evidence supporting a maternal environmental factor. Lancet *2*, 53–55.
75. Poulton, J., Harley, H.G., Dasmahapatra, J., Brown, G.K., Potter, C.G., and Sykes, B. (1995). Mitochondrial DNA does not appear to influence the congenital onset type of myotonic dystrophy. J. Med. Genet. *32*, 732–735.
76. Shaw, D.J., Chaudhary, S., Rundle, S.A., Crow, S., Brook, J.D., Harper, P.S., and Harley, H.G. (1993). A study of DNA methylation in myotonic dystrophy. J. Med. Genet. *30*, 189–192.
77. Webb, D., Muir, I., Faulkner, J., and Johnson, G. (1978). Myotonia dystrophica: obstetric complications. Am. J. Obstet. Gynecol. *132*, 265–270.
78. Barua, S., and Junaid, M.A. (2015). Lifestyle, pregnancy and epigenetic effects. Epigenomics *7*, 85–102.
79. Reik, W., and Surani, M.A. (2015). Germline and pluripotent stem cells. Cold Spring Harb. Perspect. Biol. *7*, 7.
80. Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. Nature *447*, 425–432.
81. Lucifero, D., Mertineit, C., Clarke, H.J., Bestor, T.H., and Trasler, J.M. (2002). Methylation dynamics of imprinted genes in mouse germ cells. Genomics *79*, 530–538.
82. Ly, L., Chan, D., and Trasler, J.M. (2015). Developmental windows of susceptibility for epigenetic inheritance through the male germline. Semin. Cell Dev. Biol. *43*, 96–105.
83. Maatouk, D.M., and Resnick, J.L. (2003). Continuing primordial germ cell differentiation in the mouse embryo is a cell-intrinsic program sensitive to DNA methylation. Dev. Biol. *258*, 201–208.
84. Clarke, H.J., and Vieux, K.F. (2015). Epigenetic inheritance through the female germ-line: The known, the unknown, and the possible. Semin. Cell Dev. Biol. *43*, 106–116.
85. Sun, Y.J., and Baumer, A. (1999). Nonrandom X inactivation and selection of fragile X full mutation in fetal fibroblasts. Am. J. Med. Genet. *86*, 162–164.
86. Malter, H.E., Iber, J.C., Willemsen, R., de Graaff, E., Tarleton, J.C., Leisti, J., Warren, S.T., and Oostra, B.A. (1997). Characterization of the full fragile X syndrome mutation in fetal gametes. Nat. Genet. *15*, 165–169.
87. Salat, U., Bardoni, B., Wöhrle, D., and Steinbach, P. (2000). Increase of FMRP expression, raised levels of FMR1 mRNA, and clonal selection in proliferating cells with unmethylated fragile X repeat expansions: a clue to the sex bias in the transmission of full mutations? J. Med. Genet. *37*, 842–850.
88. Willemsen, R., Bontekoe, C.J., Severijnen, L.A., and Oostra, B.A. (2002). Timing of the absence of FMR1 expression in full mutation chorionic villi. Hum. Genet. *110*, 601–605.
89. Brykczynska, U., Pecho-Vrieseling, E., Thiemeyer, A., Klein, J., Fruh, I., Doll, T., Manneville, C., Fuchs, S., Iazeolla, M., Beibel, M., et al. (2016). CGG repeat-induced FMR1 silencing depends on the expansion size in human iPSCs and neurons carrying unmethylated full mutations. Stem Cell Reports *7*, 1059–1071.
90. Sarkar, P.S., Paul, S., Han, J., and Reddy, S. (2004). Six5 is required for spermatogenic cell survival and spermiogenesis. Hum. Mol. Genet. *13*, 1421–1431.
91. Vazquez, J.A., Pinies, J.A., Martul, P., De los Rios, A., Gatzambide, S., and Busturia, M.A. (1990). Hypothalamic-pituitary-testicular function in 70 patients with myotonic dystrophy. J. Endocrinol. Invest. *13*, 375–379.
92. Cleary, J.D., Tomé, S., López Castel, A., Panigrahi, G.B., Foiry, L., Hagerman, K.A., Sroka, H., Chitayat, D., Gourdon, G., and Pearson, C.E. (2010). Tissue- and age-specific DNA replication patterns at the CTG/CAG-expanded human myotonic dystrophy type 1 locus. Nat. Struct. Mol. Biol. *17*, 1079–1087.
93. 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.
94. Martorell, L., Monckton, D.G., Gamez, J., and Baiget, M. (2000). Complex patterns of male germline instability and somatic mosaicism in myotonic dystrophy type 1. Eur. J. Hum. Genet. *8*, 423–430.
95. Massari, A., Gennarelli, M., Menegazzo, E., Pizzuti, A., Silani, V., Mastrogiacomo, I., Pagani, E., Angelini, C., Scarlato, G., Novelli, G., et al. (1995). Postzygotic instability of the myotonic dystrophy p[AGC] in repeat supported by larger expansions in muscle and reduced amplifications in sperm. J. Neurol. *242*, 379–383.
96. Jansen, G., Willems, P., Coerwinkel, M., Nillesen, W., Smeets, H., Vits, L., Höweler, C., Brunner, H., and Wieringa, B. (1994). Gonosomal mosaicism in myotonic dystrophy patients: involvement of mitotic events in (CTG)n repeat variation and selection against extreme expansion in sperm. Am. J. Hum. Genet. *54*, 575–585.
97. Klesert, T.R., Otten, A.D., Bird, T.D., and Tapscott, S.J. (1997). Trinucleotide repeat expansion at the myotonic dystrophy locus reduces expression of DMAHP. Nat. Genet. *16*, 402–406.
98. Thornton, C.A., Wymer, J.P., Simmons, Z., McClain, C., and Moxley, R.T., 3rd. (1997). Expansion of the myotonic dystrophy CTG repeat reduces expression of the flanking DMAHP gene. Nat. Genet. *16*, 407–409.
99. Inukai, A., Doyu, M., Kato, T., Liang, Y., Kuru, S., Yamamoto, M., Kobayashi, Y., and Sobue, G. (2000). Reduced expression of DMAHP/SIX5 gene in myotonic dystrophy muscle. Muscle Nerve *23*, 1421–1426.
100. Alwazzan, M., Newman, E., Hamshere, M.G., and Brook, J.D. (1999). Myotonic dystrophy is associated with a reduced level of RNA from the DMWD allele adjacent to the expanded repeat. Hum. Mol. Genet. *8*, 1491–1497.
101. Fu, Y.H., Friedman, D.L., Richards, S., Pearlman, J.A., Gibbs, R.A., Pizzuti, A., Ashizawa, T., Perryman, M.B., Scarlato, G., Fenwick, R.G., Jr., et al. (1993). Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. Science *260*, 235–238.
102. Sabouri, L.A., Mahadevan, M.S., Narang, M., Lee, D.S., Surh, L.C., and Korneluk, R.G. (1993). Effect of the myotonic dystrophy (DM) mutation on mRNA levels of the DM gene. Nat. Genet. *4*, 233–238.

103. Depardon, F., Cisneros, B., Alonso-Vilatela, E., and Montañez, C. (2001). Myotonic dystrophy protein kinase (DMPK) gene expression in lymphocytes of patients with myotonic dystrophy. Arch. Med. Res. *32*, 123–128.
104. Hofmann-Radvanyi, H., Lavedan, C., Rabès, J.P., Savoy, D., Duros, C., Johnson, K., and Junien, C. (1993). Myotonic dystrophy: absence of CTG enlarged transcript in congenital forms, and low expression of the normal allele. Hum. Mol. Genet. *2*, 1263–1266.
105. Narang, M.A., Waring, J.D., Sabourin, L.A., and Korneluk, R.G. (2000). Myotonic dystrophy (DM) protein kinase levels in congenital and adult DM patients. Eur. J. Hum. Genet. *8*, 507–512.
106. Hamshere, M.G., Newman, E.E., Alwazzan, M., Athwal, B.S., and Brook, J.D. (1997). Transcriptional abnormality in myotonic dystrophy affects DMPK but not neighboring genes. Proc. Natl. Acad. Sci. USA *94*, 7394–7399.
107. Davis, B.M., McCurrach, M.E., Taneja, K.L., Singer, R.H., and Housman, D.E. (1997). Expansion of a CUG trinucleotide repeat in the 3′ untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. Proc. Natl. Acad. Sci. USA *94*, 7388–7393.
108. Michel, L., Huguet-Lachon, A., and Gourdon, G. (2015). Sense and antisense DMPK RNA foci accumulate in DM1 tissues during development. PLoS ONE *10*, e0137620.
109. Cobo, A.M., Poza, J.J., Martorell, L., López de Munain, A., Emparanza, J.I., and Baiget, M. (1995). Contribution of molecular analyses to the estimation of the risk of congenital myotonic dystrophy. J. Med. Genet. *32*, 105–108.
110. Liu, G., Chen, X., Gao, Y., Lewis, T., Barthelemy, J., and Leffak, M. (2012). Altered replication in human cells promotes DMPK (CTG)(n) · (CAG)(n) repeat instability. Mol. Cell. Biol. *32*, 1618–1632.
111. Langley, A.R., Gräf, S., Smith, J.C., and Krude, T. (2016). Genome-wide identification and characterisation of human DNA replication origins by initiation site sequencing (ini-seq). Nucleic Acids Res. *44*, 10230–10247.