*Human Molecular Genetics, 2011, Vol. 20, No. 1* **1–15**
*doi:10.1093/hmg/ddq427*
*Advance Access published on November 1, 2010*

# Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues

**Arturo López Castel[1], Masayuki Nakamori[3], Stephanie Tomé[1,4], David Chitayat[1,2,5,6], Geneviève Gourdon[4], Charles A. Thornton[3] and Christopher E. Pearson[1,6,*]**

[1]Genetics and Genome Biology and [2]Department of Pediatrics, Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, Ontario, Canada M5G 1L7, [3]Department of Neurology, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, USA, [4]Inserm U781, Hôpital Necker-Enfants Malades, Paris 75015, France, [5]Department of Obstetrics and Gynecology, The Prenatal Diagnosis and Medical Genetics Program, Mount Sinai Hospital, Toronto, Ontario, Canada and [6]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada

Received August 17, 2010; Revised and Accepted September 27, 2010

**Myotonic dystrophy (DM1) affects multiple organs, shows age-dependent progression and is caused by CTG expansions at the DM1 locus. We determined the DM1 CpG methylation profile and CTG length in tissues from DM1 foetuses, DM1 adults, non-affected individuals and transgenic DM1 mice. Analysis included CTCF binding sites upstream and downstream of the CTG tract, as methylation-sensitive CTCF binding affects chromatinization and transcription of the DM1 locus. In humans, in a given foetus, expansions were largest in heart and smallest in liver, differing by 40–400 repeats; in adults, the largest expansions were in heart and cerebral cortex and smallest in cerebellum, differing by up to 5770 repeats in the same individual. Abnormal methylation was specific to the mutant allele. In DM1 adults, heart, liver and cortex showed high-to-moderate methylation levels, whereas cerebellum, kidney and skeletal muscle were devoid of methylation. Methylation decreased between foetuses and adults. Contrary to previous findings, methylation was not restricted to individuals with congenital DM1. The expanded repeat demarcates an abrupt boundary of methylation. Upstream sequences, including the CTCF site, were methylated, whereas the repeat itself and downstream sequences were not. In DM1 mice, expansion-, tissue- and age-specific methylation patterns were similar but not identical to those in DM1 individuals; notably in mice, methylation was present up- and downstream of the repeat, but greater upstream. Thus, in humans, the CpG-free expanded CTG repeat appears to maintain a highly polarized pattern of CpG methylation at the DM1 locus, which varies markedly with age and tissues.**

## INTRODUCTION

Myotonic dystrophy type 1 [DM1 (MIM 160900)]is an inherited multisystemic disorder with a highly variable clinical presentation (1,2). Multiple organs are affected, with muscle weakness and wasting, myotonia, cardiac conduction defects, cataracts and neuropsychological impairment being some of the most characteristic clinical features. The inheritance pattern shows an earlier age of onset and more severe symptoms in successive generations. The spectrum of DM1 ranges from mild, late-onset symptoms to severe disease in infancy, with severity and age of onset roughly correlating inversely with the increasing size of the (CTG)n expansion.

The molecular basis of DM1 involves an expansion, always larger than 50 repeats but able to reach thousands of units, of a CTG tract within the 3′-untranslated region (UTR) of the *DMPK* gene (3). At least two mechanisms of disease pathogenesis are thought to contribute to DM1. First, the expression

*To whom correspondence should be addressed at: Genetics and Genome Biology, The Hospital for Sick Children, TMDT Building 101 College St, 15th Floor, Rm 15-312 East Tower, Toronto, Ontario, Canada M5G 1L7. Tel: +1 4168138256; Fax: +1 4168134931; Email: cepearson.sickkids@gmail.com

of RNA with an expanded CUG repeat alters the activity of RNA splicing factors, resulting in mis-regulated alternative splicing of many different genes (4). Secondly, the expanded repeat can affect its own expression, that of its anti-sense transcript and genes upstream and downstream of *DMPK* (5–12). Both modes of pathogenesis are affected by CTG expansion size, and the latter is also linked to epigenetic modifications of the DM1 locus. The DM1 CTG repeat is located in a gene-rich region of the genome (13) and is embedded in a large 3.5 kb CpG island (Fig. 1A and B) (14). In addition to its presence in the 3′-UTR of the *DMPK* gene, the CTG repeat is in the promoter of the downstream *SIX5* gene. The expanded CTG repeat modifies the transcription of *SIX5* (5–7,10) and *DMWD* (13,15). This modified allelic expression of DM1 locus genes can vary between tissues (7,8,16–20), a phenomenon thought to be mediated by altered chromatin packaging (10,21,22). In addition, transcription across the anti-sense strand of the *DMPK* gene can also be modified by the expansion, a phenomenon that is regulated by binding of the chromatin insulator CTCF to sequence motifs adjacent to the DM1 CTG tract (9,11). These allelic alterations in gene expression are thought to contribute to the varied symptoms of DM1 (23–25). Chromatin packaging is altered at the DM1 locus with expanded DM1 repeats compared with the corresponding wild-type allele. Strong nucleosome positioning upon the expanded CTG repeats (26) and loss of a DNaseI and nuclease hypersensitive site some 700 bp downstream of the repeat (9,10,27) suggested a compact chromatin packaging. Histone marks associated with heterochromatic conformation were present in a DM1 cell line (11). This DM1 chromatin packaging is thought to be mediated by CpG methylation-sensitive binding of the chromatin insulator CTCF proximal to the CTG repeat. Other disease-associated repeat expansions and contractions such as fragile X syndrome [FRAXA (MIM 300624)], Friedreich ataxia [FRDA (MIM 229300)] and facioscapulohumeral muscular dystrophy [FSHD (MIM 158900)] are also known to have differential CpG methylation and chromatin packaging which affects gene expression (reviewed in 21,22).

DM1 patients display high levels of CTG instability between different tissues of the same individual, with large inter-tissue CTG length differences (28–35). This somatic instability is detectable as early as 13–16 weeks during foetal development and continues following birth to varying degrees in different tissues (28,30,36–40). Although the significance of somatic repeat heterogeneity to disease is unclear, ongoing expansions have been proposed to exacerbate DM1 symptoms as they are usually largest in clinically important tissues such as skeletal muscle, heart and brain (36,40). Recent evidence suggests that CpG methylation and CTCF binding may contribute to CTG/CAG repeat instability (41–45). However, it is unknown whether and how CpG methylation and CTG length heterogeneity may correlate between tissues and developmental periods. Although instability shows development- and tissue-specific patterns, an analysis of both foetal and adult tissues in a single study has not been performed. The use of newly developed methods will allow for improved higher-resolution analyses of somatic CTG instability profiles (46) compared with studies performed over 15 years ago using low-resolution methods (28,30,36–40).

The dynamic regulation of CpG methylation in tissue- and development-specific manners and its important roles in controlling cellular processes (47–52) make this epigenetic mark worthy of analysis at the DM1 locus. However, information of the CpG methylation status at the DM1 locus is very limited. Only two reports on CpG methylation at the DM1 locus have been published (summarized in Supplementary Material, Fig. S1). Using methylation-sensitive restriction analysis, both reports observed constitutive methylation in all individuals tested [non-affected, classical DM1 and congenital DM1 (CDM)]. Shaw *et al.* (53) found methylation (∼3300–8700 bp) upstream of the CTG tract. Steinbach *et al.* (54) also found methylation (1159–1232 bp) upstream of the repeat tract. These sites appear to coincide with recently observed methylation in non-affected cells (Supplementary Material, Fig. S1) (52). Shaw *et al.* (53) reported no methylation differences between unaffected individuals and either classical DM1 or CDM patients. In contrast, Steinbach *et al.* (54) observed complete methylation at numerous *Hpa*II and *Hha*I upstream sites, extending up to a single *Sac*II site just upstream of the CTG repeat (as close as 42 bp) for patients with large CTG expansions [up to (CTG)1833 repeats] and juvenile onset or congenital disease. In the same study, classical DM1 adult onset DNAs with similar repeat lengths did not show the same methylation. Both studies were limited, in that methylation was assessed primarily in leukocytes of post-natal individuals and did not take into account potential tissue-specific or development-specific differences. Furthermore, both studies only assessed methylation upstream of the repeat and used methyl-sensitive restriction digestion, which misses many CpG sites and is not quantitative.

Various points indicate that a deeper appreciation of the methylation status at the DM1 locus is needed: considering (i) the tissue specificity of DM1 disease symptoms; (ii) the tissue-, development- and age-specific somatic instability of the CTG repeat; (iii) the many DM1 region transcripts (sense and anti-sense) whose expression may be affected by methylation; (iv) the regulation of DNA methylation known to be dynamically modulated; make it interesting to learn the methylation status in a variety of tissues from DM1 individuals of various ages and developmental stages and (v) recent evidence suggests that high levels of non-CpG cytosine methylation can occur predominantly at CAG and CTG sequences (52) and such methylation has been reported at non-expanded repeats of certain disease loci (55). The methylation status of the expanded repeat and adjacent sequences in patient tissues is unknown.

We characterized at high resolution the CpG methylation profile of the DM1 locus, focusing on regions close to the CTG repeat (Fig. 1C and D). Methylation profiles using bisulphite sequencing were determined at each CpG site upstream and downstream of the CTG repeat including the two DM1 CTCF binding sites (9). Non-CpG methylation of the CTG/CAG repeat was assessed using a newly developed restriction assay. Both methylation and repeat length were assessed in various tissues and developmental windows in several individuals (foetuses and adult autopsies) and in transgenic DM1 mice. Tissue-, age- and DNA region-specific variations in CpG methylation were observed only in tissues with CTG expansions. Contrary to previous findings

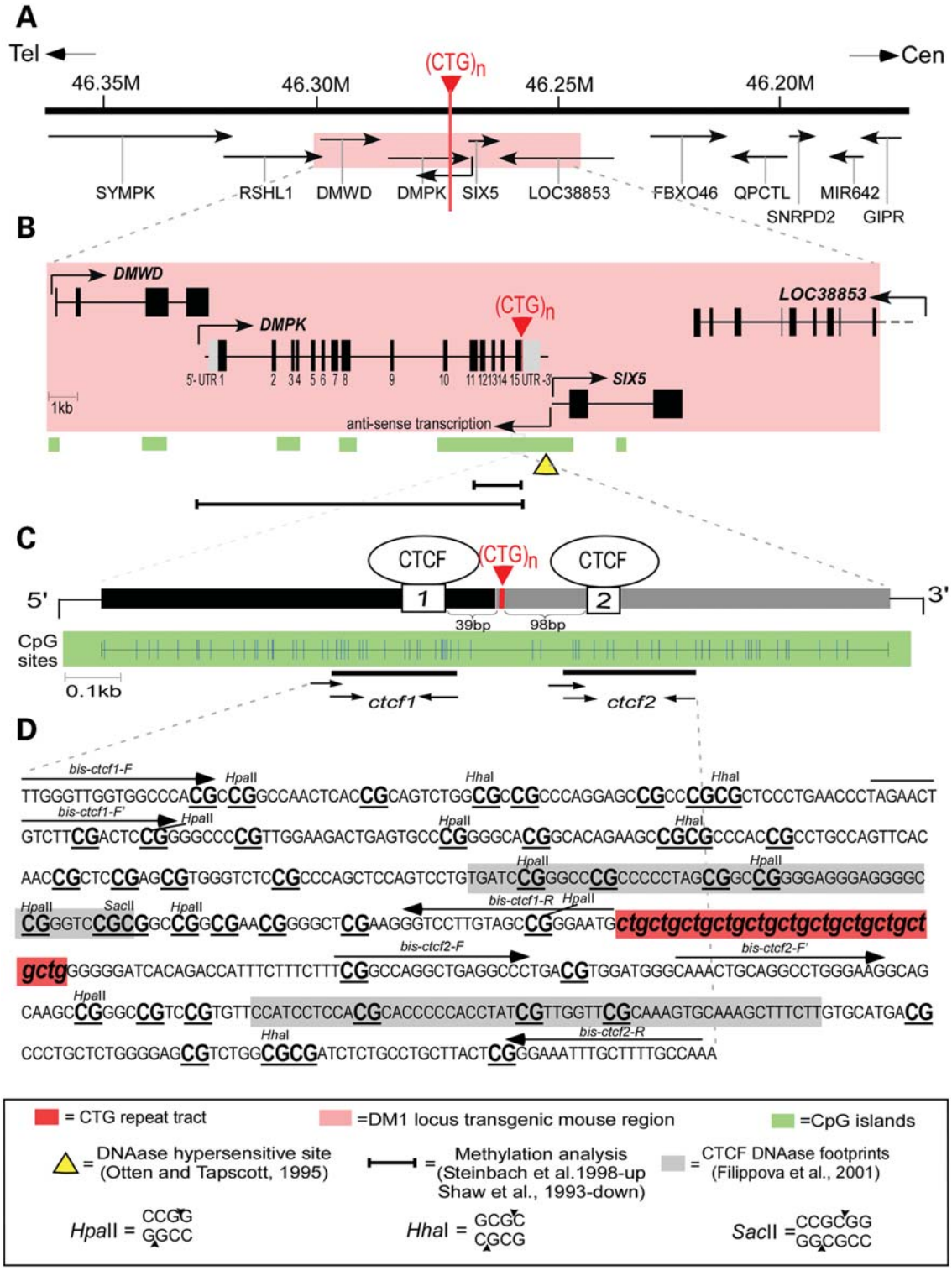


**Figure 1.** Detailed genomic and sequence location of the DNA regions analysed for CpG methylation. (**A**) View of the gene-enriched region in human chromosome 19 where the *DMPK* gene and the CTG repeat tract are enclosed. Pink shadow shows the human region (∼45 kb) single integrated in the DM1 transgenic mice genomes. (**B**) A more detailed view of the ∼45 kb region that includes the *DMPK* gene (20 or more than 300 CTG repeats in mice DM20-949 and DM300-328, respectively) and neighbouring genes. CpG islands through this region are marked in green, as derived from the UCSC Genome Browser (104). The regions previously analysed for methylation (black bars) (53,54) and the DNAase hypersensitive site detected upstream (yellow triangle) (27), coincident with the anti-sense transcription initiation reported in the DM1 locus (11), are also indicated. (**C**) The *DMPK* regions analysed by bisulphite sequencing PCR, located proximal and distal to the CTG repeat tract, are shown, as well as the CTCF binding sites 1 and 2 (grey shadows) (9). (**D**) *DMPK* sequence analysed by bisulphite sequencing PCR. PCR primers location, CTCF binding sites, CpG sites and *Hpa*II, *Hha*I and *Sac*II methylation-sensitive sites are indicated. For all sections, the CTG repeat is shown in red. *Ctcf1* and *ctcf2* primer sets for hemi-nested PCR amplification after DNA bisulphite treatment are shown.

(54), we do not find hypermethylation to be strictly limited either to juvenile/congenital samples or to tissues with the largest CTG expansions. In addition to proximal methylation being restricted to the disease allele, we observe an abrupt boundary of methylation, polarized to the region upstream of the expanded CTG tract. The expanded repeat itself and the downstream region from the repeat did not show (high levels) methylation. Furthermore, the degree of this polarized methylation varies between foetal and adult samples. Thus, the abnormal methylation at the DM1 locus cannot be generalized from a given tissue, time point, DNA region or cell lineage.

## RESULTS

### CTG length heterogeneity in DM1 patient tissues

Various post mortem tissues from human foetal, adult DM1 patients and non-DM1 controls (Table 1) were collected to assess the DNA methylation status and to determine the inter- and intra-tissue CTG length variations. For repeat length analysis, we used a high-resolution method [locked nucleic acid (LNA) probe-Southern] that takes advantage of the increased electrophoretic resolution of DM1 restriction fragments having a minimum of non-repetitive flanking sequences, reduced background hybridization and a repeat-specific LNA probe (46). Representative Southern blots are shown in Figure 2, and CTG sizes for each individual are summarized in Table 1. Short DM1 alleles were polymerase chain reaction (PCR)-amplified and sequenced.

Of the foetal tissues assessed, the largest expansions often arose in heart and muscle, whereas the smallest expansions were in liver (Fig. 2A and Table 1; Supplementary Material, Fig. S2). Intra- and inter-tissue repeat heterogeneity was not apparent for the foetus with the shortest gestational age and smallest expansion (DC-FEDM3, 14 weeks, 1600 CTG repeats). Only two tissues from this foetus were available for analysis (heart and skeletal muscle). In DC-FEDM2 and DC-FEDM1, foetuses with larger repeats or longer gestation, inter-tissue differences of 120 and 40 repeats were detected between heart and skeletal muscle, respectively (Table 1). Overall, the greatest inter-tissue CTG length differences in foetal samples were 400 and 240 repeats for foetuses DC-FEDM2 (19 weeks, 3820 CTG repeats) and DC-FEDM1 (14 weeks, 2660 CTG repeats), respectively. The intra-tissue repeat length heterogeneity was limited in all foetuses, as evident by the relatively distinct bands. These modest levels of CTG instability in foetuses of 14–19 weeks are consistent with previous studies, suggesting that ~13 weeks is the developmental window when somatic instability starts to become apparent (35).

Adult DM1 tissues displayed large inter-tissue CTG length variations, and the intra-tissue length heterogeneity was also extensive (Fig. 2B and Table 1; Supplementary Material, Fig. S3). The greatest length range was in the cerebral cortex, followed by liver and skeletal muscle, with length heterogeneity spanning as many as 2800 repeats (Table 1, Supplementary Material, Fig. S3). In adult tissues, the expansions were consistently largest in the heart and cerebral cortex (up to 6550 repeats), whereas the smallest expansions were in the cerebellum (Supplementary Material, Fig. S2). The greatest inter-tissue CTG length differences were 5770, 5240, 4900, 4890 and 1700 repeat units, for patients UR-ADM8, UR-ADM5, UR-ADM9, UR-ADM6 and UR-UDM4, respectively. The larger inter-tissue CTG length differences and intra-tissue CTG length heterogeneity observed in adult tissues are consistent with previously reported active instability occurring over the age of the DM1 individual (35–37).

Together, the foetal and adult samples support the existence of ongoing somatic CTG instability with age, as the adult samples displayed greater degrees of both inter-tissue and intra-tissue length heterogeneity. Expansion patterns between tissues were different for foetal and DM1 adults, suggesting tissues-specific rates or times of expansion. For example, comparing the foetal with adult normalized inter-tissue length differences, it is apparent that in foetuses, the liver incurs fewer expansions than heart, brain or muscle, whereas in adults, the liver has incurred similar expansions as heart, brain (cortex) and muscle (Supplementary Material, Fig. S2). Other evidence suggests that CTG expansions in the DM1 muscle occur more rapidly earlier than later in life (38). Only the expanded, mutant DM1 allele showed severe length variations. The non-expanded DM1 CTG repeat did not change dramatically in any of the DM1 foetal or adult tissues and did not show variation in control non-DM1 foetal or adult tissues (Table 1). We observed some minor products with CTG lengths just greater by a few repeat units compared with the non-expanded allele in cortex, skeletal muscle and cerebellum in the UR-ADM9 adult sample (juvenile DM1). Such limited DM1 CTG size fluctuations have been reported previously in somatic tissues of non-expanded alleles (muscle and heart) (56,57). Thus, CTG instability of the expanded allele was tissue-specific with age-specific variations.

### CpG methylation in DM1 patient tissues

We focused our methylation analysis upon regions immediately flanking unstable CTG tract, as these include the CTCF binding sites, demonstrated to epigenetically regulate expression of the DM1 locus (5,9,11). This region includes 18 CpG sites upstream and 11 downstream from the CTG repeat, respectively (Fig. 1). Only minimal methylation was present in non-DM1 foetal and adult tissue samples (Figs 3 and 4, top panels). In contrast, each of three DM1 foetuses that we examined showed prominent methylation (Fig. 3, bottom panel). Strikingly, the location of the methylation was highly polarized (methylation was present at every CpG site analysed upstream of the CTG repeat), whereas methylation downstream of the repeat was absent in two DM1 foetuses and sparse in the third (Fig. 3, bottom panel). The presence of methylation upstream but not downstream of the expanded repeat suggested that the repeat may block methylation progression. However, that the CTG repeat itself may harbour non-CpG methylation could not be excluded. The presence of non-CpG methylation was assessed through the expanded CTG/CAG repeat (Fig. 5) to better demarcate the borders of aberrant methylation. High levels of non-CpG methylation in the expanded repeat were not detected in foetal samples that showed high levels of aberrant CpG

**Table 1.** Human samples analysed

| Patient | Age of examination | Gender | Disease status | Parental transmission | Source | CTG repeat alleles (length) | |
|---|---|---|---|---|---|---|---|
| | | | | | | Short | Expanded (peak if smear) |
| DC-FEFX1 | Foetus: 14 weeks | Male | Non-DM1 | — | Skeletal muscle | 13/13 | — |
| | | | | | Kidney | | — |
| | | | | | Liver | | — |
| | | | | | Heart | | — |
| DC-FEDM1[a] | Foetus: 14 weeks | Male | (DM1/CDM) | Maternal (5/800) | Heart | 11 | 2660 |
| | | | | | Pancreas | | 2660 |
| | | | | | Skeletal muscle | | 2620 |
| | | | | | Kidney | | 2620 |
| | | | | | Skin | | 2620 |
| | | | | | Brain | | 2580 |
| | | | | | Liver | | 2420 |
| DC-FEDM2[a] | Foetus: 18–19 weeks | Male | (DM1/CDM) | n.d. | Heart | 11 | 3820 |
| | | | | | Skeletal muscle | | 3700 |
| | | | | | Kidney | | 3590 |
| | | | | | Brain | | 3530 |
| | | | | | Skin | | 3530 |
| | | | | | Liver | | 3420 |
| DC-FEDM3 | Foetus: 14 weeks | Male | (DM1/CDM) | n.d. | Heart | 13 | 1600 |
| | | | | | Skeletal muscle | | 1600 |
| ADN1 | Adult: 54 years | Male | Non-DM1 | — | Cerebellum | 11/14 | — |
| | | | | | Skeletal muscle | | — |
| | | | | | Cortex | | — |
| | | | | | Liver | | — |
| | | | | | Heart | | — |
| UR-ADM9[a] | Adult: 44 years | Female | Juvenile DM1 | n.d. | Heart | 12 | 4100–5300 (4600), 6000 |
| | | | | | Liver | | 4200–4600 (4500) |
| | | | | | Skeletal muscle | | 3700–4300 (4200) |
| | | | | | Cortex | | 4000–5500 (4900) |
| | | | | | Cerebellum | | 1100–1500 (1200), 3800–4600 (—)[b] |
| UR-ADM6[a] | Adult: 47 years | Male | Classical DM1 | n.d. | Skeletal muscle | 5 | 3700–4300 (4200) |
| | | | | | Cortex | | 2800–5300 (4800) |
| | | | | | Liver | | 3900–4900 (4500) |
| | | | | | Heart | | 3500–5300 (5200) |
| | | | | | Cerebellum | | 410, 580 |
| UR-ADM4[a] | Adult: 50 years | Female | Classical DM1 | n.d. | Heart | 10 | 4900, 6300 |
| | | | | | Skeletal muscle | | 4600–5400 (5100) |
| UR-ADM8[a] | Adult: 53 years | Male | Classical DM1 | n.d. | Heart | 14 | 4200–5300 (4900) |
| | | | | | Skeletal muscle | | 3300–4400 (3900), 5300–6100 (—)[b] |
| | | | | | Liver | | 4500–4900 (4800) |
| | | | | | Cortex | | 2200–5500 (4900) |
| | | | | | Cerebellum | | 330, 440, 4200–4500 (—)[b] |
| UR-ADM5[a] | Adult: 55 years | Male | Classical DM1 | n.d. | Heart | 4 | 5100–6550 (6000) |
| | | | | | Kidney | | 4900–5600 (5300) |
| | | | | | Liver | | 4600–5400 (5000) |
| | | | | | Pancreas | | 4400–5100 (4600) |
| | | | | | Cortex | | 2900–5700 (5400) |
| | | | | | Skeletal muscle | | 3900–4400 (4100) |
| | | | | | Cerebellum | | 1310 |
| UR-ADM3 | Adult: 80 years | Male | Minimal DM1 | n.d. | Blood | 5 | 80 |
| | | | | | Grey matter | | 177 |
| | | | | | White matter | | 105 |
| DMPK493[b] | Adult: ? | n.d. | Non-DM1 | — | Skin fibroblast Cell line | 5/25 | — |
| DMRB993[a,c] | Adult: 12 years | Female | Congenital DM1 | n.d. | Skin fibroblast Cell line | 14 | 2973 |
| 94-916[d] | Adult: 54 years | Female | Classical DM1 | Both parents | Blood | — | 60/1000 |
| 99-423[e] | Adult: 18 years | Female | Assympt. | Both parents | Blood | — | 85/90 |
| 06-2095[d] | Adult: 20 years | Female | Assympt. | Both parents | Blood | — | 75/90 |

DM1 or CDM, not diagnosed DM1 form; assympt., assymptomatic; n.d., no data.
[a]Long expanded alleles sized by LNA-Southern approach (46).
[b](—) Weak smear, no peak calculation.
[c]See Krahe *et al.* (58) and Frisch *et al.* (10).
[d]See Martorell *et al.* (60).
[e]See Cobo *et al.* (59).

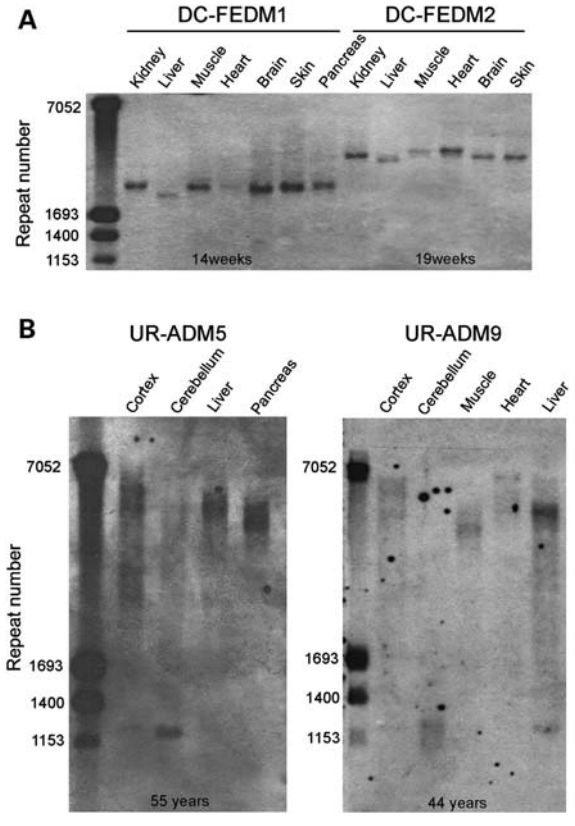


**Figure 2.** Tissue-specific CTG repeat sizing of the expanded allele in human DM1 samples. Representative LNA-Southerns (46) of some (**A**) foetal and (**B**) adult DM1 individuals. The range of CTG lengths in a given tissue was considerable and difficult to assess precisely as many were evident as smears (Table 1). As previously noted by others (7), the sensitivity to detect the range of smeared products depended on the amount of DNA loaded.

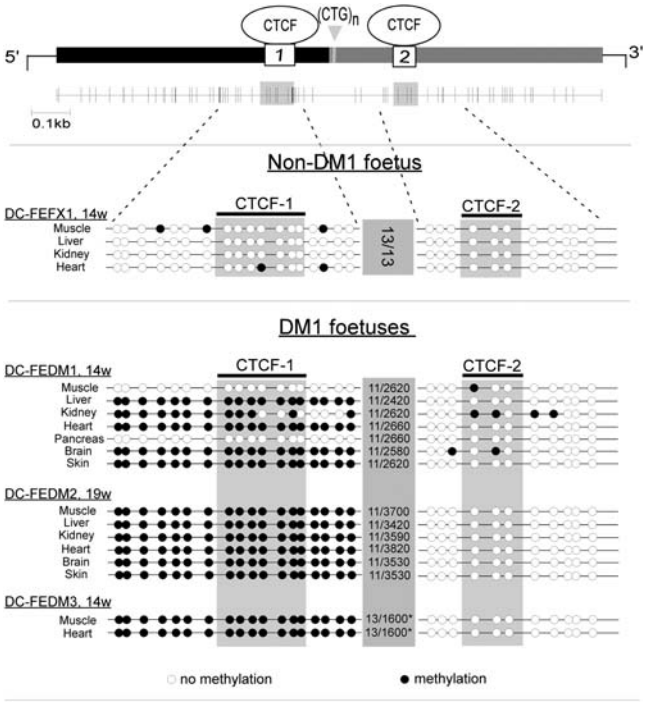


**Figure 3.** CpG methylation profiles from unborn individuals. Top panel: non-DM1 foetus analysed (muscle, liver, kidney and heart); bottom panel: DM1 foetuses analysed including different tissue panels depending on the individual. *DMPK* alleles CTG repeat lengths (inside red box) and CTCF binding sites location (grey shadows) are indicated. White dots, CpG sites devoid of methylation; black dots, CpG sites with methylation presence.

methylation upstream of the repeat (heart and brain in DC-FEDM2).

Although high levels of methylation were present in most foetal tissues, a few dramatic tissue-specific differences were also observed in DC-FEDM1, a foetus with early gestation. In contrast to highly methylated tissues, skeletal muscle and pancreas were not methylated in the same patient, suggesting DM1 tissue-specific methylation issues already in early disease developmental stages. Simultaneously, the high levels of methylation upstream of the CTG repeat are encompassing the CpG sites critical for CTCF binding in many foetal tissues.

Methylation was also assessed in adult post mortem tissues from four individuals with classical DM1 and one adult with juvenile-onset DM1. Once again, highly expanded CTG repeats in these samples were associated with abnormal methylation in every patient. Similar to foetal samples, the adult tissues exhibited a striking polarity, so that methylation was restricted to sites upstream of the expanded repeat. When compared with foetal sample, the overall levels of methylation in adult tissues were lower (compare Figs 3 and 4, bottom panels). Most adult DM1 tissues showed partial methylation (no presence of methylation in all CpG sites assessed) in the upstream region. Notably,methylation in DM1 adults did not consistently occur at specific CpG sitesor show an obvious correlation with CTG expansion size. In general, the heart, cerebral cortex and liver often displayed more methylation, whereas the cerebellum and skeletal muscle displayed minimal levels of methylation. Methylation levels were less in adult tissues compared with matching tissues of foetuses.

## Aberrant methylation is specific to the expanded DM1 allele

To test whether the methylation detected by bisulphite sequencing in DM1 samples was specific to the disease allele, we devised a methyl-specific multiplex PCR assay using the methylation-sensitive enzyme *Sac*II, for which there is one proximal site (42 bp) upstream of the CTG tract (Fig. 6), the same *Sac*II site assessed by Steinbach *et al.* (54). Our forward and reverse primers were placed on opposite sides of the CTG repeat and therefore were only capable of amplifying across the non-expanded CTG tract. One downstream and two upstream primers that did and did not cover the *Sac*II sitepermitted determination of the methylation status of the *Sac*II site in the non-expanded allele (Fig. 6A). Results revealed that in the absence of *Sac*II digestion, two PCR products, both from the non-expanded allele, were evident. Successful *Sac*II digestion, an indication of a non-methylated template, eliminated only the PCR product that encompassed the *Sac*II site. Representative examples are shown in Figure 6B. Thus, the absence of methylation on

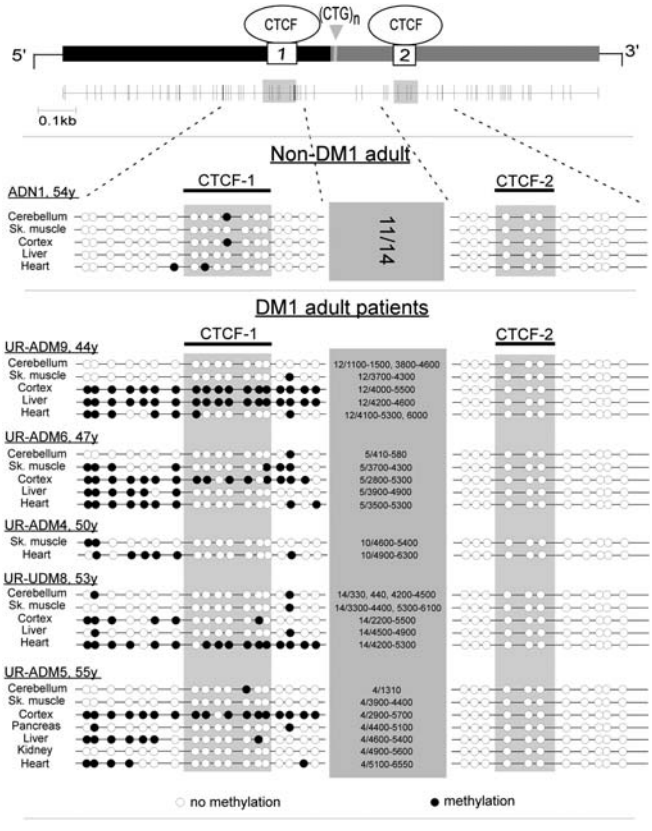


**Figure 4.** CpG methylation profiles from adult individuals. Top panel: non-DM1 adults analysed (cerebellum, skeletal muscle, cortex liver and heart); bottom panel: DM1 adults analysed including different tissue panels depending on the individual. *DMPK* alleles CTG repeat lengths (inside red box) and CTCF binding sites location (grey shadows) are indicated. White dots, CpG sites devoid of methylation; black dots, CpG sites with methylation presence.

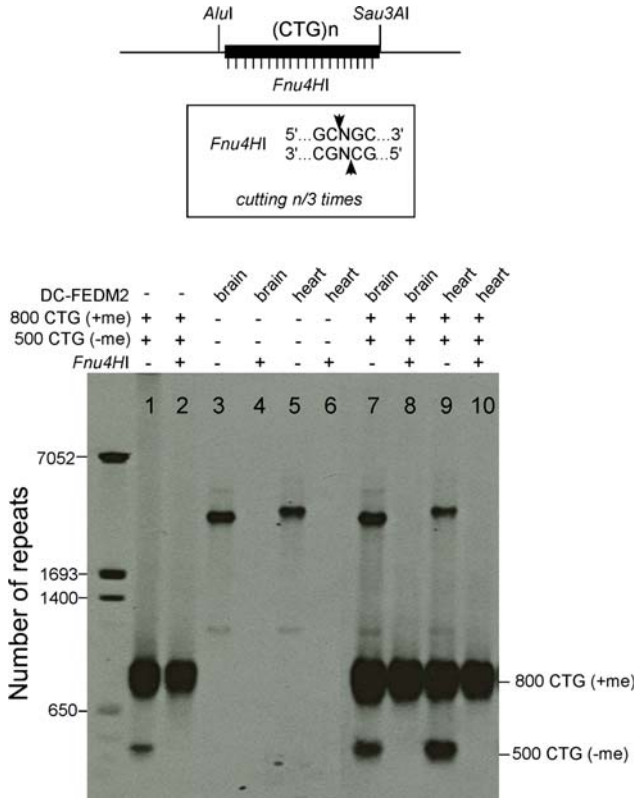


**Figure 5.** Assessment of non-CpG methylation through DM1-expanded alleles. DC-FEDM2 brain and heart samples were used by the strong difference between upstream (high methylated) and downstream (unmethylated) levels of CpG methylation (Fig. 3). DNA digestion with the non-methylation-sensitive enzymes *Alu*I and *Sau3A*I (to free the repeat) and *Fnu*4HI (sensitive to CpNpG methylation) (top panel) was combined with LNA-Southern detection (46). No presence of high levels of methylation was detected (lanes 4, 6, 8 and 10, bottom panel). DNA fragments of 800 CTG repeats (methylated with the *M.CviPI* GpC methylase) and 500 CTG repeats (unmethylated) were included as controls for the methyl-sensitive *Fnu*4HI digestion, either alone (lanes 1 and 2) or with the DC-FEDM2 genomic DNA (lanes 7–10).

the non-expanded allele supports the conclusion that the high levels of CpG methylation detected by bisulphite sequencing originated from the expanded allele.

All of the DM1 foetal and DM1 adult tissues described earlier contained large CTG tract sizes with the shortest being more than 330 repeats (cerebellum) or more than 1600 in other tissues. A DM1 patient with a small CTG expansion of 80, 105 and 177 CTG repeats in the blood, cortex white matter and cortex grey matter (UR-ADM3, Table 1) displayed essentially no methylation in any of the CpG sites tested (Fig. 7), which contrasts with the high levels of methylation in the cortex of DM1 adults with larger expansions. It is unclear whether this absence of methylation indicates a CTG expansion size dependence upon becoming methylated or of methylation upon large expansion. Due to the limited availability of DM1 patient tissues with short expansions, we further addressed a possible association of methylation with CTG expansion size by assessing the methylation status of patient DNAs derived from either blood or cultured fibroblast cells. These included DM1 expansions of 133, 300 and 2973 repeats as well as DNAs from DM1 patients with homozygous expansions that are typically in the shorter range (60/1000; 85/90 and 75/90 repeats) (Fig. 7 and Table 1) (10,58–60). Only the largest expansion (CTG)2973 showed methylation, and this was present only in the upstream region. Together, these results support the suggestion that abnormal methylation correlates with large repeat expansions. Such a conclusion is consistent with the high levels of methylation in cells with expansions of (CTG)1833 repeats, but not in those with less than (CTG)1000 repeats (54). However, repeat length alone is unlikely to be the sole determinant of methylation, as many tissues with long repeat tracts did not show methylation, whereas other tissues from the same individual did (Fig. 4).

## Methylation of the human DM1 locus in transgenic mice

Transgenic organisms are powerful tools in the study of diseases caused by unstable repeats, and it is important to appreciate how these models reflect what is occurring in patients (Supplementary Material, Table S3 in reference 3). Towards such a comparison, we assessed CpG methylation in a well-characterized DM1 transgenic mouse model, which contains a single-copy integration of ~45 kb large segment of the human DM1 locus derived from a DM1 patient (Fig. 1A and B). Mice analysed

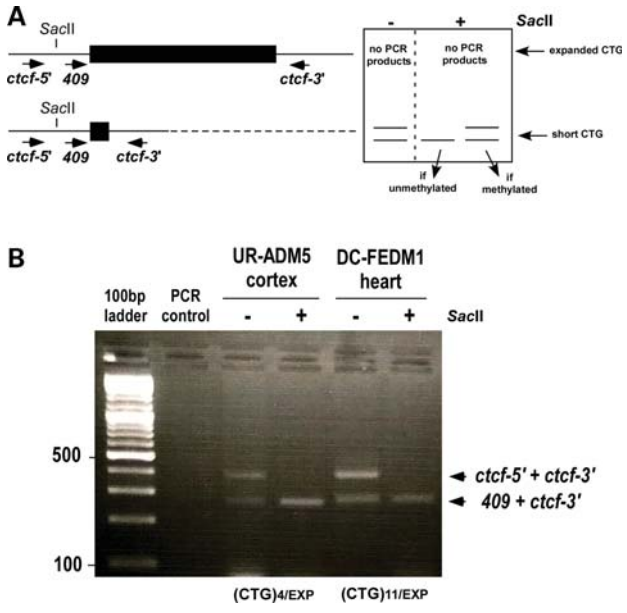


**Figure 6.** Assessment of methylation allele specificity. (**A**) *Sac*II methylation-sensitive digestion combined with multiplex PCR amplification was performed, using two upstream primers and one downstream primer to produce two PCR products from only the non-expanded CTG allele, that either did or did not encompass the methylation-sensitive *Sac*II site present upstream of the CTG tract. Following *Sac*II digestion, the elimination of only the PCR product encompassing the *Sac*II site revealed that this site in the non-expanded allele was not methylated in the DC-FEDM1 (heart) and the UR-ADM5 (cortex) samples, whichby bisulphite analysis showed high levels of upstream methylation (Figs 3 and 4, respectively). Thus, the high levels of CpG methylation detected by bisulphite sequencing were exclusively present on the expanded DM1 allele.

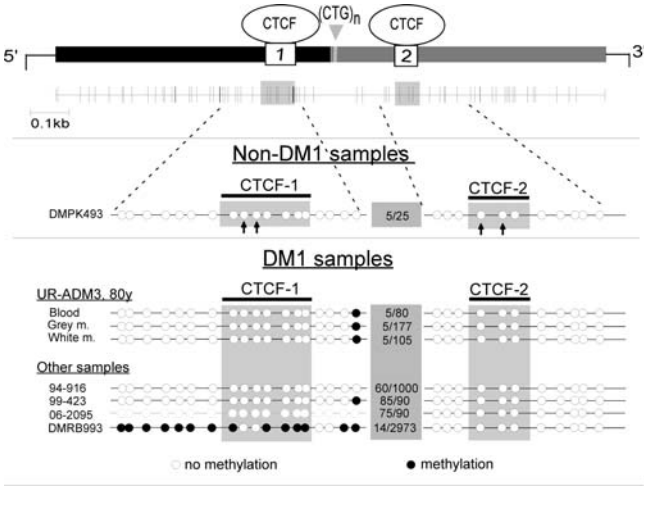


**Figure 7.** CpG methylation levels from other non-affected and DM1 samples. Methylation analysis was performed over some other samples, including fibroblast cells (DMPK493 and DMRB993, from a non-affected and DM1 individual, respectively), tissues from a DM1 individual with a short CTG repeat expansion (UR-ADM3) and blood samples from rare individuals with both alleles in the range of disease-associated lengths. More details of all these samples are given in Table 1. *DMPK* alleles, CTG repeat lengths (inside red box) and CTCF binding sites location (grey shadows) are indicated. White dots, CpG sites devoid of methylation; black dots, CpG sites with methylation presence. No data from the upstream region in the 06-2095 sample.

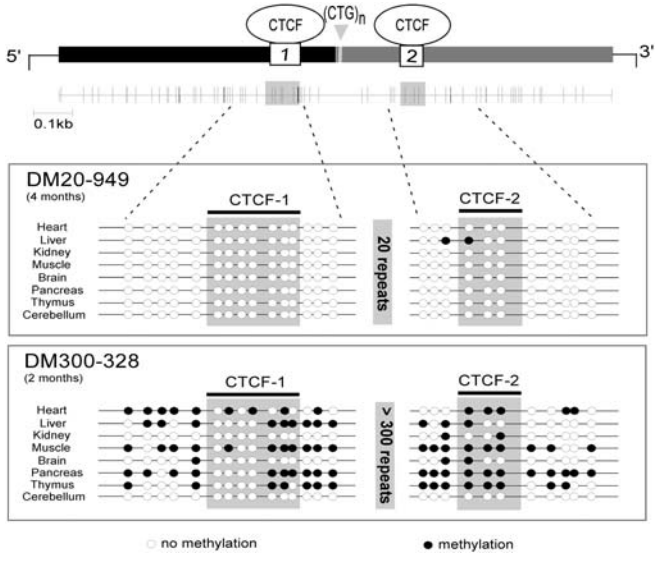


**Figure 8.** DM1 transgenic mice CpG methylation profiles. Top panel: methylation from DM20-949 (20 repeats) mouse tissues; bottom panel: methylation from DM300-328 (more than 300 repeats) mice tissues. The analysis included 15 CpG sites upstream and 11 CpG sites downstream from the CTG repeat tract. White dots, CpG sites devoid of methylation; black dots, CpG sites with methylation presence; grey shadows, CTCF binding sites location. A pool of three to eight mice was analysed.

harboured a short (DM20-949) or an expanded allele (DM300-328) with (CTG)20 or (CTG) > 300, respectively. The same CpG regions adjacent to the repeat and a similar tissue set were assessed in the mice as in the human samples.

Methylation of the human DM1 transgene in the DM1 mice showed similarities compared with the endogenous locus in human DM1 individuals: (i) both showed methylation exclusively on the expanded allele. No detectable or limited levels of methylation are present in the non-expanded line (DM20-949) (Fig. 8, top panel) contrasted with many methylated CpG sites in the expanded line (DM300-328) in the same tissues (Fig. 8, bottom panel); (ii) both humans and mice showed tissue-specific patterns of methylation. The DM300-328 cerebellum was devoid of methylation, whereas skeletal muscle and pancreas were highly methylated (Fig. 8, bottom panel) and (iii) both humans and mice showed less methylation in foetal/younger ages compared with adult ages. In young DM300-328 mice (2 months) and older (6 months), mice showed an age-dependent decrease of methylation levels (pancreas, heart and thymus in Fig. 9). But humans and mice methylation also differed in: (i) mice showed age-dependent increases of methylation (the brain, for example); (ii) methylation patterns for a given tissue did not correlate exactly between species (compare Figs 4 and 8) and (3) unlike the human locus, which showed a polarized localization of methylation, the mice showed methylation both up- and downstream of the repeat with a preference for higher levels upstream (Fig. 10A). Thus, in some ways, the DM1 mouse model recapitulates methylation features of the DM1 locus in patients. However, variations between the DM1 mice and humans are evident, and these must be considered. A more detailed description of methylation in DM1 mice is given in Supplementary Material.

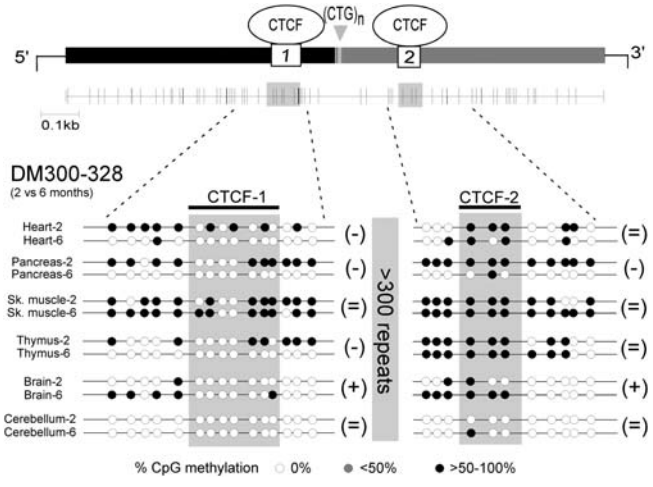


**Figure 9.** Age-specific methylation in DM300-328 mice tissues. Methylation was assessed in 2 and 6 months DM300-328 mice in the presence of more than 300 CTG repeat tracts in the *DMPK* gene. Changes in methylation levels are indicated for upstream and downstream regions using +, − and =, for increases, decreases and minimal overall changes based on the number of CpG sites methylated.

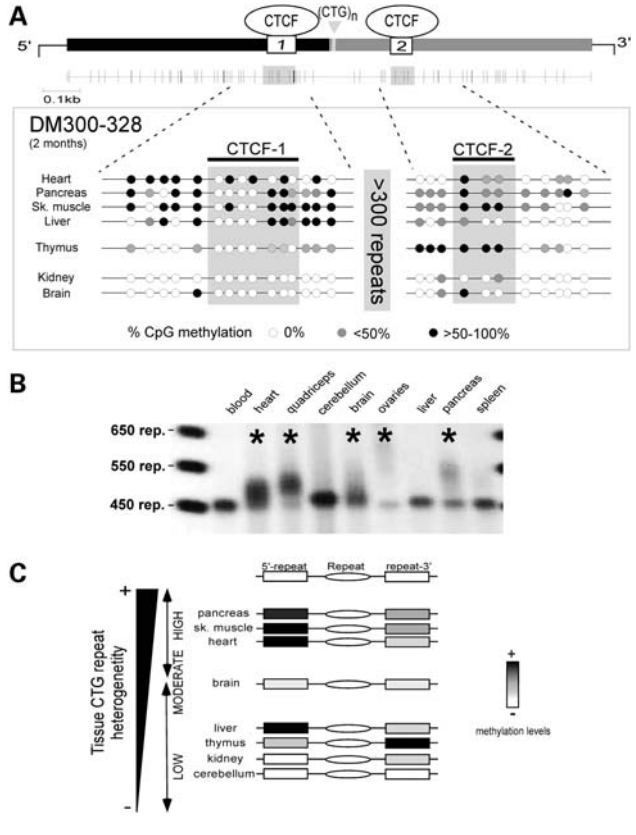


**Figure 10.** Tissue CTG repeat heterogeneity and CpG methylation levels in DM300-328 mice. (**A**) Methylation quantification (based on CpG sites' height peaks in the chromatogram trace files using 4Peaks software, see Supplementary Material, Fig. S4). CpG sites devoid of methylation (white dots), with less than 50% methylation (grey dots) and from 50 to 100% methylation (black dots) CTCF binding sites location (grey shadows). (**B**) Somatic repeat heterogeneity profiles between different DM300-328 tissues in old mice (19 months) (99). *Unstable somatic tissues. (**C**) Methylation levels and repeats heterogeneity association in the tissues analysed from DM300-328 mice.

## DISCUSSION

We mapped the sites of CpG methylation closest to the DM1 CTG repeat in a panel of human foetal and adult DM1 tissues. Methylation is present upstream of CTG expansions, but not downstream of the repeat or through the repeat itself, and not in non-expanded alleles. Upstream proximal methylation was specific to the expanded allele and present in classical and juvenile DM1 forms. Methylation is not uniform among tissues, not limited to tissues with the largest expansions, and is less conspicuous in older individuals.

Previous studies detected methylation far upstream (~3300–8700 and 1159–1232 bp) of the DM1 CTG repeat in all individuals analysed, including non-DM1 samples (Supplementary Material, Fig. S1) (53,54). High levels of methylation upstream and proximal to the repeat were also reported only in CDM (54). The most striking levels of methylation they observed were in an 'exceptional' dura matter cell line of a CDM foetus (P57F), which was compared with blood DNA of classical DM1 individuals assuming an absence of tissue specificity (54). Importantly, in our study, methylation was not limited to young, severely affectedor CDM individuals, as previously suggested (54). A major difference between our approach and others is that previous studies used predominantly DNA from blood, whereas we studied DNA from various tissues. Methylation varied between tissues and between adults and foetuses, indicating that methylation status cannot be generalized from the analysis of a single tissue, time point in lifeor cell lineage. This confirms the suspicions of tissue specificity and developmental age, raised by Shaw *et al.* (53). Unlike Steinbach *et al.* (54), we do not observe a limitation of methylation to CDM individuals. Steinbach *et al.* (54) may not have detected methylation inclassical DM1 individuals due to (i) the older DM1 patients they assessed may have lost or partially lost methylation, as observed in our aged DM1 adult patients, and (ii) the limited single-site methylation sensitivity of *Sac*II restriction digestion analysis. We extended in high resolution the region analysed and delimited abnormal CpG methylation to the region upstream of the repeat region as CpG methylation downstream of the repeat is almost non-existent in human samples.

### DM1 methylation spreading, polarity and boundary and polar elements

The presence of proximal methylation upstream of the expanded CTG repeat in DM1 patients, coupled with the far upstream methylation in non-affected individuals, raises the possibility that the mutant allele incurred methylation spreading from far upstream towards the expanded repeat. Such methylation spreading may have arisen due to the loss of a methylation boundary between these two regions. Such a phenomenon was recently reported for the *FMR1* gene region, in which methylation far upstream in non-affected individuals was extended in FRAXA patient cells to encompass *FMR1* promoter adjacent to the expanded CGG repeat (61). In FRAXA, FRAXE, FRAXF, FRAX10A, FRAX11B and FRAXA16A, the aberrant methylation included the expanded CGG repeat and downstream regions (62,63). If

methylation spreading occurred on the mutant DM1 alleles, it was limited to regions upstream of the CTG tract.

The abrupt demarcation of abnormal methylation at the DM1 locus suggests the presence of an insulator element acting as a barrier to block further spreading of methylation. The DM1 CTG repeat is flanked by CTCF binding sites, and CTCF binding at other genomic loci has been shown to serve as a boundary element between highly methylated and unmethylated genomic regions, defining regions of X-inactivation, genetic imprintingor promoter activity (64–68). In those cases, the boundaries between methylated and unmethylated DNAs were defined by an enrichment of bound CTCF protein at an unmethylated CTCF binding site (65–69). However, at the DM1 locus, it is unlikely that the upstream CTCF-1 binding site can explain the abrupt boundary of methylation that we observed to occur at that point. This is unlikely because the methylation extends over and beyond the CTCF-1 binding site, right up to the 5′-end of the CTG tract. This methylation of the CTCF-1 binding site would be expected to abrogate CTCF binding, as has been observed by loss of *in vitro* binding (9), loss of chromatin immunoprecipitation and loss of *in vivo* footprints at the CTCF-1 site in DM1 cells (9,11,54). Thus, if the upstream CTCF-1 site is not the boundary element, this leaves either the downstream CTCF-2 binding site, which remains unmethylated in all cases, or the expanded CTG tract, which may serve as the boundary element. It is noteworthy that the expanded CTG repeat itself has been demonstrated to act as an insulator element by blocking transcription enhancer activity (9), and expanded CTG repeats in the absence of CTCF binding sitesexhibit variegation of expression of linked genes (70). Furthermore, the expanded DM1 CTG repeat constitutes an extended region (several kilobases) devoid of CpG dinucleotides, thereby interrupting and dividing the large 3.5 kb CpG island of the DM1 locus, which contrasts with the expanded CpG-containing CGG repeats in disease-linked rare fragile site syndromes. Interestingly, the aberrant DM1 methylation does not extend from the flank into the expanded CTG repeat; such an observation would be consistent with the inability of pre-existing CpG methylation to stimulate non-CpG methylation (71). Based on the demonstration that CAG repeats but not CGG repeats can stall the methyltransferase DNMT1 (72,73), we speculate that the CpG-free expanded CTG tract may insulate against CpG methylation spreading to downstream regions. Alternatively, the expanded repeat may act as a nidus, signalling abnormal polarized methylation.

Other instances of polarized elements relative to the DM1 CTG repeat have been reported. Chromatin packaging is altered in a polarized manner at the mutant DM1 locus with the loss of a DNaseI and nuclease hypersensitive site some 700 bp downstream of the repeat (9,10,27). Some expanded DM1 repeats contain non-CTG interruptions, which only occurred at the 3′-end of the CTG repeat tract (74,75). Repeat tracts bearing these interruptions displayed unique inter-generational instability. Recently, the replication origin that replicates the expanded DM1 repeat in transgenic mice was localized to the region downstream of the repeat (45). The polar arrangement of CpG methylation, chromatin modifications, non-CTG interruptions and replication origins is interesting. Future studies will reveal if there is overlap in the regulation of these polarized elements at the DM1 locus.

### DM1 methylation dynamics

The DM1 disease-associated methylation appeared to depend on age. The density of methylation upstream of the expanded DM1 CTG repeat tracts was decreased in older DM1 individuals compared with foetuses. This is unlike the aberrant methylation of the FRAXA locus, which is consistent between various ages (76–78). Age-dependent methylation changes have been reported for non-disease-associated methylation, and these dynamics vary between tissues and loci and can be affected by the environment (79). Presuming that the adult DM1 individuals assessed here contained similar high levels of aberrant CpG methylation during their foetal development, as the DM1 foetuses we analysed, there must have been a loss of methylation between these two ages. Other studies in the blood DNA of DM1 individuals aged $<1-62$ years suggest an age-dependent loss of methylation (54). Our analysis cannot determine when or how the DM1-specific aberrant methylation is gained or lost. The apparent loss of DM1 methylation did not correlate with specific CpG sitesor CTG expansion sizeor patient age. The decreased levels of methylation were often coincident with the increased levels of CTG length variations. It is unclear what the significance is of this apparent loss of aberrant DM1 methylation.

### DM1 methylation and CTG repeat instability

It is tempting to speculate that the absence or loss of methylation may be somehow linked to the coincident increase in inter-tissue differences in the CTG tract length. Such a relation exists in FRAXA in which evidence suggests that the presence of DNA methylation protects against contractions of expanded CGG tracts (63,80). In contrast, in many but not all adult DM1 tissues, the degree of intra-tissue somatic instability, indicated by repeats appearing as smears or multiple sized fragments, was greater in tissues retaining high levels of methylation. This might suggest that the retention of methylation correlates with increased CTG instability. Similarly, in the DM1 mice, the highest upstream methylation levels were associated with tissues displaying the largest and most heterogeneous repeat expansions (Fig. 10B and C). However, evidence from various model systems also does not present a consistent effect, as they suggest that methylation adjacent to CTG/CAG tracts can enhance or protect against repeat instability (41–44). Whether the loss of methylation is a cause of, a result ofor unrelated to the increased DM1 CTG expansions cannot be concluded, but the analysis herein suggests that if such a relation exists it is complex.

The influence of methylation on CTG instability in DM1 individuals may vary between tissues. Interestingly, the cerebellum, which has the shortest repeats, was devoid of methylation. Cerebellum in DM1 and other repeat expansion diseases such as Huntington's disease [HD (MIM 143100)], spinocerebellar ataxias 1 and 3 (MIM 164400 and MIM 109150, respectively) and dentatorubral-pallidoluysian atrophy (MIM 125370) consistently display the shortest CTG/CAG expansion size and

the most limited length heterogeneity, even compared with lymphocytes (7,19,39,81–84). Having the shortest expansion size among tissues, the cerebellum may have escaped somatic expansions and could reflect the progenitor CTG length present in the germ cell of the transmitting parent. Alternately, the short CTG repeat in the cerebellum may even represent contractions of the progenitor size. The cerebellum may harbour trans-factors that render it less susceptible to instability (85,86). Another possibility is that the low levels of methylation in the cerebellum (87,88) may provide a *cis*-environment that is protective against CTG instability or conducive to contractions. It is often suggested that the progenitor CTG length present in the germ cell of the transmitting parentcan be estimated as the shortest allele in the blood. However, blood is known to actively incur expansions over periods of 5–7 years (31), and hence blood may not be representative of progenitor lengths. Although we did not have access to bloods of any of our patients, Steinbach *et al.* (54) reported methylation in blood of DM1 patients. It is unclear whether this absence of abnormal methylation in the cerebellum indicates a dependence of CTG expansion upon methylation (or methylation loss). It would be interesting to assess the methylation status in the foetal DM1 cerebellum, as absence of methylation during early development when other tissues display high levels of methylationcould support the notion that CTG instability may arise during loss of methylation, whereas tissues that escape methylation, such as the cerebellum, would be spared somatic instability. Unfortunately, foetal cerebellums were not available. Analysis of additional tissues/individuals is necessary to approximate a relation of methylation with CTG instability.

### Mice versus men

Methylation of the human DM1 transgene in the DM1 mice showed similarities and differences compared with the DM1 locus in human DM1 individuals. Both showed expansion-specific methylation, tissue-specific methylation and changes in methylation levels with age in foetal/younger ages compared with adult ages. However, methylation patterns for a given tissue did not correlate between species; in humans, methylation was strongly polarized, whereas the mice showed higher methylation upstream of the CTG, but was also methylated downstream. For unknown reasons, variations between the DM1 mice and humans are evident and these must be considered.

### Implications of methylation to DM1 disease

Abnormal methylation has been observed at other unstable repeat disease loci, and for some, its role in pathogenesis is known. The FRAXA, FRAXE, FRAXF, FRAX10A, FRAX11B and FRAXA16A expanded repeats are extensively methylated leading to the inactivation of the associated gene. Methylation of three CpG sites proximal to the expanded GAA repeat in FRDA patients was coincident with altered histone modifications and consistent with inactivated gene transcription (89–91). The pathogenic role is less clear in the case of the aberrant hypomethylation in the contracted D4Z4 repeats of FSHDor the hypomethylated satellite repeats in patients with immunodeficiency, centromeric instability and facial anomalies (ICF) syndrome. CpG methylation at the HDlocus within (92) and 130 kb telomeric to the HTT gene (93) showed varying levels between tissues of HD patients (94,95) and their age (96). It is unknown whether methylation has a role in HD pathogenesis. Thus, for some disease loci, the role of methylation is known and for others it is less clear.

How the tissue-specific abnormal methylation upstream of the DM1-expanded CTG repeat might contribute to disease remains unknown. Although abnormal methylation is not unique to the most severe congenital form of myotonic dystrophy as previously thought (11,54), methylation upstream of the expanded CTG tract might correlate with disease onset or clinical variability. A simple correlation of our methylation results with previous DM1 gene expression levels in cells of DM1 individuals (5–7,11,16,19,97) is not obvious. We speculate that abnormal methylation might affect the expression of the toxic-CUG *DMPK* transcript, anti-sense-*DMPK* expression and/or *Six-5*. Such effects might vary with both the apparent age-dependent degradation of methylation and the variable CTG lengths in tissues, which could differentially affect the degree of MBNL1 protein sequestration and the severity of the resulting mis-splicing.

Finally, that many CAG/CTG repeat disease loci have proximal CTCF binding sites (9) are within large CpG islands (98), where some are known to have adjacent methylation that can vary between tissues and ages of affected individuals (94–96), and that most repeat diseases present tissue-specific vulnerability, suggests that the abnormal tissue-specific polarized CpG methylation described here at the DM1 locus might be present at other disease repeat loci.

## MATERIALS AND METHODS

### Transgenic mice

Transgenic mouse lines DM20-949 (20 repeats) and DM300-328 (more than 300 repeats), both harbouring a single-copy integration of ~45 kb human region that includes the *DMPK* gene, have been described previously (99) (Fig. 1). For studies of repeat instability and methylation at the DM1 locus, the DM20-949 and DM300-328 transgenic mice have several advantages. Both lines of transgenic mice are hemizygous and harbour single-copy integrations, which simplify the analysis of repeat stability and methylation. Also, the relatively large transgene fragment, totalling ~45 kb from the human locus and including >20 kb of sequence upstream and downstream of the CTG repeat, may insulate the repeat region from the chromatin environment that may exist at the respective integration sites. Indeed, the fragment includes three genes from the DM1 locus whose regulation in transgenic mice is similar to the human locus (100–102), suggesting that these transgenes are not subject to widespread heterochromatinization. Housing and handling of mice were performed according to the French government ethical guidelines.

### Human tissues

Human tissues analysed in this study are listed in Table 1. Autopsy tissues from a non-affected individual (ADN1) were

obtained snap-frozen from the National Disease Research Interchange. Tissues from DM1 foetuses and adult individuals were obtained at autopsy.

### DNA samples

DNA extraction from mice tissues was performed by homogenization of samples, proteinase K lysis and final phenol, chloroform and ethanol precipitation steps. DNAs from up to eight different tissues were collected from 4 months DM20-949 mice and 2 months DM300-328 mice. DNAs from specified tissues were extracted from 6 months DM300-328 mice. DNAs, from a minimum of three to eight mice bearing similar sizes of repeat tracts, were used for the analysis of each DM300-328 tissue. The same DNA extraction procedure was used to obtain genomic DNA from the various human tissue collections available for this study (listed in Table 1).

### CTG repeat analysis

Mice length of the large CTG repeat allele in the DM300-328 mice was examined by PCR, as described in Tomé *et al.* (103). Human-CTG repeat lengths were assessed by Southern blot using an LNA probe [DIG-labelled (CAG)7-5′-gcAgCagcAgCagCagcAgca-3′], as described previously (46). This sizing method takes advantage of the increased electrophoretic resolution of DM1 restriction fragments having a minimum of non-repetitive flanking sequences. Non-expanded CTG allele was sized by sequencing [The Centre for Applied Genomics (TCAG), MaRS Centre, Toronto, Canada] of the products obtained after PCR amplification (forward primer 409: 5′-gaagggtccttgtagccgggaa-3′; reverse primer 407: 5′-cagagcagggcgtcatgcaca-3′; 66.5°C annealing).

### CpG methylation analysis by bisulphite-sequencing PCR

Bisulphite treatment DNA (1–2 μg) was denatured by adding 5 μl of freshly prepared 3 M NaOH in a 50 μl total volume, incubated at 42°C for 30 min and immediately placed on ice. An aliquot of 510 μl of freshly prepared sodium bisulphite (final concentration 3.3 M) and 30 μl of freshly prepared 10 mM hydroquinone was added and incubated at 55°C for 16 h. Bisulphite modification of the DNA was followed by two consecutive cleaning and recovering steps. First, samples were desulphonated, purified and eluted using columns and buffers from QIAquick PCR purification kit (Qiagen, Valencia, CA, USA). Secondly, DNA eluted in 50 μl elution buffer (10 mM Tris–Cl, pH 8.5) was denatured with 5 μl of freshly prepared 3 M NaOH, incubated at 37°C for 15 min, neutralized by adding 21 μl of ammonium acetate 7.5 M (final concentration 2 M) and ethanol precipitated. DNA was resuspended in 20–40 μl double distilled water. If DNA concentrations were low or PCR amplification repeatedly failed, bisulphite treatment was performed with only 10–100 ng of DNA using the Imprint DNA Modification kit (Sigma, Canada), optimized for low DNA quantities.

The *DMPK* gene region assessed is detailed in Figure 2. DNA was amplified by hemi-nested PCR approach using two different set of primers, *ctcf1* and *ctcf2* (Supplementary Material, Table S1), to amplify upstream (5′) and downstream (3′) from the repeat, respectively. Primers for bisulphate-modified DNA were designed using Methyl Primer Express Software v1.0 (Applied Biosystems, USA). An aliquot of 1–5 μl of bisulphate-treated DNAs was amplified in 25 or 50 μl PCR reactions with the corresponding primers. Products were resolved by running 2% agarose gel electrophoresis and visualized by ethidium bromide staining. Detection of methylation was performed by direct PCR product sequencing as follows: first, if necessary (i.e. non-specific PCR products appearance), a step of gel extraction (QIAquick PCR purification kit, Qiagen) was performed; secondly, PCR product was cleaned (QIAquick PCR purification kit, Qiagen) and thirdly, samples were sent routinely to the sequencing facility of The Hospital for Sick Children (The Centre for Applied Genomics). Additional quantification of the methylation profiles was performed in samples with only one DM1 allele in the genome (mouse model), by measuring in each CpG site the cytosine height peaks in the chromatogram traces (4Peaks software) (Supplementary Material, Fig. S4A). Broad categories were established to minimize errors introduced by the methylation quantification by direct PCR sequencing: (i) only T peak (no methylation), (ii) double peak, T > C (less than 50% methylation) and (iii) double peak, C > T or only C peak (>50–100% methylation). Cloning of PCR products (TOPO-TA 4.0 cloning kit, Stratagene, USA) and at least 10 clones sequencing were performed in some mouse samples in order to validate the direct PCR sequencing methylation results (Supplementary Material, Fig. S4B). The cloning approach was also used when direct PCR sequencing was impossible due to weak PCR products. All clones and direct PCR chromatogram traces with <95% of non-CpG cytosines converted were discarded by incomplete bisulphite modification. Each sequence was proofread against the electropherogram.

### CpG methylation analysis by *Sac*II methylation-sensitive digestion

*Sac*II methylation-sensitive digestion (unique site 42 bases upstream of the repeat), followed by multiplex PCR (*ctcf-F* and *ctcf-R* encompassing the *Sac*II site and *409* and *ctcf*-R not encompassing the *Sac*II site) (primers in Supplementary Material, Table S1), was performed in methylated samples (detected by bisulphite sequencing) in order to discriminate the presence of methylation from the non-expanded DM1 alleles.

### CpNpG methylation through expanded CTG alleles

To determine the methylation status of non-CpG methylation through the extensively expanded CTG DM1 alleles, we devised a methyl-sensitive restriction digestion assay. Briefly, we revealed digestion by *Fnu*4HI to be sensitive to cytosine methylation at CpNpG sites, and this sensitivity combined with the LNA-Southern detection (46) was used to assess the methylation of highly expanded CTG repeats at the DM1 locus (details of this assay will be outlined elsewhere).

## SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

## ACKNOWLEDGEMENTS

We acknowledge use of tissues (sample ADN1) procured by the National Disease Research Interchange (NRDI) with support from National Institutes of Health (grant no. 5U42RR006042-19). We thank Dr V.L. Funanage for providingDMPK493 and DMRB993 cell lines. We thank Dr M. Baiget for providing the homozygous DM1 samples.

*Conflict of Interest statement*. The authors declare no conflict of interest.

## FUNDING

This work was supported in the Pearson laboratory by the Canadian Institutes of Health Research, the Muscular Dystrophy Association and a post-doctoral fellowship from The Hospital for Sick Children Research Training Centre (A.L.C.). Work in the Thornton laboratory is supported by the University of Rochester Paul Wellstone Muscular Dystrophy Cooperative Research Center with support from the NIH (grant no. U54NS48843) and the Muscular Dystrophy Association and postdoctoral fellowships from the Cell Science Research Foundation and the Uehara Memorial Foundation (M.N.). Work in the Gourdon laboratory is supported by the Institut National de la Santé et de la Recherche Médicale, the Association Française contre les Myopathies, the Agence Nationale de la Recherche, and a PhD grant from the 'Ministère Français de la Recherche et de la Technologie' (S.T.).

## REFERENCES

1. Machuca-Tzili, L., Brook, D. and Hilton-Jones, D. (2005) Clinical and molecular aspects of the myotonic dystrophies: a review. *Muscle Nerve*, **32**, 1–18.
2. Schara, U. and Schoser, B.G. (2006) Myotonic dystrophies type 1 and 2: a summary on current aspects. *Semin. Pediatr. Neurol.*, **13**, 71–79.
3. López Castel, A., Cleary, J.D. and Pearson, C.E. (2010) Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat. Rev. Mol. CellBiol.*, **11**, 165–170.
4. Ranum, L.P. and Cooper, T.A. (2006) RNA-mediated neuromuscular disorders. *Annu. Rev. Neurosci.*, **29**, 259–277.
5. 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.
6. Thornton, C.A., Wymer, J.P., Simmons, Z., McClain, C. and Moxley, R.T. III (1997) Expansion of the myotonic dystrophy CTG repeat reduces expression of the flanking DMAHP gene. *Nat. Genet.*, **16**, 407–409.
7. Korade-Mirnics, Z., Tarleton, J., Servidei, S., Casey, R.R., Gennarelli, M., Pegoraro, E., Angelini, C. and Hoffman, E.P. (1999) Myotonic dystrophy: tissue-specific effect of somatic CTG expansions on allele-specific DMAHP/SIX5 expression. *Hum. Mol. Genet.*, **8**, 1017–1023.
8. 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.
9. 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.
10. Frisch, R., Singleton, K.R., Moses, P.A., Gonzalez, I.L., Carango, P., Marks, H.G. and Funanage, V.L. (2001) Effect of triplet repeat expansion on chromatin structure and expression of DMPK and neighboring genes, SIX5 and DMWD, in myotonic dystrophy. *Mol. Genet. Metab.*, **74**, 281–291.
11. 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.
12. Oude Ophuis, R.J., Mulders, S.A., van Herpen, R.E., van de Vorstenbosch, R., Wieringa, B. and Wansink, D.G. (2009) DMPK protein isoforms are differentially expressed in myogenic and neural cell lineages. *Muscle Nerve*, **40**, 545–555.
13. Alwazzan, M., Hamshere, M.G., Lennon, G.G. and Brook, J.D. (1998) Six transcripts map within 200 kilobases of the myotonic dystrophy expanded repeat. *Mamm. Genome*, **9**, 485–487.
14. 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.
15. Tachi, N., Ohya, K. and Chiba, S. (1999) Expression of the myotonic dystrophy locus-associated homeodomain protein in congenital myotonic dystrophy. *J. Child Neurol.*, **14**, 471–473.
16. Eriksson, M., Ansved, T., Edström, L., Anvret, M. and Carey, N. (1999) Simultaneous analysis of expression of the three myotonic dystrophy locus genes in adult skeletal muscle samples: the CTG expansion correlates inversely with DMPK and 59 expression levels, but not DMAHP levels. *Hum. Mol. Genet.*, **8**, 1053–1060.
17. Eriksson, M., Ansved, T., Edstrom, L., Wells, D.J., Watt, D.J., Anvret, M. and Carey, N. (2000) Independent regulation of the myotonic dystrophy 1 locus genes postnatally and during adult skeletal muscle regeneration. *J. Biol. Chem.*, **275**, 19964–19969.
18. Eriksson, M., Hedberg, B., Carey, N. and Ansved, T. (2001) Decreased DMPK transcript levels in myotonic dystrophy 1 type IIA muscle fibers. *Biochem. Biophys. Res. Commun.*, **286**, 1177–1182.
19. Gennarelli, M., Pavoni, M., Amicucci, P., Angelini, C., Menegazzo, E., Zelano, G., Novelli, G. and Dallapiccola, B. (1999) Reduction of the DM-associated homeo domain protein (DMAHP) mRNA in different brain areas of myotonic dystrophy patients. *Neuromuscul. Disord.*, **9**, 215–219.
20. Westerlaken, J.H., Van der Zee, C.E., Peters, W. and Wieringa, B. (2003) The DMWD protein from the myotonic dystrophy (DM1) gene region is developmentally regulated and is present most prominently in synapse-dense brain areas. *Brain Res.*, **971**, 116–127.
21. Wang, Y.H. (2007) Chromatin structure of repeating CTG/CAG and CGG/CCG sequences in human disease. *Front. Biosci.*, **12**, 4731–4741.
22. Kumari, D. and Usdin, K. (2009) Chromatin remodeling in the noncoding repeat expansion diseases. *J. Biol. Chem.*, **284**, 7413–7417.
23. Johnson, K.J., Boucher, C., King, S.K., Winchester, C.L., Bailey, M.E., Hamilton, G.M. and Carey, N. (1996) Is myotonic dystrophy a single-gene disorder? *Biochem. Soc. Trans.*, **24**, 510–513.
24. Harris, S., Moncrieff, C. and Johnson, K. (1996) Myotonic dystrophy: will the real gene please step forward! *Hum. Mol. Genet.*, **5**, 1417–1423.
25. Groenen, P. and Wieringa, B. (1998) Expanding complexity in myotonic dystrophy. *Bioessays*, **20**, 901–912.
26. Wang, Y.H., Amirhaeri, S., Kang, S., Wells, R.D. and Griffith, J.D. (1994) Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science*, **265**, 669–671.
27. Otten, A.D. and Tapscott, S.J. (1995) Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. *Proc. NatlAcad. Sci. USA*, **92**, 5465–5469.
28. Anvret, M., Ahlberg, G., Grandell, U., Hedberg, B., Johnson, K. and Edström, L. (1993) Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy. *Hum. Mol. Genet.*, **2**, 1397–1400.
29. 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.

30. Thornton, C.A., Johnson, K.J. and Moxley, R.T. (1994) Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. *Ann. Neurol.*, **35**, 104–107.
31. Martorell, L., Martinez, J.M., Carey, N., Johnson, K. and Baiget, M. (1995) Comparison of CTG repeat length expansion and clinical progression of myotonic dystrophy over a five year period. *J. Med. Genet.*, **32**, 593–596.
32. Monckton, D.G., Wong, L.-J.C., 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.
33. Wong, L.J., Ashizawa, T., Monckton, D.G., Caskey, C.T. and Richards, C.S. (1995) Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. *Am. J. Hum. Genet.*, **56**, 114–122.
34. 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.
35. 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.
36. Jansen, G., Willems, P., Coerwinkel, M., Nillesen, W., Smeets, H., Vits, L., Howeler, 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.
37. Wöhrle, D., Kennerknecht, I., Wolf, M., Enders, H., Schwemmle, S. and Steinbach, P. (1995) Heterogeneity of DM kinase repeat expansion in different foetal tissues and further expansion during cell proliferation *in vitro*: evidence for a casual involvement of methyl-directed DNA mismatch repair in triplet repeat stability. *Hum. Mol. Genet.*, **4**, 1147–1153.
38. 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.
39. Ishii, S., Nishio, T., Sunohara, N., Yoshihara, T., Takemura, K., Hikiji, K., Tsujino, S. and Sakuragawa, N. (1996) Small increase in triplet repeat length of cerebellum from patients with myotonic dystrophy. *Hum. Genet.*, **98**, 138–140.
40. Martorell, L., Johnson, K., Boucher, C.A. and Baiget, M. (1997) Somatic instability of the myotonic dystrophy (CTG)n repeat during human foetal development. *Hum. Mol. Genet.*, **6**, 877–880.
41. Nichol, K. and Pearson, C.E. (2002) CpG methylation modifies the genetic stability of cloned repeat sequences. *Genome Res.*, **12**, 1246–1256.
42. Gorbunova, V., Seluanov, A., Mittelman, D. and Wilson, J.H. (2004) Genome-wide demethylation destabilizes CTG.CAG trinucleotide repeats in mammalian cells. *Hum. Mol. Genet.*, **13**, 2979–2989.
43. Dion, V., Lin, Y., Hubert, L., Jr, Waterland, R.A. and Wilson, J.H. (2008) Dnmt1 deficiency promotes CAG repeat expansion in the mouse germline. *Hum. Mol. Genet.*, **17**, 1306–1317.
44. Libby, R.T., Hagerman, K.A., Pineda, V.V., Lau, R., Cho, D.H., Baccam, S.L., Axford, M.M., Cleary, J.D., Moore, J.M., Sopher, B.L. *et al.* (2008) CTCF *cis*-regulates trinucleotide repeat instability in an epigenetic manner: a novel basis for mutational hot spot determination. *PLoS Genet.*, **4**, e1000257.
45. Cleary, J.D., Tomé, S., López Castel, A., Panigrahi, G., 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.
46. Nakamori, M., Sobczak, K., Moxley, R.T. III and Thornton, C.A. (2009) Scaled-down genetic analysis of myotonic dystrophy type 1 and type 2. *Neuromuscul. Disord.*, **19**, 759–757.
47. Razin, A. and Shemer, R. (1995) DNA methylation in early development. *Hum. Mol. Genet.*, **4**, 1751–1755.
48. Shiota, K., Kogo, Y., Ohgane, J., Imamura, T., Urano, A., Nishino, K., Tanaka, S. and Hattori, N. (2002) Epigenetic marks by DNA methylation specific to stem, germ and somatic cells in mice. *Genes Cells*, **7**, 961–969.
49. Song, F., Mahmood, S., Ghosh, S., Liang, P., Smiraglia, D.J., Nagase, H. and Held, W.A. (2009) Tissue specific differentially methylated regions (TDMR): changes in DNA methylation during development. *Genomics*, **93**, 130–139.
50. Nagase, H. and Ghosh, S. (2008) Epigenetics: differential DNA methylation in mammalian somatic tissues. *FEBS J.*, **275**, 1617–1623.
51. Aranda, P., Agirre, X., Ballestar, E., Andreu, E.J., Román-Gómez, J., Prieto, I., Martin-Subero, J.I., Cigudosa, J.C., Siebert, R., Esteller, M. *et al.* (2009) Epigenetic signatures associated with different levels of differentiation potential in human stem cells. *PLoS One*, **4**, e7809.
52. Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M. *et al.* (2010) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, **462**, 315–322.
53. 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.
54. Steinbach, P., Glaser, 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.
55. Lee, J., Jang, S.J., Benoit, N., Hoque, M.O., Califano, J.A., Trink, B., Sidransky, D., Mao, L. and Moon, C. (2010) Presence of 5-methylcytosine in CpNpG trinucleotides in the human genome. *Genomics*, **96**, 67–72.
56. Ansved, T., Edström, L., Grandell, U., Hedberg, B. and Anvret, M. (1997) Variation of CTG-repeat number of the DMPK gene in muscle tissue. *Neuromuscul. Disord.*, **7**, 152–155.
57. Furutama, D., Negoro, N., Terasaki, F., Tsuji-Matsuyama, K., Sakai, R., Maeda, T., Tanaka, T., Hoshiga, M., Ishihara, T., Ohsawa, N. and Hanafusa, T. (2010) Possible *de novo* CTG repeat expansion in the DMPK gene of a patient with cardiomyopathy. *J. Clin. Neurosci.*, **17**, 408–409.
58. Krahe, R., Ashizawa, T., Abbruzzese, C., Roeder, E., Carango, P., Giacanelli, M., Funanage, V.L. and Siciliano, M.J. (1995) Effect of myotonic dystrophy trinucleotide repeat expansion on DMPK transcription and processing. *Genomics*, **28**, 1–14.
59. Cobo, A., Martinez, J.M., Martorell, L., Baiget, M. and Johnson, K. (1993) Molecular diagnosis of homozygous myotonic dystrophy in two asymptomatic sisters. *Hum. Mol. Genet.*, **2**, 711–715.
60. Martorell, L., Illa, I., Rosell, J., Benitez, J., Sedano, M.J. and Baiget, M. (1996) Homozygous myotonic dystrophy: clinical and molecular studies of three unrelated cases. *Med. Genet.*, **33**, 783–785.
61. Naumann, A., Hochstein, N., Weber, S., Fanning, E. and Doerfler, W. (2009) A distinct DNA-methylation boundary in the 5'-upstream sequence of the FMR1 promoter binds nuclear proteins and is lost in Fragile X syndrome. *Am. J. Hum. Genet.*, **85**, 606–616.
62. Hansen, R.S., Gartler, S.M., Scott, C.R., Chen, S.H. and Laird, C.D. (1992) Methylation analysis of CGG sites in the CpG island of the human FMR1 gene. *Hum. Mol. Genet.*, **1**, 571–578.
63. Pearson, C.E., Nichol Edamura, K. and Cleary, J.D. (2005) Repeat instability: mechanisms of dynamic mutations. *Nat. Rev. Genet.*, **6**, 729–742.
64. Lewis, A. and Murrell, A. (2004) Genomic imprinting: CTCF protects the boundaries. *Curr. Biol.*, **14**, 284–286.
65. Filippova, G.N., Cheng, M.K., Moore, J.M., Truong, J.P., Hu, Y.J., Nguyen, D.K., Tsuchiya, K.D. and Disteche, C.M. (2005) Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. *Dev. Cell*, **8**, 31–42.
66. Hancock, A.L., Brown, K.W., Moorwood, K., Moon, H., Holmgren, C., Mardikar, S.H., Dallosso, A.R., Klenova, E., Loukinov, D., Ohlsson, R. *et al.* (2007) A CTCF-binding silencer regulates the imprinted genes AWT1 and WT1-AS and exhibits sequential epigenetic defects during Wilms' tumourigenesis. *Hum. Mol. Genet.*, **16**, 343–354.
67. Nagarajan, R.P., Patzel, K.A., Martin, M., Yasui, D.H., Swanberg, S.E., Hertz-Picciotto, I., Hansen, R.L., Van de Water, J., Pessah, I.N., Jiang, R. *et al.* (2008) MECP2 promoter methylation and X chromosome inactivation in autism. *Autism Res.*, **1**, 169–178.
68. Ohlsson, R., Bartkuhn, M. and Renkawitz, R. (2010) CTCF shapes chromatin by multiple mechanisms: the impact of 20 years of CTCF research on understanding the workings of chromatin. *Chromosoma*, **119**, 351–360.

69. Goto, Y. and Kimura, H. (2009) Inactive X chromosome-specific histone H3 modifications and CpG hypomethylation flank a chromatin boundary between an X-inactivated and an escape gene. *Nucleic Acids Res.*, **37**, 7416–7428.
70. Saveliev, A., Everett, C., Sharpe, T., Webster, Z. and Festenstein, R. (2003) DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing. *Nature*, **422**, 909–913.
71. Hsieh, C.L. (2005) The novo methylation activity of Dnmt3a is distinctly different than that of Dnmt1. *BMC Biochem.* 6, 6.
72. Smith, S.S., Laayoun, A., Lingeman, R.G., Baker, D.J. and Riley, J. (1994) Hypermethylation of telomere-like foldbacks at codon 12 of the human c-Ha-ras gene and the trinucleotide repeat of the FMR-1 gene of fragile X. *J. Mol. Biol.*, **243**, 143–151.
73. Smith, S.S. and Baker, D.J. (1997) Stalling of human methyltransferase at single-strand conformers from the Huntington's locus. *Biochem. Biophys. Res. Commun.*, **234**, 73–78.
74. 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*, **149**, 1365–1374.
75. Braida, C., Stefanatos, R.K., Adam, B., Mahajan, N., Smeets, H.J., Niel, F., Goizet, C., Arveiler, B., Koenig, M., Lagier-Tourenne, C. *et al.* (2010) Variant CCG and GGC repeats within the CTG expansion dramatically modify mutational dynamics and likely contribute toward unusual symptoms in some myotonic dystrophy type 1 patients. *Hum. Mol. Genet.*, **19**, 1399–1412.
76. Devys, D., Biancalana, V., Rousseau, F., Boué, J., Mandel, J.L. and Oberlé, I. (1992) Analysis of full fragile X mutations in foetal tissues and monozygotic twins indicate that abnormal methylation and somatic heterogeneity are established early in development. *Am. J. Med. Genet.*, **43**, 208–216.
77. Tassone, F., Hagerman, R.J., Gane, L.W. and Taylor, A.K. (1999) Strong similarities of the FMR1 mutation in multiple tissues: postmortem studies of a male with a full mutation and a male carrier of a premutation. *Am. J. Med. Genet.*, **84**, 240–244.
78. Reyniers, E., Martin, J.J., Cras, P., Van Marck, E., Handig, I., Jorens, H.Z., Oostra, B.A., Kooy, R.F. and Willems, P.J. (1999) Postmortem examination of two fragile X brothers with an FMR1 full mutation. *Am. J. Med. Genet.*, **84**, 245–249.
79. Christensen, B.C., Houseman, E.A., Marsit, C.J., Zheng, S., Wrensch, M.R., Wiemels, J.L., Nelson, H.H., Karagas, M.R., Padbury, J.F., Bueno, R. *et al.* (2009) Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. *PLoS Genet.*, **5**, e1000602.
80. Cleary, J.D. and Pearson, C.E. (2003) The contribution of *cis*-elements to disease-associated repeat instability: clinical and experimental evidence. *Cytogenet. Genome Res.*, **100**, 25–55.
81. Telenius, H., Kremer, B., Goldberg, Y.P., Theilmann, J., Andrew, S.E., Zeisler, J., Adam, S., Greenberg, C., Ives, E.J., Clarke, L.A. *et al.* (1994) Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. *Nat Genet.*, **6**, 409–414.
82. Aoki, M., Abe, K., Tobita, M., Kameya, T., Watanabe, M. and Itoyama, Y. (1996) Reduction of CAG expansions in cerebellar cortex and spinal cord of DRPLA. *Clin. Genet.*, **50**, 199–201.
83. Tanaka, F., Sobue, G., Doyu, M., Ito, Y., Yamamoto, M., Shimada, N., Yamamoto, K., Riku, S., Hshizume, Y. and Mitsuma, T. (1996) Differential pattern in tissue-specific somatic mosaicism of expanded CAG trinucleotide repeats in dentatorubral-pallidoluysian atrophy, Machado–Joseph disease, and X-linked recessive spinal and bulbar muscular atrophy. *J. Neurol. Sci.*, **135**, 43–50.
84. De Rooij, K.E., De Koning Gans, P.A., Roos, R.A., Van Ommen, G.J. and Den Dunnen, J.T. (1995) Somatic expansion of the (CAG)n repeat in Huntington disease brains. *Hum Genet.*, **95**, 270–274.
85. Watase, K., Venken, K.J., Sun, Y., Orr, H.T. and Zoghbi, H.Y. (2003) Regional differences of somatic CAG repeat instability do not account for selective neuronal vulnerability in a knock-in mouse model of SCA1. *Hum. Mol. Genet.*, **12**, 2789–2795.
86. Roth, R.B., Hevezi, P., Lee, J., Willhite, D., Lechner, S.M., Foster, A.C. and Zlotnik, A. (2006) Gene expression analyses reveal molecular relationships among 20 regions of the human CNS. *Neurogenetics*, **7**, 67–80.
87. De Bustos, C., Ramos, E., Young, J.M., Tran, R.K., Menzel, U., Langford, C.F., Eichler, E.E., Hsu, L., Henikoff, S., Dumanski, J.P. *et al.*. (2009) Tissue-specific variation in DNA methylation levels along human chromosome 1. *Epigenet.Chromatin*, **2**, 7.
88. Ladd-Acosta, C., Pevsner, J., Sabunciyan, S., Yolken, R.H., Webster, M.J., Dinkins, T., Callinan, P.A., Fan, J.B., Potash, J.B. and Feinberg, A.P. (2007) DNA methylation signatures within the human brain. *Am. J. Hum. Genet.*, **81**, 1304–1315.
89. Greene, E., Mahishi, L., Entezam, A., Kumari, D. and Usdin, K. (2007) Repeat-induced epigenetic changes in intron 1 of the frataxin gene and its consequences in Friedreich ataxia. *Nucleic Acids Res.*, **35**, 3383–3390.
90. Herman, D., Jenssen, K., Burnett, R., Soragni, E., Perlman, S.L. and Gottesfeld, J.M. (2006) Histone deacetylase inhibitors reverse gene silencing in Friedreich's ataxia. *Nat. Chem. Biol.*, **2**, 551–558.
91. Al-Mahdawi, S., Pinto, R.M., Ismail, O., Varshney, D., Lymperi, S., Sandi, C., Trabzuni, D. and Pook, M. (2008) The Friedreich ataxia GAA repeat expansion mutation induces comparable epigenetic changes in human and transgenic mouse brain and heart tissues. *Hum. Mol. Genet.*, **17**, 735–746.
92. Flanagan, J.M., Popendikyte, V., Pozdniakovaite, N., Sobolev, M., Assadzadeh, A., Schumacher, A., Zangeneh, M., Lau, L., Virtanen, C., Wang, S.C. and Petronis, A. (2006) Intra- and interindividual epigenetic variation in human germ cells. *Am. J. Hum. Genet.*, **79**, 67–84.
93. Park, C.W., Chen, Z., Kren, B.T. and Steer, C.J. (2004) Double-stranded siRNA targeted to the huntingtin gene does not induce DNA methylation. *Biochem. Biophys. Res. Commun.*, **323**, 275–280.
94. Theilmann, J.L., Robbins, C.A. and Hayden, M.R. (1989) Methylation at the D4S95 locus and predictive testing. *Am. J. Hum. Genet.*, **45**, 477–479.
95. Pritchard, C.A., Cox, D.R. and Myers, R.M. (1989) Methylation at the Huntington disease-linked D4S95 locus. *Am. J. Hum. Genet.*, **45**, 335–336.
96. Reik, W., Maher, E.R., Morrison, P.J., Harding, A.E. and Simpson, S.A. (1993) Age at onset in Huntington's disease and methylation at D4S95. *J. Med. Genet.*, **30**, 185–188.
97. 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. NatlAcad. Sci. USA*, **94**, 7394–7399.
98. Gourdon, G., Dessen, P., Lia, A.S., Junien, C. and Hofmann-Radvanyi, H. (1997) Intriguing association between disease associated unstable trinucleotide repeat and CpG island. *Ann. Genet.*, **40**, 73–77.
99. Seznec, H., Lia-Baldini, A.S., Duros, C., Fouquet, C., Lacroix, C., Hofmann-Radvanyi, H., Junien, C. and Gourdon, G. (2000) Transgenic mice carrying large human genomic sequences with expanded CTG repeat mimic closely the DM CTG repeat intergenerational and somatic instability. *Hum. Mol. Genet.*, **9**, 1185–1194.
100. Lia, A.S., Seznec, H., Hofmann-Radvanyi, H., Radvanyi, F., Duros, C., Saquet, C., Blanche, M., Junien, C. and Gourdon, G. (1998) Somatic instability of the CTG repeat in mice transgenic for the myotonic dystrophy region is age dependent but not correlated to the relative inter-tissue transcription levels and proliferative capacities. *Hum. Mol. Genet.*, **7**, 1285–1291.
101. Seznec, H., Agbulut, O., Sergeant, N., Savouret, C., Ghestem, A., Tabti, N., Willer, J.C., Ourth, L., Duros, C., Brisson, E. *et al.* (2001) Mice transgenic for the human myotonic dystrophy region with expanded CTG repeats display muscular and brain abnormalities. *Hum. Mol. Genet.*, **10**, 2717–2726.
102. Guiraud-Dogan, C., Huguet, A., Gomes-Pereira, M., Brisson, E., Bassez, G., Junien, C. and Gourdon, G. (2007) DM1 CTG expansions affect insulin receptor isoforms expression in various tissues of transgenic mice. *Biochim. Biophys. Acta*, **1772**, 1183–1191.
103. Tomé, S., Holt, I., Edelmann, W., Morris, G.E., Munnich, A., Pearson, C.E. and Gourdon, G. (2009) MSH2 ATPase domain mutation affects CTG*CAG repeat instability in transgenic mice. *PLoS Genet.*, **5**, e1000482.
104. Karolchik, D., Baertsch, R., Diekhans, M., Furey, T.S., Hinrichs, A., Lu, Y.T., Roskin, K.M., Schwartz, M., Sugnet, C.W., Thomas, D.J. *et al.* (2003) The UCSC Genome Browser Database. *Nucleic Acids Res.*, **31**, 51–54.