with myotonia in DM model mice (Mankodi *et al.* 2002). The aberrant splicing of multiple genes is likely to cause multisystemic symptoms in DM.

Two families of proteins, 'muscleblind-like' (MBNL) and 'CUG-BP and ETR-3-like factor' (CELF), are major candidates in the pathogenesis of DM. MBNL is a family of CCCH-type zinc finger-containing RNA-binding proteins, and three isoforms are known in human: *MBNL1, MBNL2/MBLL* and *MBNL3/MBXL.* Our previous study showed that all three MBNL proteins bound to CHG/CHHG repeat RNA (where H indicates A, C and U) (Kino *et al.* 2004), and MBNL1 and MBNL2 have been shown to be colocalized with CUG/CCUG inclusions in DM1 and DM2 cells (Miller *et al.* 2000; Fardaei *et al.* 2002; Holt *et al.* 2009), giving rise to an idea that MBNL function may be disrupted in DM cells. Actually, knockout of either *Mbnl1* or *Mbnl2* caused some DM-like defects in mice, such as abnormal splicing, myotonia and histological abnormalities in skeletal muscle (Kanadia *et al.* 2003; Hao *et al.* 2008), which suggests that the loss of MBNL function because of the sequestration of CUG/CCUG repeat RNA leads to the characteristic features of DM.

CUG-BP and ETR-3-like factor proteins are another family of proteins involved in the pathogenesis of DM1. CELF proteins are also a family of highly conserved RNA-binding proteins, and six genes belong to this family: CELF1/CUGBP1/CUG-BP/BRUNOL2, CELF2/ETR-3/CUGBP2/BRUNOL3, CELF3/TNRC4/BRUNOL1, CELF4/BRUNOL4, CELF5/BRUNOL5 and CELF6/BRUNOL6. Several lines of evidence has shown that CELF1 and other members of CELF family play an important role in the pathogenesis of DM. Remarkably, CELF1 is activated in DM1 cells via PKC-mediated phosphorylation (Kuyumcu-Martinez *et al.* 2007), and transgenic mice expressing human CELF1 recapitulate some of DM-like abnormalities, especially histological muscle impairment (Timchenko *et al.* 2004; Ho *et al.* 2005).

Several studies have demonstrated that both MBNL and CELF proteins directly regulate the alternative splicing dysregulated in patients with DM. Interestingly, CELF1 and MBNL1 regulate antagonistically the alternative splicing of the cardiac troponin T exon 5 and insulin receptor exon 11 (Philips *et al.* 1998; Savkur *et al.* 2001; Ho *et al.* 2004), suggesting that both the elevation of CELF1 and the loss of MBNL1 function may be involved in the development of DM. But, notably, DNA microarray analysis revealed that HSA^{LR} mice and *Mbnl1*-knockout mice shared most of their splicing abnormalities: 128 of 156 abnormal splicing events in HSA^{LR} mice also found in *Mbnl1*-knockouts (Du *et al.* 2010). Thus, the extent to which each of MBNL and CELF family proteins can account for splicing abnormalities and the pathogenesis of DM remains unclear.

Although splicing defects in multiple genes likely cause the multisystemic symptoms in patients with DM, no abnormal splicing event has been found to explain the development of defects in DM other than that in *CLCN1*, whose abnormal splicing may cause myotonia (Lueck *et al.* 2007). Given that the MBNL and CELF families are involved in the global shift in alternative splicing that occurs during the development of skeletal and cardiac muscle (Lin *et al.* 2006; Kalsotra *et al.* 2008), additional, unknown splicing abnormalities likely exist in DM muscle.

Searches for abnormal splicing regulation on a genome-wide scale have been conducted in several lines of model mice (Kalsotra *et al.* 2008; Du *et al.* 2010), but such comprehensive studies have rarely been performed on human patients with DM. Many efforts have been made to identify alternative splicing events that are misregulated in DM, and most studies have been based on the physiological complications of patients with DM; for example, the aberrant splicing of *CLCN1, IR*, sarcoplasmic/endoplasmic reticulum calcium ATPase 1 (*SERCA1*) and ryanodine receptor 1 (*RYR1*) was investigated because patients with DM show myotonia, insulin resistance and increased muscle Ca^{2+} levels (Savkur *et al.* 2001; Charlet *et al.* 2002; Kimura *et al.* 2005). Thus, we investigated the comprehensive splicing profile of muscle tissue of individuals with DM1 using a splicing-sensitive microarray to identify novel aberrant splicing events in individuals with DM1. Specifically, we used the Affymetrix (Santa Clara, CA, USA) GeneChip Human Exon 1.0 ST Array. This array is a powerful tool for detecting cassette exons and mutually exclusive exons; with approximately four probes per exon, this microarray can analyze exon-level expression on a whole-genome scale.

In a comprehensive search for aberrant splicing events in patients with DM1, we found a novel splicing abnormality in *MYOM1* exon 17a. MYOM1 is a constituent of the sarcomeric M band, suggesting its involvement in muscle impairment in patients with DM. A subsequent splicing assay using a *MYOM1* minigene system revealed that MBNL and CELF family proteins function as *trans*-acting factors in the alternative splicing of *MYOM1* exon 17a. Our results

suggest that the MBNL family is a key regulator of *MYOM1* alternative splicing.

Results

Microarray-based comprehensive study of splicing in DM1 muscle detected novel splicing abnormalities

To investigate novel splicing defects in patients with DM1, we analyzed the splicing profiles of skeletal muscle derived from three patients with DM1 and four non-DM controls using the Human Exon 1.0 ST Array. In total, 289 probe sets displayed a significant difference (splicing index [SI] > 1 or <−1; $P < 0.05$) between the two groups, and 39 of them were designed against exons that are known or predicted to be alternatively spliced (Table S1 in Supporting Information).

Inclusion of MYOM1 exon 17a was increased in DM1 muscle

The greatest change was recorded at exon 17a in *MYOM1* (SI = 2.47), encoding a component of the sarcomeric M band. We thus focused on this exon as a candidate whose the abnormal splicing of which might cause the muscular defects in DM and analyzed the splicing regulation of *MYOM1* exon 17a.

We confirmed the aberrant splicing of *MYOM1* exon 17a in DM1 skeletal muscle by RT-PCR (Fig. 1). Because the alternative splicing of exon 17a generates two splicing isoforms in murine embryonic heart (Agarkova *et al.* 2000), we anticipated that two isoforms (A and D in Fig. 1A) would be expressed from *MYOM1* in patients with DM1. The products amplified from isoforms A and D were expected to be 479 and 191 nucleotides in length, respectively, using the primers designed for the flanking exons (17 and 18). As in mouse skeletal muscle, isoform D was exclusively expressed in non-DM muscle (Fig. 1B). Isoform D was also dominant in DM1 muscle, but the expression of isoform A was significantly increased, in agreement with our exon array analysis. Furthermore, two additional products were amplified from the DM1 samples (Fig. 1B). Sequencing revealed that these products were isoforms of *MYOM1* containing truncated forms of exon 17a: isoform B lacked 36 nucleotides at the 3′ end of exon 17a, whereas isoform C lacked 100 nucleotides. As shown in Fig. 1D, the alternative 5′ splice sites of exon 17a used in isoform B and C are consistent with the consensus sequence of 5′ splice site.

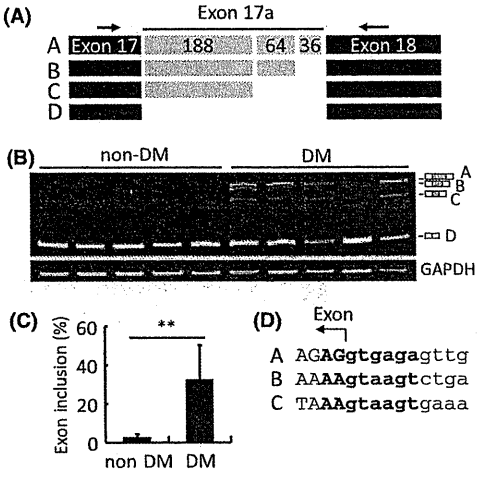

Figure 1 The inclusion of *MYOM1* exon 17a was increased in DM. (A) Schematic diagram showing the products of the alternative splicing of *MYOM1*. The number of nucleotides of exon 17a is shown: exon 17a has two internal 5′ splice sites, generating four splicing isoforms. The names of the isoforms are given on the left side of the panel. Arrows represent the primer pairs used for RT-PCR. (B) Total RNA isolated from homogenized skeletal muscle tissue was subjected to RT-PCR and electrophoretically resolved on an 8% polyacrylamide gel. The identity of the bands is indicated on the right side of the panel. GAPDH was amplified as a loading control. (C) The histogram illustrates the inclusion of exon 17a (mean ± SD of the five individuals). The percentage of inclusions of exon 17a was calculated as the ratio of the isoforms containing full length or part of exon 17a (isoforms A, B and C) to the total spliced products. The statistical significance of the results is indicated by $**P < 0.01$ (Student's *t*-test). All bands of interest were gel-isolated, cloned, and confirmed by sequencing. (D) Alternative 5′ splice sites of exon 17a.

Both MBNL1 and CELF1 suppressed exon 17a inclusion

Next, we investigated whether MBNL and CELF family proteins regulate the splicing of *MYOM1* exon 17a using a cellular splicing assay. Because the endogenous myomesin gene was not expressed in any of the cell lines we tested, including HEK293T (see Fig. 2B, no TF), HeLa and C2C12 cells (data not shown), we constructed pMYOM1, a minigene that covered exons 17–18 of human *MYOM1* and expressed it in HEK293T cells (Fig. 2A,B). When transfected into HEK293T cells, the construct produced the same four isoforms as in DM1 muscle (Fig. 2B, vector). In HEK293T cells, 44% of the spliced product contained full length or parts of exon 17a. We then cotransfected pMYOM1 with MBNL

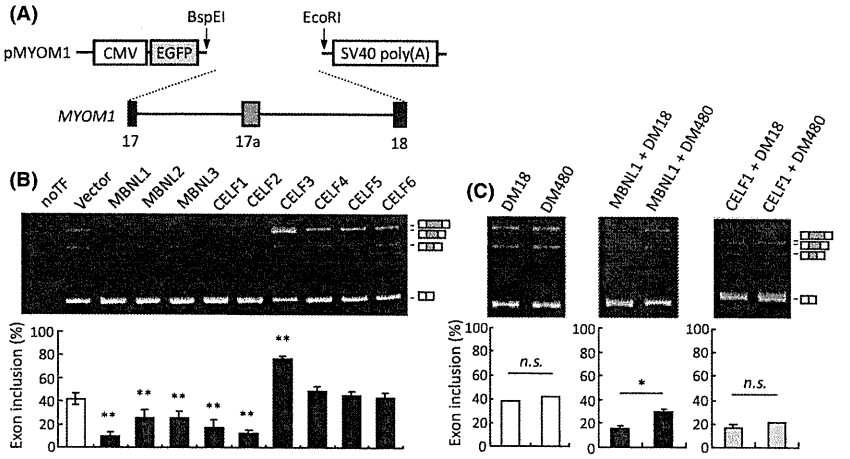

Figure 2 CUG-BP and ETR-3-like factor and MBNL family proteins regulate *MYOM1* splicing. (A) Schematic diagram of pMYOM1 minigene. Boxes and horizontal lines represent exons and introns, respectively. The genomic region from exon 17–18 was cloned into pEGFP-C1. (B) HEK293T cells were cotransfected with the pMYOM1, and the expression vector of each protein indicated. NoTF indicates cells were transfected with pEGFP-C1 vector instead of *MYOM1* minigene. Total RNA from the cells was subjected to RT-PCR and electrophoretically resolved on an 8% polyacrylamide gel. The upper panel shows representative results of our splicing assay. All three MBNL proteins, CELF1 and CELF2 significantly repressed the inclusion of exon 17a, whereas CELF3 promoted its inclusion. The percentage of inclusion of exon 17a was calculated and shown in a histogram as in Fig. 1B (n = 4). The statistical significance of the results is indicated by $^{**}P < 0.01$ (Dunnett's test). (C) The splicing regulation of MBNL1 was prevented by long CUG repeat. DM18 and DM480 indicate transfection of dystrophia myotonica-protein kinase 3′ UTR containing CUG repeat (18 and interrupted 480 repeat, respectively). The splicing pattern of pMYOM1 did not change when the minigene was cotransfected with either DM18 or DM480 in a ratio of 1:9 in HEK293T cells (left); however, when pMYOM1 was cotransfected with CUG repeat and MBNL1 in a ratio of 1:10:15, the longer CUG repeat increased the inclusion of exon 17a (middle). The splicing pattern did not change when pMYOM1 was cotransfected with CUG repeat and CELF1 (right). The inclusion rate of exon 17a was calculated and shown in a histogram as in Fig. 1B (n = 3). The statistical significance of the results is indicated by $^{*}P < 0.05$ (Student's t-test).

and CELF family proteins and examined the patterns of minigene splicing by RT-PCR (Fig. 2B). Consistent with the current model of DM pathology, all three MBNL proteins decreased the inclusion rate of exon 17a. MBNL1 changed the splicing most, with 9.8% of the spliced product containing exon 17a. Regarding the CELF proteins, CELF1 and CELF2 promoted the exclusion of exon 17a, whereas CELF3 strongly increased its inclusion. No other CELF protein had a significant effect on exon 17a splicing. The comparable expressions of the RNA proteins were confirmed in the previous study (Kino *et al.* 2009). We conducted the minigene assay on MBNL1 and CELF1 in two other cell lines: C2C12 and HeLa (Fig. S1 in Supporting Information). The effects of MBNL1 and CELF1 in the two cell lines were the same as in HEK293T; both proteins repressed the inclusion of exon 17a of p*MYOM1*, indicating that the phenomenon was not specific to HEK293T cells.

Because CUG repeat–mediated downregulation of MBNL1 function is thought to be a major cause of aberrant splicing in DM1, we tested the effect of CUG repeat RNA on the *MYOM1* alternative splicing (Fig. 2C). Unexpectedly, there was no significant difference between DM18 and DM480, which express 3′ UTR of *DMPK* containing $(CUG)_{18}$ and interrupted $(CUG)_{480}$, respectively. Because this may be due to the relatively low endogenous expression level of MBNL1 in HEK293T cells (Sen *et al.* 2010), we examined the effect of CUG repeat RNA under cotransfection of MBNL1. In the presence of abundant MBNL1, DM480 increased the inclusion of exon 17a. It is unlikely that the effect of DM480 came out as a result of MBNL1-mediated suppression of exon 17a inclusion because DM480 did not show any significant effect under CELF1 over-expression. Thus, CUG repeat should promote the exon inclusion by preventing the function of MBNL1.

Region around exon 17a is sufficient for MBNL1 and CELF1 regulation

Surprisingly, MBNL1 and CELF1 regulate the alternative splicing of *MYOM1* to the same direction. To

determine whether MBNL1 and CELF1 suppress the inclusion of exon 17a by the same mechanism, we explored the regions required for the splicing regulation of the two RNA-binding proteins. First, we examined the contribution of the intronic sequences using intronic deletion mutant minigenes (Δ1–Δ6, Fig. 3A). Compared with pMYOM1 minigene, some mutants altered their basal splicing patterns in HEK293T cells (Fig. 3B); for example, Δ5 mutant showed the increased inclusion of exon 17a with variant C being predominant, suggesting that the region 100–500 bp downstream of exon 17a might be important for the selection of 5′ splice site of intron 17a (data not shown). As for the regulation of MBNL1 and CELF1, any deletion mutants we tested were responsive to both proteins (Fig. 3B). Δ3–Δ6 mutants had only a 100-bp intronic fragment left at the most deleted end of the intron. Thus, for the responsiveness to MBNL1 and CELF1, at least one regulatory element should lie in the exons or their flanking 100-bp intronic regions.

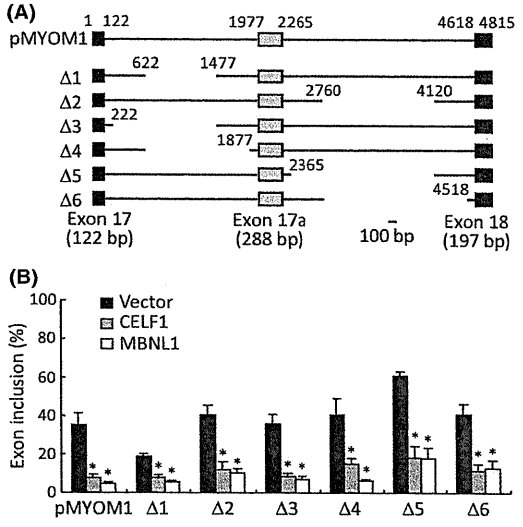

Figure 3 Intronic deletion analysis of pMYOM1. (a) A schematic diagram showing deletion mutants. Deletion series of pMYOM1 were generated by PCR-mediated mutagenesis. The numbers indicate the positions of nucleotides at the termini of exons and the junction of deleted regions. (b) Splicing analysis of pMYOM1 deletion mutants in HEK293T cells. Deletion mutants Δ1–Δ6 were cotransfected with an empty vector, MBNL1, or CELF1 in HEK293T cells, and the inclusion of exon 17a was analyzed by RT-PCR. The percentage of inclusion of exon 17a was calculated and shown in a histogram as in Fig. 1B ($n = 4$). The statistical significance of the results is indicated by $*P < 0.05$ (Dunnett's test against vector).

We then investigated the region around exon 17a using the chimeric minigene system (Kino *et al.* 2009). In this system, we used a vector that carried a sequence covering exon 1–2 of mouse tropomyosin 2 (*Tpm2*), a gene distinct from *MYOM1* (Fig. 4A). We could insert an arbitrary exon together with its flanking introns into intron 1 of the vector; for example, we inserted *Tpm2* exon 9 as a reference (Fig. 4A, *Tpm2*-ex9). When *Tpm2*-ex9 minigene was expressed in HEK293T cells, exon 9 in intron 1 was successfully recognized as an exon, and MBNL1 hardly changed the splicing pattern of *Tpm2*-ex9 (Fig. 4B). Then, we examined a series of *Tpm2*-based chimeric minigenes that contained *MYOM1* exon 17a and its flanking introns or a hybrid of the regions around *Tpm2* exon 9 and *MYOM1* exon 17a (Fig. 4A). Like *Tpm2*-ex9 minigene, the inserted exons of these chimeric minigenes were alternatively spliced. Among the minigenes, *Tpm2*-ex17a, 17a-9 and 5′ex17a were more responsive to MBNL1 than *Tpm2*-ex9 (Fig. 4B). Because these minigenes shared 158 bp of the 5′ region of *MYOM1* exon 17a, we substituted this segment in *Tpm2*-ex17a minigene with a corresponding one of *Tpm2* exon 9. The resulting minigene, Δ5′ex17a, little changed its splicing pattern when co-expressed with MBNL1 (Fig. 4C). These data suggest that a *cis*-element for MBNL1 may be located in the former half of exon 17a.

We compared the effect of CELF1 on *Tpm2*-ex17a and *Tpm2*-ex9 minigenes (Fig. 4D). This analysis revealed that *Tpm2*-ex17a, but not *Tpm2*-ex9, was responsive to CELF1, suggesting that the regulatory element(s) for CELF1 may be in the region around *MYOM1* exon 17a.

Mutation in MBNL1-recognition motifs extinguished the responsiveness to MBNL1

A recent report has described YGCY sequence as a MBNL1-recognition motif (Goers *et al.* 2010), which is consistent with our previous study that revealed MBNL1 bound to CHHG repeat RNA (where H indicates A, U and C) (Kino *et al.* 2004; Goers *et al.* 2010). Because several studies have shown that MBNL1 requires its recognition motif(s) in the target sequence to regulate splicing (Ho *et al.* 2004; Hino *et al.* 2007), we searched YGCY motifs in the MBNL1-responsive region of exon 17a and found four motifs with two UGCU sequences arranged in tandem (Fig. 5A). To investigate whether MBNL1 regulates the splicing of exon 17a via these motifs, we introduced mutations into each YGCY sequence. We

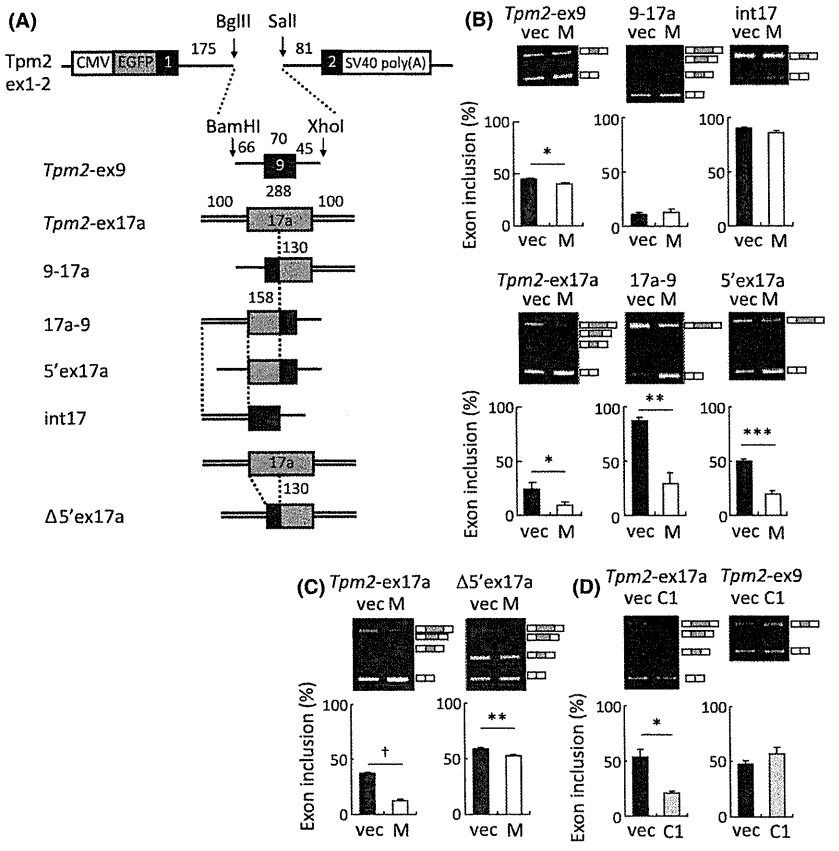

Figure 4 Splicing regulation in chimeric minigenes. (A) Structure of the *Tpm2*-based chimeric minigene. Fragments of *Tpm2* covering exons 1–2 were inserted downstream of EGFP. Test exons together with their flanking regions were inserted into intron 1 of *Tpm2*. Intronic fragments derived from *MYOM1* are indicated by double lines, whereas those derived from *Tpm2* (regions flanking exon 9) are indicated by thin lines. Exonic sequences of *MYOM1* exon 17a and *Tpm2* exon 9 are indicated by grey and black boxes, respectively. The numbers indicate the nucleotide length of each exonic and intronic sequences. (B) Splicing assay results using *Tpm2*-based chimeric minigenes in HEK293T cells. Upper bands correspond to the spliced products containing an exon inserted between *Tpm2* exon 1 and 2. 'vec' and 'M' indicate empty vector and MBNL1, respectively. Compared with *Tpm2*-ex9, *Tpm2*-ex17a, 17a-9 and 5'ex17a minigenes exhibited evident responses to MBNL1. The percentage of inclusion of exon 17a was calculated and shown in a histogram as in Fig. 1B ($n = 3$). The statistical significance of the results is indicated by $\star P < 0.05$, $\star\star P < 0.01$, $\star\star\star P < 0.001$ and $\dagger P < 0.0001$ (Student's *t*-test). (C) Splicing regulation of chimeric minigenes lacking the former half of exon 17a. Results of the splicing assay are shown as in B. The structures of minigenes are shown in A. (D) Sequence around exon 17a is sufficient for CELF1-mediated exon 17a exclusion. *Tpm2*-ex17a and *Tpm2*-ex9 minigenes were tested for their responsiveness to CELF1 (C1). The splicing assay was performed as in B, except that CELF1 was used in place of MBNL1. The responsibility to CELF1 was increased in the presence of exon 17a and its flanking introns.

mutated an additional UGCU motif in intron 17 because MBNL1 may require two or more YGCY motifs to regulate splicing by inducing and stabilizing a hairpin structure in a target RNA (Warf *et al.* 2009). When we expressed each of the mutated minigenes with or without MBNL1 in HEK293T cells, mut1 minigene showed a dramatic decrease in the responsiveness to MBNL1, whereas the other mutations did not change their splicing patterns (Fig. 5B).

CELF1 was also cotransfected with these mutated series of minigenes to discriminate the regulatory mechanism of CELF1 from that of MBNL1 (Fig. 5C). All the point-mutated minigenes retained the responsiveness to CELF1 in a comparable level (data not shown). Importantly, CELF1 decreased the inclusion of exon 17a in mut1 minigene, which lacked a *cis*-element for MBNL1. Therefore, CELF1 regulated the alternative splicing of *MYOM1* exon 17a in a distinct pathway from that of MBNL1.

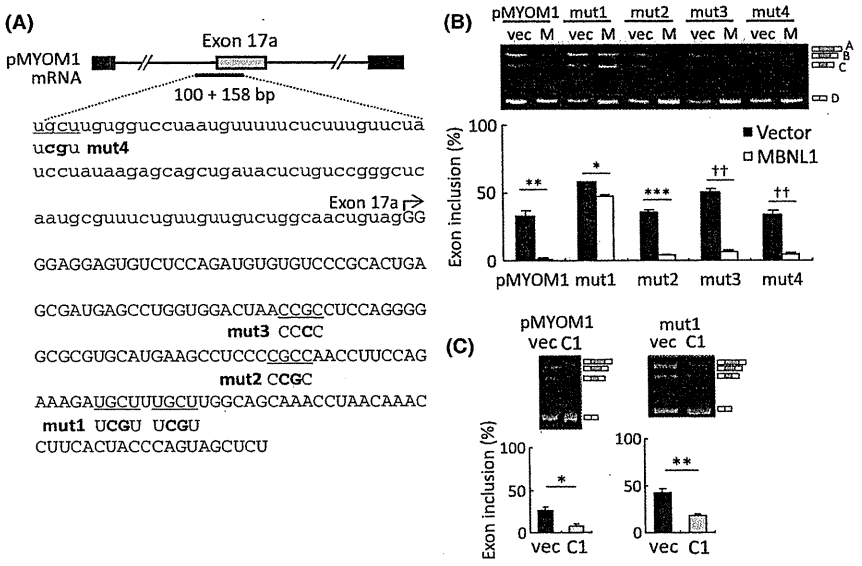

Figure 5 Analysis of MBNL1-recognition motifs in 5′ region of exon 17a. (A) Sequence of a 158-bp portion of 5′ region of exon 17a and a 100-bp portion of its upstream intron is shown. MBNL1-recognized YGCY motifs are underlined, and sequence of a mutant minigene is shown in boldface below each motifs. (B) Result of splicing assay with mutant minigenes in HEK293T cells. The 'vec' and 'M' indicate empty vector and MBNL1, respectively. The mut1 minigene showed remarkable decrease in responsiveness to MBNL1. The percentage of inclusion of exon 17a was calculated and shown in a histogram as in Fig. 1B (n = 3). (C) Same as (B) except C1 indicates CELF1. The mut1 minigene showed responsiveness to CELF1. The statistical significance of the results is indicated by $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $††P < 0.00001$ (Student's t-test).

Discussion

In this study, using splicing-sensitive microarrays, we investigated the comprehensive splicing profiles of patients with DM1 and non-DM individuals, and identified the abnormal inclusion of *MYOM1* exon 17a as a novel splicing defect in DM1. The protein encoded by *MYOM1* is a constituent of the sarcomeric M band, where it anchors the giant protein titin and maintains the thick filament lattice (Lange *et al.* 2005a,b). Like other myofibrillar proteins, MYOM1 is composed primarily of immunoglobulin-like (motif I) and fibronectin type III (motif II) domains with the arrangement II-II-I-I-I-I-I-II-II-II-II-II (Lange *et al.* 2005a). Exon 17a encodes a polypeptide with an EH-domain inserted between the third and fourth motif I domains; the EH-domain makes MYOM1 more elastic, like the PEVK domain in titin (Bertoncini *et al.* 2005; Schoenauer *et al.* 2005). However, the physiological importance of the inclusion of this exon is unclear. In DM1 muscle, we observed increased expression of isoform C (Fig. 1B). This isoform generates a truncated form of MYOM1 because of the presence of a premature termination codon in exon 18. Thus, the elevation of isoform C may cause a dominant negative effect and inhibit sarcomere formation in DM1 muscle. Because we focused on the regulatory mechanism of the alternative splicing of *MYOM1*, we did not study physiological significance of the finding. We are going to assess the pathological importance of the alternative splicing of *MYOM1*. Treatment of antisense oligonucleotide that prevents exon 17a inclusion is worth applying to DM model mice.

The alternative splicing of *MYOM1* exon 17a is strictly regulated in a developmental- and tissue-dependent manner: the expression of exon 17a is restricted to embryonic heart, soleus and extraocular muscle (Agarkova *et al.* 2000). Thus, the aberrant inclusion of exon 17a in DM1 muscle is consistent with other splicing abnormalities that show embryonic splicing patterns in DM1 muscle. MBNL1 regulates the developmental switch of alternative splicing in heart and skeletal muscle (Lin *et al.* 2006; Kalsotra *et al.* 2008), and the depletion of MBNL1 function is thought to play a major role in the aberrant transition of splicing to the embryonic or fetal patterns (Du *et al.* 2010). Our cellular splicing assay revealed that over-expression of MBNL family proteins decreased the inclusion of exon 17a, and the regulation of

MBNL1 required the UGCU motifs in the 5′ region of exon 17a (Figs 2B and 5B). As MBNL1 recognizes a UGCU motif to bind to its target pre-mRNA (Goers *et al.* 2010), it is anticipated that MBNL1 regulates the alternative splicing of exon 17a by its direct binding to *MYOM1* pre-mRNA. Although we did not test downregulation of MBNL1, the expanded CUG repeat increased the inclusion of exon 17a in a MBNL1-dependent manner (Fig. 2C), suggesting that abnormal splicing of *MYOM1* exon 17a in patients with DM1 may be due to the loss of MBNL function caused by trinucleotide repeat–mediated sequestering of MBNL proteins.

The essential UGCU motifs for MBNL1 regulation were located in the 5′ region of exon 17a. This is consistent with our previous study on the alternative splicing regulation of murine *Clcn1*, which revealed that MBNL1 bound to the 5′ end of exon 7A to repress its splicing. It has been suggested that the location of YGCY motifs in the flanking introns of the regulated exon determines whether MBNL1 promotes or suppresses splicing (Goers *et al.* 2010). In addition to this, our studies on alternative splicing of *Clcn1* and *MYOM1* revealed the importance of exonic region for MBNL1 function. Generally, in an exonic region, there are some regulatory elements named exonic splicing enhancer (ESE), which serve as binding sites for serine/arginine-rich (SR) proteins that facilitate the exon inclusion. So, we speculate that the interaction of MBNL1 with UGCU motifs may prevent SR proteins from binding to ESE in exon 17a and thus inhibit the exon inclusion in normal muscle. As shown in Fig. 5B, the mutation in the tandem UGCU motifs decreased the responsiveness to MBNL1, in that MBNL1 promoted the isoform D production; however, mut1 minigene responded to the over-expression of MBNL1 by increasing the ratio of isoform C. This result suggests that MBNL1 can influence the choice of 5′ splice sites of intron 17a. The production of isoform C was also promoted when the 5′ region of intron 17a was deleted (see Fig. 3, Δ5, data not shown). As there are several UGCU sequences in this deleted region, MBNL1 might bind to the 5′ region of intron 17a and prevent the usage of 5′ splice sites for isoform A and B.

In contrast to the result of MBNL proteins, the regulation of CELF proteins was inconsistent with the current pathomechanism of DM. Although the elevation of CELF1 and CELF2 has been observed in patients with DM1 (Timchenko *et al.* 1996), over-expression of either of the two proteins decreased the inclusion of exon 17a, that is, restored the splicing pattern in non-DM patients (Fig. 2B). Over-expression of CELF3 promoted strongly the inclusion of exon 17a; however, it is unlikely that CELF3 regulates *MYOM1* alternative splicing *in vivo* because CELF3 is expressed selectively in brain, and no expression of *CELF3* was detected by RT-PCR in DM1 cDNA used in this study (data not shown, and also see the study by Nezu *et al.* 2007). As MBNL1 and CELF1 regulate the alternative splicing of *MYOM1* to the same direction, we tried to dedicate the regulatory mechanisms of these two proteins. Remarkably, the minigene with mutations in the tandem UGCU motifs in exon 17a decreased responsiveness to MBNL1 but not to CELF1 (Fig. 5B,C), suggesting that CELF1 can regulate the *MYOM1* alternative splicing independently of the MBNL1 binding to its pre-mRNA. Furthermore, double expression of CELF1 and MBNL1 more suppressed the inclusion of exon 17a than those by the each protein (data not shown), which indicates the redundant function of MBNL1 and CELF1 in the regulation of the *MYOM1* alternative splicing. As CELF1 can repress the exon inclusion of *Tpm2*-ex17a minigene, it is likely that *cis*-element(s) for CELF1 may be located in exon 17a and its flanking introns. Actually, we found UGU-rich sequence around the boundary of intron 17 and exon 17a (Fig. 5A). Because we had previously shown that CELF1 bound specifically to UGU repeat RNA (Takahashi *et al.* 2000), the UGU-rich sequence might serve as a CELF1-binding site. When we think of the difference between CELF1 and MBNL1 in the regulation of the *MYOM1* alternative splicing, it is interesting that there were no UGU-rich regions around the *cis*-elements for MBNL1; the nearest UGU-rich region was 83 nt away from the UGCU motifs. Therefore, it is unlikely that MBNL1 and CELF1 may prevent the same SR protein from binding to *MYOM1* pre-mRNA. The identification of *cis*-element for CELF1 would be helpful to understand the regulation of the *MYOM1* alternative splicing; however, we plan to report the element in more detail in future.

Considering the fact that abnormal splicing of *MYOM1* exon 17a was observed in patients with DM1 (Fig 1B), the elevation of CELF1 and CELF2 might have little effect on exon 17a splicing in DM1 muscle. Actually, expression of the expanded CUG repeat prevented MBNL1 function but did not promote CELF1 activity (Fig. 2C). It is possible that DM480 might not express enough CUG repeat transcripts to lead hyperphosphorylation of CELF1 in our system; however, even so, the same level of

expression of DM480 sequestered MBNL1 and 'worsened' the splicing pattern of *MYOM1*. Thus, MBNL1 should be a main factor of the abnormal splicing of *MYOM1* in DM1 muscle.

In conclusion, we found the aberrant inclusion of *MYOM1* exon 17a as a novel splicing abnormality in DM1 skeletal muscle. This is an interesting example of the splicing abnormalities in DM1 because the alternative splicing of *MYOM1* is regulated by MBNL proteins and CELF1/2 to the same direction. Our results suggest that the abnormal behavior of CELF1/2 is negligible in aberrant splicing of *MYOM1* in patients with DM1; otherwise, another splicing factor may be involved with splicing abnormalities in DM1. Although comprehensive analysis using DNA microarray is a powerful tool to elucidate a broader network in DM1 pathomechanism, we only focused on one abnormal splicing event. There are several interesting candidates of splicing abnormalities in our microarray data. Surprisingly, only a few exons, including *LDB3/ZASP* exon 11, have been reported to be aberrantly spliced in patients with DM, and the other abnormal splicing events known to occur in patients with DM were not detected in this study. However, the probe signals for known abnormally spliced exons of *IR*, *CLCN1*, *SERCA1*, *MBNL1*, *MBNL2*, nebulin-related anchoring protein (*NRAP*) and dystrophin (*DMD*) showed changes consistent with previous reports, although the statistical significance of the changes did not satisfy our criterion, possibly due to the marked variability in DM abnormalities. Notably, a large portion of the mis-spliced candidates (13 exons) were related to the actin cytoskeleton, which may reflect the skeletal muscle histological abnormalities and muscle weakness of patients with DM1.

Experimental procedures

Tissue samples and RT-PCR

For RT-PCR validation, biopsied materials were obtained from the biceps brachii or quadriceps femoris of five patients with DM1 and five age-matched non-DM individuals without DM1-type histological abnormalities (Table S3 in Supporting Information). All clinical materials used in this study were obtained for diagnostic purposes with written informed consent. The studies were approved by the Ethical Committee. All biopsy samples were stored at −80 °C. Clinically, three of five patients with DM1 displayed a predominance of type 1 fibers. Clinical features of 'non-DM patients' are summarized in Table S3 in Supporting Information. Note that they were not diagnosed as any known myopathy including DM. Total RNA was extracted using the Trizol reagent (Invitrogen, Carlsbad, CA, USA) without DNase treatment. Typically, 0.5 μg of total RNA was reverse-transcribed for RT-PCR using a PrimeScript 1st Strand cDNA Synthesis kit (Takara Bio, Otsu, Japan) with random hexamer primers. Sequences of the primers used in RT-PCR for *MYOM1* (ex17a Fw and ex17a Rv) and *GAPDH* (GAPDH Fw and GAPDH Rv) are shown in Table S2 in Supporting Information. The products were electrophoretically resolved on an 8% polyacrylamide gel that was stained with ethidium bromide and analyzed using an LAS-3000 luminescence image analyzer (Fujifilm, Tokyo, Japan). The ratio of exon 17a inclusion in *MYOM1* was calculated as (17a inclusion)/(17a inclusion + 17a skipping) × 100 (see Fig. 1).

Microarray analysis

For microarray analysis, we used muscle samples from three patients with DM1 and four age-matched non-DM individuals (Table S3 in Supporting Information). Non-DM individuals were clinically diagnosed not as DM. Total RNA was extracted as described earlier, and its concentration and purity were calculated using NanoDrop ND-1000 (NanoDrop Technologies, Wilmington, DE, USA) and an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Total RNA samples were submitted to Kurabo Industries (Osaka, Japan; an authorized service provider for Affymetrix Japan K.K., Tokyo, Japan) for analysis using a GeneChip Human Exon 1.0 ST Array (Affymetrix Technologies) according to a standard protocol. The CEL files were further analyzed using exon array analyzer (Gellert *et al.* 2009), a web-based analytic tool, with the following criteria: probe set, core; algorithm, RMA; filters, default; *P*-value, <0.05; and SI, >1 or <−1. In this tool, gene level normalized intensity (NI) and SI are calculated as following: NI = (probe set intensity)/(expression level of the gene), SI = log 2 (NI_{DM1}/NI_{nonDM}) (Gellert *et al.* 2009).

Constructs

PCR-amplified fragments of *MYOM1* containing from exon 17 to exon 18 were cloned between the *Bsp*EI and *Eco*RI sites in pEGFP-C1 (Clontech, Mountain View, CA, USA). The genomic region covering exon 17 to exon 18 was divided into three fragments, and each fragment was amplified from human genomic DNA by PCR using pairs of primers MYOM1-1 to MYOM1-3 (Table S2 in Supporting Information). The first and second fragments were joined at the AccI site, and the second and third fragments at the *Xho*I site. To make intron-deleted minigenes, pairs of primers were designed to the flanking ends of the deleted regions. PCR products were gel-purified and circularized by self-ligation. Construction of pEGFP-Tpm2-ex1-2 vector and *Tpm2*-ex9 has been previously described (Kino *et al.* 2009). The chimeric minigenes of *Tpm2*-ex17a, 17a-9 and 9-17a were generated by insertion of PCR-amplified fragments into the pEGFP-Tpm2-

ex1-2 vector or *Tpm2*-ex9. For *Tpm2*-ex17a, a fragment corresponding to exon 17a and 100 bp of the both flanking introns was amplified from pMYOM1 with Tpm2-ex17a_Fw and Tpm2-ex17a_Rv, and this fragment was inserted between *Bgl*II and *Sal*I sites of pEGFP-Tpm2-ex1-2. The fragments for 17a-9 and 9-17a were amplified from *Tpm2*-ex17a with two primers complementary either to exon 17a (17a-9_R or 9-17a_F) or to pEGFP-C1 vector sequence (GFP-F or EGFP-C rv1). To insert these fragments into *Tpm2*-ex9, a *Bgl*II-recognition site in exon 9 and appropriate restriction sites in multicloning site were used. Because exon 17a would generate premature termination codons when ligated with exon 1, we inserted a cytosine nucleotide into exon 1 of *Tpm2*-ex17a and 17a-9 by PCR-based site-directed mutagenesis (Weiner & Costa 1994). Construction of the rest of the chimeric minigenes and site-directed mutagenesis of the pMYOM1 were performed using the PCR overlap extension method (Ho *et al.* 1989). A pair of primers with overlapping region was designed to the region around the junction of a chimeric minigene or the mutation site of a mutated minigene. The first round of PCR was carried out using each of the overlapping primers and each of two oligonucleotides complementary either to the 5′ or to 3′ ends of the minigene, to amplify overlapping portions of minigene from pMYOM1 or appropriate chimeric minigenes. Sequences of all the primers used in construction are shown in Table S2 in Supporting Information. The other constructs used in this study have been described previously (Kino *et al.* 2004, 2009; Nezu *et al.* 2007; Mori *et al.* 2008; Onishi *et al.* 2008). Briefly, the coding regions for the RNA-binding proteins were PCR-amplified from a human cDNA library and cloned into pSecDK, a mammalian expression vector with a myc-tag that was modified from pSecTagA (Invitrogen) to delete the Igκ chain leader sequence. The constructs have been confirmed to express intended proteins in a comparable level (Kino *et al.* 2009). DM18 and DM480 contain a fragment of the 3′ region of *DMPK* with a CTG18 and interrupted CTG480 repeats, respectively.

Cellular splicing assay

Cells transfected with plasmids for the expression of a protein and minigene were harvested 48 h post-transfection. Typically, cells were cultured in 12-well plates and transfected with 0.4 μg of protein expression plasmid (or cognate empty vector) and 0.1 μg of minigene expression plasmid using FuGENE6.0 (Roche Applied Sciences, Indianapolis, IN, USA). Total RNA was extracted with a GenElute Mammalian Total RNA Miniprep kit (Sigma-Aldrich, St. Louis, MO, USA) without DNase treatment. Typically, 1.0 μg of total RNA was reverse-transcribed with a PrimeScript 1st Strand cDNA Synthesis kit (Takara Bio) using oligo(dT) primers. PCR was performed using two oligonucleotides complementary to the EGFP sequence (GFP-F) and the last exon of minigene (ex17a Rv for pMYOM1 and deleted or mutated minigenes; Tpm2-Rv for chimeric minigenes), respectively. The products were electrophoretically resolved on an 8% polyacrylamide gel that was stained with ethidium bromide and analyzed using an LAS-3000 luminescence image analyzer.

Acknowledgements

We thank Dr H. Mitsuhashi for valuable discussions and encouragement. This work was supported by research grants from the Ministry of Health, Labor and Welfare, Japan (20B-13).

References

Agarkova, I., Auerbach, D., Ehler, E. & Perriard, J.C. (2000) A novel marker for vertebrate embryonic heart, the EH-myomesin isoform. *J. Biol. Chem.* **275**, 10256–10264.

Aslanidis, C., Jansen, G., Amemiya, C., *et al.* (1992) Cloning of the essential myotonic dystrophy region and mapping of the putative defect. *Nature* **355**, 548–551.

Bertoncini, P., Schoenauer, R., Agarkova, I., Hegner, M., Perriard, J.C. & Guntherodt, H.J. (2005) Study of the mechanical properties of myomesin proteins using dynamic force spectroscopy. *J. Mol. Biol.* **348**, 1127–1137.

Brook, J.D., McCurrach, M.E., Harley, H.G., *et al.* (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. *Cell* **68**, 799–808.

Buxton, J., Shelbourne, P., Davies, J., Jones, C., Van Tongeren, T., Aslanidis, C., de Jong, P., Jansen, G., Anvret, M., Riley, B., Williamson, R. & Johnson, K. (1992) Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature* **355**, 547–548.

Charlet, B.N., Savkur, R.S., Singh, G., Philips, A.V., Grice, E.A. & Cooper, T.A. (2002) Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. *Mol. Cell* **10**, 45–53.

Davis, B.M., McCurrach, M.E., Taneja, K.L., Singer, R.H. & 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.

Du, H., Cline, M.S., Osborne, R.J., Tuttle, D.L., Clark, T.A., Donohue, J.P., Hall, M.P., Shiue, L., Swanson, M.S., Thornton, C.A. & Ares, M. Jr (2010) Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nat. Struct. Mol. Biol.* **17**, 187–193.

Fardaei, M., Rogers, M.T., Thorpe, H.M., Larkin, K., Hamshere, M.G., Harper, P.S. & Brook, J.D. (2002) Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum. Mol. Genet.* **11**, 805–814.

Gellert, P., Uchida, S. & Braun, T. (2009) Exon Array Analyzer: a web interface for Affymetrix exon array analysis. *Bioinformatics* **25**, 3323–3324.

Goers, E.S., Purcell, J., Voelker, R.B., Gates, D.P. & Berglund, J.A. (2010) MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Res.* **38**, 2467–2484.

Hao, M., Akrami, K., Wei, K., De Diego, C., Che, N., Ku, J.H., Tidball, J., Graves, M.C., Shieh, P.B. & Chen, F. (2008) Muscleblind-like 2 (Mbnl2) -deficient mice as a model for myotonic dystrophy. *Dev. Dyn.* **237**, 403–410.

Harley, H.G., Brook, J.D., Rundle, S.A., Crow, S., Reardon, W., Buckler, A.J., Harper, P.S., Housman, D.E. & Shaw, D.J. (1992) Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* **355**, 545–546.

Harper, P.S. (2001) *Myotonic Dystrophy*, 3rd edn. London: W. B. Saunders.

Hino, S., Kondo, S., Sekiya, H., Saito, A., Kanemoto, S., Murakami, T., Chihara, K., Aoki, Y., Nakamori, M., Takahashi, M.P. & Imaizumi, K. (2007) Molecular mechanisms responsible for aberrant splicing of SERCA1 in myotonic dystrophy type 1. *Hum. Mol. Genet.* **16**, 2834–2843.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. & Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51–59.

Ho, T.H., Bundman, D., Armstrong, D.L. & Cooper, T.A. (2005) Transgenic mice expressing CUG-BP1 reproduce splicing mis-regulation observed in myotonic dystrophy. *Hum. Mol. Genet.* **14**, 1539–1547.

Ho, T.H., Charlet, B.N., Poulos, M.G., Singh, G., Swanson, M.S. & Cooper, T.A. (2004) Muscleblind proteins regulate alternative splicing. *EMBO J.* **23**, 3103–3112.

Holt, I., Jacquemin, V., Fardaei, M., Sewry, C.A., Butler-Browne, G.S., Furling, D., Brook, J.D. & Morris, G.E. (2009) Muscleblind-like proteins: similarities and differences in normal and myotonic dystrophy muscle. *Am. J. Pathol.* **174**, 216–227.

Kalsotra, A., Xiao, X., Ward, A.J., Castle, J.C., Johnson, J.M., Burge, C.B. & Cooper, T.A. (2008) A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. *Proc. Natl Acad. Sci. USA* **105**, 20333–20338.

Kanadia, R.N., Johnstone, K.A., Mankodi, A., Lungu, C., Thornton, C.A., Esson, D., Timmers, A.M., Hauswirth, W.W. & Swanson, M.S. (2003) A muscleblind knockout model for myotonic dystrophy. *Science* **302**, 1978–1980.

Kimura, T., Nakamori, M., Lueck, J.D., Pouliquin, P., Aoike, F., Fujimura, H., Dirksen, R.T., Takahashi, M.P., Dulhunty, A.F. & Sakoda, S. (2005) Altered mRNA splicing of the skeletal muscle ryanodine receptor and sarcoplasmic/endoplasmic reticulum Ca2+ -ATPase in myotonic dystrophy type 1. *Hum. Mol. Genet.* **14**, 2189–2200.

Kino, Y., Mori, D., Oma, Y., Takeshita, Y., Sasagawa, N. & Ishiura, S. (2004) Muscleblind protein, MBNL1/EXP, binds specifically to CHHG repeats. *Hum. Mol. Genet.* **13**, 495–507.

Kino, Y., Washizu, C., Oma, Y., Onishi, H., Nezu, Y., Sasagawa, N., Nukina, N. & Ishiura, S. (2009) MBNL and CELF proteins regulate alternative splicing of the skeletal muscle chloride channel CLCN1. *Nucleic Acids Res.* **37**, 6477–6490.

Kuyumcu-Martinez, N.M., Wang, G.S. & Cooper, T.A. (2007) Increased steady-state levels of CUGBP1 in myotonic dystrophy 1 are due to PKC-mediated hyperphosphorylation. *Mol. Cell* **28**, 68–78.

Lange, S., Agarkova, I., Perriard, J.C. & Ehler, E. (2005a) The sarcomeric M-band during development and in disease. *J. Muscle Res. Cell Motil.* **26**, 375–379.

Lange, S., Himmel, M., Auerbach, D., Agarkova, I., Hayess, K., Furst, D.O., Perriard, J.C. & Ehler, E. (2005b) Dimerisation of myomesin: implications for the structure of the sarcomeric M-band. *J. Mol. Biol.* **345**, 289–298.

Lin, X., Miller, J.W., Mankodi, A., Kanadia, R.N., Yuan, Y., Moxley, R.T., Swanson, M.S. & Thornton, C.A. (2006) Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum. Mol. Genet.* **15**, 2087–2097.

Liquori, C.L., Ricker, K., Moseley, M.L., Jacobsen, J.F., Kress, W., Naylor, S.L., Day, J.W. & Ranum, L.P. (2001) Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* **293**, 864–867.

Lueck, J.D., Lungu, C., Mankodi, A., Osborne, R.J., Welle, S.L., Dirksen, R.T. & Thornton, C.A. (2007) Chloride channelopathy in myotonic dystrophy resulting from loss of posttranscriptional regulation for CLCN1. *Am. J. Physiol.* **292**, C1291–C1297.

Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M. & Thornton, C.A. (2000) Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science* **289**, 1769–1773.

Mankodi, A., Takahashi, M.P., Jiang, H., Beck, C.L., Bowers, W.J., Moxley, R.T., Cannon, S.C. & Thornton, C.A. (2002) Expanded CUG repeats trigger aberrant splicing of ClC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol. Cell* **10**, 35–44.

Miller, J.W., Urbinati, C.R., Teng-Umnuay, P., Stenberg, M.G., Byrne, B.J., Thornton, C.A. & Swanson, M.S. (2000) Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J.* **19**, 4439–4448.

Mori, D., Sasagawa, N., Kino, Y. & Ishiura, S. (2008) Quantitative analysis of CUG-BP1 binding to RNA repeats. *J. Biochem.* **143**, 377–383.

Nezu, Y., Kino, Y., Sasagawa, N., Nishino, I. & Ishiura, S. (2007) Expression of MBNL and CELF mRNA transcripts in muscles with myotonic dystrophy. *Neuromuscul. Disord.* **17**, 306–312.

Onishi, H., Kino, Y., Morita, T., Futai, E., Sasagawa, N. & Ishiura, S. (2008) MBNL1 associates with YB-1 in cytoplasmic stress granules. *J. Neurosci. Res.* **86**, 1994–2002.

Philips, A.V., Timchenko, L.T. & Cooper, T.A. (1998) Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science* **280**, 737–741.

Ranum, L.P. & Cooper, T.A. (2006) RNA-mediated neuromuscular disorders. *Annu. Rev. Neurosci.* **29**, 259–277.

Ranum, L.P. & Day, J.W. (2004) Myotonic dystrophy: RNA pathogenesis comes into focus. *Am. J. Hum. Genet.* **74**, 793–804.

Savkur, R.S., Philips, A.V. & Cooper, T.A. (2001) Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat. Genet.* **29**, 40–47.

Schoenauer, R., Bertoncini, P., Machaidze, G., Aebi, U., Perriard, J.C., Hegner, M. & Agarkova, I. (2005) Myomesin is a molecular spring with adaptable elasticity. *J. Mol. Biol.* **349**, 367–379.

Sen, S., Talukdar, I., Liu, Y., Tam, J., Reddy, S. & Webster, N.J. (2010) Muscleblind-like 1 (Mbnl1) promotes insulin receptor exon 11 inclusion via binding to a downstream evolutionarily conserved intronic enhancer. *J. Biol. Chem.* **285**, 25426–25437.

Takahashi, N., Sasagawa, N., Suzuki, K. & Ishiura, S. (2000) The CUG-binding protein binds specifically to UG dinucleotide repeats in a yeast three-hybrid system. *Biochem. Biophys. Res. Commun.* **277**, 518–523.

Taneja, K.L., McCurrach, M., Schalling, M., Housman, D. & Singer, R.H. (1995) Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J. Cell Biol.* **128**, 995–1002.

Timchenko, L.T., Miller, J.W., Timchenko, N.A., DeVore, D.R., Datar, K.V., Lin, L., Roberts, R., Caskey, C.T. & Swanson, M.S. (1996) Identification of a (CUG)n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res.* **24**, 4407–4414.

Timchenko, N.A., Patel, R., Iakova, P., Cai, Z.J., Quan, L. & Timchenko, L.T. (2004) Overexpression of CUG triplet repeat-binding protein, CUGBP1, in mice inhibits myogenesis. *J. Biol. Chem.* **279**, 13129–13139.

Warf, M.B., Diegel, J.V., von Hippel, P.H. & Berglund, J.A. (2009) The protein factors MBNL1 and U2AF65 bind alternative RNA structures to regulate splicing. *Proc. Natl Acad. Sci. USA* **106**, 9203–9208.

Weiner, M.P. & Costa, G.L. (1994) Rapid PCR site-directed mutagenesis. *PCR Methods Appl.* **4**, S131–S136.

Received: 26 April 2011
Accepted: 8 June 2011

Supporting Information/Supplementary material

The following Supporting Information can be found in the online version of the article:

Figure S1 MBNL1 and CELF1 regulate *MYOM1* splicing in other cell lines.

Table S1 Microarray data

Table S2 Primer sequences

Table S3 Patient samples

Additional Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Journal of Neuroscience Research

Journal of Neuroscience Research 89:562–575 (2011)

Differential Effects of the HESR/HEY Transcription Factor Family on Dopamine Transporter Reporter Gene Expression Via Variable Number of Tandem Repeats

Kouta Kanno[1,2] and Shoichi Ishiura[1]*

[1]Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan
[2]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo, Japan

The 3′-untranslated region (UTR) of the human dopamine transporter (*DAT1*) gene contains a variable number of tandem repeats (VNTR) domain, which is thought to be associated with dopamine-related psychiatric disorders, personality, and behavior. However, the molecular and neuronal functions of polymorphisms within the VNTR domain are unknown. We previously identified the transcription factor HESR1 (HEY1) as a VNTR-binding protein. *Hesr1* knockout mice exhibit DAT up-regulation in the brain and low levels of spontaneous activity. Other members of the HESR (HEY) family, including HESR2 (HEY2) and 3 (HEYL), have similar DNA-binding domains. In this study, we analyzed the effects of HESR1, -2, and -3 on *DAT1* expression in human neuroblastoma SH-SY5Y cells using luciferase reporter assays. We found that the VNTR domain played an inhibitory role in *DAT1* reporter gene expression and that HESR1 and -2 inhibited expression via both the core promoter and the VNTR. The inhibitory effects of HESR family members on *DAT* reporter gene expression differed depending on the number of repeats in the VNTR domain. We also found that each Hesr was expressed in the dopaminergic neurons in the mouse midbrain. These results suggest that the HESR family is involved in *DAT* expression via the VNTR domain.

Key words: dopamine; genetic polymorphism; VNTR; luciferase reporter assay

The dopaminergic nervous system plays important regulatory roles in locomotion, cognition, reward, affection, and hormone release (Bannon et al., 2001; Jackson and Westlinddanielsson, 1994; Missale et al., 1998; Uhl, 2003). Thus, dopamine and its related genes are thought to be involved in neuropsychiatric disorders and behavioral traits. The human dopamine transporter (*DAT1*) gene is involved in many dopamine-related disorders. DAT levels are reduced in Parkinson's disease (PD) and elevated in attention deficit hyperactivity disorder (ADHD), Tourette's syndrome, and major depression (Madras et al., 1998; Muller-Vahl et al., 2000; Brunswick et al., 2003; Krause et al., 2003). In addition, several psychoactive drugs, including cocaine, amphetamine, and methylphenidate, are known to inhibit dopamine reuptake by DAT (Giros et al., 1991, 1992; Kilty et al., 1991; Shimada et al., 1991; Giros and Caron, 1993).

A functional genetic polymorphism has been described in the 3′-untranslated region (UTR) of exon 15 in *DAT1* (Michelhaugh et al., 2001). This 3′-UTR contains a 40-bp-long variable number of tandem repeats (VNTR) domain (Fig. 1; Vandenbergh et al., 1992; Michelhaugh et al., 2001). The polymorphism within this region is known to be associated with such neuropsychiatric disorders as ADHD, PD, alcoholism, and drug abuse (Cook et al., 1995; Ueno et al., 1999; Vandenbergh et al., 2000; Ueno, 2003; D'souza and Craig, 2008) and with modified gene expression depending on the genotype in vivo (Heinz et al., 2000; Jacobsen et al., 2000; Mill et al., 2002; D'souza and Craig, 2008) and in mammalian cell lines (Fuke et al., 2001, 2005; Inoue-Murayama et al., 2002; Miller and Madras, 2002; Greenwood and Kelsoe, 2003; Mill et al., 2005; VanNess et al., 2005; D'souza and Craig, 2008). It is expected that this region contains binding sites for interacting proteins, but, because these factors have not been described, the molecular and neuronal functions of the polymorphism are unknown (Michelhaugh et al., 2001).

We previously identified and characterized HESR1 (HEY1) as a *trans*-acting repressor of gene expression

Contract grant sponsor: Human Frontier Science Program; Contract grant sponsor: Ministry of Education, Culture, Sports, Science, and Technology of Japan; Contract grant sponsor: JSPS Research Fellowship for Young Scientists (to K.K.).

*Correspondence to: Dr. Shoichi Ishiura, Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1, Komaba, Meguro-ku, Tokyo 153-8902, Japan.
E-mail: cishiura@mail.ecc.u-tokyo.ac.jp

Received 18 August 2010; Revised 10 November 2010; Accepted 9 December 2010

Published online 2 February 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/jnr.22593

that acts through the 3′-UTR of *DAT1* via a yeast one-hybrid system (Fuke et al., 2005). In addition, we also showed that HESR1 binds directly to the region by electrophoretic mobility shift assay (EMSA) and represses the expression of the endogenous *DAT1* gene in the HEK293 cell line by RT-PCR assay (Fuke et al., 2006). In *Hers1* knockout (KO) mice, *DAT1* expression was enhanced, and the mice exhibited a reduction in spontaneous locomotor activity and exploration to novelty (Fuke et al., 2006). These findings suggest that the human *HESR1* gene and the polymorphisms could be related to many psychiatric disorders and behavioral traits. However, it is possible that other factors affect *DAT1* expression via the VNTR domain, insofar as more than one interacting factor is expected to bind this region (Michelhaugh et al., 2001). Conflicting results have been reported from studies using different cell lines, which may express different transcription factors (Fuke et al., 2001, 2005; Inoue-Murayama et al., 2002; Miller and Madras, 2002; Greenwood and Kelsoe, 2003; Mill et al., 2005; VanNess et al., 2005; D'souza and Craig, 2008). These results suggest that *DAT1* expression can be altered by cell-specific factors depending on the VNTR alleles present.

The Hesr family genes *Hesr1*, *-2*, and *-3* (*Hey1*, *Hey2*, and *HeyL*) were identified as the hairy/enhancer split-type basic helix-loop-helix (bHLH) genes. They have been shown to be direct transcriptional targets of the Notch signaling pathway, which is essential for neural development (Kokubo et al., 1999; Leimeister et al., 1999; Nakagawa et al., 1999; Henderson et al., 2001; Iso et al., 2001, 2003; Wang et al., 2002; Sakamoto et al., 2003). HESR family genes carry a bHLH domain essential for DNA binding as well as an Orange domain and YRPW motif, which mediate interaction with proteins and affect dimerization or recruitment of corepressors (Fischer and Gessler, 2007; Fig. 2). HESR proteins repress the expression of target genes by binding to E- or N-box bHLH-binding consensus sites (Nakagawa et al., 2000; Iso et al., 2001, 2003). The bHLH domain is highly conserved among HESR family members (Steidl et al., 2000), and Hesr1 and -2 repress gene expression via the same sequence (Kokubo et al., 2007). Thus, HESR1 along with HESR2 and -3 may be candidate factors regulating *DAT* expression via the VNTR. However, any roles of HESR2 and -3 in *DAT* gene expression have not yet been characterized.

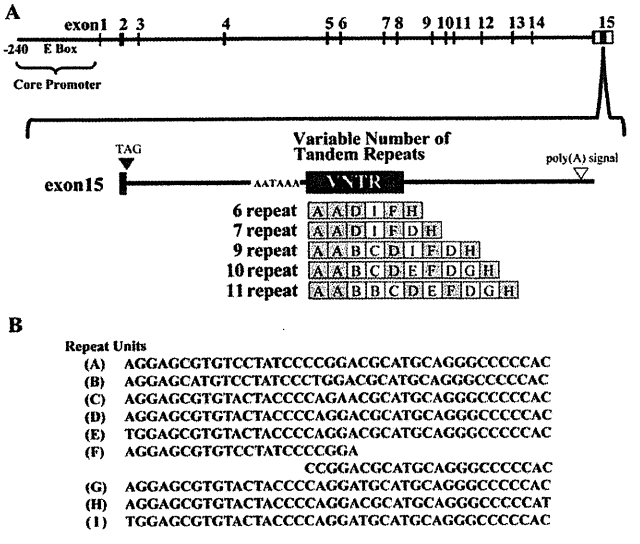

Fig. 1. Genomic structure of *DAT1* and allelic variants of the VNTR domain in exon 15. **A:** The coding region (black box), noncoding region (open box), VNTR domain, and constant parts of the repeat units (gray box) are shown. Exon 15 of *DAT1* contains a stop codon (solid arrowhead) and polyadenylation signal (open arrowhead). Upstream of the VNTR domain are six nucleotides (AATAAA) that resemble a polyadenylation signal. The allelic variants of the VNTR indicate the repeat unit type (A–I) for each allele. **B:** Nucleotide sequence of each unit of the VNTR polymorphism in the 3′-UTR of *DAT1*.

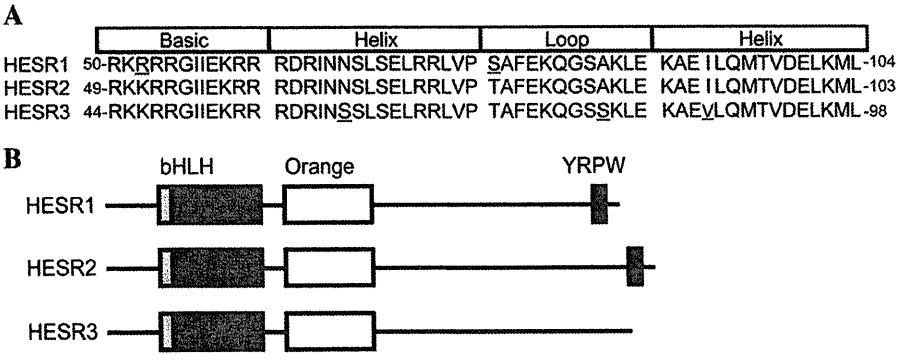

Fig. 2. Structure of HESR family members. **A:** Comparison of the primary sequences of the bHLH domain among human HESR family members (HESR1, 50–104 amino acids; HESR2, 49–103 amino acids; and HESR3, 44–98 amino acids). Those residues that differ among the family members are underscored. The primary sequences of the bHLH domain in the mouse Hesr family are identical to those in the human protein. **B:** Structure of the HESR family with the three major domains noted: bHLH domain, Orange domain, and YRPW motif.

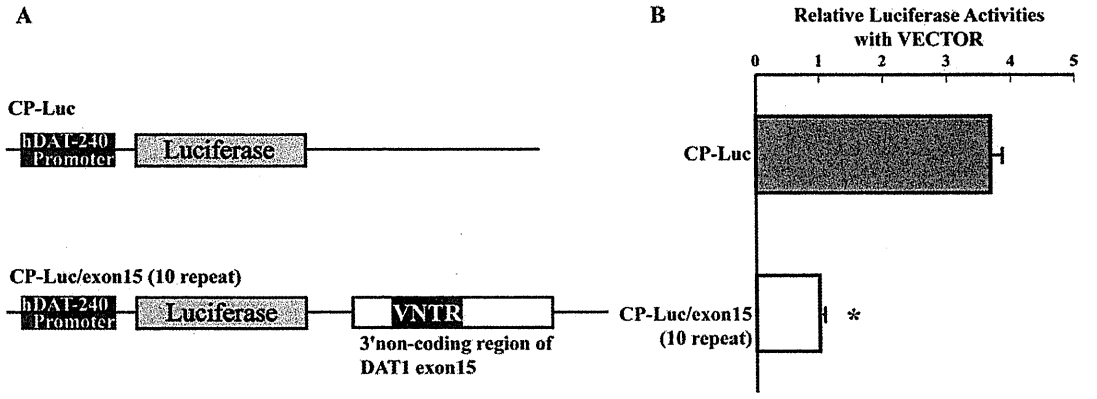

Fig. 3. Luciferase reporter vector and its activity in SY-SH5Y cells. **A:** Schematic diagram of the luciferase reporter vectors CP-Luc and CP-Luc/exon 15. CP-Luc contains only the *DAT* core promoter, whereas CP-Luc/exon 15 contains both the core promoter and the 3′-UTR with 10r, which is the most common allele. **B:** Negative regulation of gene expression through the 3′-UTR of *DAT* in SY-SH5Y cells. CP-Luc or CP-Luc/exon 15 and empty vector were transfected into the cells. Relative luciferase activity is expressed as the mean ± SEM. $*P < 0.0001$, Student's *t*-test.

In this study, we sought to clarify whether HESR2 and -3, as well as HESR1, also affect *DAT1* expression in human neuroblastoma cells via the 3′-UTR, including the VNTR region. Luciferase reporter constructs were made containing the endogenous *DAT1* core promoter and the VNTR with 6, 7, 9, 10, or 11 repeats, which is consistent with the sequences reported previously (Fuke et al., 2001). The differential effects on VNTR among HESR1, -2, and -3 were also characterized. Furthermore, we conducted immunohistochemistry for Hesrs and tyrosine hydroxylase (TH), a marker of dopaminergic neurons, in mouse midbrains to analyze localization of Hesr proteins for the investigation of HESR family function in the brain dopamine systems, because localization of HESR proteins in the adult brain has not been reported, although dopaminergic neurons localize in the specific brain regions (Bjorklund and Dunnett, 2007).

MATERIALS AND METHODS

General Procedure

First, to characterize the role of the 3′-UTR in *DAT* expression, a CP-Luc or CP-Luc/exon15 (10r) reporter vector and empty vector (control) were expressed in cells. Ten repeats is the most common allele in the population (Fuke et al., 2005; D'souza and Craig, 2008). CP-Luc contained only the *DAT* core promoter, whereas CP-Luc/exon15 contained both the core promoter and the 3′-UTR (Fig. 3). Next, the empty vector or each HESR was cotransfected with CP-Luc or CP-Luc/exon15 (10r). Finally, to compare the VNTR alleles and HESRs, each repeat (6, 7, 9, 10, or 11r) of CP-Luc/exon15 and the HESR or vector were cotransfected. The relative luciferase activity was standardized to that of the vector. In the experiment examining the repeat effect, the relative luciferase activity was standardized to that of 10r in each group. In addition, the localization of Hesrs in the mouse midbrain was detected by immunohistochemistry.

Cloning and Construction

Two kinds of luciferase reporter vectors were prepared: CP-Luc and CP-Luc/exon 15 (see Fig. 3A). CP-Luc contained the human *DAT* core promoter (–240 to +2; Fig. 1A; Kouzmenko et al., 1997; Sacchetti et al., 1999) cloned from the DAT1-8317 plasmid (Sacchetti et al., 1999), a gift from Dr. Michael J. Bannon (Wayne State University, School of Medicine, Detroit, MI), upstream of the firefly luciferase site in the modified pGL3 vector (Promega, Madison, WI). CP-Luc/exon 15 contained the human *DAT* core promoter and 3′-UTR region including the VNTR domain downstream of the luciferase site in CP-Luc. There are five kinds of VNTR alleles (6, 7, 9, 10, and 11r), which is consistent with the sequence (Fig. 1B) reported in our previous study (Fuke et al., 2001). These reporter vectors are the same as the constructs used in our previous study (Fuke et al., 2005); schematic structures of the reporter vectors are shown in Figure 3A. Additional information on these constructs is described in our previous work (Fuke et al., 2005).

All of the HESR family expression vectors were made by cloning the cDNA into myc-pcDNA modified from pcDNA 3.1+ (Invitrogen, Carlsbad, CA), an expression vector for mammalian cell lines. The Myc-tag is located upstream of the multicloning site. Myc-pcDNA was also used as a control vector (the vector described in Figs. 3–6). Myc-pcDNA and Myc-HESR1 (human) are the same as the constructs described in our previous study (Fuke et al., 2005). Myc-Hesr1 (mouse) was a gift from Dr. Hiroki Kokubo (Division of Mammalian Development, National Institute of Genetics, Mishima, Japan). Human HESR2 and -3 cDNAs were amplified from a fetal brain cDNA library (Clontech, Palo Alto, CA) by PCR with oligonucleotide primers and the high-fidelity DNA polymerase PrimeStar (TaKaRa, Shiga, Japan). Mouse Hesr2 and -3 cDNAs were amplified from cloned cDNAs in pBluescript (gifts from Dr. Hiroki Kokubo) by PCR with oligonucleotide primers and PrimeStar (TaKaRa). Each oligonucleotide primer was designed to

contain a restriction site (italicized). The primer sequences are as follows: human HESR2, Fw (XhoI) 5′-TTA*CTCGA*GAT GAAGCGCCC-3′ and Rv (ApaI) 5′-TTAGGGCCCT TAAAAAGCTCCAAC-3′; human HESR3, Fw (XhoI) 5′-TTA*CTCGA*GATGAAGCGACCC-3′ and Rv (XhoI) 5′-TTA*CTCGA*GTCAGAAAGCCCC-3′; mouse Hesr2, Fw (XhoI) 5′-TTA*CTCGA*GATGAAGCGCCCT-3′ and Rv (ApaI) 5′-ATA*GGGCCC*TTAAAAGCTGGCTCC-3′; and mouse Hesr3, Fw (XhoI) 5′-TTA*CTCGA*GATGAA GCGGCC-3′ and Rv (ApaI) 5′-TTAGGGCCCTCAGAAA GCC-3′. These amplified fragments were cloned into T-Vectors (Promega p-GEM T Easy Kit) and sequenced by the dideoxy chain termination method with CEQ DTCS and CEQ8000 (Beckman Coulter, Brea, CA). Finally, the fragments were digested with each restriction enzyme and subcloned into myc-pcDNA.

Cell Culture and Transient Transfection: Luciferase Reporter Assay

The methods used for culture, transfection, cell harvesting, and luciferase activity measurements followed the standard methods of the Dual-Luciferase Reporter Assay System (Promega). SH-SY5Y cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Sigma, St. Louis, MO) supplemented with 10% FBS at 37°C in an atmosphere of 5% CO_2. SH-SY5Y cells were plated into 24-well plates and cultured until they grew to 80% confluence before transfection. The cells were transfected with 1 μg total plasmid using Lipofectamine 2000 reagent (Invitrogen). The firefly luciferase reporter gene (0.5 μg) and each HESR or vector (0.5 μg) were coexpressed in the cells. Plasmid pRL (Promega) containing the sea pansy luciferase gene was cotransfected (20 ng) as an internal control to normalize the transfection efficiency in all experiments. After 48 hr, the cells were harvested and stored at −80°C. Luciferase activity was assayed using the Dual-Luciferase Reporter Assay System (Promega). The firefly and sea pansy luciferase activity was measured using a Centro LB 960 (Berthold, Bad Wildbad, Germany) for 10 sec after a 2-sec delay, and then the value of each sample was calculated as light units of firefly luciferase per that of sea pansy. Each HESR group and its controls were measured at the same time on a Centro LB 960.

Animals

Adult (9-week-old) male C57BL6/J mice (CLEA Japan, Tokyo, Japan) were kept under a controlled temperature (23–25°C) and photoperiod (LD 14:10, lights off at 22:00 hr). Food and water were available freely. All experiments were conducted according to the Regulations for Animal Experimentation at the University of Tokyo (Tokyo, Japan).

Tissue Preparation

The mice were deeply anesthetized using sodium pentobarbital (50 mg/kg body weight) and then perfused intracardially with 0.05 M phosphate-buffered saline (PBS; pH 7.4) followed by 4% paraformaldehyde in 0.05 M phosphate buffer (PB). The brains were removed and postfixed with the same fixative for 2 hr and immersed in 30% sucrose in 0.05 M PB for several days at 4°C. Serial coronal brain sections (20 μm) including the midbrain ventral tegmental area (VTA) and substantia nigra (SN; −3.04 to −3.49 to the bregma) were made with a cryostat and collected according to a brain map (Franklin and Paxinos, 2008). Five animals were used in this experiment.

Immunohistochemistry

Immunoperoxidase staining. Free-floating sections of the midbrain were incubated with 0.6% H_2O_2 in 10 mM PBS for 30 min at room temperature (RT) before and after rinsing with 10 mM PBS. Next, the sections were incubated with 5% normal goat serum (NGS; Vector Laboratories, Burlingame, CA), 0.4% Triton X-100, and 10 mM PBS for 1 hr at RT and then with a primary rabbit antibody against Hesr1 (working dilution 1:500; Chemicon, Temecula, CA), Hesr2 (1:1,000; Chemicon), or Hesr3 (1:500; Chemicon) containing 5% NGS, 0.4% Triton X-100, and 10 mM PBS for 3 nights at 4°C. After washing with 10 mM PBS, the sections were reacted with biotinylated goat anti-rabbit IgG (1:200; Vector Laboratories) in 5% NGS, 0.4% Triton X-100, and 10 mM PBS overnight at 4°C. The sections were rinsed with 10 mM PBS three times and then incubated in avidin-biotin-peroxidase complex (ABC) solution (Vectastain Elite ABC Kit; Vector Laboratories). Next, the sections were reacted with 0.05% 3,3′-diaminobenzidine (DAB) in 0.01% H_2O_2 and 100 mM Tris-HCl to visualize Hesr1-, -2-, or -3-immunoreactive (-ir) cells.

For single staining of the TH-ir (a marker of dopaminergic neurons) cells, similar steps were taken, except for the steps involving incubation with the primary or secondary antibody and visualization. The sections were incubated with a mouse anti-TH antibody (1:10,000; Chemicon) overnight at 4°C. The sections were then reacted with biotinylated goat anti-mouse IgG (1:200, Vector Laboratories) in 5% NGS, 0.4% Triton X-100, and 10 mM PBS overnight at 4°C after washing with 10 mM PBS three times. The sections were rinsed with 10 mM PBS three times and then incubated in ABC solution (Vector Laboratories). Next, the sections were reacted with 0.02% DAB solution to visualize TH-ir cells.

For double labeling, the same staining steps as for TH were performed after staining for each Hesr. Instead of DAB, a Vector SG Kit was used for visualization of TH after the ABC reaction. Immunostained sections mounted on slides were dehydrated through a graded ethanol series, cleared with xylene, and then coverslipped with an embedding agent.

Fluorescence immunohistochemistry for Hesrs and TH. Free-floating sections were rinsed with 10 mM PBS. The sections were then incubated with 5% normal donkey serum (NDS; Jackson Immunoresearch, West Grove, PA), 0.4% Triton X-100, and 10 mM PBS for 1 hr at RT and then with the primary rabbit antibody against Hesr1 (1:500; Chemicon), Hesr2 (1:1,000; Chemicon), or Hesr3 (1:500; Chemicon) and the primary mouse antibody for TH (1:10,000; Chemicon) in 5% NDS, 0.4% Triton X-100, and 10 mM PBS for 2 nights at 4°C. After washing with 10 mM PBS, the sections were reacted with donkey anti-rabbit IgG (1:200; Jackson Immunoresearch; Cy3) and donkey anti-mouse IgG (1:200; Jackson Immunoresearch; Cy2) in 5% NDS, 0.4% Triton X-100, and 10 mM PBS overnight at

4°C. Next, the sections were rinsed with 10 mM PBS before and after incubating with Hoechst 33342 (1:1,000; Dojindo, Tokyo, Japan) for 15 min. The immunostained sections were mounted on slides and then coverslipped. Images were captured with a digital CCD camera (DP70; Olympus, Tokyo, Japan) and the microscope software manager DP (Olympus) and analyzed in Photoshop CS4 (Adobe, San Jose, CA).

Statistical Analysis

Statistical analysis was performed in JMP 8.0 (SAS Institute). All values are reported as the mean ± SEM. Each value was standardized to that of CP-Luc/exon15 (10r) (Fig. 3), vector (Figs. 4, 5), or 10r (Fig. 6) in each group. Student's *t*-test was performed to detect statistical significance between two experimental objects. Tukey-Kramer's honestly significant difference (HSD) test was used as a post hoc test after one-way ANOVA. Tukey-Kramer's HSD test is a standard method of JMP to find significant difference among data after ANOVA. In addition, two-way ANOVA was conducted to detect interactive effects between the number of VNTR repeats and HESRs. Differences were considered significant at $P < 0.05$.

RESULTS

Lusciferase Activity of CP-Luc and CP-Luc/Exon15 (10r) in SH-SY5Y Cells

To determine the effect of the VNTR domain on *DAT1* expression, we made luciferase reporter constructs (Fig. 3). The relative luciferase activity of CP-Luc was 3.68 ± 0.15, whereas that of CP-Luc/exon 15 (10r) was 1.00 ± 0.03. A significant difference was noted between CP-Luc and CP-Luc/exon 15 (10r) ($P < 0.0001$, n = 8; Fig. 3). These results indicate that the 3′-UTR including the VNTR domain strongly inhibited luciferase expression in this reporter assay.

Effects of Human HESR and Mouse Hesr Family on Lusciferase Activity of CP-Luc

We next examined the effect of HESR family members on expression of the CP-Luc plasmid (Fig. 4A).

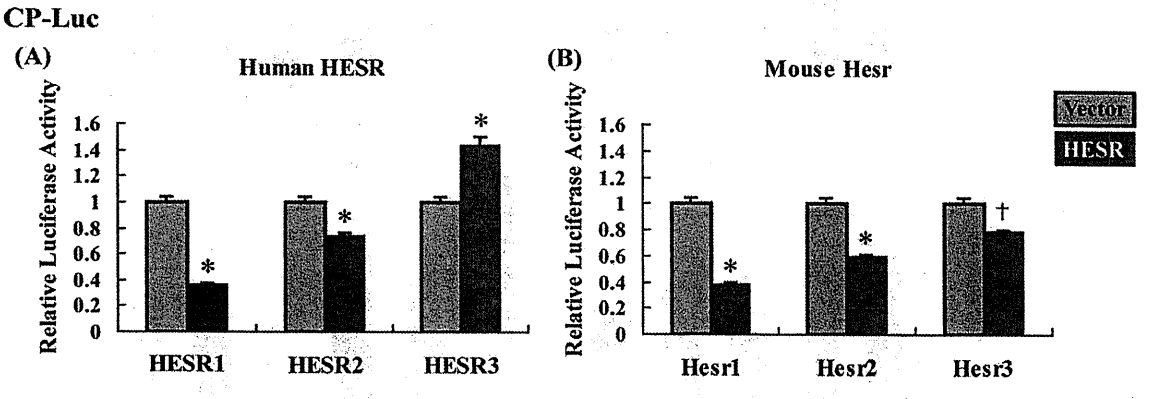

Fig. 4. Effects of the HESR family on the luciferase activity of CP-Luc. Cells were transfected with the CP-Luc reporter construct and human (**A**) and mouse (**B**) Hesrs or with empty vector. The values represent the mean ± SEM. of relative luciferase activity of the reporter construct in each group. *$P < 0.0001$ vs. vector, †$P < 0.0003$ vs. vector, Student's *t*-test.

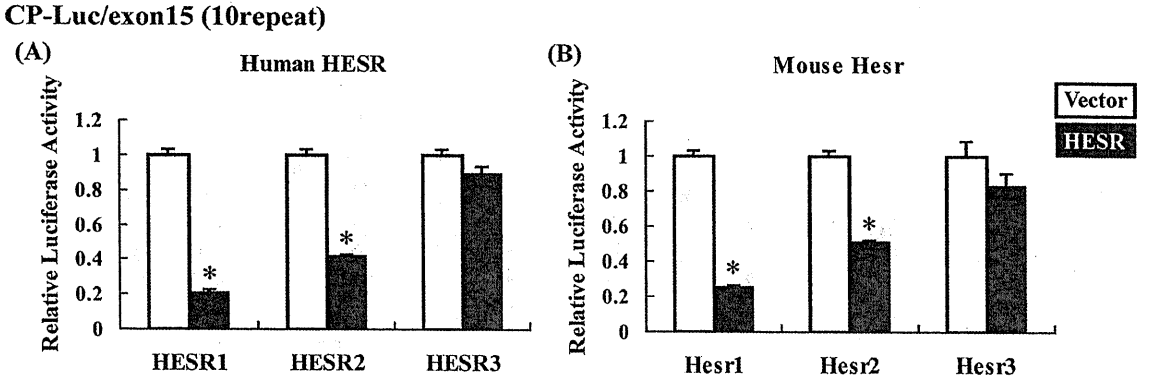

Fig. 5. Effects of the HESR family on the luciferase activity of CP-Luc/exon 15 (10r). Cells were transfected with CP-Luc/exon 15 (10r) and human (**A**) and mouse (**B**) Hesrs or with empty vector. The values represent the mean ± SEM. of relative luciferase activity of the reporter construct in each group. *$P < 0.0001$ vs. vector, Student's *t*-test.

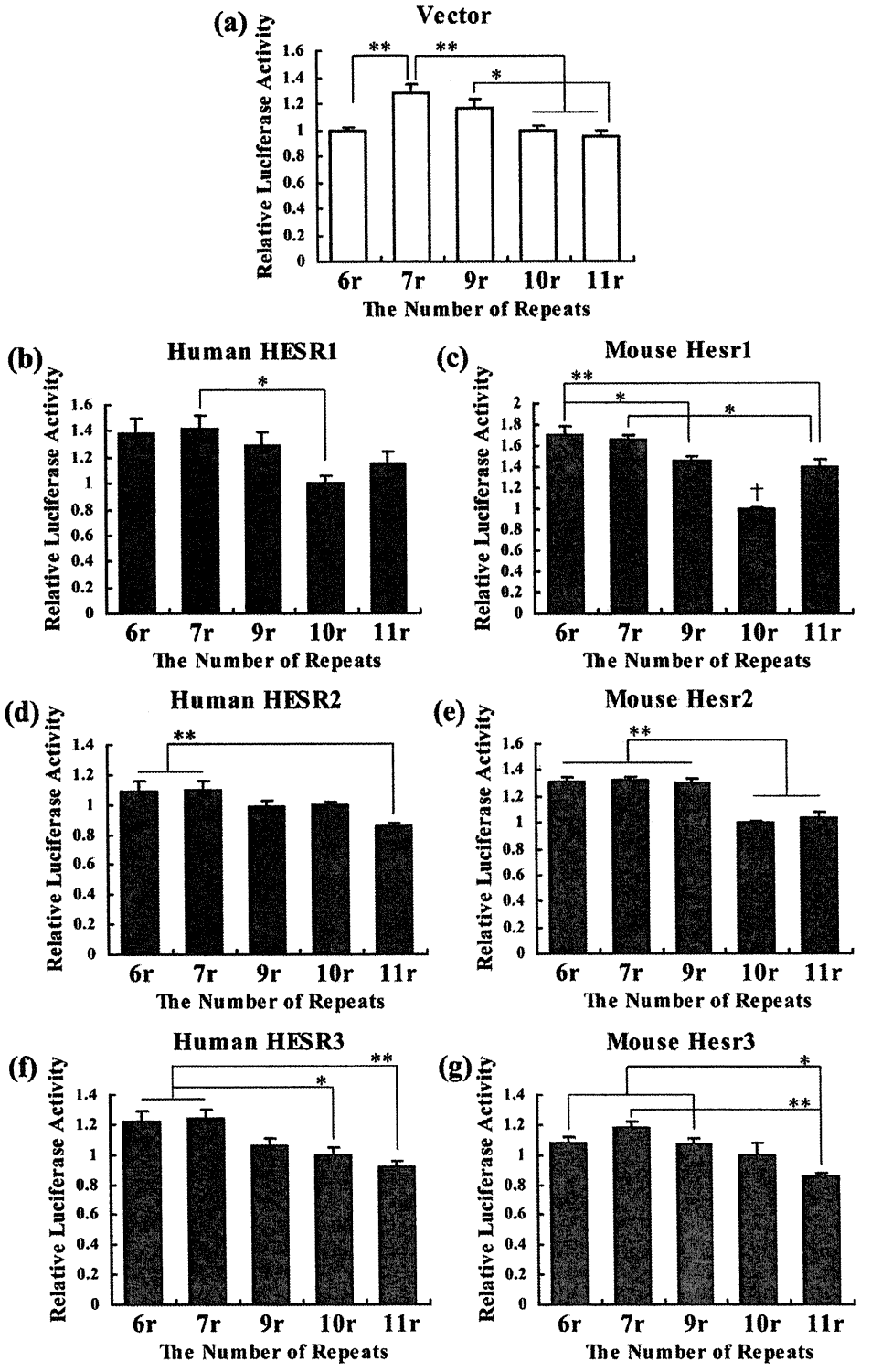

Fig. 6. Differential effects of the HESR family on the VNTR alleles. Cells were transfected with CP-Luc/exon 15 containing the indicated number of repeats and human (**b,d,f**) or mouse Hesrs (**c,e,g**) or with empty vector (**a**). Relative luciferase activity is expressed as the mean ± SEM. of relative luciferase activity of the reporter construct in each group. $*P < 0.05$, $**P < 0.01$, $\dagger P < 0.01$ vs. the other repeats, Tukey-Kramer's HSD test after one-way ANOVA.

With the expression of human HESR family members, the luciferase activity of CP-Luc relative to the vector control (1.00 ± 0.04) was as follows: HESR1, 0.36 ± 0.01; HESR2, 0.74 ± 0.02; and HESR3, 1.44 ± 0.06. The activity of CP-Luc with HESR1 or 2 was significantly lower than that with the vector control ($P < 0.0001$, n = 8 in each group), whereas that with HESR3 was higher than the control ($P < 0.0001$, n = 8 in each group).

With the expression of mouse Hesr family members (Fig. 4B), the luciferase activity of CP-Luc relative to the vector control (1.00 ± 0.04) was as follows: Hesr1, 0.39 ± 0.01; Hesr2, 0.60 ± 0.01; and Hesr3, 0.79 ± 0.01. The activity of CP-Luc with Hesr1 ($P < 0.0001$), -2 ($P < 0.0001$), or -3 ($P < 0.0003$) was significantly lower than with the vector (n = 8 in each group). These results indicate that only human HESR3 stimulated the expression of the CP-Luc construct containing the human *DAT1* core promoter, whereas expression of the other family members inhibited expression.

Effects of Human HESR and Mouse Hesr Family on Lusciferase Activity of CP-Luc/Exon 15 (10r)

To investigate the role of the VNTR domain, we cotransfected HESR genes with CP-Luc/exon 15 (10r). With the expression of human HESR family members (Fig. 5A), the luciferase activity of the CP-Luc/exon 15 (10r) relative to the vector control (1.00 ± 0.03) was as follows: HESR1, 0.21 ± 0.01; HESR2, 0.42 ± 0.01; and HESR3, 0.90 ± 0.04. These results illustrate that the luciferase activity of CP-Luc/exon 15 (10r) with HESR1 or 2 was significantly lower than with the vector ($P < 0.0001$, n = 8 in each group), although no significant decrease was observed with HESR3 ($P = 0.07$, n = 8).

The same trend was observed for mouse Hesr (Fig. 5B). With the expression of mouse Hesr family proteins, the luciferase activity of CP-Luc/exon 15 (10r) with each Hesr relative to the vector control (1.00 ± 0.07) was as follows: Hesr1, 0.25 ± 0.01; Hesr2, 0.51 ± 0.01; and Hesr3, 0.83 ± 0.09. The level of luciferase activity with Hesr1 or -2 was significantly lower than that with the vector ($P < 0.0001$, n = 8 in each group), whereas no significant decrease in activity was detected with Hesr3 ($P = 0.16$, n = 7).

Comparison of Effects of HESR Family and VNTR on Lusciferase Activity of CP-Luc/Exon15 (nr)

We investigated the effect of human and mouse HESR family members on the VNTR domain by first evaluating the effect of the repeat number. Luciferase expression from the CP-Luc/exon 15 construct is shown in Figure 6a. The following levels of activity were calculated: 6r, 0.99 ± 0.02; 7r, 1.28 ± 0.06; 9r, 1.16 ± 0.07; 10r, 1.00 ± 0.03; and 11r, 0.95 ± 0.04. One-way ANOVA indicated that there was a significant effect of the number of repeats ($F_{4,35} = 8.0167$, $P < 0.0001$). Post hoc analysis indicated that the luciferase activity level associated with 7r was significantly higher than that with 6, 10, or 11r ($P < 0.01$) and that the level of activity associated with 9r was significantly higher than that associated with 11r ($P < 0.05$; n = 8 in each group; Fig. 6a).

The effect of human HESR1 on the VNTR repeat number is shown in Figure 6b. With the expression of HESR1, CP-Luc/exon 15 activity (6r, 1.38 ± 0.11; 7r, 1.41 ± 0.11; 9r, 1.29 ± 0.11; 10r, 1.00 ± 0.06; and 11r, 1.14 ± 0.10) was compared with that for the vector alone (6r, 4.69 ± 0.11; 7r, 6.06 ± 0.28; 9r, 5.48 ± 0.35; 10r, 4.72 ± 0.16; and 11r, 4.50 ± 0.19). One-way ANOVA indicated a significant effect of the number of repeats ($F_{4,34} = 3.006$, $P < 0.032$). Post hoc analysis indicated that the luciferase activity associated with 7r was significantly higher than that associated with 10r (n = 7 for 11r, n = 8 in the other groups; Fig. 6b). In addition, two-way ANOVA indicated that there was a significant interaction between the number of repeats with HESR1 (or vector) ($F_{4,69} = 5.04$, $P < 0.002$).

We also investigated the effects of mouse Hesr1 (Fig. 6c). With the expression of Hesr1, CP-Luc/exon 15 activity (6r, 1.70 ± 0.08; 7r, 1.65 ± 0.05; 9r, 1.46 ± 0.04; 10r, 1.00 ± 0.02; and 11r, 1.41 ± 0.06) was compared with the control vector (6r, 3.90 ± 0.09; 7r, 5.04 ± 0.23; 9r, 4.56 ± 0.29; 10r, 3.92 ± 0.13; and 11r, 3.74 ± 0.16). As shown for human HERS1, one-way ANOVA indicated that there was a significant effect of the number of repeats ($F_{4,35} = 25.7587$, $P < 0.0001$). Post hoc analysis indicated that the level of luciferase activity associated with 6r was significantly higher than that associated with 9r ($P < 0.05$) or 11r ($P < 0.01$), and that the level for 7r was also higher than that for 11r (0.05). Furthermore, the lowest level of activity was observed for 10r ($P < 0.01$, n = 8 in each group; Fig. 6c). In addition, two-way ANOVA indicated that there was a significant interaction between the number of repeats with Hesr1 or the control vector ($F_{4,70} = 6.34$, $P < 0.002$).

We also investigated the effects of human HESR2 and mouse Hesr2. With the expression of human HERS2, CP-Luc/exon 15 activity (6r, 1.08 ± 0.07; 7r, 1.10 ± 0.06; 9r, 0.99 ± 0.04; 10r, 1.00 ± 0.02; and 11r, 0.85 ± 0.02) was again compared with the control vector (6r, 2.36 ± 0.05; 7r, 3.05 ± 0.14; 9r, 2.76 ± 0.17; 10r, 2.36 ± 0.08; and 11r, 2.26 ± 0.10) (Fig. 6d). One-way ANOVA indicated that there was a significant effect of the number of repeats ($F_{4,35} = 4.6640$, $P < 0.004$). Post hoc analysis indicated that the level of luciferase activity associated with 6 or 7r was significantly higher than that associated with 11r ($P < 0.01$, n = 8 in each group; Fig. 6d). In addition, two-way ANOVA indicated that there was a significant interaction between the number of repeats with HESR2 (or the control vector) ($F_{4,70} = 5.29$, $P < 0.0009$).

With the expression of mouse Hesr2 (Fig. 6e), CP-Luc/exon15 activity (6r, 1.32 ± 0.02; 7r, 1.32 ± 0.03; 9r, 1.30 ± 0.03; 10r, 1.00 ± 0.01; and 11r, 1.04 ± 0.04) was again compared with the control vector (6r, 1.94 ± 0.04; 7r, 2.51 ± 0.11; 9r, 2.26 ± 0.14; 10r,

1.95 ± 0.07; and 11r, 1.86 ± 0.08). One-way ANOVA indicated that there was a significant effect of the number of repeats ($F_{4,35} = 31.6769$, $P < 0.0001$). Post hoc analysis indicated that the level of luciferase activity for 6, 7, or 9r was significantly higher than that for 10 or 11r ($P < 0.01$, n = 8 in each group; Fig. 6e). In addition, two-way ANOVA indicated that there was a significant interaction between the number of repeats with Hesr2 (or the control vector; $F4, 70 = 4.19$, $P < 0.0045$).

Human HESR3 and mouse Hesr3 were also investigated. With human HESR3 expression, CP-Luc/exon 15 activity (6r, 1.22 ± 0.07; 7r, 1.24 ± 0.06; 9r, 1.06 ± 0.05; 10r, 1.00 ± 0.04; and 11r, 0.92 ± 0.04) was compared with the control vector (6r, 1.11 ± 0.03; 7r, 1.43 ± 0.07; 9r, 1.29 ± 0.08; 10r, 1.12 ± 0.04; and 11r, 1.06 ± 0.05). One-way ANOVA again indicated a significant effect of the number of repeats ($F_{4,35} = 7.0211$, $P < 0.0003$). Post hoc analysis indicated that the level of luciferase activity using 6r or 7r was significantly higher than that using 10 ($P < 0.05$) or 11r ($P < 0.01$; n = 8 in each group; Fig. 6f). In addition, two-way ANOVA indicated that there was a significant interaction between the number of repeats with HESR3 or control vector ($F_{4,70} = 3.14$, $P < 0.02$).

With the expression of mouse Hesr3, CP-Luc/exon 15 activity (6r, 1.08 ± 0.04; 7r, 1.18 ± 0.04; 9r, 1.07 ± 0.04; 10r, 1.00 ± 0.08; and 11r, 0.85 ± 0.03) was again compared with that of the control vector (6r, 1.15 ± 0.07; 7r, 1.44 ± 0.07; 9r, 1.28 ± 0.05; 10r, 1.20 ± 0.10; and 11r, 1.07 ± 0.05). One-way ANOVA indicated that there was a significant effect of the number of repeats ($F_{4,30} = 6.4123$, $P < 0.0007$). Post hoc analysis indicated that the level of luciferase activity for 11r was significantly lower than that for 6 ($P < 0.05$), 7 ($P < 0.01$), and 9r ($P < 0.05$; n = 7 in each group; Fig. 6g). In addition, two-way ANOVA indicated no significant interaction between the number of repeats with Hesr3 or the control vector ($F_{4,65} = 0.67$, $P = 0.62$). The value of luciferase expression from the each CP-Luc/exon 15 construct (6, 7, 9, 10, and 11r) with empty vector compared in Figure 6b–g was measured from same samples as used in Figure 6a.

Localization of Hesr Family Members in Mouse Midbrain

We next analyzed the localization of Hesr1, -2, and -3 by immunoperoxidase staining. Hesr1-, Hesr2-, and Hesr3-ir cells were observed in both dopaminergic and nondopaminergic regions in the mouse midbrain (Fig. 7a). Immunoreactivity against Hesr1 and -2 was observed mainly in the nucleus, whereas Hesr3 was detected in the cytoplasm (Fig. 7b). However, in the nondopaminergic regions, immunoreactivity against Hesr1 and -2 was detected in the nucleus and cytoplasm (Fig. 7c). We also analyzed the localization of Hesr family members in the TH-ir cells using double-fluorescence immunostaining. Hesr1-, Hesr2-, and Hesr3-TH-ir cells were observed in both the SN and the VTA (Fig. 8).

DISCUSSION

Roles of the 3′-UTR Including VNTR Domain for *DAT* Expression

The relative luciferase activity of CP-Luc was significantly higher than that of CP-Luc/exon 15 (10r) in SH-SY5Y cells (Fig. 3). This suggests that the 3′-UTR plays an inhibitory role in *DAT1* expression. It is possible that endogenous factors affect *DAT1* expression via the 3′-UTR; in fact, it is predicted that more than one factor binds to this region (Michelhaugh et al., 2001). Thus, we investigated the HESR family as novel candidate regulatory factors modulating *DAT1* expression.

Effects of HESR Family on the *DAT* Core Promoter

Human HESR1 and -2 or mouse Hesr1, -2, and -3 significantly decreased the relative luciferase activity level of CP-Luc containing the *DAT1* core promoter, whereas human HESR3 increased CP-Luc activity (Fig. 4). These results were unexpected, insofar as HESR1 has been identified as a 3′-UTR-binding protein. This suggests that the HESR family generally down-regulates *DAT* expression through the core promoter region. In particular, HESR1 showed strong inhibitory effects in this region, with a 36% decrease in luciferase activity with human HESR1 expression and a 39% decrease with mouse Hesr1. In fact, the core promoter region has an E-box known to be bHLH-binding consensus sites. Moreover, it has been reported that mouse *Hesr2* does not contain an E- or N-box and is repressed by Hesr proteins (Nakagawa et al., 2000). Therefore, it is likely that HESR family members recognize a binding site different from that recognized by other bHLH family members. In addition, a functional single nucleotide polymorphism (SNP; −67 A/T) in the promoter was reported to be associated with personality traits, ADHD, and bipolar disorder (Greenwood and Kelsoe, 2003; Ohadi et al., 2006, 2007; Shibuya et al., 2009). This SNP may be a point of interaction with HESR family members.

Human HESR3 was the only HESR family member that significantly increased luciferase activity of CP-Luc containing the *DAT1* core promoter. We previously reported that HESR1 containing the Leu94Met SNP in the second helix of the bHLH domain lacked inhibitory activity (Fuke et al., 2005). It was also demonstrated that this SNP dramatically transforms HESR1 from an androgen receptor corepressor to an activator (Villaronga et al., 2010). HESR3 is the only HESR with a single-amino-acid substitution adjacent to the Leu in HESR1 and -2 located in the second helix of the bHLH domain (Fig. 2). Thus, the second helix in HESR family members may be critical in the modulation of gene expression.

Effects of HESR Family on the 3′-UTR of DAT

As shown in Figure 5, human and mouse HESR1 and -2 significantly inhibited the luciferase activity of CP-Luc/exon 15 (10r). Taken together with the results shown in Figure 4, this result indicates that HESR1 and

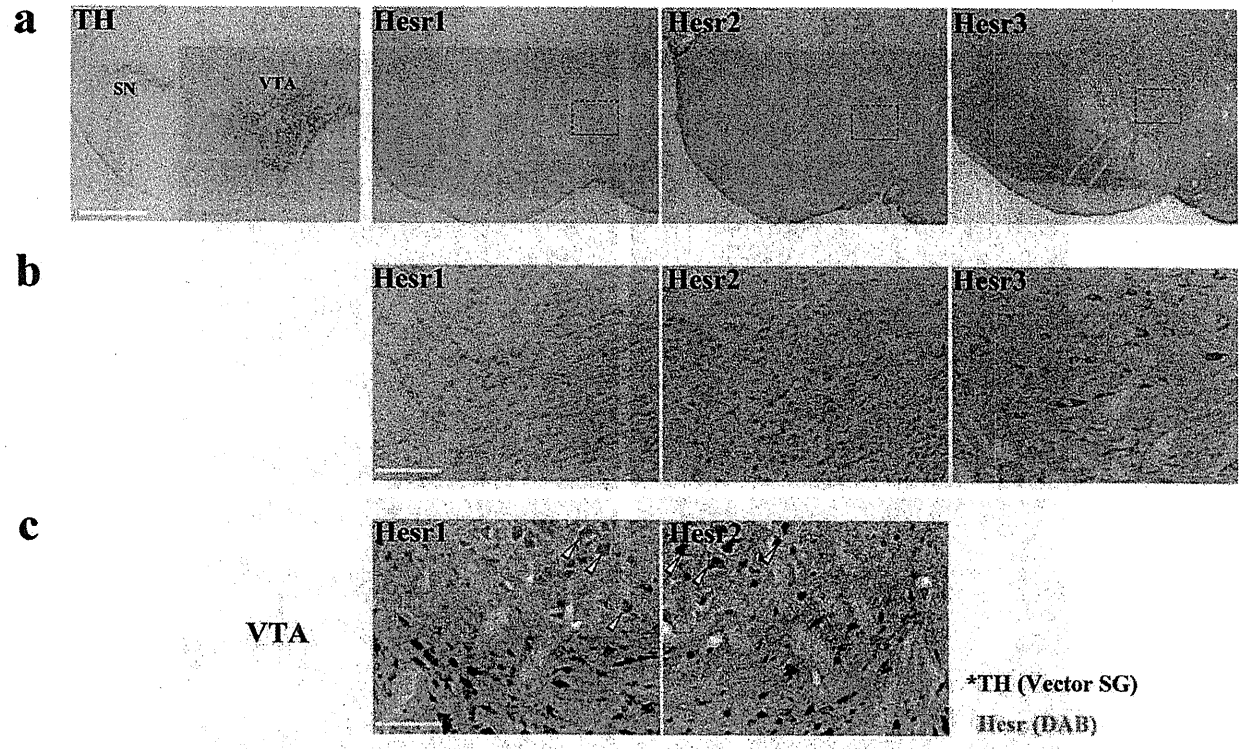

Fig. 7. Distribution of the Hesr family in mouse midbrain. Photomicrographs showing immunoperoxidase staining visualized by DAB (brown) or Vector SG (blue/gray). **a:** Localization of tyrosine hydroxylase (TH) or each Hesr family member in mouse midbrain, including the substantia nigra (SN) and ventral tegmental area (VTA). **b:** High-magnification images of the boxed areas in a. Hesr1 and -2 are localized primarily in the nucleus, whereas Hesr3 is located predominantly in the cytoplasm. **c:** Photomicrographs of the VTA and surrounding area depicting dual labeling for TH (Vector SG) and Hesr1 or 2 (DAB). Arrowheads indicate cells in which Hesr1 or -2 immunoreactivity was observed in both the nucleus and the cytoplasm. Scale bars = 500 μm in a; 100 μm in b,c.

-2 inhibit *DAT1* expression through both the core promoter and the 3′-UTR. HESR1 and -2 decreased the relative level of luciferase activity to less than 25% (HESR1) and 50% (HESR2). This degree of decrease is relatively high compared with that seen with the CP-Luc. Thus, HESR1 and -2 may have a stronger inhibitory effect on *DAT1* expression in the presence of the 3′-UTR.

We previously showed that HESR1 bound the 3′-UTR of *DAT1* directly by electromobility shift assays (Fuke et al., 2006). Because the bHLH domain is highly conserved (Steidl et al., 2000), we investigated the effect of other HESR family members to determine whether they affect gene expression by interacting with this same region. It has been proposed that more than one factor binds to this region and that the 3′-UTR, including the VNTR domain, modulates gene expression (Michelhaugh et al., 2001). Functional VNTR polymorphisms also exist in the serotonin transporter (*SERT*) gene located in intron 2, and two transcription factors, Y box-binding protein 1 (YB-1) and CCTC-binding factor (CTCF), were found to be responsible for the modulation of VNTR function (Klenova et al., 2004). This suggests that the VNTR domain functions as a modulator of gene expression (Nakamura et al., 1998) with other binding proteins.

Comparison of *DAT* Expression With the VNTR: Differential Effects of HESR Family on Each VNTR Allele

The number of repeats in the VNTR domain significantly affected the level of luciferase activity of CP-Luc/exon 15 in SH-SY5Y cells (Fig. 6a). This suggests that the VNTR domain in the *DAT* 3′-UTR is the functional sequence for *DAT* expression and is supported by in vivo neuroimaging involving SPECT (Heinz et al., 2000; Jacobsen et al., 2000; Martinez et al., 2001) and ex vivo RT-PCR analysis (Mill et al., 2002; Brookes et al., 2007). When the HESR family was transiently transfected, the number of repeats in the VNTR domain significantly affected the luciferase activity of CP-Luc/exon 15 (Fig. 6b–g). In addition, interactive effects between the number of repeats and the HESR or control vector were detected for all groups except mouse Hesr3. This suggests that the HESR