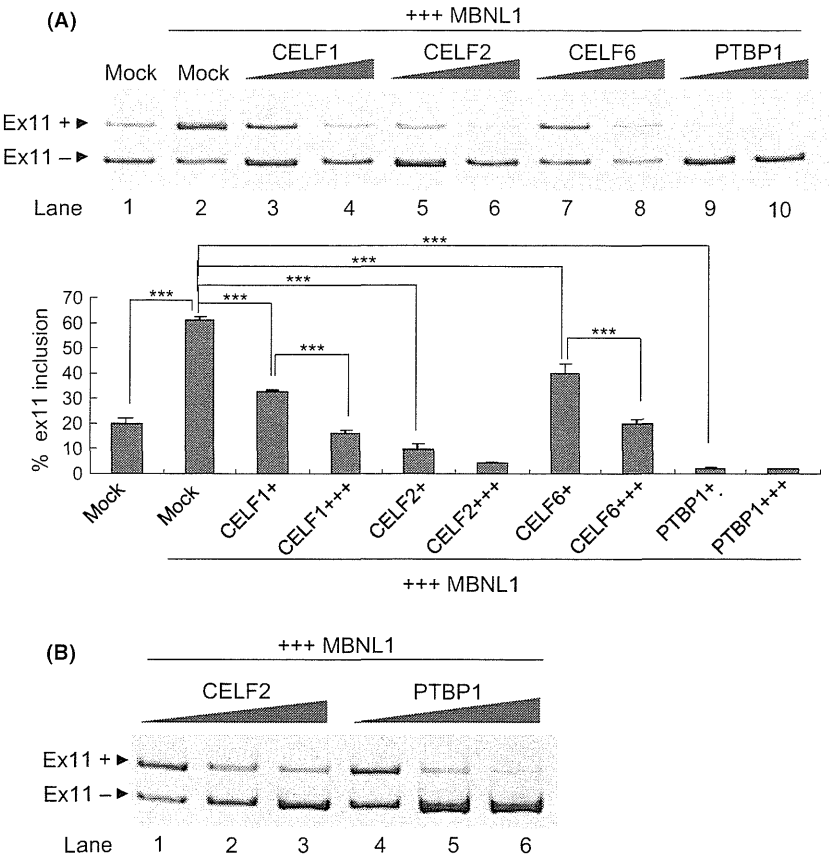


**Figure 4** MBNL1 is antagonized by CELF1, CELF2, CELF6 and PTBP1. (A) Antagonistic effects of CELF1, CELF2, CELF6 and PTBP1 on the regulation of *ABLIM1* splicing by MBNL1. The *ABLIM1* minigene (0.15 μg) was cotransfected with MBNL1 (EXP42) or MBNL1 plus the CELF family or PTBP1 into C2C12 cells. Upper panel: lane 1, mock (pSecDK) transfected (1.4 μg); lane 2–10, a constant amount of the MBNL1 construct (0.7 μg) and increasing amounts of CELF1, CELF2, CELF6 or PTBP1 constructs were cotransfected at ratios of 3 : 0 (lane 2), 3 : 1 (lane 3, 5, 7, and 9) and 3 : 3 (lane 4, 6, 8, and 10). The total amount of transfected plasmids (1.4 μg) was kept constant by adjusting the amount of mock control vector. The lower panel shows the percentage of exon 11 inclusion relative to the total transcripts (means ± SEM, $n = 3$). Statistical significance was assessed by ANOVA and Tukey's multiple comparison test (***$P < 0.001$). (B) Effects of CELF2 and PTBP1 on *ABLIM1* splicing. The *ABLIM1* minigene (0.15 μg) was cotransfected with MBNL1 (EXP42), MBNL1 plus CELF2 or MBNL1 plus PTBP1 into C2C12 cells. Lane 1–6, a constant amount of MBNL1 construct (0.7 μg) and increasing amounts of CELF2 or PTBP1 constructs were cotransfected at a ratio of 3 : 0.09 (lane 1, 4), 3 : 0.3 (lane 2, 5) and 3 : 0.6 (lane 3, 6). The total amount of transfected plasmids (1.4 μg) was kept constant by adjusting the amount of mock control.

isoform) was detected in skeletal muscles of non-DM individuals, but not in most patients with DM1 (Fig. 1A–C). Moreover, aberrant splicing of mouse *Ablim1* exon 11 also occurred in skeletal muscle of a DM mouse model, $HSA^{LR}$, which expresses expanded CUG repeats in skeletal muscle (Fig. 6A). Therefore, *ABLIM1* ex11+ may have a muscle-specific function, and the deficiency of *ABLIM1* ex11+ in patients with DM might be involved in muscle symptoms. The ABLIM1 protein contains four LIM domains at its N-terminus and a villin headpiece domain at its C-terminus. Four splicing isoforms in humans and three splicing isoforms in mouse are known (Fig. S8 in Supporting Information). ABLIM1 binds to filamentous actin (F-actin) via its villin headpiece domain and is localized to the Z-disk in cardiac muscle (Roof *et al.* 1997). Thus, it is assumed that ABLIM1 could mediate the interaction between actin filaments and a protein complex at a Z-disk. Exon 11 (84 nt) of human *ABLIM1* codes for 28 amino acids (aa), located 5 amino acid downstream of the fourth LIM domain, which mediates protein–protein

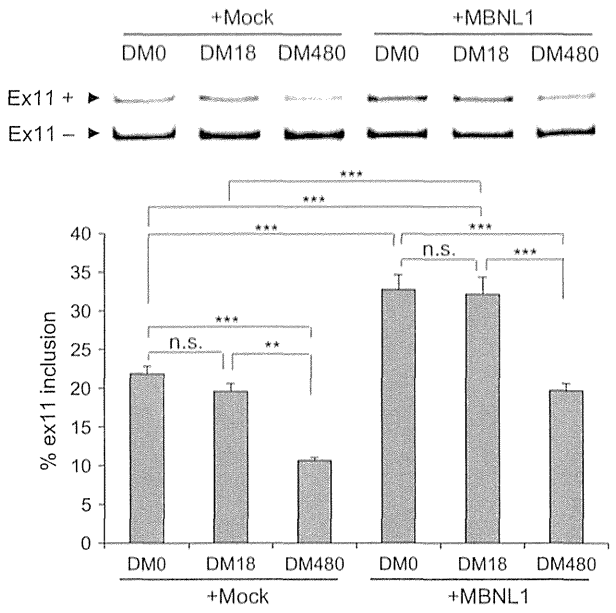


**Figure 5** Expanded CUG repeat RNAs affect *ABLIM1* splicing, and the effect is antagonized by MBNL1 in C2C12 cells. The *ABLIM1* minigene (0.15 μg) and mock (pSecDK) or MBNL1 (EXP42) construct (0.7 μg) were cotransfected with DM0, DM18 or DM480 (0.7 μg) construct into C2C12 cells. The lower panel shows the percentage of exon 11 inclusion relative to the total transcripts (means ± SEM, $n = 3$). Statistical analyses were carried out by ANOVA and Tukey's multiple comparison test by comparison with the normal (DM18) CTG repeat (**$P < 0.01$,***$P < 0.001$).

interactions (Schmeichel & Beckerle 1994). As the exclusion of exon 11, which predominantly occurs in DM1 muscle and nonmuscle tissues in normal adults, does not result in a frame shift or non-sense-mediated mRNA decay (NMD), the ex11− isoform of the ABLIM1 protein should be expressed in DM1 muscle. It is possible that the exclusion of exon 11 may affect intercellular localization of ABLIM1 by different protein–protein interactions, although no ABLIM1-interacting protein via its four LIM domains or exon 11 coding region has yet been identified. Additional studies are required to show the relationship between abnormal splicing of ABLIM1 and DM symptoms.

The splicing of *ABLIM1* exon 11 is antagonistically regulated by MBNL and CELF proteins. All three MBNL proteins promoted the inclusion of exon 11 of endogenous and the *ABLIM1* minigene in cultured cells (Figs 2, 3). Both MBNL1 and MBNL2 promoted the inclusion of endogenous *ABLIM1* exon 11 at a similar level, but MBNL2 did not promote that of the *ABLIM1* minigene to the same degree as MBNL1. We believe this is because internal regions of intron 10 and intron 11 of *ABLIM1* were excluded in the *ABLIM1* minigene (Fig. 3A). The sequence interacting with MBNL2 may be located in the deleted part. Another possibility is that MBNL2 protein expression in the minigene splicing assay with C2C12 cells was lower than that observed in endogenous *ABLIM1* splicing assay with HEK293 cells (Fig. 3B). On the other hands, CELF1, 2 and 6, and PTBP1 promoted the exclusion of the *ABLIM1* minigene exon 11 (Fig. 3B). As the ex11+ isoform was nearly undetectable in HEK293 (Fig. 2), HeLa, SH-SY5Y and undifferentiated C2C12 (data not shown) cells, it was impossible to assess the effect of CELF1, 2 and 6, and PTBP1 on endogenous *ABLIM1*. Although the inclusion of *ABLIM1* exon 11 was repressed by over-expression of CELF1, splicing of *ABLIM1* exon 11 was not changed by knockdown of Celf1 in our experiments (Fig. 3D). We presume that the high level of Celf1 is needed for exclusion of *ABLIM1* exon 11 and that endogenous Celf1 in C2C12 cells may be below this level. The abnormality of *Ablim1* splicing was also detected in MBNL knockout mice and CELF1-over-expressing mice (Kalsotra *et al.* 2008). Therefore, we conclude that MBNL and CELF regulate *ABLIM1* splicing. The results of Celf2 knockdown experiments were different from siCelf2 (No. 11) and siCelf2 (No. 12; Fig. 3D), although the degrees of Celf2 knockdown at a protein level by those siCelf2s were almost same (Fig. S6 in Supporting Information). This might be because siCelf2 affected other splicing factors by off-targets effect. Because at least siCelf2 (No. 11) significantly promoted the inclusion of *ABLIM1* ex11 compared to control, we suggest the possibility that endogenous Celf2 in C2C12 suppresses ex11 inclusion.

PTBP1 promoted exclusion of *ABLIM1* exon 11 more strongly than CELF1, 2 and 6. Therefore, we suspected that expression of PTBP1 may be increased in patients with DM. RT-PCR analysis showed that the expression of PTBP1 in patients with DM1 is similar to that of non-DM1 individuals at the RNA level (Fig. S5 in Supporting Information). However, it is still possible that PTBP1 protein is activated in DM1, similar to CELF1, which is active in DM1 following stabilization by phosphorylation, but the expression of CELF1 does not change at the RNA level. In addition, the level of PTBP1 expression appears to be correlated with *ABLIM1* exon 11 splicing because Ptbp1 expression decreases gradually

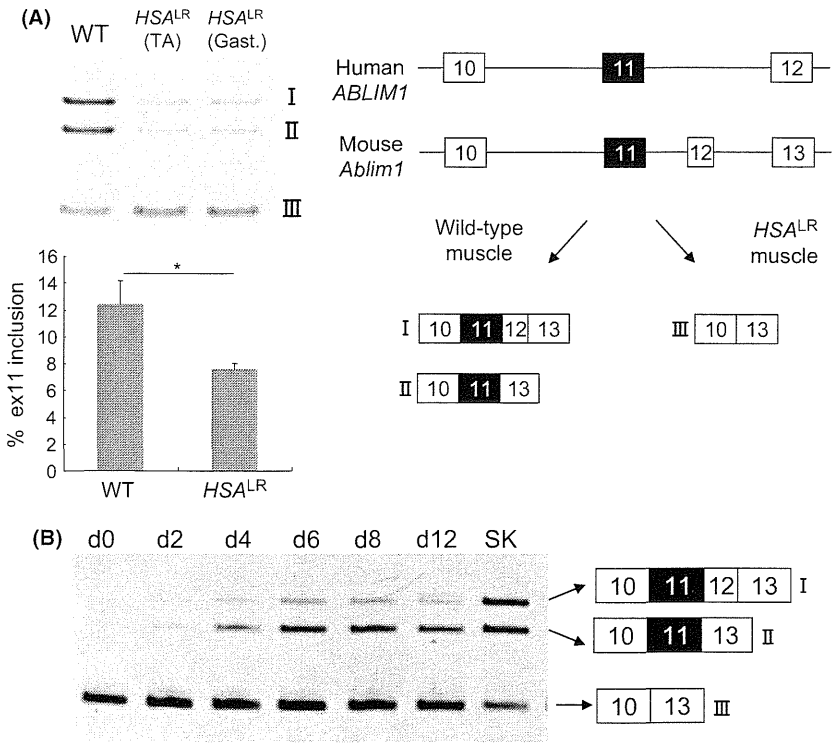


**Figure 6** Mouse *ABLIM1* splicing. (A) Expanded CUG repeat RNAs affect mouse *ABLIM1* splicing *in vivo*. $HSA^{LR}$ mice have 250 expanded CTG repeats in the 3′-UTR of the human skeletal muscle actin gene and recapitulate DM1 symptoms. Mouse *ABLIM1* exon 11 is homologous to the human *ABLIM1* exon 11, but mouse *ABLIM1* exon 12 is not homologous to any exon of human *ABLIM1*. Isoforms of *ABLIM1*-I (exons 11 and 12 inclusion isoform) and *ABLIM1*-II (exon 11 inclusion isoform) are predominant in skeletal muscle of wild-type mice compared to *ABLIM1*-III (exons 11 and 12 exclusion isoform). In contrast, *ABLIM1*-III is predominant in $HSA^{LR}$ mice. TA, *tibialis anterior*; Gast., *gastrochemius*. Exon 11 inclusion was calculated as (I + II)/(I + II + III) (means ± SEM, $n = 4$). (B) The splicing pattern of *ABLIM1* changes during C2C12 differentiation. RT-PCR of endogenous *ABLIM1* was carried out using C2C12 cells after induction of differentiation. The growth medium was replaced with differentiation medium on day 0 (d0). D2, 4, 6, 8 and 12 indicate the day after differentiation. SK indicates adult skeletal muscle of wild-type mice. *ABLIM1*-III is predominant at d0 and d2, but *ABLIM1*-I and *ABLIM1*-II are gradually increased after d4.

during C2C12 differentiation, whereas that of Mbnl1 and Fox1 increase (Bland *et al.* 2010) when *Ablim1* exon 11 inclusion increases (Fig. 6B). In other words, when PTBP1 protein expression is high, the exon 11 exclusion isoform predominates, like DM1. In contrast, it was reported that PTBP1 decreased the *TNNT2* exon 5 inclusion isoform, which is expressed mainly in the heart and skeletal muscle of patients with DM1 (Philips *et al.* 1998; Charlet *et al.* 2002a). Therefore, the role of PTBP1 in DM1 muscle remains elusive, and aberrant splicing of *ABLIM1* and other genes may be caused by an imbalance in those splicing factors.

In this study, we identified an abnormal ABLIM1 exon 11 splice variant in skeletal muscle of patients with DM1 and $HSA^{LR}$ mice (a DM1 mouse model). ABLIM1 protein is located on the Z-disk, a structure important for mechanical stability and intracellular signaling in muscle cells. The pathomechanism of muscle symptoms, such as muscle weakness and atrophy, in DM1 has not been elucidated. This study may clarify the pathological mechanism of DM1.

## Experimental procedures

### Human skeletal muscle biopsies and mouse samples

Biopsies were obtained from the biceps brachii or quadriceps femoris muscle of six patients with DM1 and seven non-DM individuals without muscular disease (Table S1 in Supporting Information). Four non-DM individuals did not show any histological abnormalities, and three non-DM individuals showed mild atrophy or only type 2 fiber atrophy. Hematoxylin and Eosin staining (H&E staining) of biopsies from a non-DM and patients with DM1 is shown in (Fig. S9 in Supporting Information). All biopsies were stored at −80 °C. Clinically, all

patients with DM1 had muscle weakness with myotonia. Four DM1 patients had congenital onset, and two DM1 patients' onset was during childhood or adolescence. Pathologically, all patients with DM1 showed immature fiber type or myopathic changes with variation in fiber size. All biopsies were acquired with informed consent. Mouse samples were obtained from the tibialis anterior (TA) muscle of $HSA^{LR}$ mice and FVB mice (WT), which were 28–33 weeks old.

## Splice analysis of human tissue samples and mouse muscle

Total RNAs were isolated from human muscle biopsies and mouse TA muscle using TRIzol reagent (Invitrogen) according to the manufacturer's protocol without DNase treatment. Human total RNA master panel II (Clontech) was used for the analysis of other human tissues. All total RNAs were stored at −80 °C. cDNAs were synthesized using a PrimeScript first-strand cDNA synthesis lit (TaKaRa) in 10-μL total volume with oligo dT primers and total RNA (0.5 μg for human biopsies, 1.0 μg for other human tissue samples and 0.2 μg for mouse samples). cDNA of fetal skeletal muscle (BioChain) was synthesized from total RNA of one donor (male, 20 weeks old). All cDNAs were stored at −20 °C. PCR was carried out using Ex Taq DNA polymerase (TaKaRa) according to the manufacturer's protocol. Primer sequences and annealing temperatures are listed in Table S2 (Supporting Information). PCR conditions were as follows: an initial denaturation step (96 °C for 2 min), quantitative cycles (96 °C for 30 s, annealing temperature for 30 s and 72 °C for 1 min) and a final extension step (72 °C for 5 min). The cycle numbers were adjusted so that the amplification was within the logarithmic phase. PCR products were resolved by electrophoresis on an 8% polyacrylamide gel strained with ethidium bromide and analyzed using a LAS3000 imager (FUJIFILM). The intensity of band signals was quantified using the MULTIGAUGE software (FUJIFILM). The ratio of exon 11 inclusion of human *ABLIM1* and mouse *Ablim1* was calculated as follows: (ex11 inclusion)/(ex11 inclusion + ex11 exclusion) and (ex11 inclusion + ex11 and ex12 inclusion)/(ex11 inclusion + ex11 and ex12 inclusion + ex11 exclusion), respectively. Mean and *P* values were determined using Student's *t*-test. PCR products were cloned into pGEM-T Easy vector (Promega) and sequenced.

## Construction of *ABLIM1* minigenes, splice factors and DMPK constructs, and siRNAs

The genomic sequence of human *ABLIM1* from exon 10 to exon 12 is too large to make a minigene construct. Therefore, we constructed the minigenes containing relevant regulatory elements of the alternative splice variant. Three genomic fragments were amplified using *Pfu Ultra* high-fidelity DNA polymerase (Stratagene) from human genomic DNA (Promega). The first fragment contained exon 10 and 343 nucleotides downstream of the exon; the second fragment contained exon 11 and 343 nucleotides upstream and downstream of the regulated exon; and the third contained exon 12 and 356 nucleotides upstream of the exon. The three fragments were amplified by PCR with primers containing restriction enzyme sites at their 5′ ends. The following primer sets were used for PCR amplification: first fragment, forward (5′-GAAGATCTGGCTCCACCGTTTGGC-3′) and reverse (5′-CGAGCTCCAAAACAGTCTGGTGAGGTC-3′); second fragment, forward (5′-CGAGCTCGGGGCATTTGTAACAACAGATT-3′) and reverse (5′-GGAATTCTTCTTGCCTATGAGGCTGGATC-3′); and third fragment, forward (5′-GGAATTCGGAGGCCATTGCAATAATCT-3′) and reverse (5′-GCGGTACCGATACTTACATAGATAGTATGACCT-3′). Three fragments were inserted into the *Bgl*II- *Sac*I, *Sac*I- *Eco*RI and *Eco*RI- *Kpn*I sites of pEGFP-C1 (Clontech) one by one. The minigene was sequenced before use in the experiments. Splice factors MBNL1, MBNL2, MBNL3, CELF1, CELF2, CELF3, CELF4, CELF5, CELF6, FOX1 and PTBP1 were cloned as described previously into pSecDK, a mammalian expression vector with a myc-tag and 6×His-tag modified from pSecTagA (Invitrogen) by deleting the Igκ chain leader sequence (Kino *et al.* 2009). Expression of these constructs was confirmed by Western blotting using anti-Myc (1 : 5000, R950-25; Invitorogen), anti-His (1 : 2000, 34670; Qiagen) or real-time PCR. CTG repeat constructs, mRFP-DMPK3′-CTG18 (DM18) and mRFP-DMPK3′-CTG 480 (DM480), were generated as described previously (Kino *et al.* 2009). mRFP-DMPK3′-CTG0 (DM0) was made by performing inverted PCR from DM18 construct using primer set: forward (5′-AAAGTCGACGGGGGATCACAGACCATTTC-3′), reverse (5′-AAAGTCGACCATTCCCGGCTACAAGGACC -3′). Expression of these constructs was confirmed by fluorescence imaging [green fluorescence protein (GFP) and red fluorescence protein (RFP)].

For knockdown experiment of the splicing factor, we used the following stealth siRNAs (Invitrogen, the sense strand sequences are given): siMBNL1 (No. 82) (HSS142882), 5′-GCUCCAGGGAGAACUGCAAAUAUCU-3′; siMBNL1 (No. 83) (HSS142883), 5′-GCAGUUGGAGAUAAAUGGACGCAAU-3′; siCELF1 (No. 15) (HSS173815), 5′-CCACUCUGUACAACCAGAAUCUUCU-3′; siCELF1 (No. 16) (HSS173816), 5′-GGUCCAGAGGGAGCCAACCUGUUCA-3′; siCELF1 (No. 47) (HSS116447), 5′-GGACAGAUUGAAGAAUGCCGGAUAU-3′; siCelf2 (No. 11) (MSS204011), 5′-CCUUAUGGAGCUGUCUACCAGAUCA-3′; siCelf2 (No. 12) (MSS204012), 5′-GCUGGAGCCACUGUCGGAUUGAAUA-3′; siPtbp1 (No. 37) (MSS276537), 5′-GGUGUGGUCAAAGGCUUCAAGUUCU-3′; siPtbp1 (No. 38) (MSS276538), 5′-CAGUGCUUCGUGGACAGCCCAUCUA-3′. The knockdown was confirmed by Western blotting using anti-MBNL1 [1 : 1000 (3A4-1E9); Sigma], anti-CELF1 [1 : 1000 (3B1); Ribonomics], anti-CELF2 [1 : 200 (1H2); Santa Cruz) and anti-PTBP1 (1 : 250, 32-4800; Invitrogen). Anti-Actin (1 : 400, A2066; Sigma) was used as internal control (Fig. S6 in Supporting information).

### Cell culture, transfection and differentiation

HEK293 and C2C12 cells were grown in Dulbecco's modified Eagle's medium (DMEM) containing 10% and 20% fetal bovine serum (FBS), respectively. C2C12 cells were subcultured before reaching 80% confluency. Before transfection, HEK293 and C2C12 cells were plated at 90–95% and 75–80% confluency on 12-well plates, respectively. C2C12 cells for knockdown experiment were plated at 50% confluency on 12-well plates. For the endogenous splicing assay, splice factor constructs or empty vectors (1.6 μg) were transfected into HEK293 cells with 4-μL Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's protocol. In the *ABLIM1* minigene splicing assay, minigenes (0.15–0.6 μg) and expression constructs or cognate empty vectors (1.3–1.4 μg) or duplex RNAs (48 pmol) were transfected into C2C12 cells with 8-μL Lipofectamine 2000 reagent (Invitrogen). C2C12 myoblasts were differentiated into myotubes by replacing medium with DMEM containing 10% horse serum (differentiation medium) on 35-mm dishes. Differentiation medium was exchanged every 2 days.

### Cell splicing assay

Forty-eight hours after transfection, total RNAs were extracted using a GenElute mammalian total RNA miniprep kit (Sigma) according to the manufacturer's protocol. cDNA synthesis was carried out as described above with total RNA (2.5 μg for endogenous splicing assay, 0.4–1.5 μg for the minigene splicing assay). For the minigene splicing assay, PCR was carried out using the pEGFP primer set: forward (5′-CATGGTCCTGCTGGAGTTCGTG-3′) and reverse (5′-GCAAGTAAAACCTCTACAAATGTGG-3′). PCR conditions were as follows: an initial denaturation step (96 °C for 2 min), 22–24 cycles (96 °C for 30 s, 60 °C for 30 s and 72 °C for 1 min) and a final extension step (72 °C for 5 min). For the endogenous splicing assay, the same primer sets as those used for biopsies were used. PCR products were sequenced.

## Acknowledgements

*HSA*$^{LR}$ mice were kindly provided by Prof. C.A. Thornton and Prof. M. P. Takahashi. This work was supported in part by an intramural research grant (26–8) for Neurological and Psychiatric Disorders of NCNP, Ministry of Health, Labor and Welfare, Japan, and JSPS KAKENHI (grant no. 249583). We would like to thank Dr Noboru Sasagawa, Dr Yoko Oma, Dr Yoshihiro Kino, Miss Zhao Yimeng, Mr Kosuke Oana.

## References

Bland, C.S., Wang, E.T., Vu, A., David, M.P., Castle, J.C., Johnson, J.M., Burge, C.B. & Cooper, T.A. (2010) Global regulation of alternative splicing during myogenic differentiation. *Nucleic Acids Res.* **38**, 7651–7664.

Boutz, P.L., Chawla, G., Stoilov, P. & Black, D.L. (2007a) MicroRNAs regulate the expression of the alternative splicing factor nPTB during muscle development. *Genes Dev.* **21**, 71–84.

Boutz, P.L., Stoilov, P., Li, Q., Lin, C.H., Chawla, G., Ostrow, K., Shiue, L., Ares, M. Jr & Black, D.L. (2007b) A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes Dev.* **21**, 1636–1652.

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* **69**, 385.

Chan, R.C. & Black, D.L. (1997) The polypyrimidine tract binding protein binds upstream of neural cell-specific c-src exon N1 to repress the splicing of the intron downstream. *Mol. Cell. Biol.* **17**, 4667–4676.

Charlet, B.N., Logan, P., Singh, G. & Cooper, T.A. (2002a) Dynamic antagonism between ETR-3 and PTB regulates cell type-specific alternative splicing. *Mol. Cell* **9**, 649–658.

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

Cooper, T.A. (2005) Use of minigene systems to dissect alternative splicing elements. *Methods* **37**, 331–340.

Dansithong, W., Paul, S., Comai, L. & Reddy, S. (2005) MBNL1 is the primary determinant of focus formation and aberrant insulin receptor splicing in DM1. *J. Biol. Chem.* **280**, 5773–5780.

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.

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.

Frank, D., Kuhn, C., Katus, H.A. & Frey, N. (2006) The sarcomeric Z-disc: a nodal point in signalling and disease. *J. Mol. Med. (Berl.)* **84**, 446–468.

Frey, N., Barrientos, T., Shelton, J.M., Frank, D., Rutten, H., Gehring, D., Kuhn, C., Lutz, M., Rothermel, B., Bassel-Duby, R., Richardson, J.A., Katus, H.A., Hill, J.A. & Olson, E.N. (2004) Mice lacking calsarcin-1 are sensitized to calcineurin signaling and show accelerated cardiomyopathy in response to pathological biomechanical stress. *Nat. Med.* **10**, 1336–1343.

Fu, Y.H., Friedman, D.L., Richards, S., Pearlman, J.A., Gibbs, R.A., Pizzuti, A., Ashizawa, T., Perryman, M.B., Scarlato, G., Fenwick, R.G. Jr & Caskey, C.T. (1993) Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science* **260**, 235–238.

Fugier, C., Klein, A.F., Hammer, C. *et al.* (2011) Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nat. Med.* **17**, 720–725.

Gooding, C., Roberts, G.C. & Smith, C.W. (1998) Role of an inhibitory pyrimidine element and polypyrimidine tract binding protein in repression of a regulated alpha-tropomyosin exon. *RNA* **4**, 85–100.

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.

Harper, P. (2001) *Myotonic Dystrophy*, 3rd ed. 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, 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.

Jin, Y., Suzuki, H., Maegawa, S., Endo, H., Sugano, S., Hashimoto, K., Yasuda, K. & Inoue, K. (2003) A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG. *EMBO J.* **22**, 905–912.

Jones, K., Wei, C., Iakova, P., Bugiardini, E., Schneider-Gold, C., Meola, G., Woodgett, J., Killian, J., Timchenko, N.A. & Timchenko, L.T. (2012) GSK3beta mediates muscle pathology in myotonic dystrophy. *J. Clin. Investig.* **122**, 4461–4472.

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 $Ca^{2+}$-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.

Koebis, M., Ohsawa, N., Kino, Y., Sasagawa, N., Nishino, I. & Ishiura, S. (2010) Alternative splicing of myomesin 1 gene is aberrantly regulated in myotonic dystrophy type 1. *Genes Cells* **16**, 961–972.

Koshelev, M., Sarma, S., Price, R.E., Wehrens, X.H. & Cooper, T.A. (2010) Heart-specific overexpression of CUGBP1 reproduces functional and molecular abnormalities of myotonic dystrophy type 1. *Hum. Mol. Genet.* **19**, 1066–1075.

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.

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.

Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., Neville, C., Narang, M., Barcelo, J., O'Hoy, K., Leblond, S., Earle-MacDonald, J., DeJong, P.J., Wieringa, B. & Korneluk, R.G. (1992) Myotonic dystrophy mutation: an unstable CTG repeat in the 3′ untranslated region of the gene. *Science* **255**, 1253–1255.

Makeyev, E.V., Zhang, J., Carrasco, M.A. & Maniatis, T. (2007) The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol. Cell* **27**, 435–448.

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.

Ohsawa, N., Koebis, M., Suo, S., Nishino, I. & Ishiura, S. (2010) Alternative splicing of PDLIM3/ALP, for alpha-actinin-associated LIM protein 3, is aberrant in persons with myotonic dystrophy. *Biochem. Biophys. Res. Commun.* **409**, 64–69.

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.

Roof, D.J., Hayes, A., Adamian, M., Chishti, A.H. & Li, T. (1997) Molecular characterization of abLIM, a novel actin-binding and double zinc finger protein. *J. Cell. Biol.* **138**, 575–588.

Salisbury, E., Sakai, K., Schoser, B., Huichalaf, C., Schneider-Gold, C., Nguyen, H., Wang, G.L., Albrecht, J.H. & Timchenko, L.T. (2008) Ectopic expression of

cyclin D3 corrects differentiation of DM1 myoblasts through activation of RNA CUG-binding protein, CUGBP1. *Exp. Cell Res.* **314**, 2266–2278.

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.

Schmeichel, K.L. & Beckerle, M.C. (1994) The LIM domain is a modular protein-binding interface. *Cell* **79**, 211–219.

Sergeant, N., Sablonniere, B., Schraen-Maschke, S., Ghestem, A., Maurage, C.A., Wattez, A., Vermersch, P. & Delacourte, A. (2001) Dysregulation of human brain microtubule-associated tau mRNA maturation in myotonic dystrophy type 1. *Hum. Mol. Genet.* **10**, 2143–2155.

Southby, J., Gooding, C. & Smith, C.W. (1999) Polypyrimidine tract binding protein functions as a repressor to regulate alternative splicing of alpha-actinin mutally exclusive exons. *Mol. Cell. Biol.* **19**, 2699–2711.

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.

Tang, Z.Z., Yarotskyy, V., Wei, L., Sobczak, K., Nakamori, M., Eichinger, K., Moxley, R.T., Dirksen, R.T. & Thornton, C.A. (2012) Muscle weakness in myotonic dystrophy associated with misregulated splicing and altered gating of Ca (V)1.1 calcium channel. *Hum. Mol. Genet.* **21**, 1312–1324.

Timchenko, N.A., Cai, Z.J., Welm, A.L., Reddy, S., Ashizawa, T. & Timchenko, L.T. (2001) RNA CUG repeats sequester CUGBP1 and alter protein levels and activity of CUGBP1. *J. Biol. Chem.* **276**, 7820–7826.

Underwood, J.G., Boutz, P.L., Dougherty, J.D., Stoilov, P. & Black, D.L. (2005) Homologues of the *Caenorhabditis elegans* Fox-1 protein are neuronal splicing regulators in mammals. *Mol. Cell. Biol.* **25**, 10005–10016.

Ward, A.J., Rimer, M., Killian, J.M., Dowling, J.J. & Cooper, T.A. (2010) CUGBP1 overexpression in mouse skeletal muscle reproduces features of myotonic dystrophy type 1. *Hum. Mol. Genet.* **19**, 3614–3622.

Zhang, J., Bahi, N., Llovera, M., Comella, J.X. & Sanchis, D. (2009) Polypyrimidine tract binding proteins (PTB) regulate the expression of apoptotic genes and susceptibility to caspase-dependent apoptosis in differentiating cardiomyocytes. *Cell Death Differ.* **16**, 1460–1468.

Zhang, L., Liu, W. & Grabowski, P.J. (1999) Coordinate repression of a trio of neuron-specific splicing events by the splicing regulator PTB. *RNA* **5**, 117–130.

*Received*: 4 June 2014
*Accepted*: 14 October 2014

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

**Figure S1** Exon array analysis of *ABLIM1* in non-DM1 and DM1 biopsies.

**Figure S2** *ABLIM1* exon 4 splicing doesn't change in patients with DM1.

**Figure S3** Endogenous *ATP2A1* splicing is regulated by MBNL and CELF families in HEK-293.

**Figure S4** The expression of FOX1 construct in C2C12 cells transfected with Mock (pSecDK) or FOX1 construct is confirmed by quantitative RT-PCR (each $n = 3$).

**Figure S5** The expression of *PTBP1* in skeletal muscle of non-DM individuals and DM1 patients is nearly on same level, confirmed by quantitative RT-PCR using same samples as ones of Fig. S2 (each $n = 6$).

**Figure S6** The knockdown of endogenous Mbnl1, Celf1, Celf2, or Ptbp1 in C2C12 cells was confirmed by Western blotting.

**Figure S7** The expression of splicing factor constructs (Fig. 4) was confirmed by Western blotting using anti-His antibody.

**Figure S8** *ABLIM1* splicing isoforms and gene structures in human and mouse according to NCBI.

**Figure S9** Hematoxylin and Eosin staining (H&E staining) of biopsies of non-DM individual (C-2) and six DM1 patients (DM).

**Table S1** Information of biopsies

**Table S2** Primers used for RT-PCR

*Human Molecular Genetics, 2015, Vol. 24, No. 3* *740–756*
*doi:10.1093/hmg/ddu492*
*Advance Access published on September 30, 2014*

# Nuclear localization of MBNL1: splicing-mediated autoregulation and repression of repeat-derived aberrant proteins

**Yoshihiro Kino[1,2,3,4], Chika Washizu[3], Masaru Kurosawa[1,2,3], Yoko Oma[5], Nobutaka Hattori[1], Shoichi Ishiura[5] and Nobuyuki Nukina[1,2,3,*]**

[1]Department of Neuroscience for Neurodegenerative Disorders, Juntendo University Graduate School of Medicine, Tokyo 113-0033, Japan, [2]CREST (Core Research for Evolutionary Science and Technology), JST, Saitama 332-0012, Japan, [3]Laboratory for Structural Neuropathology, Brain Science Institute, RIKEN, Saitama 351-0198, Japan, [4]Department of Bioinformatics and Molecular Neuropathology, Meiji Pharmaceutical University, Tokyo 204-8588, Japan and [5]Department of Life Sciences, Graduate School of Arts and Sciences, the University of Tokyo, Tokyo 153-8902, Japan

Received September 1, 2014; Revised September 1, 2014; Accepted September 22, 2014

**In some neurological diseases caused by repeat expansions such as myotonic dystrophy, the RNA-binding protein muscleblind-like 1 (MBNL1) accumulates in intranuclear inclusions containing mutant repeat RNA. The interaction between MBNL1 and mutant RNA in the nucleus is a key event leading to loss of MBNL function, yet the details of this effect have been elusive. Here, we investigated the mechanism and significance of MBNL1 nuclear localization. We found that MBNL1 contains two classes of nuclear localization signal (NLS), a classical bipartite NLS and a novel conformational NLS. Alternative splicing of exon 7 acts as a switch between these NLS types and couples MBNL1 activity and intracellular localization. Depending on its nuclear localization, MBNL1 promoted nuclear accumulation of mutant RNA containing a CUG or CAG repeat, some of which produced proteins containing homopolymeric tracts such as polyglutamine. Furthermore, MBNL1 repressed the expression of these homopolymeric proteins including those presumably produced through repeat-associated non-ATG (RAN) translation. These results suggest that nuclear retention of expanded RNA reflects a novel role of MBNL proteins in repressing aberrant protein expression and may provide pathological and therapeutic implications for a wide range of repeat expansion diseases associated with nuclear RNA retention and/or RAN translation.**

## INTRODUCTION

Expansions of repetitive sequences in non-coding regions cause neurological diseases such as myotonic dystrophy (DM) types 1 and 2 (DM1 and DM2), and spinocerebellar ataxia type 8 (SCA8) (1). The number of non-coding repeat expansion diseases is still growing, as exemplified by recent findings of the mutations causing spinocerebellar ataxia type 36 and chromosome 9p-linked amyotrophic lateral sclerosis/frontotemporal dementia (C9-ALS/FTD) (2–4). One common pathological feature of these diseases is the presence of nuclear RNA inclusions called ribonuclear inclusions or RNA foci. Notably, muscleblind-like 1 (MBNL1) is a common component of ribonuclear inclusions in DM1, DM2, SCA8 and Huntington's disease-like 2 (HDL2) (5–8). MBNL1 binds directly to CUG and CCUG-repeat RNA (9), which are expressed in these diseases. Recruitment of MBNL1 into ribonuclear inclusions is thought to cause functional depletion of MBNL1 (10,11).

MBNL proteins are evolutionarily conserved RNA-binding proteins with variable numbers of C3H-type zinc finger motifs. In vertebrates, three MBNL orthologs, MBNL1, MBNL2 and MBNL3, have been identified (12). MBNL1 and MBNL2 are expressed in a wide variety of adult tissues including brain, heart and skeletal muscle, whereas MBNL3 is expressed predominantly in the placenta (12). MBNL1 is a splicing factor that directly regulates alternative splicing of numbers of transcripts. Importantly, splicing misregulation of MBNL-regulated genes has been found in DM1 and DM2 tissues (11). Indeed,

*To whom correspondence should be addressed. Tel: +81 7069691951; Email: nnukina@juntendo.ac.jp

*Mbnl1* knockout mice manifest some DM-like phenotypes such as myotonia, cataracts and splicing misregulation, whereas overexpression of Mbnl1 in DM model mice ameliorates the disease phenotypes (13–15). MBNL2 also accumulates in RNA foci in DM tissues, and *Mbnl2* knockout mice exhibit myotonia or brain abnormalities including sleep disorder and learning deficits (10,16,17). Therefore, MBNL proteins are recognized as key mediators, in addition to pathological molecular markers, of repeat expansion diseases such as DM1.

One major class of coding repeat expansion diseases is CAG-repeat diseases, which include Huntington's disease (HD). An expansion of CAG repeats encoding a polyglutamine (polyQ) tract leads to production of an aggregation-prone protein with various toxic effects (18). Although coding CAG-repeat diseases and non-coding CTG repeat diseases have been attributed to protein and RNA gain-of-function mechanisms, respectively, recent studies have revealed overlapping features between these diseases. CAG-repeat-containing transcripts are expressed from the disease allele and produce polyQ-containing proteins in DM1, SCA8 and HDL2 (19–21), whereas toxic effects of CAG-repeat RNA have been suggested in polyQ diseases (22–25). Furthermore, a recent study reports that expanded-repeat RNA can be translated even without an ATG codon (21). In this process, RAN translation, a single repeat tract can produce homopolymeric proteins in multiple frames (21). Thus, gene products from an expanded allele are much more complex than previously recognized. Interestingly, MBNL1 binds to CAG repeats in addition to CUG and CCUG repeats and accumulates in CAG-repeat RNA foci (9,26,27). In addition, MBNL1 and its orthologs can modulate the toxicity of the CAG repeat in animal models (22,25). Thus, MBNL proteins have become potential common players in many repeat expansion diseases, not only CTG/CCTG repeat diseases.

A key issue related to both the abnormal regulation and function of MBNL1 in repeat expansion diseases is the alternative splicing of exon 7 in this protein. The inclusion of this exon is elevated in DM1, DM2 and SCA8 patients and model mice (8,11,28). Moreover, the same alteration is observed in $Mbnl1^{\Delta ex3/\Delta ex3}$ knockout mice, suggesting an autoregulation of exon 7 by MBNL1 (11,29). Furthermore, isoforms containing this exon exhibit nuclear localization (11,30,31), although the mechanism of nuclear localization of MBNL1 is still elusive. Another key issue related to MBNL1 in the nucleus is the precise effect of the RNA foci formation. As demonstrated in DM pathogenesis, nuclear interaction between MBNL proteins and mutant RNA leads to loss of MBNL function. However, it is still unclear whether the colocalization of MBNL1 and repeat RNA in the inclusions reflects some active response of MBNL1 or a passive sequestration of MBNL1. Indeed, recent studies have shown that depletion of MBNL1 reduces ribonuclear inclusions (32–34), supporting the notion that MBNL1 is involved in the process of foci formation.

In this study, we investigated the mechanism, regulation and significance of MBNL1 nuclear localization. We found that nuclear localization is a major determinant of MBNL1 function. We then focused on the regulation of nuclear localization of MBNL1 and identified novel conserved motifs that cooperatively mediate nuclear localization of MBNL1. Alternative splicing of exon 7 acts as a regulatory switch that couples MBNL1 nuclear activity and nuclear localization. Finally, we identified the significance of the interaction between MBNL1 protein and repeat RNA in the nucleus. MBNL1 promotes the nuclear retention of repeat-containing transcripts, which results in repression of aberrant protein expression from the expanded repeats, pointing to a novel aspect of RNA foci formation as well as a potential strategy for preventing protein toxicity in repeat expansion diseases.

## RESULTS

### Intracellular localization of MBNL isoforms

Mammalian MBNL proteins show complex splicing variations (Fig. 1A and Supplementary Material, Fig. S1A). We first tried to identify the determinants of MBNL1 properties in relation to splice variations. In this study, subcellular localization of protein or RNA was evaluated by automated fluorescence analysis of individual transfected cells (35). The nucleocytoplasmic index (NCI) ranges from 0 to 1 with higher values indicating stronger cytoplasmic localization in a cell. An averaged NCI value from a cell population was used as a representative value of the transfected construct. Using this system, intracellular localization of EGFP-fused MBNL1 isoforms was analyzed in comparison with the control, EGFP. In COS-7 cells, only $MBNL1_{42}$ showed nuclear localization (Fig. 1B and C). Thus, exon 7, which is included only in this isoform, must be involved in nuclear localization (Fig. 1A). The other isoforms showed more cytoplasm localization with $MBNL1_{40s}$ showing the most cytoplasmic localization (Fig. 1B and C). Protein expression of these isoforms was verified by western blot (Supplementary Material, Fig. S1C). We also confirmed that C-terminally myc-tagged MBNL1 isoforms showed subcellular localization similar to that of N-terminally EGFP-tagged ones (Supplementary Material, Fig. S5C).

Next, we compared splicing regulatory activity of these isoforms using an *Actn1* minigene (36,37). This minigene contains two mutually exclusive alternative exons, NM (non-muscle) and SM (smooth muscle) (Fig. 1D). In the basal splicing pattern, NM inclusion was predominant. When the amount of MBNL1 was increased, NM inclusion was gradually reduced, whereas SM inclusion was transiently increased and then decreased at maximum MBNL1 dose (Fig. 1D). Among MBNL1 isoforms, $MBNL1_{42}$ had the strongest effect on *Actn1* splicing (Fig. 1E), consistent with its strong nuclear localization. Isoforms 41, 40, 41 and 40s exhibited *Actn1* splicing regulation weaker than that of $MBNL1_{42}$ but stronger than isoforms 36 and 35 (Fig. 1E and Supplementary Material, Fig. S1D). Both $MBNL1_{36}$ and $MBNL1_{35}$ lacked exon 5 (Fig. 1A), suggesting that the linker region between ZnF1/2 and ZnF3/4 encoded by this exon determines the splicing regulatory activity of MBNL1. An N-terminal region of MBNL1 with four ZnF motifs (MBNL1-N) alone showed much weaker activity compared with $MBNL1_{42}$ (Fig. 1F). This reduction can be explained by its weaker nuclear localization (Fig. 1C); the activity of MBNL1-N splicing regulation was fully recovered by the addition of an NLS, whereas it was further reduced by the addition of a nuclear export signal (NES) (Fig. 1C and F). The regulatory activity of MBNL1-N was weakened when exon 5 was deleted (Fig. 1F, MBNL1-NΔex5). However, the presence of exon 5 is not an absolute requirement for splicing regulation, as

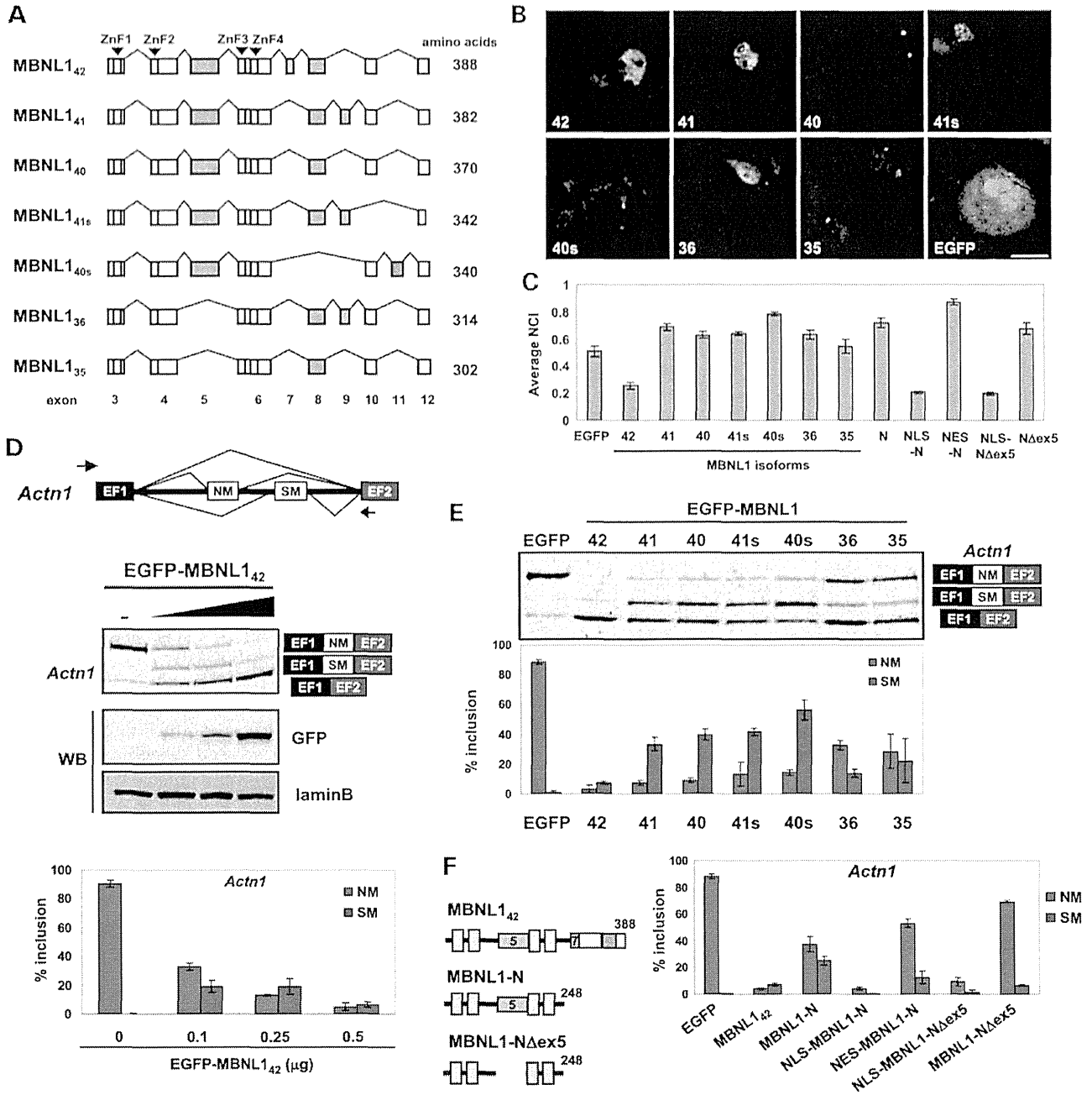


**Figure 1.** Splice variation in MBNL proteins. (**A**) Structures of MBNL isoforms. Alternative exons (5, 7, 8, 9, 10 and 11) are indicated by different colors. Blue boxes indicate zinc finger motifs. (**B**) Intracellular localization of MBNL isoforms. The expression of EGFP-fused MBNL isoforms in COS-7 cells was observed by confocal microscopy. Hoechst 33342 was used to stain the nucleus. Scale bar represents 20 μm. (**C**) Quantitative localization analysis of EGFP-fused MBNL1 isoforms. Bars indicate average NCI value (see text) of transformants in COS-7 cells. EGFP serves as a control. (**D**) Dose-dependent splicing regulation of *Actn1* minigenes by EGFP-MBNL1$_{42}$ in COS-7 cells. Structure of *Actn1* minigenes is shown at the top. Arrows indicate the position of primers used in the splicing assays. The results of splicing assay of *Actn1* and western blots of EGFP-MBNL1 and endogenous Lamin B are shown in the middle panels. Bar chart shows the quantified results of *Actn1* splicing assay (mean ± SD, $n = 3$). (**E**) Splicing regulation of *Actn1* by EGFP-fused MBNL1 isoforms in COS-7 cells. Bar chart shows quantified results (mean ± SD, $n = 3$). (**F**) *Actn1* splicing assay of the N-terminal region of MBNL1 in COS-7 cells. An NLS or NES-fused MBNL1-N was also examined. NΔex5 is a common N-terminus region of MBNL1$_{35}$ and MBNL1$_{36}$. Bar chart shows quantified results as in (E).

NLS-fused NΔex5 showed splicing regulation similar to that of MBNL1$_{42}$ or NLS-MBNL1-N (Fig. 1F). In conclusion, the splicing regulatory activity of MBNL1 can be attributed to the N-terminus region containing four ZnF motifs and a linker region, although the activity is also determined by the degree of nuclear localization that is mediated by the C-terminus including exon 7. Similarly, MBNL2 has an alternative exon corresponding to MBNL1 exon 7, which is associated with nuclear localization and stronger splicing regulation (Supplementary Material, Fig. S2A–C).

### Identification of a nuclear localization signal in the C-terminus of MBNL1

As seen earlier, nuclear localization is a major determinant for the functions of MBNL1. We sought to identify the mechanism of its nuclear localization using EGFP-fused MBNL1 mutants. In both COS-7 and Neuro2a (N2a) cells, a C-terminal region of $MBNL1_{42}$ was essential for nuclear localization, as the deletion of the C-terminus disrupted nuclear localization (1–248, Fig. 2A). In addition, the C-terminus of $MBNL1_{42}$ (C42), but not of the other isoforms (C40. C41, C40s and C41s), was sufficient for inducing nuclear localization (Fig. 2A). Deletion analysis of this region revealed that nuclear localization activity could be narrowed to the regions 276–308, which comprises exon 7 and

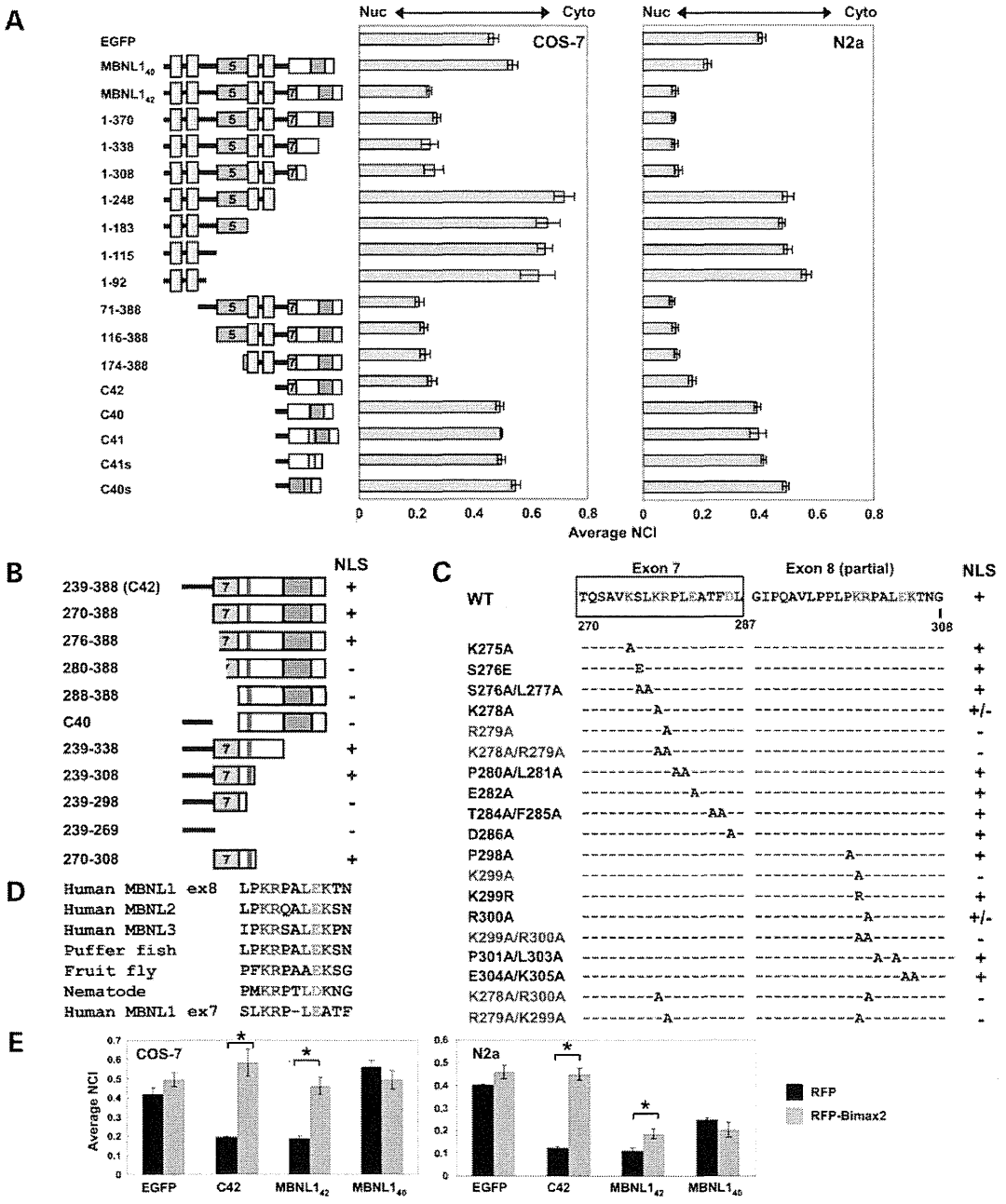


**Figure 2.** Determination of an NLS in the C-terminal region of $MBNL1_{42}$. (**A**) Intracellular localization of MBNL1 deletion mutants in COS-7 and N2a cells. Bar chart shows average NCI values of MBNL1 mutants (mean ± SD, $n = 3$). The structure of the transfected constructs is indicated at the left. (**B**) Nuclear localization of the C-terminal fragments of $MBNL1_{42}$ fused with EGFP. '+' indicates enhanced nuclear localization of the MBNL1 mutant. The results are based on the analysis of Supplementary Material, Figure S3A. (**C**) Point mutation analysis of $MBNL1_{42}$. Mutations were introduced into the full-length $MBNL1_{42}$ fused with EGFP. The results are based on the analysis of Supplementary Material, Figure S3B. (**D**) Evolutionarily conserved motif of exon 8. (**E**) The effect of RFP-fused Bimax2 peptide on the localization of MBNL1. Bar chart shows average NCI values in the presence and absence of RFP-Bimax2 (mean ± SD, $n = 3$). $^*P < 0.05$, two-tailed unpaired $t$-test ($n = 3$).

a part of exon 8 (Fig. 2B and Supplementary Material, Fig. S3A). Consistently, a peptide covering 270–308 was sufficient for driving nuclear localization (Fig. 2B and Supplementary Material, Fig. S3A). Point mutation analysis of this region in the context of full-length MBNL1$_{42}$ revealed the importance of four basic amino acids, K278, R279, K299 and R300 (Fig. 2C and Supplementary Material, Fig. S3B). Single point mutation of either R279A or K299A was sufficient for disrupting nuclear localization. In the case of K278A and R300A, single mutation showed little or no effect, whereas the simultaneous mutation of K278A and R300A completely disrupted nuclear localization (K278A/R300A, Fig. 2C and Supplementary Material, Fig. S3B). We also examined the effect of NLS mutations on the splicing activity. Both R279A and K299A mutants showed weakened splicing regulation of MBNL1, even though the protein expression of these mutants was equivalent to that of wild type (WT) (Supplementary Material, Fig. S3C and D).

All MBNL1 isoforms analyzed in this study, except for MBNL1$_{40s}$, contain a highly conserved motif in exon 8, KRPALE (Fig. 2D, Supplementary Material, Fig. S4A, and pink boxes in Fig. 2B). A corresponding motif in the fruit fly is involved in nuclear localization (30). However, this motif was not sufficient for mammalian MBNL1 nuclear localization, as demonstrated by the non-nuclear localization of C40 and 288–388, which contain this motif (Fig. 2B and Supplementary Material, Fig. S3A). The identified residues essential for nuclear localization were partly similar to a classical bipartite NLS (biNLS), which is typically summarized as KRX$_{(10\text{-}12)}$KR and acts through the importin α/β pathway. Co-expression of red fluorescent protein (RFP)-fused Bimax2 peptide, a specific inhibitor of the importin α/β pathway (38), prevented nuclear localization of C42 (Fig. 2E). Therefore, the NLS in the C-terminus of MBNL1 appears to be a bipartite NLS, even though the spacing of two KR motifs (19 amino acids) is larger than a typical biNLS (10–12 amino acids). In conclusion, both exon 7 and the conserved motif of exon 8 provide KR motifs that constitute a biNLS only when exon 7 is included in the alternative splicing.

### The second NLS activity of MBNL1 involves multiple conserved motifs

During the above-mentioned analysis, we noticed a cell-type-dependent difference in nuclear localization of MBNL1. In N2a cells, MBNL1$_{40}$ showed significant nuclear localization despite the absence of exon 7 (Fig. 2A). Interestingly, Bimax2 disrupted nuclear localization of MBNL1$_{42}$ in COS-7 but only partially in N2a cells (Fig. 2E). Furthermore, nuclear localization of MBNL1$_{40}$ in N2a was not affected by Bimax2 (Fig. 2E). These results suggested that MBNL1 contains another NLS independent of the importin α/β pathway. We performed deletion analysis of MBNL1$_{40}$ (identical to 1-388Δex7 in Fig. 3A, a designation based on MBNL1$_{42}$ to maintain the amino acid numbering). In N2a cells, nuclear localization of MBNL1$_{40}$ required at least parts of both N- and C-terminus regions (Fig. 3A). Comparison of 1-308Δex7 and 1-298Δex7 demonstrated the necessity of the region around the conserved motif in exon 8 (pink box, Fig. 3A). In addition, deletion of an N-terminal region containing ZnF1/2 resulted in the loss of nuclear localization, as shown by 71-388Δex7 and 71-308Δex7 (Fig. 3A). Thus, the region around the conserved motif in exon 8 and the N-terminus region were essential for nuclear localization. We then tested point mutations of the conserved motif on exon 8 in the context of MBNL1$_{40}$. Similar to the case of MBNL1$_{42}$, mutation of K299A and R300A disrupted the nuclear localization of MBNL1$_{40}$ (Fig. 3B). However, other residues in the conserved motif were also essential for nuclear localization as revealed by the moderate to severe effects of P301D, A302D, L303A and E304A (Fig. 3B). This stood in contrast to the case of MBNL1$_{42}$, in which mutations of these residues did not affect nuclear localization (Fig. 2C).

Next, the involvement of the N-terminal region was examined. ZnF1/2 is a very highly conserved motif through evolution. We noticed that flanking regions of ZnF1/2 are also highly conserved and contain motifs, RD/KWL and KxQL/NGR (Fig. 3C and Supplementary Material, Fig. S4A). These motifs are not found in the flanking regions of ZnF3/4 of MBNL1 or other C3H ZnF proteins. Deletion of the first 13 amino acids from the N-terminus disrupted nuclear localization (Δ13, Fig. 3C). This could be partly explained by the depletion of the RD/KWL motif as shown by its mutation, mut(9-14) (Fig. 3C). Deletion of the KxQL/NGR motif, Δ(76-85), moderately weakened nuclear localization of MBNL1 (Fig. 3C). Interestingly, ZnF1/2, but not ZnF3/4, was found to contribute to nuclear localization (ΔZnF1/2 and ΔZnF3/4, Fig. 3C). Therefore, the N-terminal region of MBNL1 that comprises ZnF1/2 and its flanking motifs and the C-terminal conserved motif in exon 8 are essential for a second NLS activity of MBNL1. In contrast to the biNLS that comprises a linear motif, the second NLS was mediated by multiple and discrete regions of MBNL1. This is consistent with the recent classification of a conformational NLS (conNLS) (39).

We then analyzed whether the conNLS can act in the context of MBNL1$_{42}$ in N2a cells. A mutation of the KR motif in exon 7, which disrupts the biNLS, weakened but still preserved nuclear localization [1-388(K278A/R279A), Fig. 3D], suggesting that conNLS activity is present in MBNL1$_{42}$. However, deletion of a C-terminus region including the conserved motif of exon 8 also weakened but still preserved nuclear localization (Fig. 3D, compare 1–308, 1–298 and EGFP). This result was unexpected as this deletion was predicted to disrupt both biNLS and conNLS. Further deletion of exon 7 resulted in the loss of nuclear localization (1–269, Fig. 3D). Thus, exon 7 is involved in nuclear localization even in the absence of exon 8. As exon 7 itself was not sufficient for inducing nuclear localization (239–269, Fig. 2B), this effect of exon 7 requires the presence of the N-terminal region. We noticed that the sequence around the KR motif in exon 7 (KRPLE) is similar to that of the conserved motif in exon 8, KRPALE (Fig. 2D). When the KR motif of exon 7 was mutated in the context of 1–298, the mutant, 1–298 (K278A/R279A), exhibited loss of nuclear localization similar to 1–269 (Fig. 3D). Therefore, it appears that exon 7 contains another copy of the conserved motif and constitutes a conNLS together with the N-terminus region. In conclusion, MBNL1 has two classes of NLS (Fig. 3E). One is the biNLS consisting of two KR motifs embedded in conserved motifs of exons 7 and 8. The other is the conNLS, which involves the N-terminus region and either one of the conserved motifs in exon 7 or exon 8. Thus, MBNL1$_{42}$ possesses three possible NLS combinations, biNLS, conNLS using exon 7 and conNLS using exon 8. Most other isoforms contain only one conNLS using exon 8. MBNL1$_{40s}$ lacks any known NLS.

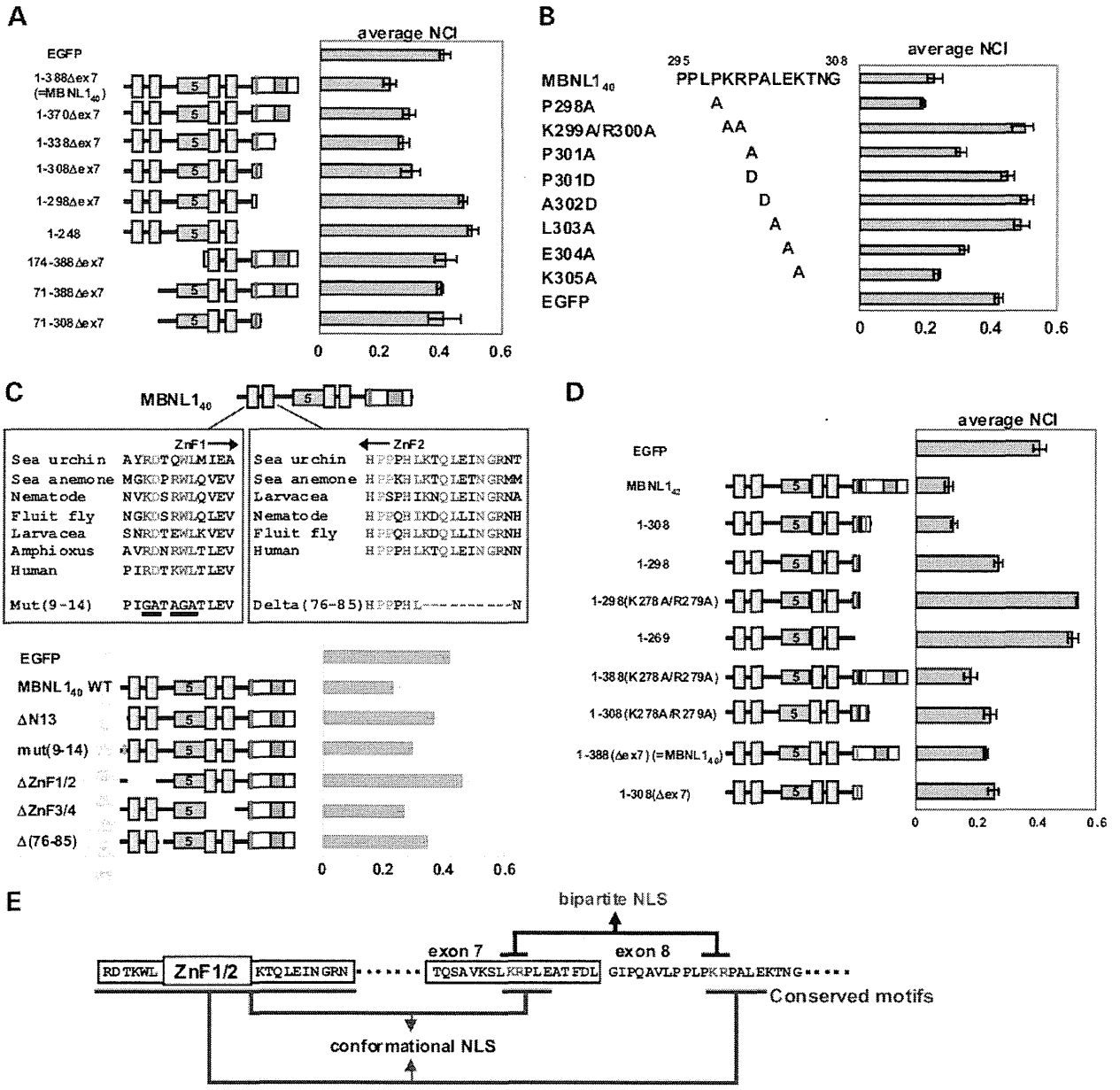


**Figure 3.** Identification of the second NLS activity of MBNL1. (**A**) Deletion analysis of MBNL1$_{40}$. Average NCI values of MBNL1$_{40}$ deletion mutants in N2a cells are indicated (mean ± SD, $n = 3$). 1-388Δex7 is identical to MBNL1$_{40}$. (**B**) The effects of point mutations on the localization of MBNL1$_{40}$. Mutations were introduced in the conserved motif of exon 8. Bar chart shows average NCI values as in (A). (**C**) Mutation analysis of conserved motifs in the N-terminus region of MBNL1. Evolutionary conservation of motifs flanking ZnF1/2 is shown at the top. Bar chart shows the nucleocytoplasmic localization of N-terminal mutants. The structure of the mutants is shown at the left. (**D**) The conserved motifs in exon 7 act as a part of a conNLS. Bar chart shows average NCI values as in (A). The structure of the mutants is shown at the left. (**E**) The mechanism of nuclear localization of MBNL1. See text for the detail.

## Splicing-mediated coupling of MBNL activity and subcellular localization

As demonstrated earlier, exon 7 is an important determinant of the localization and function of MBNL1. In both N2a cells and adult mouse striatum, we detected Mbnl1 isoforms with and without exon 7 (Supplementary Material, Fig. S1A). Endogenous Mbnl1 was localized in the nucleus and cytoplasm (Supplementary Material, Fig. S5A). To analyze splicing regulation of this exon, we analyzed a minigene covering exons from 6 to 8 of mouse *Mbnl1*. Overexpression of the N-terminal regions of all three Mbnl paralogs strongly reduced the inclusion of exon 7 (Fig. 4A). In contrast to the case of *Actn1*, the N-terminus of Mbnl proteins did not require an NLS for strong repression of *Mbnl1* exon 7, suggesting the high sensitivity of this exon to Mbnl proteins. Co-expression of the *Mbnl1* minigene and a DMPK construct with an interrupted CTG480, but not CTG18, increased the inclusion of exon 7 (Fig. 4B). We also observed a small but significant increase in exon 7 inclusion when the

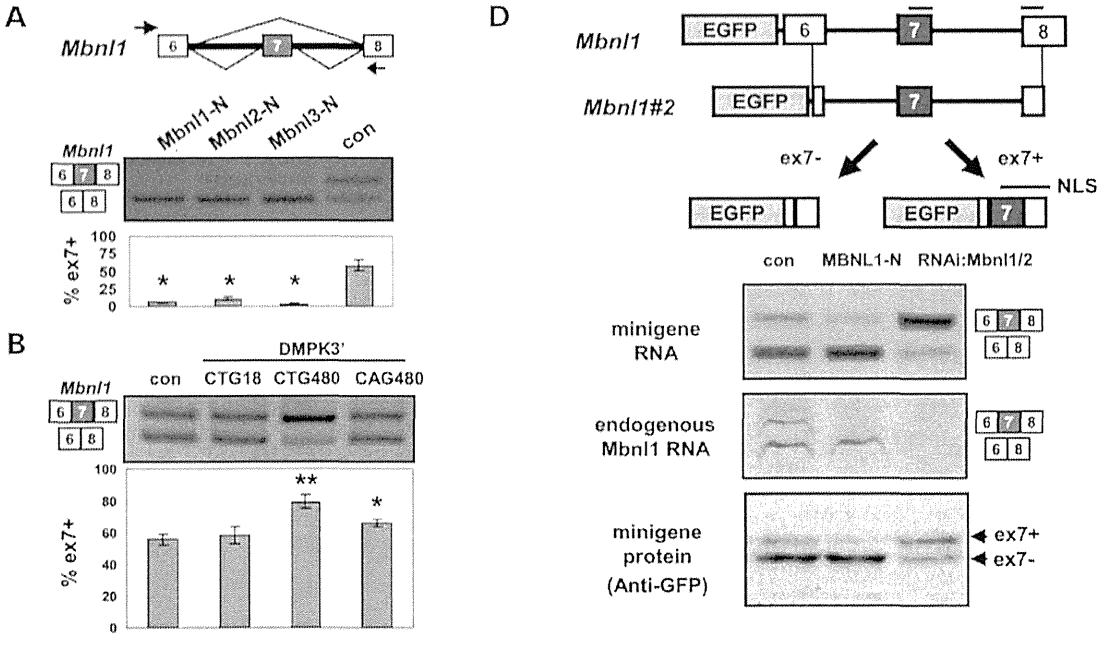


**Figure 4.** Splicing regulation of an *Mbnl1* minigene. (**A**) Structure of an *Mbnl1* minigene covering exons 6–8 (top). *Mbnl1* minigene is regulated by the N-terminus region of murine Mbnl paralogs in N2a cells (middle and bottom). Bar chart shows quantified results of the splicing assay (mean ± SD, $n = 3$). (**B**) Results of splicing assay using the *Mbnl1* minigene and DMPK3′ constructs with a normal or an expanded repeat. Results are shown as in (A). (**C**) The effect of RNAi-mediated depletion of endogenous Mbnl1 in N2a cells. Splicing assay of *Mbnl1* minigene was performed using cells depleted of Mbnl1 and/or Mbnl2 as indicated. Bar chart shows quantified results of the splicing assay (mean ± SD, $n = 3$). (**D**) An *Mbnl1* fluorescent minigene to examine the relationship between splicing pattern and localization. The *Mbnl1#2* minigene contains shortened exons 6 and 8. An N2a-based cell line stably expressing *Mbnl1#2* was established (line C3). Overexpression of MBNL1-N and simultaneous knockdown of endogenous Mbnl1 and Mbnl2 were examined. The splicing pattern of minigene mRNA (top) and endogenous Mbnl1 mRNA (middle) were detected by RT-PCR. Minigene protein was detected by western blot using anti-GFP antibody (bottom). (**E**) Quantified results of EGFP localization in the *Mbnl1#2* cell line C3 when MBNL1 was overexpressed or depleted. Overexpression of RFP-Bimax2 was also tested as a control of biNLS inhibition. Bar chart shows average NCI of RFP-positive cells (mean ± SD, $n = 3$). $^*P < 0.05$, Tukey's test for multiple comparison ($n = 3$).

CAG480 construct was overexpressed (Fig. 4B). RNAi-mediated knockdown of endogenous Mbnl1 induced a moderate increase in the inclusion of exon 7, whereas knockdown of Mbnl2 alone exhibited little effect (Fig. 4C and Supplementary Material, Fig. S2D). Importantly, simultaneous knockdown of both Mbnl1 and Mbnl2 increased exon 7 inclusion more strongly than that of Mbnl1 alone (miMbnl1/2, Fig. 4C and Supplementary Material, Fig. S2D). These results demonstrated that alternative splicing of *Mbnl1* exon 7 is regulated by MBNL proteins in a dose-dependent manner. We confirmed intracellular association between EGFP-MBNL1 and endogenous *Mbnl1* transcripts by ribonucleoprotein immunoprecipitation analysis using a cell line stably expressing EGFP-MBNL1$_{40}$ (Supplementary Material, Fig. S4B). Thus, MBNL1 associates with its transcript and autoregulates alternative splicing.

As the *Mbnl1* minigene contains a region corresponding to the biNLS identified earlier, we expected that the localization of EGFP fused upstream of the minigene might be changed according to the splicing pattern of exon 7. However, the EGFP fluorescence of this construct was not strong enough to be detected clearly. We made another shortened minigene (Mbnl1#2, Fig. 4D), which still contained the biNLS region and showed stronger fluorescence than the original minigene. A N2a-based cell line stably expressing this minigene showed EGFP fluorescence in both nucleus and cytoplasm (Supplementary Material, Fig. S4C). Patterns of minigene-derived protein or RNA were altered when MBNL1 was overexpressed or knocked down (Fig. 4D). The patterns of the transcript and the protein expressed from the minigene were altered depending on the dose of MBNL1 (Fig. 4D). Notably, cytoplasmic localization of EGFP

was enhanced when MBNL1 was transfected, whereas nuclear localization was enhanced when endogenous Mbnl1 and Mbnl2 were knocked down (Fig. 4E and Supplementary Material, Fig. S4C). Thus, alternative splicing of *Mbnl1* exon 7 directly reflects the activity of MBNL proteins and alters nucleocytoplasmic localization of MBNL1. This mechanism can be regarded as a dynamic negative feedback where an increase in nuclear MBNL activity leads to a decrease in the production of nuclear isoform containing exon 7.

### MBNL1 promotes nuclear accumulation of RNA with expanded CUG and CAG repeats

Next, we analyzed the effect of MBNL1 on the localization of transcripts containing expanded CUG or CAG repeats. It has been known that CUG repeat-containing RNA is retained in the nucleus and co-aggregates with MBNL1 (10). As reported previously (32), knockdown of MBNL proteins greatly reduced RNA foci in DM1 fibroblasts (Supplementary Material, Fig. S11B and C), indicating an essential role of MBNL proteins in RNA foci formation. We then used constructs that contain an interrupted repeat similar to the ones widely used in DM1 studies (Supplementary Material, Fig. S6A). CUG-repeat RNA in the context of a DMPK fragment (DMPK3′-CTG480) was visualized by fluorescence in situ hybridization (FISH). Unexpectedly, many cells showed RNA localization in the cytoplasm (Fig. 5A). In contrast, co-transfection of $MBNL1_{42}$ with the repeat constructs greatly enhanced nuclear accumulation of repeat RNA mainly as foci (arrows, Fig. 5B and Supplementary Material, Fig. S7A and B). We quantified the relative nucleocytoplasmic distribution of repeat RNA based on the FISH signal. Nuclear localization of CUG repeat was enhanced by $MBNL1_{42}$ compared with EGFP (Fig. 5C). Mutation analysis revealed that the enhancement of nuclear RNA localization was dependent on the composition of zinc finger motifs (comparison of MBNL1-N versus 71–248), linker region ($MBNL1_{40}$ versus $MBNL1_{35}$ and 71–388 versus 174–388) and nuclear localization (NLS-MBNL1-N versus NES-MBNL1-N) (Fig. 5C and Supplementary Material, Fig. S8A, and see also Fig. 2A for the localization of MBNL1 mutants). We also tested non-coding repeats in a DMPK-independent context, in which CTG or CAG repeat is downstream of a CMV promoter and no ATG codon is located between the promoter and the repeat. As with DMPK3′-CTG480, CAG240 did not show an efficient nuclear accumulation in the absence of $MBNL1_{42}$ overexpression (Fig. 5D). Thus, expanded RNA alone does not efficiently accumulate in the nucleus in the absence of a sufficient amount of MBNL1. Comparative analysis of CAG and CTG constructs with various lengths (Supplementary Material, Fig. S8B) revealed that (i) CUG-repeat RNA tends to show stronger nuclear localization than the length-matched CAG repeat, (ii) $MBNL1_{42}$ enhanced nuclear localization of both CAG and CUG repeats, whereas MBNL1-N showed weak effects on CAG repeats and (iii) even the same size of repeat (CTG480) can have differential degree of nuclear localization depending on the context (V5 versus DMPK′3). We also analyzed mutual effects of repeat RNA and MBNL1 on their nucleocytoplasmic localization. We used MBNL1 variants with partial cytoplasmic localization, $MBNL1_{40}$ (containing a conNLS) and MBNL1-N (containing no NLS). While these proteins enhanced nuclear localization of RNA, that of these proteins were not altered (Fig. 5E). Finally, knockdown of Mbnl proteins increased cytoplasmic localization of CUG repeat (Fig. 5F). These results suggest that the nuclear retention of CUG-repeat RNA can be actively caused by MBNL1 in the nucleus. Interestingly, we noticed that most endogenous Mbnl1 was located outside of RNA foci that were induced by MBNL overexpression (Supplementary Material, Fig. S7C). We also examined a CAA repeat as a control. Although CAA108 occasionally formed nuclear foci, EGFP-$MBNL1_{42}$ did not colocalize with them (Supplementary Material, Fig. S9A). Moreover, both overexpression and knockdown of MBNL1 did not affect the localization of CAA-repeat RNA (Supplementary Material, Fig. S9B and C). Thus, the effect of MBNL1 is dependent on repeat sequences.

### MBNL1 suppresses the expression of aberrant proteins containing an expanded CUG or CAG repeat

A recent study reports protein expression from expanded repeats in the absence of an initiation codon, in which the repeats tract could be translated in all three frames (21). We examined whether any proteins are expressed from the non-coding repeat constructs used in Figure 5D (CTG240 and CAG240). CAG240 produced a high-molecular-weight protein detected by antibody against V5-tag, which is located downstream of the repeat tract (Fig. 6A). We also examined different lengths of CAG and CTG repeats. Interestingly, CAG constructs expressed V5-positive proteins in a length-dependent manner with peak expression at 360 repeats. In contrast, CTG constructs did not express proteins detected by anti-V5 (Fig. 6A). Owing to 5-base interruptions for every 20 CAG or CTG repeats, the interrupted CAG and CTG repeat tracts encode $(Q20\text{-}S22\text{-}A20)_n$ and $(C20\text{-}A20\text{-}L21)_n$ polypeptides, respectively, if translated (Supplementary Material, Fig. S6A). The large protein expressed from the CAG construct was detected by 1C2 antibody, indicating the presence of polyQ tracts. In the absence of an apparent start codon upstream of the repeat tract, the reading frame of the repeat is not clear. We examined three different reading frames of the V5 tag downstream of the repeat tract (Fig. 6B); reading frame 1 was used in the above-mentioned experiments. CAG-repeat constructs with different reading frames produced high-molecular-weight proteins detected by anti-V5 and 1C2 antibodies (Fig. 6B). These constructs showed some difference in the pattern of lower-molecular-weight proteins (corresponding to 40–70 kDa) (Fig. 6B). In contrast, little or no protein was expressed from the CTG constructs with different reading frames of the V5 tag (Fig. 6B). Next, we tested the effect of MBNL1 on the expression of the protein from the CAG construct. When the CAG360 construct was co-expressed with $MBNL1_{42}$ or $MBNL2_{42}$, the anti-V5-reactive product was decreased compared with control (Fig. 6C). $MBNL1_{42}$ enhanced nuclear localization of CAG360 mRNA (Supplementary Material, Fig. S8B). Simultaneous knockdown of both Mbnl1 and Mbnl2 increased the expression of CAG360 (Fig. 6D). Therefore, nuclear retention of repeat-containing RNA induced by MBNL1 can result in the reduction of aberrant proteins expressed from the repeat RNA.

### MBNL1 suppresses the expression of polyglutamine-containing proteins

As shown earlier, MBNL1 can alter the localization of seemingly non-coding CAG repeats and reduce the expression of

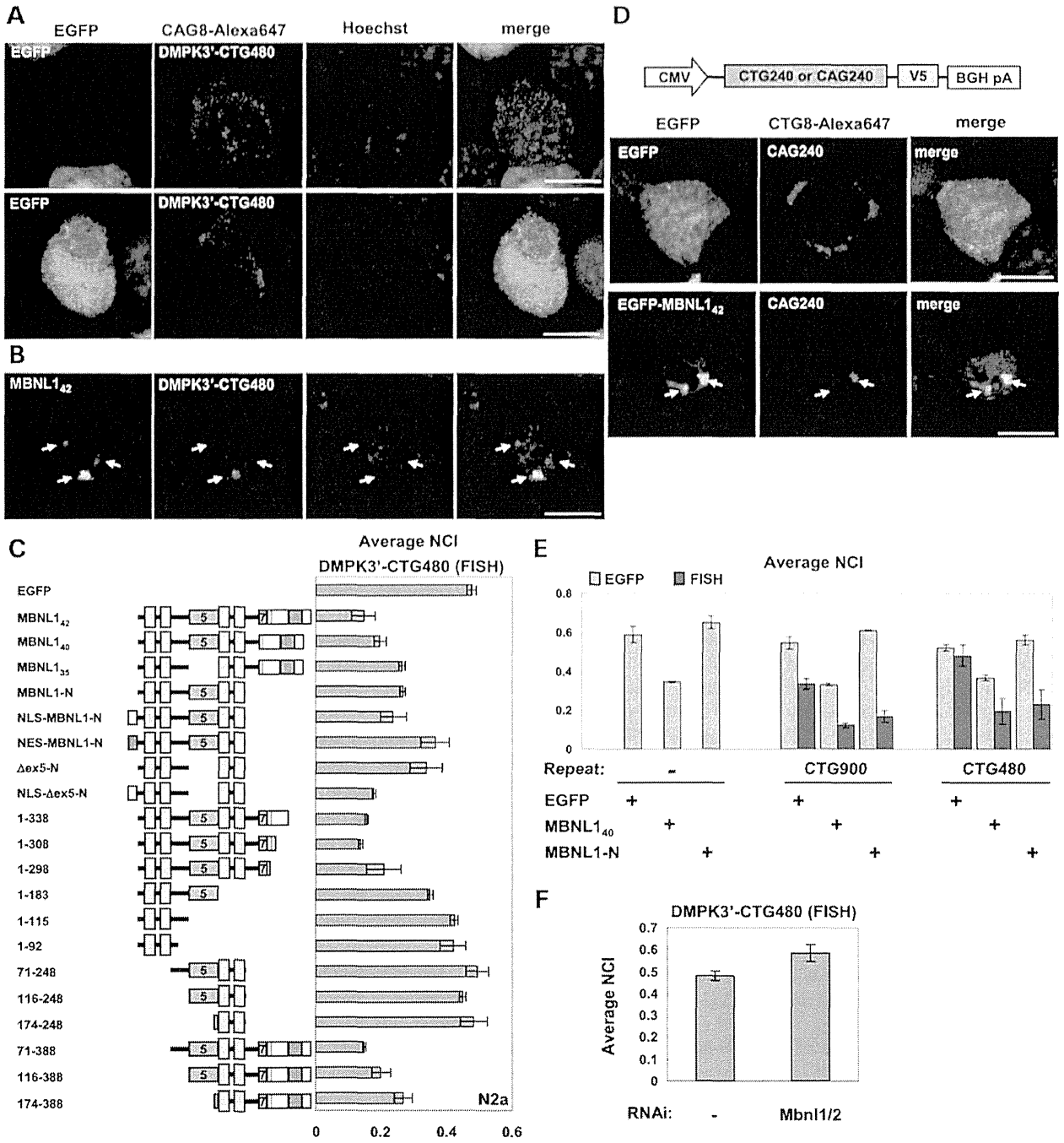


**Figure 5.** MBNL1 induces nuclear retention of repeat RNA. (**A** and **B**) Intracellular localization of CUG repeat in the context of the 3′ region of DMPK 3′. DMPK3′-CTG480 was transfected with either EGFP (A) or EGFP-MBNL1$_{42}$ (B) in N2a cells, and the localization of repeat RNA was detected by FISH. (**C**) Quantitative analysis of CUG-repeat RNA detected by FISH. DMPK3′-CTG480 was transfected with EGFP or EGFP-MBNL1 mutants as indicated. Bars indicate average NCI (mean ± SD, $n = 3$). See also Supplementary Material, Figure S8A. (**D**) Structure of non-coding repeat construct (top). Long interrupted CTG or CAG repeat was inserted between CMV promoter and a V5 tag. Intracellular localization of CAG240 RNA co-transfected with EGFP (middle) or EGFP-MBNL1$_{42}$ (bottom). Nuclear staining with Hoechst33342 is included in the merged panels. (**E**) Quantitative analysis of nucleocytoplasmic localization of MBNL1 variants and repeat RNA in N2a cells co-expressing them. For RNA expression, repeat tracts in the DMPK3′ context were transfected with EGFP or EGFP-fused MBNL1 variants. Intracellular localization of EGFP and FISH signals in GFP/Alexa-546 double-positive cells was quantified (mean ± SD, $n = 3$). As a control, the localization of proteins was quantified in the absence of RNA constructs. (**F**) The effect of depletion of endogenous Mbnl1/2 on the RNA localization of DMPK3′-CTG480. Bar chart shows RNA localization as in (C). In (B and D), arrows indicate RNA foci. Scale bar indicates 10 μm.

polyQ-containing proteins. We next asked whether MBNL1 can alter the expression from apparently coding CAG repeats. For this purpose, the effect of MBNL1 on the localization of huntingtin exon 1 RNA was tested. Similar to the results in Figure 5, MBNL1 enhanced the nuclear localization of CAG-containing mRNA of huntingtin, as seen by FISH analysis

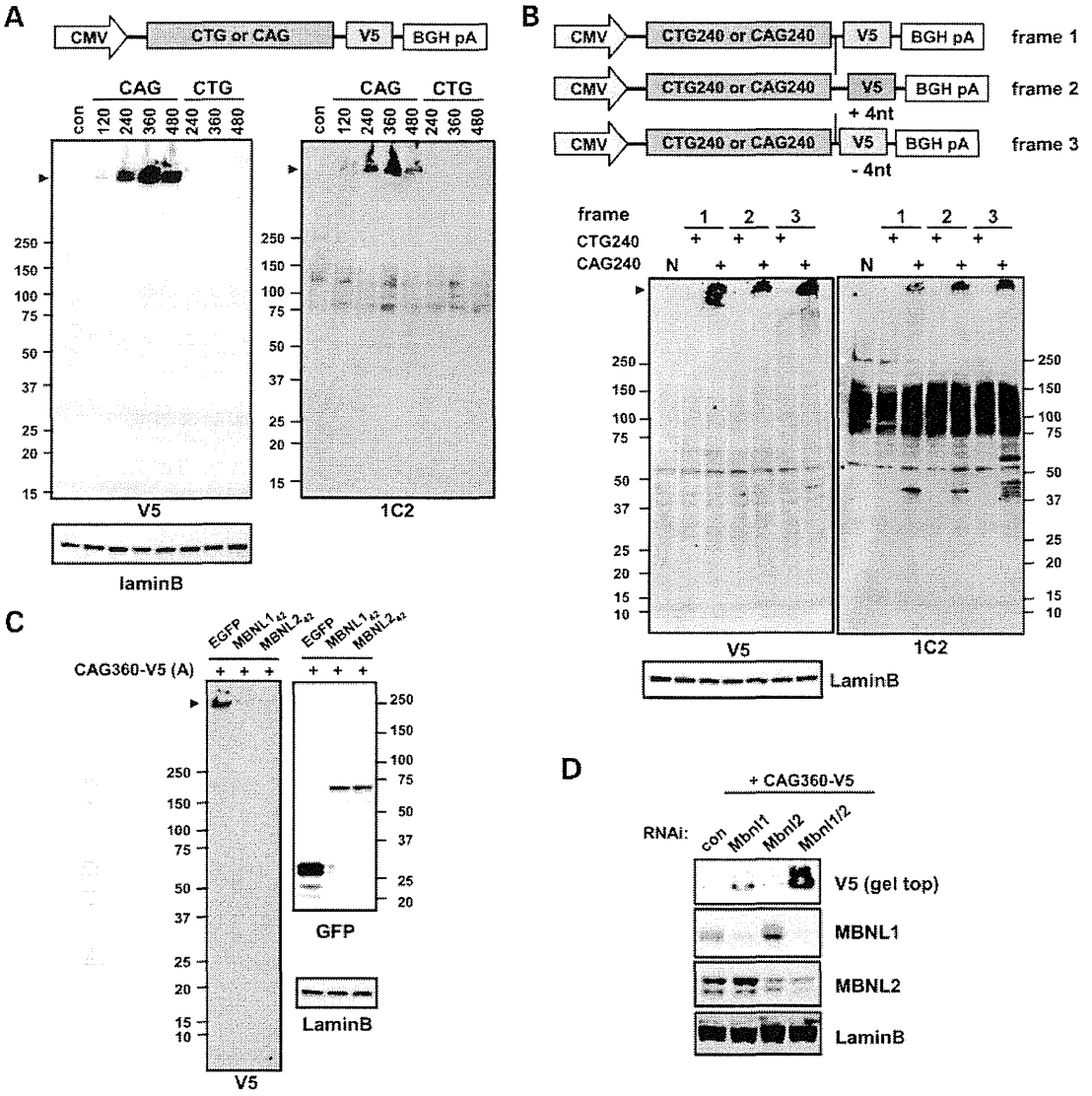


**Figure 6.** MBNL1 induces nuclear retention of repeat RNA. (**A**) Structure of non-coding repeat constructs. A long, interrupted CTG or CAG repeat was inserted between the CMV promoter and a V5 tag. There is no ATG codon located between the CMV promoter and the repeat tract. (**B**) Three different reading frames of V5 tag relative to the repeat tract were prepared. Repeat-derived proteins expressed from the CAG or CTG constructs. Protein expression was detected using anti-V5 and 1C2 antibodies. Lamin B served as a loading control. (**C**) Western analysis of the expression of CAG-repeat-derived proteins co-transfected with EGFP, EGFP-MBNL1$_{42}$ or EGFP-MBNL2$_{42}$. (**D**) Western analysis of the aberrant protein from the CAG360 construct with or without RNAi treatment for Mbnl1 and/or Mbnl2.

(Fig. 7A and B). When protein expression was examined, co-expression of EGFP-MBNL1 greatly reduced mutant huntingtin protein fused with RFP or V5 tag detected mainly at the gel top (Fig. 7C and Supplementary Material, Fig. S6B). Reduction in protein expression was much milder for huntingtin with CAG18 (Fig. 7C). Knockdown of both Mbnl1 and Mbnl2 resulted in the accumulation of mutant huntingtin (Fig. 7D). Depletion of Mbnl1/2 also led to cytoplasmic localization of mutant huntingtin mRNA and increased the cytotoxicity (Fig. 7E and F). RNAi-resistant mutant of MBNL1$_{42}$, MBNL1(res)$_{42}$, but not WT MBNL1$_{42}$ counteracted the effect of Mbnl1/2 knockdown (Supplementary Material, Fig. S10). The expression of RFP-fused CAG78 encoding Q78 was reduced by the expression of MBNL1 (Supplementary Material, Fig. S6C). Furthermore, MBNL1 reduced the expression of EGFP-fused CAG78 in different reading frames, which expressed polyglutamine, polyalanine or polyserine (Supplementary Material, Fig. S6D). Thus, the repressive effect of MBNL1 was independent of the context of huntingtin and was not restricted to the reading frame of polyQ. In conclusion, MBNL1 has a repressive effect on the expression of CAG-repeat-derived proteins.

In primary cultured mouse cortical neurons, we observed RNA foci and RNA localization in the nucleus and cytoplasm when DMPK3′-CTG480 or mutant huntingtin exon 1 was transfected with EGFP, whereas co-expression of MBNL1$_{42}$ enhanced the localization of repeat RNA in the nuclear foci (Fig. 8A and Supplementary Material, Fig. S11A). We also analyzed the striatum of HD model mice, the R6/2 strain. Transcripts

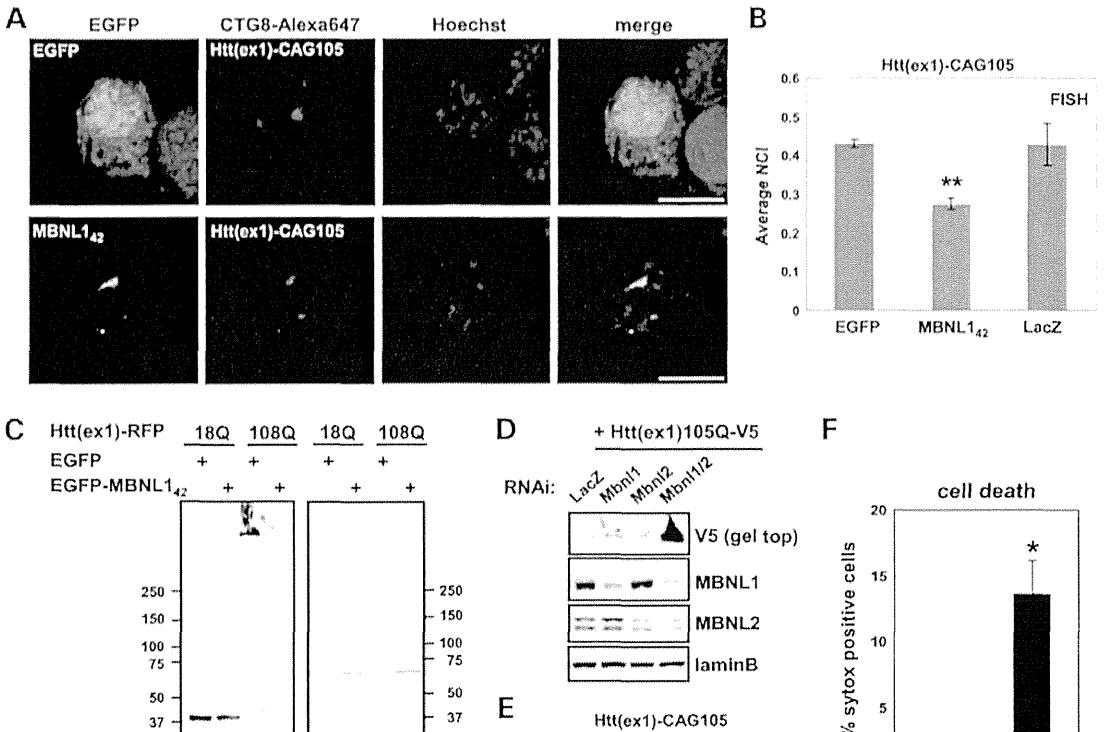


**Figure 7.** MBNL1 represses the expression of expanded huntingtin exon 1. (**A**) Nucleocytoplasmic localization of huntingtin exon 1 mRNA was altered by EGFP-MBNL1$_{42}$. The huntingtin construct contains CAG105 and a V5 tag sequence. The localization of huntingtin mRNA was detected by FISH using Alexa-647-labeled CTG8 oligonucleotide. Scale bar indicates 10 μm. (**B**) Quantitative localization analysis of (A). Bar chart indicates average NCI of the CAG-repeat RNA of huntingtin detected by FISH (mean ± SD, $n = 3$). $^{**}P < 0.005$ in comparison with EGFP (Tukey's multiple test). (**C**) Western analysis of the effect of EGFP-MBNL1 on the expression of RFP-fused huntingtin exon 1 with CAG18 or CAG108 in N2a cells. Transfected proteins were detected using anti-GFP and anti-RFP. (**D**) Western analysis of the effect of RNAi-mediated knockdown of MBNL1 and/or MBNL2 on the expression of huntingtin exon 1 with CAG105 in N2a cells. Huntingtin was detected by anti-V5 antibody. Knockdown of endogenous Mbnl1 and Mbnl2 was confirmed by specific antibodies. Lamin B serves as a loading control. (**E**) Huntingtin mRNA localization was altered by RNAi-mediated knockdown of Mbnl1 and Mbnl2 ($^{*}P < 0.05$, two-tailed $t$-test; $n = 3$). (**F**) The effect of depletion of Mbnl proteins on the cytotoxicity of mutant huntingtin. RNAi vectors expressing RFP were transfected with mutant huntingtin and treated with sytox green to label dying and dead cells. The fraction of cells with both green and red signals in the total transfected cells (red positive cells) were quantified (mean ± SD). $^{*}P = 0.022$ (two-tailed unpaired $t$-test, $n = 6$). RNAi treatment of Mbnl1/2 itself did not show cytotoxicity (data not shown).

from Huntingtin transgene were detected by FISH in the transgenic (TG) mice but not in the WT control (Fig. 8B). Mutant RNA were localized mostly in the cytoplasm, and we hardly observed RNA foci in the nucleus (Fig. 8B). Mbnl proteins were distributed in both nucleus and cytoplasm, and the staining patterns were similar between WT and TG mice, except for slight reduction of Mbnl2 staining in TG mice reflecting reduction in its mRNA level (Fig. 8C, Supplementary Material, Fig. S2G). Alternative splicing of Mbnl1 exon 7 was not altered in TG mice (Fig. 8D), suggesting that the function of Mbnl proteins was preserved in TG mice. In cells derived from striatal neurons of HD knock-in mice (StHdh 111/111), we did not detect specific FISH signals above background, as we observed diffuse staining in StHdh 7/7 that may reflect CAG-containing transcripts other than huntingtin (Fig. 8E, upper panels). However, overexpression of MBNL1$_{42}$ manifested RNA foci in StHdh 111/111 cells (Fig. 8E, lower panels). These results suggest that the expression of expanded-repeat transcripts is not sufficient for RNA foci formation in some cell-types. However, overexpression of MBNL proteins could induce RNA foci formation and/or nuclear RNA retention in such cells.

## DISCUSSION

### Interplay of conserved motifs determines nuclear localization of MBNL1

Though previous studies indicated that exon 7 of MBNL1 is involved in nuclear localization (11,31,40), the mechanism has been uncharacterized. Here, we clarified the role of exon 7 in nuclear localization. Two clusters of KR residues in the C-terminus region were essential for the function of the NLS.

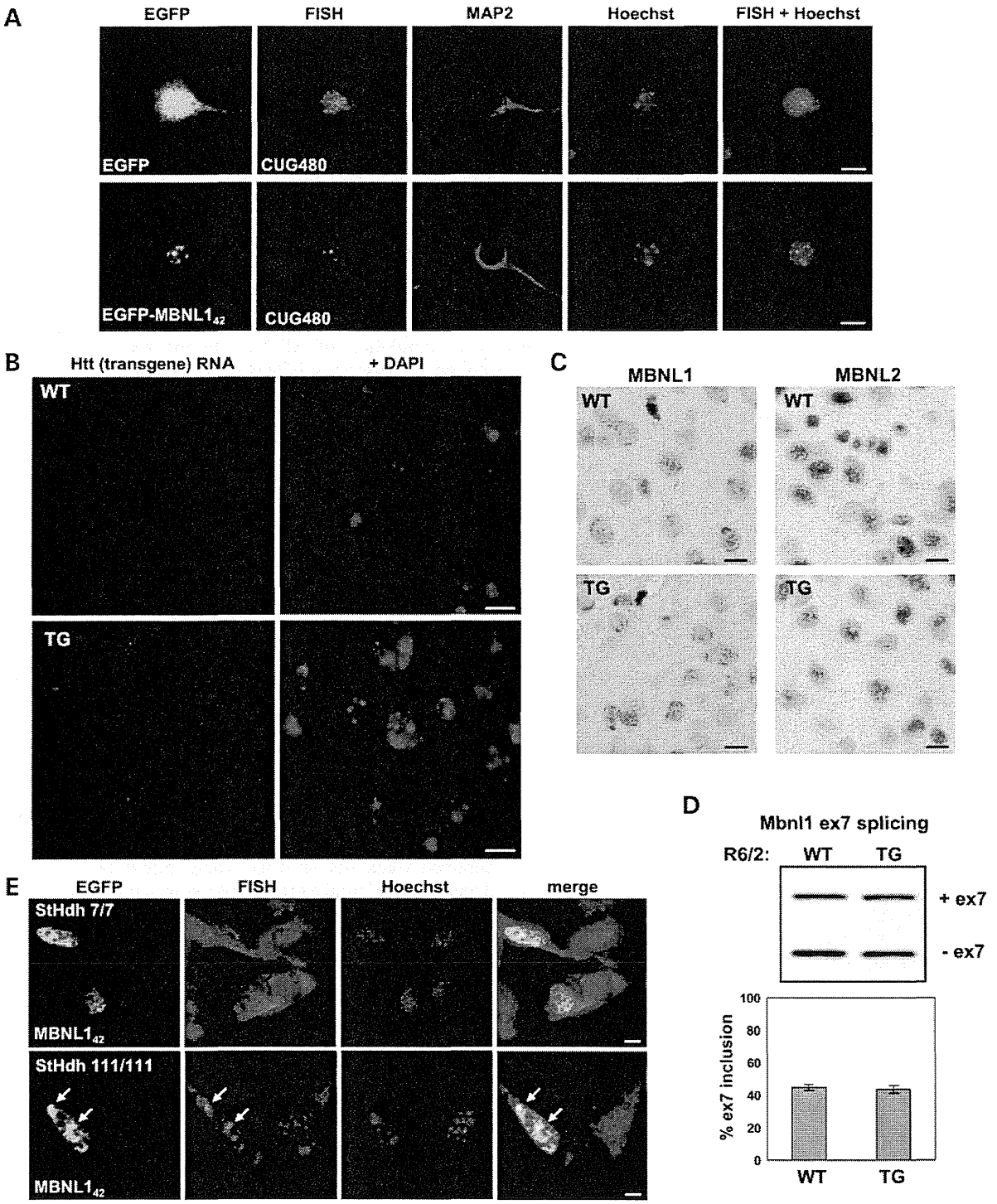


**Figure 8.** Localization of expanded RNA and MBNL proteins in cultured neurons and disease model cells. (**A**) Subcellular localization of mRNA of DMPK3′-CTG480 in primary cultured cortical neurons. The localization of repeat RNA was detected by FISH using Cy3-labeled CAG8. MAP2 serves as a marker of neurons. (**B**) FISH analysis of huntingtin transgene in the striatum of R6/2 mice. Mutant huntingtin mRNA was detected using DIG-labeled antisense RNA of human huntingtin. (**C**) Immunohistochemical analysis of Mbnl proteins in the striatum of R6/2 mice. (**D**) Splicing analysis of *Mbnl1* exon 7 in the striatum of R6/2 mice. The inclusion of exon 7 was not different between WT and TG mice ($n = 4$). (**E**) FISH analysis of StHdh cells transfected with EGFP-MBNL1$_{42}$. RNA foci were detected in StHdh 111/111 cells expressing EGFP-MBNL1 (arrows), but not in untransfected cells. Scale bars indicate 10 μm.

The activity of this NLS was inhibited by a specific inhibitor of the importin alpha/beta pathway (Bimax2), clearly demonstrating that the two KR clusters constitute a biNLS. We also found the second NLS that is dependent on the cell type. This latter NLS involves multiple conserved regions of the protein (Fig. 3E). None of the conserved motifs, namely KRPALE, ZnF1/2, RD/KWL and KxQL/NGR, was sufficient for the NLS activity by itself. Thus, the second NLS is consistent with a conNLS (39). It is currently unclear which import pathway mediates the activity of the conNLS. As ZnF1/2 is involved in RNA-binding activity (9), the status of RNA binding of MBNL1 might be related to the activity of the conNLS. We also found that MBNL1 contains two KRP(A)LE motifs in exons 7 and 8 and that each of these motifs can act as a part of the conNLS. Thus, $MBNL1_{42}$ can utilize two combinations of the conNLS in addition to the biNLS. Therefore, the inclusion of exon 7 acts as a switch for multiplying the pathways of nuclear import, ensuring efficient nuclear localization. Finally, the residues in the KRP(A)LE motifs are differently required in the context of biNLS and conNLS (Figs 2C and 3B), providing a unique example of motif usage in two different import pathways.

### Alternative splicing of exon 7 couples nuclear activity and intracellular localization of MBNL1

It has been suggested that MBNL1 autoregulates exon 7 splicing (11,29). Mouse studies also suggested cross-regulation among MBNL proteins (11,41). Here, we show that alternative splicing of exon 7 directly couples activity of MBNL proteins and nuclear localization of MBNL1 at the cellular level. Overexpression of any Mbnl paralogs repressed exon 7 inclusion, whereas simultaneous knockdown of Mbnl1 and Mbnl2 was required for efficient skipping of exon 7 (Fig. 4A and C). MBNL2 has an alternative exon corresponding to exon 7 of MBNL1, which is involved in the nuclear localization as well as splicing activity of MBNL2 (Supplementary Material, Fig. S2B and C) and is misregulated in DM1 (11). These explain why knockdown of either Mbnl1 or Mbnl2 showed weak effects on the splicing regulation and the repression of repeat-derived proteins (Figs 4C, 6D and 7D), because a reduction in one MBNL protein would be compensated by the increased nuclear isoforms of the other MBNL protein. Importantly, we showed that the NLS activity of MBNL1 closely reflects its splicing activity in a dynamic manner using a cell line expressing a fluorescent minigene, demonstrating that depletion of Mbnl proteins induced nuclear accumulation of the *Mbnl1* minigene (Fig. 4D and E). Moreover, co-expression of repeat RNA and MBNL1 *per se* did not induce translocation of cytoplasmic MBNL1 into the nucleus (Fig. 5E), suggesting that an extensive nuclear accumulation of MBNL1 observed in disease conditions may require enhanced production of nuclear isoforms. Altogether, we established an activity-dependent nuclear localization of MBNL1, in which MBNL isoforms containing exon 7 are predominantly produced when the nuclear amount or activity of MBNL proteins is low, whereas isoforms without exon 7 are predominantly produced after the amount of nuclear MBNL proteins reaches a certain amount. This mechanism would keep the amount of nuclear MBNL protein to a certain level in an autoregulatory feedback loop (Fig. 9A) and suggests the status of exon 7 splicing as an indicator of MBNL activity. Moreover, the increase of nuclear isoforms in disease conditions would accelerate the interaction between MBNL1 and CUG-repeat RNA in the nucleus (Fig. 9B).

### Repression of aberrant protein expression by MBNL1

Previous reports suggest that knockdown of MBNL proteins reduces RNA foci formation (32–34). We have shown that MBNL1 promotes the accumulation of expanded-repeat-containing RNA in the nucleus. As a consequence of nuclear RNA retention, MBNL1 represses the expression of mutant protein from transcripts containing an expanded CAG repeat (Figs 6 and 7). The effect of MBNL1 was observed for both coding and non-coding repeat tracts. To date, mislocalization or accumulation of MBNL1 in the nuclear foci has been thought to cause a loss of MBNL1 function leading to misregulation of alternative splicing (Fig. 9B). We propose that such accumulation of MBNL1 has another aspect, the repression of cytoplasmic transport of mutant RNA and its subsequent translation (Fig. 9B). This process might be regarded as an RNA quality control mechanism that limits the production of aberrant proteins from non-physiological transcripts with certain structural properties. Thus, MBNL1 is not simply a victim of expanded repeats but can be a guardian against protein toxicity. As MBNL proteins have cytoplasmic roles (42,43), we do not rule out the contribution of cytoplasmic MBNL1 in repressing repeat-derived proteins.

We observed protein expression from non-coding repeats with interruptions, which have not been examined previously (21). We observed the expression of polyQ-containing proteins from interrupted CAG repeats, which were repressed by MBNL proteins (Fig. 6C). Importantly, similar results were obtained for mutant huntingtin with a pure CAG repeat (Fig. 7C), suggesting that some important properties of pure repeats are retained in interrupted repeats. We hardly observed protein expression from interrupted CTG repeats even when Mbnl proteins were depleted (data not shown). There might be an orientation-dependent preference in the expression of these repeats. Consistently, RAN translation is dependent on the context surrounding the repeat (21). Alternatively, the homopolymeric proteins may have different susceptibility to degradation.

### RNA-binding proteins as potential modifiers of repeat expansion diseases

Our results suggest MBNL proteins as potential modifiers of polyQ diseases. Interestingly, MBNL1-containing RNA foci were recently detected in the fibroblasts of human HD and spinocerebellar ataxia type 3 (24,44). However, RNA foci were hardly detected in HD model mice and cells derived from HD knock-in mice, in our hands. There might be cell-type-dependent differences in the interaction between MBNL proteins and repeat RNA. Mbnl2 is a major MBNL protein in the mouse brain (Supplementary Material, Fig. S2F). We observed a slight reduction in the expression of Mbnl2 (Fig. 8C and Supplementary Material, Fig. S2G). However, alternative splicing of *Mbnl1* exon 7 was unchanged, suggesting that the activity of Mbnl proteins was not markedly altered in HD mice. In contrast, MBNL proteins are found in RNA foci in DM1 and SCA8 that is accompanied by splicing misregulation of *Mbnl1* exon 7 (8,11,28). These