平成 29 年度 博士論文

Matrin3 binds directly to intronic pyrimidine-rich sequences and controls alternative splicing

（Matrin3 はイントロン内のピリミジンが豊富な領域に直接結合し、選択的スプライシングを制御する）

植村　有里

大阪大学大学院 生命機能研究科 時空生物学講座 病因解析学研究室
大阪大学大学院 医学系研究科 神経遺伝子学教室

Abstract

Matrin3 is an RNA binding protein that is localized in the nuclear matrix and causative mutations in the *Matrin3* gene have been recently identified in the patients with amyotrophic lateral sclerosis (ALS). Therefore, it is pivotal to clarify the target RNAs and the function of Matrin3, especially in neuronal cells, as a first step to elucidate the disease mechanism. However, although various roles in RNA metabolism have been reported for Matrin3, in vivo target RNAs to which Matrin3 binds directly have not been investigated comprehensively so far. Here, I show that Matrin3 binds predominantly to intronic regions of pre-mRNAs. Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP) analysis using human neuronal cells revealed that Matrin3 recognized pyrimidine-rich sequences as binding motifs, including the polypyrimidine tract, a splicing-regulatory element. I further validated that Matrin3 bound directly to RNAs with pyrimidine-rich sequences by in vitro binding assay. Splicing-sensitive microarray analysis combined with PAR-CLIP data demonstrated that depletion of Matrin3 preferentially increased the inclusion of cassette exons that were adjacent to introns that contained Matrin3-binding sites. Given that a part of Matrin3 proteins are known to bind to polypyrimidine tract binding protein 1 (PTBP1), one of the splicing regulators

that bind to polypyrimidine tract, I further performed PAR-CLIP analysis against PTBP1 to examine the overlapping of the binding sites between Matrin3 and PTBP1. Then, I found that although most of the genes targeted by PTBP1 were also bound by Matrin3, Matrin3 could control alternative splicing in a PTBP1-independent manner, at least in part. These findings suggest that Matrin3 is a splicing regulator that targets intronic pyrimidine-rich sequences.

Table of Contents

I. Introduction

RNA-binding proteins (RBPs) regulate various aspects of RNA processing, such as splicing, editing, modification, transport and stability. Recently, it has been estimated that more than 1,100 RBPs are expressed in human tissues (Castello *et al.* 2013). These RBPs act on each target RNA cooperatively or competitively to ensure the precise regulation of each processing step. Indeed, mutations that lead to genetic diseases have been identified in approximately 10% of the genes that encode RBPs (Castello *et al.* 2013). Consequently, it is important to elucidate the functions of each RBP by, for instance, identifying the target RNAs and the binding sites of each RBP comprehensively. However, the number of RBPs whose functions or targets have been identified is very limited.

One of the most powerful methods to identify the targets of RBPs with high resolution is Crosslinking and Immunoprecipitation (CLIP), together with related methods, such as Photoactivatable-Ribonucleoside-Enhanced CLIP (PAR-CLIP) (Hafner *et al.* 2010). In PAR-CLIP, 4-thiouridine (4-SU) is added to the cell culture medium and is incorporated into RNAs in place of uridine. This results in strong crosslinking between the 4-SU in the RNA and bound RBPs. In addition, crosslinked 4-SU is converted into cytosine (C) rather than thymidine (T) during reverse

transcription, which thereby marks the binding site of the protein at a single-nucleotide resolution (Hafner *et al.* 2010; Lebedeva *et al.* 2011; Yokoshi *et al.* 2014; Li *et al.* 2015). After developing PAR-CLIP, the functions of RBPs have been identified by using this method. For example, PAR-CLIP against HuR, a highly conserved essential RBP, revealed that HuR bound predominantly to 3'UTR of mRNA and intron (Lebedeva *et al.* 2011). The binding sites of HuR in 3'UTR were highly overlapped with microRNA seeds, which indicated the involvement of HuR in microRNA processing. Furthermore, splicing pattern alaysis combined with PAR-CLIP date revealed that HuR-binding in intronic region regulates alternative splicing (Lebedeva *et al.* 2011). Another example is FMRP, a RBP encoded in Fragile X-mental Retardation 1 (*FMR1*) gene and loss of *FMR1* causes Fragile X syndrome (FXS) (Pieretti *et al.* 1991; Verkerk *et al.* 1991) Intriguingly, point mutations have been identified in one of two RNA binding domains as a cause of FXS, which was implicated to affect RNA binding ability. The PAR-CLIP analysis against wild-type and mutant FMRP successfully identified two RNA binding sequence motifs, which correspond to each RNA binding domains (Ascano *et al*. 2012). Then, in vitro binding assay using target RNAs containing these motifs as substrates revealed that mutant FMRP showed decreased affinity to natural target RNA. Furthermore, it was found that mutant FMRP failed to suppress translation from target

RNAs, which suggested that FMPR regulates translation efficiency (Ascano *et al*. 2012). Lastly, my labmate has successfully elucidated the function of Ataxin-2, which promotes mRNA stability by direct binding to target mRNAs, through the comprehensive identification of its targets by using PAR-CLIP (Yokoshi *et al.* 2014). Ataxin-2 bound predominantly to AU-rich sequences located in 3'UTR. Transcriptome-wide analysis of RNA expression after knockdown or overexpression of Ataxin-2 combined with PAR-CLIP date revealed that Ataxin-2 stabilized mRNA and promotes protein expression by binding directly to target mRNA. These examples suggest that PAR-CLIP is a powerful method for identification of as-yet-unknown targets and functions of RBPs.

Matrin3 is an RNA/DNA binding protein that was originally identified as one of the 12 nuclear matrix proteins, which were isolated from rat liver and separated by two-dimentional-Polyaclylamide Gel Electrophoresis (2D-PAGE) system (Belgrader *et al.* 1991; Nakayasu & Berezney 1991). This 2D-PAGE experiment showed that Matrin3 is slightly acidic 125 kDa protein (Belgrader *et al.* 1991; Nakayasu & Berezney 1991). On the other hand, another group identified DNA binding protein P130 from rat liver nuclei, which bound to highly repetitive DNA component (Hibino *et al.* 1992). Lately, P130 was turned out to be identical to Matrin3 (Hibino *et al.* 2006). Cloning and

sequencing of rat *Matrin3* gene revealed that rat Matrin3 was composed of 845 amino acids and highly conserved between human and rat with 96% identity (Belgrader *et al.* 1991). Furthermore, it was suggested that Matrin3 might be subjected to post-translational modification because its apparent molecular weight was higher than its theoretical molecular weight (95.3 kDa). Indeed, Matrin3 has many putative phosphorylation sites and the molecular weight of Matrin3 was slightly decreased after the treatment with phosphatase, suggesting that phosphorylation partially contributes to the molecular weight of Matrin3, which is higher than expected (Hibino *et al.* 1998). Intriguingly, the treatment with kinase inhibitor led to increasing amount of Matrin3 in the cytoplasm, suggesting that phosphorylation status modulates intracellular localization of Matrin3 (Hisada-Ishii *et al.* 2007). However, the molecular weight of Matrin3 treated with the phosphatase was still higher than that calculated, which suggests that Matrin3 is likely subjected to other post-translational modification besides phospholylation.

Matrin3 contains several functional domains, two zinc finger DNA binding domains (ZF1 and ZF2), two RNA recognition motifs (RRM1 and RRM2), nuclear localization signal (NLS), nuclear export signals (NES), and membrane retention signal (MRS) (**Fig. 1A**) (Hibino *et al.* 2006). Among these domains, NLS is reported to be

essential for cell proliferation, given that deletion of NLS in chicken cells leads to loss of Matrin3 from the nucleus and reduced cell number (Hisada-Ishii *et al.* 2007).

In vitro binding assay using rat Matrin3 protein showed that Matrin3 binds to highly repetitive DNA fragment and deletion of both ZF1 and ZF2 depleted DNA binding activity but not affected RNA binding affinity, which suggests that Matrin3 binds to DNA through ZF domains (Hibino *et al*. 1998; Hibino *et al.* 2006). Intriguingly, it was reported that the treatment of Matrin3 with the phosphatase reduces its DNA binding ability, which suggests that DNA binding is affected by phospholylation status of Matrin3 (Hibino *et al*. 1998). As a DNA binding protein, Matrin3 binds to matrix or scaffold attachment region (MAR/SAR) of genomic DNA and possibly modulates the activity of downstream promoters and promote transcription (Hibino *et al.* 2000). In addition, it has been reported that Matrin3 participates in various aspects of DNA metabolism, including DNA damage response, chromatin organization and transcriptional regulation (Zeitz *et al.* 2009; Salton *et al.* 2010; Skowronska-Krawczyk *et al.* 2014) (**Table 1**). In detail, Matrin3 is recruited to double strand brake of DNA through association with Protein-associated splicing factor (PSF) and $p54^{nrb}$ and knockdown of Matrin3 affects the repairment of the brake and cell cycle, which suggests that Matrin3 is involved in DNA damage response (Salton *et al.* 2010). Apart

from DNA damage response, Matrin3 might act in chromatin organization, given that yeast two-hybrid assay demonstrated that Matrin3 associates with chromatin remodeling factor CHD3 (Zeitz *et al*. 2009). In addition, Matrin3 plays a role in transcription as a DNA binding protein since Matrin3 is reported to localize to the enhancer regions together with POU-homeodomain transcription factor (Pit1) and this association is required for transcription of Pit1-target genes (Skowronska-Krawczyk *et al.* 2014).

In addition to the functions as a DNA binding protein, Matrin3 is reported to contribute to cell survival (**Table1**). For instance, knockdown of Matrin3 reduces cell proliferation and leads to necrosis of endothelial cells, which suggests that Matrin3 contributes to cell proliferation (Przygodzka *et al.* 2011). Importantly, Matrin3 was identified as one of the targets phospholylated by cAMP-dependent protein kinase (PKA) following activation of NMDA receptor in rat brain (Giordano *et al*. 2005). Phospholyalation of Matrin3 was followed by its degradation and inhibition of phosphorlylation prevented degradation of Matrin3 and cell death, which suggested phospholylation of Matrin3 as a rapid way of transferring information from synapses through NMDA receptors (Giordano *et al*. 2005). The proteome analysis revealed that Matrin3 associated with Calmodulin and Caspase-3/8, which might also contribute to rapid degradation of Matrin3 after NMDA receptor activation through cleavage of

Matrin3 within its RRM (Valencia *et al.* 2007). Matrin3 is also involved in the maintenance of interface between lamin and nuclear matrix by interacting with Lamina A/C (**Table1**). The mutation in *LMNA* gene that encodes Lamin A has been reported in certain myophathies and this mutation reduced the ability for the binding to Matrin3, which led to the increased distance between Lamin A and Matrin3 and reduced biophysical properties (Depreux *et al.* 2015).

Finally, it has been reported that Matrin3 also participates in various aspects of RNA metabolism, including mRNA stabilization, gene silencing and the retention of hyper-edited RNA, viral gene expression by binding to other proteins containing RBPs, such as DHX9, hnRNPK, $p54^{nrb}$, PSF, Argonaute 1 and 2 (AGO 1 and 2) and HIV-1 Rev (Zhang & Carmichael 2001; Höck *et al.* 2007; Valencia *et al.* 2007; Zeitz *et al.* 2009; Kula *et al.* 2011; Przygodzka *et al.* 2011; Salton *et al.* 2011; Yedavalli & Jeang 2011) (**Table 1)**. For instance, DHX9 and hnRNPK, both of which are RBPs, were identified as binding partners of Matrin3 and might regulate the stability of certain mRNAs by binding to Matrin3 in an RNA-dependent manner (Salton *et al.* 2011). Matrin3 was also identified as a member of the complex together with $p54^{nrb}$ and PSF (Zhang & Carmichael 2001). The complex binds to RNAs that contain abundant inosines as a consequence of RNA editing and this association seems to be essential to retain

hyperedited RNAs in the nucleus. Furthermore, Matrin3 was also reported to bind to AGO1 and 2 complex in an RNA dependent manner, thereby regulating the efficiency of microRNA-mediated gene-silencing (Höck *et al.* 2007). Finally, it was recently reported that Matrin3 regulates alternative splicing in a polypyrimidine tract binding protein 1 (PTBP1)-dependent and -independent manner in HeLa cells (Joshi *et al.* 2011; Coelho *et al.* 2015). As mentioned above, although many functions in RNA metabolism have been reported for Matrin3 by interacting other RBPs and mostly in an RNA-dependent manner, in vivo target RNAs that Matrin3 directly binds to have not been identified. Indeed, Matrin3 contains two RRMs and can bind to albumin mRNA fragment via RRM1 and RRM2 in vitro (Hibino *et al.* 2006). However, in vivo target RNAs and binding motifs of Matrin3 remain to be investigated comprehensively by the CLIP method so far.

Of note, causative mutations for amyotrophic lateral sclerosis (ALS) have been identified in the *Matrin3* gene (Johnson *et al.* 2014). ALS is one of the neurodegenerative diseases and characterized by selective degeneration of both upper and lower motor neurons, which leads to progressive muscle wasting and atrophy. Approximately 90% of ALS cases are sporadic and the rest is familial (Ling *et al*. 2013). The finding that ubiquitinated TDP-43 protein, which is a RBP predominantly localized

in nucleus under the normal condition, is accumulated in the cytoplasmic includions and cleared out from the nucleus in the affected regions of the most cases with both sporadic and familial ALS (Arai *et al*. 2006, Neumann *et al*. 2006) brought us the hypothesis that RNA metabolism is involved in the ALS pathology. In addition, although it is extremely rare, the point mutations in TAR DNA-binding protein (*TARDBP*) gene that encodes TDP-43 have been identified in the patients with familial ALS. Subsequently, many causative mutations have been identified in the genes encoding RBPs, such as Fused in Sarcoma (FUS) and hnRNPA1 (Iguchi *et al.* 2013), which further reinforced the possible involvement of RNA dysmetabolism in ALS pathology. TDP-43 is a 43 kDa RNA/DNA binding protein that contains two RRMs. The CLIP experiment against TDP-43 in mouse brain revealed that it targeted more than 6,000 RNAs and bound preferentially to intronic UG-rich sequence (Polymenidou *et al*. 2011; Tollervey *et al*. 2011). Knockdown of TDP-43 affected the expression level of target mRNAs and altered splicing pattern of a number of cassette exons (Polymenidou *et al*. 2011; Tollervey *et al*. 2011). In addition, TDP-43 is involved in various steps of RNA metabolism including regulation of RNA stability, transport, and microRNA biogenesis (Buratti & Baralle 2001; Kawahara & Mieda-Sato 2012; Lagier-Tourenne *et al*. 2012). Given that TDP-43 is cleared out from the nucleus in the affected neurons in the patients

(Arai *et al*. 2006, Neumann *et al*. 2006), the loss of the functions is thought to contribute to ALS pathology. FUS is another RBP whose gene is mutated in the patients with familial ALS, although its frequency is also quite rare (Kwiatkowski *et al*. 2009; Vance *et al*. 2009). The CLIP analysis agains FUS identified that it targeted more than 5,500 RNAs in human brains and had preference to GUGGU motif (Lagier-Tourenne *et al*. 2012). FUS also acts as a splicing regulator in the nucleus and is accumulated in the cytoplasmic inclusions in the affected regions of familial ALS patients (Kwiatkowski *et al*. 2009; Vance *et al*. 2009). Importantly, although both TDP-43 and FUS regulates splicing, the number of the common target sites is limited (45 genes) (Ling *et al*. 2013). Consequently, in terms of specifying the disease mechanism, it is pivotal to identify the target RNAs and functions of Matrin3, especially in neuronal cells.

In the present study, I performed PAR-CLIP for Matrin3 using human neuroblastoma-derived SH-SY5Y cells. I revealed that Matrin3 bound predominantly to intronic regions of pre-mRNAs with pyrimidine-rich sequences. An in vitro binding assay confirmed that Matrin3 could bind directly to RNAs with pyrimidine-rich sequences. Furthermore, combinatorial analysis of data obtained using a splicing-sensitive microarray with the PAR-CLIP data demonstrated that Matrin3 knockdown preferentially increased the inclusion of exons that contained

Matrin3-binding sites in adjacent introns. I found that this splicing control did not affect gene expression greatly and was, at least in part, independent of PTBP1.

II. Results

Matrin3 binds predominantly to intronic regions in pre-mRNAs

To identify the target RNAs to which Matrin3 binds directly in vivo, I performed PAR-CLIP analysis of endogenous Matrin3 using SH-SY5Y cells. After 4-SU had been added to the cell culture medium, the cells were cultured for 14 to 16 h and then irradiated with UV to crosslink the bound proteins with 4-SU-incorporated RNAs. Matrin3-RNA complexes were immunoprecipitated from cell lysate that had been prepared after the crosslinking and then the RNAs that were bound to Matrin3 were fragmented by treatment with micrococcal nuclease (MNase) and radiolabeled. After the samples had been separated by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE), the autoradiographic analysis demonstrated that certain RNAs were bound directly to Matrin3 (**Fig. 1B**). The RNA fragments were purified from the gel and reverse transcribed to generate a cDNA library, which was subjected to deep sequencing. I performed the PAR-CLIP analysis in duplicate and obtained 193,798,031 and 278,640,634 sequence reads for Replicate 1 and Replicate 2, respectively (**Table 2**). After the low quality and short reads (<20 nt) had been removed, the remaining sequence reads were aligned with the human reference genome (hg19) with allowance of up to two mismatches. Given that 4-SU crosslinked with protein is converted into C

during the reverse transcription (RT) step, thereby indicating the binding site of proteins at a single nucleotide resolution, the proportion of T-to-C mismatches among all the possible mismatches was calculated. It was found that ~80% of them were T-to-C in Replicate 1 (**Fig. 1C**), and similar results were also obtained for Replicate 2. This confirmed that the majority of the sequence reads was derived from RNAs bound directly to Matrin3. Finally, I identified 1,882,819 and 1,767,968 clusters of Matrin3-binding sites in 21,419 and 20,677 genes for Replicate 1 and Replicate 2, respectively, using PARalyzer (Corcoran *et al.* 2011) (**Table 2**). More than 92% of the target genes overlapped between Replicate 1 and Replicate 2 (**Fig. 1D**), and therefore I used the data from Replicate 1 for the subsequent analysis. Although 15% and 12% of the clusters were mapped to intergenic and repeat regions, respectively, the majority (72%) of the clusters were annotated to regions that encompassed protein-coding genes (mRNA). Furthermore, most of these clusters (90%) were located within intronic regions (**Fig. 1E**). These results suggest that nuclear Matrin3 binds predominantly to intronic regions in pre-mRNAs.

Gene Ontology (GO) analysis of the 13,012 protein-coding genes that contained Matrin3-binding sites revealed that Matrin3 targeted various types of gene broadly and there was no preference for the specific categories (**Table 3**). However, many of the

highest-ranked genes with abundant binding sites were involved in neuronal functions, including limbic system-associated membrane protein (*LSAMP*), autism susceptibility candidate 2 (*AUTS2*), neuron navigator 2 (*NAV2*), and calcium voltage-gated channel auxiliary subunit alpha 2delta1 (*CACNA2D1*) (**Table 4**). Furthermore, I examined 16 ALS-related genes (Taylor *et al.* 2016) and found that 14 genes contained Matrin3-binding sites, including Fused in sarcoma (*FUS*), *hnRNPA1*, and TAR DNA-binding protein (*TARDBP*), which encodes TDP-43 (**Table 5**). Among these genes, *MATR3*, which encodes Matrin3, contained the greatest number of Matrin3 binding sites, which suggests that Matrin3 expression might be autoregulated.

Pyrimidine-rich sequences are the preferred motifs targeted by Matrin3

Next, I examined the motif sequences that were recognized by Matrin3 by using cERMIT, which analyzes the sites that show a T-to-C conversion to identify binding motifs (Georgiev *et al.* 2010). An analysis of 7-mer motifs revealed that long pyrimidine-rich sequences were the top candidates for Matrin3 targets (**Fig. 2A** and **Table 6**). The polypyrimidine tract is a well-characterized motif that is composed of 10–20 nucleotides that are rich in pyrimidines and is located approximately 3–40 bp upstream of a 3' splicing site (3'ss) to control splicing (Reed 1996). Consequently, I

analyzed the position of the Matrin3-crosslinked sites that contained T-to-C conversions around 3'ss. This analysis demonstrated that Matrin3-binding sites were relatively highly enriched in the regions 5–15 bp upstream of 3'ss, which correspond to polypyrimidine tracts (**Fig. 2B**).

To confirm that Matrin3 binds directly to RNAs that contain pyrimidine-rich sequences, I planned to perform electrophoretic mobility shift assays (EMSAs) (Kawahara & Mieda-Sato 2012; Yokoshi *et al.* 2014). However, I failed to prepare recombinant Matrin3 by either expression in *E.coli* or using a baculovirus expression system. Consequently, I performed reverse EMSAs by radiolabeling proteins instead of RNAs (Filion *et al.* 2006). In this reverse assay, non-denatured proteins can migrate into the gel only when they are bound to nucleic acids, which have a negative charge. First, we validated this assay system using TDP-43, an RNA-binding protein, which was shown previously to bind to a UG-rich sequence, leading to a band shift in an EMSA, but did not bind to miR-143-3p (Kawahara & Mieda-Sato 2012). As expected, radiolabeled TDP-43 showed no migration in the absence of RNA (**Fig. 2C**). However, TDP-43 migrated efficiently into the gel when it was mixed with 12-UG-repeat RNA, but not with miR-143-3p. These data validated the application of reverse EMSA for RNA and RBPs. Then, I applied this reverse EMSA to Matrin3. After confirming the

expression of radiolabeled wild-type (WT) Matrin3 by autoradiography (**Fig. 2D**), I performed a reverse EMSA with RNA that contained pyrimidine-rich sequences derived from a certain region in Intron 1 (chr13; 102,901,094–102,901,248) of Fibroblast growth factor 14 (FGF14), which was one of the highest-ranking target sites of Matrin3 (**Table 4**). I termed this sequence FGF14a (**Fig. 2E**). I observed that Matrin3 migrated in the reverse EMSA in a manner that depended on the dose of FGF14a RNA (**Fig. 2F**). In contrast, this migration was abolished when cold Matrin3 was added as a competitor prior to the addition of radiolabeled Matrin3 (**Fig. 2G**), which suggests significant binding of Matrin3 to FGF14a. The binding efficiency was reduced when the length of the pyrimidine-rich sequence was shortened (FGF14b) (**Fig. 2H**). Furthermore, I could not observe substantial binding of Matrin3 to RNA in the absence of the pyrimidine-rich sequence (FGF14c). Next, I prepared radiolabeled Matrin3 without RNA recognition motifs (ΔRRM1+2), and this failed to bind to any FGF14 RNA fragments (**Fig. 2D**,**H**). Collectively, these results suggest that Matrin3 can bind directly to RNA that contains pyrimidine-rich sequences.

Matrin3 controls alternative splicing without affecting gene expression

Given that Matrin3 targets intronic pyrimidine-rich sequences in pre-mRNAs, I knocked

down Matrin3 by using small interfering RNA (siRNA) and analyzed how depletion of Matrin3 affected the gene expression and the splicing pattern by using a microarray (**Fig. 3A**). Although knockdown of Matrin3 significantly affected the expression level of 1,339 genes, which did not include Matrin3 mRNA *per se*, most of the changes were subtle or moderate. The expression of only 78 genes was either upregulated or downregulated more than 2-fold. Integration of the PAR-CLIP data revealed that only 14 of these 78 genes contained Matrin3-binding sites. Furthermore, I could observe no significant differences in the cumulative distributions of the expression changes between the genes with Matrin3-binding sites and those without binding sites (**Fig. 3B,C**). From these results, I concluded that binding of Matrin3 largely does not affect the expression of individual mRNAs.

Next, to examine the extent to which Matrin3 affects alternative splicing, I counted the number of cassette exons whose signal intensity was changed significantly after depletion of Matrin3 by setting the filter criteria as |splicing index value| ≥2 (Yao *et al.* 2016). In addition, given that signal intensity was also affected by the alteration of gene expression, I limited the subject to the genes whose expression level was decreased or increased less than 2-fold after Matrin3 knockdown. This analysis demonstrated that depletion of Matrin3 affected both exon inclusion and skipping to a similar extent.

Namely, 577 and 521 cassette exons were included and excluded significantly, respectively, after Matrin3 knockdown (**Fig. 4**). However, when I limited the subject to exons that were adjacent to introns that contained Matrin3-binding sites, exon inclusion predominated (~70%) (**Fig. 3D**). Matrin3-binding sites were enriched at both ends of the introns adjacent to these included exons, as compared with the excluded exons, and the density of the binding sites was highest in the region that corresponded to the polypyrimidine tract (**Fig. 3E**). Taken together, the data suggested that Matrin3 controls alternative splicing and an enrichment of Matrin3-binding sites in adjacent introns predominantly promotes exon skipping.

The target sites of Matrin3 are largely also recognized by PTBP1

PTBP1 is one of the RBPs that bind to polypyrimidine tract and regulate splicing (Valcarcel *et al.* 1996; Wagner & Garcia-Blanco 2001; Xue *et al.* 2009; Llorian *et al.* 2010). To evaluate the degree to which the target binding sites and genes of Matrin3 overlap with those of PTBP1, we performed PAR-CLIP analysis of endogenous PTBP1 in SH-SY5Y cells in duplicate (**Fig. 5A**). This yielded 80,471 and 77,832 clusters of PTBP1-binding sites from 202,953,476 and 247,390,487 sequence reads for Replicate 1 and Replicate 2, respectively (**Table 2**), which was only ~40% of the number of clusters

obtained for Matrin3. This difference might be due to differences in the amount of RNA fragments extracted from the gels to create the cDNA libraries or simply due to the different scale of gene regulation by the two proteins. Given that the targeted genes overlapped greatly between Replicate 1 and Replicate 2 (**Fig. 5B**), I used the data from Replicate 1 for the subsequent analysis. The motif analysis demonstrated that pyrimidine-rich motifs were the top target sequences recognized by PTBP1 as expected, and the motifs were very similar to those targeted by Matrin3 (**Fig. 5C** and **Table 7**). In addition, 96% of the PTBP1 target genes (**Fig. 5D**) and 77% of the PTBP1-binding sites (**Fig. 5E**) were also targeted by Matrin3. These results suggest that Matrin3 and PTBP1 largely share binding sites and favor pyrimidine-rich sequences.

Matrin3-mediated regulation of alternative splicing is partly independent of PTBP1

Given that a region of the Matrin3 protein binds to PTBP1 (Joshi *et al.* 2011; Coelho *et al.* 2015), I examined whether Matrin3-dependent control of alternative splicing required PTBP1. For this purpose, I knocked down PTBP1 in SH-SY5Y cells by using siRNA. However, this resulted in compensatory upregulation of PTBP2, a neuronal homologue of PTBP1, as reported previously (Boutz *et al.* 2007), which is usually

expressed at low levels in SH-SY5Y cells (**Fig. 6A**). Consequently, I knocked down both PTBP1 and PTBP2 and compared the pattern of splicing with that in Matrin3-depleted cells in which I had confirmed that the amount of PTBP1 was not affected (**Fig. 3A**). I selected three cassette exons that contained Matrin3-binding sites but no PTBP1-binding sites in the adjacent intron(s), namely Exon 14 (chr3; 9,478,534–9,478,572) of SET domain containing 5 (*SETD5*), Exon 20 (chr14; 76,203,909–76,203,950) of Tubulin Tyrosin Ligase Like 5 (*TTLL5*) and Exon 5a (chr12; 56,308,060–56,308,150) of Within BGCN (*WIBG*). Semi-quantitative RT-PCR analysis demonstrated that the inclusion of these exons was upregulated significantly only when Matrin3 was knocked down (**Fig. 6B,C,D**). In contrast, the inclusion of Exon 7 (chr7; 44,695,917–44,695,961) of Oxoglutarate Dehydrogenase (*OGDH*), which contains binding sites for both Matrin3 and PTBP1 in the adjacent introns, was promoted significantly after the knockdown of either Matrin3 or the PTBPs (**Fig. 6E**). I could not observe any significant alteration in the splicing pattern for Exon 32 (chr7; 73,480,274–73,480,327) of elastin (*ELN*) after the depletion of either Matrin3 or PTBPs (**Fig. 6F**). This gene contains no binding sites for either Matrin3 or PTBP1 in the adjacent introns. Collectively, the data suggest that Matrin3 controls alternative splicing likely through direct binding to adjacent introns in a manner that is independent of

PTBP1, at least in part.

III. Discussion

Although it has been reported that Matrin3 can bind directly to RNA in vitro (Hibino *et al.* 2006; Yamazaki *et al.* 2014) and iCLIP data for Matrin3 were utilized to study alternative splicing (Coelho *et al.* 2015), the in vivo targets of Matrin3 remain to be investigated comprehensively. In the present study, I performed PAR-CLIP analysis for Matrin3 for the first time and successfully identified transcriptome-wide target RNAs and binding motifs of endogenous Matrin3. This analysis revealed that the majority of the binding sites of Matrin3 are located in intronic regions of pre-mRNAs. Furthermore, I found that Matrin3 preferentially recognized pyrimidine-rich sequences as binding motifs, which I confirmed by an in vitro binding assay. This result might not be surprising, given that much of the core sequences in the RRMs of Matrin3 and PTBP1 are identical (Matsushima *et al.* 1996). Indeed, previous CLIP and PAR-CLIP analyses for HeLa cells (Xue *et al.* 2009; Ling *et al.* 2016) and our PAR-CLIP analysis for SH-SY5Y cells identified pyrimidine-rich sequences as the binding motifs of PTBP1 and I found that 77% of PTBP1-binding clusters were shared by Matrin3.

I revealed that Matrin3 controlled alternative splicing in a manner that was independent of PTBPs, at least in part, and acted predominantly as a repressor, as shown previously (Coelho *et al.* 2015). However, although Coelho *et al* (Coelho *et al.* 2015)

observed uniformly elevated binding of Matrin3 throughout introns adjacent to repressed exons, with no discrete peaks or enriched motifs, I found a discrete peak near the 3'ss, which corresponds to the location of the polypyrimidine tract. This discrepancy might be attributed to the differences in methodology or the quality of the CLIP libraries. The mechanism that underlies Matrin3-mediated splicing regulation remains unknown. However, it has been reported that PTBP1 and hnRNP C bind to pyrimidine-rich sequences and regulate alternative splicing by competing with U2AF65, which also binds to pyrimidine-rich sequences and promotes exon inclusion (Singh *et al.* 1995; Sauliere *et al.* 2006; Zarnack *et al.* 2013). Indeed, U2AF65 shows a binding peak near the 3'ss of cassette exons (Zarnack *et al.* 2013; Coelho *et al.* 2015), which is similar to the peak that we observed for Matrin3. Therefore, although further study is needed to elucidate how Matrin3, PTBP1 and hnRNP C cooperatively or competitively share the same binding sites to regulate splicing, it is possible that Matrin3 also controls alternative splicing by competing with U2AF65 (**Fig. 7**).

It has been reported that point mutations in the *Matrin3* gene cause ALS and distal myopathy (Senderek *et al.* 2009; Johnson *et al.* 2014). The four point mutations originally identified (Ser85Cys, Phe115Cys, Pro154Ser and Thr622Ala) (**Fig. 8**) have no effect on the localization of Matrin3 in Chinese Hamster Ovary (CHO) cells

(Gallego-Iradi *et al.* 2015). However, Matrin3 is found to be redistributed in the cytoplasm in some neurons of the spinal cord in transgenic mice overexpressing wild-type Matrin3, which exhibit muscle atrophy and weakness (Moloney *et al.* 2016). Therefore, although it still remains unknown whether either loss of function or gain of function is the main cause of these diseases, identification of the target RNAs and the function of Matrin3 could be the first step to elucidate the disease mechanism. Indeed, by using neuronal cells for PAR-CLIP analysis, I identified various genes that are involved in neuronal functions and those that are mutated in familial ALS as targets of Matrin3 (**Table 5**). I further found that Matrin3 participates in the control of splicing of these genes. It is known that the nervous system shows high levels of alternative splicing (Grosso *et al.* 2008) and aberrant splicing was observed in the patients with ALS (Ling *et al.* 2015; Prudencio *et al.* 2015). Intriguingly, two more point mutations in *Matrin3* gene have been recently identified (Ala62Thre and Val394Met) (**Fig. 8**) by whole-exsome sequencing (Lin *et al.* 2015; Leblond *et al.* 2016) and Val394Met is located in RRM1, possibly affecting RNA-binding ability. Therefore, future studies should clarify whether Val394Met mutation affects splicing activity of Matrin3. In addition, it also remains undetermined whether disease-associated point mutations in *Matrin3*, which are located outside the RRMs, affect the regulation of splicing by

Matrin3. There is a possibility that these point mutations affect binding ability to other proteins, leading to aberrant splicing. Indeed, dysregulation of splicing is induced by ALS-linked point mutations in *TARDBP* and *FUS* (Arnold *et al.* 2013; Reber *et al.* 2016) and Matrin3 associates with TDP-43 and FUS (Ling *et al.* 2010; Salton *et al.* 2011; Johnson *et al.* 2014; Kamelgarn *et al.* 2016; Yamaguchi & Takanashi 2016). Therefore, future studies should address how Matrin3 and TDP-43 or FUS coordinately regulate alternative splicing and how disease-associated point mutations in *Matrin3* affect binding to TDP-43 or FUS in terms of the elucidation of disease pathogenesis.

IV. Experimental procedures

Cell culture

SH-SY5Y cells were maintained in Dulbecco's Modified Eagle's Medium: Nutrient mixture F-12 (DMEM/F-12; Thermo Fisher Scientific) supplemented with 15% fetal bovine serum (FBS; Thermo Fisher Scientific) as described previously (Kawahara & Mieda-Sato 2012). Cells were maintained at 37°C in the presence of 5% CO_2.

Antibodies

The primary antibodies used for PAR-CLIP and western blotting were as follows: goat polyclonal anti-Matrin3 antibody (sc-55723; Santa Cruz), goat polyclonal anti-hnRNP I (PTBP1) antibody (sc-16547; Santa Cruz), mouse polyclonal anti-PTBP2 antibody (H00058155-A01, Abnova) and mouse monoclonal anti-GAPDH antibody (M171-3; MBL). Normal goat IgG (sc-2028; Santa Cruz) was used as a control in the PAR-CLIP.

Western blot analysis

Total cell lysates were separated by using SDS-PAGE, transferred to a polyvinylidene difluoride (PVDF) membrane (Bio-Rad), and immunoblotted with primary antibodies using the SNAP i.d.® 2.0 Protein Detection System (Merck Millipore).

PAR-CLIP analysis

PAR-CLIP analysis was performed as described previously (Hafner *et al.* 2010; Lebedeva *et al.* 2011; Yokoshi *et al.* 2014; Li *et al.* 2015) with some modifications. Briefly, SH-SY5Y cells were cultured on 60 15-cm culture dishes in the presence of 100 μM 4-SU (Sigma) for 14–16 h. After the medium had been aspirated, the cells were crosslinked on ice by irradiation with 365 nm UV at 150 mJ/cm^2 in a UV Crosslinker (FS-1500L, Funakoshi). The cells were collected and resuspended in two volumes of NP-40 Lysis buffer (50 mM HEPES pH 7.5, 150 mM KCl, 2 mM EDTA, 0.5% NP-40, 0.5 mM DTT) with 1× complete, EDTA-free Protease Inhibitor Cocktail (PI; Roche). After incubation with rotation for 30 min at 4°C, the cell homogenates were homogenized 15 times with a tight pestle in an ice-cold Dounce tissue grinder (Wheaton). Then, whole cell lysate was collected by centrifugation at 20,000 × g for 30 min at 4°C and treated with 5 U/ml RNase T1 (Thermo Fisher Scientific) for 15 min at 22°C. After the antibody for either Matrin3 or PTBP1 or normal goat IgG had been conjugated to magnetic Protein G Dynabeads (Thermo Fisher Scientific), 10 ml of the whole cell lysate were incubated with the beads (25 μl of beads and 6 μg of antibody per ml of cell lysate) at 4°C for 1 h with rotation. Then, after the beads had been resuspended with Micrococcal Nuclease (MNase) buffer (50 mM Tris-HCl pH 7.4, 5

mM $CaCl_2$), 0.2 gel U/μl MNase (NEB) were added and the mixture was incubated at 37°C for 5 min. The beads were treated with 0.5 U/μl Calf Intestinal Alkaline Phosphatase (CIP) in Dephospho buffer (50 mM Tris-HCl pH 7.9, 0.1 M NaCl, 10 mM $MgCl_2$, 1 mM DTT) with 1× PI at 37°C for 10 min with gentle shaking. Then, a 3'-linker (L3; 5'-p-gugucagucacuuccagcgg-puromycin-3') that had been synthesized by Dharmacon was ligated to the RNA fragments in Ligase buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 0.2 mg/ml BSA, 10 mM ATP) with 0.5 U/μl T4 RNA Ligase (Thermo Fisher Scientific) through overnight incubation at 16°C in a Thermomixer comfort (Eppendorf). To radiolabel the RNA fragments, the beads were incubated with [γ-^{32}P]-ATP (0.5 μCi/μl) and 1 U/μl T4 Polynucleotide Kinase (NEB) in PNK buffer (70 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 5 mM DTT) with 1× PI at 37°C for 30 min. After three washes with Wash buffer (50 mM Tris-HCl pH 7.4, 0.3 M NaCl, 0.05% NP-40, 1 mM DTT), the beads were resuspended in 50 μl of 1× NuPAGE LDS Sample Buffer (Thermo Fisher Scientific) and boiled at 95°C for 10 min to elute protein-RNA complexes. The eluted samples were loaded onto NuPAGE Novex 4–12% Bis-Tris gels (Thermo Fisher Scientific) and separated by size. Gels were exposed for 1 h and the radiolabeled RNA fragments were visualized on a phosphorimager (GE Healthcare). After the bands that corresponded to the target proteins had been excised,

the protein-RNA complexes were electroeluted using a D-tube Dialyzer (Merck Millipore). Then, the samples were treated with 2× Proteinase K buffer (100 mM Tris-HCl pH 7.5, 100 mM NaCl, 20 mM $MgCl_2$) that contained 2 mg/ml Proteinase K (Thermo Fisher Scientific) at 55°C for 30 min and the RNA fragments were recovered by phenol extraction and 2-propanol precipitation with glycogen. The purified RNAs were ligated with 2 μM 5'-linker (L5; 5'-OH-agggaggacgaugcgg-OH-3'), which had been synthesized by Hokkaido System Science (Sapporo, Japan), in Ligase buffer with 0.5 U/μl T4 RNA Ligase and 1× PI at 16°C overnight. The 5'-linker-ligated RNAs were loaded onto 10% TBE-Urea gels (Thermo Fisher Scientific) and separated by size. The gels were exposed for 1 h and the radiolabeled RNAs were visualized on a phosphorimager. Fragments of RNA between 60 and 100 nt in length were excised and the RNAs were eluted by incubation at 65°C for 2 h in RNase-free water and recovered by phenol extraction and 2-propanol precipitation with glycogen. After treatment with 0.1 U/μl DNase I (Thermo Fisher Scientific) at 37°C for 20 min, the denatured RNAs were reverse transcribed into cDNAs using 0.5 μM P3 primer (5'-ccgctggaagtgactgacac-3') and SuperScript III (Thermo Fisher Scientific) in accordance with the manufacturer's instructions. The resultant cDNAs were amplified by PCR using AccuPrime Pfx SuperMix (Thermo Fisher Scientific) and the primers

DSFP5 (5'-aatgatacggcgaccaccgactatggatacttagtcagggaggacgatgcgg-3') and DSFP3 (5'-caagcagaagacggcatacgaccgctggaagtgactgacac-3'). The PCR products were loaded onto 10% TBE gels (Thermo Fisher Scientific) and stained with SYBR Safe DNA Gel Stain Dye (Thermo Fisher Scientific). After fragments between 125 and 150 bp in length had been excised and incubated in Diffusion buffer (0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, 0.1% SDS) at 50°C for 30 min, the PCR products were purified using a QIAEX II Gel Extraction Kit (QIAGEN). They were then subjected to deep sequencing using a HiSeq 2000 Sequencing System (Illumina) with the primer SSP1 (5'-ctatggatacttagtcagggaggacgatgcgg-3') at BGI (Beijing, China) and Macrogen (Kyoto, Japan).

Mapping of the Reads

The reads that were obtained by deep sequencing were edited by removing reads that contained ambiguous bases as described previously (Yokoshi *et al.* 2014). After the 3' adaptor sequence had been stripped away by an in-house Perl script, reads that were less than 20 nt in length were removed. The remaining reads were collapsed using fastx_collapser in FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), and then mapped to the reference genome (hg19) by using the Bowtie algorithm with allowance

of up to two mismatches. For each read, only the best mismatch-stratum was reported for up to ten different locations (Yokoshi *et al.* 2014). The binding sites were identified using PARalyzer (Corcoran *et al.* 2011). I defined a cluster as a site that was represented by at least five reads with T-to-C conversions at two or more locations.

Annotation of the Clusters

Annotation data for lincRNA, misc_RNA, Mt_rRNA, Mt_tRNA, rRNA, tRNA, snoRNA, snRNA, repetitive elements, piwi-interacting RNA, microRNA and protein-coding transcripts were downloaded from multiple sources described previously (Ascano *et al.* 2012: Yokoshi *et al.* 2014). Then, the clusters that had been assigned to the category of protein-coding transcript were divided further into coding sequences (CDS), 5' untranslated region (UTR), 3'UTR and intron. The clusters that had been assigned to multiple categories were assigned a single annotation as described previously (Ascano *et al.* 2012: Yokoshi *et al.* 2014). I identified the genes that harbored the clusters by matching the sequences of each cluster with the gene database GRCh37. p13, which was downloaded from Ensembl using the BioMart tool.

Identification of Motifs

To identify motifs in the clusters generated by PARalyzer, the cERMIT tool (Georgiev *et al.* 2010) was downloaded and used as described previously (Ascano *et al.* 2012; Yokoshi *et al.* 2014).

Gene Ontology Analysis

Matrin3 target genes were subjected to GO analysis using PantherDB (Thomas *et al.* 2006).

siRNA knockdown

The siRNA duplexes for Matrin3 (5'-gggaugauugggaagaaaaTT-3') were synthesized by GeneDesign (Osaka, Japan). Pre-designed siRNAs for PTBP1 and PTBP2 were obtained from Dharmacon (ON-TARGETplus siRNA). Negative control siRNA was obtained from Bioneer (AccuTarget Negative Control siRNA). Transfection of siRNAs was performed using a Neon® Transfection System 100 μl Kit (Thermo Fisher Scientific). After trypsinization, 5×10^6 SH-SY5Y cells were washed twice in Dulbecco's phosphate-buffered saline (DPBS) and resuspended in 100 μl of Buffer R that contained 0.2 μM siRNA. For the double transfection of PTBP1 and PTBP2

siRNAs, 100 μl of Buffer R that contained 0.2 μM siRNA for each PTBP was used. Then, electroporation was performed with the following parameters: pulse voltage, 1,100 V; pulse width, 50 ms; pulse number, 1. Twenty-four hours after transfection, the cell culture medium was replaced with fresh DMEM/F-12 medium, and the cells were incubated for an additional 24 h and then harvested to extract either total RNA or cell lysate.

Preparation of cell lysates

To prepare cell lysates, the collected cells were washed twice in DPBS. After centrifugation at 5,000 × g for 5 min at 4°C, five volumes of lysis buffer 1 (20 mM Tris-HCl pH 7.9, 25% glycerol, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM PMSF, and 0.5 mM DTT) with 1×PI were added to the pellet. After transfer to a 1.5-ml tube, the sample was repeatedly frozen in liquid nitrogen and thawed on ice five times. After centrifugation at 15,000 × g for 20 min at 4°C, the supernatant was transferred to a new 1.5-ml tube and stored at -80ºC until use.

RNA preparation

Total RNA was extracted using TRIZOL reagent followed by further purification with a

PureLink RNA Mini Kit in accordance with the manufacturer's instructions (Thermo Fisher Scientific).

Microarray Analysis

Transcriptome-wide splicing analysis was performed using purified total RNAs and a GeneChip® Human Transcriptome Array 2.0 (Affymetrix) at TaKaRa (Shiga, Japan). After the signals had been normalized using Affymetrix GeneChip Command Console Software 4.0 and Expression Console Software 1.4.1, alternative splicing and differential expression at the gene level were analyzed using Affymetrix Transcriptome Analysis Console Software 3.0.

Analysis of the splicing pattern

Purified total RNAs were reverse transcribed into cDNAs using random hexamer primers and Superscript III in accordance with the manufacturer's instructions. The resultant cDNAs were amplified by PCR with Phusion Hot Start II High-Fidelity DNA Polymerase (Thermo Scientific) and the following primers: SETD5_FW (5'-agcacacacctacaagcatcac-3'), SETD5_DW (5'-agcttcttcatactggtcagtcc-3'), TTLL5_FW (5'-tgtgtgccaagatcctgc-3'), TTLL5_DW (5'-cagaccaagcacagaaccac-3'), WIBG_FW

(5'-ttaggaggtctgggcgagaag3'), WIBG_DW (5'-gggcacatatccttctttcac3'), OGDH_FW (5'-aggggtcaccacattgcaaa3'), OGDH_DW (5'-ctgattcctgtcccccgatg3'), ELN_ FW (5'-gtatacctccagctgcagccgc-3') and ELN_DW (5'-ctgggaaaatgggagacaatccg-3'). PCR products were separated on Novex 10% TBE gels (Thermo Fisher Scientific) and the intensity of each band was quantified using Image J software (http://rsb.info.nih.gov/ij/) after staining with SYBR Safe DNA Gel Stain Dye.

Synthesis of expression constructs for Matrin3 and its deletion mutants

A full-length expression construct for human Matrin3 tagged with a HaloTag at the N-terminus was obtained from Promega and the coding sequence was transferred to the expression vector p3xFLAG-CMV10 (Sigma). The resultant construct was termed p3xFLAG-CMV10-Matrin3 WT. The expression construct for the mutant in which both RRMs were deleted, which was termed p3xFLAG-CMV10-Matrin3 ΔRRM1+2, was generated by a two-step PCR extension method as described previously (Kawahara & Mieda-Sato 2012). In brief, two Matrin3 fragments were amplified with the primers hMatrin3_FW1_del (5'-ttgagttatggtaggatggcag-3') and hMatrin3_dRRM1_RRM2_DW1 (5'-tcctcagaaccagttttttatacatgtgtc-3') and with the primers pCMV10_DW1 (5'-tattaggacaaggctggtgggc-3') and hMatrin3_dRRM1_RRM2_FW1

(5'-cctgcaaggacctagacacatgtataaaaa-3'), respectively. After gel purification, a second round of amplification was then carried out with the primers hMatrin3_FW1_del and pCMV10_DW1 using a mixture of the two fragments as the template. The resultant PCR products were digested with AvrII and BamHI-HF (NEB) and then inserted into p3xFLAG-CMV-10-Matrin3 that had been linearized with the same restriction enzymes. Finally, the regions that corresponded to Matrin3 WT and ΔRRM1+2 were amplified by PCR and inserted into the pTNT vector (Promega) after digestion with KpnI and NotI (NEB).

Preparation of Matrin3- and TDP-43-target RNAs

To generate the expression constructs for Matrin3-target RNAs, which were used for reverse EMSA, I amplified two fragments that corresponded to a certain intronic region of FGF14 (**Fig. 2E**) using human genomic DNA (Novagen) as the template with the primers FGF14ab_FW (5'-atatgtaatcagacatcaag-3') and FGF14a_DW (5'-catattacatgcagaaagaa-3') and with the primers FGF14ab_FW and FGF14b_DW (5'-acgaaagctgaaaaaagcat-3'), respectively. The PCR products were gel extracted and inserted into the vector pCR BluntII-TOPO vector (Thermo Fisher Scientific). The resultant constructs were termed FGF14a and FGF14b. To construct FGF14c (**Fig. 2E**),

the synthesized complementary oligonucleotides (5'-atatgtaatcagacatcaag-3' and 5'-cttgatgtctgattacatat-3') were mixed in the annealing buffer (10 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA), incubated at 95°C for 15 min, and cooled down gradually. The annealed oligonucleotides were inserted into pCR BluntII-TOPO. After the direction of the insertion had been confirmed by sequencing, the construct was linearized with the restriction enzyme SpeI (NEB) and in vitro transcription was performed with T7 RNA polymerase using a Riboprobe® In Vitro Transcription System (Promega) at 37°C for 1 h. After treatment with RQ1 RNase-Free DNase (Promega), the resultant RNA fragments were recovered by phenol extraction and 2-propanol precipitation with glycogen. After the concentration had been measured using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific), the final concentration was adjusted to 1–100 ng/μl and the RNAs were stored at -80°C until use. The target RNAs of TDP-43, namely miR-143-3p and UG12, were synthesized by Hokkaido System Science (Sapporo, Japan), as described previously (Kawahara & Mieda-Sato 2012).

Reverse EMSA assay

The Matrin3 WT, ΔRRM1+2 and TDP-43 proteins were synthesized using a TnT® T7

Quick Coupled Transcription/Translation System (Promega) with EasyTag™ L-[^{35}S]-Methionine (PerkinElmer). The 25-μl reaction mixtures, which contained Quick Master Mix (20 μl), [^{35}S]-Methionine (1 μl), PCR enhancer (1.5 μl), and plasmid DNA (450 ng), were incubated at 30°C for 90 min. Cold non-radiolabeled Matrin3 WT was synthesized in a reaction mixture that did not contain [^{35}S]-Methionine. Aliquots of the reaction mixtures were loaded onto NuPAGE Novex 4–12% Bis-Tris gels (Thermo Fisher Scientific) and the gels were exposed overnight to visualize the radiolabeled proteins on a phosphorimager. Then, the reverse EMSA was carried out as described previously (Filion *et al.* 2006) with modifications. In brief, the radiolabeled proteins (3 μl for Matrin3 and 1.5 μl for TDP-43) were incubated at 25°C for 25 min with the indicated amounts of RNA (0.5–50 ng for Matrin3 and 50 ng for TDP-43) in a 20-μl reaction mixture that contained 10 mM Tris-HCl pH 8.0, 25 mM KCl, 10 mM NaCl, 1 mM $MgCl_2$, 10% glycerol and 0.5 mM DTT. The reaction products were loaded on a 6% DNA Retardation Gel (Thermo Fisher Scientific) and separated by electrophoresis at 200 V for 2 h or 20 min for Matrin3 and TDP-43, respectively, with cooling. Then, the gels were exposed overnight and visualized using a phosphorimager.

In the competition assay, the indicated amount of non-radiolabeled Matrin3 protein (1.0–5.0 × the volume of radiolabeled Matrin3 protein) was incubated with 50

ng of FGF14a RNA at 25°C for 10 min prior to the addition of radiolabeled protein.

Statistical analysis

The statistical analyses of the differences in splicing patterns and expression levels using the microarray data were carried out using one-way ANOVA and Welch's *t*-test, respectively. All other analyses were carried out using the Mann–Whitney *U*-test. All values are displayed as the mean ± standard error of the mean (SEM). Statistical significance is displayed as $p<0.05$ (*).

Accession numbers

The PAR-CLIP data used in this study are available through the DNA Data Bank of Japan (DDBJ) under accession numbers DRA005124 and DRA005839. Microarray data are available through the Gene Expression Omnibus (GEO) under accession number GSE86967.

V. References

Arai, T., Hasegawa, M., Akiyama, H., Ikeda, K., Nonaka, T., Mori, H., Mann, D., Tsuchiya, K., Yoshida, M., Hashizume, Y. & Oda, T. (2006) TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun*. **351**, 602-611.

Arnold, E.S., Ling, S.C., Huelga, S.C. *et al.* (2013) ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43. *Proc. Natl Acad. Sci. USA* **110**, E736-E745.

Ascano, M., Mukherjee, N., Bandaru, P., Miller, J.B., Nusbaum, J.D., Corcoran, D.L., Langlois, C., Munschauer, M., Dewell, S., Hafner, M., Williams, Z., Ohler, U. & Tuschl, T. (2012) FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature* **492**, 382-386.

Belgrader, P., Dey, R. & Berezney, R. (1991) Molecular-Cloning of Matrin-3 - a 125-Kilodalton Protein of the Nuclear Matrix Contains an Extensive Acidic Domain. *J. Biol. Chem.* **266**, 9893-9899.

Buratti, E. & Baralle, F.E. (2001) Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. *J. Biol. Chem.* **276**, 36337-36343.

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

Castello, A., Fischer, B., Hentze, M.W. & Preiss, T. (2013) RNA-binding proteins in Mendelian disease. *Trends Genet.* **29**, 318-327.

Coelho, M.B., Attig, J., Bellora, N., Konig, J., Hallegger, M., Kayikci, M., Eyras, E., Ule, J. & Smith, C.W.J. (2015) Nuclear matrix protein Matrin3 regulates alternative splicing and forms overlapping regulatory networks with PTB. *EMBO J.* **34**, 653-668.

Corcoran, D.L., Georgiev, S., Mukherjee, N., Gottwein, E., Skalsky, R.L., Keene, J.D. & Ohler, U. (2011) PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data. *Genome Biol.* **12**, R79.

Depreux, F.F., Puckelwartz, M.J., Augustynowicz, A., Wolfgeher, D., Labno, C.M., Pierre-Louis, D., Cicka, D., Kron, S.J., Holaska, J. & McNally, E.M. (2015) Disruption of the lamin A and matrin-3 interaction by myopathic LMNA mutations. *Hum. Mol. Genet.* **24**, 4284-4295.

Filion, G.J., Fouvry, L. & Defossez, P.-A. (2006) Using reverse electrophoretic mobility shift assay to measure and compare protein-DNA binding affinities. *Anal. Biochem.* **357**,

156-158.

Gallego-Iradi, M.C., Clare, A.M., Brown, H.H., Janus, C., Lewis, J. & Borchelt, D.R. (2015) Subcellular Localization of Matrin 3 Containing Mutations Associated with ALS and Distal Myopathy. *PLoS One* **10**, e0142144.

Georgiev, S., Boyle, A.P., Jayasurya, K., Ding, X., Mukherjee, S. & Ohler, U. (2010) Evidence-ranked motif identification. *Genome Biol.* **11,** R19.

Giordano, G., Sanchez-Perez, A.M., Montoliu, C., Berezney, R., Malyavantham, K., Costa, L.G., Calvete, J.J. & Felipo, V. (2005) Activation of NMDA receptors induces protein kinase A-mediated phosphorylation and degradation of matrin 3. Blocking these effects prevents NMDA-induced neuronal death. *J. Neurochem*. **94**, 808-818.

Grosso, A.R., Gomes, A.Q., Barbosa-Morais, N.L., Caldeira, S., Thorne, N.P., Grech, G., von Lindern, M. & Carmo-Fonseca, M. (2008) Tissue-specific splicing factor gene expression signatures. *Nucleic Acids Res.* **36**, 4823-4832.

Höck, J., Weinmann, L., Ender, C., Rüdel, S., Kremmer, E., Raabe, M., Urlaub, H. & Meister, G. (2007) Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep* **8**, 1052-1060.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.C., Munschauer, M., Ulrich, A., Wardle, G.S.,

Dewell, S., Zavolan, M. & Tuschl, T. (2010) Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. *Cell* **141**, 129-141.

Hibino, Y., Nakamura, K., Asano, S. & Sugano, N. (1992) Affinity of a highly repetitive bent DNA for nuclear scaffold proteins from rat liver. *Biochem. Biophys. Res. Commun.* **184**, 853-858.

Hibino, Y., Ohzeki, H., Hirose, N. & Sugano, N. (1998) Involvement of phosphorylation in binding of nuclear scaffold proteins from rat liver to a highly repetitive DNA component. *Biochim. Biophys. Acta.* **1396**, 88-96.

Hibino, Y., Ohzeki, H., Sugano, N. & Hiraga, K. (2000) Transcription modulation by a rat nuclear scaffold protein, P130, and a rat highly repetitive DNA component or various types of animal and plant matrix or scaffold attachment regions. *Biochem. Biophys. Res. Commun.* **279**, 282-287.

Hibino, Y., Usui, T., Morita, Y., Hirose, N., Okazaki, M., Sugano, N. & Hiraga, K. (2006) Molecular properties and intracellular localization of rat liver nuclear scaffold protein P130. *Biochim. Biophys. Acta.* **1759**, 195-207.

Hisada-Ishii, S., Ebihara, M., Kobayashi, N. & Kitagawa, Y. (2007) Bipartite nuclear localization signal of matrin 3 is essential for vertebrate cells. *Biochem. Biophys. Res. Commun.* **354**, 72-76.

Iguchi, Y., Katsuno, M., Ikenaka, K., Ishigaki, S. & Sobue, G. (2013) Amyotrophic lateral sclerosis: an update on recent genetic insights. *J. Neurol.* **260**, 2917-2927.

Johnson, J.O., Pioro, E.P., Boehringer, A. *et al.* (2014) Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nat. Neurosci.* **17**, 664-666.

Joshi, A., Coelho, M.B., Kotik-Kogan, O., Simpson, P.J., Matthews, S.J., Smith, C.W. & Curry, S. (2011) Crystallographic analysis of polypyrimidine tract-binding protein-Raver1 interactions involved in regulation of alternative splicing. *Structure* **19**, 1816-1825.

Kamelgarn, M., Chen, J., Kuang, L., Arenas, A., Zhai, J., Zhu, H. & Gal, J. (2016) Proteomic analysis of FUS interacting proteins provides insights into FUS function and its role in ALS. *Biochim. Biophys. Acta.* **1862**, 2004-2014.

Kawahara, Y. & Mieda-Sato, A. (2012) TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proc. Natl Acad. Sci. USA* **109**, 3347-3352.

Kula, A., Guerra, J., Knezevich, A., Kleva, D., Myers, M.P. & Marcello, A. (2011) Characterization of the HIV-1 RNA associated proteome identifies Matrin 3 as a nuclear cofactor of Rev function. *Retrovirology* **8**. 60

Kwiatkowski, T.J., Jr., Bosco, D.A., Leclerc, A.L. et al. (2009) Mutations in the

FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* **323**, 1205-1208.

Lagier-Tourenne, C., Polymenidou, M., Hutt, K.R. et al. (2012) Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat. Neurosci.* **15**, 1488-1497.

Lebedeva, S., Jens, M., Theil, K., Schwanhausser, B., Selbach, M., Landthaler, M. & Rajewsky, N. (2011) Transcriptome-wide Analysis of Regulatory Interactions of the RNA-Binding Protein HuR. *Mol. Cell* **43**, 340-352.

Leblond, C.S., Gan-Or, Z., Spiegelman, D. *et al.* (2016) Replication study of MATR3 in familial and sporadic amyotrophic lateral sclerosis. *Neurobiol. Aging* **37**, 209 e217-209 e221.

Li, Q., Uemura, Y. & Kawahara, Y. (2015) Cross-Linking and Immunoprecipitation of Nuclear RNA-Binding Proteins. *Nuclear Bodies and Noncoding RNAs: Methods Mol. Biol.*, 247-263.

Lin, K.P., Tsai, P.C., Liao, Y.C., Chen, W.T., Tsai, C.P., Soong, B.W. & Lee, Y.C. (2015) Mutational analysis of MATR3 in Taiwanese patients with amyotrophic lateral sclerosis. *Neurobiol. Aging* **36**, 2005 e2001-2004.

Ling, J.P., Chhabra, R., Merran, J.D., Schaughency, P.M., Wheelan, S.J., Corden, J.L. &

Wong, P.C. (2016) PTBP1 and PTBP2 Repress Nonconserved Cryptic Exons. *Cell Rep.* **17**, 104-113.

Ling, J.P., Pletnikova, O., Troncoso, J.C. & Wong, P.C. (2015) TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science* **349**, 650-655.

Ling, S.-C., Albuquerque, C.P., Han, J.S., Lagier-Tourenne, C., Tokunaga, S., Zhou, H. & Cleveland, D.W. (2010) ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS. *Proc. Natl Acad. Sci. USA* **107**, 13318-13323.

Ling, S.C., Polymenidou, M. & Cleveland, D.W. (2013) Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79**, 416-438.

Llorian, M., Schwartz, S., Clark, T.A., Hollander, D., Tan, L.-Y., Spellman, R., Gordon, A., Schweitzer, A.C., de la Grange, P. & Ast, G. (2010) Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. *Nat. Struct. Mol. Biol.* **17**, 1114-1123.

Matsushima, Y., Ohshima, M., Sonoda, M. & Kitagawa, Y. (1996) A family of novel DNA-binding nuclear proteins having polypyrimidine tract-binding motif and arginine/serine-rich motif. *Biochem. Biophys. Res. Commun.* **223**, 427-433.

Moloney, C., Rayaprolu, S., Howard, J., Fromholt, S., Brown, H., Collins, M., Cabrera,

M., Duffy, C., Siemienski, Z., Miller, D., Swanson, M.S., Notterpek, L., Borchelt, D.R. & Lewis, J. (2016) Transgenic mice overexpressing the ALS-linked protein Matrin 3 develop a profound muscle phenotype. *Acta Neuropathol. Commun.* **4**, 122.

Nakayasu, H. & Berezney, R. (1991) Nuclear Matrins - Identification of the Major Nuclear Matrix Proteins. *Proc. Natl. Acad. Sci. USA* **88**, 10312-10316.

Neumann, M., Sampathu, D.M., Kwong, L.K. et al. (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130-133.

Pieretti, M., Zhang, F.P., Fu, Y.H., Warren, S.T., Oostra, B.A., Caskey, C.T. & Nelson, D.L. (1991) Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell* **66**, 817-822.

Polymenidou, M., Lagier-Tourenne, C., Hutt, K.R. et al. (2011) Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* **14**, 459-468.

Prudencio, M., Belzil, V.V., Batra, R., Ross, C.A., Gendron, T.F., Pregent, L.J., Murray, M.E., Overstreet, K.K., Piazza-Johnston, A.E. & Desaro, P. (2015) Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat. Neurosci.* **18**, 1175.

Przygodzka, P., Boncela, J. & Cierniewski, C.S. (2011) Matrin 3 as a key regulator of endothelial cell survival. *Exp. Cell Res.* **317**, 802-811.

Reber, S., Stettler, J., Filosa, G., Colombo, M., Jutzi, D., Lenzken, S.C., Schweingruber, C., Bruggmann, R.m., Bachi, A. & Barabino, S.M. (2016) Minor intron splicing is regulated by FUS and affected by ALS‚Äêassociated FUS mutants. *EMBO J.* **35**, 1504-1521.

Reed, R. (1996) Initial splice-site recognition and pairing during pre-mRNA splicing. *Curr. Opin. Genet. Dev.* **6**, 215-220.

Salton, M., Elkon, R., Borodina, T., Davydov, A., Yaspo, M.L., Halperin, E. & Shiloh, Y. (2011) Matrin 3 Binds and Stabilizes mRNA. *PLoS One* **6**. e23882.

Salton, M., Lerenthal, Y., Wang, S.Y., Chen, D.J. & Shiloh, Y. (2010) Involvement of Matrin 3 and SFPQ/NONO in the DNA damage response. *Cell Cycle* **9**, 1568-1576.

Sauliere, J., Sureau, A., Expert-Bezancon, A. & Marie, J. (2006) The polypyrimidine tract binding protein (PTB) represses splicing of exon 6B from the beta-tropomyosin pre-mRNA by directly interfering with the binding of the U2AF65 subunit. *Mol. Cell. Biol.* **26**, 8755-8769.

Senderek, J., Garvey, S.M., Krieger, M., Guergueltcheva, V., Urtizberea, A., Roos, A., Elbracht, M., Stendel, C., Tournev, I. & Mihailova, V. (2009) Autosomal-dominant

distal myopathy associated with a recurrent missense mutation in the gene encoding the nuclear matrix protein, matrin 3. *Am. J. Hum. Genet.* **84**, 511-518.

Singh, R., Valcarcel, J. & Green, M.R. (1995) Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science* **268**, 1173.

Skowronska-Krawczyk, D., Ma, Q., Schwartz, M., Scully, K., Li, W.B., Liu, Z.J., Taylor, H., Tollkuhn, J., Ohgi, K.A., Notani, D., Kohwi, Y., Kohwi-Shigematsu, T. & Rosenfeld, M.G. (2014) Required enhancer-matrin-3 network interactions for a homeodomain transcription program. *Nature* **514**, 257-261.

Taylor, J.P., Brown, R.H., Jr. & Cleveland, D.W. (2016) Decoding ALS: from genes to mechanism. *Nature* **539**, 197-206.

Thomas, P.D., Kejariwal, A., Guo, N., Mi, H., Campbell, M.J., Muruganujan, A. & Lazareva-Ulitsky, B. (2006) Applications for protein sequence‚Äìfunction evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Res.* **34**, W645-W650.

Tollervey, J.R., Curk, T., Rogelj, B. et al. (2011) Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat Neurosci.* **14**, 452-458.

Valcarcel, J., Gaur, R.K., Singh, R. & Green, M.R. (1996) Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA

[corrected]. *Science* **273**, 1706-1709.

Valencia, C.A., Ju, W. & Liu, R. (2007) Matrin 3 is a Ca2+/calmodulin-binding protein cleaved by caspases. *Biochem. Biophys. Res. Commun.* **361**, 281-286.

Vance, C., Rogelj, B., Hortobagyi, T. et al. (2009) Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* **323**, 1208-1211.

Verkerk, A.J., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F.P. & et al. (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65**, 905-914.

Wagner, E.J. & Garcia-Blanco, M.A. (2001) Polypyrimidine tract binding protein antagonizes exon definition. *Mol. Cell. Biol.* **21**, 3281-3288.

Xue, Y., Zhou, Y., Wu, T., Zhu, T., Ji, X., Kwon, Y.-S., Zhang, C., Yeo, G., Black, D.L. & Sun, H. (2009) Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. *Mol. Cell* **36**, 996-1006.

Yamaguchi, A. & Takanashi, K. (2016) FUS interacts with nuclear matrix-associated protein SAFB1 as well as Matrin3 to regulate splicing and ligand-mediated transcription.

Sci. Rep. **6**, 35195.

Yamazaki, F., Kim, H.H., Lau, P., Hwang, C.K., Iuvone, P.M., Klein, D. & Clokie, S.J.H. (2014) pY RNA1-s2: A Highly Retina-Enriched Small RNA That Selectively Binds to Matrin 3 (Matr3). *PLoS One* **9,** e88217.

Yao, Y., Shang, J., Song, W., Deng, Q., Liu, H. & Zhou, Y. (2016) Data for the gene expression profiling and alternative splicing events during the chondrogenic differentiation of human cartilage endplate-derived stem cells under hypoxia. *Data Brief* **7**, 1438-1442.

Yedavalli, V.S.R.K. & Jeang, K.T. (2011) Matrin 3 is a co-factor for HIV-1 Rev in regulating post-transcriptional viral gene expression. *Retrovirology* **8**. 61

Yokoshi, M., Li, Q., Yamamoto, M., Okada, H., Suzuki, Y. & Kawahara, Y. (2014) Direct binding of Ataxin-2 to distinct elements in 3'UTRs promotes mRNA stability and protein expression. *Mol. Cell* **55**, 186-198.

Zarnack, K., König, J., Tajnik, M., Martincorena, I.i., Eustermann, S., Stévant, I., Reyes, A., Anders, S., Luscombe, N.M. & Ule, J. (2013) Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell* **152**, 453-466.

Zeitz, M.J., Malyavantham, K.S., Seifert, B. & Berezney, R. (2009) Matrin 3:

Chromosomal Distribution and Protein Interactions. *J. Cell. Biochem.***108**, 125-133.

Zhang, Z. & Carmichael, G.G. (2001) The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* **106**, 465-475.

VI. Figures

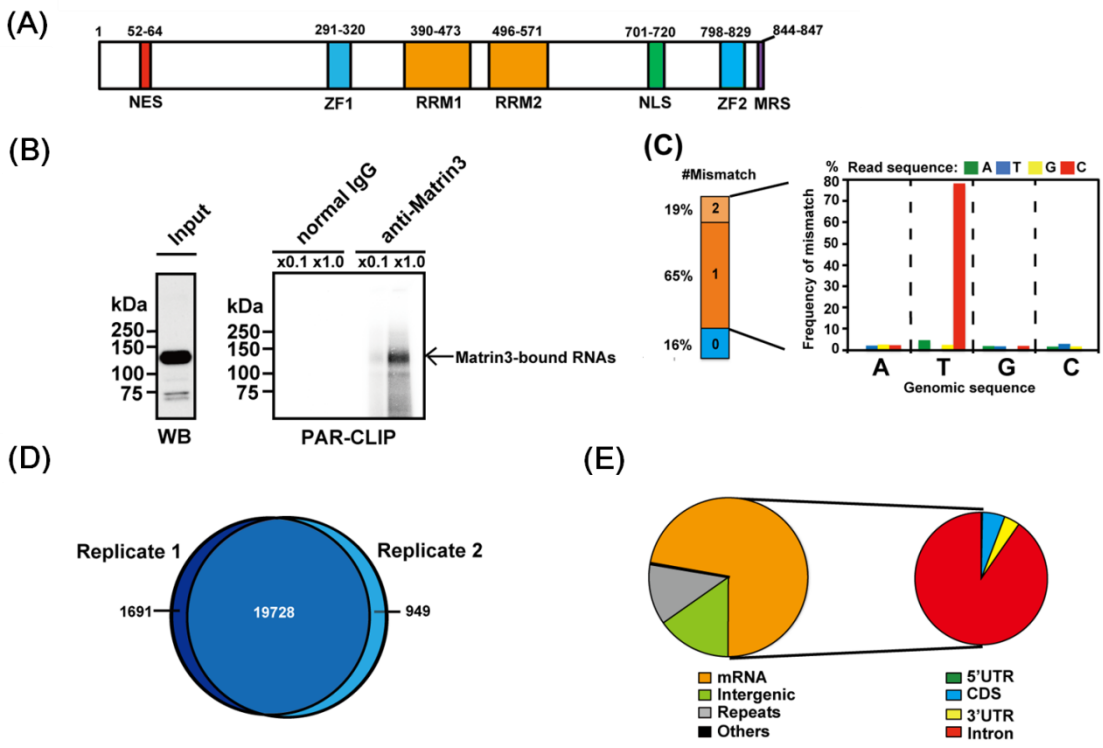

Figure 1 Identification of target RNAs of Matrin3 in SH-SY5Y cells using PAR-CLIP. (A) Schematic diagram of human Matrin3. NES, nuclear export signal; ZF, zinc finger; RRM, RNA recognition motif; NLS, nuclear localization signal; MRS, membrane retention signal. (B) Western blot (WB) analysis (left panel) and autoradiographic analysis (right panel) for immunoprecipitated Matrin3, which was crosslinked with radiolabeled RNA fragments. The samples were loaded on SDS gels without dilution (×1.0) or after dilution 1/10 (×0.1). (C) The number of mismatches (left panel) and the frequency of each nucleotide mismatch (right panel) in PAR-CLIP reads of Replicate 1 aligned to the human genomic sequence. (D) Confirmation of the reproducibility of the

PAR-CLIP analysis for Matrin3. A Venn diagram shows the numbers of overlapping and nonoverlapping genes between Replicate 1 and Replicate 2 of the PAR-CLIP analysis for Matrin3. (E) Distribution of the binding sites of Matrin3 in different types of RNA (left panel) and in pre-mRNAs (right panel).

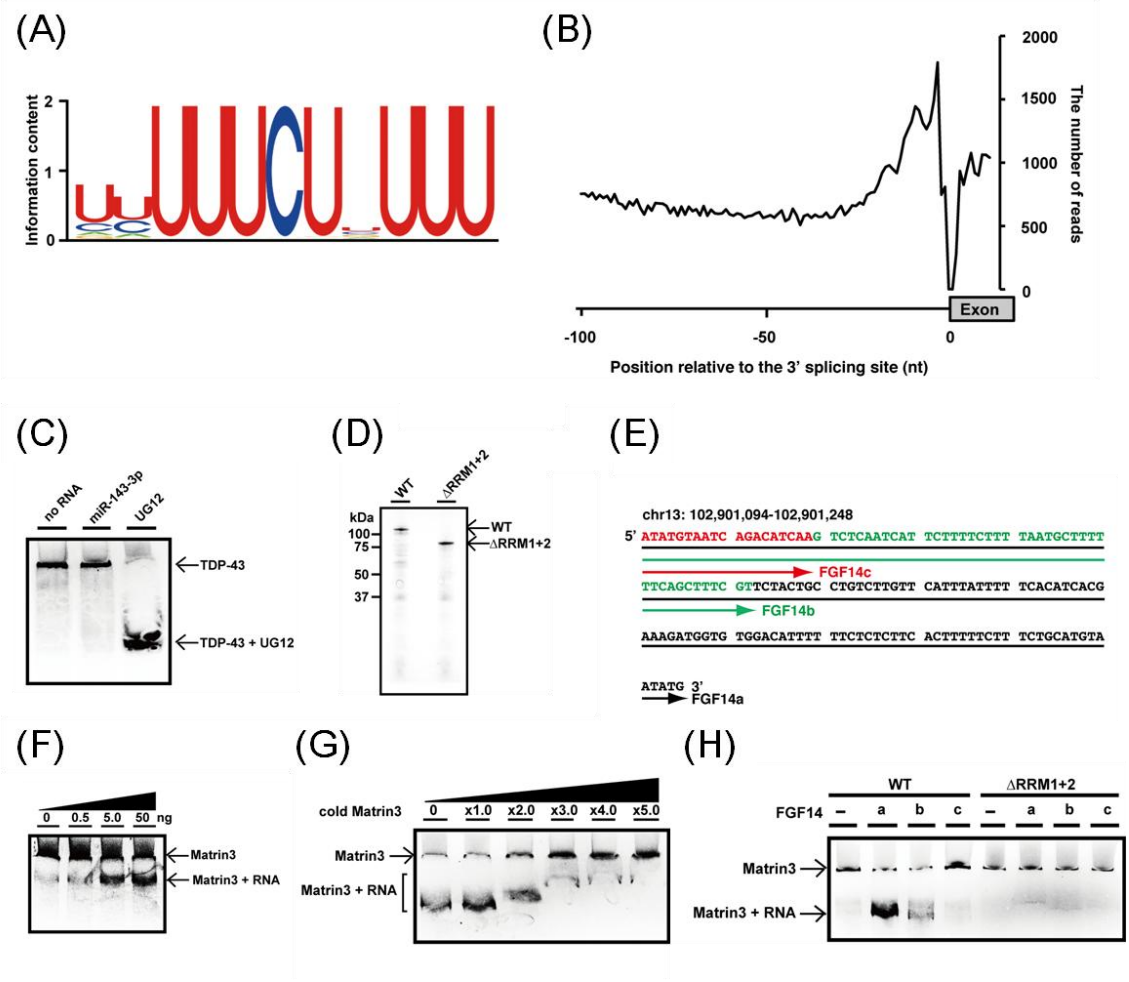

Figure 2 Analysis of Matrin3 recognition motifs. (A) A representative Matrin3-binding motif. (B) The location of the crosslinked sites that contained a T-to-C conversion relative to the 3' splicing site. (C) Results of a reverse EMSA with TDP-43 and either 50 ng of the indicated RNA or no RNA. We demonstrated previously by using EMSA that 12-UG-repeat RNA (UG12) binds efficiently to TDP-43, whereas miR-143-3p does not (Kawahara & Mieda-Sato 2012). (D) Phosphorimage of SDS gels that resolved

radiolabeled wild-type (WT) Matrin3 and mutant Matrin3 in which both RRMs were deleted (ΔRRM1+2); both proteins were synthesized by in vitro transcription/translation. (E) The partial sequence of FGF14 Intron 1 used for reverse EMSA. The nucleotide sequences underlined correspond to each RNA fragment. (F) Results of a reverse EMSA assay with WT Matrin3 and the indicated amount of FGF14a RNA. (G) Results of a reverse EMSA with radiolabeled Matrin3 and 50 ng of FGF14a RNA in the presence of the indicated amount of non-radiolabeled Matrin3 as a competitor. (H) Results of a reverse EMSA with either WT Matrin3 or ΔRRM1+2 and either 50 ng of the indicated RNA (FGF14a, FGF14b or FGF14c) or no RNA (shown as -).

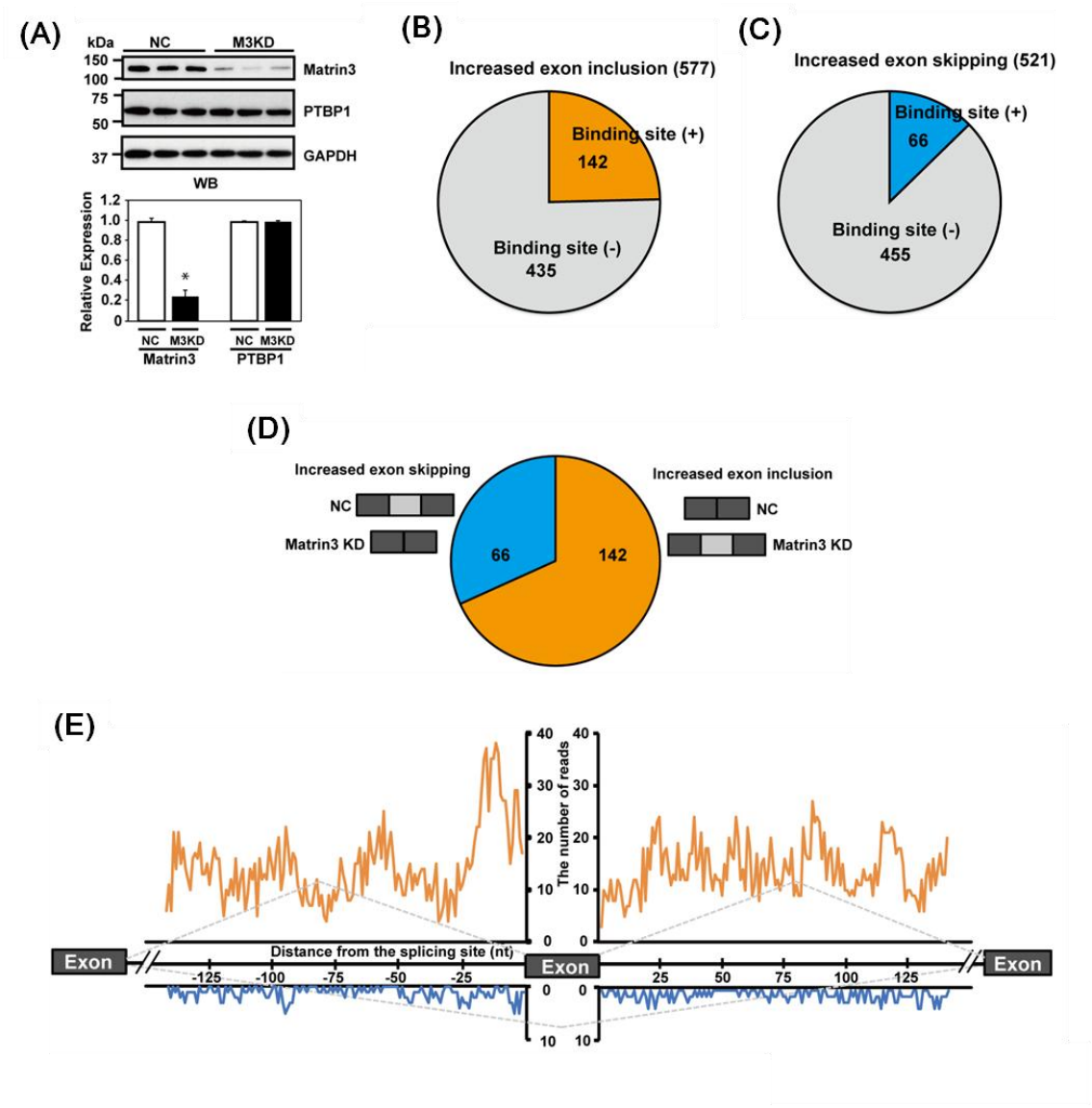

Figure 3 The effect of Matrin3 knockdown on alternative splicing. (A) The change in Matrin3 and PTBP1expression after knockdown of Matrin3 (M3KD) in SH-SY5Y cells were detected by western blot (WB) analysis. GAPDH was detected as a reference. NC, negative control. The relative expression levels of Matrin3 and PTBP1 normalized to GAPDH were quantified and are displayed at the bottom [mean ± SEM (n=3); *p<0.05]. The mean value for the relative expression of Matrin3 and PTBP1 in NC samples was

set as one. (B, C) Analysis of the changes in splicing pattern after Matrin3 knockdown. The number of exons whose inclusion (B) or skipping (C) was increased significantly in response to Matrin3 knockdown is indicated, in relation to whether the adjacent introns contained Matrin3-binding sites as identified by the PAR-CLIP analysis. (D) The number of cassette exons which are adjacent to introns that contain Matrin3-binding sites and whose inclusion or skipping is increased significantly in response to Matrin3 knockdown are shown. (E) The plots show the position and number of sequence reads derived from Matrin3-bound RNAs that are located in exons which are adjacent to introns that contain Matrin3-binding sites and whose inclusion (upper) or skipping (lower), as indicated with dashed lines, is increased significantly in response to Matrin3 knockdown. The center of each read is set as the position.

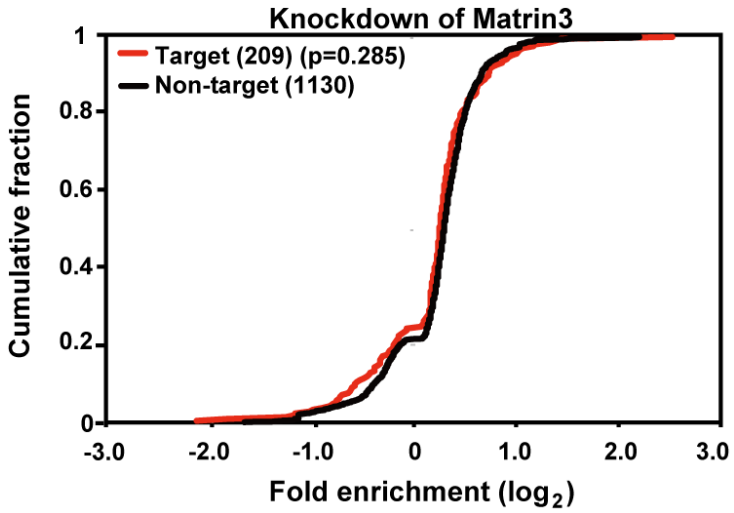

Figure 4 Cumulative distribution plots of changes in gene expression after Matrin3 knockdown. The genes whose expression level was altered significantly in response to depletion of Matrin3 are divided into two categories, Target and Non-target, depending on the presence of the binding sites that had been identified by PAR-CLIP analysis. The change in expression of each mRNA is defined as the mean value of three independent experiments. The number of genes in each category and the p-values calculated between two categories (Welch's *t*-test) are indicated in parentheses.

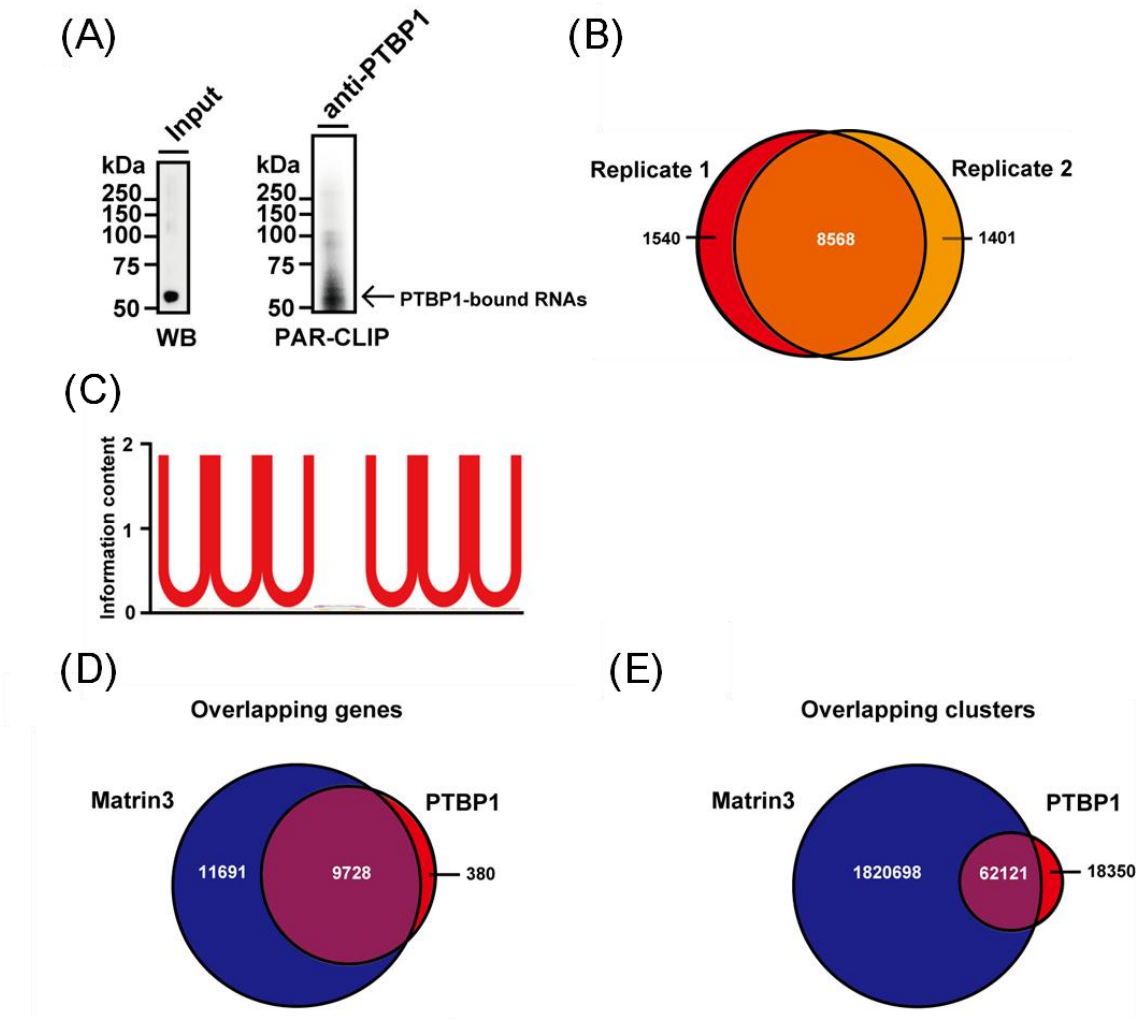

Figure 5 Comparison of target RNAs between Matrin3 and PTBP1. (A) Western blot (WB) analysis (left panel) and autoradiographic analysis (right panel) for immunoprecipitated PTBP1, which was crosslinked with radiolabeled RNA fragments. (B) Confirmation of the reproducibility of the PAR-CLIP analysis for PTBP1. A Venn diagram shows the numbers of overlapping and nonoverlapping genes between Replicate 1 and Replicate 2 of the PAR-CLIP analysis for PTBP1. (C) A representative

PTBP1-binding motif. (D, E) A Venn diagram that shows the number of overlapping and nonoverlapping genes (D) and clusters (E) targeted by Matrin3 and PTBP1.

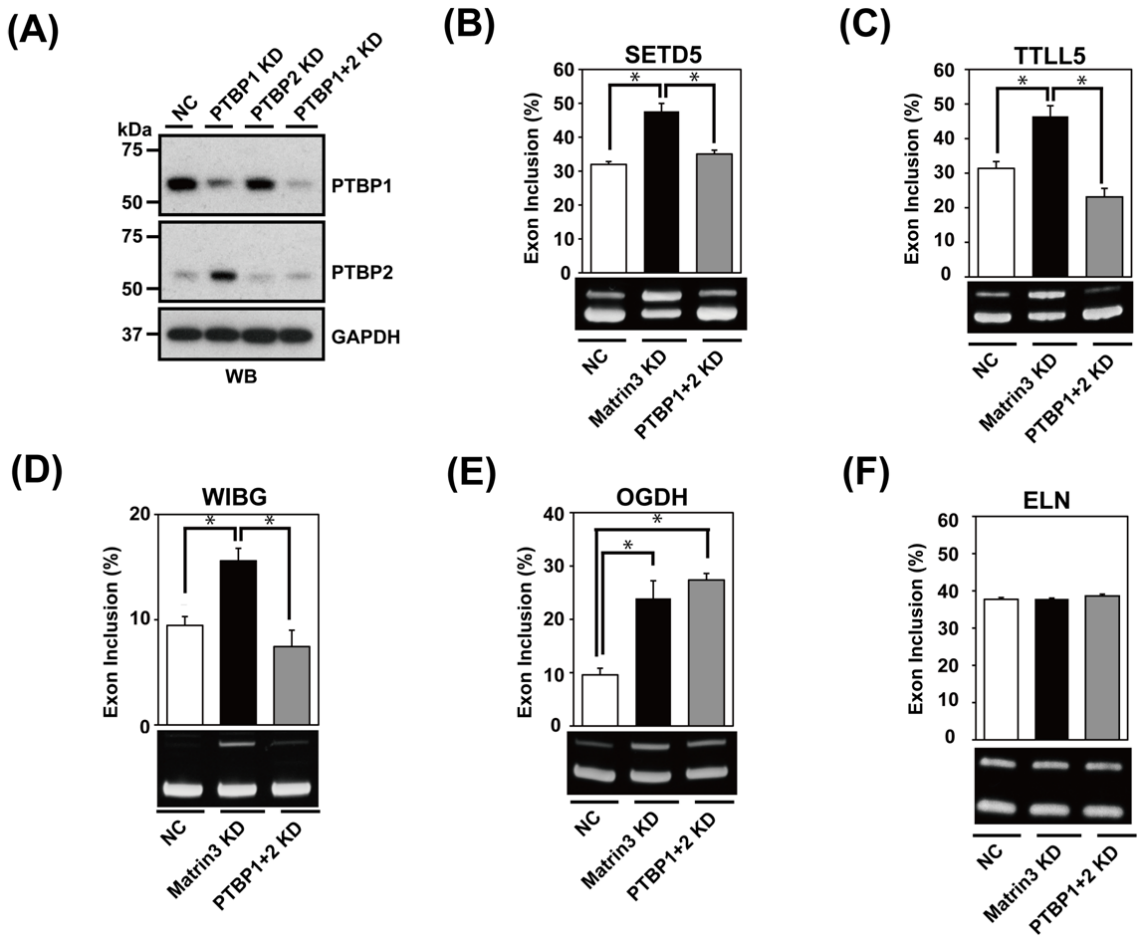

Figure 6 Validation of the changes in splicing pattern that occurred after knockdown of either Matrin3 or PTBPs. (A) The changes in the expression of PTBP1 and PTBP2 after transfection of the indicated siRNAs were detected by western blot (WB) analysis. GAPDH was detected as a reference. NC, negative control. (B–F) The changes in splicing patterns that occurred after transfection of the indicated siRNAs were quantified by RT-PCR that targeted the cassette exons of the *SETD5* (B), *TTLL5* (C), *WIBG* (D), *OGDH* (E), and *ELN* (F) genes and are shown as the percentage of exon inclusion [mean ± SEM (n=3); *p<0.05].

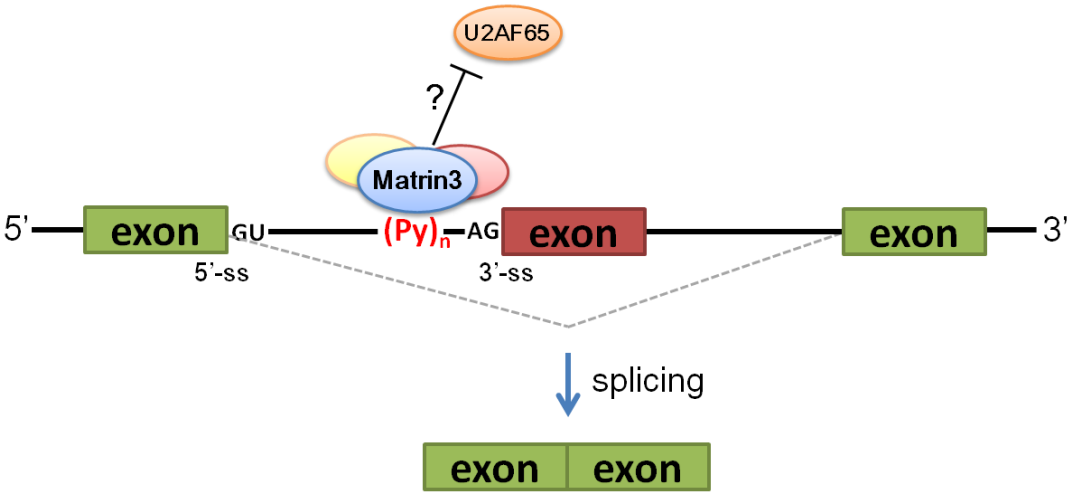

Figure 7 A proposed model for splicing regulation by Matrin3. Matrin3 directly binds to pyrimidine-rich sequences including polypyrimidine tract ($(Py)_n$) and predominantly promotes skipping of a cassette exon (shown in brown), which may be achieved by competition with U2AF65 that also targets polypyrimidine tract and promotes exon inclusion.

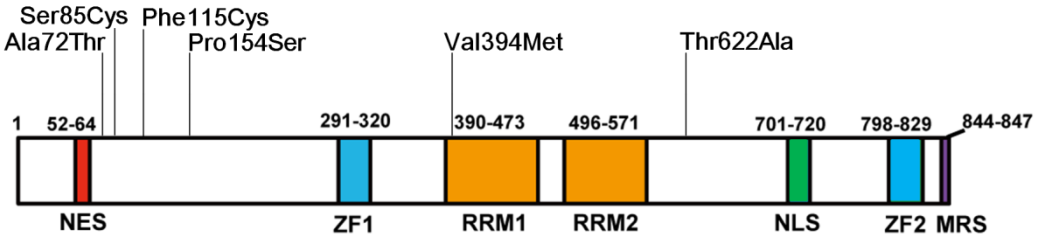

Figure 8 Amino acid substitutions due to ALS-associated point mutations in *Matrin3* gene. Six amino acid substitutions (Ala72Thr, Ser85Cys, Phe115Cys, Pro154Ser, Val394Met, Thr622Ala) due to reported point mutations are mapped in the schematic diagram of Matrin3 (Johnson *et al.* 2014; Lin *et al.* 2015; Leblond *et al.* 2016).

VII. Tables

Table 1 Documented functions of Matrin3 and its interacting proteins.

Documented Function		**Interacting Proteins**	**Reference**
DNA metabolism	Chromatin organization	CHD3, BAZ1A, SMARCA4	Zeitz et al, 2009
	DNA Damage response	PSF, $p54^{nrb}$	Salton et al, 2010
	Transcriptional regulation	Pit1, hnRNPU	Skowronska-Krawczyk et al, 2014
RNA metabolism	mRNA stabilization	DHX9, hnRNPK	Salton et al, 2011
	Gene silencing	AGO1 & AGO2 complex	Höck et al, 2007
	Retention of hyper-edited mRNA	PSF, $p54^{nrb}$	Zhang & Carmichael, 2001
	Viral gene expression	HIV-1 Rev	Yedavalli et al, 2011 Kula et al, 2011
Others	Mediation of neuronal cell death in response to NMDA-type glutamate receptor activation	Calmodulin, Caspases-3/8, DAXX, HIPK1,BRE	Valencia et al, 2007 Przygodzka et al, 2011 Giordano et al. 2005
	Organization of lamina and nuclear matrix	LaminA/C	Depreux et al, 2015

Table 2 General mapping information for PAR-CLIP data for Matrin3 and PTBP1.

Target	**Matrin3**		**PTBP1**	
	Replicate 1	Replicate 2	Replicate 1	Replicate 2
Total reads	193,798,031	278,640,634	202,953,476	247,390,487
After removal of low quality reads	132,272,315	133,931,562	104,731,104	101,939,930
After removal of short reads (<20nt)	131,889,195	133,329,613	104,086,909	100,835,941
Number of reads mapped by Bowtie[1)]	56,619,237	60,066,857	10,883,848	13,412,692
Number of clusters identified by PARalyzer[2)]	1,882,819	1,767,968	80,471	77,832

[1)]Allowance of two mismatches and removal of reads with multiple genomic hits (>10);

[2)]A cluster is defined as a site represented by more than five reads with T-to-C conversions at two or more locations.

Table 3 Gene ontology (GO) analysis for the target mRNAs of Matrin3.

GO Term	Targets (%)	Genome[1)] (%)	Fold enrichment	P-value	Description
GO:0008152	48.8	39.5	1.2	1.43E-99	metabolic process
GO:0044238	40.5	32.9	1.2	9.95E-72	primary metabolic process
GO:0006139	20.4	16.2	1.3	9.47E-35	nucleobase-containing compound metabolic process
GO:0019538	16.1	12.9	1.3	3.98E-25	protein metabolic process
GO:0051179	14.7	12.1	1.2	7.61E-17	localization
GO:0006810	14.3	11.8	1.2	7.58E-16	transport
GO:0015031	8.0	6.1	1.3	9.08E-16	protein transport
GO:0016070	13.8	11.3	1.2	2.23E-15	RNA metabolic process
GO:0006886	7.9	6.1	1.3	2.48E-15	intracellular protein transport
GO:0006464	7.2	5.5	1.3	4.36E-14	cellular protein modification process
GO:0007049	8.2	6.4	1.3	2.07E-13	cell cycle
GO:0042995	30.1	27.3	1.1	9.38E-11	cellular process
GO:0016192	5.5	4.3	1.3	5.54E-09	vesicle-mediated transport
GO:0007067	3.4	2.4	1.4	5.92E-09	mitosis
GO:0065007	17.0	14.9	1.1	6.38E-09	biological regulation
GO:0071840	6.5	5.2	1.3	7.20E-09	cellular component organization or biogenesis
GO:0006397	2.9	2.0	1.4	9.38E-09	mRNA processing
GO:0006366	10.7	9.1	1.2	3.96E-08	transcription from RNA polymerase II promoter
GO:0006351	10.7	9.1	1.2	4.84E-08	transcription, DNA-dependent
GO:0006996	2.1	1.5	1.5	8.22E-08	organelle organization
GO:0000398	2.4	1.7	1.4	1.25E-07	mRNA splicing, via spliceosome
GO:0016043	6.1	4.9	1.2	1.67E-07	cellular component organization
GO:0006412	2.6	1.9	1.4	2.49E-07	translation
GO:0006325	1.8	1.2	1.5	7.72E-07	chromatin organization
GO:0006468	3.3	2.5	1.3	4.71E-06	protein phosphorylation
GO:0006259	2.7	2.0	1.3	6.59E-06	DNA metabolic process

[1)]background fraction (all ENSEMBL) for a specific GO term

Table 4 Top 100 genes targeted by Matrin3.

Ensembl annotation	#clusters	Gene name	Gene full name	Chr.	Gene type
ENSG00000185565	12063	LSAMP	limbic system-associated membrane protein	chr3	protein_coding
ENSG00000158321	6252	AUTS2	autism susceptibility candidate 2	chr7	protein_coding
ENSG00000131558	4604	EXOC4	exocyst complex component 4	chr7	protein_coding
ENSG00000148219	4555	ASTN2	astrotactin 2	chr9	protein_coding
ENSG00000204442	4322	FAM155A	family with sequence similarity 155, member A	chr13	protein_coding
ENSG00000102466	4190	FGF14	fibroblast growth factor 14	chr13	protein_coding
ENSG00000144642	3854	RBMS3	RNA binding motif, single stranded interacting protein 3	chr3	protein_coding
ENSG00000157168	3835	NRG1	neuregulin 1	chr8	protein_coding
ENSG00000184347	3552	SLIT3	slit homolog 3 (Drosophila)	chr5	protein_coding
ENSG00000183098	3515	GPC6	glypican 6	chr13	protein_coding
ENSG00000166833	3351	NAV2	neuron navigator 2	chr11	protein_coding
ENSG00000184903	3235	IMMP2L	IMP2 inner mitochondrial membrane peptidase-like (S. cerevisiae)	chr7	protein_coding
ENSG00000113448	3217	PDE4D	phosphodiesterase 4D, cAMP-specific	chr5	protein_coding
ENSG00000133019	2983	CHRM3	cholinergic receptor, muscarinic 3	chr1	protein_coding
ENSG00000153944	2953	MSI2	musashi RNA-binding protein 2	chr17	protein_coding
ENSG00000135298	2852	BAI3	brain-specific angiogenesis inhibitor 3	chr6	protein_coding
ENSG00000149256	2779	TENM4	teneurin transmembrane protein 4	chr11	protein_coding
ENSG00000100154	2637	TTC28	tetratricopeptide repeat domain 28	chr22	protein_coding
ENSG00000108018	2607	SORCS1	sortilin-related VPS10 domain containing receptor 1	chr10	protein_coding
ENSG00000171094	2583	ALK	anaplastic lymphoma receptor tyrosine kinase	chr2	protein_coding
ENSG00000196628	2580	TCF4	transcription factor 4	chr18	protein_coding
ENSG00000196277	2561	GRM7	glutamate receptor, metabotropic 7	chr3	protein_coding
ENSG00000185420	2558	SMYD3	SET and MYND domain containing 3	chr1	protein_coding
ENSG00000184588	2475	PDE4B	phosphodiesterase 4B, cAMP-specific	chr1	protein_coding
ENSG00000117114	2417	LPHN2	latrophilin 2	chr1	protein_coding
ENSG00000239268	2368	RP11-384F7.2	-	chr3	lincRNA
ENSG00000197157	2362	SND1	staphylococcal nuclease and tudor domain containing 1	chr7	protein_coding
ENSG00000185008	2314	ROBO2	roundabout, axon guidance receptor, homolog 2 (Drosophila)	chr3	protein_coding
ENSG00000196782	2271	MAML3	mastermind-like 3 (Drosophila)	chr4	protein_coding
ENSG00000112541	2249	PDE10A	phosphodiesterase 10A	chr6	protein_coding
ENSG00000182168	2223	UNC5C	unc-5 homolog C (C. elegans)	chr4	protein_coding
ENSG00000198597	2133	ZNF536	zinc finger protein 536	chr19	protein_coding

ENSG00000186153	2096	WWOX	WW domain containing oxidoreductase	chr16	protein_coding
ENSG00000066032	2058	CTNNA2	catenin (cadherin-associated protein), alpha 2	chr2	protein_coding
ENSG00000184305	2057	CCSER1	coiled-coil serine-rich protein 1	chr4	protein_coding
ENSG00000146555	2003	SDK1	sidekick cell adhesion molecule 1	chr7	protein_coding
ENSG00000151276	1992	MAGI1	membrane associated guanylate kinase, WW and PDZ domain containing 1	chr3	protein_coding
ENSG00000107518	1958	ATRNL1	attractin-like 1	chr10	protein_coding
ENSG00000183580	1952	FBXL7	F-box and leucine-rich repeat protein 7	chr5	protein_coding
ENSG00000181722	1948	ZBTB20	zinc finger and BTB domain containing 20	chr3	protein_coding
ENSG00000144868	1940	TMEM108	transmembrane protein 108	chr3	protein_coding
ENSG00000187391	1937	MAGI2	membrane associated guanylate kinase, WW and PDZ domain containing 2	chr7	protein_coding
ENSG00000049618	1916	ARID1B	AT rich interactive domain 1B (SWI1-like)	chr6	protein_coding
ENSG00000184005	1911	ST6GALNAC3	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 3	chr1	protein_coding
ENSG00000189283	1906	FHIT	fragile histidine triad	chr3	protein_coding
ENSG00000153956	1856	CACNA2D1	calcium channel, voltage-dependent, alpha 2/delta subunit 1	chr7	protein_coding
ENSG00000170921	1854	TANC2	tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 2	chr17	protein_coding
ENSG00000140443	1840	IGF1R	insulin-like growth factor 1 receptor	chr15	protein_coding
ENSG00000182985	1822	CADM1	cell adhesion molecule 1	chr11	protein_coding
ENSG00000185532	1820	PRKG1	protein kinase, cGMP-dependent, type I	chr10	protein_coding
ENSG00000143473	1820	KCNH1	potassium voltage-gated channel, subfamily H (eag-related), member 1	chr1	protein_coding
ENSG00000152061	1813	RABGAP1L	RAB GTPase activating protein 1-like	chr1	protein_coding
ENSG00000272168	1805	CASC15	cancer susceptibility candidate 15 (non-protein coding)	chr6	lincRNA
ENSG00000067715	1785	SYT1	synaptotagmin I	chr12	protein_coding
ENSG00000132549	1779	VPS13B	vacuolar protein sorting 13 homolog B (yeast)	chr8	protein_coding
ENSG00000230590	1709	FTX	FTX transcript, XIST regulator (non-protein coding)	chrX	lincRNA
ENSG00000148498	1687	PARD3	par-3 family cell polarity regulator	chr10	protein_coding
ENSG00000138696	1663	BMPR1B	bone morphogenetic protein receptor, type IB	chr4	protein_coding
ENSG00000242593	1641	RP5-921G16.1	-	chr7	antisense
ENSG00000150672	1634	DLG2	discs, large homolog 2 (Drosophila)	chr11	protein_coding
ENSG00000197555	1628	SIPA1L1	signal-induced proliferation-associated 1 like 1	chr14	protein_coding
ENSG00000100320	1610	RBFOX2	RNA binding protein, fox-1 homolog (C. elegans) 2	chr22	protein_coding
ENSG00000160145	1595	KALRN	kalirin, RhoGEF kinase	chr3	protein_coding
ENSG00000172915	1589	NBEA	neurobeachin	chr13	protein_coding
ENSG00000134138	1580	MEIS2	Meis homeobox 2	chr15	protein_coding
ENSG00000144278	1573	GALNT13	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 13 (GalNAc-T13)	chr2	protein_coding
ENSG00000153317	1567	ASAP1	ArfGAP with SH3 domain, ankyrin repeat and PH domain 1	chr8	protein_coding
ENSG00000117020	1558	AKT3	v-akt murine thymoma viral oncogene homolog 3	chr1	protein_coding

ENSG00000152127	1544	MGAT5	mannosyl (alpha-1,6-)-glycoprotein beta-1,6-N-acetyl-glucosaminyltransferase	chr2	protein_coding
ENSG00000157985	1541	AGAP1	ArfGAP with GTPase domain, ankyrin repeat and PH domain 1	chr2	protein_coding
ENSG00000048052	1536	HDAC9	histone deacetylase 9	chr7	protein_coding
ENSG00000174891	1522	RSRC1	arginine/serine-rich coiled-coil 1	chr3	protein_coding
ENSG00000091129	1520	NRCAM	neuronal cell adhesion molecule	chr7	protein_coding
ENSG00000144036	1518	EXOC6B	exocyst complex component 6B	chr2	protein_coding
ENSG00000204406	1518	MBD5	methyl-CpG binding domain protein 5	chr2	protein_coding
ENSG00000156052	1508	GNAQ	guanine nucleotide binding protein (G protein), q polypeptide	chr9	protein_coding
ENSG00000145996	1506	CDKAL1	CDK5 regulatory subunit associated protein 1-like 1	chr6	protein_coding
ENSG00000249859	1501	PVT1	Pvt1 oncogene (non-protein coding)	chr8	processed_transcript
ENSG00000171723	1494	GPHN	gephyrin	chr14	protein_coding
ENSG00000142599	1476	RERE	arginine-glutamic acid dipeptide (RE) repeats	chr1	protein_coding
ENSG00000134775	1475	FHOD3	formin homology 2 domain containing 3	chr18	protein_coding
ENSG00000106554	1475	CHCHD3	coiled-coil-helix-coiled-coil-helix domain containing 3	chr7	protein_coding
ENSG00000058091	1474	CDK14	cyclin-dependent kinase 14	chr7	protein_coding
ENSG00000176406	1468	RIMS2	regulating synaptic membrane exocytosis 2	chr8	protein_coding
ENSG00000182185	1456	RAD51B	RAD51 paralog B	chr14	protein_coding
ENSG00000184349	1453	EFNA5	ephrin-A5	chr5	protein_coding
ENSG00000156639	1452	ZFAND3	zinc finger, AN1-type domain 3	chr6	protein_coding
ENSG00000137968	1444	SLC44A5	solute carrier family 44, member 5	chr1	protein_coding
ENSG00000145743	1444	FBXL17	F-box and leucine-rich repeat protein 17	chr5	protein_coding
ENSG00000164418	1437	GRIK2	glutamate receptor, ionotropic, kainate 2	chr6	protein_coding
ENSG00000118922	1425	KLF12	Kruppel-like factor 12	chr13	protein_coding
ENSG00000064042	1425	LIMCH1	LIM and calponin homology domains 1	chr4	protein_coding
ENSG00000119522	1417	DENND1A	DENN/MADD domain containing 1A	chr9	protein_coding
ENSG00000113391	1403	FAM172A	family with sequence similarity 172, member A	chr5	protein_coding
ENSG00000162849	1396	KIF26B	kinesin family member 26B	chr1	protein_coding
ENSG00000196482	1392	ESRRG	estrogen-related receptor gamma	chr1	protein_coding
ENSG00000147862	1387	NFIB	nuclear factor I/B	chr9	protein_coding
ENSG00000182621	1362	PLCB1	phospholipase C, beta 1 (phosphoinositide-specific)	chr20	protein_coding
ENSG00000165659	1362	DACH1	dachshund homolog 1 (Drosophila)	chr13	protein_coding
ENSG00000185483	1354	ROR1	receptor tyrosine kinase-like orphan receptor 1	chr1	protein_coding

Table 5 Matrin3-binding sites in ALS-related genes.

Gene name	Protein name	#clusters	Locus
MATR3	Matrin3	178	5q31.2
DCTN1	Dynactin subunit 1	74	2p13
C9orf72	C9orf72	58	9p21-22
VCP	Valosin Containing Protein	43	9p13.3
FUS	FUS	41	16p11.2
TBK1	TANK binding kinase 1	37	12q14.1
SQSTM1	Sequestosome 1	25	5q35
UBQLN2	Ubiquilin 2	22	Xp11.23-Xp13.1
PFN1	Profilin 1	20	17p13.2
HNRNPA1	hnRNP A1	18	12q13.1
TARDBP	TDP-43	14	1p36.22
SOD1	Superoxide dismutase 1	7	21q22.1
OPTN	Optineurin	7	10p15-p14
CHCHD10	Coiled-coil-helix-coiled-coil-helix domain-containing protein 10	1	22q11.23
ANG	Angiogenin	0	14q11
TUBA4A	Tubulin Alpha 4A	0	2q36.1

Table 6 Top five 7-mer motifs of Matrin3 inferred by the cERMIT tool.

Cluster	Motif	Score(#Targets, %Targets, %cumulative)
1	UUUCUNU	67.82(9073, 13, 13)
	UNUUUCU	66.30(8646, 12, 19)
	UUCUNUU	64.97(8366, 12, 24)
	UCUNUUU	61.68(6882, 10, 27)
	WUUUUCU	57.08(5530, 8, 28)
	UYUUUCU	56.23(6002, 8, 28)
	UWUUUCU	56.13(5012, 7, 28)
	YUUUUCU	56.10(5809, 8, 29)
	KUUUUCU	52.73(5097, 7, 30)
	CWUUUCU	49.54(3868, 5, 31)
	YCUUUCU	47.66(4049, 5, 31)
	UUUUUCK	47.34(3746, 5, 32)
	KCUUUCU	45.99(3706, 5, 32)
	WCUUUCU	44.99(3731, 5, 33)
	cumulative % = 33	
2	UNUCUUU	67.37(8659, 12, 12)
	UUNUCUU	66.54(8740, 12, 18)
	UCUUUNU	63.62(7949, 11, 23)
	UCUUUYU	53.23(5463, 8, 23)
	UCUUUUY	51.14(5114, 7, 24)
	UCUUUUK	49.57(4722, 6, 25)
	UCUUUUW	49.51(4197, 6, 26)
	UCUUUCW	48.01(3970, 5, 26)
	UCUUUCY	46.13(4046, 5, 27)
	cumulative % = 39	
3	UUUNCUU	65.31(8412, 12, 12)
	cumulative % = 42	
4	UUUNUCU	65.15(7995, 11, 11)
	cumulative % = 43	
5	UUUCNUU	65.09(7709, 11, 11)
	UUUUCNU	62.32(7689, 11, 16)
	UUCNUUU	60.62(7409, 10, 20)
	cumulative % = 46	

Table 7 Top five 7-mer motifs of PTBP1 inferred by the cERMIT tool.

Cluster	Motif	Score(#Targets, %Targets, %cumulative)
1	UUUNUUU	20.02(6708, 9, 9)
	cumulative % = 9	
2	UUUUNCU	18.68(6028, 8, 8)
	cumulative % = 15	
3	UUUUCYU	18.26(4490, 6, 6)
	UUUCYUU	15.69(4213, 6, 9)
	UUUUUCY	14.59(4079, 5, 12)
	cumulative % = 18	
4	UUUMUUU	17.76(3951, 5, 5)
	UUUUMUU	16.65(3950, 5, 8)
	UUMUUUU	13.84(3604, 5, 10)
	cumulative % = 20	
5	UUUCUYU	17.72(4321, 6, 6)
	UUUUCUY	14.39(4229, 6, 9)
	UKUUUCU	14.19(3797, 5, 11)
	YUUUUCU	14.05(4269, 6, 12)
	WUUUUCU	13.52(4050, 5, 13)
	KUUUUCU	13.50(3729, 5, 13)
	UWUUUCU	13.38(3621, 5, 14)
	UYUUUCU	12.65(4143, 5, 15)
	cumulative % = 26	

VIII. *Acknowledgements*

I thank Prof. Toru Nakano and Prof. Yukio Kawahara for helpful discussion. I thank all the members in the laboratory of RNA Biology and Neuroscience and staff in the Center for Medical Research and Education, Graduate School of Medicine, Osaka University for technical support. This work was supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan for Scientific Research on Innovative Areas (15H01471) and for Scientific Research (B) (16H05320) to Prof. Yukio Kawahara and by grants from Takeda Medical Research Foundation to Prof. Yukio Kawahara.

IX. *Achievement*

Publications

1) Yuri Uemura, Takuya Oshima, Munetaka Yamamoto, Charles Jourdan Reyes, Pedro Henrique Costa Cruz, Toshiharu Shibuya, Yukio Kawahara, Matrin3 binds directly to intronic pyrimidine-rich sequences and controls alternative splicing. ***Genes to Cells.*** (2017), **9**:785-798.

2) Quan Li, Yuri Uemura, Yukio Kawahara. Cross-Linking and Immunoprecipitation of Nuclear RNA-Binding Proteins. ***Methods in Molecular Biology.*** (2015), 1262: 247-263. (論文執筆に寄与しました)

3) 植村有里, 河原行郎「RNA 代謝と神経変性」***Brain Medical.*** (2014), 26(3): 209-215.

4) Yuri Uemura, Noriko Nakagawa, Taisuke Wakamatsu, Kwang Kim, Gaetano Thomas Montelione, John Francis Hunt, Seiki Kuramitsu, Ryoji Masui. Crystal structure of the ligand-binding form of nanoRNase from Bacteroides fragilis, a member ofthe DHH/DHHA1 phosphoesterase family of proteins. ***FEBS Letters.*** (2013), 587(16):2669-2674.

5) Taisuke Wakamatsu, Kwang Kim, Yuri Uemura, Noriko Nakagawa, Seiki Kuramitsu, Ryoji Masui. Role of RecJ-like protein with 5'-3' exonuclease activity in

oligo(deoxy)nucleotide degradation. ***Journal of Biological Chemistry*****.** (2011), 286(4):2807-16. (細菌の増殖速度を計測しました)

Presentations

1) 植村有里, 山本宗隆, Charles Jourdan Reyes, 河原行郎. Matrin-3 によるスプライシング制御機構の解析. 第 17 回日本 RNA 学会年会, 札幌, 2015 年 7 月 15 日

2) Yuri Uemura, Noriko Nakagawa, Taisuke Wakamatsu, Gaetano Thomas Montelione, John Francis Hunt, Seiki Kuramitsu, and Ryoji Masui. Crystal structure and reaction mechanism of short chain DNA/RNA specific exonuclease. 第 86 回日本蛋白質科学会年会, 鳥取, 2013 年 9 月 13 日

3) Yuri Uemura, Taisuke Wakamatsu, Noriko Nakagawa, John Francis Hunt, Gaetano Thomas Montelione, Seiki Kuramitsu, and Ryoji Masui. Structure and reaction mechanism of nanoRNA-specific exonuclease. 第 13 回日本蛋白質科学会年会, 鳥取, 2013 年 6 月 14 日

4) Yuri Uemura, Noriko Nakagawa, Taisuke Wakamatsu, Gaetano Thomas Montelione, John Francis Hunt, Ryoji Masui, and Seiki Kuramitsu. Reaction mechanism of short chain DNA/RNA specific exonuclease. 第 85 回日本生化学会大会, 福岡, 2012 年 12 月 13 日

5) Yuri Uemura, Noriko Nakagawa, Taisuke Wakamatsu, Atsuhiro Shimada, Ryoji Masui, and Seiki Kuramitsu. Analysis of the unique catalytic mechanism of exonuclease TTHA0118 from *Thermus thermophilus* HB8. The 2nd Annual Meeting for Whole-Organism Science Society Joint Meeting with The 11th Annual Meeting of Structural-Biological Whole Cell Project of *Thermus thermophilus* HB8. 理化学研究所 SPring-8, 2012 年 9 月 28 日

6) Yuri Uemura, Taisuke Wakamatsu, Atsuhiro Shimada, Ryoji Masui, and Seiki Kuramitsu. Molecular and cellular analysis of exonuclease TTHA0118 from *Thermus thermophilus* HB8. The 1st Annual Meeting for Whole-Organism Science Society Joint Meeting with The 10th Annual Meeting of Structural-Biological Whole Cell Project of *Thermus thermophilus* HB8. 理化学研究所 SPring-8, 2011 年 8 月 19 日