Age-related demethylation of the TDP-43 autoregulatory region in the human motor cortex

Yuka Koike[1], Akihiro Sugai[2], Norikazu Hara[3], Junko Ito[4], Akio Yokoseki[5], Tomohiko Ishihara[1], Takuma Yamagishi[1], Shintaro Tsuboguchi[1], Mari Tada[6], Takeshi Ikeuchi[3], Akiyoshi Kakita[4], Osamu Onodera[1]

[1] Department of Neurology, Brain Research Institute, Niigata University

[2] Department of Molecular Neuroscience, Center for Bioresource-based Researches, Brain Research Institute, Niigata University

[3] Department of Molecular Genetics, Center for Bioresource-based Researches, Brain Research Institute, Niigata University

[4] Department of Pathology, Brain Research Institute, Niigata University

[5] Department of Inter-Organ Communication Research, Niigata University Graduate School of Medical and Dental Sciences

[6] Department of Pathology Neuroscience, Center for Bioresource-based Researches, Brain Research Institute, Niigata University

Corresponding author: Akihiro Sugai and Osamu Onodera

E-mail: akihiro.sugai@bri.niigata-u.ac.jp; onodera@bri.niigata-u.ac.jp

1-757 Asahimachi-dori, Chuo-Ku, Niigata-City, Niigata 951-8585, Japan

Tel and Fax: +81 25 227 0663

A running title: Age-related demethylation prevents TDP-43 autoregulation

Keywords: amyotrophic lateral sclerosis, *TARDBP*, DNA methylation, alternative splicing, epigenetic age

Abstract

In amyotrophic lateral sclerosis (ALS), TAR DNA-binding protein 43 (TDP-43) forms aggregates in the motor cortex of the aging brain. This aggregate formation may be triggered by the increase in TDP-43 levels with aging. However, the amount of TDP-43 is autoregulated by the alternative splicing of the *TARDBP* 3'UTR, and its relationship with aging remains unresolved. Since DNA methylation is altered during aging, we hypothesized that 3'UTR methylation is also altered in the aging motor cortex, disrupting this autoregulatory system and increasing TDP-43 levels. We found that DNA demethylation in the autoregulatory region of TDP-43 reduced alternative splicing and increased TDP-43 expression. Furthermore, in the human motor cortex, we found that this region was demethylated with age and that the expression of TDP-43 increased. The dysregulation of TDP-43 autoregulation by age-related DNA demethylation in the motor cortex may explain the contribution of aging and system selectivity in ALS.

Introduction

In neurodegenerative diseases, in which aging is the primary risk factor[1], specific proteins aggregate selectively in the central nervous system. It has been speculated that an increase in the amount of causative proteins contributes to this aggregate formation[2,3]. Thus, one pathological hypothesis of system selectivity in neurodegenerative diseases is that the amount of disease-causing protein is increased in the impaired system. In fact, alpha-synuclein levels have been shown to be increased in neuronal systems affected by alpha-synucleinopathy[4]. However, the association between the amount of the causative protein and aging is obscure.

Sporadic amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease that selectively impairs primary motor neurons in the motor cortex and alpha-motor neurons in the spinal cord, in which aging is a major risk factor, similar to other sporadic neurodegenerative diseases[5]. In more than 95% of sporadic ALS cases, the nuclear protein TDP-43 accumulates in lesions[6]. In patients with sporadic ALS, TDP-43 forms aggregates in the cytoplasm and is depleted from the nucleus in diseased cells[7,8]. Increased intracellular concentrations of TDP-43 have been speculated to be a factor in this aggregate formation[9,10]. However, the mechanism linking TDP-43, motor neuron selectivity, and aging has not yet been elucidated.

The level of TDP-43 in cells is strictly autoregulated via alternative splicing of the *TARDBP* 3'UTR[11–13]. Excess nuclear TDP-43 binds to the *TARDBP* pre-mRNA 3'UTR and induces its splicing to produce nonsense-mediated mRNA decay-sensitive *TARDBP* mRNA and reduce the level of TDP-43[12,14]. On the other hand, when the level of TDP-43 in the nucleus is reduced, this splicing is repressed, and *TARDBP* mRNA levels are increased. Therefore, *TARDBP* mRNA expression is continually increased in

ALS-affected cells with reduced nuclear TDP-43 levels[14]. Upon the formation of cytoplasmic TDP-43 aggregates, which is presumed to reduce the TDP-43 level in the nucleus, either by entrapping newly generated TDP-43 or by inhibiting the nuclear transport of TDP-43, the progression of the disease accelerates[15,16]. However, it is unclear what triggers the disruption of this autoregulatory mechanism leading to increased TDP-43 expression and TDP-43 aggregate formation in the motor cortex with age.

Aging alters the methylation status of DNA. The DNA methylation status defines epigenetic age separate from chronological age. In age-related diseases, this epigenetic age contributes more than chronological age to mortality and morbidity[17]. Indeed, in sporadic ALS, an accelerated epigenetic age based on DNA methylation results in a younger age of onset[18]. DNA methylation not only regulates gene expression but also affects alternative splicing[19–23]. Based on these facts, we hypothesized that the DNA methylation state of the *TARDBP*-3'UTR changes with age in the motor cortex, which in turn alters the splicing of the *TARDBP*-3'UTR in the direction of increased TDP-43 expression.

To explore this hypothesis, we used the dCas9 system to selectively demethylate the CpG region of the alternative splicing-related site of the *TARDBP* 3'UTR. We found that the demethylation of this region inhibited alternative splicing related to TDP-43 autoregulation and increased the expression of *TARDBP* mRNA. Furthermore, we found in the human motor cortex that this region was demethylated with aging and that the degree of demethylation was correlated with the level of *TARDBP* mRNA.

Results

CpG sites are clustered around the alternative splicing site in the *TARDBP* 3'UTR

TDP-43 binds to the 3'UTR of *TARDBP* pre-mRNA and induces the alternative splicing of intron 6 (exonic intron) and intron 7 (3'UTR intron), leading to nonsense-mediated mRNA decay[12,14]. We found that 15 CpG sites were clustered around the 5' splicing site of alternative intron 7. An analysis of publicly available data revealed that these CpG sites are moderately to highly methylated in the human prefrontal cortex (Fig. 1a)[24,25]. At methylated DNA sites, RNA polymerase II may stall, eliciting splicing[20,21,26]. Indeed, we found from publicly available data that the CpG cluster region of the *TARDBP* 3'UTR is enriched in RNA polymerase II[27] (Fig. 1b), which may be associated with the alternative splicing of intron 7[28]. We found that the region is also enriched in H3K27ac marks (Fig. 1b)[27], a characteristic of regions where DNA methylation changes with age[29,30]. We thus hypothesized that the mRNA level of *TARDBP* increases with age via DNA demethylation in this region.

Demethylation of the *TARDBP* 3'UTR suppresses alternative splicing and increases *TARDBP* expression

To test our hypothesis, we first examined whether the demethylation of this region suppressed splicing. To manipulate the DNA methylation status at these 15 CpG sites, we applied the dCas9 system with four guide RNAs (Fig. 2a). The TET1-*TARDBP*-target vector[31] or the DNMT3A-*TARDBP*-target vector[32] was applied for demethylation or methylation, respectively, at the 15 CpG sites (Fig. 2b).

CpGs 1–15 in the 3'UTR were highly methylated in HEK293T cells (Fig. 2c; control). The transfection of the TET1-*TARDBP*-target vector with each guide RNA demethylated these CpG sites (Figure 2c). The 3'UTR CpG 10–15 sites were more efficiently demethylated than 3'UTR CpG 1–9, which was negatively correlated with CpG density (Figure 2c-e, Supplementary Figure 1a, b). However, the transfection of the DNMT3A-*TARDBP*-target vector did not alter the DNA methylation percentage (Fig. 2c).

Then, we asked whether the heterogeneity of each epiallele was altered upon region-specific TET1-dependent demethylation, as age-related DNA methylation changes have been reported to increase the heterogeneity of DNA methylation[33] and decrease the correlation between the methylation status of each CpG pair[34]. The methylation pattern of cells transfected with the TET1-*TARDBP*-target vector showed a higher epipolymorphism score than that of the control at CpGs 10–15 (Fig. 2f-h, Supplementary Fig. 1c, d)[35,36]. However, the methylation linkage diagram revealed that the correlation of the methylation status of each CpG pair was increased by TET1 (Fig. 2i). These results indicate the existence of an interaction in the demethylation of each of the 3'UTR CpG 10–15 sites, which is not consistent with the age-related demethylation patterns that have been reported thus far[34].

We then examined whether demethylation in the region suppresses the alternative splicing of *TARDBP* mRNA (Fig. 3a). In cells transfected with the TET1-*TARDBP*-target vector with guide RNA1, the alternative splicing of intron 6 and intron 7 was suppressed (Fig. 3b, c), with a concomitant 1.85-fold increase in the level of unspliced canonical mRNA (Fig. 3d). These results demonstrate that the demethylation of 3'UTR CpGs 10–15, which was sensitive to TET1, suppresses the alternative splicing of intron 6 and intron 7 and increases *TARDBP* mRNA levels (Fig. 3e).

DNA methylation of *TARDBP* in the human brain

We next explored the DNA methylation status of the *TARDBP* gene in the human brain by examining the motor, occipital, and cerebellar cortexes of seven patients with sporadic ALS and eight control patients without brain disease (Supplementary Table 1). We investigated 22 CpG sites in the *TARDBP* promoter region that were predicted by PROMOTER SCAN[37] and PromoterInspector[38] and the 15 CpG sites in the *TARDBP* 3'UTR (Fig. 4a).

In the examined brain regions, the CpG sites in the promoter region were hypomethylated, while 3'UTR CpGs 1–15 were moderately to highly methylated (Fig. 4b-j). The percentage of DNA methylation was lower in 3'UTR CpGs 10–15 than in 3'UTR CpGs 1–9 (Fig. 4, Supplementary Fig. 2a) and was positively correlated with the CpG density (Supplementary Fig. 2b). The interindividual variability in the percentage of methylation was greater in 3'UTR CpGs 10–15 than in 3'UTR CpGs 1–9 (Fig. 4, Supplementary Fig. 2c), which was also associated with the surrounding CpG density (Supplementary Fig. 2d)[39]. The cerebral cortex exhibited a lower percentage of 3'UTR CpG 10–15 methylation and greater interindividual variability than the cerebellum (Fig. 4, Supplementary Fig. 2c). In these brain regions, the percentage of DNA methylation at each CpG did not differ between the ALS patients and the controls (Fig. 4b-j).

Characteristics of *TARDBP* 3'UTR DNA methylation in the motor cortex

To evaluate the tissue- and age-related methylation patterns in the motor cortex, we further recruited 3 ALS patients and 3 controls, resulting in a total of 10 ALS patients and 11 controls (Supplementary Table 1). As a noncentral nervous system tissue, we analyzed DNA obtained from the livers of ALS patients. Among 3'UTR CpGs 1–9, the motor cortex exhibited more epialleles with demethylated CpGs than the others (Fig. 5a), with high pairwise correlations for each CpG site (Fig. 5c). In 3'UTR CpGs 10–15, the motor and occipital cortex presented more epialleles with demethylated CpGs (Fig. 5b), with high pairwise correlations for each CpG site (Fig. 5c) compared to those in the cerebellum. The epipolymorphism scores of nervous tissues were higher than those for the liver (Supplementary Fig. 3a-f). These features were similar in the ALS and control groups. Each tissue was classified by principal component analysis incorporating these indices (Fig. 5d). The motor cortex was also characterized by high interindividual variability (Fig. 5d).

Age-related demethylation of the *TARDBP* 3'UTR increases *TARDBP* mRNA levels in the motor cortex

Next, we investigated the association between the DNA methylation status of the 3'UTR CpG in the motor cortex and age (Fig. 6a-d). In the motor cortex, we found that the percentage of DNA methylation at 3'UTR CpGs 10–15 was inversely correlated with age (Fig. 6c, d). The percentage of DNA methylation was approximately 80% in individuals in their 50s but decreased to nearly 50% in individuals in their 80s. In the elderly, the proportion of epialleles with five or more methylated sites among 3'UTR CpGs 10–15 was extremely low (Figure 6e). In contrast, there was no correlation with age in the other tissues or at other CpG sites (Fig. 6a, b, Supplementary Fig. 4). Moreover, there was no correlation with age in ALS patients (an interaction was found in covariance analysis, $p = 0.028$) (Fig. 6c). In the ALS group, the degree of DNA methylation was as low as approximately 65% at 45 years of age, which was similar to

that in individuals in their 60s and 70s.

The epipolymorphism score of 3'UTR CpGs 10–15 was positively correlated with age in the motor cortex of the controls (Fig. 6f, Supplementary Fig. 5), consistent with the age-related increase in the heterogeneity of DNA methylation[33]. In addition, the methylation linkage diagram of the two youngest (middle-aged) and two oldest (elderly) individuals showed generally lower pairwise correlations in the elderly (Fig. 6g, h), suggesting stochastic demethylation of the epialleles with aging[34], as opposed to regulated demethylation by specific enzymes (Fig. 6i).

Finally, we investigated the association between the DNA methylation status of the 3'UTR CpGs and *TARDBP* mRNA levels in the motor cortex. *TARDBP* mRNA was negatively correlated with the percentage of DNA methylation at 3'UTR CpGs 10–15 in the control group (Fig. 6j). In contrast, this relationship was not observed in the ALS group (an interaction was observed in covariance analysis, $p = 0.005$).

Discussion

In this study, we first showed that the 3'UTR CpGs of *TARDBP* were methylated and that the 3'UTR CpG 10–15 region was sensitive to TET1. Second, a demethylated state of 3'UTR CpGs 10–15 suppressed the splicing of intron 7 of the *TARDBP* 3'UTR and increased *TARDBP* mRNA levels. Third, the 3'UTR CpG methylation pattern of *TARDBP* differed among brain regions, with demethylation being observed in the motor cortex. Fourth, we showed that demethylation in this region was accelerated by aging only in the motor cortex. Finally, we showed that in the motor cortex, the degree of methylation was inversely correlated with the amount of *TARDBP* mRNA. These findings indicate that the age-dependent demethylation of the *TARDBP* 3-UTR increases *TARDBP* mRNA levels in the human motor cortex. We propose that this may partly explain why aging is a risk factor and why the motor cortex is affected in ALS.

We showed that the demethylation of 3’UTR CpGs 10–15 suppressed the splicing of intron 7. DNA methylation sites are enriched within alternative splice sites and can either enhance or suppress exon inclusion[22,23,26]. DNA methylation causes RNA polymerase II pause by modulating proteins such as CCCTC-binding factor[20,21] and methyl-CpG binding protein 2[23,26]. The halting of RNA polymerase II increases the splicing of intron 7[28]. Thus, we speculate that the demethylation of 3'UTR CpGs 10–15 may promote the elongation of RNA polymerase II and suppress the alternative splicing of intron 7. The suppression of the intron 7 splicing results in an increase in canonical *TARDBP* mRNA levels. Our findings indicate that the DNA methylation status of 3’UTR CpGs 10–15 regulates the amount of TDP-43.

Next, we discuss the relationship between this age-related demethylation in the motor cortex and the motor cortex selectivity of ALS. We found age-dependent demethylation of 3'UTR CpGs 10–15 in the motor cortex but not in the occipital and cerebellar cortexes. This result is consistent with previous studies showing that the age-related state of DNA methylation differs from brain region to brain region[40]. In particular, it is consistent with results showing less of an effect of age-related methylation in the occipital lobe and cerebellum[41]. The differential expression of DNMT and TET may explain the specific DNA methylation profiles associated with aging in various brain regions[42]. Indeed, the expression of TET1, TET3, and DNMT3A2 is also affected by neural activity[43–45]. The unique profile of *TARDBP* 3'UTR DNA methylation in the motor cortex may contribute to susceptibility to TDP-43 pathology during brain aging in individuals with ALS. On the other hand, this region was found to be uniformly demethylated in ALS patients, and this did not change with age. This may reflect the changes in the constituent cells due to degeneration.

We will also discuss the limitations of this study. First, our analysis targeted 15 CpGs located at the splicing site of intron 7. However, intron 6 also contributes to the autoregulation of TDP-43[13,14]. Therefore, we cannot exclude the possibility that the DNA methylation status of CpGs in the neighborhood of the splicing site of intron 6 may be affected. Second, the present analysis included a mixture of multiple cell types. Therefore, it is unclear whether the observed age dependency was due to senescence of each cell itself or to changes in constituent cells due to aging[36]. In the future, methylation analysis at the single-cell level is expected to be performed to examine the methylation changes in each cell type[19,46]. Third, the limited number of brain regions analyzed restricts our ability to discuss tissue specificity.

In summary, our findings suggest that age-related demethylation may affect gene expression levels via alternative splicing. In the central nervous system in particular, diverse mRNA isoforms produced via alternative splicing contribute to higher-order mechanisms and pathological mechanisms[47]. Therefore, the identification of DNA methylation sites that regulate these processes and the elucidation of the underlying molecular mechanisms may contribute to understanding the pathogenesis of not only sporadic ALS but also other sporadic neurodegenerative diseases.

Materials & Methods

Targeted manipulation of DNA methylation

A TET1-*TARDBP*-target vector incorporating dCas9-peptide repeats and the catalytic domain of the DNA demethylase (scFv-TET1) was applied to demethylate the target DNA region (Fig. 2b)[31]. A DNMT3A-*TARDBP*-target vector incorporating dCas9 and DNMT3A, the catalytic domain of DNA

methyltransferase, was applied for the methylation of the target DNA region (Fig. 2b)[32].

To construct the TET1-*TARDBP*-target vector, pPlatTET-gRNA2 (Addgene plasmid 82559; Addgene, Cambridge, MA, USA) was cut at the AflII site, and a guide RNA fragment was inserted using Gibson assembly master mix (New England Biolabs, Ipswich, MA, USA). Four guide RNAs were selected using Benchling (https://benchling.com/). To construct the DNMT3A-*TARDBP*-target vector, pdCas9-DNMT3A-EGFP (Addgene plasmid 71666) was cut at a BbsI site, and a guide RNA fragment was inserted in the same manner. The sequences of the guide RNAs are described in Supplementary Table 2. As a control vector, pHRdSV40-scFv-GCN4-sfGFP-VP64-GB1-NLS (Addgene plasmid 60904) was used. The plasmid vectors (0.265 pmol) were transfected using Lipofectamine 3000 (Invitrogen, Waltham, MA, USA) into HEK293T cells cultured in Dulbecco's modified Eagle's medium (Gibco, Waltham, MA, USA) containing 10% fetal bovine serum in 24-well plates. After 48 hours, GFP-positive cells were isolated using FACS Diva (BD Biosciences, San Jose, CA, USA) software on a FACSAria II system (BD Biosciences).

Analyzed human tissues

Frozen autopsy brains of sporadic ALS patients without ALS-causing gene mutations and control patients without brain disease were pathologically confirmed at the Department of Pathology, Brain Research Institute, Niigata University (Supplementary Table 1). Although a variant (2,076 G> A) reported in a family lineage presenting with ALS and frontotemporal degeneration with elevated expression of *TARDBP*[48] causes the loss of the sixth of the 15 CpG sites in the 3'UTR (Fig. 1a), we found no variants during the amplicon sequencing of the patients included in this study. The Institutional Ethical Review Board of Niigata University approved this study, which investigated postmortem tissues autopsied with written informed consent from families.

Analysis of DNA methylation status

Genomic DNA from 2.5×10^5 HEK293T cells was purified using a Nucleo Spin Tissue XS kit (MACHEREY-NAGEL, Düren, Nordrhein-Westfalen, Germany). From 25 mg of human postmortem tissue, genomic DNA was purified using a DNeasy Blood & Tissue kit (QIAGEN, Venlo, Limburg, Netherlands). According to the manufacturer's protocols, the purified genomic DNA was bisulfite treated with the EpiTect Fast DNA Bisulfite kit (QIAGEN). In both HEK293T cells and human postmortem tissues, bisulfite amplicon sequencing was performed using primer sets A and B (as shown in Fig. 2a), which target the 1st–9th (3'UTR CpGs 1–9) and the 10th–15th (3'UTR CpGs 10–15) CpG sites, respectively. The bisulfite-treated DNA was amplified by nested PCR (Supplementary Table 2) using KAPA HiFi HS Uracil+ ReadyMix (Kapa Biosystems, Woburn, MA, USA). For first and second rounds

of PCR, the thermal cycling conditions were as follows: 95°C for 3 minutes, then 30 cycles of 98°C for 20 seconds, 60°C for 15 seconds and 72°C for 15 seconds, and a final step at 72°C for 1 minute. The PCR products were first diluted 1/100 with DNase-free water and then used as templates for the second round of PCR. Second-round PCR products were purified with an Agencourt AMpure XP kit (BECKMAN COULTER Life Sciences, Brea, CA, USA). To distinguish each sample, dual-index sequences were added to the amplified bisulfite-treated products via an additional PCR step. The index sequences were defined by the combination of primers with TruSeq HT index 1 (D 7XX) and TruSeq HT index 2 (D 5XX) (Illumina, San Diego, CA, USA). Index PCR was performed using KAPA HiFi HS ReadyMix (Kapa Biosystems) under the following thermal cycling conditions: 95°C for 3 minutes, then 8 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 30 seconds, with a final step at 72°C for 5 minutes. The index PCR products were also purified with an Agencourt AMpure XP kit (BECKMAN COULTER Life Sciences).

The amplicons were sequenced on the Illumina MiSeq platform in 2 × 251 bp paired-end mode using MiSeq Reagent Nano Kit v2 500 cycles (Illumina). The sequenced reads were mapped to the human reference genome hg19 using the methylation analysis tool Bismark v0.18.1 (Babraham Bioinformatics, Babraham, Cambridgeshire, UK)[49]. Bismark detected the methylation rates of CpG sites by dividing the total number of mapped reads by the number of reads representing the methylation state at a given CpG site. Furthermore, we generated the pileup of the sequence data at targeted CpG sites using GATK's pileup command (Aaron et al., 2009). For each fragment derived from paired-end reads, we scanned the sequence corresponding to the targeted CpG sites and counted the number of fragments for individual sequence patterns.

The epipolymorphism score and methylation linkage diagram were calculated based on a previous report[35]. The CpG density at each CpG site was calculated by kernel density estimation (bandwidth = 10.73). Principal component analysis was performed using the average methylation rates, average pairwise correlations, and epipolymorphism scores for 3'UTR CpGs 1–9 and 3'UTR CpGs 10–15. The ‘pcaMethods’ R package[51] was used to apply the nipals algorithm with unit variance scaling.

RNA analysis of HEK293T cells

From 2.5×10^5 HEK293T cells, total RNA was extracted using a Nucleo Spin RNA XS kit (MACHEREY-NAGEL). First-strand cDNA was synthesized from total RNA using Prime Script Reverse Transcriptase (TAKARA Bio, Kusatsu, Japan). RT-PCR was performed using LA Taq (TAKARA Bio). The primer pairs employed for RT-PCR are listed in Supplementary Table 2. The thermal cycling conditions were as follows: 94°C for 2 minutes, then 35 cycles of 94°C for 30 seconds, 55°C for 45 seconds, and 72°C for 1 minute and 30 seconds (primer pair F1-R1) or 2 minutes and 30 seconds (primer

pair F1-R2), with a final step at 72°C for 5 minutes. The amplified products were separated by electrophoresis on a 2% agarose gel and quantified by using Image Quant TL analysis software (GE Healthcare, Chicago, IL, USA). Quantitative real-time PCR was performed on the TP-850 Real-Time PCR Detection System (TAKARA Bio) using SYBR Green Premix ExTaq II. The primer pairs employed for quantitative real-time PCR are listed in Supplementary Table 2. The thermal cycling conditions were as follows: 95°C for 30 seconds, then 40 cycles of 95°C for 15 seconds and 60°C for 30 seconds. The delta-delta CT method was used for quantitative evaluation.

RNA analysis of the human motor cortex

One patient with an unknown postmortem interval (control 9) (Supplement Table 1) was excluded from the RNA analysis. Total RNA was extracted from frozen brain tissue (100-200 mg) using a mirVana miRNA isolation kit (Applied Biosystems, Waltham, MA, USA). RNA quality was evaluated based on the RNA integrity number as determined using Tape Station 2200 (Agilent Technologies, Santa Clara, CA, USA), and samples with RNA integrity numbers below 6.0 were also excluded. First-strand cDNA was synthesized from total RNA using SuperScript VILO MasterMix (Invitrogen). Droplet digital PCR was performed using QX 200 ddPCR Supermix for Probes (No dUTP) on a QX 200 Droplet Digital PCR System (Bio Rad, Hercules, CA, USA). The primer pairs employed for PCR are listed in Supplementary Table 2. The thermal cycling conditions for the probe assays were as follows: 95°C for 10 minutes, then 40 cycles of 94°C for 30 seconds and 55°C for 2 minutes (ramping rate reduced to 2°C per second), and a final step at 98°C for 10 minutes.

Statistics

For each experiment, the corresponding statistics test, which was two-tailed, is indicated in each figure legend. The number of samples for each group is indicated in each figure legend. Statistical significance was considered to exist at $p < 0.05$.

Data availability

All relevant data are available from the corresponding author upon reasonable request.

References

1. Hou, Y. *et al.* Ageing as a risk factor for neurodegenerative disease. *Nat. Rev. Neurol.* **15**, 565–581 (2019).
2. Ross, C. A. & Poirier, M. A. Protein aggregation and neurodegenerative disease. *Nat. Med.* **10**, S10 (2004).

3. Tanaka, M. & Komi, Y. Layers of structure and function in protein aggregation. *Nat. Chem. Biol.* **11**, 373–377 (2015).

4. Henderson, M. X. *et al.* Spread of α-synuclein pathology through the brain connectome is modulated by selective vulnerability and predicted by network analysis. *Nat. Neurosci.* **22**, 1248–1257 (2019).

5. Niccoli, T., Partridge, L. & Isaacs, A. M. Ageing as a risk factor for ALS/FTD. *Hum. Mol. Genet.* **26**, R105–R113 (2017).

6. Al-Chalabi, A., Van Den Berg, L. H. & Veldink, J. Gene discovery in amyotrophic lateral sclerosis: Implications for clinical management. *Nat. Rev. Neurol.* **13**, 96–104 (2017).

7. Arai, T. *et al.* 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 (2006).

8. Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).

9. Tsai, K. J. *et al.* Elevated expression of TDP-43 in the forebrain of mice is sufficient to cause neurological and pathological phenotypes mimicking FTLD-U. *J. Exp. Med.* **207**, 1661–1673 (2010).

10. Xu, Y.-F. *et al.* Wild-type human TDP-43 expression causes TDP-43 phosphorylation, mitochondrial aggregation, motor deficits, and early mortality in transgenic mice. *J. Neurosci.* **30**, 10851–10859 (2010).

11. Ayala, Y. M. *et al.* TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J.* **30**, 277–288 (2011).

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

13. Sugai, A. *et al.* Non-genetically modified models exhibit TARDBP mRNA increase due to perturbed TDP-43 autoregulation. *Neurobiol. Dis.* **130**, 104534 (2019).

14. Koyama, A. *et al.* Increased cytoplasmic TARDBP mRNA in affected spinal motor neurons in ALS caused by abnormal autoregulation of TDP-43. *Nucleic Acids Res.* **44**, 5820–5836 (2016).

15. Sugai, A. *et al.* Robustness and vulnerability of the autoregulatory system that maintains nuclear TDP-43 levels: a trade-off hypothesis for ALS pathology based on in silico data. *Front. Neurosci.* **12**, 28 (2018).

16. Woerner, A. C. *et al.* Cytoplasmic protein aggregates interfere with nucleo-cytoplasmic transport of protein and RNA. *Science* **17**, 1–8 (2015).

17. Horvath, S. & Raj, K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat. Rev. Genet.* **19**, 371–384 (2018).

18. Zhang, M. *et al.* DNA methylation age acceleration is associated with ALS age of onset and survival. *Acta Neuropathol.* **139**, 943–946 (2020).

19. Linker, S. M. *et al.* Combined single-cell profiling of expression and DNA methylation reveals splicing regulation and heterogeneity. *Genome Biol.* **20**, 1–14 (2019).

20. Marina, R. J. *et al.* TET-catalyzed oxidation of intragenic 5-methylcytosine regulates CTCF-dependent alternative splicing. *EMBO J.* **35**, 335–355 (2016).

21. Shukla, S. *et al.* CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* **479**, 74–9 (2011).

22. Yearim, A. *et al.* HP1 Is Involved in Regulating the Global Impact of DNA Methylation on Alternative Splicing. *Cell Rep.* **10**, 1122–1134 (2015).

23. Maunakea, A. K., Chepelev, I., Cui, K. & Zhao, K. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res.* **23**, 1256–1269 (2013).

24. Song, Q. *et al.* A reference methylome database and analysis pipeline to facilitate integrative and comparative epigenomics. *PLoS One* **8**, (2013).

25. Zeng, J. *et al.* Divergent whole-genome methylation maps of human and chimpanzee brains reveal epigenetic basis of human regulatory evolution. *Am. J. Hum. Genet.* **91**, 455–465 (2012).

26. Lev Maor, G., Yearim, A. & Ast, G. The alternative role of DNA methylation in splicing regulation. *Trends Genet.* **31**, 274–280 (2015).

27. Oki, S. *et al.* Ch IP □Atlas: a data□mining suite powered by full integration of public Ch IP □seq data . *EMBO Rep.* **19**, 1–10 (2018).

28. Avendaño-Vázquez, S. E. *et al.* Autoregulation of TDP-43 mRNA levels involves interplay between transcription, splicing, and alternative polyA site selection. *Genes Dev.* **26**, 1679–1684 (2012).

29. Cole, J. J. *et al.* Diverse interventions that extend mouse lifespan suppress shared age-associated epigenetic changes at critical gene regulatory regions. *Genome Biol.* **18**, 1–16 (2017).

30. Reynolds, L. M. *et al.* Age-related variations in the methylome associated with gene expression in human monocytes and T cells. *Nat. Commun.* **5**, (2014).

31. Morita, S. *et al.* Targeted DNA demethylation in vivo using dCas9–peptide repeat and scFv–TET1 catalytic domain fusions. *Nat. Biotechnol.* 1–9 (2016).

32. Vojta, A. *et al*. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res.* **44**, 5615–5628 (2016).

33. Hernando-Herraez, I. *et al*. Ageing affects DNA methylation drift and transcriptional cell-to-cell variability in mouse muscle stem cells. *Nat. Commun.* **10**, 1–11 (2019).

34. Heyn, H. *et al*. Distinct DNA methylomes of newborns and centenarians. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 10522–10527 (2012).

35. Landan, G. *et al*. Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues. *Nat. Genet.* **44**, 1207–1214 (2012).

36. Scherer, M. *et al*. Quantitative comparison of within-sample heterogeneity scores for DNA methylation data. *Nucleic Acids Res.* 1–13 (2020).

37. Prestridge, D. Predicting Pol II promoter sequences using transcription factor binding sites. *J. Mol. Biol.* **249**, 923–932 (1995).

38. Scherf, M., Klingenhoff, A. & Werner, T. Highly specific localization of promoter regions in large genomic sequences by PromoterInspector: A novel context analysis approach. *J. Mol. Biol.* **297**, 599–606 (2000).

39. Palumbo, D., Affinito, O., Monticelli, A. & Cocozza, S. DNA Methylation variability among individuals is related to CpGs cluster density and evolutionary signatures. *BMC Genomics* **19**, 1–9 (2018).

40. Hernandez, D. G. *et al*. Distinct DNA methylation changes highly correlated with chronological age in the human brain. *Hum. Mol. Genet.* **20**, 1164–1172 (2011).

41. Horvath, S. *et al*. The cerebellum ages slowly according to the epigenetic clock. *Aging (Albany. NY).* **7**, 294–306 (2015).

42. Unnikrishnan, A. *et al*. The role of DNA methylation in epigenetics of aging. *Pharmacol. Ther.* 172–185 (2020).

43. Li, X. *et al*. Neocortical Tet3-mediated accumulation of 5-hydroxymethylcytosine promotes rapid behavioral adaptation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 7120–7125 (2014).

44. Oliveira, A. M. M., Hemstedt, T. J. & Bading, H. Rescue of aging-associated decline in Dnmt3a2 expression restores cognitive abilities. *Nat. Neurosci.* **15**, 1111–1113 (2012).

45. Rudenko, A. *et al*. Tet1 is critical for neuronal activity-regulated gene expression and memory extinction. *Neuron* **79**, 1109–1122 (2013).

46. Luo, C. *et al*. Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. *Science* **357**, 600–604 (2017).

47. Thalhammer, A., Jaudon, F. & Cingolani, L. A. Emerging Roles of Activity-Dependent Alternative Splicing in Homeostatic Plasticity. *Front. Cell. Neurosci.* **14**, 1–9 (2020).

48. Gitcho, M. a *et al.* *TARDBP* 3'-UTR variant in autopsy-confirmed frontotemporal lobar degeneration with TDP-43 proteinopathy. *Acta Neuropathol.* **118**, 633–645 (2009).

49. Krueger, F. & Andrews, S. R. Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).

50. Aaron McKenna, Matthew Hanna, Eric Banks, Andrey Sivachenko, Kristian Cibulskis, Andrew Kernytsky, Kiran Garimella, David Altshuler, Stacey Gabriel, M. D. and M. A. D. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data Aaron. *Genome Res.* **20**, 254–260 (2009).

51. Stacklies, W., Redestig, H., Scholz, M., Walther, D. & Selbig, J. pcaMethods - A bioconductor package providing PCA methods for incomplete data. *Bioinformatics* **23**, 1164–1167 (2007).

Acknowledgments

This research was supported by a grant-in-aid for Scientific Research on Innovative Areas (Brain Protein Aging and Dementia Control; 26117006) from MEXT; grants-in-aid for Scientific Research (A) (26250017 and 25253065), Scientific Research (C) (17K09751), and Young Scientists (19K23961) from JSPS; a grant-in-aid from the Research Committee of CNS Degenerative Diseases and Comprehensive Research on Disability, Health, and Welfare (13230021) from the Japanese Ministry of Health, Labour, and Welfare of Japan; the Tsubaki Memorial Foundation; the Takeda Science Foundation; and SERIKA FUND.

Author Contributions

YK, AS and OO designed this study and wrote the paper. YK conducted the experiments. YK and AS analyzed the results. NH and TI contributed to the experiments and the analyses using the MiSeq platform (Illumina). JI, MT and AK contributed to the provision of human autopsied tissues. AY, TI, TY and ST contributed to the interpretation of the results. All authors critically revised the draft and approved the final version.

Competing interest

The authors declare no competing interests.

Figure legends

Fig. 1 *TARDBP* 3'UTR CpG sites. (a) *TARDBP* DNA methylation status of the human prefrontal cortex obtained from MethBase. *TARDBP* transcripts, ALS-causing gene variants (OMIM Allelic Variants) and SNP positions as indicated by the Single Nucleotide Polymorphism Database (dbSNP) are also shown

with the UCSC browser. (b) CpG (circles) located around the 5' site of intron 7 in magnified views of exons 5 to 6 and 3'UTR, ChIP-seq data for H3K27ac (SRX2161995) and RNA polymerase II (SRX100455) in K562 cells, as indicated by ChIP-Atlas, and eCLIP data for TDP-43 in K562 cells from the ENCODE database (ENCSR584TCR).

Fig. 2 Targeted manipulation of the DNA methylation state of the *TARDBP* 3'UTR. (a) Schematic of the CpG sites of the *TARDBP* gene, the target regions of guide RNAs (gRNA1, 2, 3, 4), and primer sets A and B used for bisulfite amplicon sequencing. (b) Structure of the vectors used in the experiment. (c) Percent DNA methylation at 3'UTR CpGs 1–15 in HEK293T cells (mean ± SD, n=3). (d)(e) Proportions of the number of sequence reads classified by the number of CpGs methylated among nine CpG sites (3'UTR CpGs 1–9) or six CpG sites (3'UTR CpGs 10–15). (f) Schematic of epipolymorphisms showing methylation profiles with the same mean methylation level but different methylation patterns within the CpG region. Black and white circles represent methylated CpGs and demethylated CpGs, respectively. (g)(h) Scatter plot showing the epipolymorphism score for the DNA methylation percentage. (i) Methylation linkage diagram of 3'UTR CpGs 1–9 and 3'UTR CpGs 10–15, showing the correlation of all CpG site pairs in the target region. Green circles indicate the mean percentage of DNA methylation at each CpG site.

Fig. 3 Effect of the demethylation of *TARDBP* 3'UTR CpGs on the alternative splicing and mRNA expression of *TARDBP*. (a) Primer sets for the RT-PCR of *TARDBP* mRNA (arrows: F1/R1, F1/F2) and for quantitative real-time PCR (arrowheads). (b)(c) Analysis of alternative splicing by using primer set F1/R1 (b) or F1/F2 (c) (mean ± SD, n=3, Dunnett's test). (d) Quantitative real-time PCR analysis of *TARDBP* mRNA using RPLP1 and RPLP2 as reference genes (mean ± SD, n=3, Dunnett's test). *$p<0.05$, ****$p<0.001$. (e) Effect of the demethylation of *TARDBP* 3'UTR CpGs 10–15 on the TDP-43 autoregulation mechanism. Green circles indicate the percent methylation of each CpG site. The binding of TDP-43 to pre-mRNA (red) triggers alternative splicing, and this spliced isoform is degraded via nonsense-mediated mRNA decay. The amount of TDP-43 in the nucleus determines the ratio of these isoforms. However, the demethylation of the 3'UTR CpGs reduces the efficiency of TDP-43 autoregulation by attenuating alternative splicing and increasing canonical mRNA levels.

Fig. 4 Percentages of *TARDBP* gene DNA methylation in the human brain. (a) Schematic of the DNA methylation status of the human prefrontal cortex obtained from the methylomic database and the CpG region analyzed by bisulfite amplicon sequencing. (b)-(j) DNA methylation percentages of each CpG site in the *TARDBP* promoter region (b)(e)(h) and *TARDBP* 3'UTR (c)(d)(f)(g)(i)(j) in 7 ALS patients and 8 controls. Mean (large circles) ± SD.

Fig. 5 DNA methylation status of *TARDBP* among different brain regions and liver tissues. (a)(b)

Proportions of the number of sequence reads classified by the number of CpGs methylated among nine CpG sites (3'UTR CpGs 1–9) or six CpG sites (3'UTR CpGs 10–15). (c) Methylation linkage diagram of 3'UTR CpGs 1–9 and 3'UTR CpGs 10–15, showing the correlation of all CpG site pairs in the target region. Green circles indicate the mean percentage of DNA methylation at each CpG site. (d) Principal component analysis using mean methylation percentages, mean pairwise correlations, and epipolymorphism scores in 3'UTR CpGs 1-9 and 3'UTR CpGs 10–15. Ten ALS patients and 11 controls were included in the analysis of 3'UTR CpGs 10–15 in the motor cortex, five ALS patients were included in the analysis of the liver, and seven ALS patients and eight controls were included in the other analyses.

Fig. 6 Effect of aging on the DNA methylation status and mRNA expression of *TARDBP* in the motor cortex. (a)-(c) Scatter plot showing the mean DNA methylation percentage in the promoter region (a), 3'UTR CpGs 1–9 (b), and 3'UTR CpGs 10–15 (c) according to age at autopsy (Pearson's correlation test). (d) Scatter plot showing the percentage of DNA methylation at each CpG site among 3'UTR CpGs 10–15 according to age at autopsy in the controls (Pearson's correlation test). (e) Proportions of the number of sequence reads classified by the number of methylated CpGs among the six CpG sites (3'UTR CpGs 10–15) in the controls and ALS patients. (f) Correlation between age at autopsy and the epipolymorphism scores of 3'UTR CpGs 10–15 (Pearson's correlation test). (g) (h) Methylation linkage diagram of 3'UTR CpGs 10–15 in two of the youngest and two of the oldest individuals analyzed (g) and heat map (h) showing the extent of pairwise correlation for each CpG pair. (i) For each DNA methylation profile, the effects of the targeted manipulation of DNA methylation, differences according to brain region, and the effect of aging are shown as fold changes (log2) relative to the experimental control, liver, and middle-aged group results, respectively. (j) Scatter plot showing the association between *TARDBP* mRNA expression relative to that of the reference gene RPLP1 and the percentage of DNA methylation in 3'UTR CpGs 10–15 in 8 ALS patients and 8 controls. *$p<0.05$.

Fig.1

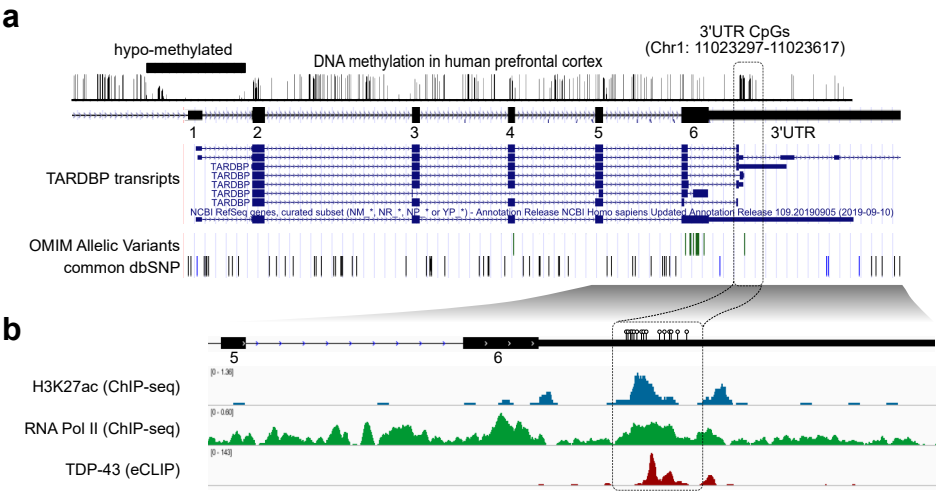

Fig. 1 TARDBP 3'UTR CpG sites. (a) TARDBP DNA methylation status of the human prefrontal cortex obtained from MethBase. TARDBP transcripts, ALS-causing gene variants (OMIM Allelic Variants) and SNP positions as indicated by the Single Nucleotide Polymorphism Database (dbSNP) are also shown with the UCSC browser. (b) CpG (circles) located around the 5' site of intron 7 in magnified views of exons 5 to 6 and 3'UTR, ChIP-seq data for H3K27ac (SRX2161995) and RNA polymerase II (SRX100455) in K562 cells, as indicated by ChIP-Atlas, and eCLIP data for TDP-43 in K562 cells from the ENCODE database (ENCSR584TCR).

Fig.2

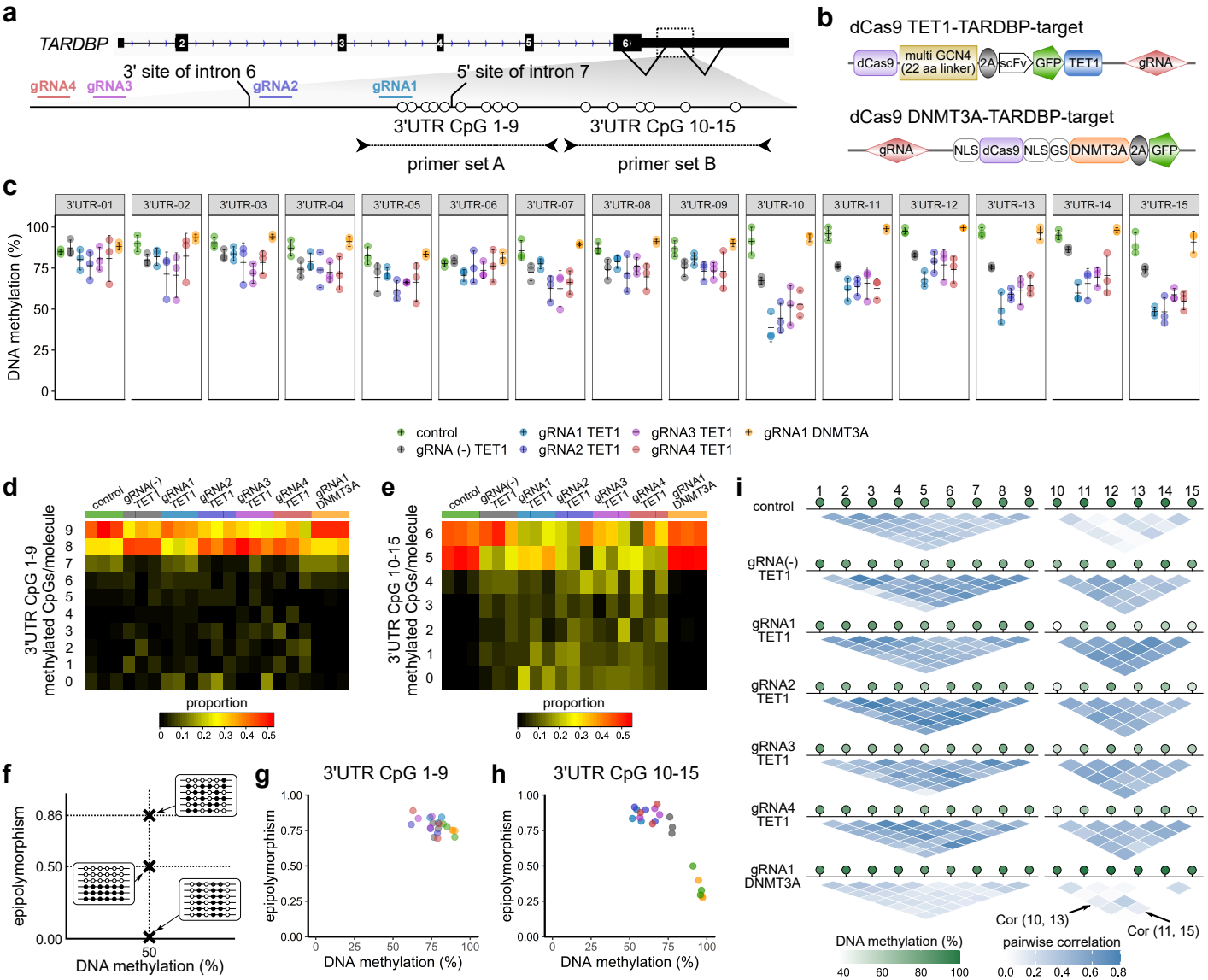

Fig. 2 Targeted manipulation of the DNA methylation state of the TARDBP 3'UTR. (a) Schematic of the CpG sites of the TARDBP gene, the target regions of guide RNAs (gRNA1, 2, 3, 4), and primer sets A and B used for bisulfite amplicon sequencing. (b) Structure of the vectors used in the experiment. (c) Percent DNA methylation at 3'UTR CpGs 1–15 in HEK293T cells (mean ± SD, n=3). (d)(e) Proportions of the number of sequence reads classified by the number of CpGs methylated among nine CpG sites (3'UTR CpGs 1–9) or six CpG sites (3'UTR CpGs 10–15). (f) Schematic of epipolymorphisms showing methylation profiles with the same mean methylation level but different methylation patterns within the CpG region. Black and white circles represent methylated CpGs and demethylated CpGs, respectively. (g)(h) Scatter plot showing the epipolymorphism score for the DNA methylation percentage. (i) Methylation linkage diagram of 3'UTR CpGs 1–9 and 3'UTR CpGs 10–15, showing the correlation of all CpG site pairs in the target region. Green circles indicate the mean percentage of DNA methylation at each CpG site.

Fig.3

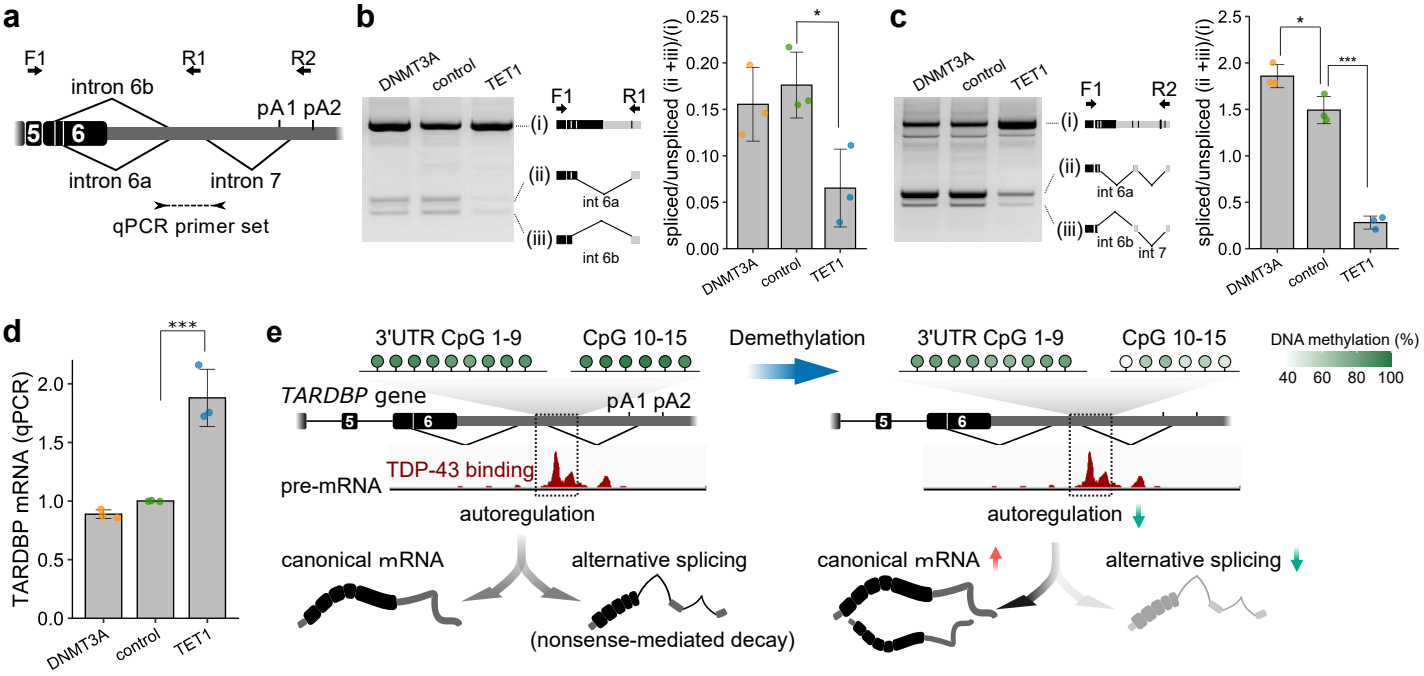

Fig. 3 Effect of the demethylation of TARDBP 3'UTR CpGs on the alternative splicing and mRNA expression of TARDBP. (a) Primer sets for the RT-PCR of TARDBP mRNA (arrows: F1/R1, F1/F2) and for quantitative real-time PCR (arrowheads). (b)(c) Analysis of alternative splicing by using primer set F1/R1 (b) or F1/F2 (c) (mean ± SD, n=3, Dunnett's test). (d) Quantitative real-time PCR analysis of TARDBP mRNA using RPLP1 and RPLP2 as reference genes (mean ± SD, n=3, Dunnett's test). *p<0.05, ****p<0.001. (e) Effect of the demethylation of TARDBP 3'UTR CpGs 10–15 on the TDP-43 autoregulation mechanism. Green circles indicate the percent methylation of each CpG site. The binding of TDP-43 to pre-mRNA (red) triggers alternative splicing, and this spliced isoform is degraded via nonsense-mediated mRNA decay. The amount of TDP-43 in the nucleus determines the ratio of these isoforms. However, the demethylation of the 3'UTR CpGs reduces the efficiency of TDP-43 autoregulation by attenuating alternative splicing and increasing canonical mRNA levels.

Fig.4

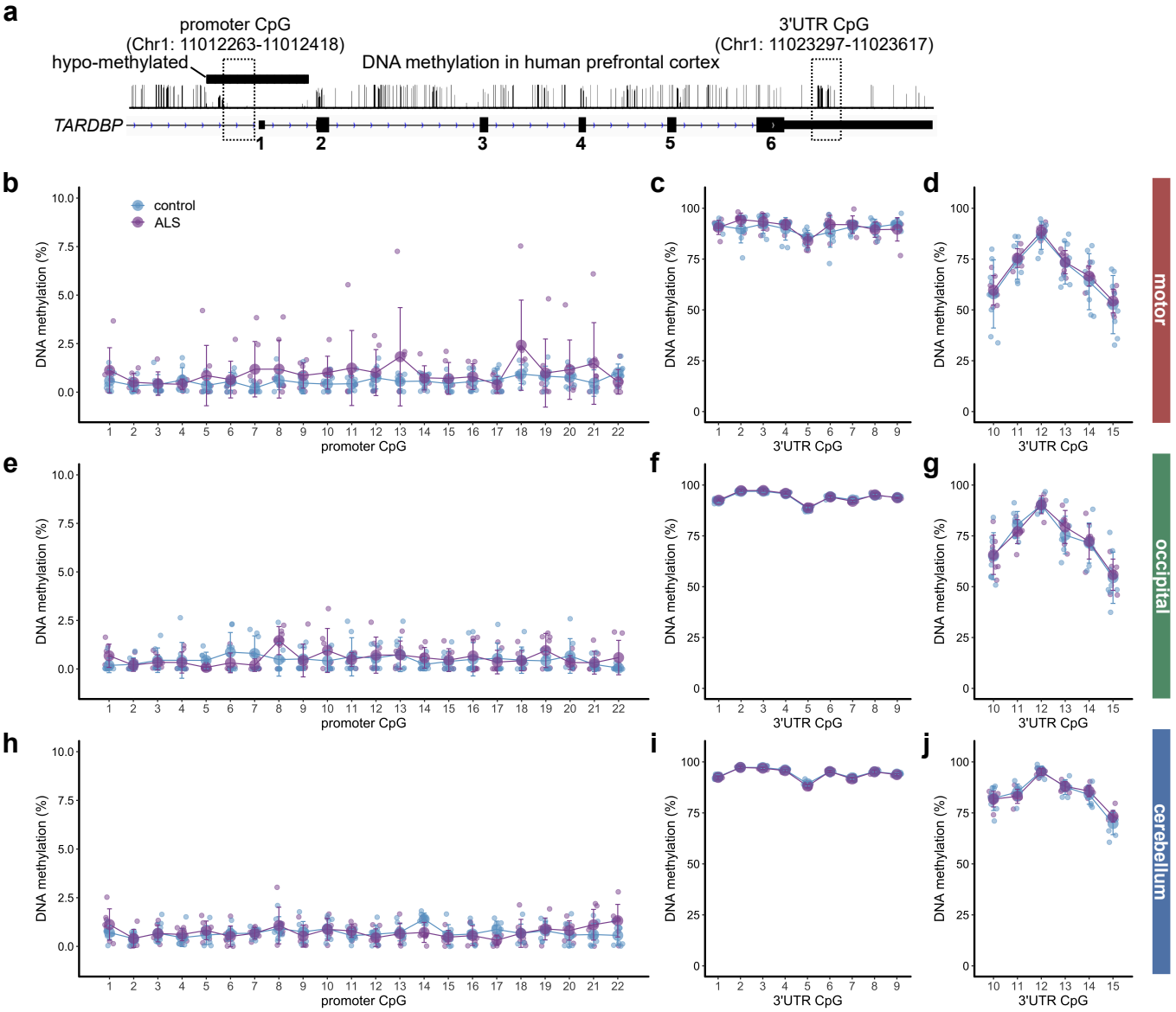

Fig. 4 Percentages of TARDBP gene DNA methylation in the human brain. (a) Schematic of the DNA methylation status of the human prefrontal cortex obtained from the methylomic database and the CpG region analyzed by bisulfite amplicon sequencing. (b)-(j) DNA methylation percentages of each CpG site in the TARDBP promoter region (b)(e)(h) and TARDBP 3'UTR (c)(d)(f)(g)(i)(j) in 7 ALS patients and 8 controls. Mean (large circles) ± SD.

Fig.5

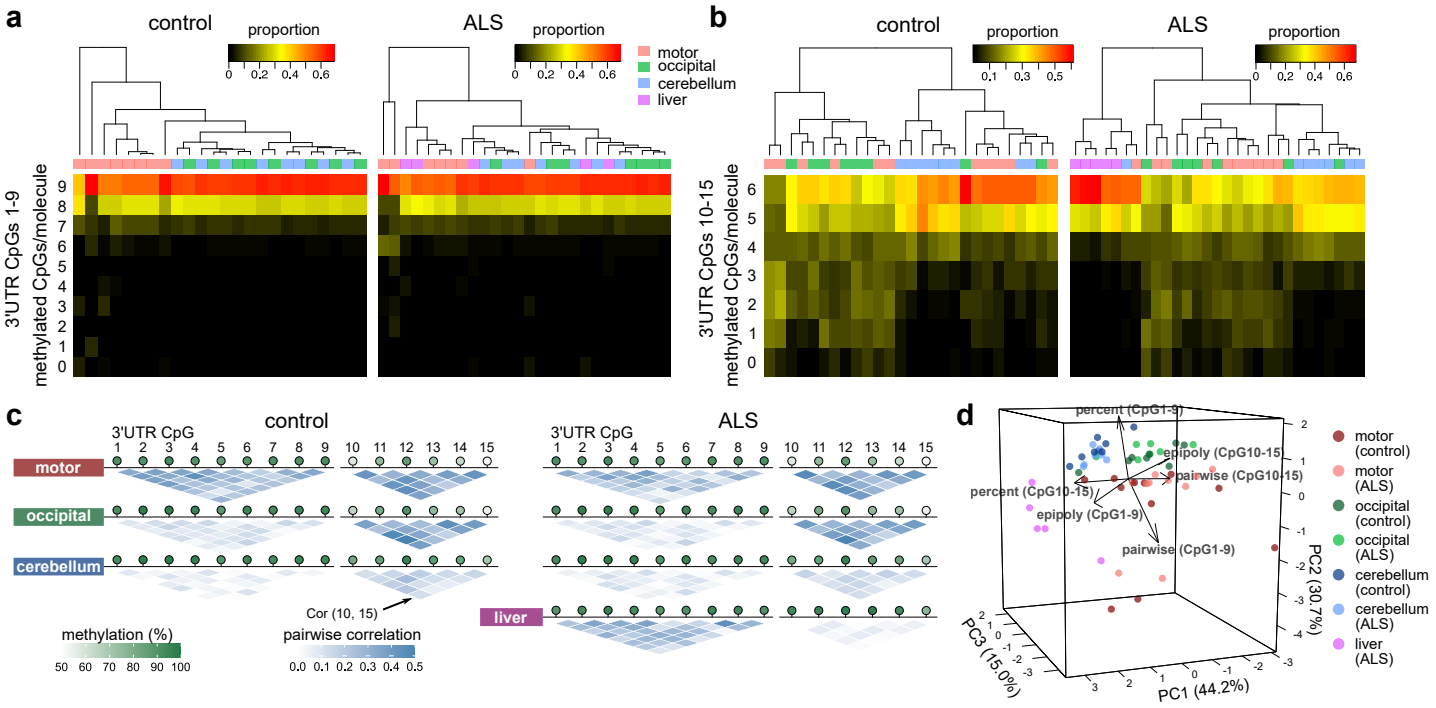

Fig. 5 DNA methylation status of TARDBP among different brain regions and liver tissues. (a)(b) Proportions of the number of sequence reads classified by the number of CpGs methylated among nine CpG sites (3'UTR CpGs 1–9) or six CpG sites (3'UTR CpGs 10–15). (c) Methylation linkage diagram of 3'UTR CpGs 1–9 and 3'UTR CpGs 10–15, showing the correlation of all CpG site pairs in the target region. Green circles indicate the mean percentage of DNA methylation at each CpG site. (d) Principal component analysis using mean methylation percentages, mean pairwise correlations, and epipolymorphism scores in 3'UTR CpGs 1-9 and 3'UTR CpGs 10–15. Ten ALS patients and 11 controls were included in the analysis of 3'UTR CpGs 10–15 in the motor cortex, five ALS patients were included in the analysis of the liver, and seven ALS patients and eight controls were included in the other analyses.

Fig.6

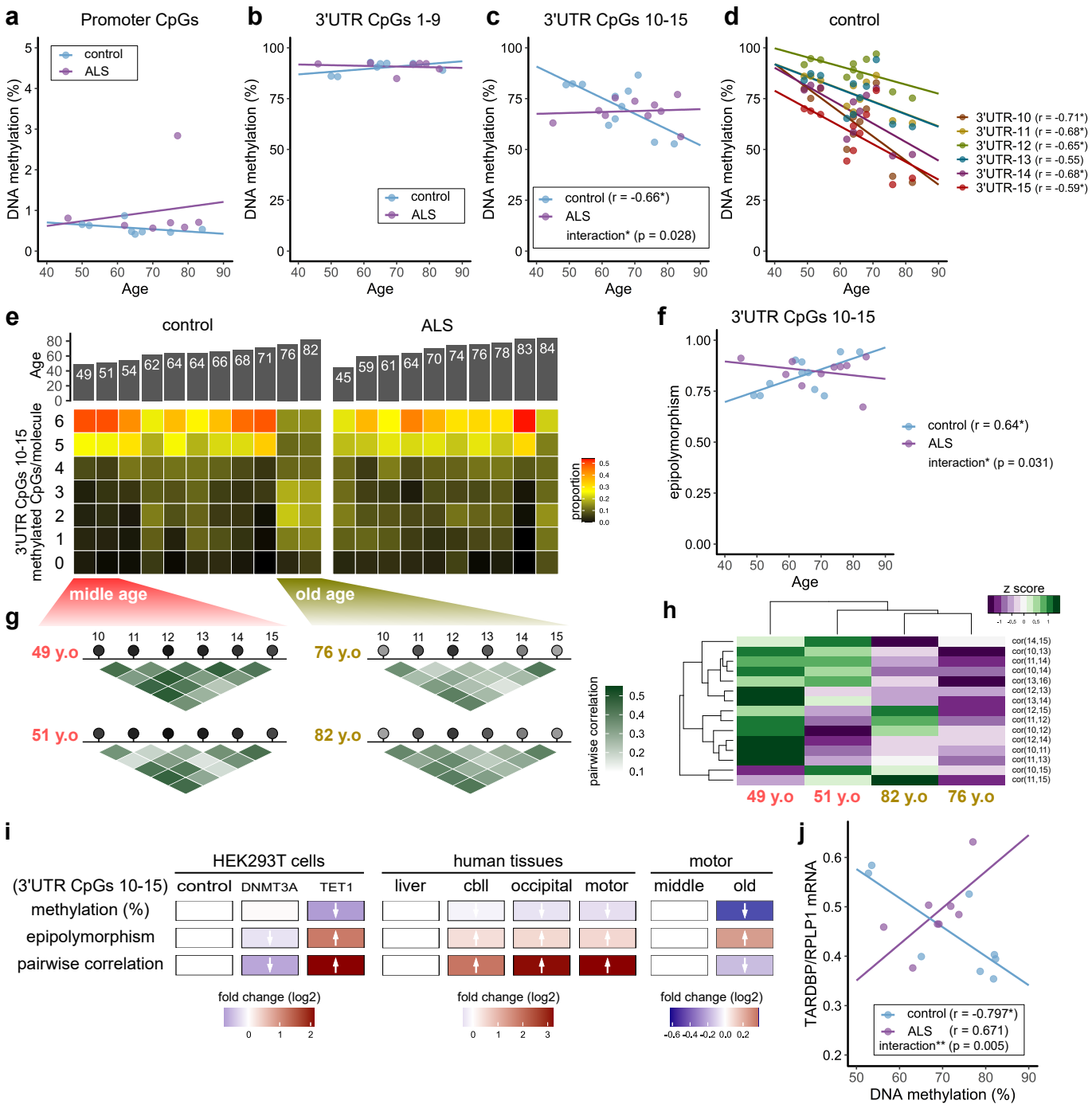

Fig. 6 Effect of aging on the DNA methylation status and mRNA expression of TARDBP in the motor cortex. (a)-(c) Scatter plot showing the mean DNA methylation percentage in the promoter region (a), 3'UTR CpGs 1–9 (b), and 3'UTR CpGs 10–15 (c) according to age at autopsy (Pearson's correlation test). (d) Scatter plot showing the percentage of DNA methylation at each CpG site among 3'UTR CpGs 10–15 according to age at autopsy in the controls (Pearson's correlation test). (e) Proportions of the number of sequence reads classified by the number of methylated CpGs among the six CpG sites (3'UTR CpGs 10–15) in the controls and ALS patients. (f) Correlation between age at autopsy and the epipolymorphism scores of 3'UTR CpGs 10–15 (Pearson's correlation test). (g) (h) Methylation linkage diagram of 3'UTR CpGs 10–15 in two of the youngest and two of the oldest individuals analyzed (g) and heat map (h) showing the extent of pairwise correlation for each CpG pair. (i) For each DNA methylation profile, the effects of the targeted manipulation of DNA methylation, differences according to brain region, and the effect of aging are shown as fold changes (log2) relative to the experimental control, liver, and middle-aged group results, respectively. (j) Scatter plot showing the association between TARDBP mRNA expression relative to that of the reference gene RPLP1 and the percentage of DNA methylation in 3'UTR CpGs 10–15 in 8 ALS patients and 8 controls. *p<0.05.