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 Research, Brain Research Institute, Niigata University

[3] Department of Molecular Genetics, Center for Bioresource-based Research, 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 Research, 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

Running title: Age-related demethylation prevents TDP-43 autoregulation

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

Abstract

In amyotrophic lateral sclerosis (ALS), TAR DNA-binding protein 43 (TDP-43), which is encoded by *TARDBP*, forms aggregates in the motor cortex. This aggregate formation may be triggered by an increase in the TDP-43 levels with aging. However, the amount of TDP-43 is autoregulated by the alternative splicing of the *TARDBP* 3'UTR, and how this autoregulation is affected by aging remains to be elucidated. We found that DNA demethylation in the autoregulatory region of the *TARDBP* 3'UTR reduced alternative splicing and increased *TARDBP* expression. Furthermore, in the human motor cortex, we found that this region was demethylated with age, resulting in an increased expression of *TARDBP*. The acceleration of DNA demethylation in the motor cortex was associated with the onset age of ALS. Collectively, the dysregulation of TDP-43 autoregulation by age-related DNA demethylation in the motor cortex may explain the contribution of aging and motor system selectivity in ALS.

Introduction

In neurodegenerative diseases with aging as a 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 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 unclear.

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, and similar to other sporadic neurodegenerative diseases, aging is a major risk factor[5]. In more than 95% of sporadic ALS cases, the nuclear protein TDP-43 forms aggregates in the cytoplasm and is depleted from the nucleus in diseased cells[6,7,8]. Increased intracellular concentrations of TDP-43 have been speculated to be a factor in this aggregate formation[9,10]. However, to date, the mechanism linking TDP-43, motor neuron selectivity, and aging has not been elucidated.

Nuclear TDP-43 binds the *TARDBP* pre-mRNA 3'UTR, and the level of TDP-43 is strictly autoregulated via its alternative splicing[11–13]. An increased level of nuclear TDP-43 promotes its splicing to produce nonsense-mediated mRNA decay-sensitive *TARDBP* mRNA and reduce the level of TDP-43[12,14]. However, when the level of TDP-43 in the nucleus is reduced, this splicing is repressed, and the *TARDBP* mRNA levels are increased. Therefore, *TARDBP* mRNA expression is continually increased in ALS-affected cells with reduced nuclear TDP-43 levels[14]. Following the formation of cytoplasmic

TDP-43 aggregates, which are presumed to reduce the TDP-43 level in the nucleus by either entrapping newly generated TDP-43 or inhibiting the nuclear transport of TDP-43, the progression of the disease accelerates[15,16]. Some mutations in the *TARDBP* gene can alter alternative splicing and affect the autoregulatory mechanism[17,18] However, in sporadic ALS, the factor triggering the disruption of this autoregulatory mechanism leading to increased TDP-43 expression and TDP-43 aggregate formation in the motor cortex is unclear.

Aging alters the methylation status of DNA. The DNA methylation status defines the epigenetic age separate from the chronological age. In age-related diseases, this epigenetic age contributes more than chronological age to mortality and morbidity[19]. Indeed, an accelerated epigenetic age based on DNA methylation is associated with the age of onset of ALS patients with *C9orf72* expansion repeats[20] and sporadic ALS patients[21]. DNA methylation not only regulates gene expression but also affects alternative splicing[22–26]. 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 and increased the expression of *TARDBP* mRNA. Furthermore, in the human motor cortex, this region was demethylated with aging, and the degree of demethylation was correlated with the level of *TARDBP* expression. An accelerated DNA methylation age of the *TARDBP* 3'UTR in the ALS motor cortex was associated with a younger age of onset, suggesting its involvement in ALS pathogenesis.

Results

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

TDP-43 binds 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)[27,28]. At methylated DNA sites, RNA polymerase II may stall, eliciting splicing[23,24,29]. Indeed, by analyzing publicly available data, we found that the CpG cluster region of the *TARDBP* 3'UTR is enriched in RNA polymerase II[30] (Fig. 1b), which may be associated with the alternative splicing of intron 7[31]. We found that this region is also enriched in H3K27ac marks (Fig. 1b)[30], a characteristic of regions where DNA methylation changes with age[32,33]. Thus, we 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 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[34] or the DNMT3A-*TARDBP*-target vector[35] 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 the 3'UTR CpG 1–9 sites, which were negatively correlated with the 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 investigated 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 heterogeneity in DNA methylation[36] and decrease the correlation between the methylation status of each CpG pair[37]. The methylation pattern of the 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)[38,39]. Furthermore, 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 that TET1 affects each CpG on the allele as a cluster rather than stochastically.

Then, we examined whether demethylation in the region suppresses the alternative splicing of *TARDBP* mRNA (Fig. 3a). In the 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 are sensitive to TET1, suppresses the alternative splicing of intron 6 and intron 7 and increases the *TARDBP* mRNA levels (Fig. 3e).

DNA methylation of *TARDBP* in the human brain

Subsequently, we 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: Case; ALS 1–7, control 1–8). We investigated 22 CpG sites in the *TARDBP* promoter region that were predicted by PROMOTER SCAN[40]

and PromoterInspector[41] 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 in 3'UTR CpGs 10–15 was lower than that 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 in 3'UTR CpGs 10–15 was greater than that in 3'UTR CpGs 1–9 (Fig. 4, Supplementary Fig. 2c) and was also associated with the surrounding CpG density (Supplementary Fig. 2d)[42]. 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 examined the motor cortex of 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 other regions (Fig. 5a) with high pairwise correlations in 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 in each CpG site (Fig. 5c) compared to those in the cerebellum. The epipolymorphism scores of the nervous tissues were higher than those of the liver (Supplementary Fig. 3a-f). These features were similar in the ALS and control groups. Each tissue was classified by a 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 in the motor cortex

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

the ALS group, the degree of DNA methylation was as low as approximately 65% at the age of 45 years and 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. 6e, Supplementary Fig. 5), which is consistent with the age-related increase in the heterogeneity of DNA methylation[36]. 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. 6f, g), suggesting stochastic demethylation of the epialleles with aging[37] as opposed to regulated demethylation by specific enzymes (Fig. 6h).

Demethylation of the *TARDBP* 3'UTR increases *TARDBP* expression in the motor cortex

Then, we investigated the association between the DNA methylation status of the 3'UTR CpGs and *TARDBP* expression in the motor cortex. The expression level of *TARDBP* mRNA tended to increase with DNA demethylation of 3'UTR CpGs 10–15 (Fig. 7a). In particular, the analysis of only the control group without brain disease showed a significant correlation (Fig. 7a). We also examined the TDP-43 protein levels in the motor cortex of 11 controls and 5 ALS patients whose tissue was used in the DNA methylation analysis (Fig. 7b). In the RIPA-soluble fraction, there was no association between the DNA methylation of 3'UTR CpGs 10–15 and the TDP-43 levels (Fig. 7b, c). However, in the RIPA-insoluble urea-soluble fraction, the full-length TDP-43 protein levels (43 kDa) were inversely correlated with the percentage of DNA methylation of 3'UTR CpGs 10–15 (Fig. 7b, d). The TDP-43 C-terminal fragment (CTF) in the urea-soluble fraction also showed a similar trend with a significant correlation in the analysis of the ALS group only (Fig. 7b, e).

Accelerated DNA demethylation of the *TARDBP* 3'UTR is associated with the onset age of ALS

Finally, we investigated the relationship between the DNA methylation status of the *TARDBP* 3'UTR and the clinical features of ALS patients. As shown in Fig. 6, *TARDBP* 3'UTR CpGs 10–15 showed demethylation in the middle-aged ALS group compared to the control group, suggesting that the relationship between epigenetic age rather than chronological age and clinical features should be considered. Therefore, we calculated the acceleration of the DNA methylation age in the ALS cases after estimating the epigenetic age based on the DNA methylation of 3'UTR CpG 10 in the control motor cortex (Fig. 8a). As a result, the acceleration of the DNA methylation age was significantly inversely correlated with the onset age of ALS (Fig. 8b). No correlation between DNA methylation age acceleration and disease duration was observed (Fig. 8c).

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 alternative splicing of the *TARDBP* 3'UTR and increased the expression of *TARDBP*. Third, the 3'UTR CpG methylation pattern differed among different brain regions. Fourth, the 3'UTR was demethylated by aging in the motor cortex, and the acceleration of demethylation in ALS is related to the age of onset. Thus, we propose that age-dependent demethylation of the *TARDBP* 3'UTR in the motor cortex may partially explain why aging is a risk factor of and 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 with alternative splice sites and can either enhance or suppress exon inclusion[25,26,29]. In general, DNA methylation causes RNA polymerase II to pause by modulating proteins, such as CCCTC-binding factor[23,24] and methyl-CpG binding protein 2[26,29]. Regarding the *TARDBP* 3'UTR, excess TDP-43 halts RNA polymerase II and increases the alternative splicing of intron 7[31]. However, CCCTC-binding factor and methyl-CpG binding protein are not enriched in the *TARDBP* 3'UTR in public databases[30]. Therefore, we consider the possibility that proteins other than these known factors may also be involved in the mechanism linking DNA methylation and splicing.

Next, we discuss the relationship between this age-related demethylation in the motor cortex and motor cortex selectivity in ALS. We found age-dependent demethylation of 3'UTR CpGs 10–15 in the motor cortex but not the occipital and cerebellar cortexes. This result is consistent with previous studies showing that the age-related state of DNA methylation differs across brain regions[43]. In particular, our results are consistent with results showing less of an effect of age-related methylation in the occipital lobe and cerebellum[44]. The differential expression of DNMT and TET may explain the specific DNA methylation profiles associated with aging in various brain regions[45]. Indeed, the expression of TET1, TET3, and DNMT3A2 is also affected by neural activity[46–48]. The accelerated demethylation of the *TARDBP* 3'UTR in the motor cortex of the ALS patients analyzed in this study was associated with the age of onset. Thus, 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.

We 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 near the splicing site of intron 6 may be affected. Second, the present analysis included a mixture of multiple cell types. Therefore, whether the observed age dependency was due to the senescence of each cell or changes in constituent cells due to aging remains unclear[39]. Furthermore, it may be difficult to properly evaluate the motor cortex derived from ALS patients, who exhibit advanced neuronal loss and gliosis with whole tissues. In the future, methylation analyses at the single-cell level should be performed to examine the

methylation changes in each cell type[22,49]. Third, the limited number of brain regions analyzed restricts our ability to emphasize tissue specificity. Finally, an analysis of a larger number of autopsy brains from patients without brain disease and ALS patients of varying ages should validate our conclusions in the future.

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[50]. Therefore, the identification of the 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 and 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)[34]. A DNMT3A-*TARDBP*-target vector incorporating dCas9 and DNMT3A, which is the catalytic domain of DNA methyltransferase, was applied for the methylation of the target DNA region (Fig. 2b)[35].

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, the 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

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. The sample size was first set to 8 ALS patients and 8 controls following previous reports investigating the relationship between the accelerated age of DNA methylation and the age of onset in ALS autopsy brains[20,21]. During the analysis, one ALS patient was excluded due to the detection of a mutation in *TBK1*. To analyze the relationship between the DNA methylation levels of CpGs 10-15 in the motor cortex and age, we included the motor cortex of 3 ALS patients and 3 controls and analyzed the data of 10 ALS cases and 11 controls (Figs. 5-8) (Supplementary Table 1: Case; ALS 8–10, control 9–11). Regarding the five ALS patients whose frozen liver tissue was preserved, we also conducted a liver analysis (Fig. 5). Although a variant (2,076 G>A) reported in a family lineage presenting with ALS and frontotemporal degeneration with an elevated expression of *TARDBP*[51] causes the loss of CpG 6 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 the 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). Using 25 mg of human postmortem tissue, genomic DNA was purified using a DNeasy Blood and Tissue kit (QIAGEN, Venlo, Limburg, Netherlands). According to the manufacturer's protocols, the purified genomic DNA was bisulfite treated with an 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). In the first and second rounds of PCR, the thermal cycling conditions were as follows: 95°C for 3 minutes, 30 cycles at 98°C for 20 seconds, 60°C for 15 seconds, 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. The 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, 8 cycles at 95°C for 30 seconds, 55°C for 30 seconds, 72°C for 30 seconds, and 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 an 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)[52]. Bismark detected the methylation rates of the 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 the targeted CpG sites using GATK's pileup command[53]. For each fragment derived from paired-end reads, we scanned the sequence corresponding to the targeted CpG sites and counted the number of fragments in individual sequence patterns.

The epipolymorphism score and methylation linkage diagram were calculated based on a previous report[38]. The CpG density at each CpG site was calculated by a kernel density estimation (bandwidth=10.73). A principal component analysis was performed using the average methylation rates, average pairwise correlations, and epipolymorphism scores of 3'UTR CpGs 1–9 and 3'UTR CpGs 10–15. The 'pcaMethods' R package[54] was used to apply the nipals algorithm with unit variance scaling.

RNA analysis of HEK293T cells

From 2.5×10^5 HEK293T cells, the total RNA was extracted using a Nucleo Spin RNA XS kit (MACHEREY-NAGEL). First-strand cDNA was synthesized from the 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, 35 cycles at 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 a 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, 40 cycles at 95°C for 15 seconds and 60°C for 30 seconds. The delta-delta CT method was used for the quantitative evaluation.

RNA analysis of the human motor cortex

One patient with an unknown postmortem interval (control 9) was excluded from the RNA analysis (Supplement Table 1). The total RNA was extracted from frozen brain tissue (100-200 mg) using a mirVana miRNA isolation kit (Applied Biosystems, Waltham, MA, USA). The 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 the 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, 40 cycles at 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.

Protein extraction from the human motor cortex and western blot analysis

The frozen motor cortex remaining after the DNA methylation analysis was homogenized in 10 volumes of ice-cold RIPA buffer (50 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with a protease inhibitor cocktail (Sigma). After sonication on ice (20%/10 sec, twice), 400 μl of lysates were centrifuged at 100,000×g at 4°C for 30 min. The supernatant was used as a RIPA soluble fraction. To ensure the removal of all RIPA-soluble proteins, the pellet was dissolved in RIPA buffer, sonicated (20%/10 sec, twice), and centrifuged at 100,000×g for 30 min at 4°C. The insoluble pellet was dissolved in 100 μl urea buffer (7 M urea, 2 M thiourea, 4% CHAPS, and 30 mM Tris, pH 8.5) supplemented with a protease inhibitor cocktail, sonicated (20%/10 sec, twice), and centrifuged at 100,000×g at 22°C for 30 min. The supernatant was used as the urea-soluble fraction. The protein concentration of the RIPA-soluble fraction was measured using a BCA protein assay kit (Thermo). The protein concentration of the urea-soluble fraction was assumed to have the same ratio as the RIPA-soluble fraction. The supernatant was resuspended in Laemmli Sample Buffer (BioRad) and treated at 96°C for 5 min. The proteins were separated by SDS-PAGE using a 10% polyacrylamide gel (Super SepTM Ace, Wako) and transferred to a PVDF membrane (Millipore). The blot was immersed in the primary antibody overnight at 4°C. The antibodies used for immunoblotting included a rabbit anti-TDP-43 (C-terminal) polyclonal antibody (Proteintech 12892-1-AP, 1:5000 for the RIPA-soluble fraction, 1:2000 for the urea-soluble fraction) and a mouse anti-GAPDH monoclonal antibody (MBL). HRP-conjugated secondary antibodies (Dako) were used. The bands were detected using Immobilon Western Chemiluminescent HRP Substrate (Millipore) and quantitatively analyzed using Amersham Imager 680 (GE Healthcare).

Statistics

For each experiment, the corresponding statistical test, which was two-tailed, is indicated in each figure legend. The number of samples per 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**, 1–15 (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. White, M. A. *et al*. TDP-43 gains function due to perturbed autoregulation in a Tardbp knock-in mouse model of ALS-FTD. *Nat. Neurosci.* 1–12 (2018).

18. Fratta, P. *et al.* Mice with endogenous TDP□43 mutations exhibit gain of splicing function and characteristics of amyotrophic lateral sclerosis. *EMBO J*. **37**, e98684 (2018).

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

20. Zhang, M. *et al.* DNA methylation age-acceleration is associated with disease duration and age at onset in C9orf72 patients. *Acta Neuropathol.* **134**, 271–279 (2017).

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

22. 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).

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

24. Shukla, S. *et al.* CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* **479**, 74–79 (2011).

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

26. 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).

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

28. 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).

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

30. 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).

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

32. 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).

33. 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).

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

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

36. 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).

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

38. 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).

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

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

41. 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).

42. 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).

43. 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).

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

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

46. 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).

47. 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).

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

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

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

51. 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).

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

53. 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).

54. 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 and AS conducted the experiments. YK and AS analyzed the results. NH and TI contributed to the experiments and 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 interests

The authors have no competing interests to declare.

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 of H3K27ac (SRX2161995) and RNA polymerase II (SRX100455) in K562 cells as indicated by ChIP-Atlas and eCLIP data of 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 the guide RNAs (gRNA1, 2, 3, and 4), and primer sets A and B used for the 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 of 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 alternative splicing and mRNA

expression of *TARDBP*. (a) Primer sets for the RT-PCR of *TARDBP* mRNA (arrows: F1/R1, F1/F2) and 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 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 3'UTR CpGs reduces the efficiency of TDP-43 autoregulation by attenuating alternative splicing and increasing the 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 8 controls and 7 ALS patients. Mean (large circles)±SD.

Fig. 5 DNA methylation status of *TARDBP* in 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 the mean methylation percentages, mean pairwise correlations, and epipolymorphism scores in 3'UTR CpGs 1-9 and 3'UTR CpGs 10–15. Eleven controls and 10 ALS patients 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 of the *TARDBP* 3'UTR in the motor cortex. (a)(b) Scatter plot showing the percentage of DNA methylation at each CpG site among *TARDBP* 3'UTR CpGs 10–15 according to age at autopsy (Pearson's correlation test) in the control (a) and ALS (b) groups. (c) Scatter plot showing the average DNA methylation of 3'UTR CpGs 10–15 according to age at autopsy in the controls (Pearson's correlation test). (d) 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. (e) Correlation between age at autopsy and the epipolymorphism scores of 3'UTR CpGs 10–15 (Pearson's correlation test). (f) (g) Methylation linkage diagram of 3'UTR CpGs 10–15 in two of the youngest and two of the oldest individuals analyzed (f) and heat map (g) showing the extent of the

pairwise correlation of each CpG pair. (h) 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.

Fig. 7 Effect of DNA methylation on *TARDBP* expression in the motor cortex. (a) Scatter plot showing the association between the percentage of DNA methylation of 3'UTR CpGs 10–15 and *TARDBP* mRNA expression relative to the reference gene *RPLP1* in 8 controls and 7 ALS patients (a total of 15 cases). (b) Western blot analysis of the RIPA-soluble fraction and urea-soluble fraction from the motor cortex of 11 controls and 5 ALS patients (a total of 16 cases) using an anti-TDP-43 antibody and anti-GAPDH antibody. (c) Scatter plot showing the association between the percentage of DNA methylation of 3'UTR CpGs 10–15 and RIPA-soluble full-length TDP-43 protein (43 kDa) expression relative to the reference protein GAPDH. (d)(e) Scatter plot showing the association between the percentage of DNA methylation of 3'UTR CpGs 10–15 and the urea-soluble full-length TDP-43 protein (43 kDa) (d) and the urea-soluble C-terminal fragment (CTF) of TDP-43 (e). Pearson's correlation test, *$p<0.05$.

Fig. 8 Association between DNA methylation age acceleration and the clinical features in ALS. (a) ‘Epigenetic age’ is estimated based on the regression coefficient calculated from the chronological age and the DNA methylation percentages at *TARDBP* 3'UTR CpG 10 in the control motor cortex (11 controls). The difference between the epigenetic age and the linear regression model is defined as ‘DNA methylation age acceleration’ in the ALS cases. Ten ALS patients were included in the analysis in relation to the clinical features. (b) Scatter plot showing the age at disease onset according to DNA methylation age acceleration at 3'UTR CpG 10. (c) Scatter plot showing the acceleration of the DNA methylation age at 3'UTR CpG 10 according to the disease duration. Pearson's correlation test.

Fig.1

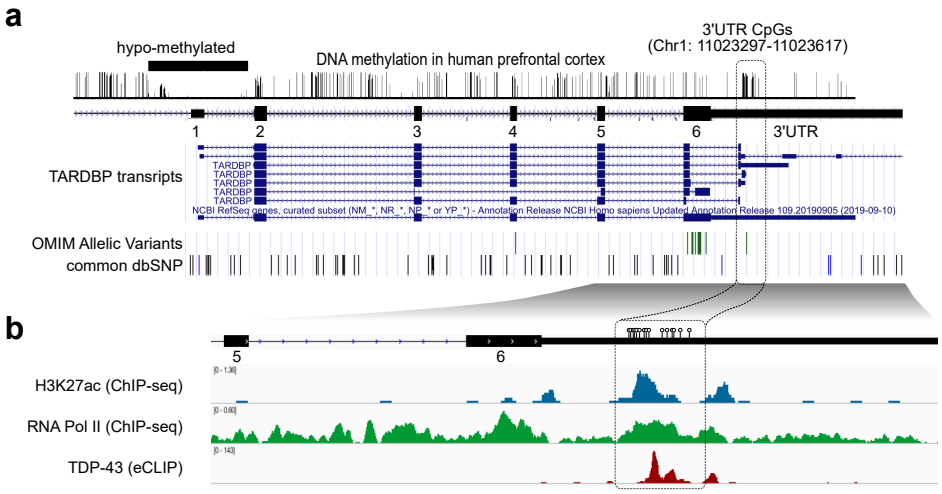

Fig. 1 TARDBP 3'UTR CpG sites. (a) TARDBP DNA methylation status of the human prefrontal cortex obtained from MethBase. TARDBP transcripts, ALS-causing gene variants (OMIM allelic variants) and SNP positions as indicated by the Single Nucleotide Polymorphism Database (dbSNP) are also shown with the UCSC 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 of H3K27ac (SRX2161995) and RNA polymerase II (SRX100455) in K562 cells as indicated by ChIP-Atlas and eCLIP data of TDP-43 in K562 cells from the ENCODE database (ENCSR584TCR).

Fig.2

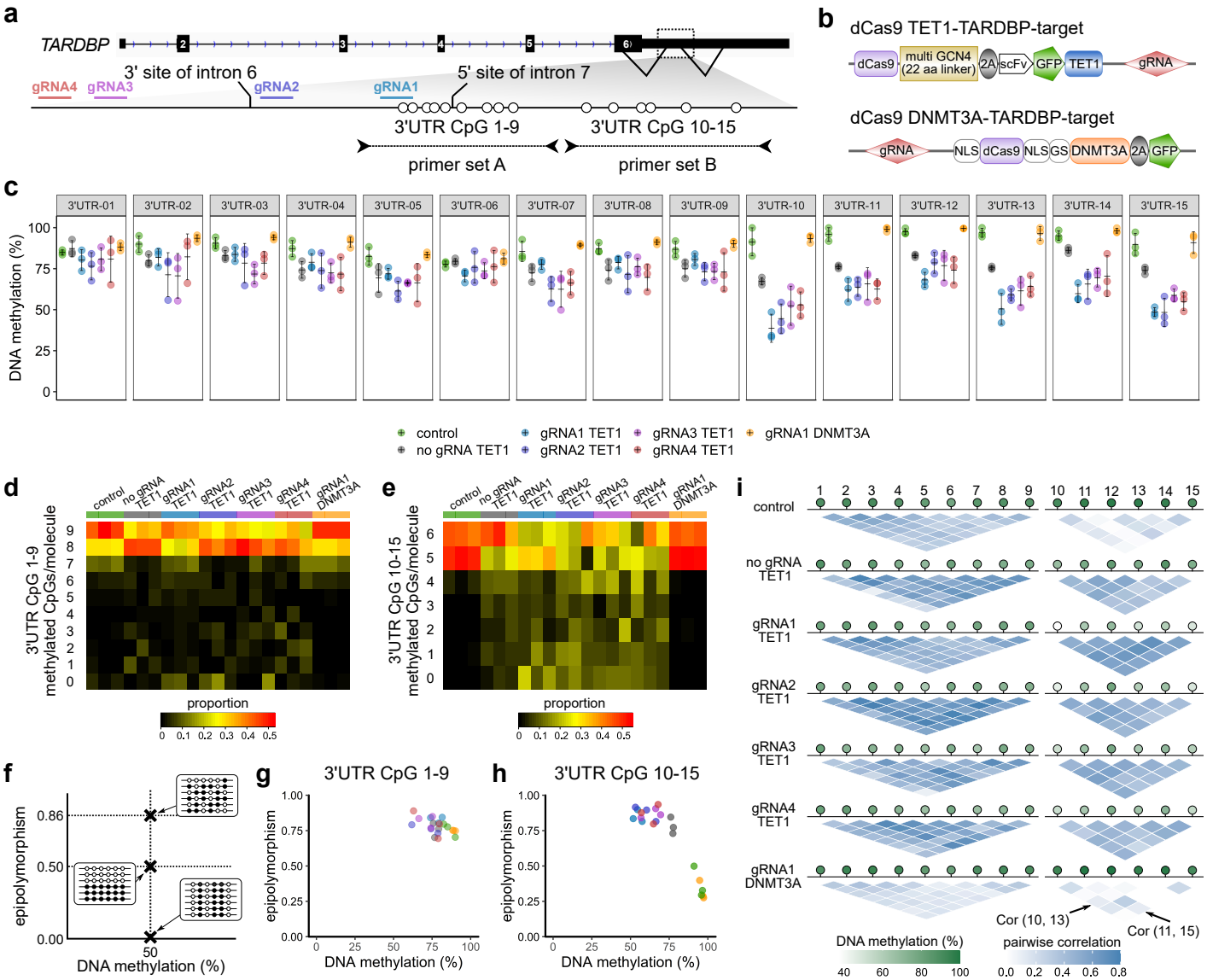

Fig. 2 Targeted manipulation of the DNA methylation state of the TARDBP 3'UTR. (a) Schematic of the CpG sites of the TARDBP gene, the target regions of the guide RNAs (gRNA1, 2, 3, and 4), and primer sets A and B used for the bisulfite amplicon sequencing. (b) Structure of the vectors used in the experiment. (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 of 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

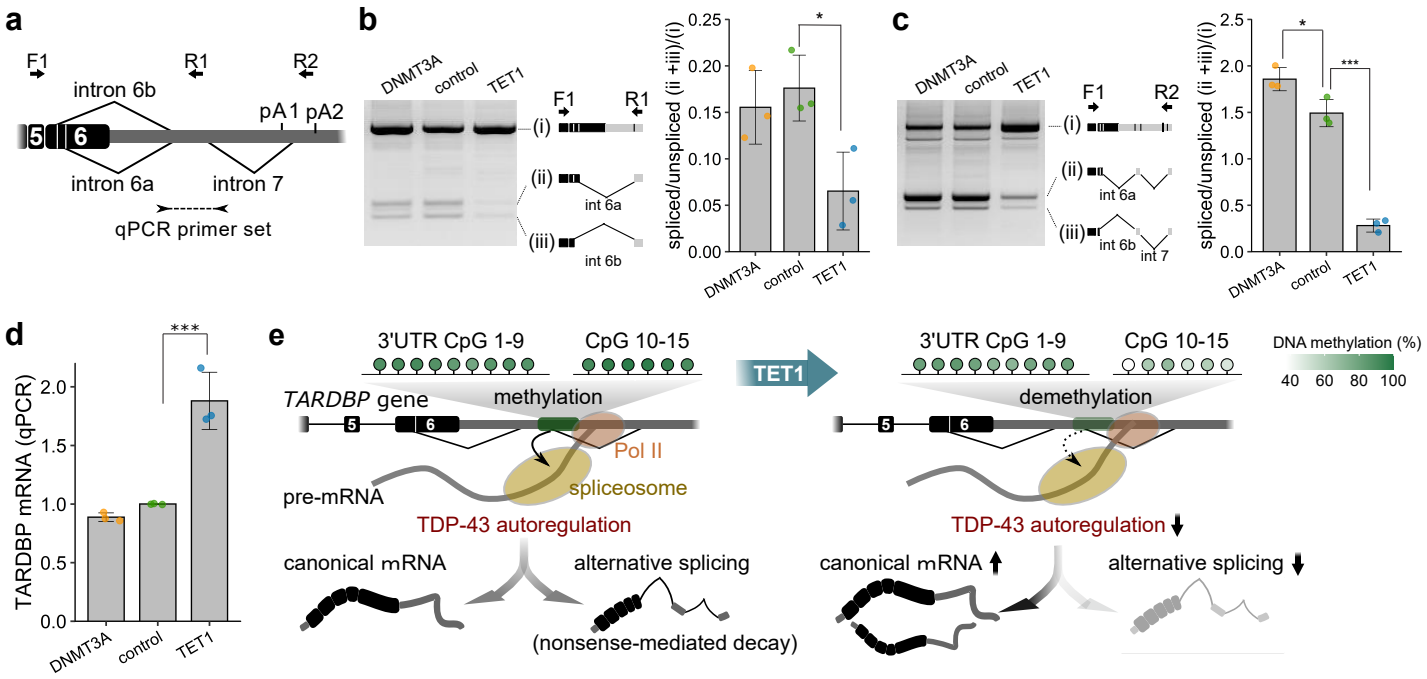

Fig. 3 Effect of the demethylation of TARDBP 3'UTR CpGs on alternative splicing and mRNA expression of TARDBP. (a) Primer sets for the RT-PCR of TARDBP mRNA (arrows: F1/R1, F1/F2) and quantitative real-time PCR (arrowheads). (b)(c) Analysis of alternative splicing by using primer set F1/R1 (b) or F1/F2 (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 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 3'UTR CpGs reduces the efficiency of TDP-43 autoregulation by attenuating alternative splicing and increasing the canonical mRNA levels.

Fig.4

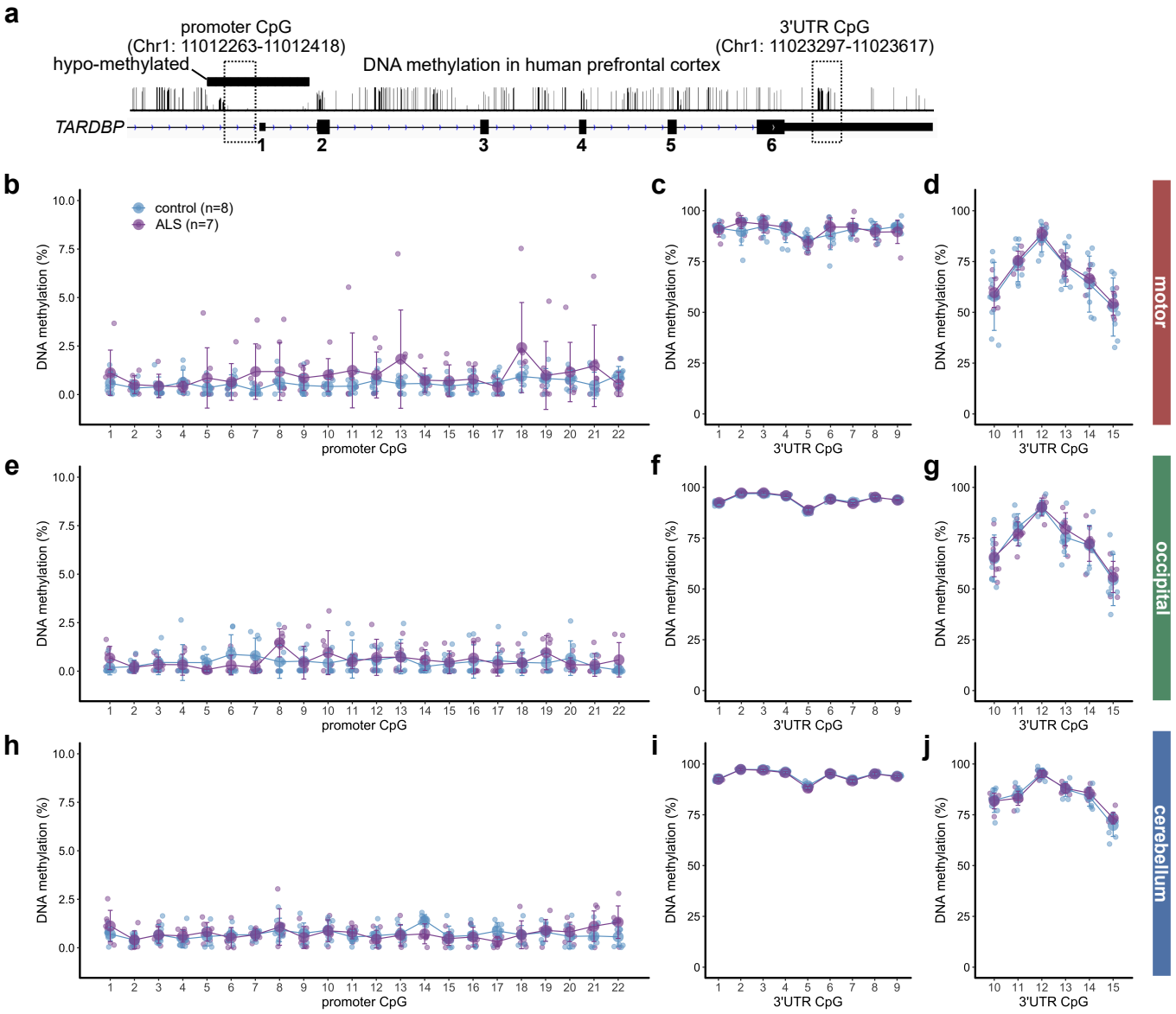

Fig. 4 Percentages of TARDBP gene DNA methylation in the human brain. (a) Schematic of the DNA methylation status of the human prefrontal cortex obtained from the methylomic database and the CpG region analyzed by bisulfite amplicon sequencing. (b)-(j) DNA methylation percentages of each CpG site in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

Fig.5

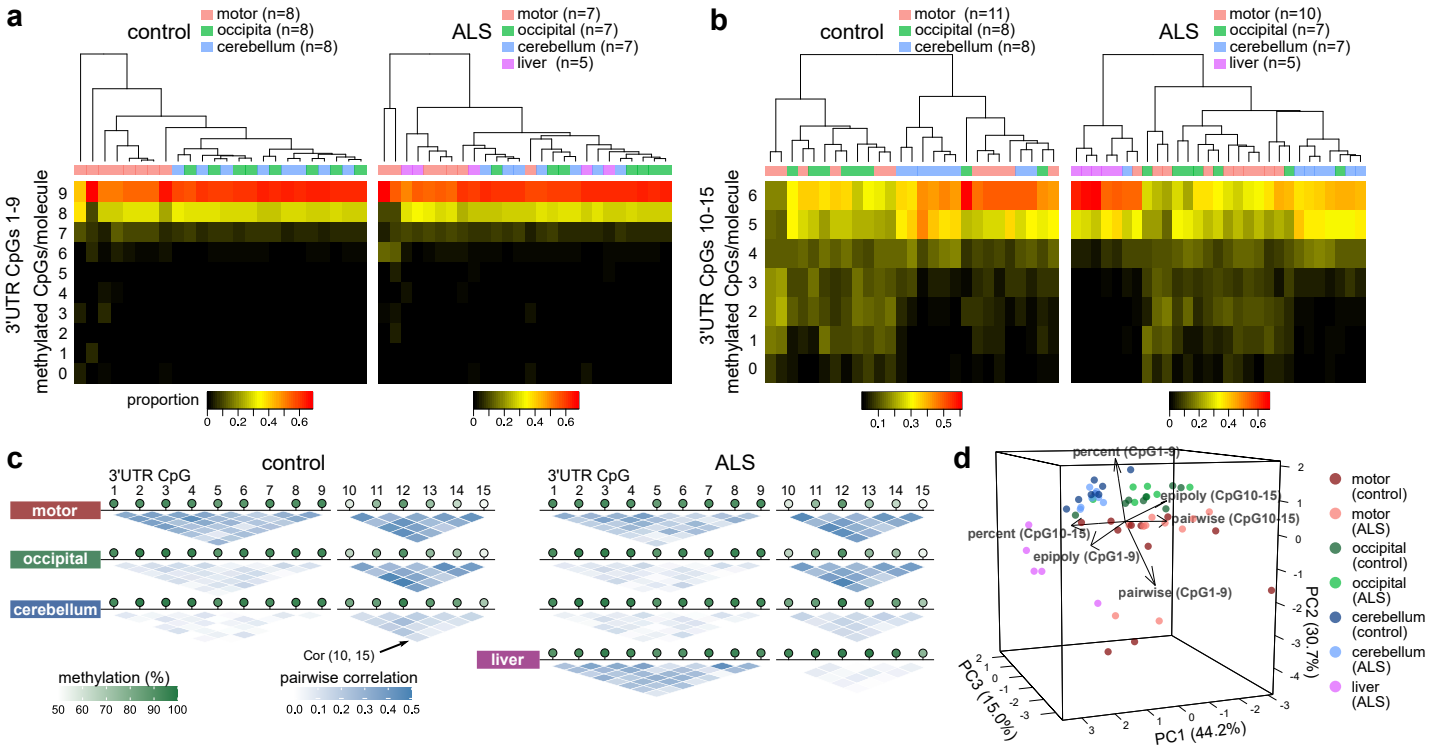

Fig. 5 DNA methylation status of TARDBP in different brain regions and liver tissues. (a)(b) Proportions of the number of sequence reads classified by the number of CpGs methylated among nine CpG sites (3'UTR CpGs 1–9) or six CpG sites (3'UTR CpGs 10–15). (c) Methylation linkage diagram of 3'UTR CpGs 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 the mean methylation percentages, mean pairwise correlations, and epipolymorphism scores in 3'UTR CpGs 1-9 and 3'UTR CpGs 10–15. Eleven controls and 10 ALS patients 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

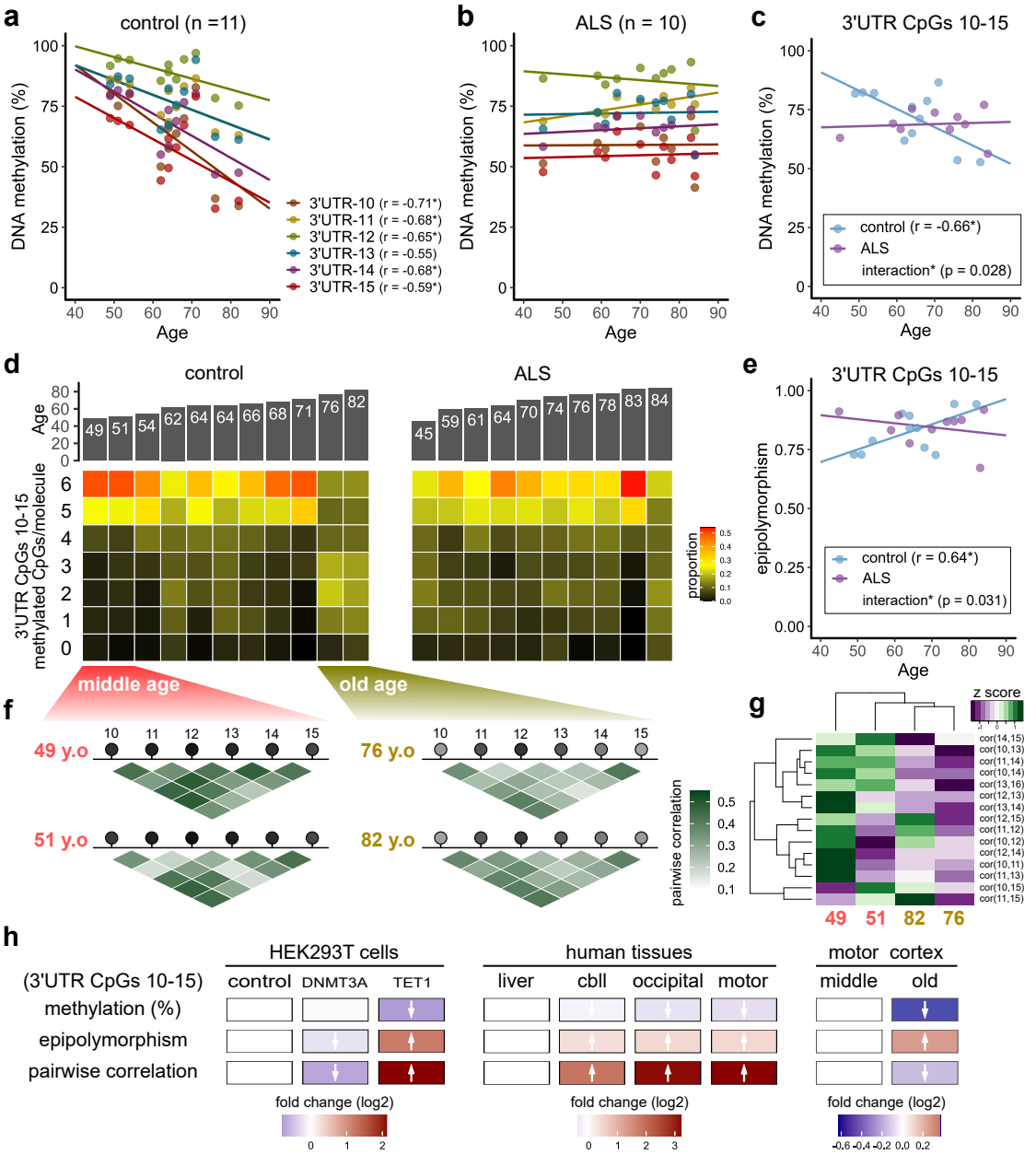

Fig. 6 Effect of aging on the DNA methylation status of the TARDBP 3'UTR in the motor cortex. (a)(b) Scatter plot showing the percentage of DNA methylation at each CpG site among TARDBP 3'UTR CpGs 10–15 according to age at autopsy (Pearson's correlation test) in the control (a) and ALS (b) groups. (c) Scatter plot showing the average DNA methylation of 3'UTR CpGs 10–15 according to age at autopsy in the controls (Pearson's correlation test). (d) 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. (e) Correlation between age at autopsy and the epipolymorphism scores of 3'UTR CpGs 10–15 (Pearson's correlation test). (f) (g) Methylation linkage diagram of 3'UTR CpGs 10–15 in two of the youngest and two of the oldest individuals analyzed (f) and heat map (g) showing the extent of the pairwise correlation of each CpG pair. (h) 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.

Fig.7

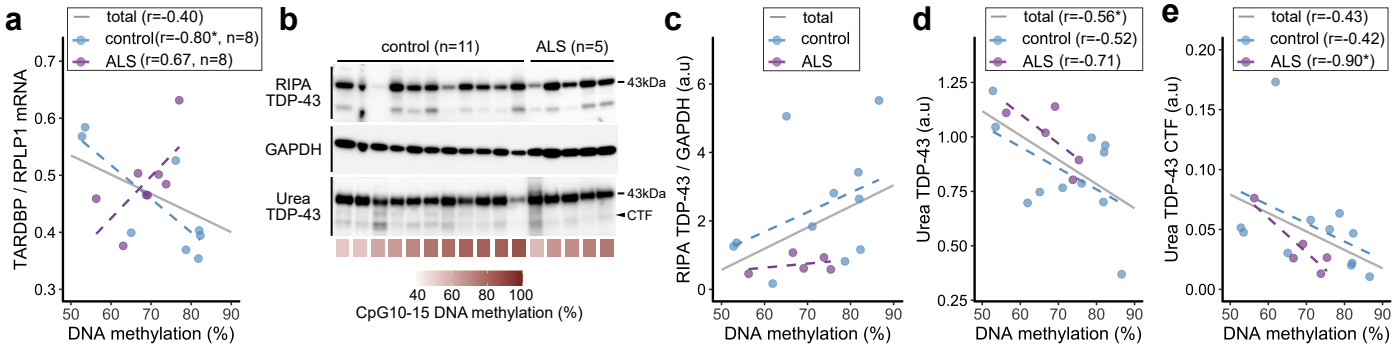

Fig. 7 Effect of DNA methylation on TARDBP expression in the motor cortex. (a) Scatter plot showing the association between the percentage of DNA methylation of 3'UTR CpGs 10–15 and TARDBP mRNA expression relative to the reference gene RPLP1 in 8 controls and 7 ALS patients (a total of 15 cases). (b) Western blot analysis of the RIPA-soluble fraction and urea-soluble fraction from the motor cortex of 11 controls and 5 ALS patients (a total of 16 cases) using an anti-TDP-43 antibody and anti-GAPDH antibody. (c) Scatter plot showing the association between the percentage of DNA methylation of 3'UTR CpGs 10–15 and RIPA-soluble full-length TDP-43 protein (43 kDa) expression relative to the reference protein GAPDH. (d)(e) Scatter plot showing the association between the percentage of DNA methylation of 3'UTR CpGs 10–15 and the urea-soluble full-length TDP-43 protein (43 kDa) (d) and the urea-soluble C-terminal fragment (CTF) of TDP-43 (e). Pearson's correlation test, *p<0.05.

Fig.8

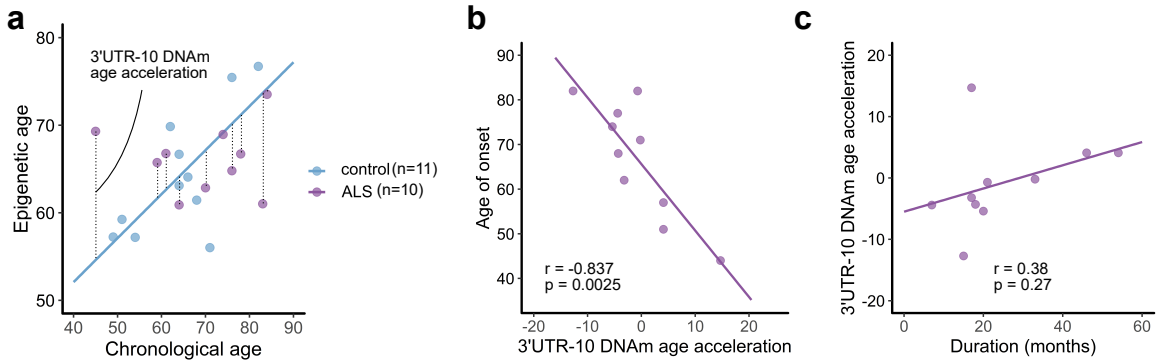

Fig. 8 Association between DNA methylation age acceleration and the clinical features in ALS. (a) 'Epigenetic age' is estimated based on the regression coefficient calculated from the chronological age and the DNA methylation percentages at TARDBP 3'UTR CpG 10 in the control motor cortex (11 controls). The difference between the epigenetic age and the linear regression model is defined as 'DNA methylation age acceleration' in the ALS cases. Ten ALS patients were included in the analysis in relation to the clinical features. (b) Scatter plot showing the age at disease onset according to DNA methylation age acceleration at 3'UTR CpG 10. (c) Scatter plot showing the acceleration of the DNA methylation age at 3'UTR CpG 10 according to the disease duration. Pearson's correlation test.