*2438–2457* *Nucleic Acids Research, 2017, Vol. 45, No. 5* *Published online 6 December 2016*
*doi: 10.1093/nar/gkw1197*

# Binding of MBD proteins to DNA blocks Tet1 function thereby modulating transcriptional noise

**Anne K. Ludwig[1,†], Peng Zhang[1,†], Florian D. Hastert[1], Stephanie Meyer[1], Cathia Rausch[1], Henry D. Herce[1], Udo Müller[2], Anne Lehmkuhl[1], Ines Hellmann[3], Carina Trummer[2], Christian Storm[4], Heinrich Leonhardt[2] and M. Cristina Cardoso[1,*]**

[1]Cell Biology and Epigenetics, Department of Biology, Technische Universität Darmstadt, 64287 Darmstadt, Germany, [2]Human Biology and BioImaging, Department of Biology II, LMU Munich, 82152 Martinsried, Germany, [3]Anthropology and Human Genomics, Department Biology II, LMU Munich, 82152 Martinsried, Germany and [4]Chemical Plant Ecology, Department of Biology, Technische Universität Darmstadt, 64287 Darmstadt, Germany

Received June 16, 2016; Revised November 03, 2016; Editorial Decision November 16, 2016; Accepted November 20, 2016

## ABSTRACT

**Aberrant DNA methylation is a hallmark of various human disorders, indicating that the spatial and temporal regulation of methylation readers and modifiers is imperative for development and differentiation. In particular, the cross-regulation between 5-methylcytosine binders (MBD) and modifiers (Tet) has not been investigated. Here, we show that binding of Mecp2 and Mbd2 to DNA protects 5-methylcytosine from Tet1-mediated oxidation. The mechanism is not based on competition for 5-methylcytosine binding but on Mecp2 and Mbd2 directly restricting Tet1 access to DNA. We demonstrate that the efficiency of this process depends on the number of bound MBDs per DNA molecule. Accordingly, we find 5-hydroxymethylcytosine enriched at heterochromatin of Mecp2-deficient neurons of a mouse model for Rett syndrome and Tet1-induced re-expression of silenced major satellite repeats. These data unveil fundamental regulatory mechanisms of Tet enzymes and their potential pathophysiological role in Rett syndrome. Importantly, it suggests that Mecp2 and Mbd2 have an essential physiological role as guardians of the epigenome.**

## INTRODUCTION

Methylation of DNA is generally accepted to be decisively involved in regulating gene expression (1). In mammals, 5-methylcytosine (5mC) accounts for 1% of all DNA bases and is primarily found as symmetrical methylation of CpG dinucleotides (2). A minor proportion of 5mC is localized within so-called CpG islands at the 5′ ends of many genes, including those, responsible for genomic imprinting and X-inactivation (3). The vast majority of methylated cytosines, however, are found in repetitive, endoparasitic sequences (4), whose transcriptional activity must be repressed to prevent translocations, gene disruption and chromosomal instability (5,6). The methylome is read and translated by conserved families of proteins, such as the methyl-CpG binding domain proteins (7). All members (of which the five best studied ones are Mecp2, Mbd1, Mbd2, Mbd3 and Mbd4) share a common protein motif, the methyl-CpG-binding domain (MBD) (8), which enables all family members except for Mbd3 to selectively bind to single methylated CpG dinucleotides (9). Moreover, all MBD proteins with the exception of Mbd4 have been described to function in transcriptional repression in part by recruiting silencing complexes such as histone deacetylases (HDACs) (1,10).

Mecp2, the founding member of the MBD protein family, is highly expressed in brain and was shown to mediate silencing of neuronal genes by the recruitment of the Sin3a–HDAC chromatin remodeling complex via its transcriptional repression domain, abbreviated TRD (10,11). In addition, Mecp2 was described to link methylated DNA with the nuclear receptor corepressor (NCoR), as well as the silencing mediator of retinoic acid and thyroid receptor (SMRT) in a neuronal activity dependent manner (12,13). Unlike its name suggests, Mecp2 binds preferentially, but not exclusively to methylated DNA (9,14,15). In addition to its core methyl-CpG binding domain (MBD), Mecp2 contains various non-sequence specific interaction sites for double-stranded DNA, including the TRD domain and, based on their relative location to the MBD and TRD, the so-called intervening domain (ID), as well as the C-terminal domain alpha (CTD alpha) (14). Upon binding to DNA, the ID and TRD domains of Mecp2, which constitute a large proportion of the extensively disordered protein, acquire secondary structure and stabilize Mecp2-chromatin complexes. Accordingly, deletion of these DNA binding do-

*To whom correspondence should be addressed. Tel: +49 6151 16 21882; Fax: +49 6151 16 21880; Email: cardoso@bio.tu-darmstadt.de
†These authors contributed equally to this work as first author.

mains were shown to considerably increase the fraction of unbound Mecp2 molecules within the cell nucleus (14,16). Besides this, MBD-based binding affinity was described to highly depend on the density of methylated CpG sites (15) and, thus, might vary extensively among different cell types. In mouse cells, Mecp2 was described to highly accumulate at densely methylated pericentric heterochromatin (17). As a consequence of homo- and hetero-interactions with itself and Mbd2 (18), as well as its multivalent DNA and 5mC binding ability, Mecp2 induces large-scale chromatin reorganization (19) accompanied by dampening transcriptional noise of highly methylated repetitive elements (20).

More recently, three mammalian enzymes (TET1-3) named after the ten-eleven translocation (t(10;11)(q22;23)) identified in a few cases of acute myeloid and lymphocytic leukemia (21–23), were shown to catalyze the conversion of 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxycytosine (5caC) in an iterative, Fe(II)- and oxoglutarate dependent oxidation reaction (23–25). This may either result in the erasure of the repressing methylcytosine mark with the aid of deaminases and enzymes of the base excision repair system (26), or the stable genomic integration of the oxidized cytosine derivatives as additional epigenetic information (27). Consequently, TET proteins have been proposed to play a key role in the long sought mechanism of active DNA demethylation (23), as well as in diversifying the epigenetic landscape, whose composition is dynamically regulated during development and in disease (27).

DNA hypo- as well as hypermethylation as a consequence of miss- or nonfunctioning 5mC writers, readers and modifiers, have been implicated in many malignancies including neurological and autoimmune disorders and cancer (28). Mutations in the X-linked *MECP2* gene cause Rett-syndrome (29,30), a debilitating neurological disease that, at a molecular level, is characterized by increased expression and retrotransposition of repetitive elements (20,31).

By dissecting the interplay of 5mC readers and modifiers, we test the hypothesis of whether the anomalous transcriptional response observed in Rett patients is due to unconfined access of TET proteins to their substrate 5mC. In accordance with this, our data unveil a molecular mechanism by which Mecp2 and Mbd2 protect 5mC from Tet1 mediated oxidation *in vivo* and *in vitro* and provide definite indications of aberrant Tet activity in a mouse model for Rett syndrome, which lacks the aforementioned MBD-based defense system.

## MATERIALS AND METHODS

### Plasmids

Mammalian expression constructs coding for GFP-tagged mouse Mbd2-, mouse Mbd3- and rat Mecp2 full length proteins, rat Mecp2 deletion mutants (Mecp2G.9: aa 163–310 and Mecp2Y.5: aa 77–162), as well as for human Mecp2 deletion mutant R111G were previously described (8,18,19,32,33).

For construction of the mCherry-tagged catalytic active (Tet1CD: aa 1365–2057) and inactive (Tet1CDmut: aa 1365–2057, H1652Y, D1654A) domain of mouse Tet1, Np95 was replaced from the mammalian expression vector pCAG-mCherry-Np95-IB (34) by Tet1CD (27) and Tet1CDmut (35), respectively using AsiSI and NotI sites.

For the expression of GFP, the commercial vector pEGFP-N1 (Clontech; Mountain View, CA, USA) was used.

Insect expression constructs coding for GFP-tagged mouse Mbd2- and rat Mecp2 full length proteins, as well as for rat Mecp2 deletion mutants Mecp2G.9 (aa 163–310) and Mecp2Y.5 (aa 77–162) were previously described (18,32,36).

For construction of the His-tagged catalytic domain of mouse Tet1, an N-terminal Histag, enterokinase- and AsiSI cutting site were amplified (fwd primer: 5′-gcc cga att cat gag cca tc-3′, rev primer: 5′-ccc ggc ggc cgc tta-3′) from an oligo (gcc cga att cat gag cca tca tca tca tca tca tga cga cga cga caa gag cga tcg cat gtc aac cag gag gga agc tta agc ggc cgc cgg g) and inserted into the commercial transfer vector pFBDM of the MultiBac Expression System (37) using EcoRI and NotI sites. The catalytic active domain of mouse Tet1 (aa 1365–2057) was then cut from the mammalian expression vector (27) and inserted into the modified pFBDM transfer vector using AsiSI and NotI sites.

For the generation of the GFP- and mCherry-tagged catalytic active domain of mouse Tet1, GFP- and cherry-tagged Tet1CD, were cut from the mammalian expression vector pCAG-GFP-Tet1CD (27) and pCAG-cherry-Tet1CD (described above), respectively and inserted into the modified pFBDM transfer vector (described above) using BamHI and NotI sites.

### Cell culture and transfection

C2C12 mouse myoblasts (38) were cultured using standard conditions described previously (39).

C2C12 cells were grown to 70% confluence on 16 mm glass cover slips in 6-well plates and transfected 3 hours post seeding using poly-ethylenimine (PEI, 1 mg/ml in $ddH_2O$, pH 10; Sigma-Aldrich, St. Louis, MO, USA) (40).

HEK 293-EBNA cells (Invitrogen; Paisley PA4 9RF, UK) were grown as previously described (41).

For the isolation of genomic DNA, HEK cells were seeded in 6-well plates at a target density of $5 \times 10^5$ cells per well and transfected 24 hours post seeding using poly-ethylenimine (PEI, 1mg/mL in $ddH_2O$, pH 7; Sigma-Aldrich, St. Louis, MO, USA).

For the production of GFP-tagged Tet1CD (27) and Mecp2 (19) proteins, HEK cells were grown to 70% confluence and transiently transfected using poly-ethylenimine (PEI, 1mg/ml in $ddH_2O$, pH 7; Sigma-Aldrich, St. Louis, MO, USA).

Sf9 insect cells (Invitrogen, Paisley PA4 9RF, UK) used for protein production were cultivated and transfected as previously described (18).

V6.5 wild type and triple Tet-knockout mouse embryonic stem cells (42) were maintained under serum-free and feeder-free conditions on Geltrex-coated flasks in N2B27 medium (50% neurobasal (Life Technologies, Carlsbad, California, USA) and 50% DMEM/F12 medium (Life Technologies, Carlsbad, California, USA) containing 2 mM L-glutamine (Life Technologies, Carlsbad, CA, USA), 0.1 mM β-mercaptoethanol, N2 supplement (Life Technolo-

gies, Carlsbad, CA, USA), B27 serum-free supplement (Life Technologies, Carlsbad, CA, USA), 100 U/ml Penicillin-Streptomycin, 1000 U/ml LIF and 2i (1 μM PD032591 and 3 μM CHIR99021 (Axon Medchem, Groningen, Netherlands)).

Mouse tail fibroblast (MTF) lox/y and MTF -/y (*Mecp2* knockout) (43) cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum. These cells are also heterozygous for *Mbd2*.

Primary neurons were isolated from brain of adult C57BL/6 mice. Whole brains were removed from mice under sterile conditions, cut into small pieces, put into 10 ml HBSS (Hank's Balanced Salt Solution) and washed by centrifugation (200 × g, 1 min). After centrifugation, HBSS was discarded and the brain pellet resuspended in 5 ml 0.25% trypsin solution supplemented with 150 units DNaseI. After incubation for 20 min at 37°C in a waterbath, cells were centrifuged (200 × g, 1 min) again. Trypsin was discarded and the pellet resupended in 5 ml FBS for 2 min at 37°C. After centrifugation (200 × g, 1 min), FBS was removed and the pellet resuspended in 5 ml Neurobasal medium (Thermo Fisher Scientific, Waltham, MA, USA) supplemented with B-27 (Thermo Fisher Scientific, Waltham, MA, USA). Next, the brain suspension was triturated by passing it through a 10 ml serological plastic pipette for 15 times and subsequently through a flamed-tip glass Pasteur pipette for 15 times. Following centrifugation (400 × g, 1 min), the pellet was resuspended in 10 ml fresh neurobasal medium. Big parts were allowed to settle for around 30 s and the supernatant was distributed in dishes, set up with laminin (Sigma, St. Louis, MO, USA) coated glass coverslips. Cells were cultured for 10 days by replacing 50% of the culture medium every two to three days before they were used for immunostaining.

### Mice

Dissected brains of male Mecp2 knockout mice (43) (∼ postnatal day 40) (WL.B6.129P2(C)-Mecp2tml.1Bird>/J; Charles River Laboratories International, Inc., Wilmington, MA 01887, USA) used for immunostainings were kindly provided by AM Bischoff, lab of D. Richter (Universitätsmedizin Göttingen, Germany).

Dissected brains of male Mecp2 knockout mice (43) (∼ postnatal day 40) (WL.B6.129P2(C)-Mecp2tml.1Bird>/J; Charles River Laboratories International, Inc., Wilmington, MA 01887, USA) used for RNA isolation were kindly provided by the laboratory of Adrian Bird.

Brains of wild type mice (∼ postnatal day 40) (C57BL/6N; Charles River Laboratories International, Inc., Wilmington, MA 01887, USA) were used as control.

### Protein preparation

GFP-tagged Tet1CD proteins (used in Supplementary Figure S9) were prepared from whole cell lysates of HEK cells 36 h post-transfection using 50 mM $NaH_2PO_4$; pH 7.5, 150 mM NaCl, 10 mM imidazole, 0.5% Tween-20, 0.5 mM EDTA, 2 mM $MgCl_2$, 0.5 mM $CaCl_2$, 1 mM PMSF, 1 μg/μl DNaseI and 1× mammalian protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO, USA). Following centrifugation, supernatant was added to pre-equilibrated (50 mM $NaH_2PO_4$; pH 7.5, 150 mM NaCl, 10 mM imidazole, 0.5 mM EDTA and 0.05% Tween-20) Ni-NTA beads that were coupled to the GFP-binding protein (GFP-Trap, ChromoTek, Planegg-Martinsried, Germany) (44,45) and incubated for 2 h at 4°C on a rotary shaker. To remove unbound proteins, beads were centrifuged and washed with wash buffer (50 mM $NaH_2PO_4$; pH 7.5, 300 mM NaCl, 10 mM imidazole and 0.1% Tween-20). Elution was performed using wash buffer containing 250 mM imidazole. Elution buffer was exchanged to 20 mM Tris–HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 1 mM DTT and 100 ng/μl BSA using PD-10 desalting columns (GE Healthcare, Freiburg, Germany).

GFP-tagged Mecp2 R111G proteins were prepared from whole cell lysates of HEK cells 24 h post-transfection using re-suspension buffer (32) containing 1 M NaCl and protease inhibitors in following concentrations: AEBSF 1 mM (AppliChem, Darmstadt, Germany), E64 10 μM (AppliChem, Darmstadt, Germany), Pepstatin A 1 μM (Sigma-Aldrich, St. Louis, MO, USA) and Aprotinin 2 ng/ml (Sigma-Aldrich, St. Louis, MO, USA). Cells were disrupted by syringe treatment (3 × 20 gauge, 3 × 21 gauge) followed by incubation on ice for 10 min.

Proteins were eluted by the addition of 4 M $MgCl_2$, pH 4.4 and subsequent incubation on ice for 10 min. Elution buffer was exchanged to 1× PBS using Amicon Ultra centrifugal filter units (Sigma-Aldrich, St. Louis, MO, USA).

All of the other GFP-, YFP- and mCherry-tagged proteins were purified from Sf9 insect cells as previously described (18,32) with following exceptions: The re-suspension buffer (32) was supplied with protease inhibitors in concentrations as described above. For the purification of Tet proteins, the sodium chloride (NaCl) concentration of the re-suspension buffer was decreased to 0.5 M. Cells were disrupted by syringe treatment (3 × 20 gauge, 3 × 21 gauge) followed by incubation on ice for 10 min. Proteins were eluted by the addition of 4 M $MgCl_2$, pH 4.4 and subsequent incubation on ice for 10 minutes. Elution buffer was exchanged to 1× PBS using Amicon Ultra centrifugal filter units (Sigma-Aldrich, St. Louis, MO, USA).

His-tagged proteins were purified from Sf9 insect cells using TALON ion metal affinity chromatography (Clontech Laboratories, Inc., CA, USA) according to the manufacturer's instructions with following changes. The re-suspension buffer contained 50 mM $NaH_2PO_4$, 300 mM NaCl and 10 mM imidazole, pH 8.0 and was supplied with protease inhibitors as described above. The elution buffer contained 50 mM $NaH_2PO_4$, 300 mM NaCl and 150 mM imidazole, pH 8.0. Elution buffer was exchanged to 1× PBS using Amicon Ultra centrifugal filter units (Sigma-Aldrich, St. Louis, MO, USA).

### Western blot analysis

HEK cells were lysed in re-suspension buffer (32) containing 1 M NaCl for 20 min on ice and whole protein lysates were blotted as described before (46) on a nitrocellulose membrane (GE Healthcare, München, Germany). Visualization of the immunoreactive bands was achieved by ECL plus Western Blot Detection reagent (GE Healthcare, München, Germany). The following antibodies were used:

monoclonal mouse anti GFP (Roche, Mannheim, Germany), polyclonal rabbit anti PCNA (Santa Cruz Biotechnology, Heidelberg, Germany), monoclonal rat anti RFP (47), Alexa488 conjugated goat anti mouse IgG (The Jackson Laboratory, Bar Harbor, USA), cy5 conjugated donkey anti rabbit IgG (The Jackson Laboratory, Bar Harbor, USA) and TexasRed conjugated donkey anti rat IgG (The Jackson Laboratory, Bar Harbor, USA).

### Genomic DNA preparation

For the preparation of genomic DNA (gDNA), sorted HEK-EBNA, as well as MTF lox/y and -/y cells were pelleted (10 min, 2000 rpm, 4°C) and incubated overnight at 50°C in TNES buffer (10 mM Tris; pH 7.5, 400 mM NaCl, 10 mM EDTA, 0.6% SDS) supplemented with 1 mg/ml Proteinase K (Carl Roth, Karlsruhe, Germany) (48). RNA was removed by the addition of 0.6 mg/ml RNase A (Qiagen, Hilden, Germany) for 30 min at 37°C. gDNA was extracted by the addition of 6 M NaCl at a final concentration of 1.25 M and vigorous shaking (48). After centrifugation (15 min, 13200 rpm, RT), gDNA was precipitated from the supernatant by the addition of 100% ice cold ethanol followed by incubation at –20°C for 1 h and subsequent centrifugation (10 min, 13 200 rpm, 4°C). After a washing step in 70% ethanol, gDNA was air dried and solved in $ddH_2O$. Isolated gDNA from HEK cells was fragmented (<2000 bp) by sonication using the Biorupter TM UCD-200 (Diagenode, Seraing Ougrée, Belgium). The concentration of gDNA was measured on a TECAN infinite M200 plate reader (Tecan Group Ltd., Maennedorf, Switzerland).

### RNA preparation

For sorted mouse C2C12 myoblasts, total RNA was isolated using the RNeasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instruction. To remove traces of genomic DNA, RNA was treated with RNase-free recombinant DNaseI (Macherey Nagel, Dueren, Germany) for 90 min at 37°C and further purified with the Qiagen RNeasy Mini Kit. To assess the concentration and purity of RNA, the ratio of absorbance at 260 and 280 nm was measured on a TECAN infinite M200 plate reader (Tecan Group Ltd., Maennedorf, Switzerland).

For Mecp2 y/- and wild type mouse brain, total RNA was isolated and treated with RNase-free DNaseI as previously described (49).

### Flow cytometry

C2C12 mouse myoblasts transiently expressing mCherry-tagged Tet1CD/Te1CDmut and GFP-tagged Mecp2 proteins, as well as HEK-EBNA cells transiently co-expressing high protein levels of mCherry-Tet1CD and GFP-tagged MBD proteins (Mecp2, Mbd2, IDTRD and MBD, respectively) were respectively trypsinized, re-suspended in PBS and separated from untransfected cells by fluorescent-activated cell sorting (FACS) on a S3e Cell Sorter (Bio-Rad Laboratories, Hercules, CA, USA) equipped with 488 and 561 nm excitation lasers and 525 ± 30 and 586 ± 25 nm emission filters, respectively. Sorted populations were either processed for RNA- (C2C12 cells) or gDNA preparation (HEK-EBNA cells), respectively.

### Real-time quantitative reverse transcription-polymerase chain reaction of major satellite repeats

For C2C12 mouse myoblasts, 20–200 ng of total RNA were used for cDNA synthesis using 200 units M-MuLV reverse transcriptase (NEB, Frankfurt, Germany), 0.01 OD units random primer of the Prime-It II Random Primer Labeling Kit (Stratagene, La Jolla, CA, USA), 0.5 mM dNTPs (Carl Roth, Karlsruhe, Germany) and 40 units recombinant ribonuclease inhibitor RNaseOUT (Invitrogen, Paisley PA4 9RF, UK) in a total reaction volume of 20 μl. Cycles were set to 10 min at 25°C, 60 min at 42°C and 20 min at 65°C.

For Mecp2 y/- and wt mouse brain, cDNA was synthesized as described previously (49) and kindly provided by Congdi Song.

Equal amounts of cDNA (0.5 ng) were used for real-time PCR with Platinum SYBR Green qPCR SuperMix-UDG w/ROX (Invitrogen, Paisley PA4 9RF, UK) on a StepOnePlus Real-Time PCR System (Applied Biosystems, Darmstadt, Germany) according to the manufacturer's instruction. UDG was inactivated for 2 min at 50°C and cDNA was denatured for 10 min at 95°C. Cycle parameters were set to 40 cycles of 15 s at 95°C and 45 s at 60°C. Specificity of amplification products was confirmed by melting curve analysis.

Gene expression level were normalized to Gapdh and calculated using the comparative CT method (ΔΔCT method).

Primers for quantitative real-time PCR contained the following sequences: Gapdh forward: 5′-CCA TAC ATA CAG GTT TCT CCA G-3′, Gapdh reverse: 5′-CTG GAA AGC TGT GGC GTG ATG G-3′, MajSat forward (20): 5′-GGC GAG AAA ACT GAA AAT CAC G-3′, MajSat reverse (20): 5′-AGG TCC TTC AGT GTG CAT TTC-3′.

### Radioactive beta-glucosyltransferase (BGT) assay

The radioactive BGT assay was performed as described previously with following exceptions (50):

Reference DNA fragments (375 bp) containing 100% hmC (except primer sites) were prepared by PCR, using a 5-hydroxymethylcytosine dNTP Mix (Zymo Research, Freiburg, Germany), and Taq DNA polymerase (Cardoso Lab, Darmstadt, Germany). As template, gDNA isolated from HEK-EBNA cells was used. Primers for PCR contained the following sequences: 5′-ATC CCA CAC CTG GCT CAG AGG G-3′ and 5′-GTC AGG GGT CAG GGA CCC ACT TGA GGA-3′. Cycles were set to: 94°C for 2 min, 40× (94°C for 15 s, 62°C for 30 s, 72°C for 40 s), 72°C for 10 min.

PCR products were purified by gel electrophoresis followed by silica column purification using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany).

Reactions contained 50 mM potassium acetate, 20 mM Tris acetate (pH 7.9), 10 mM magnesium acetate, 1 mM DTT, 2.8 μM 'cold' UDP-glucose (Sigma Aldrich, St. Louis, MO, USA), 86 nM UDP-[3H]glucose (glucose-6-3H; 60 Ci/mmol; Hartmann Analytic GmbH), 1μg DNA substrate and 75 nM recombinant β-glucosyltransferase in a

total volume of 50 μl. Reactions were incubated for 1 h at 37°C and terminated by heating at 65°C for 10 min. DNA was purified from the reaction mixture using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). Remaining radioactivity was measured using a Liquid Scintillation Analyzer Tri-Carb 2800TR (PerkinElmer, Waltham, Massachusetts, USA) with quench indicating parameter set on tSIE/AEC (transformed spectral index of the external standard/automatic efficiency control) in 2 ml of Rotiszint Eco Plus scintillation liquid (Carl Roth, Karlsruhe, Germany) in Snaptwist vials (Zinsser Analytic, Frankfurt, Germany). Samples were measured for 1 min or until the 2σ value reached 2%.

### *In vitro* oxidation and protection assay

Reference DNA fragments (375 bp) containing 100% 5mC (except primer sites) were prepared by PCR, using 5-methyl-dCTP (NEB, Frankfurt, Germany). Genomic DNA isolated from HEK-EBNA cells was used as template with primers: 5′-ATC CCA CAC CTG GCT CAG AGG G-3′ and 5′-GTC AGG GGT CAG GGA CCC ACT TGA GGA-3′, Q5® High-Fidelity DNA Polymerase (NEB, Frankfurt, Germany) and the following cycling profile: 98°C for 2 min, 40× (98°C for 15 s, 62°C for 30 s, 72°C for 60 s), 72°C for 2 min. PCR products were purified by gel electrophoresis followed by silica column purification using the QIAquick PCR purification kit (Qiagen, Hilden, Germany). For in vitro oxidation and protection assays, DNA fragments were incubated with MBD- and Tet1 proteins at 37°C in Tet oxidation buffer (10 μM $Fe(NH_4)_2(SO_4)_2.6H_2O$, 100 mM NaCl, 50 mM HEPES (pH 8), 1.2 mM adenosine triphosphate (ATP), 2.5 mM dithiothreitol (DTT), 1 mM a-ketoglutarate (aKG) and 2 mM L-ascorbic acid). Following 120 min of Tet1 incubation, the reaction was stopped by the addition of 20 μg of proteinase K at 50°C for 2 h.

### Slot blotting

gDNA samples and *in vitro* oxidation products were respectively denatured at 99°C for 10 min and placed quickly on ice for 5 min. Denatured gDNA was mixed with ice-cold 20× saline–sodium citrate (SSC) buffer at a final concentration of 4.8× SSC and blotted on a nitrocellulose membrane (Bio-Rad Laboratories, Hercules, CA, USA), which was pre-equilibrated in 20× SSC. After air-drying, the membrane was blocked with 3% milk in PBST (PBS containing 0.1% Tween) for 30 min at room temperature (RT), followed by incubation with either mouse anti 5mC (1:1000, Eurogentec, Seraing, Belgium) or rabbit anti 5hmC (1:5000, Active Motif, La Hulpe, Belgium) antibodies for 1 h at RT. The membrane was washed 3× for 10 min with PBST, before it was incubated with horseradish peroxidase (HRP)-conjugated anti mouse IgG (1:5000, GE Healthcare, Freiburg, Germany) or anti rabbit IgG (1:5000, Sigma Aldrich, St. Louis, MO, USA) antibody for 1 h at RT. After three washing steps, remaining signals were detected using Amersham ECL detection reagent (GE Healthcare, Freiburg, Germany) and imaged on a Fuji LAS-1000 imager (FUJI Film, Minato, Tokio, Japan).

### Quantification of 5hmC using a methyl sensitive restriction assay

ATTO550-labeled 42 bp-long, double-stranded oligonucleotides (GGA TGA TGA CTC TTC TGG TCmC GGA TGG TAG TTA AGT GTT GAG) (Eurofins MWG Operon, Ebersberg, Germany) containing a central methylated CpG site were diluted in Tet reaction buffer (50 mM Tris–HCl; pH 7.5, 75 μM Fe(II), 2 mM sodium ascorbate, 1 mM di-sodium-ketoglutarate) (24,35). Following incubation with purified GFP-Tet1CD and Mecp2-GFP, the reaction was heat-inactivated for 2 min at 95°C. Subsequently, oligonucleotides were digested using MspI at 37°C for 30 min. DNA was separated on a denaturing 17% polyacrylamide gel and imaged using the Typhoon TRIO Imager (GE Healthcare, Freiburg, Germany). Quantification was performed with ImageJ.

### Competitive DNA binding assay

Gel mobility shift assays (EMSA) were performed as described previously (http://www.nature.com/nmeth/journal/v2/n7/abs/nmeth0705-557.html) with following modifications. GFP-tagged MBD and cherry-tagged Tet1CD proteins were incubated with ATTO647N labeled 42 bp-long, double-stranded oligonucleotides containing a single methylated CpG dinucleotide (5′-CTC AAC AAC TAA CTA CCA TmCGG ACC AGA AGC GTC ATC ATGG -3′) in binding buffer composed of 20 mM HEPES pH 7.9, 1 mM EDTA, 3 mM $MgCl_2$, 2 mM DTT, 4% glycerol and 0.1% Triton X100 for 1.5 h at 37°C. Samples were separated on a non-denaturing 4.5% polyacrylamide gel (30%, 29:1 acrylamide:bisacrylamide), which was pre-run for 2 h at 4°C. Fluorescent signals were detected using a Storm 860 Molecular Imager (GMI, Ramsey, Minnesota, USA) and a TECAN infinite M200 plate reader (Tecan Group Ltd., Maennedorf, Switzerland), respectively.

### Immunofluorescence staining of cells

Cells were fixed for 10 min in 4% formaldehyde and permeabilized for 20 min with 0.5% Triton X-100. For detection of genomic 5hmC, endogenous Tet1 proteins, as well as NeuN, cells were incubated following formaldehyde fixation with ice-cold methanol for 5 min. After RNaseA treatment (10 μg/ml) for 30 min at 37°C, cells were washed and blocked for 30 min in 0.2% fish skin gelatin (Sigma Aldrich, St. Louis, MO, USA) at 37°C. Genomic 5hmC was detected using a rabbit anti-5hmC antibody (1:250; Active Motif, La Hulpe, Belgium) in conjunction with 25 U/ml DNaseI (Sigma Aldrich, St. Louis, MO, USA) for 70 min at 37°C. Endogenous Tet1 and NeuN proteins were detected using a rat anti Tet1 5D8 antibody (51) (1:4) and a mouse anti NeuN (1:50, Merck Millipore, Darmstadt, Germany) in conjunction with 25 U/ml DNaseI (Sigma Aldrich, St. Louis, MO, USA) for 70 min at 37°C. To stop DNaseI digestion, cells were washed with PBS containing 1 mM EDTA and 0.01% Tween. Following incubation with the secondary AMCA conjugated donkey anti rabbit IgG antibody (1:100; The Jackson Laboratory, Bar Harbor, USA), or the cy3 conjugated anti mouse IgG (1:500; The Jackson Laboratory, Bar Harbor, USA) and Alexa488 conjugated donkey

anti rat IgG antibody (1:500; The Jackson Laboratory, Bar Harbor, USA) for 45 min at RT, cells were mounted in Vectashield Medium (Vector Labs, Burlingame, CA, USA).

For immunofluorescence staining of Mecp2, fixed and permeabilized cells were blocked for 30 min in 0.2% fish skin gelatin (Sigma Aldrich, St. Louis, MO, USA). The primary rabbit anti Mecp2 antibody (32) (1:2) was applied for 1 h at RT. After three washing steps using PBST containing 0.01% Tween, the secondary donkey anti rabbit IgG cy3 (1:500, The Jackson Laboratory, Bar Harbor, USA) was applied for 45 min at RT. Following three washing steps in PBST containing 0.01% Tween, DNA was counterstained for 10 min with 1 μg/ml DAPI (Sigma, St. Louis, MO, USA), washed in PBS and mounted in Vectashield Medium (Vector Labs, Burlingame, CA, USA).

### Immunofluorescence staining of tissues

Brains of Mecp2 wild type and knockout mice were fixed for 24 h in 10% neutral-buffered formalin (Sigma, St. Louis, MO, USA) at 4°C. Fixed tissues were dehydrated (30 min 70% ethanol (Sigma, St. Louis, MO, USA), 45 min 70% ethanol, 60 min 96% ethanol, 45 min 96% ethanol, 45 min 100% ethanol, 45 min 100% ethanol, 60 min xylol (Sigma, St. Louis, MO, USA), 30 min xylol), embedded in paraffin (60 min paraffin (Carl Roth, Karlsruhe, Germany), 45 min paraffin, 60 min paraffin) and sectioned at a thickness of 6 μm. Following dewaxing in xylol (3 × 5 min) and rehydration (5 min 100% ethanol, 5 min 96% ethanol, 5 min 90% ethanol, 5 min 80% ethanol, 5 min 70% ethanol, 5 min dd$H_2O$), brain sections were incubated for 30 min at 100°C/1 bar overpressure in 10 mM sodium citrate buffer, pH 6 (Carl Roth, Karlsruhe, Germany). Sections were encircled using a hydrophobic immuno-pen (Merck Millipore, Darmstadt, Germany) and blocked for 30 min in PBS containing 4% BSA. Primary antibodies (rabbit anti 5hmC (1:1000, Active Motif, La Hulpe, Belgium), mouse anti Mecp2 8D11 (32) (1:2), mouse anti 5mC (1:100, Eurogentec, Seraing, Belgium), mouse anti NeuN (1:100, Merck Millipore, Darmstadt, Germany) and rat anti Tet1 5D8 (51) (1:2), respectively) were applied overnight in PBS supplemented with 1% BSA at 4°C. After three washing steps using PBS containing 0.1% Tween, secondary antibodies (donkey anti rabbit IgG cy3 (1:500, The Jackson Laboratory, Bar Harbor, USA), donkey anti rat IgG cy3 (1:500, The Jackson Laboratory, Bar Harbor, USA) and donkey anti mouse IgG cy3 (1:500, The Jackson Laboratory, Bar Harbor, USA)) were respectively applied for 1 h at RT. Following three washing steps in PBS containing 0.1% Tween, DNA was counterstained for 10 min with 1 μg/ml DAPI (Sigma, St. Louis, MO, USA), washed in PBS and mounted in Vectashield Medium (Vector Labs, Burlingame, CA, USA).

### Major satellite RNA FISH

For the detection of major satellite RNA transcripts, cDNA probes were amplified and labeled from genomic DNA of mouse myoblasts by PCR (major satellite fwd: 5′ AAAATGAGAAACATCCACTTG 3′, major satellite rev: 5′ CCATGATTTTCAGTTTTCTT 3′) and biotin dUTP. Brain sections were prepared as described for immunofluorescence staining of tissues. Following rehydration in water, sections were hybridized for 1 h at RT and 12 h at 4°C. After three washing steps in water, slides were incubated for 1 h at RT with Alexa-488 conjugated streptavidin (1:500, Invitrogen, Paisley PA4 9RF, UK). To remove unbound Streptavidin, slides were washed in water before DNA was counterstained for 10 min with 1 μg/ml DAPI (Sigma, St. Louis, MO, USA). Brain sections were rinsed in PBS and mounted in Vectashield Medium (Vector Labs, Burlingame, CA, USA). All reagents used for RNA FISH were supplemented with 1x ProtectRNA RNase inhibitor (Sigma, St. Louis, MO, USA). As control, equivalent slides were treated in parallel with RNase A before signal detection.

### Microscopy

Images of transiently transfected, anti 5hmC stained C2C12 mouse myoblasts were acquired using the Operetta automated imaging system with a 20× long/0.45 NA objective (PerkinElmer, UK), a xenon fiber-optic as light source, 360–400, 460–490 and 560–580 nm excitation- and 410–480, 500–550 and 590–640 emission filters, respectively. Representative images of the same cells were acquired using a Leica TCS SP5 II confocal laser scanning microscope (Leica Microsystems, Wetzlar, Germany) with a Plan-Apochromatic 100×/1.44 NA oil objective and 405, 488 and 561 nm excitation lasers.

For the analysis of Mecp2-, 5mC-, 5hmC-, NeuN- and Tet1 levels, colocalization studies of Mecp2 and 5hmC, as well as for the detection of major satellite RNA FISH signals in Mecp2 y/- and wild-type mouse brain, 3D z-stacks were acquired using the Operetta automated imaging system with a 2× long/0.08 and 20× long/0.45 NA objective (PerkinElmer, UK), a xenon fiber optic as light source, 360–400, 460–490 and 560–580 nm excitation- and 410–480, 500–550 and 590–640 emission filters, respectively.

### Image analysis and quantification

Fluorescence intensity histogram quantification of images acquired on the Operetta automated imaging system (Figures 1B and 3B) was performed using the Harmony 3.5.1 software (PerkinElmer, UK). Nuclei were detected based on Tet1CD signals and further selected pursuant to morphology properties (area and roundness). For each nucleus, mean intensities for Tet1CD- and MBD proteins (Mecp2, IDTRD, MBD, Mbd2 and Mbd3, respectively), as well as for 5hmC were calculated. After background subtraction, nuclei were binned according to (i) Tet1CD signal (subgroup) and (ii) to MBD signal (sub-subgroup). For each independent experiment, mean 5hmC level were averaged per sub-subgroup and normalized to highest 5hmC level of Tet1CD + GFP transfected cells. To automate this procedure, a routine was written in the programming language R.

For calculation of mean Mecp2-, 5mC- and 5hmC level at pericentric heterochromatin of Mecp2 y/- and wild-type mouse brain (Figure 5), mid-optical sections of the 5mC channel were used to generate chromocenter (CC) masks. Therefore, images were processed using a median filter and

thresholded in three successive steps using the basic algorithm. For the generation of binary chromocenter masks, all pixels below the final threshold were set to 0 and all pixels above the final threshold were set to 1. Total Mecp2-, 5mC- and 5hmC- signals overlapping with the chromocenter mask were calculated and divided by the total number of pixels corresponding to the area of chromocenters. To automate this procedure, a routine was written in the programming language python (http://code.google.com/p/priithon/).

To determine the accumulation of 5hmC at pericentric heterochromatin in neurons of Mecp2 y/- and wild type mouse pontes (Supplementary Figure S16D), mean 5hmC signals at chromocenters were divided by mean 5hmC signals within the nucleoplasm. Therefore, binary chromocenter masks were generated as described above. Nucleoplasm masks were prepared by subtracting a second chromocenter mask with larger surface area from a nuclear mask. For this purpose, chromocenter masks were generated as described above, except that mid-optical sections of the 5mC channel were filtered only once. Binary nuclear masks were prepared by filtering and thresholding the DAPI channel as described earlier. To further improve the nuclear mask, holes were filled using the fill holes algorithm and background was removed via the watershed algorithm. To calculate mean 5hmC signals at both, chromocenters and nucleoplasm, total 5hmC signals overlapping with either of the two masks, were divided by the total number of pixels corresponding to the area of chromocenter- and nucleoplasm masks, respectively. To automate this procedure, a routine was written in the programming language python.

Colocalization of 5hmC and chromocenters in neurons of Mecp2 y/- and wild type mouse brain (Figure 6A) was assessed by the H-coefficient (52) as previously described (53) and by line profiles generated with ImageJ (http://rsb.info.nih.gov/ij/).

RNA level of major satellite DNA in neurons of Mecp2 y/- and wild type mouse brain (Figure 6B) were calculated manually by measuring nuclear RNA FISH signals along a line through pericentric heterochromatin (50 pixel in length) using ImageJ (http://rsb.info.nih.gov/ij/) (Supplementary Figure S16A).

For quantification of mean nuclear Tet1 level in wild type and Mecp2 y/- brain (Supplementary Figure S14), binary nuclear masks were generated as described above. Total Tet1 signals overlapping with the nuclear mask were calculated and divided by the total number of pixels corresponding to the nuclear area. To automate this procedure, a routine was written in the programming language python.

### RNA-seq library preparation

Total RNA was isolated from wild type and triple Tet-knockout mouse embryonic stem cells (V6.5) in biological quadruplicates using the nucleospin triprep kit (Macherey Nagel, Düren, Germany). 50 ng RNA was reverse transcribed. cDNA was pre-amplified as described elsewhere (54). One ng of cDNA was used as input for tagmentation by the Nextera XT Sample Preparation Kit (Illumina, San Diego, CA, USA), where a second amplification round was performed for 12 cycles. For each sample, 2.5 ng of final library were pooled.

### RNA-seq and data analysis

One hundred base pairs single end reads were sequenced on an Illumina HighSeq 1500. Libraries were barcoded and mixed before sequencing. The resulting reads were mapped to the Mouse genome build mm10 using STAR version STAR 2.5.1b (55) with the specific settings: –outFilterMultimapNmax 100 –outFilterMismatchNmax 4 –winAnchorMultimapNmax → 100. The junction annotation was taken from ensembl GRCm38.75 and the index was created as recommended using the option –sjdbOverhang → 99.

The resulting bam-files were then processed using TEtranscript (56) to obtain read count tables for transcripts and transposons, using the TE annotation as provided by the authors of TEtranscript (http://labshare.cshl.edu/shares/mhammelllab/www-data/TEToolkit/TE_GTF/mm10rmskTE.gtf.gz). Normalization and differential expression analysis was done using DESeq2 (57).

### Statistical analyses

For Figures 1A and 3A, Tet1CDmut + GFP and mock were excluded from statistical tests as the mean values were at the background level. Homogeneity of variance was tested beforehand with Levene's test (using the median). The Levene's test did not indicate heterogeneous variances ($P > 0.1$) between the groups. Hence, we conducted repeated measures ANOVA for the three replicates in both experiments, which showed highly significant results ($F_{4,8} = 46.1$, $P < 0.001$ and $F_{4,8} = 22.7$, $P < 0.001$). Therefore, post-hoc pairwise *t* tests with false discovery rate correction (58) were performed.

For Figures 1B and 3B and Supplementary Figure S4, we performed Welch's ANOVA which gave a highly significant result ($F_{6,97} = 331.4$, $P < 0.001$), since the variances were heterogeneous (Levene's test: $P < 0.001$). Then, we performed post-hoc pairwise *t* tests with non-pooled standard deviations and false discovery rate correction.

For Figure 3B (lower row), since the variances were not significantly different (Levene's test: $P = 0.19$), we compared the means with a *t* test.

For Figures 2, 3C and 6E, we performed independent two-sample Student's *t*-test.

For Figures 5 and 6A and 6C, as well as Supplementary Figures S14 and S16, we performed a Wilcoxon signed-rank test.

All statistical tests were conducted with R (https://www.r-project.org/).

## RESULTS AND DISCUSSION

### Mecp2 and Mbd2 protect 5-methylcytosine from Tet1 mediated oxidation in a concentration dependent manner

Considering that MBD and Tet proteins share a common substrate, we aimed at clarifying whether binding of MBD proteins to methylated DNA protects epigenetically

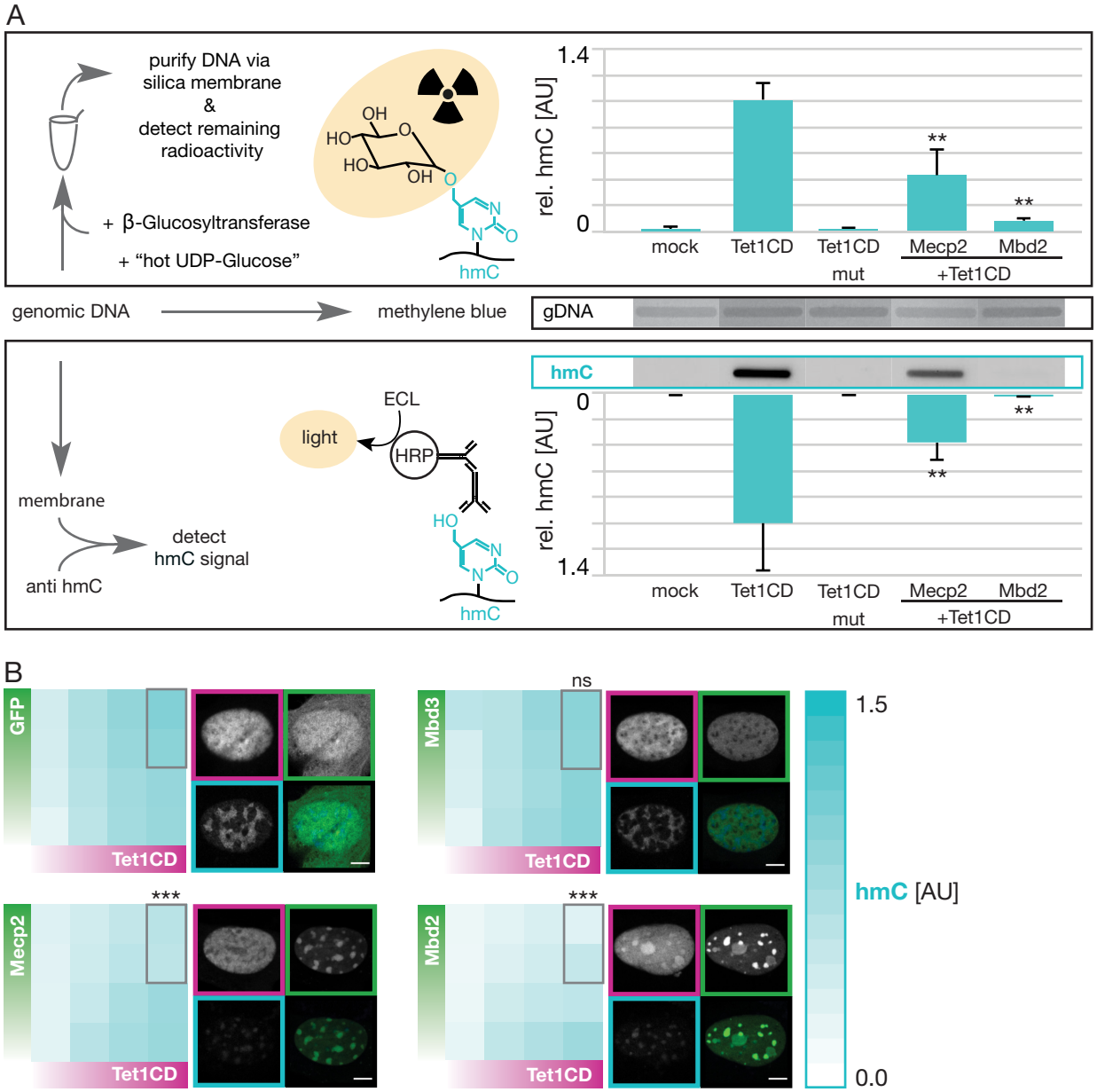



**Figure 1.** Impact of methyl-CpG binding domain proteins on the efficiency of Tet1 mediated 5mC oxidation. (**A**) Radioactive (top) and immunological (bottom) assay to determine 5hmC levels in genomic DNA (gDNA) of transiently transfected HEK cells. Schemes (left) illustrate the workflow and mode of 5hmC detection. Histograms (right) show relative 5hmC levels of three independent experiments +SD. Tet1CD + Mecp2 and Tet1CD + Mbd2 differed significantly (**$P < 0.01$; post-hoc pairwise $t$ test) from Tet1CD (see methods for details). gDNA quantities were monitored by methylene blue staining. Tet1CD corresponds to Tet1 catalytic domain and Tet1CDmut is the catalytic domain of Tet1 containing two point mutations that abolish binding of the co-factor $Fe^{2+}$ (see also Supplementary Figure S2). Full blots are shown in Supplementary Figure S1. (**B**) *In situ* staining and quantification of genomic 5hmC levels in transiently transfected C2C12 mouse myoblasts. Images were acquired on an automated high throughput imaging system with a 20×, 0.45 NA objective. Gradient heat maps show relative 5hmC ( = cyan) signals as a function of increasing Tet1CD ( = magenta) and MBD ( = green) protein expression levels depicted by the green and magenta gradient bars. Shown are mean values of five (Tet1CD + GFP, $n = 12\,798$), four (Tet1CD + Mbd2, $n = 2598$) or two (Tet1CD + Mecp2, $n = 4760$; Tet1CD + Mbd3, $n = 6449$) independent experiments, respectively. For statistical tests, 5hmC signals of cells with high Tet1CD and high Mecp2/Mbd2/Mbd3 protein levels (framed in grey) were used. Tet1CD + Mecp2 and Tet1CD + Mbd2 differed highly significantly (***$P < 0.001$; post-hoc pairwise $t$ test) from Tet1CD + GFP. No significant difference was detected for Tet1CD + Mbd3 ($P = 0.34$). Confocal images of mid optical sections of the same samples represent transiently transfected C2C12 mouse myoblasts (Tet1CD = magenta; GFP/MBD proteins = green) immunostained for 5hmC (5hmC = cyan). Scale bar, 5 μm.

silenced regions from Tet- mediated DNA demethylation. To this end, we either radioactively (Figure 1A, top), or immunologically (Figure 1A, bottom and Supplementary Figure S1) labeled and subsequently quantified global 5hmC levels in genomic DNA of FACS sorted HEK cells expressing comparable levels of the catalytic active (Tet1CD) and inactive domain (Tet1CDmut) of Tet1 alone, or in combination with Mecp2 and Mbd2, respectively (Figure 1A and Supplementary Figure S2). When compared to mock and Tet1CDmut transfected cells, both, the radioactive (Figure 1A, top), as well as the immunological (Figure 1A, bottom) assay revealed increased 5hmC levels in genomic DNA of cells expressing the catalytic active domain of Tet1 alone. Coexpression of Mecp2 or Mbd2, however, significantly decreased global 5hmC levels by at least 50%, demonstrating reduced Tet1 effectiveness in the presence of substrate-competitive proteins.

Further single-cell analysis (Supplementary Figure S3) of transiently transfected mouse myoblasts (Figure 1B) and HEK cells (Supplementary Figure S4) immunostained for 5hmC revealed a correlation between Tet1CD protein and 5hmC levels in a subpopulation of cells containing low Mecp2 (Figure 1B, bottom, left and Supplementary Figure S4, bottom, left) and Mbd2 (Figure 1B, bottom, right and Supplementary Figure S4, top, right) protein amounts, respectively. The remainder cells of the population, characterized by high expression levels of the Mbd2 and Mecp2 proteins, in contrast, showed no longer any correlation between Tet1CD protein levels and the occurrence of its oxidation product. Instead, 5hmC levels anti-correlated with increasing levels of Mecp2 (Figure 1B, bottom, left and Supplementary Figure S4, bottom, left) and Mbd2 (Figure 1B, bottom, right and Supplementary Figure S4, top, right), respectively, indicating that protection of 5mC from Tet1 catalyzed oxidation highly depends on the MBD protein concentration. In contrast to Mecp2 and Mbd2, even the highest expression levels of GFP (Figure 1B, top, left and Supplementary Figure S4, top, left) and Mbd3 (Figure 1B, top, right) proved insufficient to repress Tet1 activity. As both proteins are not capable of binding to (methylated) DNA (for Mbd3, see Supplementary Figure S5A), this suggests that direct interaction with (methylated) DNA is a prerequisite for the effective conservation of 5mC.

To determine, whether the levels of mCherry-Tet1CD obtained through overexpression in mouse myoblasts and human embryonic kidney cells are within the physiological range of endogenous Tet1 in primary mouse neurons and mouse embryonic stem cells (ESCs), we stained all of the four cell types for Tet1 and quantified the resulting immunofluorescent signals (Supplementary Figure S6). To allow a direct comparison to the Tet1 expression levels plotted in Figure 1B and Supplementary Figure S4, transiently transfected mouse myoblasts and HEK cells were binned according to the ectopic Tet1CD signal (e.g. group 1 of Supplementary Figure S6 corresponds to the first column of all heatmaps in Figure 1B and Supplementary Figure S4, respectively). We found that bins 1 and 2 of mouse myoblasts, as well as bin 1 of HEK cells express (combined ectopic+endogenous) Tet1 levels comparable to mouse ESCs. As shown previously, the level of overexpressed Mecp2 in mouse myoblasts is in the range of endogenous physiological Mecp2 levels per mouse neuronal cell nucleus (36,59). Accordingly, cells expressing low Tet1CD protein levels do not create an artificial phenotype and, thus, reflect the situation *in vivo*.

Finally, we immunologically quantified 5hmC levels in genomic DNA of Mecp2 lox/y and -/y mouse tail fibroblasts (MTF). Compared to the floxed control group (lox/y), we detected increased 5hmC levels in the corresponding Mecp2 knockout cell line (-/y), indicating that Mecp2 represses the formation of 5hmC *in vivo* (Supplementary Figure S7).

### Prior binding of Mecp2 and Mbd2 to 5-methylcytosine enhances blocking of Tet1 catalyzed 5-hydroxymethylcytosine formation *in vitro*

As described above, Tet1 mediated 5hmC formation is impaired by Mecp2 and Mbd2 *in vivo*. To gain a deeper understanding of the protective mechanism, we next sought to determine, whether the chronological order of DNA binding by MBD and Tet1 proteins would influence the extent of 5mC protection. Since at the cellular level the chronological access of MBD- and Tet1CD proteins to DNA is difficult to control, we further investigated its influence on 5mC protection in *in vitro* experiments. Therefore, various conceivable binding scenarios were systematically mimicked on a molecular level. Briefly, same molar ratios of Tet1CD and MBD proteins were incubated simultaneously or consecutively with a PCR fragment containing multiple methylated cytosines. Following 2 h of Tet1CD incubation, DNA was blotted on a membrane to then immunologically detect the amount of remaining unoxidized 5mC, as a measure of 5mC protection by MBD proteins (Figure 2A and Supplementary Figure S8).

Altogether, 5mC levels were comparatively high in any of the samples containing, in addition to Tet1CD, Mecp2 and Mbd2 (Figure 2B and C), respectively, indicating restricted Tet1CD activity in the presence of 5mC binding proteins, which is in accordance with our *in vivo* data (Figure 1). In more detail, when compared to the fully unprotected control group (DNA + Tet1CD) (Figure 2B and C, first column), 5mC levels were increased by a factor of 1.9 (Mecp2) and 1.6 (Mbd2) in samples allowing simultaneous access of Tet1CD and MBD proteins to their common substrate 5mC (Figure 2B and C, second column). Incubation of MBD proteins with the methylated DNA prior to the addition of Tet1CD enzymes, yielded 5mC signals 3.9 (Mecp2) and 2.1 (Mbd2) fold higher relative to control samples without MBD proteins (Figure 2B and C, third column). Delayed addition of methylated DNA to pre-incubated Tet1CD and MBD proteins resulted in relative 5mC levels of 2.1 (Mecp2) and 1.3 (Mbd2) (Figure 2B and C, fourth column). Accordingly, among all tested conditions, early incubation of MBD proteins with methylated DNA before the addition of Tet1CD enzymes revealed the highest 5mC signals and, thus, the best possible protection against Tet1CD catalyzed oxidation (Figure 2B and C, third column).

As Mecp2 can bind to a single methylated CpG site (mCpG), we further tested whether the protection against Tet oxidation could take place on single mCpG containing substrates. Therefore, we measured the degree of protection

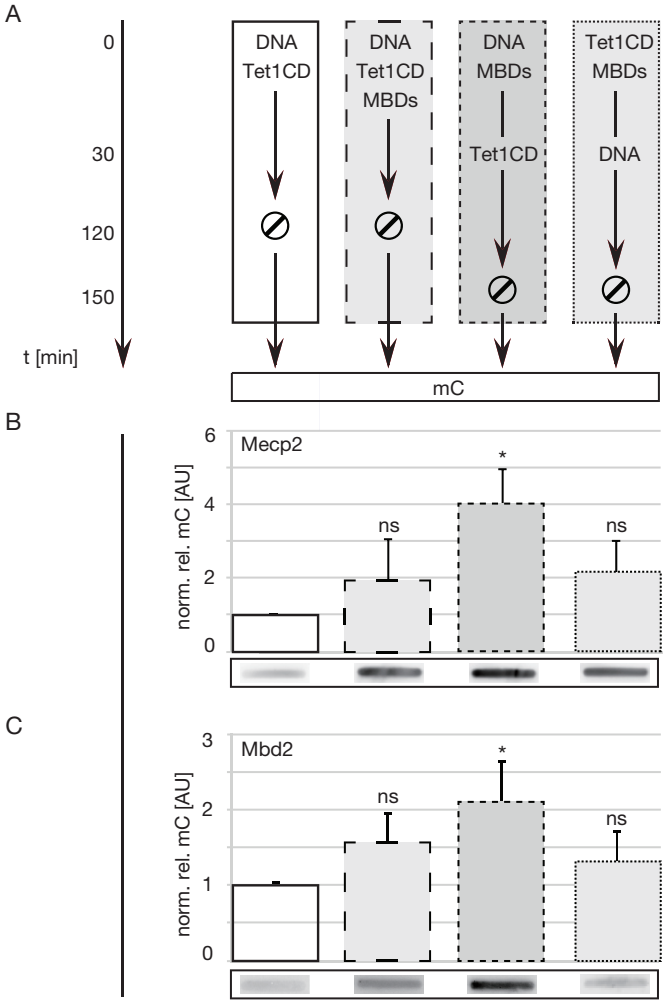



**Figure 2.** Influence of the chronological DNA binding order on the protecting ability of MBD proteins. Quantification of remaining 5mC levels in 20 pmol of double-stranded DNA containing multiple 5mC nucleotides after simultaneous or successive incubation with 20 pmol of Tet1CD (catalytic domain) and 20 pmol of MBD proteins by slot blot. (**A**) Experimental setup illustrating the incubation order and time of proteins and oligos prior slot blotting. (**B** and **C**) Histograms show relative 5mC levels of 5mC containing PCR product after incubation with (**B**) Tet1CD and Mecp2 ($n = 4$) and (**C**) Tet1CD and Mbd2 ($n = 4$). Shown are mean values + SD. Significant differences were detected for pre-incubation of Mecp2 and Mbd2 with DNA (*$P < 0.05$; ns = non significant; post-hoc pairwise $t$ tests). Full blots are shown in Supplementary Figure S8.

over time using a methyl sensitive restriction endonuclease assay (35). Consistent with our previous results, single 5mC containing, double-stranded oligonucleotides that were pre-incubated with Mecp2 lost the least amount of 5mC. Even after 2.5-h treatment with Tet1CD, the presence of Mecp2 (previously incubated with DNA) protected 80% of 5mC from oxidation versus only 20% of 5mC surviving in the absence of Mecp2 (Supplementary Figure S9).

On this basis, we conclude that binding of MBD proteins to DNA, especially prior binding provides the greatest contribution towards preserving the methylation status of CpG dinucleotides. Alternatively, protein-protein interactions, which could have formed most effectively by pre-incubation of Tet1CD and MBD proteins, do not play a role in protecting DNA from Tet1CD driven oxidation.

### Direct binding to DNA is sufficient to effectively prevent 5mC oxidation by Tet1

As indicated earlier, Mecp2 contains various different interaction sites for DNA. While the IDTRD domain of Mecp2 was shown to bind both, methylated and unmethylated DNA with similar affinity (9,10,15), the MBD domain of Mecp2 has a preference for methylated CpG dinucleotides (14,60).

To test whether and which of the above-mentioned binding mode is responsible for the conservation of 5mC, we tested the protecting ability of both Mecp2 subdomains using the battery of assays employed before (Figure 1 and Figure 2). Quantification of 5hmC levels in genomic DNA of human HEK cells revealed that both subdomains of Mecp2 avert Tet1CD mediated 5mC oxidation to a similar extent, indicating that the adverse impact on Tet1CD activity does not directly correlate with 5mC affinity (Figure 3A). To further test this conclusion, we measured the protective effect of full length Mecp2 proteins carrying an R111G mutation in their MBD domain. While mutation of arginine 111 abolishes binding to 5mC *in vitro* and to methylated heterochromatin *in vivo* (61), the mutant protein is still able to interact with unmethylated DNA in a sequence unspecific manner. Accordingly, it shifts 5mC containing DNA in the absence, but less efficiently in the presence of poly dI:dC competitor DNA (Supplementary Figure S5B). As cells expressing this mutant Mecp2 variant also exhibited low 5hmC levels (Supplementary Figure S10A left), we deduce that sequence unspecific DNA interactions, considerably contribute to defending 5mC from Tet1CD mediated oxidation. Additional 5mC recognition by a functional MBD, as demonstrated by wild type Mecp2, improved the protecting ability only marginally (Supplementary Figure S10A left). Indeed, we found that two proteins specifically binding major satellite DNA sequences independent of methylation, the polydactyl zinc finger MaSat (62), as well as the transcription activator-like effector protein msTALE (63), repressed the formation of 5hmC *in situ* (Supplementary Figure S10A, right). Thus, 5mC recognition by the MBD is unlikely to *per se* play a major role in the protective mechanism.

To validate and extend these results, we immunostained and quantified genomic 5hmC levels in single C2C12 mouse myoblasts (Figure 3B) and HEK cells (Supplementary Figure S4). Similar to C2C12 and HEK cells coexpressing Mecp2 and Mbd2 (Figure 1B and Supplementary Figure S4), genomic 5hmC content correlated with Tet1CD protein levels in a subpopulation of cells containing low IDTRD protein amounts (Figure 3B, top, left and Supplementary Figure S4, bottom, middle; rows 3 and 4 from top to bottom). Cells, characterized by high expression levels of IDTRD proteins, in contrast, showed no longer any correlation between Tet1CD protein levels and its oxidation product. Instead, 5hmC levels anti-correlated with increasing levels of IDTRD (Figure 3B, top, left and Supplementary Figure S4, bottom, middle; rows 1 and 2 from top to bottom). In C2C12 cells, the MBD domain of Mecp2 repressed merely the catalytic activity of a small number of

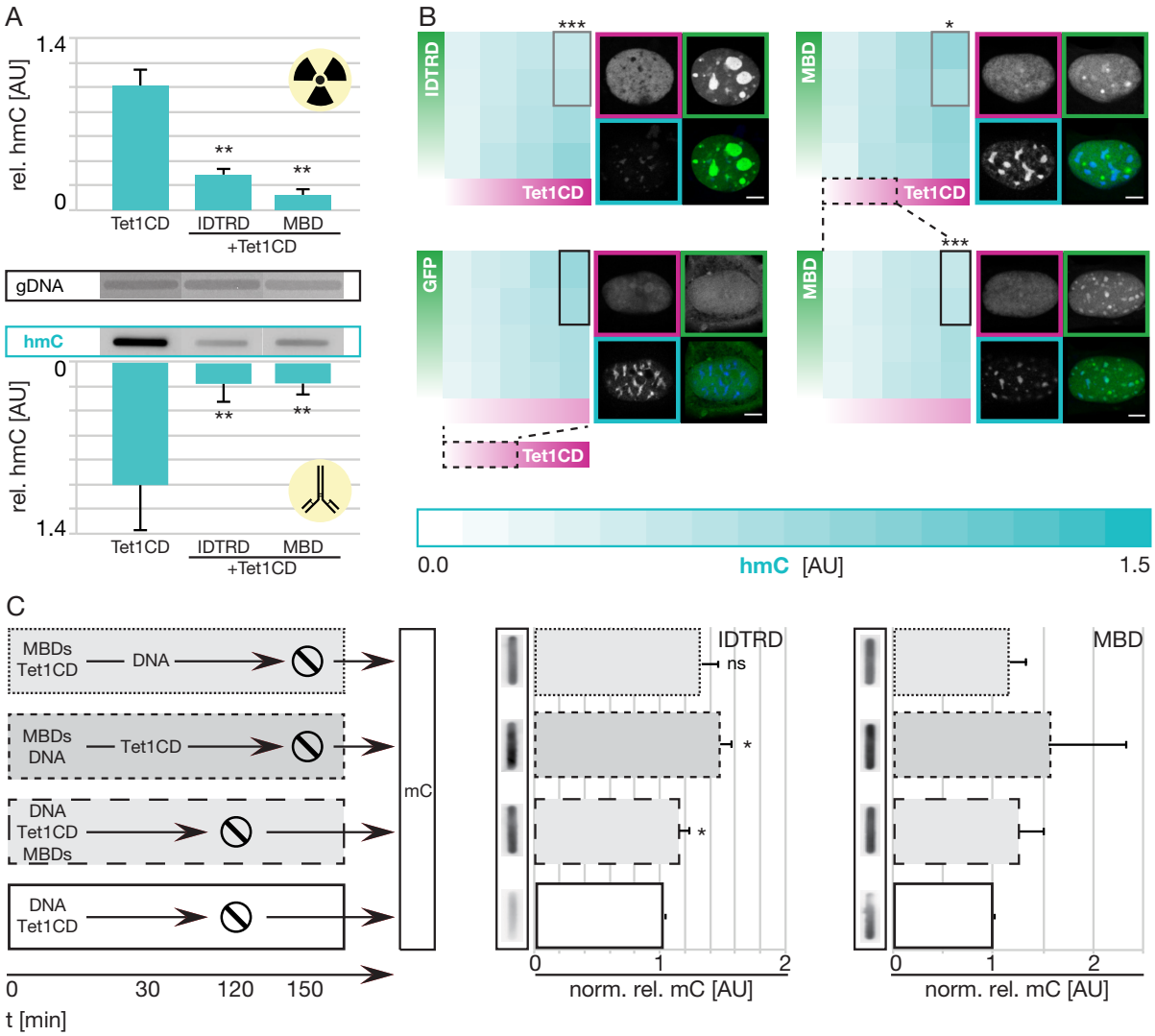



**Figure 3.** Effect of different DNA binding modes on Tet1 activity. (**A**) Radioactive (top) and immunological (bottom) assay to determine 5hmC levels in genomic DNA (gDNA) of transiently transfected HEK cells. Histograms show relative 5hmC levels of three independent experiments + SD. Tet1CD + IDTRD and Tet1CD + MBD differed significantly (**$P < 0.01$; post-hoc pairwise *t* tests) from Tet1CD + GFP. gDNA quantities were monitored by methylene blue staining. Tet1CD corresponds to Tet1 catalytic domain and IDTRD and MBD correspond to the subdomains of Mecp2 (Supplementary Figure S2). Full blots are shown in Supplementary Figure S8. (**B**) *In situ* staining and quantification of genomic 5hmC levels in transiently transfected C2C12 mouse myoblasts. Images were acquired on an automated high throughput imaging system with a 20×, 0.45 NA objective. Gradient heat maps show relative 5hmC ( = cyan) signals as a function of Tet1CD ( = magenta) and Mecp2 subdomain ( = green) levels depicted by the green and magenta gradient bars. Shown are mean values of two (Tet1CD+IDTRD, $n = 6495$; Tet1CD + MBD, $n = 1800$) independent experiments. (**B, upper row**) For statistical tests, 5hmC signals of cells with high Tet1CD and high IDTRD/MBD protein levels (framed in gray) were used. Tet1CD + IDTRD differed highly significantly (***$P < 0.001$; post-hoc pairwise *t* tests) from Tet1CD + GFP (see Figure 1B). Weakly significant difference was detected for Tet1CD + MBD (*$P < 0.05$; post-hoc pairwise *t* tests). (**B, lower row**). For statistical tests, 5hmC signals of cells with low Tet1CD and high MBD protein levels were used. From these, we selected the values framed in black. Highly significant differences were detected between both groups (***$P < 0.001$; *t* test). Confocal images of mid optical sections of the same samples represent transiently transfected C2C12 mouse myoblasts (Tet1CD = magenta; GFP/MBD proteins = green) immunostained for 5hmC (5hmC = cyan). Scale bar, 5 μm. (**C**) Quantification of remaining 5mC levels in double-stranded methylated DNA after simultaneous or successive incubation with Tet1CD proteins and Mecp2 subdomains by slot blot. Scheme illustrates the incubation order and time of proteins and methylated DNA prior slot blotting. Histograms show relative 5mC signals after incubation with Tet1CD and IDTRD ($n = 4$), as well as Tet1CD and MBD ($n = 5$), respectively. Shown are mean values + SD. Significant differences were detected for pre-incubation of IDTRD with DNA and for simultaneous incubation of IDTRD, Tet1CD and DNA (*$P < 0.05$; ns = non significant; post-hoc pairwise *t* tests). For MBD, a similar trend was detected, however, without statistical significance.

Tet1CD molecules (Figure 3B, top, right (columns 1 and 2 from left to right) and Figure 3B, bottom). In the presence of high Tet1CD protein amounts, even the highest MBD protein concentrations failed to repress the formation of 5hmC by Tet1CD (Figure 3B, top, right (columns 3 and 4 from left to right)). In HEK cells, which expressed ectopic proteins at a substantially higher level per cell than the previously analyzed C2C12 cells (Supplementary Figure S11 and Supplementary Figure S6), however, MBD protein levels were sufficient to avert the catalytic activity of low to high Tet1CD protein levels (Supplementary Figure S4, bottom, right). Accordingly, we conclude that the extent of 5mC protection substantially depends on the concentration of MBD and IDTRD molecules per cell as it determines the coverage of DNA in a sequence-unspecific manner (see also Figure 4).

Similar to the MBD domain of Mecp2, Mbd2 has been shown to preferentially bind 5mC (64,65). In contrast, though, Mbd2 was more efficient than the MBD in protecting 5mC from oxidation (Figures 1 and 3). To test whether binding kinetics *in vivo* may contribute to 5mC protection, we performed fluorescence recovery after photobleaching (FRAP) experiments. Whereas the MBD showed fast recovery at pericentric heterochromatin with halftimes of 2 s, Mbd2 recovered 15-fold slower after photobleaching (30 s) (Supplementary Figure S10B and S10C). Hence, we propose that long retention times of Mbd2 at methylated cytosines improve the efficiency of 5mC protection. However, since the IDTRD subdomain depicted similar fast recovery kinetics (2 s) like the MBD domain (Supplementary Figure S10B and 10C) but was more efficient in protecting 5mC, we deduce that additional sequence-unspecific DNA binding parameters (e.g. stoichiometry) must play a role.

Finally, *in vitro* 5mC oxidation studies using a PCR fragment containing multiple methylated cytosines showed that similar to Mecp2 and Mbd2, prior binding of MBD and IDTRD to DNA additionally strengthens the conservation of 5mC (Figure 3C and Supplementary Figure S8).

In summary, these data highlight the complexity of the MBD based 5mC protection mechanism, which achieves best performance through prior and long lasting coverage of DNA in a sequence-unspecific manner. It also differs from previous reports (66) suggesting that 5mC binding *per se* is critical to protect from oxidation.

### Binding of Mecp2 to DNA impairs the DNA binding ability of Tet1CD *in vitro*

Tet-mediated oxidation of 5mC was recently described as a complex, multistep process, initiated by the binding of Tet proteins to DNA via hydrophobic interactions, followed by recognition of 5mC in CpG context, base flipping and oxidation (67,68).

To investigate how prior binding of MBD proteins to DNA protects 5mC from Tet catalyzed oxidation (Figures 1-3), we next analyzed which of the above-mentioned step is affected. To this end, we tested the DNA binding ability of Tet proteins, considered as the first step towards 5mC oxidation, in the presence of low and high concentrations of 5mC specific (MBD, Figure 4B and Supplementary Figure S12B), as well as sequence-unspecific (IDTRD, Figure 4C and Supplementary Figure S12B) Mecp2 DNA binding domains by electrophoretic mobility shift assays (EMSAs) (Figure 4 and Supplementary Figure S12A).

To validate that both Mecp2 subdomains bind to DNA under the given reaction conditions, we initially verified their efficiency to form complexes with short double-stranded (ds), single 5mC-containing DNA in the absence of Tet1CD proteins (Figure 4A–C, first and second row). While incubation of DNA with low substoichiometric MBD (Figure 4B, first row) and IDTRD (Figure 4C, first row) protein concentrations, resulted in a single, slow migrating band, increasing protein amounts (Figure 4B-C, second row) gave rise to an additional high molecular supershift, originating from additive accumulation of proteins to an already bound DNA molecule. Hence, our data prove the suitability of the present reaction conditions. Furthermore, it indicates that at high protein levels, where most of the unbound DNA substrate is depleted, multiple binding of either Mecp2 subdomain to the same DNA fragment is promoted (Figure 4A–C, compare first and second rows). As the ratio of high molecular weight versus low molecular weight shifted DNA in the IDTRD is higher than with the MBD, we conclude that the IDTRD of Mecp2 is more efficient in fully covering DNA molecules than the MBD (Figure 4B and C, second row).

Addition of Tet1CD molecules (Supplementary Figure S12B) to single 5mC containing dsDNA, which was pre-incubated with low protein amounts of the MBD or IDTRD, respectively, resulted in two discrete prominent DNA shifts (Figure 4A–C, third row). While the fast migrating DNA co-localized with MBD (Figure 4B, third row) and IDTRD (Figure 4C, third row) protein signals, respectively, the high molecular DNA band coincided with protein signals for the catalytic domain of Tet1 (Figure 4B and C, third row). Consequently, our data indicate that in the presence of low competitive protein concentrations and excess availability of uncovered DNA substrate, Tet1CD binds to DNA without compromising efficiency (Figure 4B and C, third row). Pre-incubation of dsDNA with a higher number of IDTRD molecules, in contrast, greatly diminished Tet1CD signals (arrowhead), which instead strongly colocalized with signals derived from IDTRD proteins (arrow, Figure 4C, fourth row). Thus, we conclude that under the present reaction conditions, under which most of the DNA substrate is covered by IDTRD molecules (Figure 4C, second row), binding of Tet1 proteins to their common substrate is almost entirely averted. Similar to the IDTRD, however, as a result of lower DNA coverage, less significant, higher protein level of MBD (arrow) reduced the formation of Tet1-DNA complexes (arrowhead) (Figure 4B, fourth row). Similar results were obtained with equimolar amounts of MBD/IDTRD and Tet1CD proteins (Supplementary Figure S13).

In summary, these data demonstrate that the amount of free Tet1CD enzyme highly correlates with the number of bound methyl-CpG binding domain proteins per DNA molecule. Hence, we conclude that binding of MBD proteins to DNA protects 5mC from oxidation by restricting access of Tet1CD enzymes to DNA, whereby any further steps of the oxidation procedure are inhibited. Besides this, we deduce that the efficiency of Tet DNA binding inhibi-

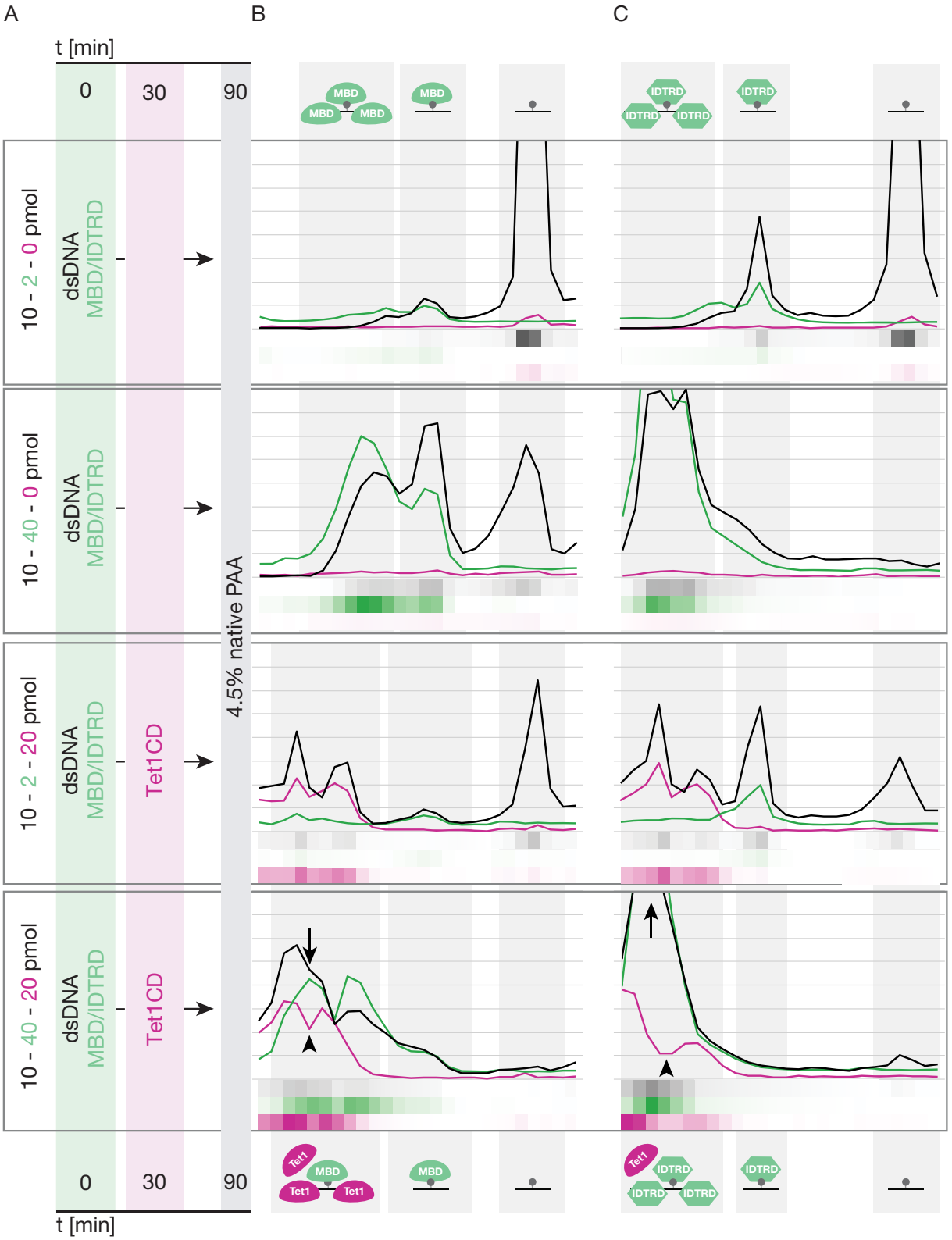



**Figure 4.** Impact of 5mC-specific and sequence-unspecific DNA binding proteins on the DNA binding ability of Tet1CD proteins. Electrophoretic mobility shift assay (EMSA) to determine the binding ability of fluorescently tagged Tet1CD to double-stranded, single mC containing DNA (ATTO647 labeled) in the presence of low and high amounts of (**B**) 5mC specific (fluorescently tagged MBD) and (**C**) sequence-unspecific (fluorescently tagged IDTRD) DNA binding domain proteins, respectively. (**A**) Experimental setup illustrating the amount, as well as the incubation order and time of proteins and DNA prior EMSA. (**B**) Separation of MBD-dsDNA (top, $n = 3$), MBD-Tet1CD-dsDNA (bottom, $n = 3$), as well as (**C**) IDTRD-dsDNA (top, $n = 3$) and IDTRD-Tet1CD-dsDNA (bottom, $n = 3$) complexes via electrophoresis through a native polyacrylamide gel visualized using a fluorescent plate reader (see also Supplementary Figure S12). Running direction is from left (– pole) to right (+ pole).

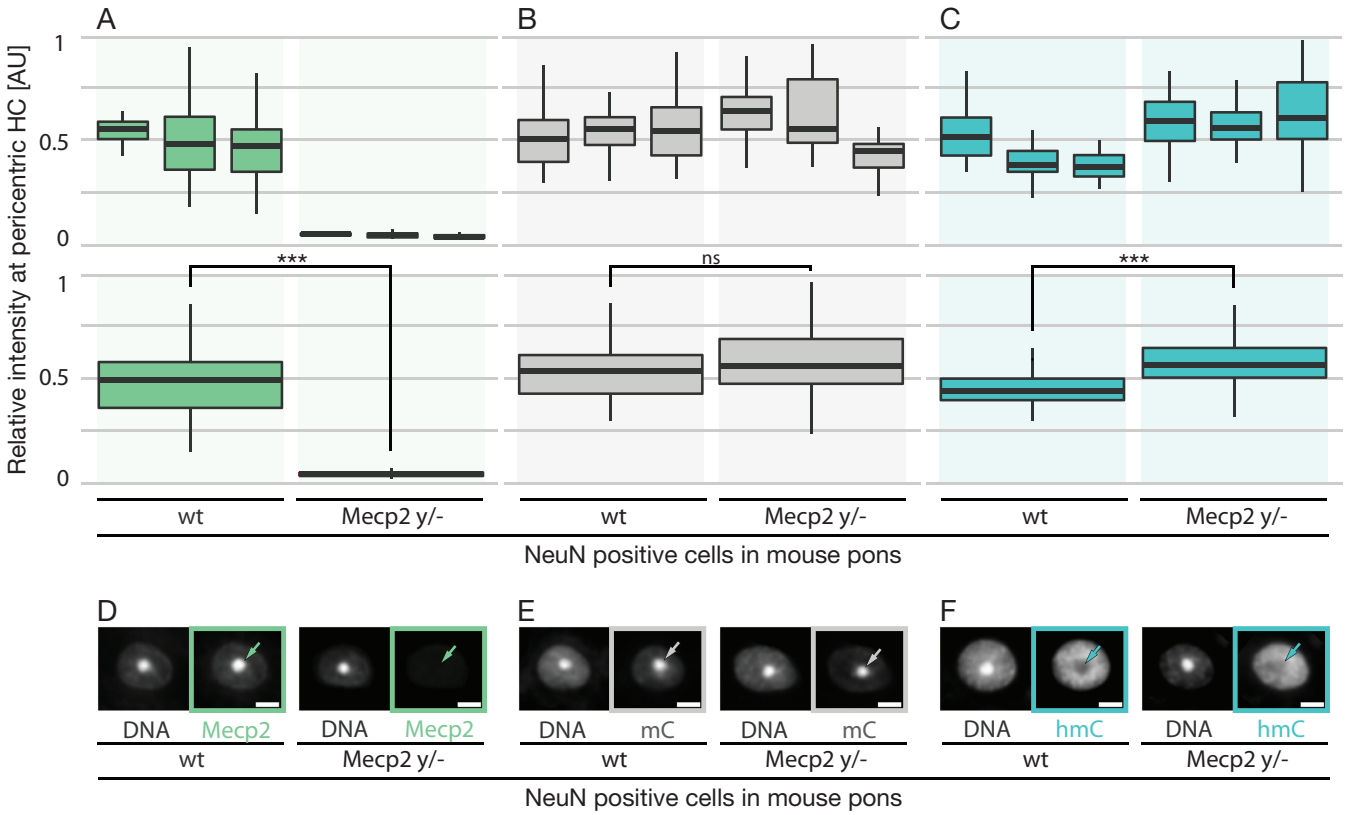



**Figure 5.** Epigenetic composition of pericentric heterochromatin in a mouse model for Rett syndrome. Immunostaining and quantification of Mecp2, 5mC and 5hmC levels in NeuN positive cells of wild type and Mecp2 knockout (y/-) mouse pontes, respectively. (**A–C**) Box plots represent the distribution of (**A**) Mecp2 (wt1 $n = 18$, wt2 $n = 38$, wt3 $n = 47$, KO1 $n = 18$, KO2 $n = 39$, KO3 $n = 31$; from left to right), (**B**) 5mC (wt1 $n = 42$, wt2 $n = 39$, wt3 $n = 44$, KO1 $n = 34$, KO2 $n = 29$, KO3 $n = 20$; from left to right) and (**C**) 5hmC (wt1 $n = 42$, wt2 $n = 39$, wt3 $n = 44$, KO1 $n = 34$, KO2 $n = 29$, KO3 $n = 20$; from left to right) levels at pericentric heterochromatin (HC) in neurons of three individual (top) and combined (bottom) wild type and Mecp2 y/- mouse pontes, respectively ($n$= number of cells). Plotted is the median, as well as the first and third quartiles. Whiskers extend to 1.5 times the interquartile range. $P$ values were calculated by Wilcoxon signed-rank test (***$P < 0.001$, ns = non significant). (**D–F**) Mid-confocal optical sections of NeuN positive cells of wild type and Mecp2 y/- mouse pontes immunostained for (**D**) Mecp2, (**E**) 5mC and (**F**) 5hmC, respectively. Arrows point to pericentric heterochromatin. DNA was counterstained with DAPI. Scale bar, 5μm.

tion, which is proportional to DNA coverage by MBD proteins, may differ from cell type to cell type, as the binding mode of MBD proteins is strongly affected by DNA methylation density and binding partners (14).

### The Tet oxidation product 5hmC is enriched in neurons of a mouse model for Rett syndrome

As we found that Mecp2 represses Tet1-mediated 5mC oxidation *in vivo* (Figure 1) and *in vitro* (Figure 2), we next tested whether the previously reported transcriptional increase of repetitive elements in Mecp2 knockout brain (20), may be considered as pathophysiological consequence of unconfined Tet activity. To address this hypothesis, we analyzed genomic 5mC- and 5hmC levels in the pons (Supplementary Figures S14A and S15) of a mouse model for Rett syndrome ($Mecp2^{-/ytm1.1Bird}$), which was previously identified as brain region partially responsible for the devastating breathing disturbances observed in Rett patients (69). Since in wild type brain Mecp2 was primarily found at pericentric heterochromatin of neurons (Figure 5A and Supplementary Figure S14B and C), we consequently focused our analysis to these chromatin regions, which in mouse cells assemble into higher order aggregates known as chromocenters (CC) (Supplementary Figure S15).

While knockout of Mecp2 had little effect on the distribution of pericentric 5mC levels (Figure 5B), the amount of the Tet oxidation product 5hmC was significantly increased at chromocenters of Mecp2 deficient neurons of the pons (Figure 5C), which is in agreement with previous data (66). Using LC–MS it has been shown in different brain regions, that 4.5% of all cytosines are methylated and 5mC levels do not change between regions. Moreover, 5hmC levels were shown to vary between 0.3% and 0.6%, with an average 0.45%, i.e. 10 times lower than 5mC (70). Using similar methods, Wu and colleagues showed that the distribution of 5hmC, but not of 5mC, varies between tissues. They, furthermore, showed that in Tet1 knockdown ES cells, although the 5hmC decreased to less than half of the control cells, 5mC did not change (71), which is similar to our results. As we measured a change of 40% for 5hmC in Mecp2 knockout relative to wild type neurons, we would expect maximally 4% change of 5mC levels in a pure population of neuronal cells. According to Münzel *et al.* (70) the maximally expected change of 4% lies within the experimental error rate (±5% for 5mC and 5hmC) and is, therefore, most

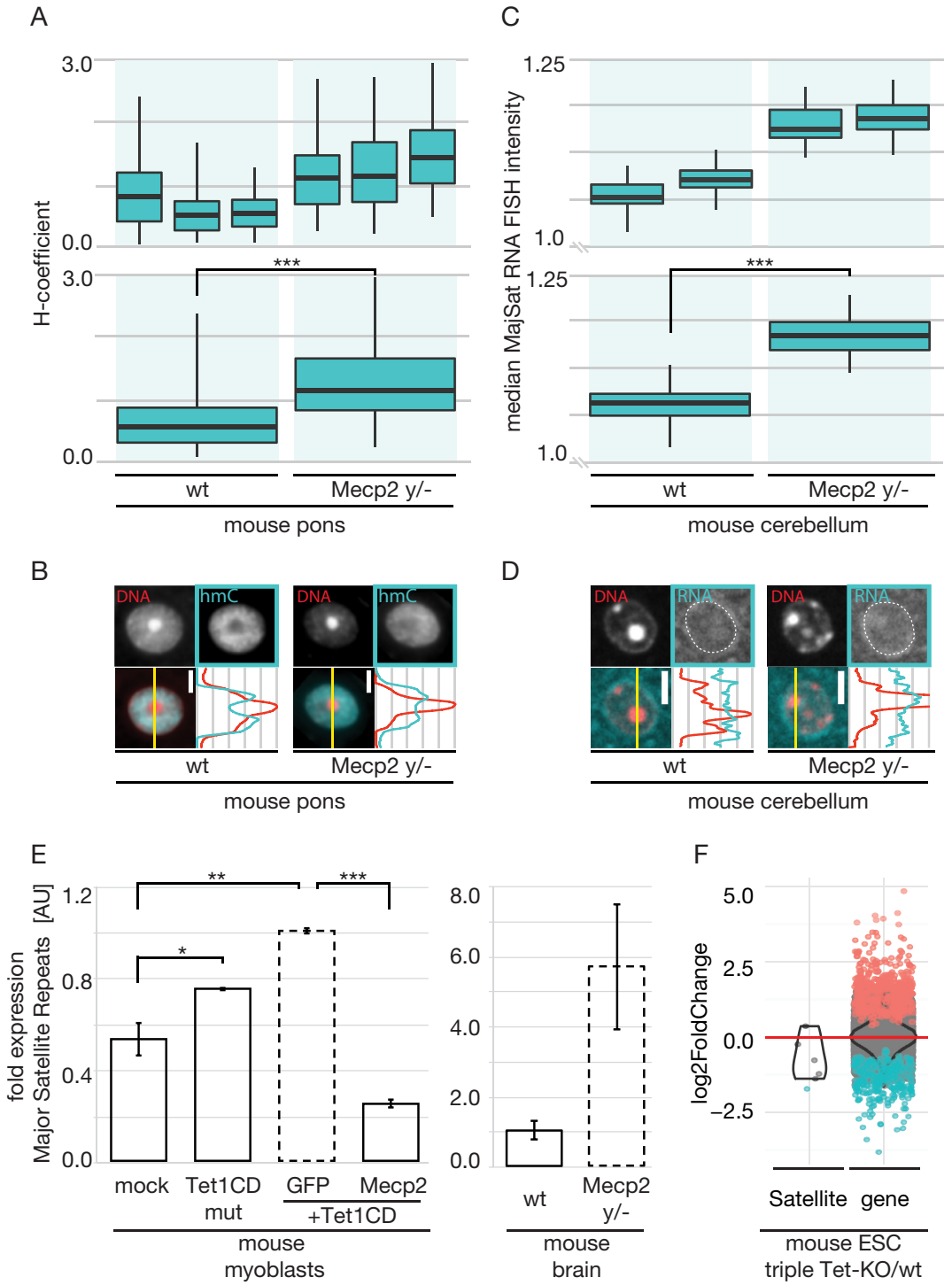



**Figure 6.** Correlation of subnuclear 5hmC distribution and major satellite expression. (**A** and **B**) Immunostaining and colocalization analysis of 5hmC with pericentric heterochromatin. (**A**) Box plots show the median 5hmC colocalization with pericentric heterochromatin in neurons of three individual (top) and combined (bottom) wild type and Mecp2 y/- mouse pontes (wt1 $n = 42$, wt2 $n = 39$, wt3 $n = 44$, KO1 $n = 34$, KO2 $n = 29$, KO3 $n = 20$; from left to right; $n$= number of cells), as well as the first and third quartiles. Whiskers extend to 1.5 times the interquartile range. $P$ values were calculated by Wilcoxon signed-rank test (***$P < 0.001$). (**B**) Line intensity plots of DNA (red) and 5hmC (cyan) distribution through pericentric heterochromatin in neurons of wild type and Mecp2 y/- mouse pontes, respectively. DNA was counterstained with DAPI. Scale bar, 5μm. (**C** and **D**) Detection and quantification of major satellite transcripts by Fluorescent In Situ Hybridization (FISH). (**C**) Box plots show the median major satellite RNA FISH signal in neurons of two individual (top) and combined (bottom) wild type and Mecp2 y/- mouse cerebella (wt1 $n = 29$, wt2 $n = 26$, KO1 $n = 30$, KO2 $n = 30$; from left to right; $n$= number of cells), as well as the first and third quartiles. Whiskers extend to 1.5 times the interquartile range. $P$ values were calculated by Wilcoxon signed-rank test (***$P < 0.001$). (**D**) Line intensity plots of DNA (red) and major satellite RNA (cyan) distribution through pericentric heterochromatin in neurons of wild type and Mecp2 y/- mouse pontes, respectively. DNA was counterstained with DAPI. Scale bar, 5μm. (**E**) RT-qPCR analysis of major satellite RNA transcript levels in transiently transfected C2C12 mouse myoblasts (left) and brain of wild type and Mecp2 y-/ mice (right), respectively.

likely not detectable as our results indicate. Importantly, it should be noted that the increase in 5hmC was not due to enhanced Tet1 expression as wild type and Mecp2 deficient neurons had comparable Tet1 levels (Supplementary Figure S14D and E).

In conclusion, our data demonstrate on the subcellular level that knockout of Mecp2 results in increased neuronal 5hmC levels. We cannot exclude that other secondary effects may contribute to the increase of 5hmC at chromocenters *in vivo*. However, expression of Mecp2 in cells that do not normally express endogenous Mecp2 (mouse myoblasts, Figure 1B), decreases Tet-mediated oxidation of 5mC. The only difference between both sets of cells is the presence or not of Mecp2. Hence, according to our cell data (Figure 1A and B), unconfined access of Tet1 proteins to pericentric heterochromatin, which is occupied when Mecp2 is present, is very likely the dominant mechanism that allows 5hmC accumulation at pericentric heterochromatin.

Previous data showed that both, Mecp2 protein and 5hmC levels are high in neurons. To address this apparent contradictory coexistence, we furthermore analyzed the expression of Tet1 in different cell types and found high levels of the 5mC oxygenase Tet1 in NeuN positive compared to surrounding glial cells (Supplementary Figure S14A–C). Furthermore, we found the degree of colocalization between pericentric heterochromatin and 5hmC considerably increased as a consequence of Mecp2 depletion. While in wild-type brain, 5hmC is anti-correlated with DNA dense chromocenters, this is not the case for Mecp2 deficient neurons (Figure 6A, top). Similar results were obtained from line intensity plots of 5hmC distribution through pericentric heterochromatin (Figure 6B), as well as accumulation studies of 5hmC at chromocenters (Supplementary Figure S16D). Accordingly, 5hmC is indeed abundant in neurons of wild-type mice, however, only at sites of low Mecp2 accumulation. Therefore, Mecp2 has a local protective effect at pericentric heterochromatin.

### In the absence of Mecp2, Tet1 reactivates expression of major satellite repeats

Next, we tested whether hypomethylation of chromocenters (Figures 5 and 6A, B), which were previously described to be rich in major satellite repeats (72), leads to reactivation of these epigenetically silenced elements. Hence, we labeled and subsequently quantified (Supplementary Figure S16A and B) their RNA transcript levels by fluorescence in situ hybridization (FISH) in single cells of Mecp2 knockout mouse cerebella (Figure 6C). Compared to wild type, mean major satellite RNA FISH signals were significantly increased in nuclei of Mecp2 deficient cells (Figure 6C). Moreover, line intensity profiles of RNA FISH levels across chromocenters of the same nuclei, showed accumulation of major satellite transcripts directly at and in close proximity to pericentric heterochromatin (Figure 6D). To ensure that the observed transcriptional increase of major satellite DNA is not limited to the analyzed brain region, we additionally confirmed its elevated expression levels in whole Mecp2 y/- mouse brain by reverse transcription quantitative polymerase chain reaction (RT-qPCR) (Figure 6E, right). Moreover, we made use of C2C12 mouse myoblasts, which show no detectable levels of Mecp2 and Mbd2 (Supplementary Figure S16C, top; (19)) and, thus, allowed us to directly test the effect of Tet proteins on the expression of DNA repeats. Hence, mouse myoblasts ectopically expressing the catalytic active domain of Tet1 were sorted by flow cytometry and the transcriptional levels of major satellite repeats were quantified by RT-qPCR. When compared to mock treated cells, major satellite RNA transcripts were increased in mouse myoblasts, congenitally lacking Mecp2 (Supplementary Figure S16C, top) and ectopically expressing Tet1CD (Figure 6E, left). Coexpression of Mecp2, however, abolished Tet1CD-mediated reactivation of major satellite repeats and reduced major satellite transcription by half when compared to mock treated cells (Figure 6E, left). While transcription level of major satellite repeats almost doubled upon ectopic expression of the catalytically active Tet1 domain, overexpression of the inactive variant resulted in an increase of only 40% (Figure 6E, left). Accordingly, we conclude that the induction of major satellite expression requires at least in part the catalytic activity of Tet1 and, thus, results from increased 5hmC levels. As overexpression of both, the catalytically active and inactive domain of Tet1 leads to decondensation of pericentric heterochromatin (Zhang *et al.*, submitted), we furthermore deduce that the 40% increase of major satellite expression in cells expressing mutant Tet1CD, might be partially caused by reorganization of chromatin to a more open and, thus, accessible state.

Finally, we analyzed expression of satellite elements in triple Tet-knockout (KO) and wild type (wt) mouse embryonic stem cells (ESC) by RNA-seq. The median log2-fold change was -0.99, indicating that the expression of most genomic satellite sequences is down-regulated upon Tet1/2/3 depletion (Figure 6F). Taken together, our data demonstrate that in the absence of Mecp2, Tet1 reactivates the expression of epigenetically silenced (major) satellite repeats, which in turn might compromise genome stability (73,74). Therefore, we suggest that unrestricted Tet activity may be part of a pathogenic cascade in Rett syndrome, which is initiated by Mecp2 gene mutations that reduce or abolish DNA binding.

In the present study, we demonstrate that prior binding of methyl-CpG binding domain proteins Mecp2 and Mbd2 to DNA protects 5mC from Tet1CD mediated oxidation in a concentration dependent manner, thereby regulating chromatin composition (Figure 7A).

←

Expression levels are relative to Tet1CD transfected cells (left), or wild type mouse brain (right). Shown are average values from ≥ two biological replicates each measured from one (left), or two (right) independent cDNA synthesis reactions, respectively. Error bars represent ± SD. *P* values were calculated by an independent two-sample student's *t*-test (*$P < 0.05$, **$P = 0.01$, ***$P < 0.001$). (**F**) Violin-plot of the $\log_2$-fold changes of the triple Tet-knockout (KO) relative to wild type (wt) mouse embryonic stem cells (ESC (v6.5)) for all genes and all satellites. Negative values indicate a down-regulation in the knockout cells relative to the wild type, positive values an up-regulation. Significant elements are marked in color. The red line is at zero, i.e. the expected value if expression were identical in the wild type and knockout. Triple Tet-knockout: $P = 4.84 \times 10^{-2}$; genes: $P = 5 \times 10^{-15}$.

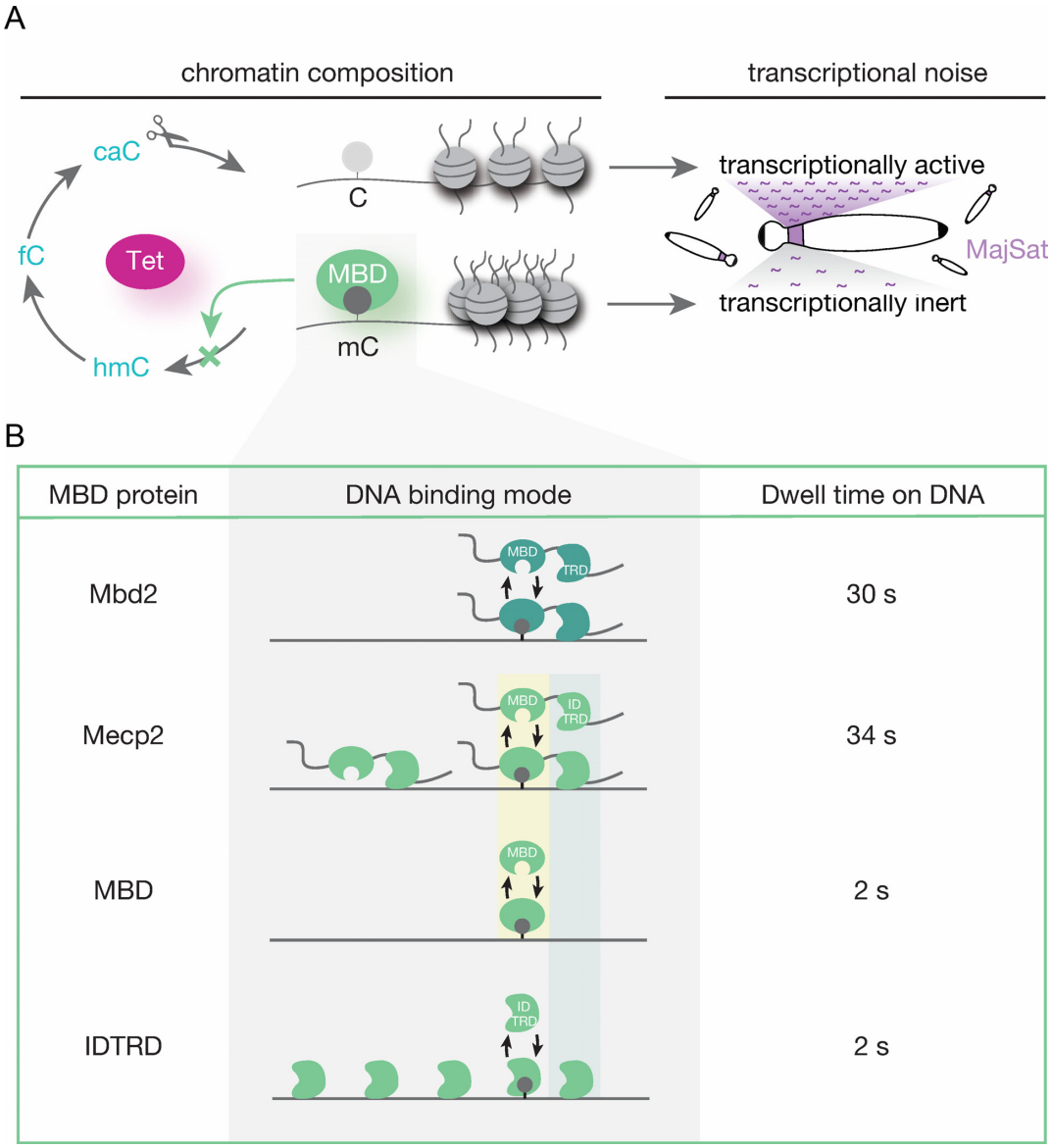



**Figure 7.** Mecp2 and Mbd2 preserve chromatin composition and thus genomic integrity by insulating 5mC from Tet1 activity. Scheme summarizing the main conclusions drawn from our studies. (**A**) We show that Mecp2 and Mbd2 protect 5mC from Tet1-mediated oxidation in a concentration dependent manner *in vivo* and *in vitro*. (**B**) The protection mechanism is not based on competition for 5mC *per se* but rather on sequence unspecific coverage of DNA and correlates with the respective MBD protein dwell time on DNA. (**A**) As a biological consequence, we measured increased 5hmC level in neurons of a mouse model for Rett syndrome with concomitant reactivation of epigenetically silenced pericentric DNA repeats.

The underlying molecular mechanism relies on competitive, sequence unspecific coverage of DNA and is affected by the respective MBD protein dwell time on DNA (Figure 7B). Accordingly, Tet binding to its substrate and, consequently, 5mC modification are inhibited and chromatin composition maintained. Hence, we infer that Tet1 activity is likely to vary according to tissue and cell specific distribution of methylated CpG sites, as it influences the binding affinity of MBD proteins (15). Furthermore, we propose that the quantity of methyl-CpG binding domain molecules per cell must be precisely regulated to accurately control Tet1 activity. Indeed, either duplication of the *MECP2* gene with increased respective protein level or mutant MECP2 proteins with impaired DNA binding, are both observed in Rett patients (61,75).

As a biological consequence, we measured increased 5hmC at pericentric heterochromatin in neurons of Mecp2 deficient mice with concomitant reactivation of epigenetically silenced major satellite repeats (Figure 7A). Compensatory effects by Mbd2 cannot come into play as its expression levels are significantly reduced in Mecp2 knockout brain (49). As Tet1 reactivates transcription of major satellite repeats in the absence of Mecp2 and Mbd2 proteins, we conclude that the transcriptional noise increase in Rett animal models (20) is likely to result, at least in part, from unconfined Tet activity and, thus, provide a potential Tet-induced pathophysiological pathway in Rett syndrome.

Since almost all mature, postmitotitc neurons were shown to express abundant levels of methyltransferases Dnmt1 and Dnmt3a (76,77), we propose that stabilization or reversion of Rett symptoms upon delivery of Mecp2 (78) results from re-methylation and subsequent binding and protection of 5mC by the exogenous wild type Mecp2 protein.

In summary, these data provide mechanistic insights into the regulation of Tet1 activity by methyl-CpG binding domain proteins and argue for a role of the MBD proteins as guardians of the epigenome.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We are indebted to Anna M. Bischoff (University of Goettingen, Germany) for providing Mecp2 y/- brain, Congdi Song (LMU Munich, Germany) for providing cDNA from Mecp2 y/- brain, Aleksandra Szwagierczak (LMU Munich, Germany) for providing plasmids, Adrian Bird (Wellcome Trust Centre for Cell Biology, Edinburgh, UK) and Rudolf Jaenisch (Whitehead Institute for Biomedical Research, Cambridge, USA) for providing cell lines, Bodo Laube (Technische Universität Darmstadt) for help with the primary neuron cultures, Corella S. Casas-Delucchi, Annette Becker, Alexander Rapp, Bianca Bertulat and Manuela Milden for experimental advice and discussions.
*Author contributions:* A.K.L., P.Z., S.M., U.M., F.D.H., C.R., C.T. and A.L. performed experiments. A.K.L., P.Z., H.D.H., U.M. and I.H. analyzed the data. C.S. performed statistical tests for Figures 1 and 3. H.L. and M.C.C. conceived the project, provided expertise and feedback. A.K.L., P.Z. and M.C.C. conceived and designed the experiments and wrote the manuscript.

## FUNDING

China Scholarship Council (to P.Z; Deutsche Forschungsgemeinschaft [DFG CA 198/7 to M.C.C. and SFB1064/A17 to H.L.]. Funding for open access charge: DFG.
*Conflict of interest statement.* None declared.

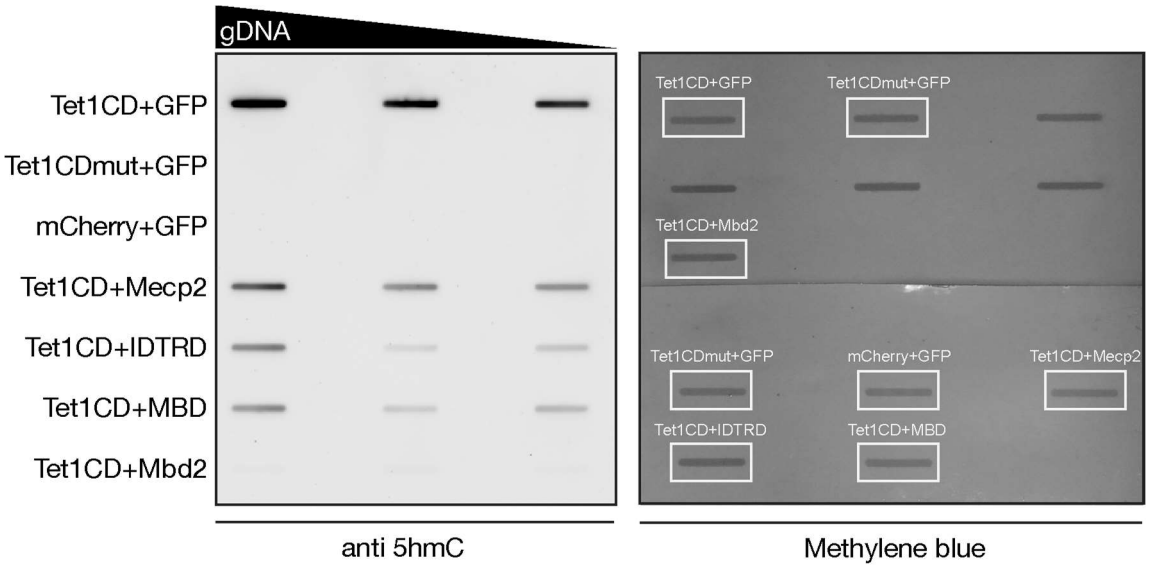
## REFERENCES

1. Bird,A.P. and Wolffe,A.P. (1999) Methylation-induced repression–belts, braces, and chromatin. *Cell*, **99**, 451–454.
2. Ehrlich,M., Gama-Sosa,M.A., Huang,L.H., Midgett,R.M., Kuo,K.C., McCune,R.A. and Gehrke,C. (1982) Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res.*, **10**, 2709–2721.
3. Bird,A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev.*, **16**, 6–21.
4. Goll,M.G. and Bestor,T.H. (2005) Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.*, **74**, 481–514.
5. Esteller,M. (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat. Rev. Genet.*, **8**, 286–298.
6. Peng,J.C. and Karpen,G.H. (2007) H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. *Nat. Cell Biol.*, **9**, 25–35.
7. Jaenisch,R. and Bird,A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.*, **33**(Suppl), 245–254.
8. Hendrich,B. and Bird,A. (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.*, **18**, 6538–6547.
9. Fraga,M.F., Ballestar,E., Montoya,G., Taysavang,P., Wade,P.A. and Esteller,M. (2003) The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties. *Nucleic Acids Res.*, **31**, 1765–1774.
10. Nan,X., Campoy,F.J. and Bird,A. (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell*, **88**, 471–481.
11. Jones,P.L., Veenstra,G.J., Wade,P.A., Vermaak,D., Kass,S.U., Landsberger,N., Strouboulis,J. and Wolffe,A.P. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.*, **19**, 187–191.
12. Ebert,D.H., Gabel,H.W., Robinson,N.D., Kastan,N.R., Hu,L.S., Cohen,S., Navarro,A.J., Lyst,M.J., Ekiert,R., Bird,A.P. *et al.* (2013) Activity-dependent phosphorylation of MeCP2 threonine 308 regulates interaction with NCoR. *Nature*, **499**, 341–345.
13. Lyst,M.J., Ekiert,R., Ebert,D.H., Merusi,C., Nowak,J., Selfridge,J., Guy,J., Kastan,N.R., Robinson,N.D., de Lima Alves,F. *et al.* (2013) Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. *Nat. Neurosci.*, **16**, 898–902.
14. Ghosh,R.P., Nikitina,T., Horowitz-Scherer,R.A., Gierasch,L.M., Uversky,V.N., Hite,K., Hansen,J.C. and Woodcock,C.L. (2010) Unique physical properties and interactions of the domains of methylated DNA binding protein 2. *Biochemistry-US*, **49**, 4395–4410.
15. Hansen,J.C., Ghosh,R.P. and Woodcock,C.L. (2010) Binding of the Rett syndrome protein, MeCP2, to methylated and unmethylated DNA and chromatin. *Iubmb Life*, **62**, 732–738.
16. Kumar,A., Kamboj,S., Malone,B.M., Kudo,S., Twiss,J.L., Czymmek,K.J., LaSalle,J.M. and Schanen,N.C. (2008) Analysis of protein domains and Rett syndrome mutations indicate that multiple regions influence chromatin-binding dynamics of the chromatin-associated protein MECP2 in vivo. *J. Cell Sci.*, **121**, 1128–1137.
17. Lewis,J.D., Meehan,R.R., Henzel,W.J., Maurer-Fogy,I., Jeppesen,P., Klein,F. and Bird,A. (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*, **69**, 905–914.
18. Becker,A., Allmann,L., Hofstatter,M., Casa,V., Weber,P., Lehmkuhl,A., Herce,H.D. and Cardoso,M.C. (2013) Direct homo- and hetero-interactions of MeCP2 and MBD2. *PLoS One*, **8**, e53730.
19. Brero,A., Easwaran,H.P., Nowak,D., Grunewald,I., Cremer,T., Leonhardt,H. and Cardoso,M.C. (2005) Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation. *J. Cell Biol.*, **169**, 733–743.
20. Skene,P.J., Illingworth,R.S., Webb,S., Kerr,A.R., James,K.D., Turner,D.J., Andrews,R. and Bird,A.P. (2010) Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Mol. Cell*, **37**, 457–468.
21. Kriaucionis,S. and Heintz,N. (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*, **324**, 929–930.
22. Lorsbach,R.B., Moore,J., Mathew,S., Raimondi,S.C., Mukatira,S.T. and Downing,J.R. (2003) TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia*, **17**, 637–641.
23. Tahiliani,M., Koh,K.P., Shen,Y., Pastor,W.A., Bandukwala,H., Brudno,Y., Agarwal,S., Iyer,L.M., Liu,D.R., Aravind,L. *et al.* (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, **324**, 930–935.
24. Ito,S., D'Alessio,A.C., Taranova,O.V., Hong,K., Sowers,L.C. and Zhang,Y. (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*, **466**, 1129–1133.
25. Ito,S., Shen,L., Dai,Q., Wu,S.C., Collins,L.B., Swenberg,J.A., He,C. and Zhang,Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, **333**, 1300–1303.
26. Guo,J.U., Su,Y., Zhong,C., Ming,G.L. and Song,H. (2011) Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell*, **145**, 423–434.
27. Spruijt,C.G., Gnerlich,F., Smits,A.H., Pfaffeneder,T., Jansen,P.W., Bauer,C., Munzel,M., Wagner,M., Muller,M., Khan,F. *et al.* (2013)

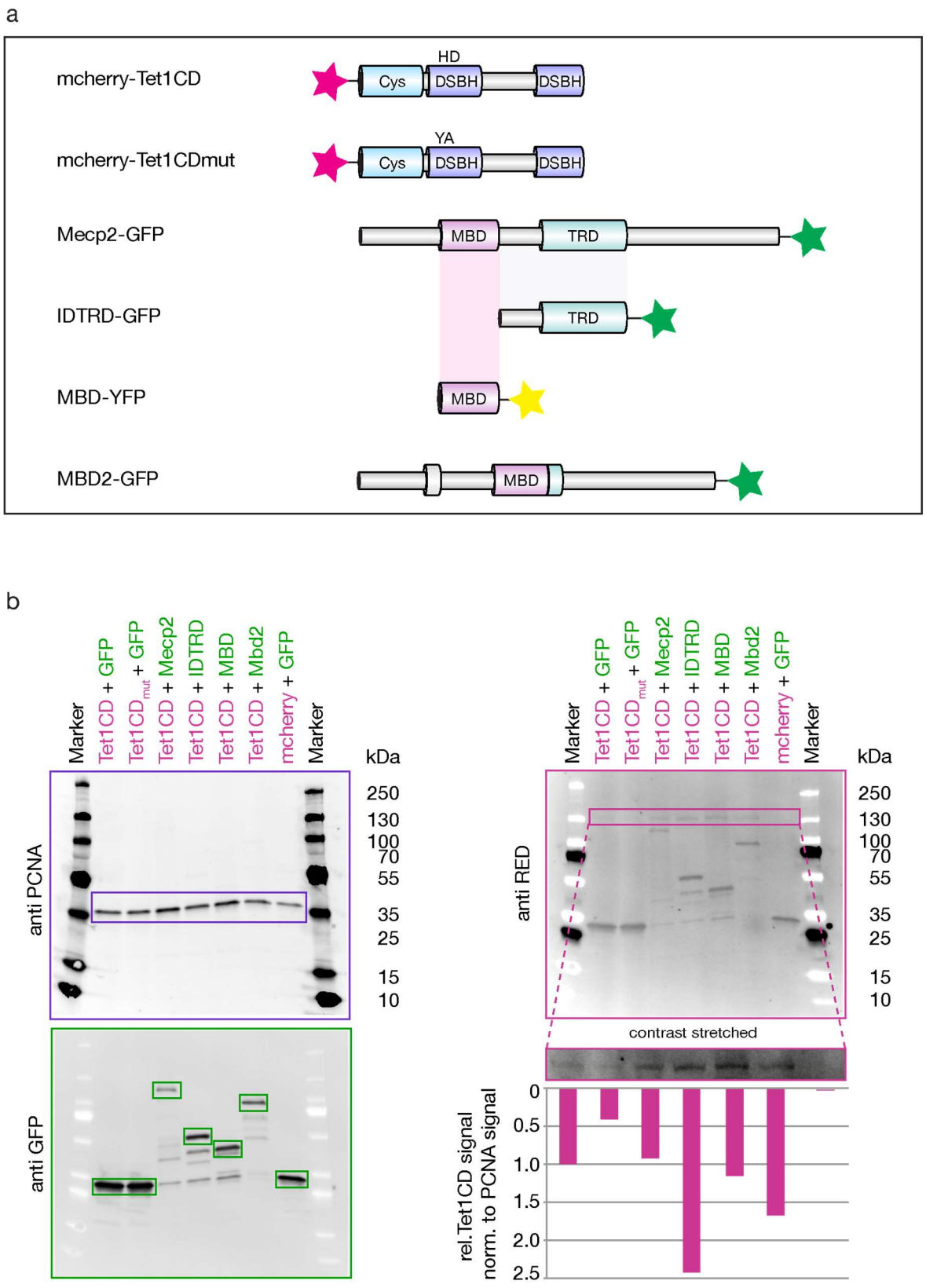
Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell*, **152**, 1146–1159.
28. Portela,A. and Esteller,M. (2010) Epigenetic modifications and human disease. *Nat. Biotechnol.*, **28**, 1057–1068.
29. Amir,R.E., Van den Veyver,I.B., Wan,M., Tran,C.Q., Francke,U. and Zoghbi,H.Y. (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.*, **23**, 185–188.
30. Kriaucionis,S. and Bird,A. (2003) DNA methylation and Rett syndrome. *Hum. Mol. Genet.*, **12**(Spec No 2), R221–R227.
31. Muotri,A.R., Marchetto,M.C., Coufal,N.G., Oefner,R., Yeo,G., Nakashima,K. and Gage,F.H. (2010) L1 retrotransposition in neurons is modulated by MeCP2. *Nature*, **468**, 443–446.
32. Jost,K.L., Rottach,A., Milden,M., Bertulat,B., Becker,A., Wolf,P., Sandoval,J., Petazzi,P., Huertas,D., Esteller,M. *et al.* (2011) Generation and characterization of rat and mouse monoclonal antibodies specific for MeCP2 and their use in X-inactivation studies. *PLoS One*, **6**, e26499.
33. Kudo,S., Nomura,Y., Segawa,M., Fujita,N., Nakao,M., Schanen,C. and Tamura,M. (2003) Heterogeneity in residual function of MeCP2 carrying missense mutations in the methyl CpG binding domain. *J. Med. Genet.*, **40**, 487–493.
34. Qin,W., Leonhardt,H. and Spada,F. (2011) Usp7 and Uhrf1 control ubiquitination and stability of the maintenance DNA methyltransferase Dnmt1. *J. Cell. Biochem.*, **112**, 439–444.
35. Muller,U., Bauer,C., Siegl,M., Rottach,A. and Leonhardt,H. (2014) TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation. *Nucleic Acids Res.*, **42**, 8592–8604.
36. Becker,A., Zhang,P., Allmann,L., Meilinger,D., Bertulat,B., Eck,D., Hofstaetter,M., Bartolomei,G., Hottiger,M.O., Schreiber,V. *et al.* (2016) Poly(ADP-ribosyl)ation of methyl CpG binding domain protein 2 regulates chromatin structure. *J. Biol. Chem.*, **291**, 4873–4881.
37. Berger,I., Fitzgerald,D.J. and Richmond,T.J. (2004) Baculovirus expression system for heterologous multiprotein complexes. *Nat. Biotechnol.*, **22**, 1583–1587.
38. Yaffe,D. and Saxel,O. (1977) Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature*, **270**, 725–727.
39. Cardoso,M.C., Leonhardt,H. and Nadal-Ginard,B. (1993) Reversal of terminal differentiation and control of DNA replication: cyclin A and Cdk2 specifically localize at subnuclear sites of DNA replication. *Cell*, **74**, 979–992.
40. Casas-Delucchi,C.S., Becker,A., Bolius,J.J. and Cardoso,M.C. (2012) Targeted manipulation of heterochromatin rescues MeCP2 Rett mutants and re-establishes higher order chromatin organization. *Nucleic Acids Res.*, **40**, e176.
41. Agarwal,N., Hardt,T., Brero,A., Nowak,D., Rothbauer,U., Becker,A., Leonhardt,H. and Cardoso,M.C. (2007) MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. *Nucleic Acids Res.*, **35**, 5402–5408.
42. Dawlaty,M.M., Breiling,A., Le,T., Barrasa,M.I., Raddatz,G., Gao,Q., Powell,B.E., Cheng,A.W., Faull,K.F., Lyko,F. *et al.* (2014) Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. *Dev. Cell*, **29**, 102–111.
43. Guy,J., Hendrich,B., Holmes,M., Martin,J.E. and Bird,A. (2001) A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat. Genet.*, **27**, 322–326.
44. Pichler,G., Leonhardt,H. and Rothbauer,U. (2012) Fluorescent protein specific Nanotraps to study protein-protein interactions and histone-tail peptide binding. *Methods Mol. Biol.*, **911**, 475–483.
45. Rothbauer,U., Zolghadr,K., Muyldermans,S., Schepers,A., Cardoso,M.C. and Leonhardt,H. (2008) A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Mol. Cell. Proteomics: MCP*, **7**, 282–289.
46. Mortusewicz,O., Rothbauer,U., Cardoso,M.C. and Leonhardt,H. (2006) Differential recruitment of DNA Ligase I and III to DNA repair sites. *Nucleic Acids Res.*, **34**, 3523–3532.
47. Rottach,A., Kremmer,E., Nowak,D., Leonhardt,H. and Cardoso,M.C. (2008) Generation and characterization of a rat monoclonal antibody specific for multiple red fluorescent proteins. *Hybridoma (Larchmt)*, **27**, 337–343.
48. Miller,S.A., Dykes,D.D. and Polesky,H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.*, **16**, 1215.
49. Song,C., Feodorova,Y., Guy,J., Peichl,L., Jost,K.L., Kimura,H., Cardoso,M.C., Bird,A., Leonhardt,H., Joffe,B. *et al.* (2014) DNA methylation reader MECP2: cell type- and differentiation stage-specific protein distribution. *Epigenet. Chromatin*, **7**, 17.
50. Szwagierczak,A., Bultmann,S., Schmidt,C.S., Spada,F. and Leonhardt,H. (2010) Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res.*, **38**, e181.
51. Bauer,C., Gobel,K., Nagaraj,N., Colantuoni,C., Wang,M., Muller,U., Kremmer,E., Rottach,A. and Leonhardt,H. (2015) Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT). *J. Biol. Chem.*, **290**, 4801–4812.
52. Herce,H.D., Casas-Delucchi,C.S. and Cardoso,M.C. (2013) New image colocalization coefficient for fluorescence microscopy to quantify (bio-)molecular interactions. *J. Microsc.*, **249**, 184–194.
53. Casas-Delucchi,C.S., van Bemmel,J.G., Haase,S., Herce,H.D., Nowak,D., Meilinger,D., Stear,J.H., Leonhardt,H. and Cardoso,M.C. (2012) Histone hypoacetylation is required to maintain late replication timing of constitutive heterochromatin. *Nucleic Acids Res.*, **40**, 159–169.
54. Picelli,S., Björklund,Å.K., Faridani,OR., Sagasser,S., Winberg,G. and Sandberg,R. (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods*, **10**, 1096–1098.
55. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
56. Jin,Y., Tam,O.H., Paniagua,E. and Hammell,M. (2015) TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics*, **31**, 3593–3599.
57. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
58. Benjamini,Y. and Hochberg,Y. (1995) Controlling the false discovery rate — a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B Met.*, **57**, 289–300.
59. Skene,P.J., Illingworth,R.S., Webb,S., Kerr,A.R., James,K.D., Turner,D.J., Andrews,R. and Bird,A.P. (2010) Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Mol. Cell*, **37**, 457–468.
60. Khrapunov,S., Warren,C., Cheng,H., Berko,E.R., Greally,J.M. and Brenowitz,M. (2014) Unusual characteristics of the DNA binding domain of epigenetic regulatory protein MeCP2 determine its binding specificity. *Biochemistry-US*, **53**, 3379–3391.
61. Agarwal,N., Becker,A., Jost,K.L., Haase,S., Thakur,B.K., Brero,A., Hardt,T., Kudo,S., Leonhardt,H. and Cardoso,M.C. (2011) MeCP2 Rett mutations affect large scale chromatin organization. *Human Mol. Genet.*, **20**, 4187–4195.
62. Lindhout,B.I., Fransz,P., Tessadori,F., Meckel,T., Hooykaas,P.J. and van der Zaal,B.J. (2007) Live cell imaging of repetitive DNA sequences via GFP-tagged polydactyl zinc finger proteins. *Nucleic Acids Res.* **35**, e107.
63. Thanisch,K., Schneider,K., Morbitzer,R., Solovei,I., Lahaye,T., Bultmann,S. and Leonhardt,H. (2014) Targeting and tracing of specific DNA sequences with dTALEs in living cells. *Nucleic Acids Res.*, **42**, e38.
64. Menafra,R., Brinkman,A.B., Matarese,F., Franci,G., Bartels,S.J., Nguyen,L., Shimbo,T., Wade,P.A., Hubner,N.C. and Stunnenberg,H.G. (2014) Genome-wide binding of MBD2 reveals strong preference for highly methylated loci. *PLoS One*, **9**, e99603.
65. Desai,M.A., Webb,H.D., Sinanan,L.M., Scarsdale,J.N., Walavalkar,N.M., Ginder,G.D. and Williams,D.C. Jr (2015) An intrinsically disordered region of methyl-CpG binding domain protein 2 (MBD2) recruits the histone deacetylase core of the NuRD complex. *Nucleic Acids Res.*, **43**, 3100–3113.
66. Szulwach,K.E., Li,X., Li,Y., Song,C.X., Wu,H., Dai,Q., Irier,H., Upadhyay,A.K., Gearing,M., Levey,A.I. *et al.* (2011) 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat. Neurosci.*, **14**, 1607–1616.

67. Hashimoto,H., Pais,J.E., Zhang,X., Saleh,L., Fu,Z.Q., Dai,N., Correa,I.R. Jr, Zheng,Y. and Cheng,X. (2014) Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA. *Nature*, **506**, 391–395.
68. Hu,L., Li,Z., Cheng,J., Rao,Q., Gong,W., Liu,M., Shi,Y.G., Zhu,J., Wang,P. and Xu,Y. (2013) Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. *Cell*, **155**, 1545–1555.
69. Ramirez,J.M., Ward,C.S. and Neul,J.L. (2013) Breathing challenges in Rett syndrome: lessons learned from humans and animal models. *Respir. Physiol. Neurobiol.*, **189**, 280–287.
70. Munzel,M., Globisch,D., Bruckl,T., Wagner,M., Welzmiller,V., Michalakis,S., Muller,M., Biel,M. and Carell,T. (2010) Quantification of the sixth DNA base hydroxymethylcytosine in the brain. *Angew. Chem.*, **49**, 5375–5377.
71. Wu,H., D'Alessio,A.C., Ito,S., Wang,Z., Cui,K., Zhao,K., Sun,Y.E. and Zhang,Y. (2011) Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev.*, **25**, 679–684.
72. Jones,K.W. (1970) Chromosomal and nuclear location of mouse satellite DNA in individual cells. *Nature*, **225**, 912–915.
73. Hansen,R.S., Wijmenga,C., Luo,P., Stanek,A.M., Canfield,T.K., Weemaes,C.M. and Gartler,S.M. (1999) The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 14412–14417.
74. Peters,A.H., O'Carroll,D., Scherthan,H., Mechtler,K., Sauer,S., Schofer,C., Weipoltshammer,K., Pagani,M., Lachner,M., Kohlmaier,A. *et al.* (2001) Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell*, **107**, 323–337.
75. Ramocki,M.B., Peters,S.U., Tavyev,Y.J., Zhang,F., Carvalho,C.M., Schaaf,C.P., Richman,R., Fang,P., Glaze,D.G., Lupski,J.R. *et al.* (2009) Autism and other neuropsychiatric symptoms are prevalent in individuals with MeCP2 duplication syndrome. *Ann. Neurol.*, **66**, 771–782.
76. Watanabe,D., Uchiyama,K. and Hanaoka,K. (2006) Transition of mouse de novo methyltransferases expression from Dnmt3b to Dnmt3a during neural progenitor cell development. *Neuroscience*, **142**, 727–737.
77. Goto,K., Numata,M., Komura,J.I., Ono,T., Bestor,T.H. and Kondo,H. (1994) Expression of DNA methyltransferase gene in mature and immature neurons as well as proliferating cells in mice. *Differentiation*, **56**, 39–44.
78. Garg,S.K., Lioy,D.T., Cheval,H., McGann,J.C., Bissonnette,J.M., Murtha,M.J., Foust,K.D., Kaspar,B.K., Bird,A. and Mandel,G. (2013) Systemic delivery of MeCP2 rescues behavioral and cellular deficits in female mouse models of Rett syndrome. *J. Neurosci.*, **33**, 13612–13620.

**Supplementary Figures**



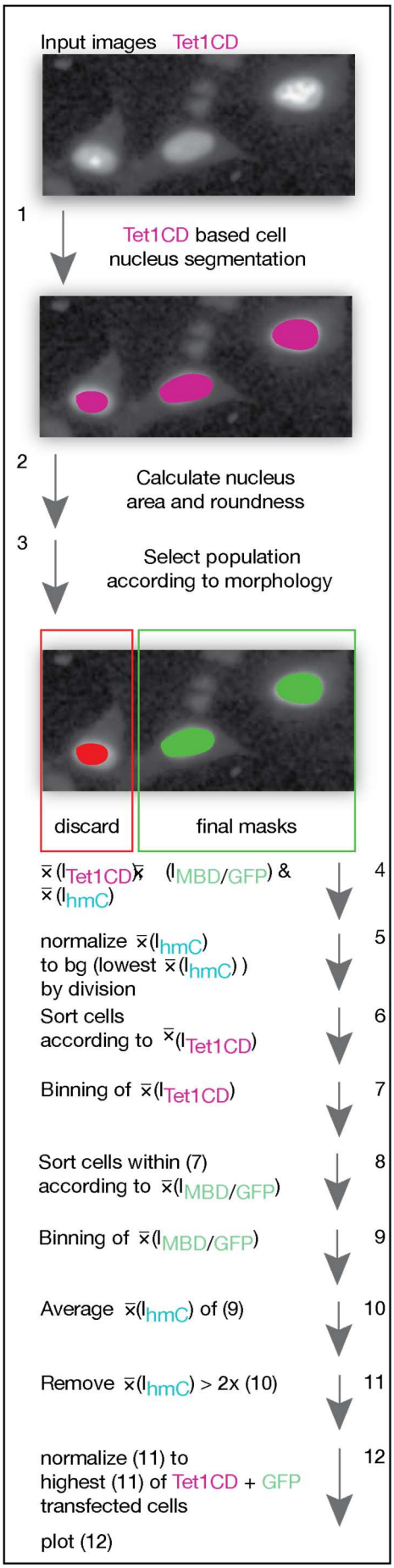

**Supplementary Figure 1. Full slot blots used for the quantification of 5hmC in genomic DNA (gDNA) of transiently transfected HEK cells (Figure 1a, bottom).** (**left**) anti 5hmC. (**right**) Methylene blue staining.




**Supplementary Figure 2. Western blot analysis of Tet1CD and MBD protein levels of transiently transfected, FACS sorted HEK cells**

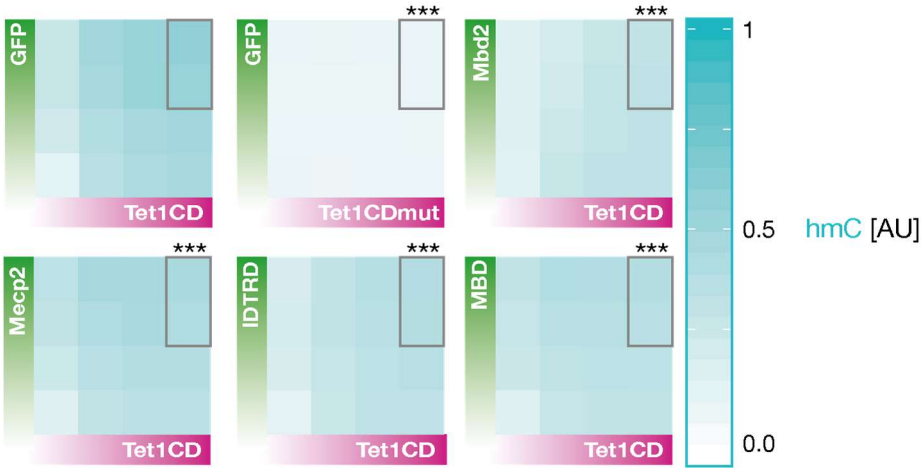
(**a**) Representative schemes of proteins that were used for *in vitro* and *in vivo* studies. Tet1CD corresponds to the Tet1 catalytic domain (aa 1365-2057) and Tet1CDmut is the catalytically inactive

domain of Tet1 (aa 1365-2057) containing two point mutations (H1652Y, D1654A) that abolish binding of the co-factor $Fe^{2+}$. Magenta colored star=mCherry; green colored star=GFP; yellow colored star=YFP. Cys=cysteine rich region; DSBH=double stranded beta helix; MBD=methyl CpG binding domain; TRD=transcriptional repression domain. (**b**) Shown are GFP or YFP-tagged Mecp2, IDTRD, MBD and Mbd2 (green, anti GFP), as well as, mCherry-tagged Tet1CD and Tet1CDmut (magenta, anti RED) proteins of transiently transfected HEK cells after enrichment for double-transfected cells by Fluorescent Activated Cell Sorting (FACS). Proliferating cell nuclear antigen (PCNA) was used as a loading control. Tet1CD expression levels are similar in cells coexpressing MBD proteins and comparatively low in cells coexpressing GFP. Accordingly, decreased genomic 5hmC levels observed in cells coexpressing MBD proteins (**Fig. 1 and Fig. 3**) are not due to lower Tet1CD expression levels.


Input images Tet1CD
1
Tet1CD based cell nucleus segmentation
2
Calculate nucleus area and roundness
3
Select population according to morphology
discard
final masks
x̄(I Tet1CD) x̄ (I MBD/GFP) & x̄(I hmC)
4
normalize x̄(I hmC) to bg (lowest x̄(I hmC) ) by division
5
Sort cells according to x̄(I Tet1CD)
6
Binning of x̄(I Tet1CD)
7
Sort cells within (7) according to x̄(I MBD/GFP)
8
Binning of x̄(I MBD/GFP)
9
Average x̄(I hmC) of (9)
10
Remove x̄(I hmC) > 2x (10)
11
normalize (11) to highest (11) of Tet1CD + GFP transfected cells
12
plot (12)

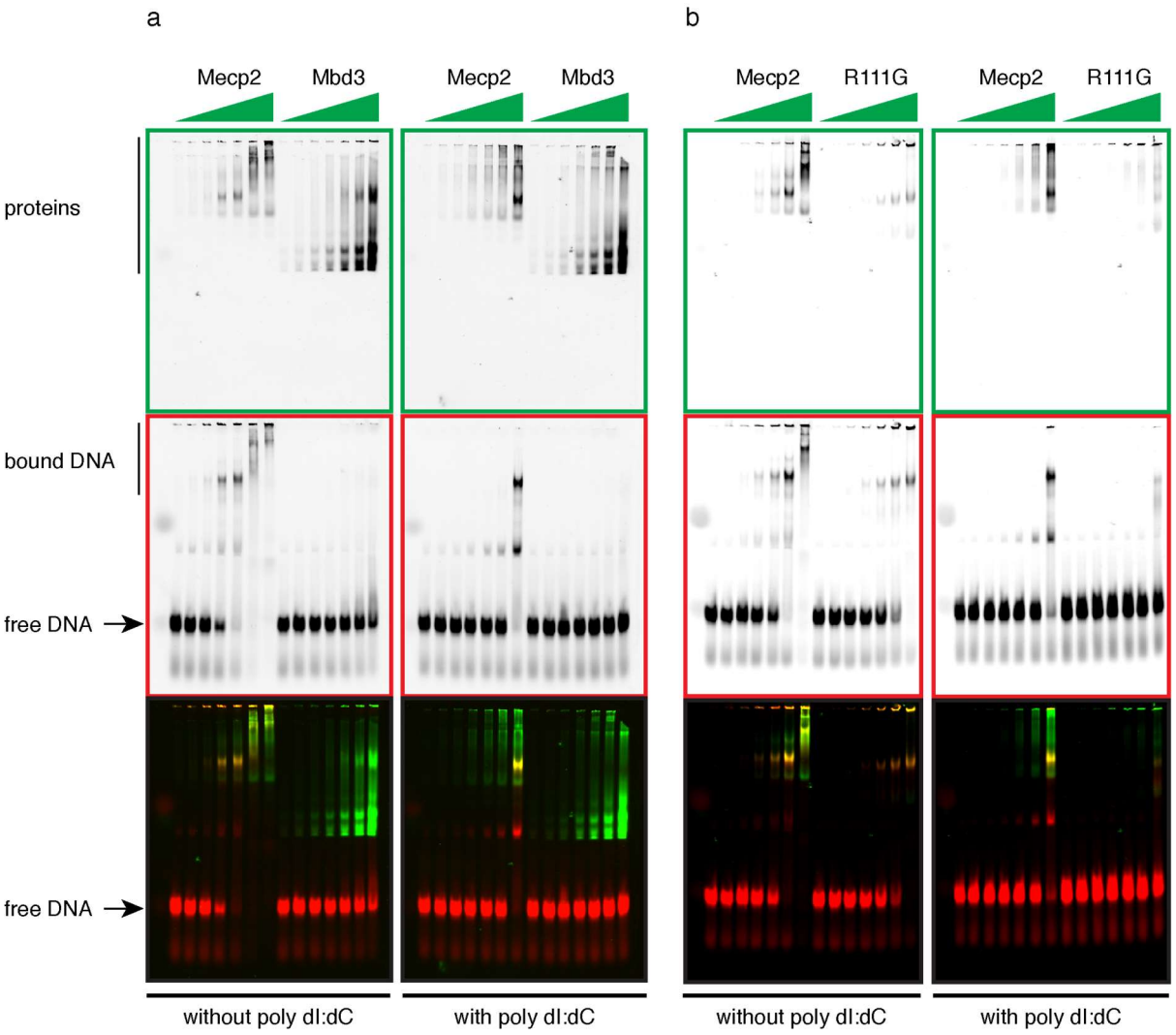
**Supplementary Figure 3. Summary of the different steps used for quantification and normalization of genomic 5hmC levels in transiently transfected C2C12 mouse myoblasts**

(**1**) Cell nuclei segmentation according to Tet1 signals. (**2**) Nuclei area and roundness calculation. (**3**) Selection of cell populations based on morphology properties. (**4**) Calculation of mean Tet1CD-, MBD- and 5hmC intensities. (**5**) Normalization of mean hmC signals to background. (**6**) Sort cells according to Tet1 signals. (**7**) Bin cells based on mean Tet1 intensities. (**8**) Sort cells within subgroups (7) according to mean MBD intensities. (**9**) Bin cells within subgroups (7) based on MBD intensities. (**10**) Average mean 5hmC intensities of each sub-subgroup (9). (**11**) Remove outliers. (**12**) Normalize mean 5hmC signals of each sub-subgroup (10) to highest 5hmC level of Tet1CD + GFP transfected cells and plot in form of a heat map as a function of Tet1CD- and MBD protein expression levels.

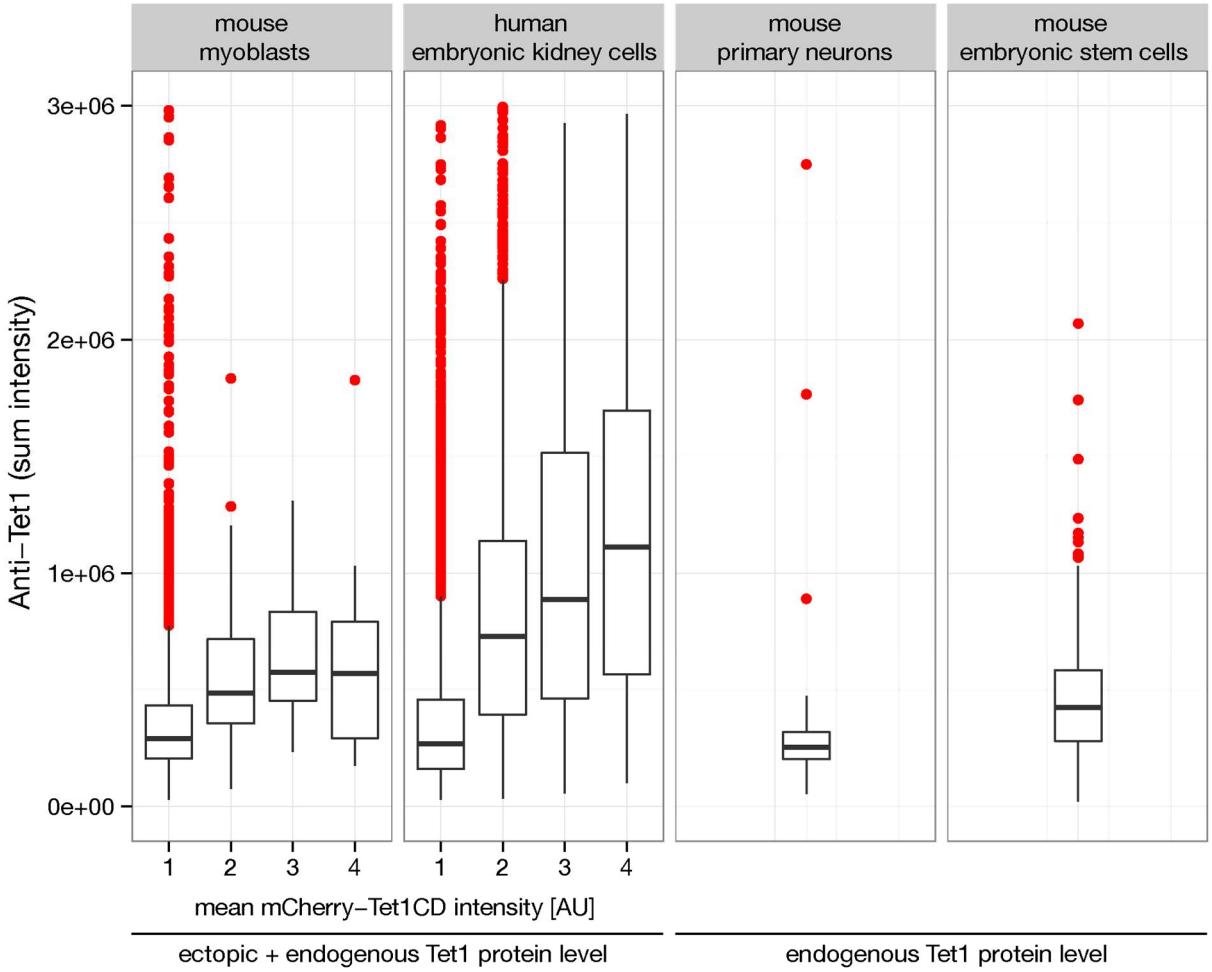



**Supplementary Figure 4. Effect of methyl-CpG binding domain proteins on Tet1CD activity in human embryonic kidney cells**

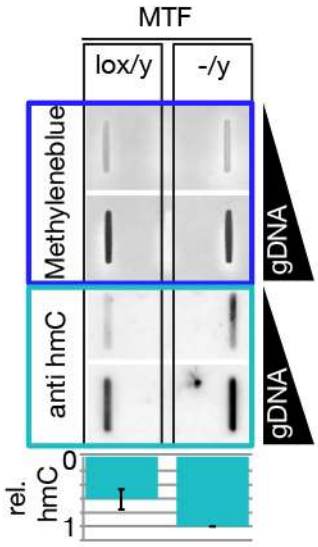
*In situ* staining and quantification of genomic 5hmC levels in transiently transfected human embryonic kidney (HEK) cells. Images were acquired on an automated high throughput imaging system with a 20x, 0.45 NA objective. Gradient heat maps show relative 5hmC (=cyan) signals as a function of increasing Tet1CD (=magenta) and MBD (=green) protein expression levels depicted by the green and magenta gradient bars. Shown are mean values of three (Tet1CD+GFP, n=38840; Tet1CDmut+GFP, n=53761; Tet1CD+Mbd2, n=39572; Tet1CD+Mecp2, n=41568; Tet1CD+IDTRD, n=30919; Tet1CD+MBD, n=32957) independent experiments, respectively. For statistical tests, 5hmC signals of cells with high Tet1CD and high Mecp2/Mbd2/Mbd3 protein levels (framed in grey) were used. All samples differed highly significant (***; $p < 0.001$; post-hoc pairwise *t* test) from Tet1CD+GFP.



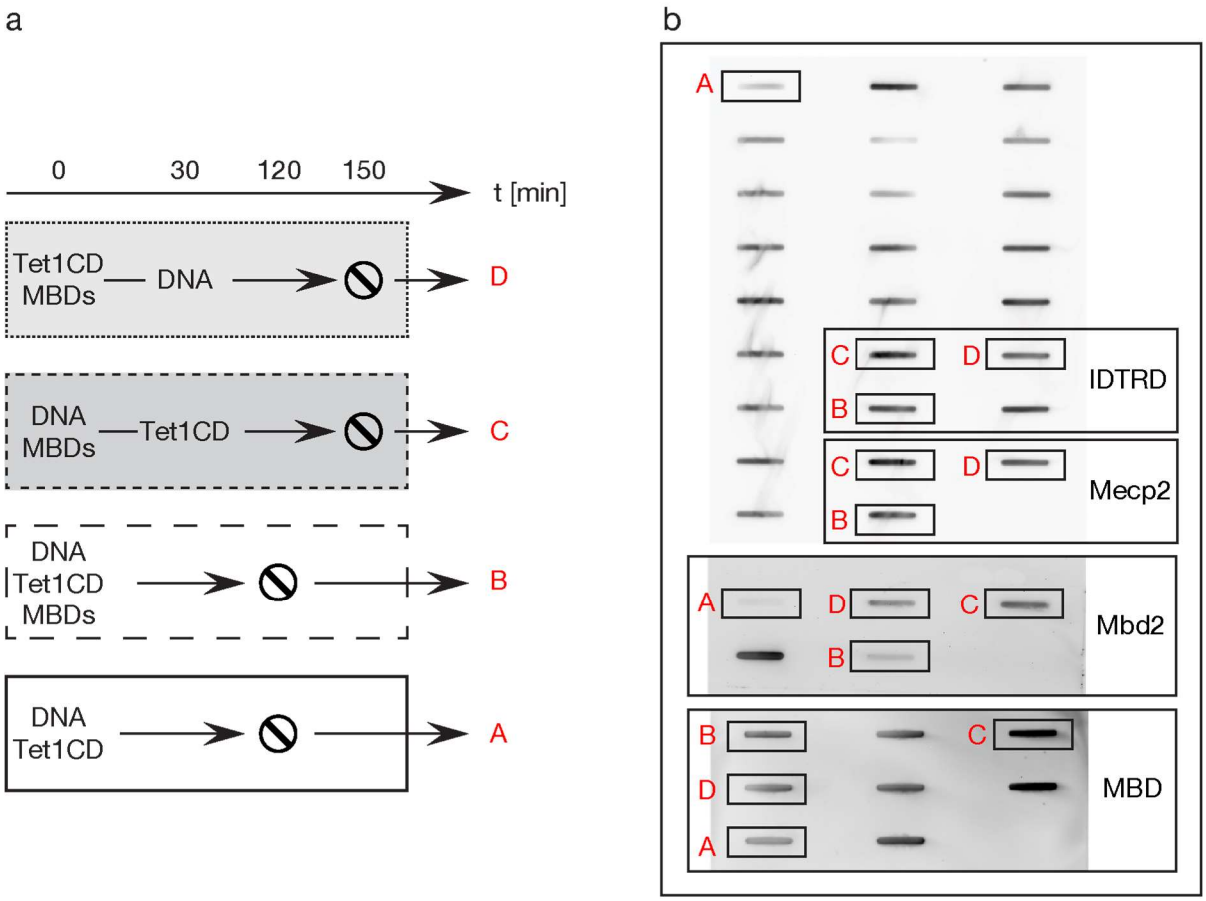

**Supplementary Figure 5. Binding properties of Mecp2, Mbd3 and Mecp2$_{R111G}$ to methylated DNA.** (**a**) Electrophoretic mobility shift assay (EMSA) of methylated DNA (red) with Mecp2 and Mbd3 (green) in the absence (left) and presence (right) of poly di:dC. (b) Electrophoretic mobility shift assay (EMSA) of methylated DNA (red) with Mecp2 and Mecp2$_{R111G}$ (green) in the absence (left) and presence (right) of poly di:dC.




**Supplementary Figure 6. Estimation of the level of mCherry-Tet1CD in transfected mouse myoblasts and human embryonic kidney cells.** *In situ* staining and quantification of Tet1 in mCherry-Tet1CD transfected mouse myoblasts (n=2704) and human embryonic kidney cells (n=11025), as well as untransfected primary mouse neurons (n=102) and mouse embryonic stem cells (n=373). Images were acquired on an automated high throughput imaging system with a 20x, 0.45 NA objective. Mouse myoblast and human embryonic kidney cells were binned (as in Figure 1b and 3b) according to the ectopic mCherry-Tet1CD expression levels. Box plots represent the total nuclear Tet1 signal (endogenous + ectopic), as well as the first and third quartiles. Whiskers extend to 1.5 times the interquartile range.

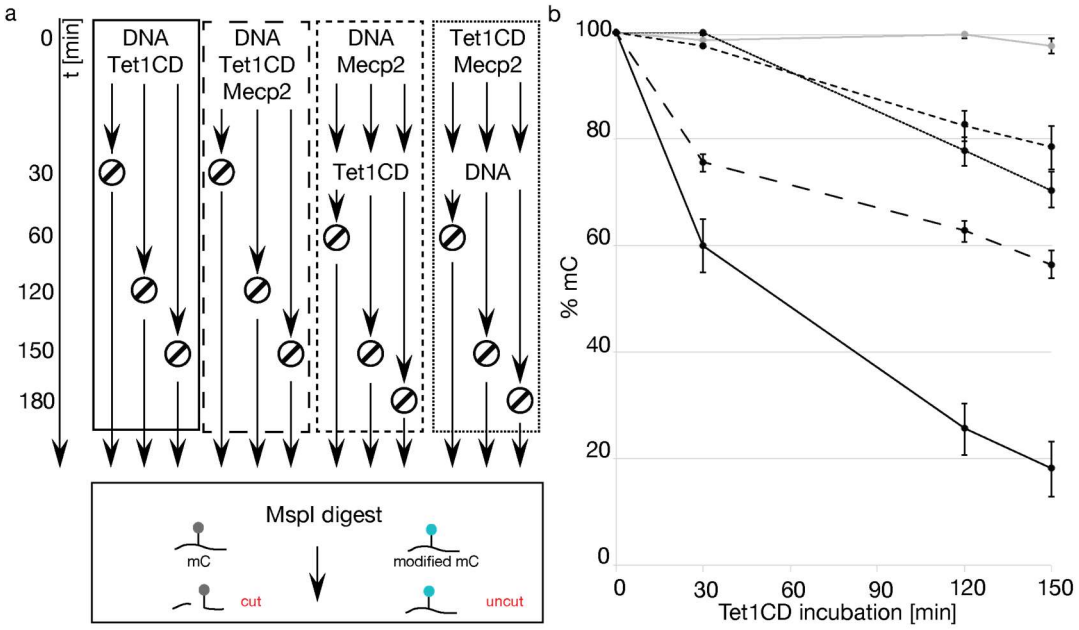



**Supplementary Figure 7. Genomic 5hmC level in the presence and absence of Mecp2 in mouse tail fibroblast (MTF) cells.** Immunological assay to determine 5hmC levels in genomic DNA (gDNA) of wild type (lox/y) and Mecp2 knockout (-/y) MTF cells. gDNA quantities were monitored by methylene blue staining.



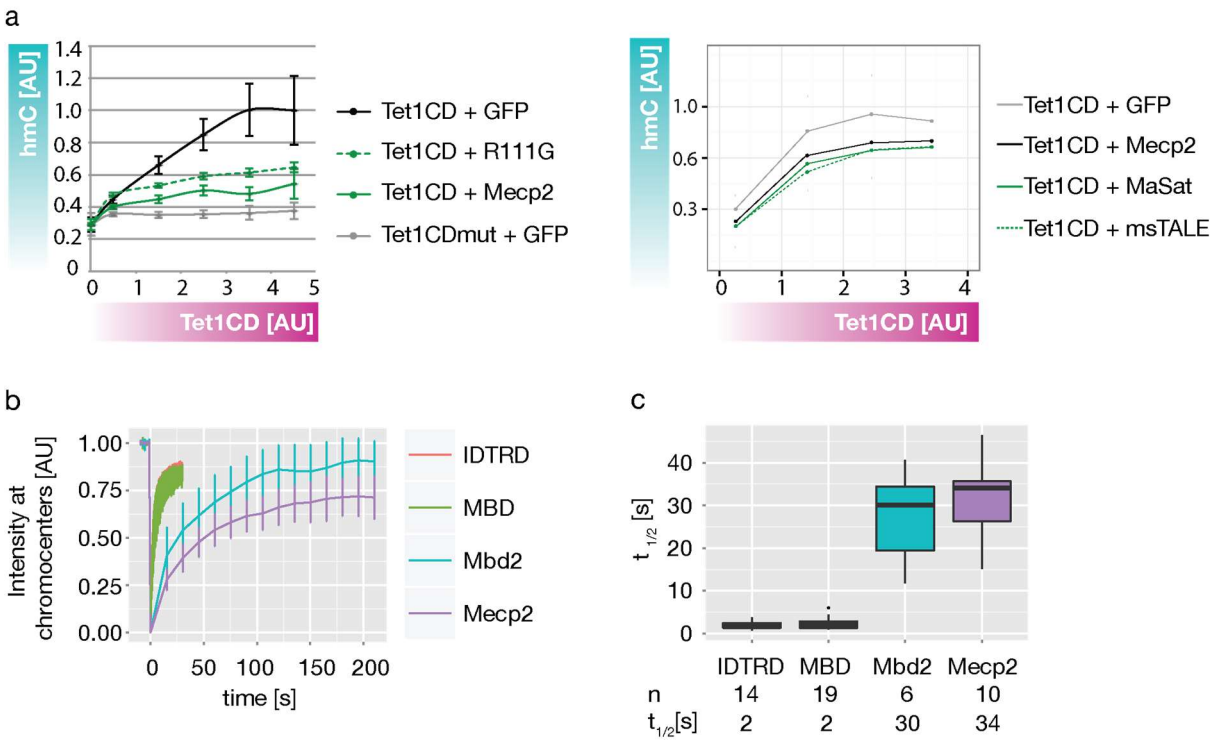

**Supplementary Figure 8. Full slot blots used for the quantification of remaining 5mC in double-stranded DNA after simultaneous or successive incubation with Tet1CD and MBD proteins**

**(see Fig. 2 and Fig. 3c).** (**a**) Experimental setup illustrating the incubation order and time of proteins and methylated PCR product prior slot blotting. To be able to assign each of the four groups to the individual slots of the slot blots, they were labeled as A, B, C and D, respectively. (**b**) One representative full slot blot for IDTRD, Mecp2, Mbd2 and MBD. Unmarked slots are unrelated to the figure.



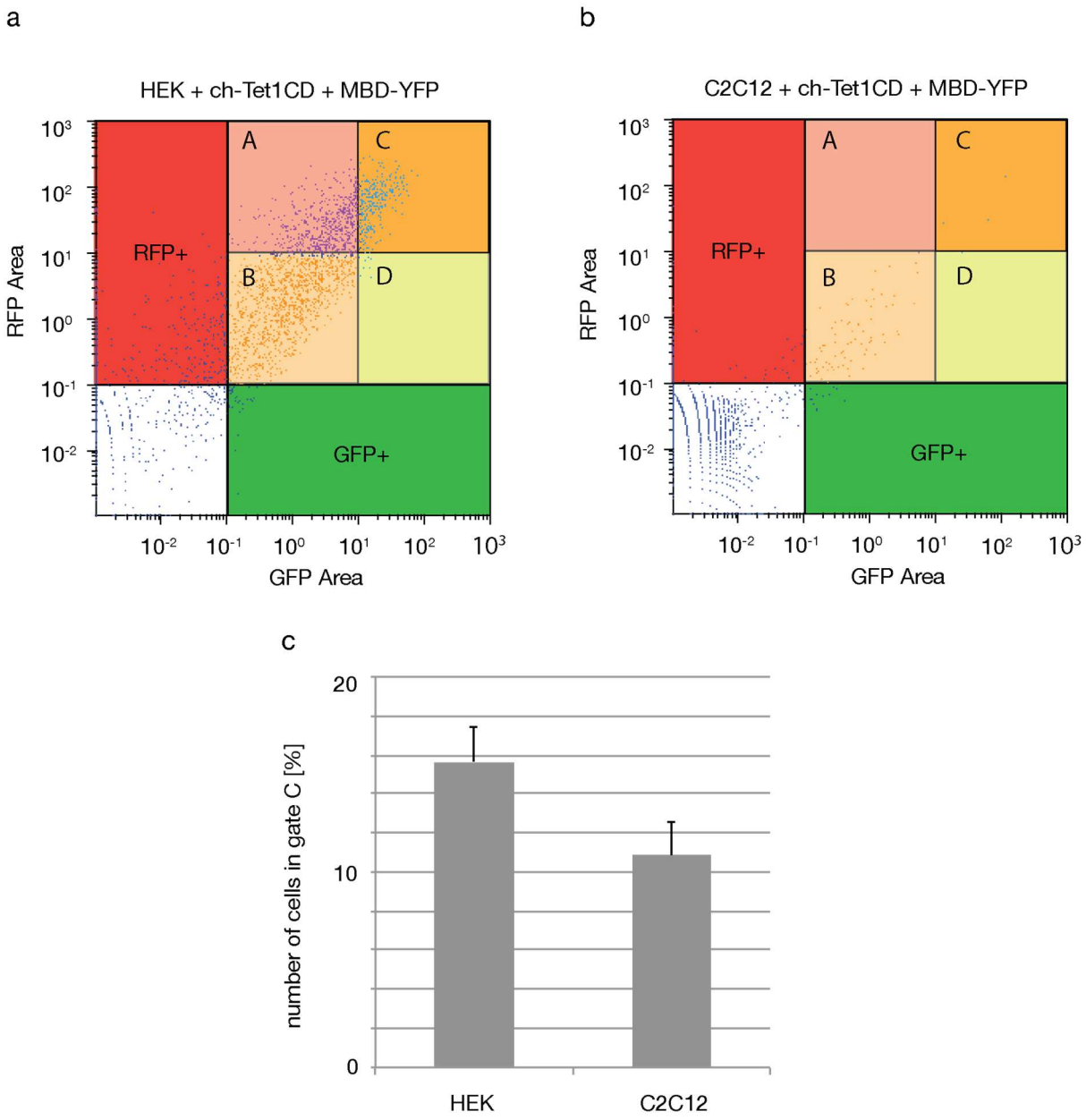

**Supplementary Figure 9. Influence of the chronological DNA binding order on the protecting ability of MBD proteins over time**

Quantification of remaining 5mC levels in single 5mC-containing oligos after simultaneous and successive incubation with Tet1CD- and Mecp2 proteins by MspI digest. (**a**) Experimental setup illustrating the incubation order and time of proteins and oligos prior MspI restriction digest. When Tet1 oxidizes 5mC in the context of CCGG, the cleavable MspI site becomes uncleavable. (**b**) Diagram shows relative 5mC levels of single 5mC containing oligos after incubation with Tet1CD and Mecp2 (n=2). Incubation of Tet1CD with methylated oligos in the absence of its cofactor α-ketoglutarat was used as a negative control (grey line). Shown are mean values +/-SD.




**Supplementary Figure 10. Contribution of direct and long-lasting 5mC binding to the protection of 5mC from Tet1CD catalyzed oxidation**

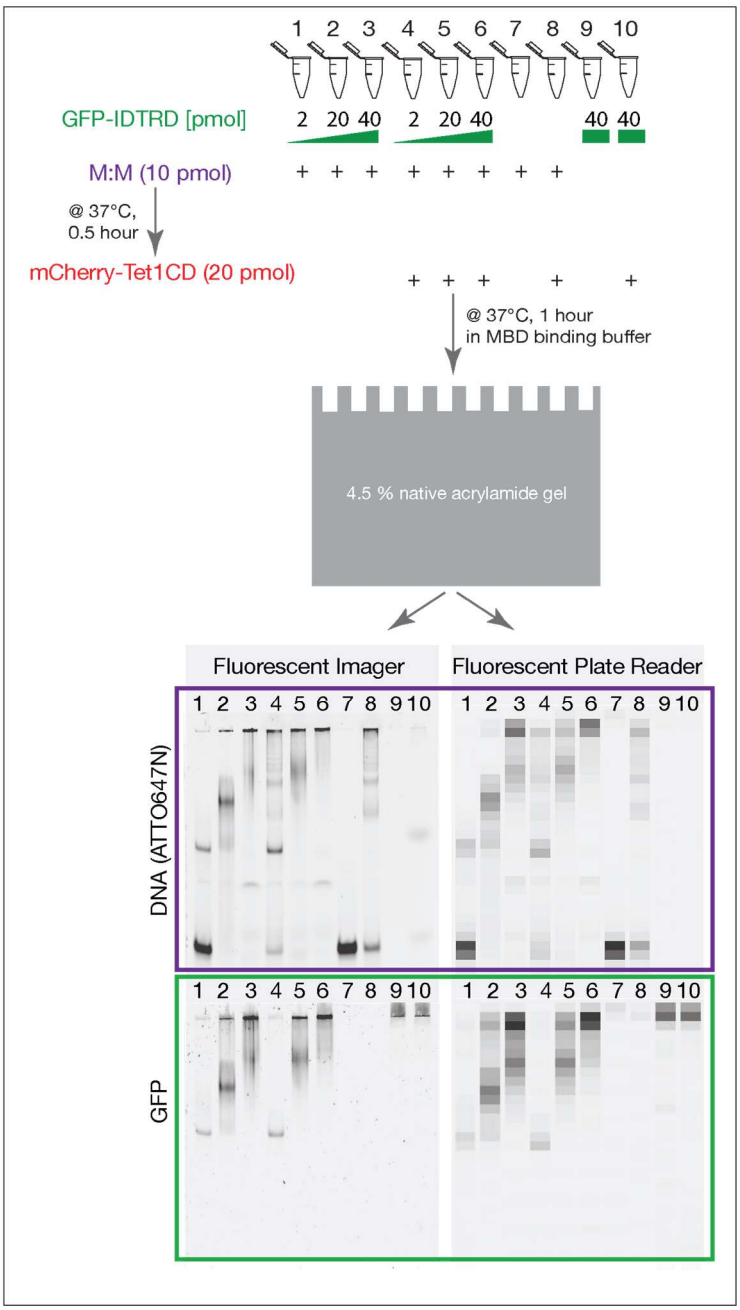
(**a**) *In situ* staining and quantification of genomic 5hmC levels in transiently transfected C2C12 mouse myoblasts. Shown are relative 5hmC (=cyan) signals as a function of Tet1CD (=magenta) level represented as gradient bars. (**a, left**) Error bars represent standard deviation. Mean values of two independent experiments are plotted (Tet1CD+GFP, n=2255; Tet1CD+Mecp2, n=2351; Tet1CD+ R111G, n=3244; n=number of cells). (**a, right**) Mean values of three independent experiments are plotted (Tet1CD+GFP, n=17741; Tet1CD+Mecp2, n=27130; Tet1CD+MaSat, n=15149; Tet1CD+msTALE, n=18593) (**b**) Accumulation kinetics of Mbd2, Mecp2, MBD and IDTRD to pericentric heterochromatin in C2C12 mouse myoblasts. (**c**) Plateau level of accumulation curves. Box plots show the median accumulation of MBD proteins to pericentric heterochromatin (IDTRD, n=14; MBD, n=19; Mbd2, n=6; Mecp2, n=10; n=number of cells), as well as the first and third quartiles. Whiskers extend to 1.5 times the interquartile range.



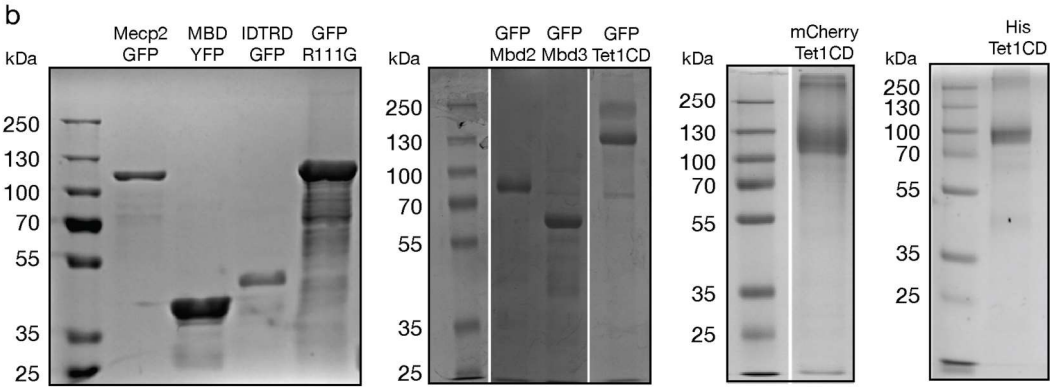

**Supplementary Figure 11. Expression levels of Tet1CD and MBD in HEK versus C2C12 cells**

Flow cytometry analysis of (**a**) HEK and (**b**) C2C12 cells ectopically coexpressing mCherry-tagged Tet1CD and the YFP-tagged MBD domain of Mecp2. Gate A: cells expressing high Tet1CD and low MBD protein levels. Gate B: cells expressing low Tet1CD and low MBD protein levels. Gate C: cells expressing high Tet1CD and high MBD protein levels. Gate D: Cells expressing low Tet1CD and high MBD protein levels. Two independent experiments were performed. Data of one representative experiment are shown. (**c**) Bar diagrams represent the number of cells expressing high Tet1CD and high MBD protein levels (Gate C). Shown are mean values and standard deviation of two independent experiments.

a

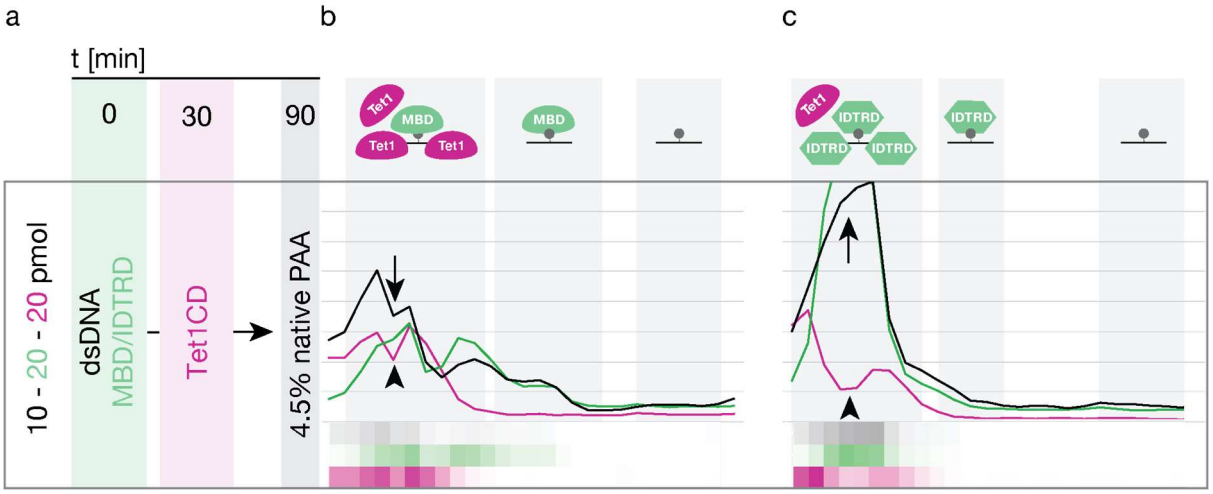

1 2 3 4 5 6 7 8 9 10
GFP-IDTRD [pmol]
2 20 40 2 20 40 40 40
M:M (10 pmol)
+ + + + + + + +
@ 37°C,
0.5 hour
mCherry-Tet1CD (20 pmol)
+ + + + +
@ 37°C, 1 hour
in MBD binding buffer
4.5 % native acrylamide gel
Fluorescent Imager
Fluorescent Plate Reader
DNA (ATTO647N)
GFP
1 2 3 4 5 6 7 8 9 10

b


kDa
Mecp2 GFP
MBD YFP
IDTRD GFP
GFP R111G
250
130
100
70
55
35
25
GFP Mbd2
GFP Mbd3
GFP Tet1CD
mCherry Tet1CD
His Tet1CD

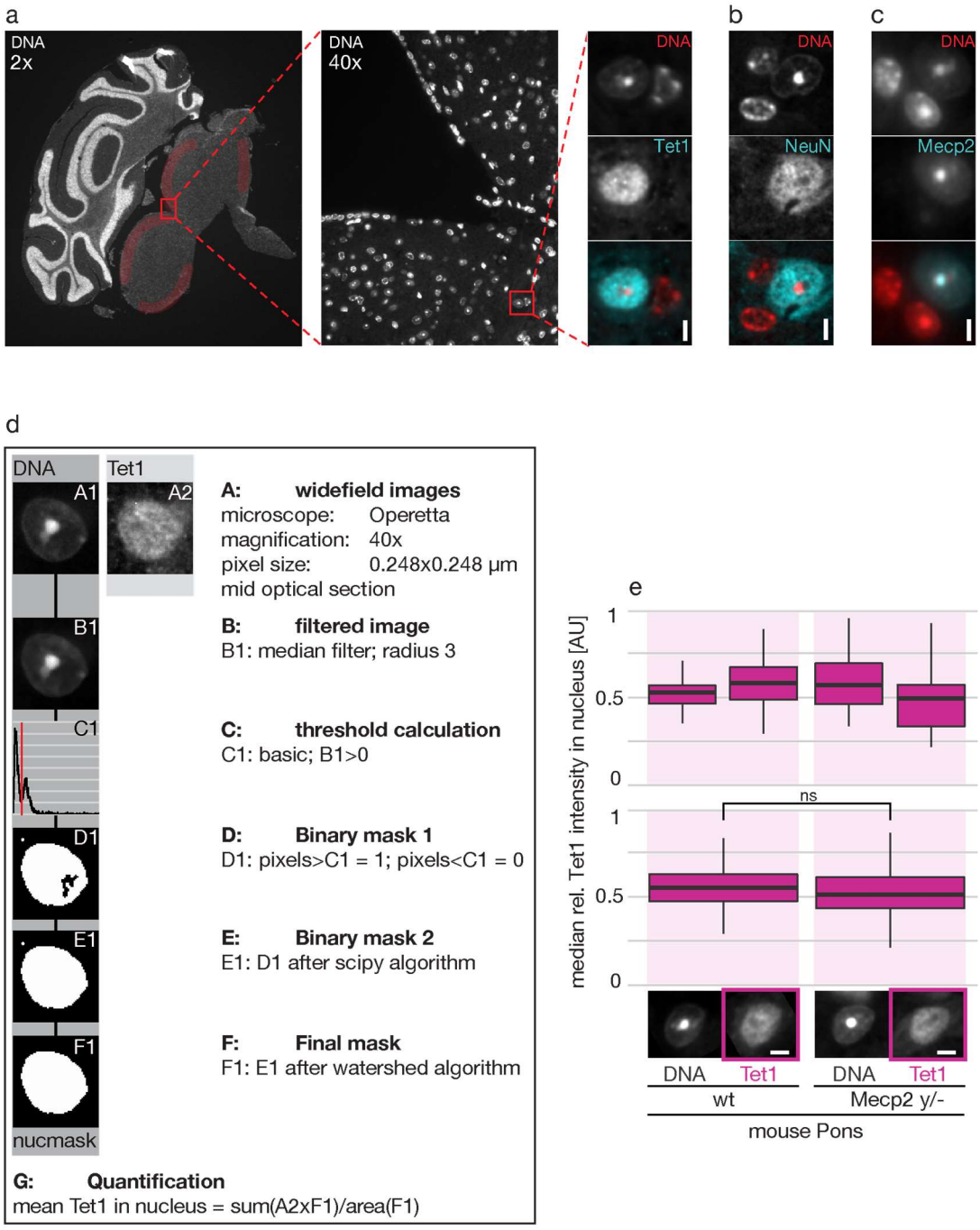
**Supplementary Figure 12. Preparation of various controls required for the implementation and interpretation of Electrophoretic Mobility Shift Assays**

(**a**) Schemes (top) illustrate the workflow of Electrophoretic Mobility Shift Assays (EMSA) fluorescent imager compared to a fluorescent plate reader. Both detection methods lead to the same result and are thus equally suitable for the analysis of EMSA. (**b**) Separation of purified, fluorescently tagged proteins via electrophoresis through a denaturing polyacrylamide gel visualized by Coomassie Brilliant Blue staining.




**Supplementary Figure 13. Impact of 5mC-specific and sequence-unspecific DNA binding proteins on the DNA binding ability of Tet1CD proteins**

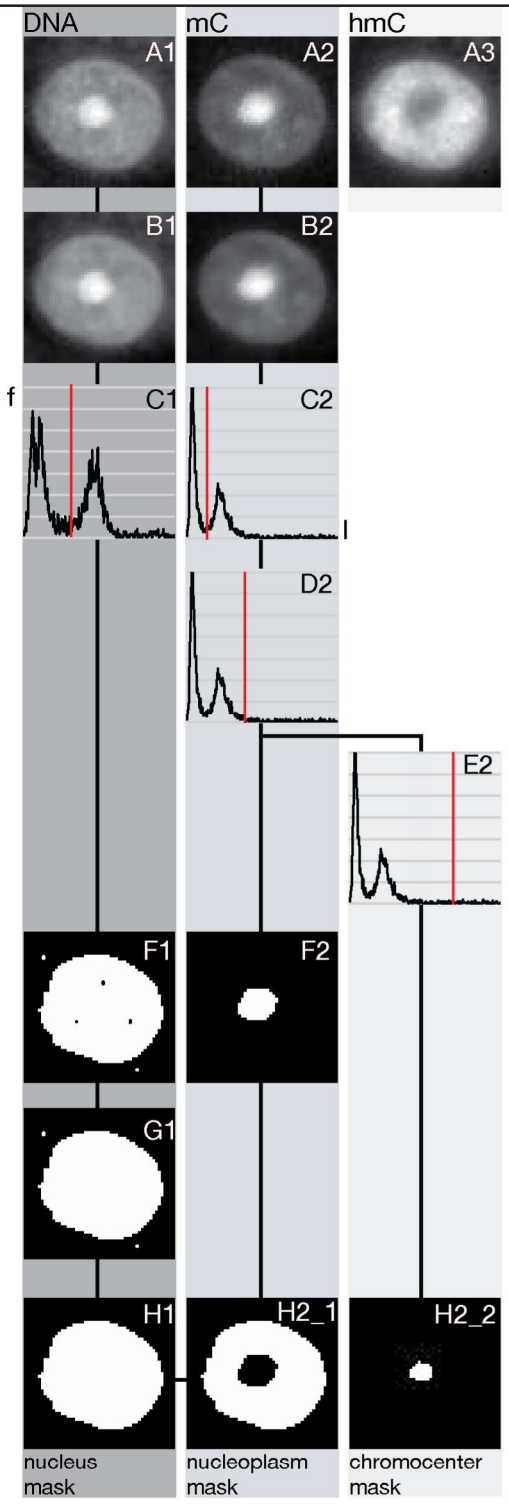
Electrophoretic mobility shift assay (EMSA) to determine the binding ability of fluorescently tagged Tet1CD to double-stranded, single mC containing DNA (ATTO647 labeled) in the presence of equimolar amounts of (**b**) 5mC specific (fluorescently tagged MBD) and (**c**) sequence-unspecific (fluorescently tagged IDTRD) DNA binding domain proteins, respectively. (**a**) Experimental setup illustrating the amount, as well as the incubation order and time of proteins and DNA prior to EMSA. (**b**) Separation of MBD-Tet1CD-dsDNA (n=3), as well as (**c**) IDTRD-Tet1CD-dsDNA (n=3) complexes via electrophoresis through a native polyacrylamide gel visualized using a fluorescent plate reader (see also **Suppl. Fig. 12**). (**b and c**) Running direction is from left (- pole) to right (+ pole). Arrow points to MBD and IDTRD protein/DNA complexes. Arrowhead points to Tet1CD/DNA complexes.




**Supplementary Figure 14. Selection of neural cells for the quantification of Tet1, Mecp2, 5mC and 5hmC levels in wild type and Mecp2 knockout mouse pontes.** (**a**) Representative low (2x) and high (40x) magnification scan of a coronal cross section through mouse brain used for quantification of nuclear Tet1 protein levels in single neural cells within the red-shaded regions of wild type and Mecp2 y/- pons. Tet1 proteins were detected immunologically using a Tet1-specific antibody. Neuronal cells were selected according to morphology properties, which were determined on (**b**) equivalent cross sections stained for Neuronal nuclear antigen (NeuN). 5mC, 5hmC and (**c**) MeCP2 levels were stained and quantified in a similar manner. (**a-c**) DNA was counterstained with DAPI.

Scale bar, 5 µm. (**d**) Visual summary of the steps used for quantification of nuclear Tet1 protein levels in neural cells of wild type and Mecp2 knockout mouse pontes. (**e**) Quantification result of (**d**). Box plots represent the distribution of nuclear Tet1 levels in neurons of two individual (top) and combined (bottom) wild type and Mecp2 y/- mouse pontes (wt1 n=29, wt2 n=64, KO1 n=56, KO2 n=43; from left to right; n=number of cells), respectively. Plotted is the median, as well as the first and third quartiles. Whiskers extend to 1.5 times the interquartile range. *P* values were calculated by Wilcoxon signed-rank test (ns = non significant). Mid-confocal optical sections represent NeuN positive cells of wild type and Mecp2 y/- mouse pontes immunostained for Tet1. DNA was counterstained with DAPI. Scale bar, 5 µm.

a




**A: widefield images**
microscope: Operetta
magnification: 40x
pixel size: 0.248x0.248 µm
mid optical section

**B: filtered images**
B1: median filter; radius 3
B2: median filter; radius 3

**C: threshold calculation 1**
C1: basic; B1>0
C2: basic; B2>0

**D: threshold calculation 2**
D2: basic; B2>C2

**E: threshold calculation 3**
E2: basic; B2>D2

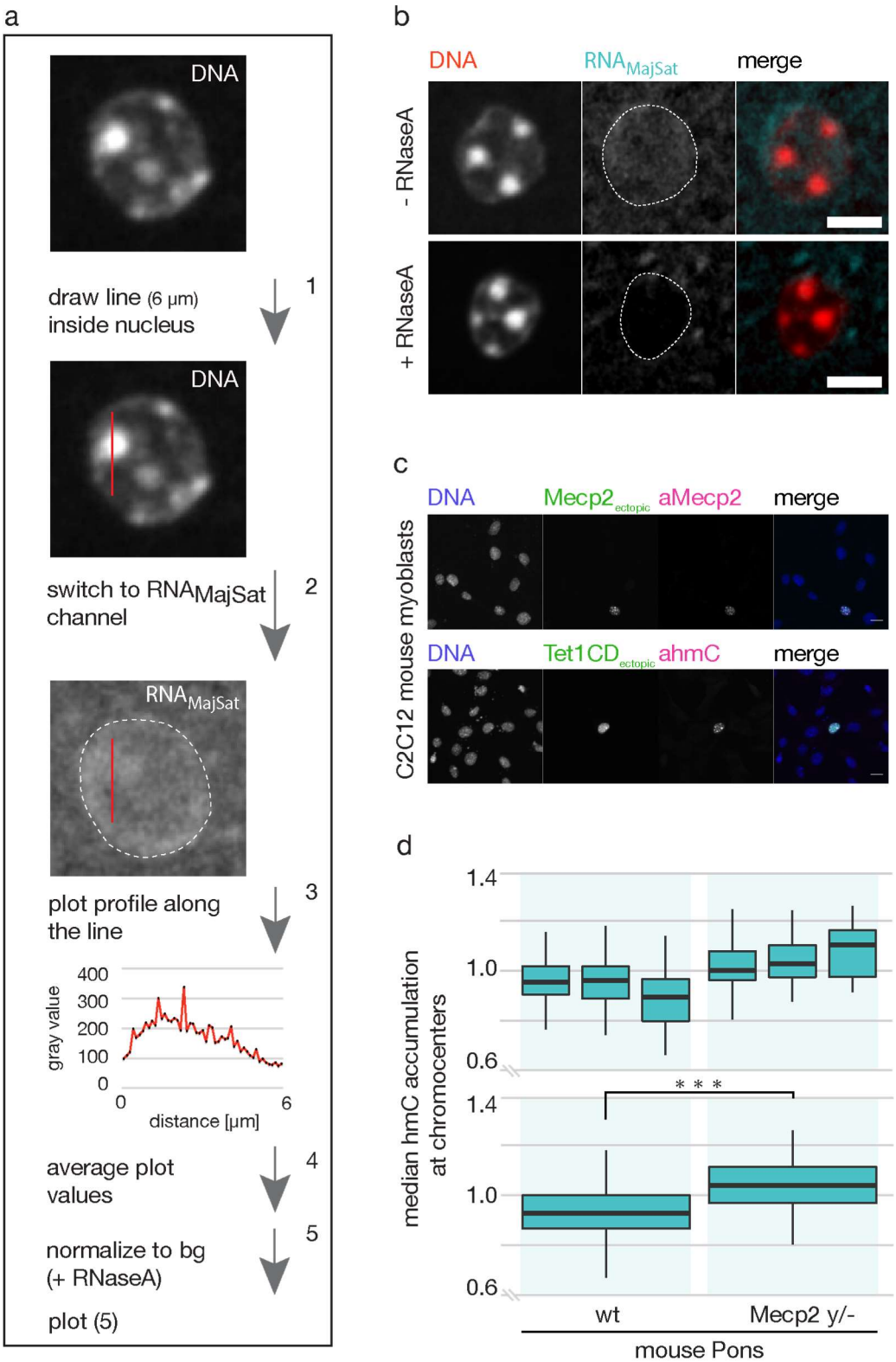
**F: Binary masks 1**
F1: pixels>C1 = 1; pixels<C1 = 0
F2: pixels>D2 = 1; pixels<D2 = 0

**G: Binary masks 2**
G1: F1 after scipy algorithm

**H: Final masks**
H1: G1 after watershed algorithm
H2_1: H1-F2
H2_2: pixels>E2 = 1; pixels<E2 = 0

**I: Quantifications**
For Figure 5:
mean hmC at CC = sum(A3xH2_2)/area(H2_2)
mean mC at CC = sum(A2xH2_2)/area(H2_2)
For Supplementary Figure 16d:
accumulation of hmC at CC = (sum(A3xH2_2)/area(H2_2))/(sum(A3xH2_1)/area(H2_1))

**Supplementary Figure 15. Visual summary of the steps used for quantification of Mecp2, 5mC and 5hmC levels at various different nuclear regions in neural cells of wild type and Mecp2 knockout mouse pontes.** (**a**) Representative summary illustrating the quantification of 5mC and 5hmC levels at pericentric heterochromatin (chromocenter, CC), as well as the accumulation of 5hmC at chromocenters. DNA was counterstained with DAPI. For the quantification of Mecp2 levels, mouse pontes were immunostained for 5mC and Mecp2.




**Supplementary Figure 16. Detection and quantification of major satellite transcripts via RNA FISH and RT qPCR, respectively.** (**a**) Visual summary of the different steps used for quantification of Major Satellite RNA transcripts detected by RNA FISH. (**b**) Mid-optical sections of RNaseA treated and untreated neural cell nuclei hybridized with a probe specific for major satellite repeats. The lack of signal in RNaseA treated cells proves specificity of the probe for major satellite RNA. (**c**) Mid-optical sections of transiently transfected (Mecp2 and Tet1CD, shown in green) C2C12 mouse myoblasts

immunostained for Mecp2 and 5hmC (shown in magenta), respectively. The lack of signal in untransfected cells verifies the absence of Mecp2 and 5hmC in C2C12 mouse myoblasts. (**b**,**c**) DNA was counterstained with DAPI. Scale bar, 5 µm. (**d**) Accumulation of 5hmC at chromocenters in neurons of two individual (top) and combined (bottom) wild type and Mecp2 y/- mouse pontes (wt1 n=29, wt2 n=26, KO1 n=30, KO2 n=30; from left to right; n=number of cells), respectively. Plotted is the median, as well as the first and third quartiles. Whiskers extend to 1.5 times the interquartile range. *P* values were calculated by Wilcoxon signed-rank test (***; $P < 0.001$).