

# CMTR1 is recruited to transcription start sites and promotes ribosomal protein and histone gene expression in embryonic stem cells

Shang Liang<sup>1</sup>, Joana C. Silva<sup>1</sup>, Olga Suska<sup>1</sup>, Radoslaw Lukoszek<sup>1,2</sup>, Rajaei Almohammed<sup>1</sup> and Victoria H. Cowling<sup>1,\*</sup>

<sup>1</sup>Centre for Gene Regulation and Expression, School of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, UK and <sup>2</sup>MRC Protein Phosphorylation and Ubiquitylation Unit, School of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, UK

Received June 17, 2021; Revised January 09, 2022; Editorial Decision February 03, 2022; Accepted February 09, 2022

## ABSTRACT

**CMTR1 (cap methyltransferase 1) catalyses methylation of the first transcribed nucleotide of RNAPII transcripts (N1 2'-O-Me), creating part of the mammalian RNA cap structure. In addition to marking RNA as self, N1 2'-O-Me has ill-defined roles in RNA expression and translation. Here, we investigated the gene specificity of CMTR1 and its impact on RNA expression in embryonic stem cells. Using chromatin immunoprecipitation, CMTR1 was found to bind to transcription start sites (TSS) correlating with RNAPII levels, predominantly binding at histone genes and ribosomal protein (RP) genes. Repression of CMTR1 expression resulted in repression of RNAPII binding at the TSS and repression of RNA expression, particularly of histone and RP genes. In correlation with regulation of histones and RP genes, CMTR1 repression resulted in repression of translation and induction of DNA replication stress and damage. Indicating a direct role for CMTR1 in transcription, addition of recombinant CMTR1 to purified nuclei increased transcription of the histone and RP genes. CMTR1 was found to be upregulated during neural differentiation and there was an enhanced requirement for CMTR1 for gene expression and proliferation during this process. We highlight the distinct roles of the cap methyltransferases RNMT and CMTR1 in target gene expression and differentiation.**

## INTRODUCTION

The RNA cap is added to the initiating nucleotide of RNA pol II (RNAPII) transcripts. It protects RNA from degradation by blocking exonucleases during transcription and identifies RNA as cellular in origin or 'self', protecting

it from degradation by proteins of the innate immunity pathways (1–3). The most prevalent cap structure, denoted <sup>m7</sup>GpppX<sub>m</sub>, includes 7-methylguanosine (<sup>m7</sup>G) and first nucleotide ribose O2-methylation (X<sub>m</sub> or N1 2'-O-Me) (4,5). N1 2'-O-Me prevents RNA interactions with proteins of the innate immune pathways, e.g. IFIT proteins and RIG-I, which bind to the cap in the absence of N1 2'-O-Me and direct RNA degradation (3,6,7). Proteins involved in splicing, nuclear export, RNA processing and translation initiation are also recruited to RNA via the cap (8,9). The role of N1 2'-O-Me in these gene expression processes is unclear; it may influence the recruitment of cap binding proteins which mediate processing events, including translation, although mechanistic information is lacking (10,11). Recent findings demonstrated superior translation of capped RNA with N1 2'-O-Me in certain cell types, although this did not correlate with binding of the translation initiation factor eIF4E, indicating the influence of other factors (12).

Following transcription initiation, a series of enzymes catalyse RNA cap formation, starting with guanosine linkage to the first transcribed nucleotide by a 5' to 5' triphosphate bridge (1,8). This cap guanosine is methylated at the N-7 position, the first and second transcribed nucleotides are methylated on the ribose at the O-2 position, and if the first nucleotide is adenosine it may also be methylated at the N-6 position (13). In mammals, cap guanosine N-7 methylation and first transcribed nucleotide ribose O-2 methylation are catalysed by the cap methyltransferases RNMT and CMTR1, respectively, to form the predominant cap structure, <sup>m7</sup>GpppX<sub>m</sub> (4,5).

The N1 2'-O-Me methyltransferase, CMTR1 (ISG95, FTSJD2, KIAA0082) (14,15), was first identified in screens for interferon stimulated genes and later was characterized as a cap methyltransferase (16–19). As an interferon stimulated gene (ISG), CMTR1 has been found to be important for the expression of a subset of other ISGs (20). *In vivo*, CMTR1 is required for neuromorphogenesis and brain

\*To whom correspondence should be addressed. Tel: +44 1382 386997; Fax: +44 1382 386997; Email: v.h.cowling@dundee.ac.uk

development, in particular promoting Camk2a expression (21). The mechanisms directing these gene-specific impacts of CMTR1 are not known. CMTR1 has a central methyltransferase core surrounded by several protein:protein interaction domains (14,15,22,23). It is recruited directly to RNAPII C-terminal domain (CTD) via an interaction with the WW domain and this interaction increases when the CTD is phosphorylated on serine-5 of the heptad repeats (22,23). Approximately half of the CMTR1 in the nucleus is present in a complex with the helicase DHX15 which influences RNA translation (23,24). DHX15 represses the catalytic activity of CMTR1 but does not bind to RNAPII or influence the RNAPII–CMTR1 interaction.

In order to investigate the role of CMTR1 in gene expression, ChIP-seq analysis was performed in murine embryonic stem cells (ESCs). The highest levels of CMTR1 were found bound to the ribosomal protein (RP) genes and histone genes. For these genes, CMTR1 binding correlated with phosphorylation of Serine-5 of RNAPII CTD at the transcription start site (TSS). Repression of CMTR1 resulted in loss of RNAPII at the TSS and repression of ribosomal protein and histone gene expression. Conversely, the addition of recombinant CMTR1 to isolated nuclei resulted in increased transcription of histone and RP genes. In correlation with regulation of histones and RP genes, repression of CMTR1 resulted in induction of markers of DNA replication stress and damage, and repression of RNA translation. There was enhanced dependency on CMTR1 for gene expression and cell proliferation when ESC were induced to differentiate.

## MATERIALS AND METHODS

### Cell culture and differentiation

Murine ESCs (Embryonic Stem cells) were cultured in Glasgow's Minimum Essential Medium (Life Technologies), 10% Knockout Serum Replacement (Invitrogen), 1% Non-Essential Amino Acids (Life Technologies), 2 mM L-Glutamine (Life Technologies), 1 mM sodium pyruvate (Life Technologies), 0.35  $\mu$ M  $\beta$ -mercaptoethanol (Sigma), 100 U/ml recombinant LIF (DSTT, Division of Signal Transduction Therapies, University of Dundee). ESCs were cultured in 0.1% gelatin-coated flasks at 37°C/5% CO<sub>2</sub>. For LIF withdrawal-induced differentiation,  $1 \times 10^6$  cells were seeded on 0.1% gelatin-coated 10 cm plates in ESC medium minus LIF with medium replacement every two days. For neural differentiation,  $1-1.6 \times 10^6$  cells were seeded on 0.1% gelatin-coated 10 cm plates in N2B27 medium composed of 50% Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12)/50% Neurobasal Medium (Life Technologies), supplemented with 1% B-27 supplement (Life Technologies), 0.5% N-2 supplement (Life Technologies), 2 mM L-Glutamine (Life Technologies) and 0.35  $\mu$ M  $\beta$ -mercaptoethanol (Sigma) (25). Medium was replaced every day.

### CRISPR/Cas9 genome targeting

Paired guide RNAs (gRNAs) targeting the end of *Rnmt* and *Cmtr1* gene exon 2 were designed (Supplementary Figure S1). List of gRNAs used for editing (Table 1):

**Table 1.** CRISPR guides

CRISPR guide target	Sequence
<i>Rnmt</i> exon 2 HA-GFP KI as	CAGCGAAAGCGTCAGTCGCTTCT
<i>Rnmt</i> exon 2 HA-GFP KI s	ATGAGCCCCGCTGCTGCTAGCGGT
<i>Cmtr1</i> exon 2 HA-GFP KI as	TCCAGAATGTACAGCTCCCCTCA
<i>Cmtr1</i> exon 2 HA-GFP KI s	AGCGAATTGGGGAGCTTGCCCGT
<i>Cmtr1</i> exon 2 KD as	CCTTTTGCTGCATGATTGACAGA
<i>Cmtr1</i> exon 2 KD s	CCACATCTGACGATGAACCTCTC

Antisense gRNAs were cloned into pX335-Cas9-D10A. Sense gRNAs were cloned into pBARED-Puro-U6. ESC cells at ~70% confluency were transfected using Lipofectamine 2000 (Life Technologies) and 1  $\mu$ g gRNAs vectors, 3  $\mu$ g donors. Cells were selected using 1.5  $\mu$ g/ml puromycin for 2 days. HA-GFP-Rnmt/*Cmtr1* KI cells were sorted and GFP positive cells retained. Cells were plated at reduced density in medium supplied with 20% pre-conditioned medium. Single cell colonies were expanded and protein and DNA analysed. For HA-GFP KI ESCs, clones showing no wild-type RNMT/CMTR1 band, but with product approximately 30 kDa larger than the wild-type protein were selected for the genotyping. For *Cmtr1*<sup>KD</sup> ESCs, clones exhibiting significant CMTR1 protein reduction were selected for genotyping. About 200 ng genomic DNA was used in genotyping. Primers were designed to ensure correct insertion of HA-GFP encoding cassette for HA-GFP KI ESCs and knockout of *Cmtr1* for *Cmtr1*<sup>KD</sup> ESCs (Table 2). For HA-GFP-CMTR1 KI ESCs, one *Cmtr1* allele has the designed mutation to create the fusion protein and the other allele has a deletion which prevents expression. Similarly the *Cmtr1*<sup>KD</sup> cells have one allele with the KD mutation and one allele with a deletion which prevents expression. To generate ESCs stably expressing HA-CMTR1 in the *Cmtr1*<sup>KD</sup> background,  $1 \times 10^6$  *Cmtr1*<sup>KD</sup> cells were transfected with 2  $\mu$ g pPyPCAGGS vector (empty-vector or HA-CMTR1) using Lipofectamine 2000 (Life Technologies). Successfully transfected cells were selected with puromycin (1  $\mu$ g/ml).

### siRNA transfection

siRNA oligos were from Dharmacon, siGENOME Mouse collection (Table 3). Transfections performed with 50  $\mu$ M siRNAs using Lipofectamine RNAiMAX (Life Technologies). For RNA-seq/protein analysis, siRNA transfection mix was added to  $3-4 \times 10^5$  freshly seeded cells on a 6-well plate. For nuclear run-on and ChIP-seq, siRNA transfection mix was added to  $3-4 \times 10^6$  freshly seeded cells on a 10 cm plate. Cells were cultured for 36 h (medium replaced after 12 h).

### Proteasome inhibition with MG132

$2-3 \times 10^5$  cells were seeded on 6-well plates one day before the treatment. Culture medium was replaced with ESC medium supplemented with 5  $\mu$ M of MG132 (Calbiochem) or mESC medium with DMSO (as controls). MG132-treated cells were collected after 2, 4 and 6 h of treatment.

**Table 2.** PCR primers

Target	Forward primer	Reverse primer
<b>ChIP-qPCR targets</b>		
<i>Actin</i> TSS	CCGAAAGTTGCCTTTTATGG	AAGGAGCTGCAAAGAAGCTG
<i>Sox2</i> TSS	CCCATTATTCCTGACAGC	GGGCTCCAAACTTCTCTCT
<i>Hist2h4</i> TSS	CAATCTGGTCCGATCCTCTCA	CCCTTCCACCCTTCCCTCT
<i>Hba-a1</i> TSS	GGGTAGAGCAAGCACAAACC	AGAAGTGTCTTGGACCTGCA
<i>Rpl10a</i> TSS	CTTCAGCCCTTGTCCCAATG	TCAAATAGAGCCCGCGTTC
<i>Hist1h3f</i> TSS	TCCGGATAGTAAACTGCTCCC	TCCTCAGCCACATTGTTACCA
<i>Hist1h2ai</i> TSS	TGAGTTATGTCTGGACGCGG	GCTAAGGAGTAGTGGCCCTTG
<i>Rps27</i> TSS	CGGCTCTGGTTACTGTCCTA	CCTTGGTCCAGGTTTCGACA
<i>Rps29</i> TSS	CAACAATGGCCCTTCCCTCT	CGTCTGAAGGCAAGATGGGT
<b>RT-qPCR targets</b>		
Actin	CGCCACCAGTTCGCCAT	CTTTCACATGCCGAGC
Cmtr1	CCCAGAAGTCCGTCTATGCT	CATCTCCCCACTCCCAGAAG
Rnmt	GCAGGCGGATACGATGCTTA	CTTTGCCATTATCAGTTCAAAGCTA
Ram	AGTCGGGTGGTTGAGGGATT	TTAAGATCCTTCATGGGACGCAC
Sox2	CAGGAGTTGTCAAGGCAGAGA	CTTAAGCCTCGGGTCCAAA
Oct4	GAGACTTTGCAGCCTGAGGG	CTTTCATGTCCTGGGACTCCTC
Nanog	AAAGGATGAAGTGAAGCGG	GTGCTGAGCCCTTCTGAATC
Klf4	AGAACAGCCACCCACACTTG	GTGGTAAGGTTTCTCGCCTGT
Brac	GCCAAAGAAAGAAACGACCAC	ATTGGGAATACCCCGCTG
Pax6	CCATCACCAATCAGCATAGGAATC	AGGAGTTGCTGGTGAGAGTT
Sox1	GCGCCATGAAATCACCG	GCCCCGGCTGCCATAAA
Sox17	TCCTGGAAGCGCTAAGGAAA	GTGATTGTGCGGAGCAAGTC
Gapdh	AAGGGCTCATGACCACAGTC	ATCACGCCACAGCTTCCA
Poldip3	TCCCTCCCAAGCAGATGAAG	GCGGTGTAGGCATCATTCTG
Srsf2	AAGTCCAGATCTGCCGAAG	ACCTAGACTTGGACTCTCGC
Hist1h3f	TCCGGATAGTAAACTGCTCCC	TCCTCAGCCACATTGTTACCA
Hist1h2ai	TGAGTTATGTCTGGACGCGG	GCTCCGAGTAGTGGCCCTTG
Hist2h2bb	ACAAGGTCTGAAGCAAGTG	GCTTGTGTAAATGCGCCAGA
Rps9	AACGTGAGGTCTGGAGGGTC	GGTCTGCAGGCGTCTCTTA
Rps27	TGCAGAGCCCCAATTCCTAC	AGCAGCCAACACACAAGACT
Rps29	GCGTTGTATTCTGGGTAGCG	ACCCATCTTGCCCTCAGACG
<b>NRO-qPCR targets</b>		
Actin	CAGCCAACTTTACGCCTAGC	GACCCTAGTGTGTCCCAAG
Hist1h3f	TCCGGATAGTAAACTGCTCCC	TCCTCAGCCACATTGTTACCA
Hist1h2ai	TGAGTTATGTCTGGACGCGG	GCTCCGAGTAGTGGCCCTTG
Rpl10a	GAGAACGCGGGCTCTATTTG	AGCGCCACGATAACAACCTAC
Rps27	CGGCTCTGGTTACTGTCCTA	CCTTGGTCCAGGTTTCGACA
Rps29	CAACAATGGCCCTTCCCTCT	CGTCTGAAGGCAAGATGGGT
Nanog	TCCCCACAGTTTGCCTAGTT	TTCTCGGGATGAAAACTGC
<b>CRISPER/Cas9 genome editing screening</b>		
<i>Rnmt</i> exon 2 KI	ACATTGCTTCAGAATCGTGTCCG	GGGTGTCTCAGGTAGACAGATGAT
<i>Cmtr1</i> exon 2 KI	GTTTGCTGCAGGTTCTCTACAGG	CAGCTGAGTATTTCTGATTCCAAC
<i>Cmtr1</i> exon 2 KD	ATTCTGAGAGCACCTCATTTCCTAG	TTTAAGGCAGTGAGGAGGTACC

**Table 3.** siRNA

siRNA	Cat. No	Target sequence	Application
CMTR1 si1	D-060995-01-0005	CAAAGGAGACCAUGAAUUU	WB, RNAPII ChIP-seq, RNA-seq
CMTR1 si2	D-060995-04-0005	GAUGGUUUAUGAAGAUGAA	WB, RNA-seq
Ct siRNA	D-001210-04-20	AUGAACGUGAAUUGCUCAA	WB, RNAPII ChIP-seq, RNA-seq

### Cell lysate preparation and western blot

2–3 × 10<sup>6</sup> cells were seeded on 10 cm plates two days before lysis. For protein analysis, cell lysates were prepared in 10 mM Tris-HCl pH 7.0, 50 mM NaCl, 30 mM sodium pyrophosphate, 50 mM NaF, 5 μM ZnCl<sub>2</sub>, 10% glycerol, 0.5% Triton X-100, 1 mM DTT, 10 mM leupeptin, 1 mM pepstatin, 0.1 mg/ml aprotinin and phosphatase inhibitor 2 and 3 cocktail (Sigma) as previously described (26). 1 μg/μl protein lysates in Laemmli sample buffer supplemented with 0.025 M DTT were boiled for 5 min. 6–10 μg protein was resolved on a polyacrylamide gel. PVDF (Millipore) was used for electroblotting. Membranes were blocked for 1 h before overnight incubation with primary antibodies at 4°C

(Table 4). Membranes were washed and HRP-conjugated secondary antibodies (Thermo Fisher) added for 1 h before imaging using SuperSignal West Pico Chemiluminescent Substrate (Thermo Fisher).

### Immunoprecipitation

For immunoprecipitation using endogenous antibody, cell lysates were prepared as previously described (26). 2 μg of anti-CMTR1 (Sigma HPA029979) or anti-RNMT (DSTT) antibody or IgG control (DSTT) was added to 1–2 mg of lysates for overnight incubation at 4°C before the incubation with protein G Dynabeads (Life Technologies) for additional 2 h (Table 4). Beads were washed and eluted as de-

**Table 4.** Antibodies

Antibody antigen	Type	Supplier/ Cat. no
RNMT	Sheep polyclonal	DSTT
CMTR1	Rabbit polyclonal	Sigma HPA029979
RAM	Sheep polyclonal	DSTT
Sox2	Mouse monoclonal	CST 4900
Oct4	Rabbit polyclonal	Abcam ab18976
Nanog	Rabbit polyclonal	Abcam ab80892
Klf4	Goat polyclonal	R&D AF3158
Sox1	Rabbit polyclonal	CST 4194
DNMT3B	Rabbit polyclonal	Abcam ab122932
GFP	Mouse monoclonal	Roche 11814460001
HA	Mouse monoclonal	BioLegend 901501
RNAPII total (Rpb1)	Rabbit monoclonal	CST 14958
RNAPII S5P	Rat monoclonal	Chromotek 3E8
RNAPII S2P	Rat monoclonal	Chromotek 3E10
Actin	Mouse monoclonal	Abcam ab3280
DHX15	Rabbit polyclonal	Abcam ab254591
UBC	Rabbit polyclonal	Dako Z0458
Chk1	Mouse monoclonal	Santa Cruz sc-56291
p-Chk1	Rabbit monoclonal	CST 2348
$\gamma$ H2AX	Rabbit polyclonal	CST 2577

scribed (27). For HA immunoprecipitation, cells were lysed in PBS, 1 M NaCl, 0.5% v/v NP-40 as described (27). About 20  $\mu$ l anti-HA magnetic beads (Pierce) were added to lysates for overnight incubation at 4°C. Beads were washed and eluted in TE/1% SDS for 30 min with rotation at room temperature (27).

#### Immunofluorescent staining

Immunofluorescence staining was performed using polyclonal rabbit anti-GFP (1/250, MBL International 598), polyclonal rabbit anti-RNMT (1/250, DSTT) and polyclonal sheep anti-CMTR1 (1/50, DSTT) antibodies (Table 4). Cells plated on Ibidi 8-well chamber slides were fixed with 4% paraformaldehyde for 20 min, permeabilized with 1% Triton/PBS for 5 min and blocked with 5% donkey serum/PBS for 1 h at room temperature. Primary antibodies diluted in either blocking buffer or 2.5% BSA/PBS were incubated with cells overnight at 4°C. Immunofluorescence labelling was performed using appropriate secondary antibodies conjugated to Alexa fluor 594 dyes (Life Technologies). Stained cells were mounted in Ibidi mounting medium with DAPI prior to imaging with a Leica SP8 confocal microscope.

#### Annexin V staining

$10^6$  cells were seeded on 10 cm plates in mESC medium or mESC medium without LIF. Every 24 h, cells were trypsinized, washed and resuspended in Fixable Viability Dye eFluor 780 (eBioscience, Life Technologies) for 10 min on ice. Cells were stained with 5  $\mu$ l of APC-Annexin V (Biolegend) in 45  $\mu$ l of 1 $\times$  Annexin V binding buffer (10 mM Hepes [pH 7.4], 140 mM NaCl, 2.5 mM CaCl<sub>2</sub>) for 10 min at room temperature, protected from light. Cells were subject to flow cytometry with a BD LSR Fortessa Cell Analyser System and analysed using FlowJo software.

#### Alkaline phosphatase staining

Cells were seeded on 6-well plates at a concentration of 70 cells/cm<sup>2</sup>. After 3 days, cells were washed in PBS and medium was replaced with 10-fold decreasing concentrations of LIF (from 100 to 0 units/ml). After 6 days, cells were fixed and stained with Alkaline Phosphatase Staining Kit (Sigma). Colonies were counted under a Zeiss Axiovert 40C microscope and scored into three distinct categories: stained (pluripotent), mixed and unstained (differentiated).

#### <sup>35</sup>S cysteine and methionine cell labelling

60–70% confluent cells from 6-well plates were used for the labelling. Cells were incubated with mESC medium supplemented with 0.8 Mbq/ml of <sup>35</sup>S Protein Labeling Mix (Perkin Elmer) at 37°C with 5% CO<sub>2</sub> for 30 min. Cells were lysed, lysates were loaded into SDS-PAGE gels as described in the previous section. Then the gel was fixed with 30% methanol and 7% acetic acid for 40 min. The fixed gel was dried using a vacuum gel dryer for 1 h. Then the dried gel was exposed to a phosphor plate before the scanning using Typhoon imager (GE).

#### Detection of mRNA cap structures by mass spectrometry

Performed according to (4).

#### ChIP

2–3  $\times 10^6$  cells were seeded on a 10 cm plate one or two day(s) before the ChIP experiment. Cells were cross-linked with 1% formaldehyde for 10 min. Glycine was added to quench the reaction. The cell lysates were prepared as described (27). Chromatin was sonicated in Bioruptor (Diagenode) at high power for 20 min (30 s on/off, 20 cycles). About 25–50  $\mu$ g sonicated chromatin was used for each IP (500  $\mu$ l). About 50  $\mu$ l (10%) from each diluted extract for IP was taken as the inputs. For ChIP using anti-HA beads, HA IPs from WT E14 cells were used as control IPs. 20  $\mu$ l of anti-HA magnetic beads (Pierce) were added to lysates from KI cells (as IPs) and E14 cells (as control IPs) for 4°C overnight incubation. For ChIP using endogenous anti-CMTR1 antibody, 2  $\mu$ g of anti-CMTR1 antibody (Sigma HPA029979) was added to lysates from *Cmtr1* KI cells and E14 cells. For RNAPII ChIP, 2  $\mu$ g of anti-Rpb1 (CST 14958) antibody was added to lysates from E14 (*Cmtr1*<sup>WT</sup>), *Cmtr1*<sup>KD</sup> or *Cmtr1*<sup>KD</sup> + HA-CMTR1 cells. The endogenous antibodies were added for 4°C overnight reaction before the incubation with protein G Dynabeads (Life Technologies) for additional 2 h. The beads were washed and eluted as described (27) before the proteinase K treatment and reverse crosslinking at 42°C overnight. The eluted DNA was purified by the QIAquick PCR purification kit (QIAGEN).

#### RNA extraction and cDNA synthesis

RNA was extracted using TRIzol Reagent (Thermo Fisher) and concentration determined using a NanoDrop 2000 spectrophotometer. Genomic DNA was removed using the

TURBO DNA-free Kit (Thermo Fisher). 400–500 ng RNA was used in 20  $\mu$ l cDNA synthesis reactions using iScript cDNA Synthesis Kit (Bio-Rad). Three biological replicates were used for the PCR analysis or sequencing (Table 2).

### Nuclear run-on assay

$3 \times 10^6$  cells were seeded on a 10 cm plate one day before the experiment. Nuclear run-on was performed as previously described (27). About 1  $\mu$ g recombinant CMTR1 protein was added to nuclei from a 10 cm plate and incubated for 20 min at 30°C, with 1 mM ATP, 1 mM CTP, 1 mM GTP, 0.5 mM BrUTP, 0.5 mM UTP and 1 U/ml RNasin. After nuclear run-on and RNA extraction, Br-U IP was performed with 8–12  $\mu$ g RNA. Following IP, nascent RNA was purified and subjected to cDNA synthesis and qPCR analysis (Table 2).

### qPCR

5  $\mu$ l qPCR reaction contained 1  $\mu$ l DNA template (undiluted IP or 10  $\times$  diluted input for ChIP-qPCR and nuclear run-on qPCR or 2  $\times$  diluted cDNA for RT-qPCR), 2.5  $\mu$ l SsoFast EvaGreen supermix (Bio-Rad), 0.5  $\mu$ l primer pairs (see list below). Each qPCR reaction was performed in triplicate on CFX384 Touch Real-Time PCR Detection System (Bio-Rad), 3 min denaturation at 95°C, followed by 40 cycles of 10 s at 95°C, 20 s at 58°C and 10 s at 72°C and final incubation at 72°C for 5 min. For ChIP-qPCR, ratio of Ct value of IP to 1% input was calculated to determine ChIP efficiency. For RT-qPCR and nuclear run-on qPCR, the Ct values of target genes were normalized to corresponding Ct value of Actin and or Gapdh in each treatment condition to obtain relative transcript levels (Table 2).

### Library preparation and sequencing

For ChIP-seq, libraries were prepared using NEBNext® Ultra™ II DNA Library Prep Kit and NEBNext® Multiplex Oligos (Index Primers Set 1) for Illumina. For RNA-seq, ribosomal RNA depletion was performed on 1  $\mu$ g RNA using the NEBNext® rRNA Depletion Kit v2 (Human/Mouse/Rat). RNA-seq library preparation was carried out using NEBNext® Ultra™ II Directional RNA Library Prep Kit and indexed with NEBNext Multiplex Oligos for Illumina (96 Unique Dual Index Primer Pairs). Sequencing was performed with a read length of 2  $\times$  75 bp on Illumina NextSeq 500 at the Tayside Centre for Genomic Analysis.

### Bioinformatics analysis

The quality of sequencing reads was checked using FastQC and adaptors were trimmed using Cutadapt. Trimmed reads were aligned to mouse genome assembly GRCm38/mm10 using STAR (28). Library complexity of mapped reads was assessed using preseq (29), before removal of PCR duplicates using Picard. Correlation of reads between the replicates was analysed using deepTools (30). For ChIP-seq, only uniquely mapped reads were used for data analysis. Multi-mapped reads were removed using SAMtools (31). HA IPs

from WT cells were used as controls for anti-HA ChIP-seq analysis. Inputs were used as controls for endogenous RNAPII ChIP-seq analysis. For genome-wide data visualization, metagene profiles and heatmaps were generated by ngs.plot (32). For data visualization of individual genes, bigwig files were generated by deepTools (30) and loaded to Integrative Genomics Viewer (33). For analysis of read density near TSS, reads aligned to TSS  $\pm$  500 bp were counted using featureCounts (34). Read counts were normalized to library size and differential binding analysed using QLFTest in EdgeR to identify target genes. For gene ontology analysis, webtool Enrichr was used (35). Protein interaction networks were constructed by STRING (36). For RNA-seq, transcripts were quantified using featureCounts. Only transcripts expressed above 1 read per million in >3 replicates were used in the following analysis. Read counts were normalized to the library size and differential expression was analysed using exactTest in EdgeR.

## RESULTS

### CMTR1 is recruited to the transcription start site (TSS)

We aimed to define and compare the target gene specificity of the cap methyltransferases RNMT and CMTR1 in murine embryonic stem cells (ESCs). To facilitate a comparison of the RNMT and CMTR1 genomic binding sites, CRISPR-Cas9 HA-GFP knock-in (KI) ESC lines were generated in which the endogenous *Rnmt* and *Cmtr1* genes were expressed as N-terminal HA-GFP fusion proteins (Supplementary Figure S1a, b). Compared to the untagged endogenous proteins in control ESCs, HA-GFP-RNMT and HA-GFP-CMTR1 were expressed at a reduced level (Supplementary Figure S1b–d). The localization of HA-GFP-RNMT and HA-GFP-CMTR1 was assessed by fluorescence microscopy and immunofluorescence analysis to be equivalent to the endogenous proteins (Supplementary Figure S2). As published previously, endogenous RNMT in ESC was nuclear and HA-GFP-RNMT was localized equivalently (26). Endogenous CMTR1 in ESC and HA-GFP-CMTR1 were also nuclear. Interaction of HA-GFP-RNMT and HA-GFP-CMTR1 with the known binding partners, RAM and DHX15, respectively, was maintained in the HA-GFP KI cells lines indicating that the tag was not interfering with complex formation (Supplementary Figure S3). Although, HA-GFP-RNMT and HA-GFP-CMTR1 were expressed at reduced levels compared to the endogenous protein, in part because they were expressed from a single allele, they are not inherently unstable; neither the endogenous proteins or HA-GFP fusions were stabilized by treatment with the proteasome inhibitor, MG132 (Supplementary Figure S4). Furthermore, the expression level of key pluripotency-associated transcription factors Sox2 and Klf4 were comparable in control and engineered ESCs (Supplementary Figure S1c).

To compare the genomic recruitment pattern of RNMT and CMTR1 as equivalently as possible, Chromatin immunoprecipitation-DNA sequencing (ChIP-seq) assays were performed using anti-HA antibodies to immunoprecipitated (IP) HA-GFP-RNMT and HA-GFP-CMTR1 in ESC lines, with parental ESCs used as a negative control (Supplementary Figure S1d). Following DNA sequenc-

ing, CMTR1 was found to be enriched at the 5' end of genes, peaking proximal to the TSS (transcription start site) (Supplementary Figure S1a, S5, S6, Supplementary Table S1). More than 700 genes had significant CMTR1 binding within 500 bp of the TSS (designated 'CMTR1-enriched genes', Figure 1B, Supplementary Table S1). In contrast, RNMT was not significantly enriched at genomic loci; only 20 genes, including coding and non-coding, were found to have RNMT enrichment (Figure 1B and Supplementary Table S1). Most of the RNMT-enriched genes (13 out of 20) colocalized with CMTR1 near the TSS, but lower numbers of reads were produced from the RNMT ChIP, consistent with reduced genomic binding compared to CMTR1 (Figure 1C). RNMT was immunoprecipitated equivalently to CMTR1 via the HA tag (Supplementary Figure S1d).

ChIP PCR performed with an anti-HA antibody was used to verify that HA-GFP-CMTR1 bound significantly to the TSS whereas HA-GFP-RNMT did not (Supplementary Figure S7a). ChIP-PCR performed using a polyclonal anti-CMTR1 antibody was used to confirm that tagged and endogenous CMTR1 enriched equivalently at the TSS of individual genes, despite the HA-GFP-CMTR1 being expressed at a reduced level compared to the endogenous protein (Supplementary Figure S7b). Since the majority of RNAPII-transcribed RNA receives a 7-methylguanosine cap during transcription, RNMT may be recruited predominantly via the guanosine cap intermediate and/or its interaction with RNA (27,37). The limited binding of RNMT to the TSS in ESC is consistent with previous work performed in HeLa cells in which RNMT had to be overexpressed for RNMT binding to the TSS to be detected (38) (see Discussion).

### CMTR1 colocalizes and correlates with RNAPII binding near the TSS

We analysed the identity of genes with CMTR1 enrichment using Gene Ontology (GO) analysis. The 768 CMTR1-enriched genes predominantly encoded RNA/DNA-binding proteins and were associated with biological processes which included transcription, RNA metabolism and translation (Figure 2A). The 20% most CMTR1-enriched genes were histones, ribosomal proteins and RNA processing and transport factors, indicating potential common features of these genes or their binding proteins which promote CMTR1 recruitment (Figure 2B,C).

Previous studies demonstrated that CMTR1 binds directly to RNAPII CTD phosphorylated on Serine-5 (RNAPII Ser5P) (23), and therefore we investigated the correlation between RNAPII and CMTR1 chromatin binding (Figure 3 and Supplementary Figure S8). RNAPII recruitment was determined by RNAPII ChIP-seq (Supplementary Table S1). RNAPII Ser5P recruitment was determined using a previous dataset (39). Genome-wide, colocalization of CMTR1 with total RNAPII and RNAPII Ser5P was observed, adjacent to the TSS, in coding and non-coding genes (Figure 3A, B). Moreover, the TSS read density in the CMTR1 ChIP samples correlated with total RNAPII and RNAPII Ser5P in expressed genes (Figure 3C and Supplementary Figure S8), consistent with CMTR1 recruitment to

RNAPII Ser5P during the early stages of transcription. Notably, CMTR1-enriched genes, including histones and ribosomal proteins, displayed high promoter-proximal levels of total RNAPII and RNAPII Ser5P (Figure 3C,D and Supplementary Figure S8).

In addition to RNAPII, the TSS may hold other factors which aid recruitment of CMTR1. We inspected the interactome of CMTR1 for such auxiliary factors but did not find a protein which correlated with CMTR1 binding better than RNAPII (40). The helicase DHX15 interacts with CMTR1 and influences gene expression (23,24), although we found this interaction to be minimal in ESC (Supplementary Figure S3b). DHX15-CMTR1 complexes are excluded from RNAPII interactions and DHX15 does not influence the interaction of CMTR1 with RNAPII (23). The hypomethylated guanosine cap may contribute to TSS recruitment. A peak of RNAPII CTD Ser5P was also observed downstream of the transcription end site (TES), which did not correlate with CMTR1 binding (Figure 3A and Supplementary Figure S9).

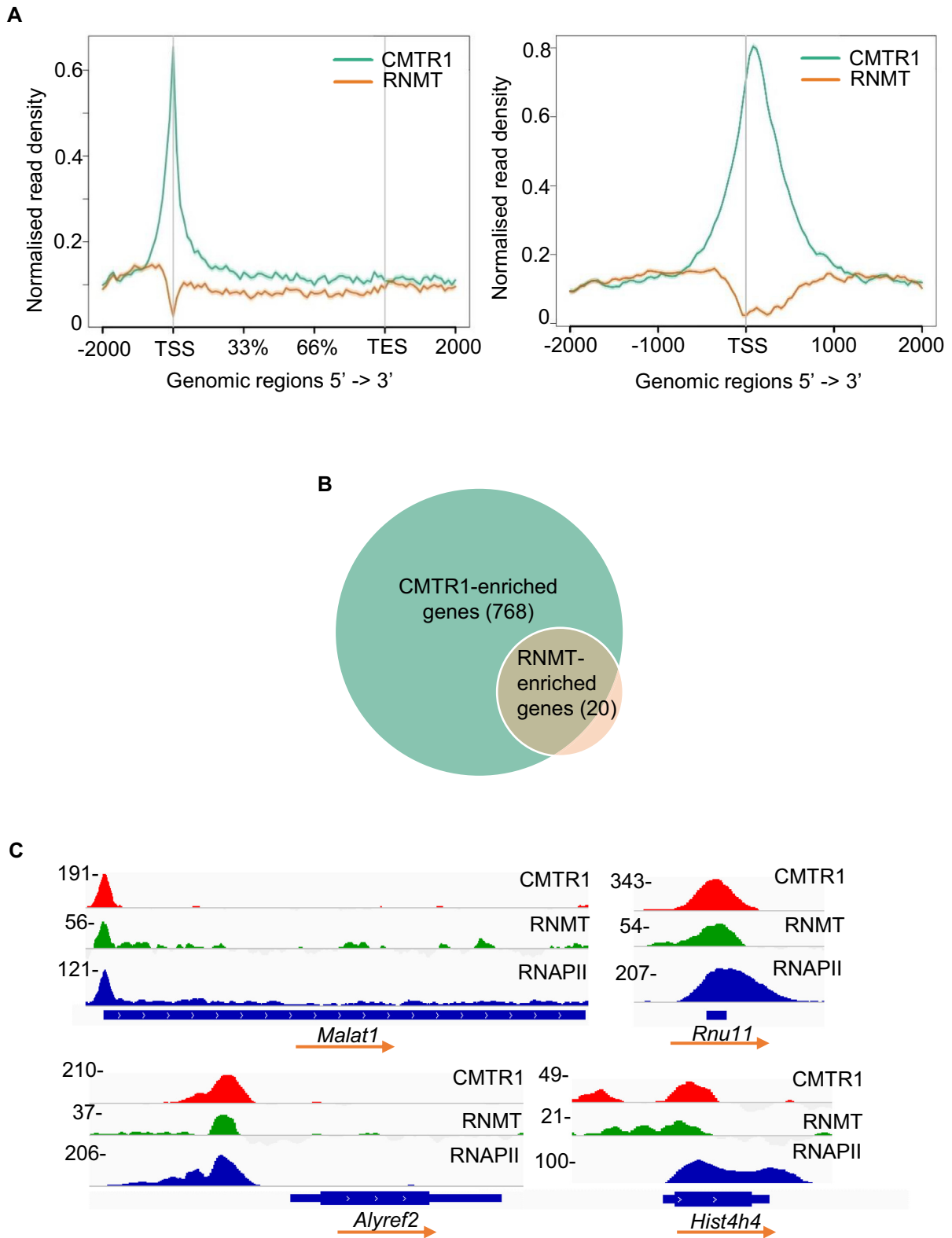
### CMTR1 influences RNAPII TSS levels

Since m7G cap formation is associated with a transcription checkpoint, we investigated the impact of CMTR1 on RNAPII recruitment and retention during the early stages of transcription. RNAPII ChIP-seq was performed following knockdown of endogenous CMTR1 using siRNA. Metagene plots revealed a global reduction in RNAPII binding at the TSS and gene body (Figure 4A), indicating that CMTR1 was influencing RNAPII recruitment or retention. In single gene analysis, out of genes with detectable TSS RNAPII binding, 7299 genes had a reduction in promoter-proximal RNAPII levels following CMTR1 siRNA transfection, whereas 3385 genes had increased RNAPII (Figure 4B,C and Supplementary Table S1).

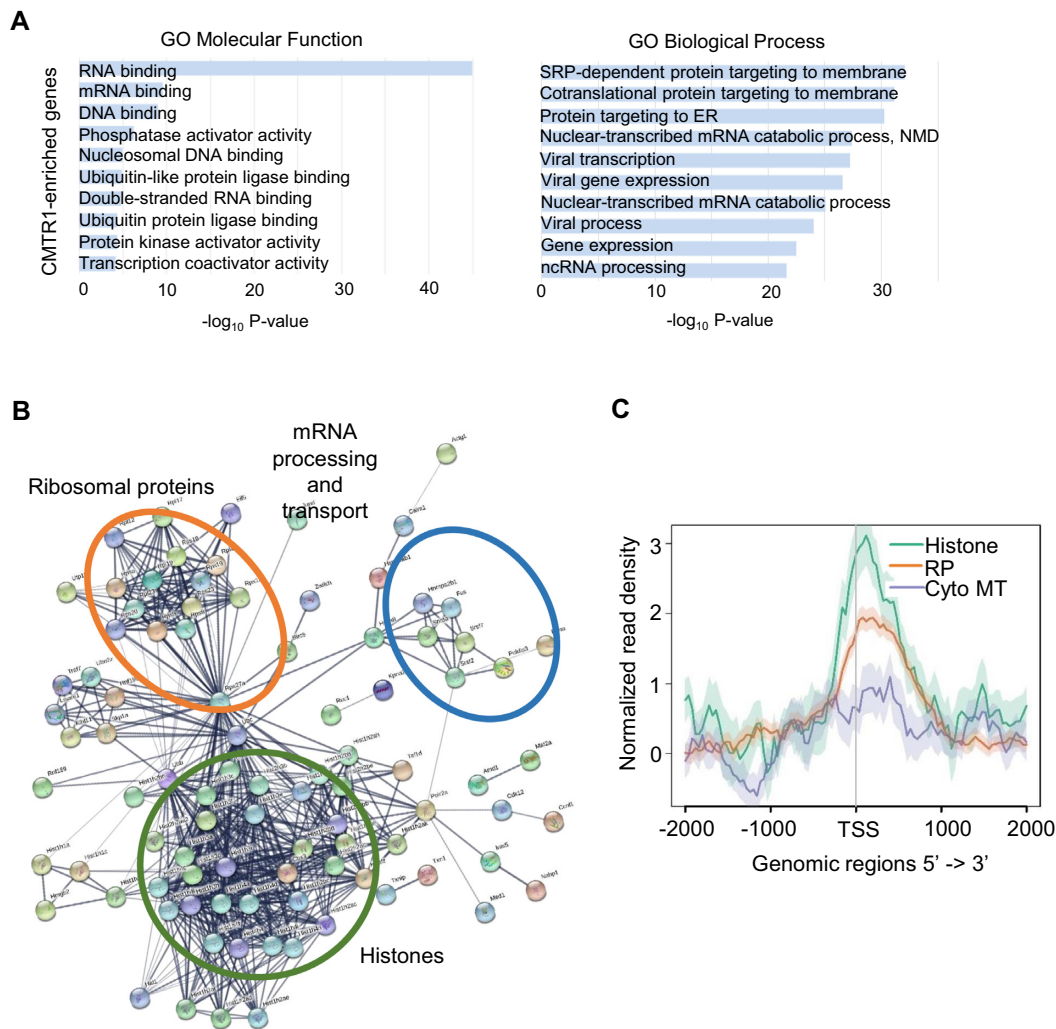
Although a reduction in CMTR1 expression resulted in reduced average RNAPII binding, for single genes the decrease in RNAPII binding did not correlate with TSS CMTR1 levels (Figure 4B). However, for the two groups of genes with the highest CMTR1 binding, the histones and ribosomal protein genes, following CMTR1 repression RNAPII TSS binding was predominantly repressed (Figure 4B,C). In particular for histone genes, the decline in TSS RNAPII levels correlated with CMTR1 binding indicating a potential role for CMTR1 in regulating their transcription directly (Figure 4B). Inspecting RNAPII binding to individual genes in cells transfected with CMTR1 siRNA or control confirmed these findings (Figure 4D). The histone genes had some of the highest levels of RNAPII binding and this was reduced for almost all when CMTR1 was repressed. RNAPII binding to ribosomal protein genes was also largely reduced following CMTR1 repression. Overall, there was a good correlation of RNAPII binding to other genes in cells transfected with CMTR1 or control siRNA, indicating that CMTR1 is not influencing RNAPII recruitment to the majority of genes.

### CMTR1 has gene specific impacts on transcript levels

Since CMTR1 repression resulted in changes in TSS RNAPII levels, we investigated the impact on the transcrip-



**Figure 1.** CMTR1 is recruited to transcription start sites (TSS). **(A)** Metaplots of normalized genome-wide CMTR1 and RNMT ChIP-seq signals  $\pm 2$  kb around gene body (left) and TSS (right). Read densities (reads per kilobase per million, RPKM) were normalized to library size and control IPs to obtain the fold changes in ChIP read densities relative to controls. **(B)** Venn diagram of number of CMTR1 and RNMT-enriched genes. Mapped reads  $\pm 500$  bp around TSS were normalized to library size to obtain RPKM. Enriched genes were defined as those with significant binding above control (fold change  $\geq 1.5$ ,  $P$ -value  $\leq 0.05$ , FDR adjusted  $P$ -value  $\leq 0.05$ ). **(C)** Integrative Genomics Viewer (IGV) depictions of normalized CMTR1, RNMT and RNA polymerase II (RNAPII) read densities on sample genes. ChIP read densities were normalized to library size and control. Transcription direction indicated.



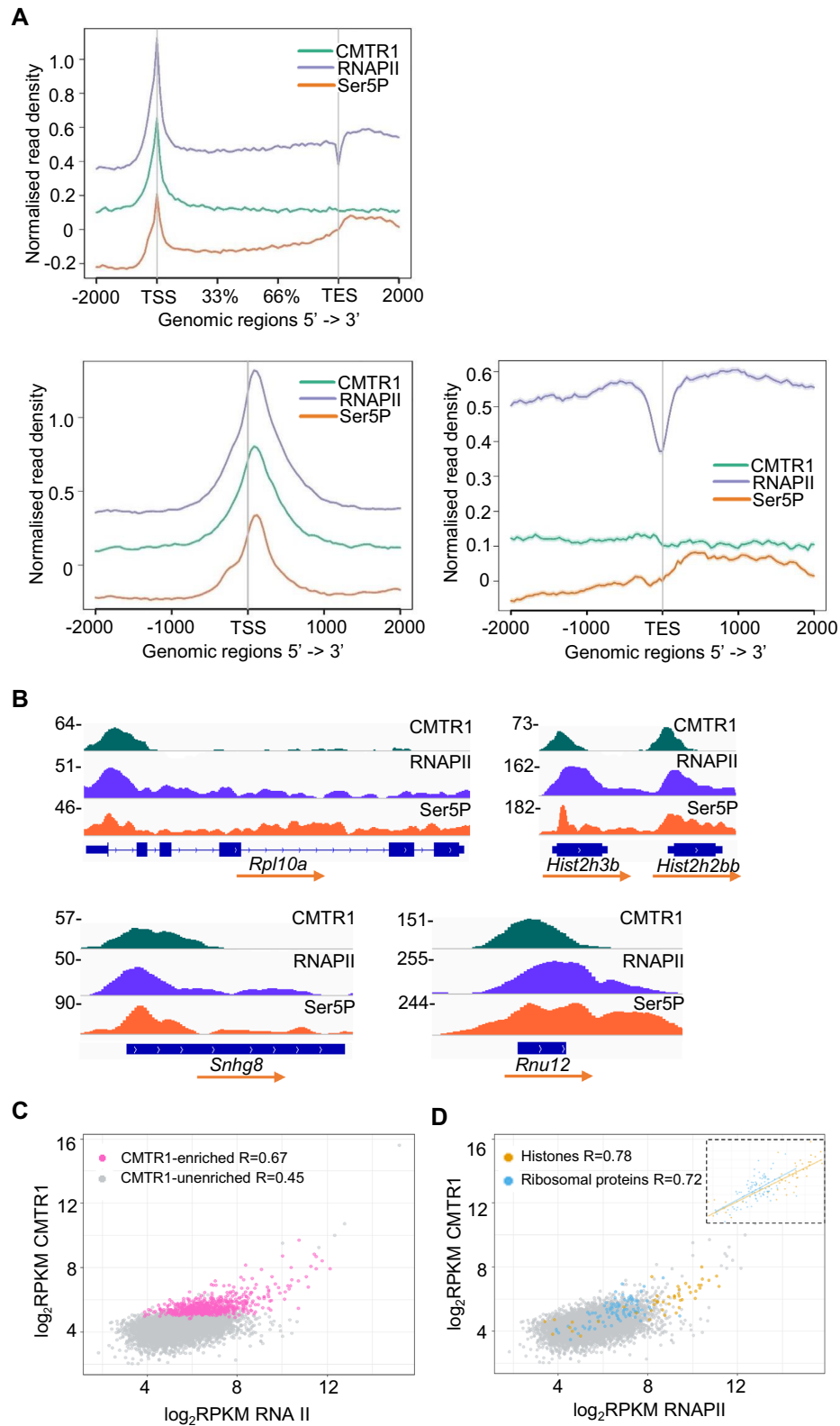
**Figure 2.** Histone and ribosomal protein genes have highest level of CMTR1 binding (A) Gene ontology (GO) analysis of CMTR1-enriched genes (768 genes). Top 10 GO Molecular Function and GO Biological Process terms reported by significance. (B) Protein interaction network of the top 20% protein-encoding CMTR1-enriched genes (154 genes). Lines indicate established interactions from experiments and databases from STRING functional protein association networks. Disconnected nodes were hidden. (C) Metaplots of normalized CMTR1 ChIP-seq signals  $\pm 2$  kb around TSS in histone (68 genes), ribosomal protein (RP, 100 genes) and cytoplasmic microtubule (Cyto MT, 70 genes).

tome by performing RNA sequencing analysis on ESCs in which CMTR1 had been repressed by transfection of CMTR1 siRNAs (Supplementary Figure S10 and Supplementary Table S1). In control ESCs, there were correlations between TSS CMTR1 binding and transcript levels for the histone and ribosomal protein genes (Supplementary Figure S11). For the CMTR1-enriched genes, ~60% were down-regulated following CMTR1 knockdown (Figure 5A). The genes with the highest CMTR1 binding, encoding histones and ribosomal proteins, were significantly repressed following CMTR1 repression, consistent with reduced RNAPII binding (Figures 5B and 4B). 87/100 ribosomal protein genes and 54/68 histone genes were repressed following CMTR1 knock-down (Supplementary Table S1). The correlation coefficients between fold change in transcript level and CMTR1 binding in histone and ribosomal proteins were most significant for these gene sets (Figure 5B). In addition, the fold change in transcript level following CMTR1 repression also correlated with RNAPII TSS

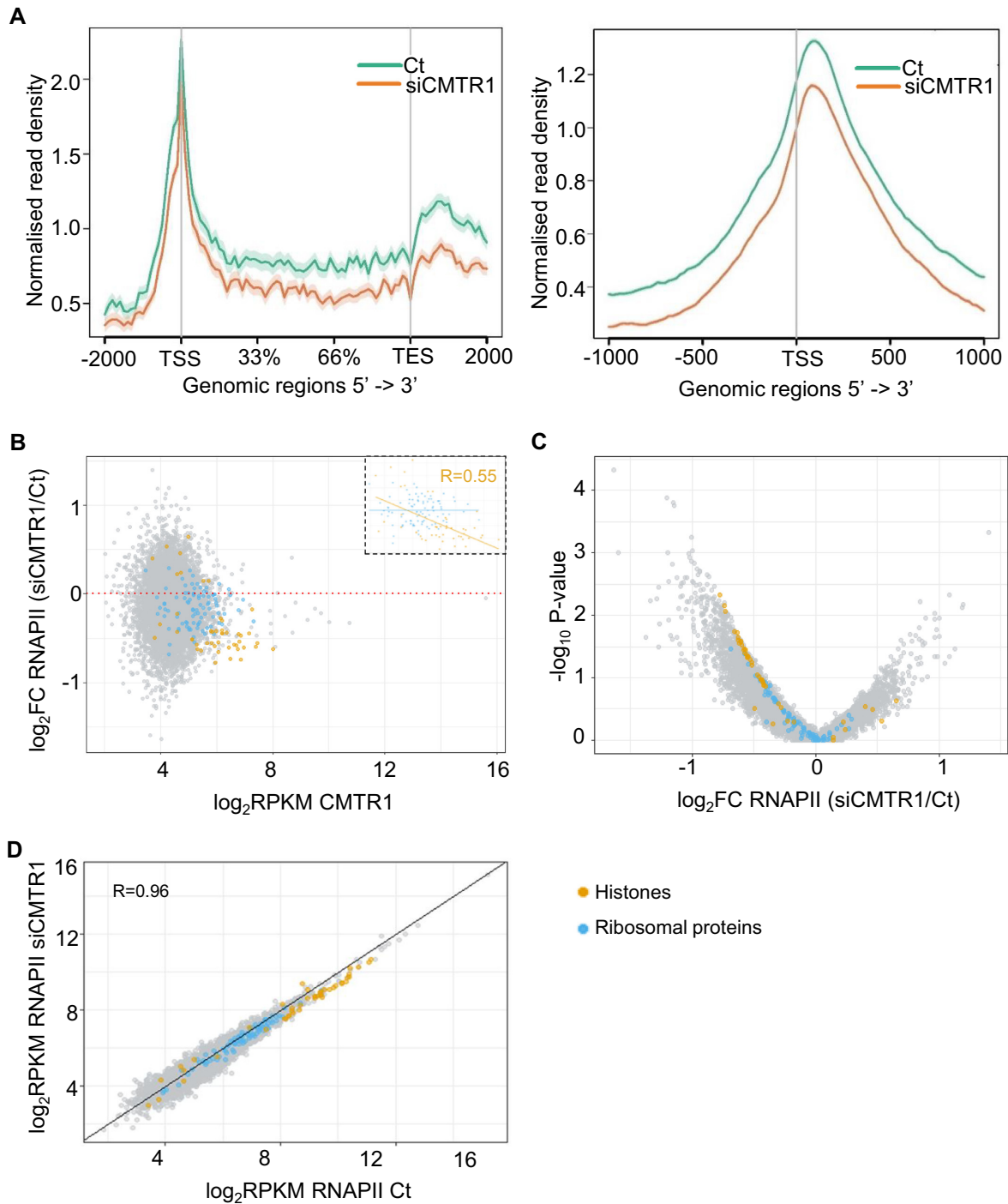
binding, for the histone and ribosomal protein genes (Supplementary Figure S12). For controls, we analysed some highly expressed but CMTR1-unenriched genes, including pluripotency-associated and splicing-associated genes. The fold changes of these transcripts did not correlate with CMTR1 or RNAPII binding, suggesting that the influence of CMTR1 on these genes is likely indirect (Figure 5C, D and Supplementary Figure S11, S12).

Since we observed that the histones and ribosomal protein genes have the highest levels of CMTR1 recruitment (Supplementary Table S1, Figures 2B, 3D), and RNAPII and transcript levels were significantly repressed on CMTR1 repression (Figures 4B, 5B), we investigated whether their transcription can be directly influenced by CMTR1. Nascent transcription of histone genes and ribosomal protein genes in isolated ESC nuclei was analysed in nuclear run-on assays (Figure 5E, F). Nascent transcription of histone genes and ribosomal protein genes was reduced in ESCs in which CMTR1 was repressed (Figure 5E).

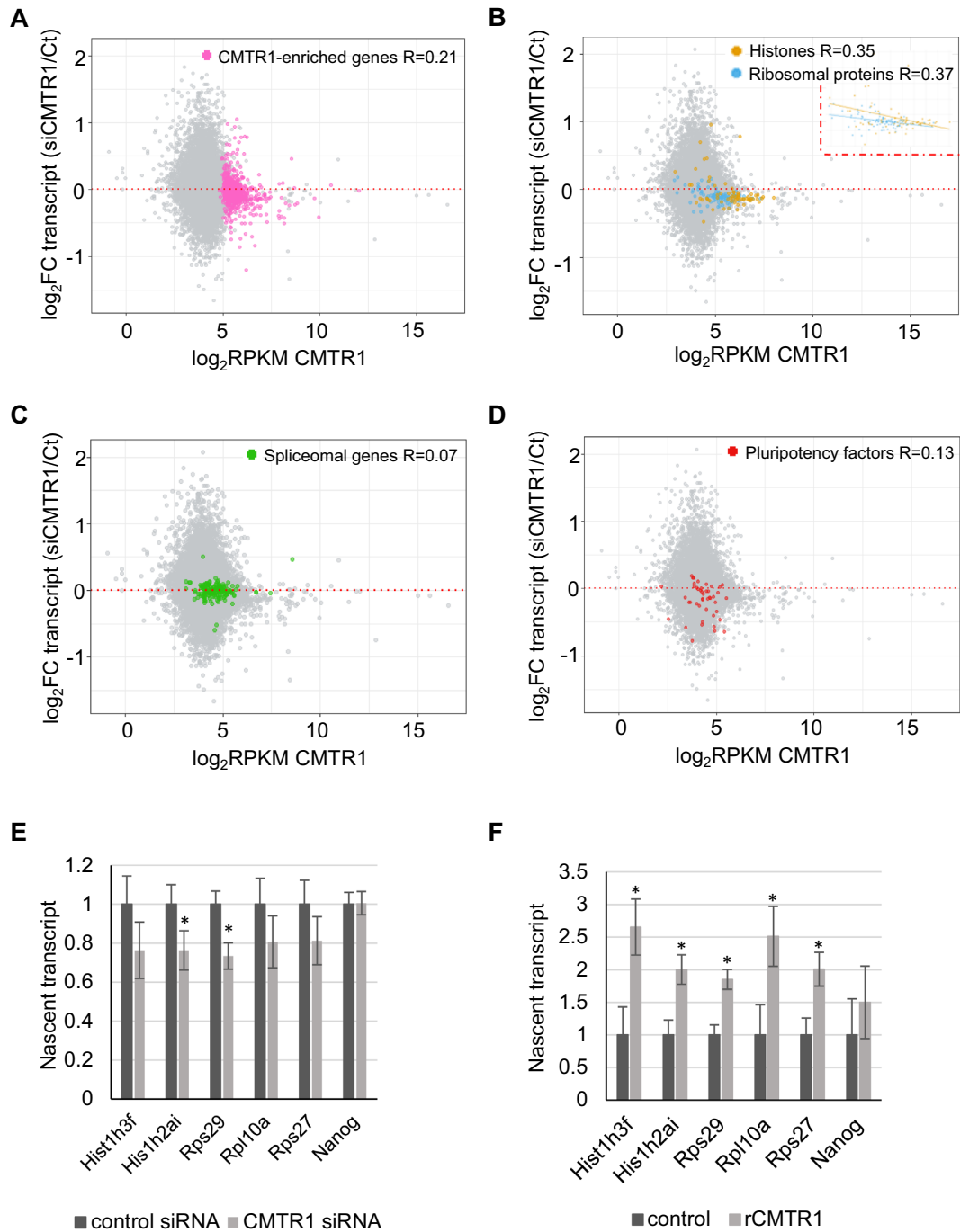




**Figure 3.** CMTR1 TSS binding correlates with RNAPII binding. (A) Metaplots of normalized genome-wide CMTR1, total RNAPII and RNAPII Ser5P CTD ChIP-seq signals  $\pm 2$  kb around gene body, TSS and TES. (B) IGV depictions of normalized CMTR1, RNAPII and Ser5P read densities on sample genes. (C and D) Dot plots indicating CMTR1 and RNAPII read densities ( $\log_2$ RPKM)  $\pm 500$  b around TSS for expressed genes. (C) CMTR1-enriched genes highlighted (pink, 768 genes). (D) Histones (orange, 68 genes) and ribosomal protein (blue, 100 genes) highlighted.



**Figure 4.** A reduction in CMTR1 expression results in loss of RNAPII TSS binding ESCs were transfected with CMTR1 siRNA 1 (siCMTR1) or control (Ct) for 36 h. **(A)** Metaplots of normalized genome-wide RNAPII ChIP-seq signals in gene body  $\pm 2$  kb (left) and TSS  $\pm 1$  kb (right). **(B)** Dot plot of CMTR1 read densities ( $\log_2$ RPKM) in control ESC and  $\log_2$  fold change in RNAPII ChIP signal (TSS  $\pm 500$  b) following transfection with CMTR1 siRNA 1. **(C)** Volcano plot indicating the relationship between  $\log_2$ FC and  $-\log_{10}$  FDR adjusted  $P$ -value in RNAPII ChIP (TSS  $\pm 500$  b) following transfection with CMTR1 siRNA 1. **(D)** Dot plot indicating the relationship between RNAPII binding in cells transfected with CMTR1 siRNA and control. Histones (orange) and ribosomal protein (blue) genes are labelled.



**Figure 5.** CMTR1 promotes expression of ribosomal protein and histone genes (A–D). Dot plots of CMTR1 binding (TSS ± 500 b) in control cells and changes in RNA levels following CMTR1 siRNA transfection. Highlighted: (A) CMTR1-enriched genes (768 genes), (B) histones genes (68 genes) and ribosomal protein genes (100 genes), (C) spliceosomal genes (154 genes) and (D) pluripotency genes (45 genes). Pearson correlation coefficients (R) indicated. (E and F) RT-qPCR analysis of nascent RNA following nuclear run-on assay ( $N = 3$ ). (E) Nuclei were prepared from ESC transfected with CMTR1 siRNA or control. (F) recombinant (r) CMTR1 protein was added to ESC nuclei. For each replicate, RNA level was normalized to the corresponding nascent Actin RNA level. For charts, average and standard deviation of three experiments presented. “\*” indicates a  $P$ -value < 0.05 for the Student’s  $T$  test, performed relative to control.

Conversely, addition of recombinant CMTR1 to ESC nuclei stimulated nascent transcription of histone and ribosomal protein genes (Figure 5F). As a control, transcription of the pluripotency factor *Nanog* was investigated, which was not significantly regulated by CMTR1 repression or addition (Figure 5E,F). These results confirmed that CMTR1 can have a direct impact on nascent transcription of histone and ribosomal protein genes.

### CMTR1 is upregulated during neural differentiation of ESC

To determine the biological impact of CMTR1, we investigated its role in ESCs in the pluripotent state and during differentiation. To induce differentiation, the growth factor LIF (Leukemia Inhibitory Factor) was withdrawn from the medium resulting in morphological changes associated with differentiation, suppression of pluripotency factors (Oct4, *Nanog*, Sox2 and Klf4), and an increase in markers of differentiation (Brachyury, Sox17 and Sox1) (Figure 6A). During LIF withdrawal-induced differentiation, CMTR1 protein was upregulated independently of changes in CMTR1 RNA level, indicating a post-transcriptional mechanism (Figure 6B). CMTR1 is likely to be upregulated either by increases in translation rate or by decreases in protein turnover. ESCs were also differentiated towards a neural fate by culture in N2B27 medium, with associated repression of pluripotency factors and induction of the neural marker, *Pax6* (Figure 6C). CMTR1 protein was also upregulated by a post-transcriptional mechanism during neural differentiation (CMTR1 RNA level was maintained) (Figure 6D).

To investigate the role of CMTR1 in differentiation, we used CRISPR technology in an attempt to delete the *Cmtr1* gene. Despite taking several approaches, we were unable to create a *Cmtr1* knock-out cell. However, we were able to achieve mutation of the *Cmtr1* gene in which 6 bp were deleted and 2 bp were mutated; this created an in-frame deletion of 2 amino acids, Ser 30 and Asp31, and mutation of Asp32 to Cys32 and Glu33 to Val33. We named this mutant *Cmtr1<sup>KD</sup>* because the expression of CMTR1 was significantly reduced (Figure 7A and Supplementary Figure S13). The *Cmtr1<sup>KD</sup>* mutation is not predicted to have a major impact on CMTR1 protein function because it is in a disordered region which is not part of the catalytic domain or a site of known protein:protein interaction (14,15,23). CAP-MAP mass spectrometry analysis of mRNA in the *Cmtr1<sup>KD</sup>* cell line indicated a reduction in the mature caps, <sup>m7</sup>GpppG<sub>m</sub>, <sup>m7</sup>GpppA<sub>m</sub> and <sup>m7</sup>Gppp<sup>m6</sup>A<sub>m</sub> and an increase in the immature caps lacking N1 2'-O-methylation, <sup>m7</sup>GpppG, <sup>m7</sup>GpppA and <sup>m7</sup>Gppp<sup>m6</sup>A compared to control ESCs (Figure 7B) (4). This reduction in N1 2'-O-Me is consistent with a reduction in CMTR1 expression and indicates the absence of a completely redundant methyltransferase for this modification.

### CMTR1 is required for histone and ribosomal protein gene expression during differentiation

Initially, we investigated whether there was concordance between observations made in CMTR1 siRNA treated cells and *Cmtr1<sup>KD</sup>* ESCs. Following CMTR1 siRNA transfection there is a rapid reduction in CMTR1, whereas in

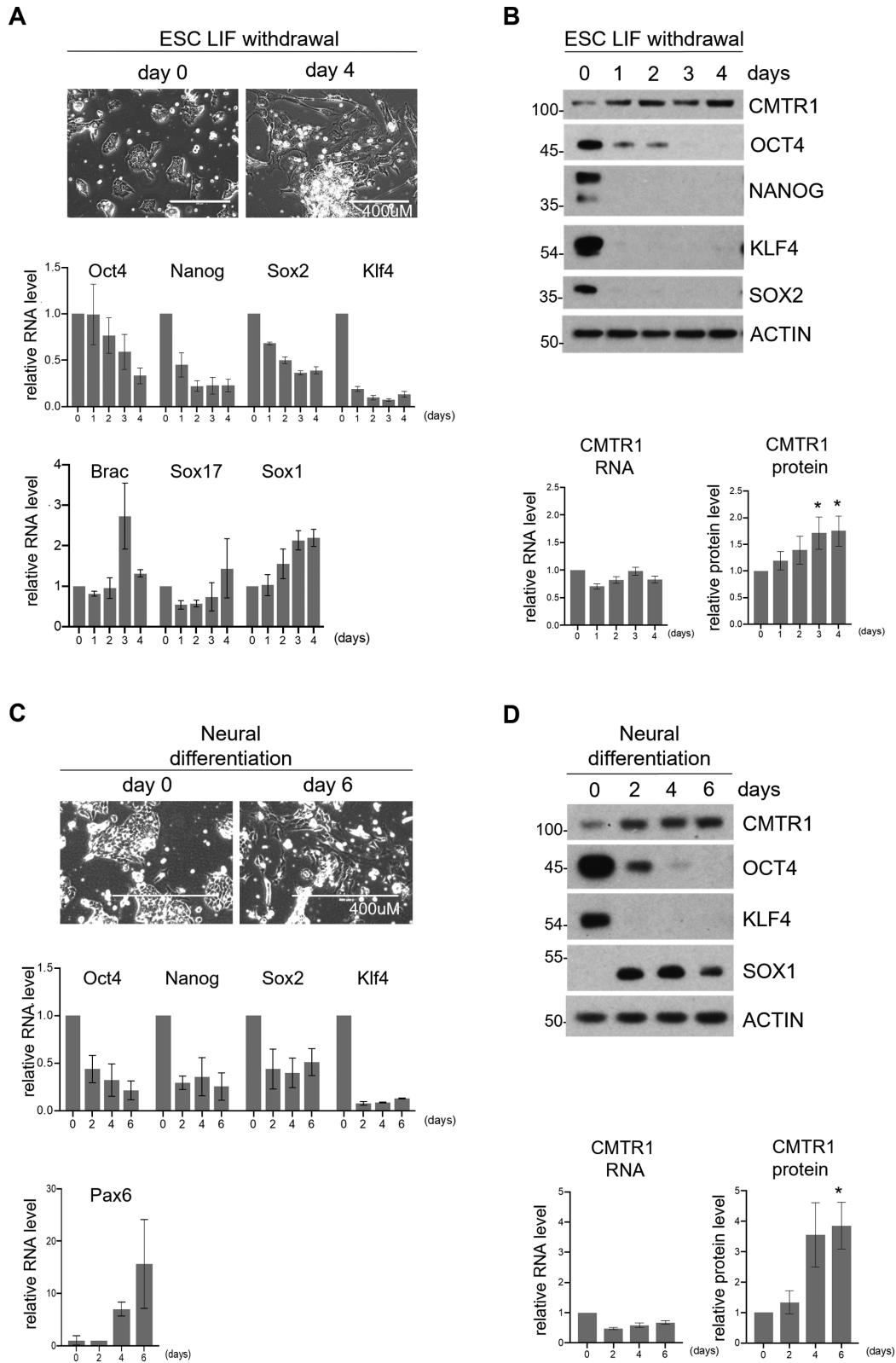
*Cmtr1<sup>KD</sup>* ESCs there is long-term, minimal CMTR1 expression. As established previously, histone and RP RNA expression was reduced on CMTR1 siRNA transfection (Figure 7C, Supplementary Figure S14, Supplementary Table S1). In the *Cmtr1<sup>KD</sup>* ESC, there was also a reduction in the histones and RPs, Hist1h3f, Hist1h2ai and Rps29, compared to *Cmtr1<sup>WT</sup>* ESCs, when cultured in LIF (Day 0) (Figure 7D and Supplementary Figure S15). Other histones and RPs were not repressed in *Cmtr1<sup>KD</sup>* ESC when cultured in LIF (Hist2h2bb, Rps9, Rps27). On neural differentiation of *Cmtr1<sup>KD</sup>* ESC, all histones and RP genes investigated were now repressed compared to the controls, suggesting that there is an increased requirement for CMTR1 as cells differentiate and CMTR1 is upregulated (see Discussion). Expression of HA-CMTR1 in the *Cmtr1<sup>KD</sup>* ESC was sufficient to rescue histone and RP expression. Three control genes were not significantly regulated in the *Cmtr1<sup>KD</sup>* ESC; *Gapdh*, *Srsf2*, *Poldip3* (Figure 7D). Confirming that there is a transcriptional component to CMTR1-regulated RNA expression, ChIP-PCR was used to confirm that RNAPII binding to histone and RP genes is also largely repressed in *Cmtr1<sup>KD</sup>* ESC compared to controls, and expression of HA-CMTR1 resulted in rescue of RNAPII binding for most genes investigated (Figure 7E).

To investigate the impact of CMTR1-dependent RP gene expression, net translation rates were approximated by determining the rate of <sup>35</sup>S cysteine and methionine incorporation into cellular proteins. In *Cmtr1<sup>KD</sup>* ESC, translation was significantly reduced compared to *Cmtr1<sup>WT</sup>* ESC and this could be rescued by expression of HA-CMTR1 (Figure 8A). A reduction in histones may lead to DNA replication stress and damage (41,42). A marker of DNA replication stress, phospho-Chk1, was upregulated on repression of CMTR1 (Figure 8B). During differentiation, a marker of DNA replication stress and damage, gamma-H2AX, was also upregulated on repression of CMTR1 (Figure 8C).

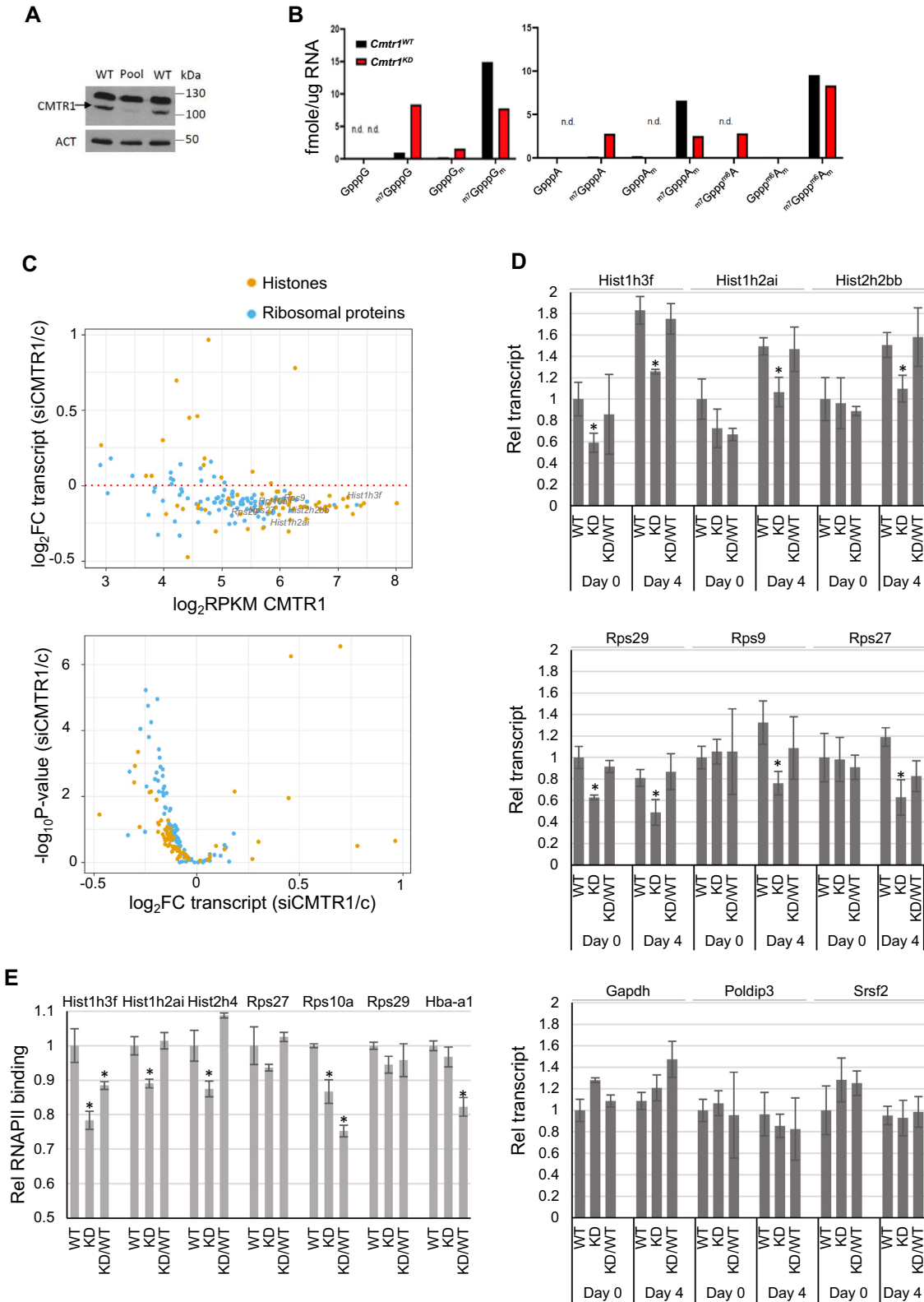
### CMTR1 is required for proliferation during differentiation

Since a reduction in protein synthesis and an increase in DNA replication stress is observed in *Cmtr1<sup>KD</sup>* ESC, we investigated the impact of CMTR1 repression on cell proliferation, a process dependent on histones for DNA replication and RPs for protein synthesis. When cultured in LIF, there was a reduction in cell proliferation in *Cmtr1<sup>KD</sup>* ESC and this could be rescued by expression of HA-CMTR1 (Figure 8D and Supplementary Figure S15B). When LIF was removed to initiate differentiation, *Cmtr1<sup>KD</sup>* ESC proliferation was further reduced, accompanied by apoptosis (Figure 8D,E, Supplementary Figure S15B,C). Cell proliferation was also severely attenuated in *Cmtr1<sup>KD</sup>* ESC undergoing neural differentiation (Figure 8F).

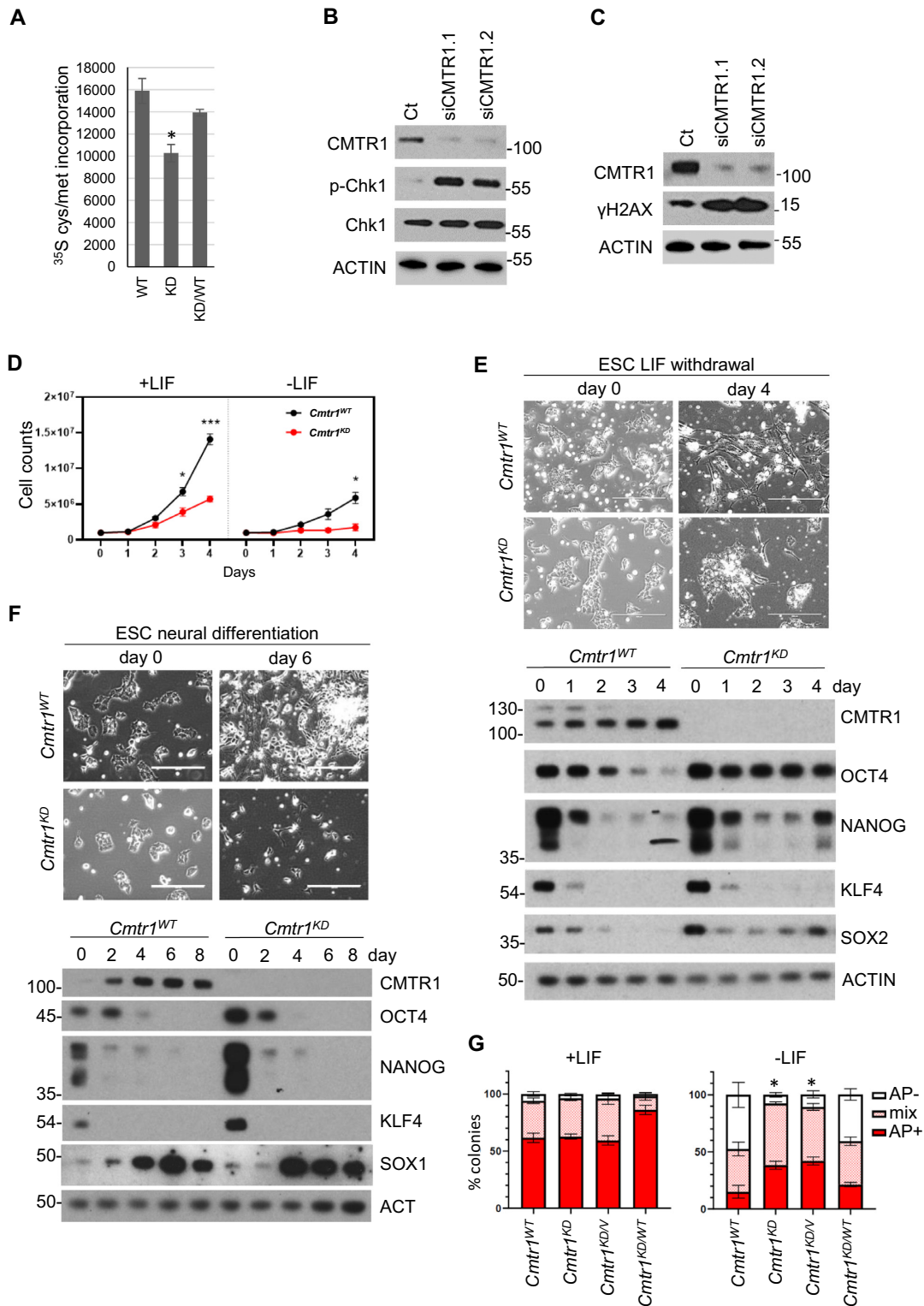
Since defects in gene expression and proliferation resulting from CMTR1 repression were exaggerated during differentiation, we investigated the regulation of pluripotency factors and markers during this process. The expression of pluripotency factors was repressed on CMTR1 siRNA transfection, however on long term inhibition of CMTR1 in *Cmtr1<sup>KD</sup>* ESC, pluripotency factors were increased (Supplementary Figure S16A, Figure 8E,F). It appears that compensatory mechanisms are occurring to maintain pluripo-



**Figure 6.** CMTR1 protein expression increases during ESC differentiation. Differentiation was induced in ESC by LIF withdrawal. **(A)** Phase contrast images (top) and RT-PCR analysis of pluripotency-associated (Oct4, Nanog, Sox2, Klf4) and differentiation-associated (Brac, Sox17, Sox1) genes presented (bottom). **(B)** Representative western blot analysis of CMTR1 and pluripotency associated proteins (top) and quantification of CMTR1 RNA and protein levels (bottom). Neural differentiation was induced in ESC. **(C)** Phase contrast images (top) and RT-PCR analysis of pluripotency-associated (Oct4, Nanog, Sox2, Klf4) and neural (Pax6) genes presented (bottom). **(D)** Representative western blot analysis of CMTR1 and pluripotency associated proteins (top) and quantification of CMTR1 RNA and protein levels (bottom). For charts, average and standard deviation of 3 + experiments presented. For CMTR1 protein chart, “\*” indicates Student’s *T* test *P*-value < 0.05. RNA levels were normalized to *Actin* and *Gapdh* and relative to day 0. Proteins levels were quantified relative to day 0 and normalized to Actin by densitometry using ImageJ.



**Figure 7.** CMTR1 is required for histone and RP gene expression. (A) Western blot analysis of CMTR1 protein expression in *Cmnt1*<sup>WT</sup> and *Cmnt1*<sup>KD</sup> ESCs. Band verified by CMTR1 IP and KO. (B) CAP-MAP RNA cap structure analysis in *Cmnt1*<sup>WT</sup> and *Cmnt1*<sup>KD</sup> ESCs. ‘N.d.’, indicates not detected. (C) Regulation of histone and ribosomal protein genes following CMTR1 siRNA transfections (siCMTR1) in ESCs was determined by RNA seq analysis. Dot plots indicates transcript level in control cells and Log<sub>2</sub>FC following CMTR1 siRNA transfection. Volcano plots demonstrates Log<sub>2</sub>FC expression and EdgeR exactTest -log<sub>10</sub> FDR adjusted *P*-value. (D) RT-PCR analysis of genes indicated relative to Actin RNA in *Cmnt1*<sup>WT</sup>, *Cmnt1*<sup>KD</sup> and *Cmnt1*<sup>KD</sup>/HA-CMTR1 ESCs, in cells cultured in LIF (day 0) and neuro differentiated cells (day 4). RNA levels were normalized to *Actin* and relative to day 0. (E) RNAPII ChIP-PCR in *Cmnt1*<sup>WT</sup>, *Cmnt1*<sup>KD</sup> and *Cmnt1*<sup>KD</sup>/HA-CMTR1 ESCs. Unexpressed gene *Hba-a1* was used as a negative control, (D, E) *N* = 3. Average and standard deviation presented. \**\** Student’s *T* test *P*-value < 0.05, relative to control.



**Figure 8.** CMTR1 is required for histone and ribosomal protein gene expression during differentiation. (A) *Cmtr1<sup>WT</sup>*, *Cmtr1<sup>KD</sup>* and *Cmtr1<sup>KD</sup>*/HA-CMTR1 ESCs were cultured in <sup>35</sup>S cysteine/methionine for 30 min and incorporation into cellular proteins was determined (N = 3). Average and standard deviation is presented. Student's *T* test; \* indicates *P*-value < 0.05 relative to control. (B,C) Western blot analysis of ESC cultured in (B) LIF and (C) without LIF for 2 days following CMTR1 siRNA transfection; N = 3. (D) *Cmtr1<sup>WT</sup>* and *Cmtr1<sup>KD</sup>* ESC were cultured in and without LIF for the days indicated. Cell number was counted; N = 3. Average and standard deviation presented. Student's *T* test; \* indicates *P*-value < 0.05 relative to control timepoint. (E) *Cmtr1<sup>WT</sup>* and *Cmtr1<sup>KD</sup>* ESC were induced to differentiate (E) by LIF withdrawal, (F) towards a neural fate. Phase contrast images and representative western blot analysis. (G) The ESC lines indicated were cultured without LIF for 6 days, then stained with alkaline phosphatase. Resultant colonies were scored as alkaline phosphatase (AP) negative (-), positive (+) or mixed; N = 4. Average and standard deviation presented. Student's *T* test performed on data from AP- colonies; \* indicates *P*-value < 0.05 relative to control timepoint.

tency factor expression on stable repression of CMTR1. However, following LIF withdrawal, *Cmtr1*<sup>KD</sup> ESC exhibited repression of pluripotency factors, equivalent to WT ESCs, although Oct4 was only mildly reduced (Figure 8E, Supplementary Figure S16B). 4 days into the differentiation programme, the *Cmtr1*<sup>KD</sup> ESCs re-expressed pluripotency factors, and expression of HA-CMTR1 rescued this phenotype (Figure 8E, Supplementary Figures S15A, S16B). This re-expression of pluripotency factors in the differentiating *Cmtr1*<sup>KD</sup> ESC culture is consistent with the differentiating cells having reduced proliferation and undergoing apoptosis (Supplementary Figure S15B,C). To validate these findings, we induced neural differentiation and found that *Cmtr1*<sup>KD</sup> ESC were also unable to proliferate during differentiation (Figure 8F). For the few *Cmtr1*<sup>KD</sup> cells which survived induction of neural differentiation (Figure 8F, micrographs), the markers of pluripotency, Oct4, Nanog, Klf4, were repressed on day 4 equivalently to WT cells and the intermediate neural marker, Sox1, was upregulated equivalently to WT cells (Figure 8F).

To further investigate the role of CMTR1 in ESC proliferation and differentiation, an Alkaline phosphatase (AP) colony assay was performed. AP staining is used as a marker of pluripotency during differentiation; when LIF is withdrawn and ESCs differentiate, there is a gradual increase in cell colonies which are negative for AP staining (Figure 8G and Supplementary Figure S17) (43). In *Cmtr1*<sup>WT</sup> ESCs, over a titration of LIF withdrawal, AP negative colonies increased, whereas in *Cmtr1*<sup>KD</sup> ESC, AP negative colonies were minimal (Supplementary Figure S17). Expression of HA-CMTR1 rescued this defect in differentiation (Figure 8G). This confirmed that there are differentiation defects in *Cmtr1*<sup>KD</sup> ESC.

## DISCUSSION

First transcribed nucleotide ribose O-2 methylation (N1 2'-O-Me) is present on mRNA in mammalian cells and by labelling cellular RNA as self, it protects it from degradation by innate immunity pathways (2,3). The contribution of N1 2'-O-Me to cap function in gene expression is less well defined. Here we report that a key function of the N1 2'-O-Me methyltransferase, CMTR1, in embryonic stem cells (ESCs) is to promote histone and ribosomal protein gene expression. We found that CMTR1 is recruited to the TSS, with highest binding at histone genes, ribosomal protein genes and mRNA processing and transport genes. Repression of CMTR1 using siRNA transfections or genome editing resulted in a reduction in RNAPII binding to histone and ribosomal protein genes and a reduction in their expression. Thus, we observed that CMTR1 controls RNA levels in a gene specific manner in ESCs, correlating with RNAPII binding and expression level. In neurons, CMTR1 was found to control expression of *Camk2a* (21), and in the innate immune response, CMTR1 was found to be required for upregulation of certain interferon stimulated genes (20). In both scenarios and presented here in ESCs, CMTR1 is controlling the expression of some of the most highly expressed genes in the cell.

To investigate the impact of CMTR1 in ESCs, lines were created with minimal CMTR1 expression (*Cmtr1*<sup>KD</sup>). The *Cmtr1*<sup>KD</sup> cells had reduced proliferation when cultured

in LIF, but when induced to differentiate, either by LIF removal or under conditions of neural differentiation, proliferation was severely attenuated compared to controls and cell differentiation was incomplete. As discussed, some ribosomal protein genes and histone genes were repressed in pluripotent conditions (in LIF), but they were more severely and consistently repressed during differentiation. Why ESCs exhibit enhanced dependency on CMTR1 for gene expression during differentiation is not clear, but it may be that the condensation of chromatin during differentiation requires increased CMTR1 to maintain interaction with RNAPII. Of note, the marker of DNA replication damage and stress, gamma-H2AX, was only upregulated on CMTR1 repression during differentiation, reinforcing that differentiation exaggerates the requirement for CMTR1.

Which genes mediate the proliferative defect in ESCs? Here we have established that the histone and RP genes are the most CMTR1-dependent and their CMTR1 dependency correlates with an impact on DNA replication stress and damage, and RNA translation, respectively. Although we have not conclusively proven that the defects in histone and RP genes are responsible for the proliferative defect observed when CMTR1 is repressed (it is not possible to restore expression of such a large complement of genes), it is likely that repression of these genes and the correlating defects in protein synthesis and DNA replication make a major contribution to the proliferative defect. For example, the ~40% reduction in protein synthesis in *Cmtr1*<sup>KD</sup> cells is unlikely to support the protein synthesis demands of cell proliferation. In addition, many other genes are CMTR1 responsive (Supplementary Table S1), which may contribute to the proliferative defect in *Cmtr1*<sup>KD</sup>.

In ESCs, the pluripotency factors were repressed following transient repression of CMTR1 using siRNA, but there was little correlation between CMTR1 binding and level of repression following CMTR1 knockdown. In addition, the *Nanog* gene was unresponsive to CMTR1 in the nuclear run-on assay emphasizing that its dependency on the enzyme is likely to involve an indirect mechanism. In *Cmtr1*<sup>KD</sup> ESCs, the pluripotency factors were elevated, potentially in a compensatory mechanism. During differentiation of *Cmtr1*<sup>KD</sup> cells induced by LIF withdrawal, most pluripotency genes were repressed with normal kinetics during initial stages of differentiation but were re-expressed in the differentiating culture in the small population of cells which escaped the proliferation defect and apoptosis. The emergence of AP negative colonies during differentiation was also repressed in *Cmtr1*<sup>KD</sup> cells, indicating a defect in differentiation. During neural differentiation, the attenuation of proliferation was severe, with few cells surviving.

The impact of CMTR1 and N1 2'-O-Me on RNA levels is likely to result from direct and indirect mechanisms. For histones and ribosomal protein genes, the more CMTR1 is bound to the TSS, the more the RNA is repressed in response to CMTR1 knockdown, indicating a relatively direct mechanism connecting CMTR1 and RNA levels. Nuclear run-on assays were used to focus on the role of CMTR1 in transcription. For the ribosomal protein genes and histone genes, the impact of CMTR1 on transcript appeared direct; in isolated nuclei CMTR1 knockdown results in a reduction in their transcription and addition of recombinant CMTR1 increases their transcription. RNMT and its product the



<sup>m7</sup>G cap are linked to a transcriptional checkpoint involving co-transcriptional stability and protein:protein interactions (1,27,37). For CMTR1 and its product N1 2'-O-Me, it is likely that analogous influences on co-transcriptional stability and transcriptional mechanisms are governing CMTR1-dependent mRNA levels. In ESCs, the expression of other genes, including pluripotency factors, is also repressed in response to CMTR1 repression. Of note, we found that genes involved in mRNA expression, processing and transport are CMTR1 targets in ESCs, which are likely to have indirect impacts on mRNA levels.

We report that the other major cap methyltransferase, RNMT, also binds at the TSS but it was only detected above significance thresholds at 20 genes. Unlike the other capping enzymes, RNMT does not have a CTD interaction domain and a direct interaction with RNAPII has not been demonstrated (22,23,44–46). Since the majority of nascent RNA has a <sup>m7</sup>G cap, and a redundant N-7 cap guanosine methyltransferase is not observed in RNMT depleted extracts, RNMT must be recruited to nascent guanosine caps during transcription via a different mechanism (37). A previous study demonstrated that RNMT binds to RNA in correlation with the expression of that gene, without apparent preference for motifs or RNA regions (27). RNMT is more highly expressed than the other capping enzymes and has much higher activity in *in vitro* assays; this and its affinity for RNA may waive the requirement for a CTD interaction motif on which the other capping enzymes are dependent to facilitate interaction with the guanosine cap substrate (26,27). RNMT may act as many other enzymes; a transient interaction with the guanosine cap substrate, facilitated by an affinity for the TSS and associated proteins/RNAs, is sufficient for N-7 guanosine cap methylation to occur.

The majority of mRNA has a <sup>m7</sup>GpppX<sub>m</sub> cap in ESCs and therefore the action of both CMTR1 and RNMT appears to be required to produce stable mRNA and viable cells (4,5). However, CMTR1 and RNMT have different target genes and quite distinct biological roles in ESCs. Whereas CMTR1 is upregulated during differentiation and is required for histone and RP gene expression and proliferation, RNMT-RAM (active RNMT complex) is repressed during differentiation and this repression is required for the process of differentiation (26). In ESCs, high levels of RNMT-RAM are required for pluripotency-associated gene expression and repression of RNMT-RAM is required for repression of these genes during differentiation (26). These differences in RNMT and CMTR1 function during differentiation may be facilitated by differences in their mechanism of recruitment to their substrate, guanosine-capped RNA, which direct differences in their major RNA targets. The differential regulation of RNMT and CMTR1 during ESC differentiation directs them to control different genes and biological pathways during this process.

## DATA AVAILABILITY

All genomic and transcriptomic data described here are deposited in Gene Expression Omnibus under accession number GSE175631.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank the Cowling lab and Alabert lab members for advice, assistance and reagents. We thank the Tayside Centre for Genomic Analysis and the MRC PPU Reagents and Services for CRISPR strategy design.

## FUNDING

China Scholarship Council (to S.L.); Wellcome Trust [097642/Z/11/Z to O.S., J.C.S.]; Medical Research Council [MR/K024213/1 to V.H.C.]; Lister Research Prize Fellowship (to V.H.C.); Royal Society Wolfson Research Merit Award [WRM\R1\180008 to V.H.C.]. Funding for open access charge: Medical Research Council.

*Conflict of interest statement.* None declared.

## REFERENCES

- Pelletier, J., Schmeing, T.M. and Sonenberg, N. (2021) The multifaceted eukaryotic cap structure. *Wiley Interdiscip. Rev. RNA*, **12**, e1636.
- Ramanathan, A., Robb, G.B. and Chan, S.H. (2016) mRNA capping: biological functions and applications. *Nucleic Acids Res.*, **44**, 7511–7526.
- Inesta-Vaquera, F. and Cowling, V.H. (2017) Regulation and function of CMTR1-dependent mRNA cap methylation. *Wiley Interdiscip. Rev. RNA*, **8**, e1450.
- Galloway, A., Atrih, A., Grzela, R., Darzynkiewicz, E., Ferguson, M.A.J. and Cowling, V.H. (2020) CAP-MAP: cap analysis protocol with minimal analyte processing, a rapid and sensitive approach to analysing mRNA cap structures. *Open Biol.*, **10**, 190306.
- Wang, J., Alvin Chew, B.L., Lai, Y., Dong, H., Xu, L., Balamkundu, S., Cai, W.M., Cui, L., Liu, C.F., Fu, X.Y. *et al.* (2019) Quantifying the RNA cap epitranscriptome reveals novel caps in cellular and viral RNA. *Nucleic Acids Res.*, **47**, e130.
- Leung, D.W. and Amarasinghe, G.K. (2016) When your cap matters: structural insights into self vs non-self recognition of 5' RNA by immunomodulatory host proteins. *Curr. Opin. Struct. Biol.*, **36**, 133–141.
- Schuberth-Wagner, C., Ludwig, J., Bruder, A.K., Herzner, A.M., Zillinger, T., Goldeck, M., Schmidt, T., Schmid-Burgk, J.L., Kerber, R., Wolter, S. *et al.* (2015) A conserved histidine in the RNA sensor RIG-I controls immune tolerance to N1-2'-O-Methylated self RNA. *Immunity*, **43**, 41–51.
- Galloway, A. and Cowling, V.H. (2018) mRNA cap regulation in mammalian cell function and fate. *Biochim. Biophys. Acta Gene Regul. Mech.*, **1862**, 270–279.
- Topisirovic, I., Svitkin, Y.V., Sonenberg, N. and Shatkin, A.J. (2011) Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip. Rev. RNA*, **2**, 277–298.
- Kuge, H., Brownlee, G.G., Gershon, P.D. and Richter, J.D. (1998) Cap ribose methylation of c-mos mRNA stimulates translation and oocyte maturation in xenopus laevis. *Nucleic Acids Res.*, **26**, 3208–3214.
- Kruse, S., Zhong, S., Bodi, Z., Button, J., Alcocer, M.J., Hayes, C.J. and Fray, R. (2011) A novel synthesis and detection method for cap-associated adenosine modifications in mouse mRNA. *Sci. Rep.*, **1**, 126.
- Sikorski, P.J., Warminski, M., Kubacka, D., Ratajczak, T., Nowis, D., Kowalska, J. and Jemielity, J. (2020) The identity and methylation status of the first transcribed nucleotide in eukaryotic mRNA 5' cap modulates protein expression in living cells. *Nucleic Acids Res.*, **48**, 1607–1626.
- Langberg, S.R. and Moss, B. (1981) Post-transcriptional modifications of mRNA. Purification and characterization of cap I and cap II RNA (nucleoside-2'-)-methyltransferases from hela cells. *J. Biol. Chem.*, **256**, 10054–10060.

14. Belanger, F., Stepinski, J., Darzynkiewicz, E. and Pelletier, J. (2010) Characterization of hMTr1, a human cap1 2'-O-ribose methyltransferase. *J. Biol. Chem.*, **285**, 33037–33044.
15. Smietanski, M., Werner, M., Purta, E., Kaminska, K.H., Stepinski, J., Darzynkiewicz, E., Nowotny, M. and Bujnicki, J.M. (2014) Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. *Nat. Commun.*, **5**, 3004.
16. Su, A.I., Pezacki, J.P., Wodicka, L., Brideau, A.D., Supekova, L., Thimme, R., Wieland, S., Bukh, J., Purcell, R.H., Schultz, P.G. *et al.* (2002) Genomic analysis of the host response to hepatitis c virus infection. *Proc. Natl. Acad. Sci. USA*, **99**, 15669–15674.
17. Geiss, G.K., Carter, V.S., He, Y., Kwieciszewski, B.K., Holzman, T., Korth, M.J., Lazaro, C.A., Fausto, N., Bumgarner, R.E. and Katze, M.G. (2003) Gene expression profiling of the cellular transcriptional network regulated by alpha/beta interferon and its partial attenuation by the hepatitis c virus nonstructural 5A protein. *J. Virol.*, **77**, 6367–6375.
18. Guerra, S., Lopez-Fernandez, L.A., Pascual-Montano, A., Munoz, M., Harshman, K. and Esteban, M. (2003) Cellular gene expression survey of vaccinia virus infection of human hela cells. *J. Virol.*, **77**, 6493–6506.
19. Kato, A., Homma, T., Batchelor, J., Hashimoto, N., Imai, S., Wakiguchi, H., Saito, H. and Matsumoto, K. (2003) Interferon-alpha/beta receptor-mediated selective induction of a gene cluster by CpG oligodeoxynucleotide 2006. *BMC Immunol.*, **4**, 8.
20. Williams, G.D., Gokhale, N.S., Snider, D.L. and Horner, S.M. (2020) The mRNA cap 2'-O-Methyltransferase CMTR1 regulates the expression of certain interferon-stimulated genes. *Mosphere*, **5**, e00202-20.
21. Lee, Y.L., Kung, F.C., Lin, C.H. and Huang, Y.S. (2020) CMTR1-Catalyzed 2'-O-Ribose methylation controls neuronal development by regulating camk2 alpha expression independent of RIG-I signaling. *Cell Rep.*, **33**, 108269.
22. Haline-Vaz, T., Silva, T.C. and Zanchin, N.I. (2008) The human interferon-regulated ISG95 protein interacts with RNA polymerase II and shows methyltransferase activity. *Biochem. Biophys. Res. Commun.*, **372**, 719–724.
23. Inesta-Vaquera, F., Chaugule, V.K., Galloway, A., Chandler, L., Rojas-Fernandez, A., Weidlich, S., Peggie, M. and Cowling, V.H. (2018) DHX15 regulates CMTR1-dependent gene expression and cell proliferation. *Life Sci. Alliance*, **1**, e201800092.
24. Toczylowska-Socha, D., Zielinska, M.M., Kurkowska, M., Astha Almeida, C.F., Stefaniak, F., Purta, E. and Bujnicki, J.M. (2018) Human RNA cap1 methyltransferase CMTr1 cooperates with RNA helicase DHX15 to modify RNAs with highly structured 5' termini. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **373**, 20180161.
25. Wongpaiboonwattana, W. and Stavridis, M.P. (2015) Neural differentiation of mouse embryonic stem cells in serum-free monolayer culture. *J. Visual. Exp.*, **99**, e52823.
26. Grasso, L., Suska, O., Davidson, L., Gonatopoulos-Pournatzis, T., Williamson, R., Wasmus, L., Wiedlich, S., Peggie, M., Stavridis, M.P. and Cowling, V.H. (2016) mRNA cap methylation in pluripotency and differentiation. *Cell Rep.*, **16**, 1352–1365.
27. Varshney, D., Lombardi, O., Schweikert, G., Dunn, S., Suska, O. and Cowling, V.H. (2018) mRNA cap methyltransferase, RNMT-RAM, promotes RNA pol II-Dependent transcription. *Cell Rep.*, **23**, 1530–1542.
28. 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.
29. Daley, T. and Smith, A.D. (2013) Predicting the molecular complexity of sequencing libraries. *Nat. Methods*, **10**, 325–327.
30. Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A. and Manke, T. (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.*, **42**, W187–W191.
31. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Proc, G.P.D. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
32. Shen, L., Shao, N.Y., Liu, X.C. and Nestler, E. (2014) ngs.plot: quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics*, **15**, 284.
33. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24–26.
34. Liao, Y., Smyth, G.K. and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.
35. Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R. and Ma'ayan, A. (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinf.*, **14**, 128.
36. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P. *et al.* (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.*, **43**, D447–D452.
37. Galloway, A., Kaskar, A., Ditsova, D., Atrih, A., Yoshikawa, H., Gomez-Moreira, C., Suska, O., Warminski, M., Grzela, R., Lamond, A.I. *et al.* (2021) Upregulation of RNA cap methyltransferase RNMT drives ribosome biogenesis during t cell activation. *Nucleic Acids Res.*, **49**, 6722–6738.
38. Aregger, M. and Cowling, V.H. (2013) Human cap methyltransferase (RNMT) N-terminal non-catalytic domain mediates recruitment to transcription initiation sites. *Biochem. J.*, **455**, 67–73.
39. Strikoudis, A., Lazaris, C., Trimarchi, T., Galvao Neto, A.L., Yang, Y., Ntziachristos, P., Rothbart, S., Buckley, S., Dolgalev, I., Stadtfeld, M. *et al.* (2016) Regulation of transcriptional elongation in pluripotency and cell differentiation by the PHD-finger protein phf5a. *Nat. Cell Biol.*, **18**, 1127–1138.
40. Simabuco, F.M., Pavan, I.C.B., Pestana, N.F., Carvalho, P.C., Basei, F.L., Campos Granato, D., Paes Leme, A.F. and Zanchin, N.I.T. (2019) Interactome analysis of the human Cap-specific mRNA (nucleoside-2'-O)-methyltransferase 1 (hMTr1) protein. *J. Cell. Biochem.*, **120**, 5597–5611.
41. Hogan, A.K. and Foltz, D.R. (2021) Reduce, retain, recycle: mechanisms for promoting histone protein degradation versus stability and retention. *Mol. Cell Biol.*, **41**, e0000721.
42. Saldívar, J.C., Cortez, D. and Cimprich, K.A. (2017) The essential kinase ATR: ensuring faithful duplication of a challenging genome. *Nat. Rev. Mol. Cell Biol.*, **18**, 622–636.
43. Marti, M., Mulero, L., Pardo, C., Morera, C., Carrio, M., Laricchia-Robbio, L., Esteban, C.R. and Izpisua Belmonte, J.C. (2013) Characterization of pluripotent stem cells. *Nat. Protoc.*, **8**, 223–253.
44. Akichika, S., Hirano, S., Shichino, Y., Suzuki, T., Nishimasu, H., Ishitani, R., Sugita, A., Hirose, Y., Iwasaki, S., Nureki, O. *et al.* (2019) Cap-specific terminal n (6)-methylation of RNA by an RNA polymerase II-associated methyltransferase. *Science*, **363**, eaav0080.
45. McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S. and Bentley, D.L. (1997) 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.*, **11**, 3306–3318.
46. Wen, Y., Yue, Z. and Shatkin, A.J. (1998) Mammalian capping enzyme binds RNA and uses protein tyrosine phosphatase mechanism. *Proc. Natl. Acad. Sci. USA*, **95**, 12226–12231.