GENETICS

Dual DNA demethylation mechanisms implement epigenetic memory driven by the pioneer factor PAX7

Juliette Harris[1], Alexandre Mayran[1], Arthur Gouhier[1], Yves Gauthier[1], Nawal Hajj Sleiman[2], Samir Merabet[2], Michael Dukatz[3], Pavel Bashtrykov[3], Albert Jeltsch[3], Haig Djambazian[4], Shu-Huang Chen[4], Aurelio Balsalobre[1], Jacques Drouin[1]*

Pioneer transcription factors have the unique ability to open chromatin at enhancers to implement new cell fates. They also provide epigenetic memory through demethylation of enhancer DNA, but the underlying mechanisms remain unclear. We now show that the pioneer paired box 7 (PAX7) triggers DNA demethylation using two replication-dependent mechanisms, including direct PAX7 interaction with the E3 ubiquitin-protein ligase (UHRF1)-DNA methyltransferase 1 (DNMT1) complex that is responsible for DNA methylation maintenance. PAX7 binds to UHRF1 and prevents its interaction with DNMT1, thus blocking activation of its enzyme activity. The ten-eleven translocation DNA dioxygenase (TET) DNA demethylases also contribute to the replication-dependent loss of DNA methylation. Thus, PAX7 hijacks UHRF1 to block activation of DNMT1 after replication, leading to loss of DNA methylation by dilution, and the process is assisted by the action of TET demethylases.

INTRODUCTION

In mammals, the long-term stability of gene expression and cell identity is largely ensured by the patterns of DNA methylation at CpG dinucleotides (*1*). These patterns provide epigenetic memory that relies on CpG methylation and demethylation at regulatory sequences within the globally highly methylated genome. Demethylated regulatory regions such as promoters and enhancers are associated with chromatin marks of transcriptional activity and DNA accessibility. The inverse relationship between DNA methylation and gene expression was studied extensively, in particular in the context of CpG-rich promoter regions (*2*). Distant regulatory regions such as enhancers also exhibit CpG demethylation following transcriptional activity (*3*).

In development, regulatory elements involved in implementation of new cell fates undergo localized DNA methylation changes (*4*, *5*). Pioneer transcription factors (PFs) act as master regulators of cell fate specification, and they fulfill this role by opening chromatin at enhancers that were previously within repressed and inaccessible chromatin (*6*). PFs also trigger demethylation of CpGs within their target sites (*7*–*9*), thus implementing long-term epigenetic memory of their crucial action (*10*).

Robust mechanisms are in place to ensure the stability of the epigenome and its maintenance after DNA replication. These mechanisms involve the maintenance of chromatin marks associated with either repressed or active chromatin (*11*) as well as the maintenance of DNA methylation. Reprogramming of cell identity by PFs must therefore counter these maintenance systems to implement a remodeled epigenome. Maintenance of CpG methylation is ensured by the E3 ubiquitin-protein ligase (UHRF1)-DNA methyltransferase 1 (DNMT1) complex that contains the maintenance DNA methyltransferase DNMT1 and its regulatory cofactor UHRF1 (*12*, *13*). Upon DNA replication, only the matrix strand of DNA remains methylated, leading to hemimethylated CpGs. Recognition of hemimethylated CpGs by UHRF1 activates DNMT1 for methylation of cytosines on newly synthesized DNA to maintain the methylation state (*14*). There is extensive genomic DNA demethylation during early embryogenesis that results from blockade of the UHRF1-DNMT1 maintenance system (*15*), and it is followed by de novo genome-wide DNA methylation, setting up long-term CpG methylation patterns (*16*).

The action of some PFs was associated with recruitment of the DNA demethylases ten-eleven translocation DNA dioxygenases (TET1/2/3) (*8*). TET enzymes lead to DNA demethylation through successive oxidations of methylcytosine (5mC) into 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fmC), and 5-carboxylcytosine (5caC) (*17*). In the present work, we investigated the mechanism used by the PF paired box 7 (PAX7) for demethylation of its target enhancer DNA. We find that PAX7-dependent DNA demethylation relies on two replication-dependent mechanisms involving blockade of the UHRF1-DNMT1 methylation maintenance complex together with the TET demethylases acting after replication.

RESULTS

PAX7 triggers progressive and localized DNA demethylation

The PF PAX7 reprograms AtT-20 cells from a pituitary corticotrope to melanotrope-like phenotype (*18*). This results from chromatin opening and activation of a new enhancer repertoire (*9*, *19*). This process (Fig. 1A) is accompanied by various epigenetic modifications at PAX7 target enhancers, such as gain of the active marks H3K4me1 and H3K27ac and loss of the repressive mark H3K9me2. As revealed by whole-genome bisulfite sequencing (WGBS), PAX7 also induces localized DNA demethylation (*9*) at pioneered enhancers (Pioneer) from >80 to <40% methylated CpGs (Fig. 1B). This demethylation contrasts with the maintenance of DNA methylation patterns at other types of PAX7 targets that include already active and accessible enhancers where PAX7 acts as a transcription factor (Constitutive sites) or sites within closed chromatin where PAX7 recruitment does not trigger chromatin remodeling (Resistant sites,

[1]Laboratoire de génétique moléculaire, Institut de recherches cliniques de Montréal, Montréal H2W1R7, Canada. [2]Institut de Génomique Fonctionnelle de Lyon, CNRS UMR5242, École Normale Supérieure de Lyon, Université Lyon I, Lyon 69007, France. [3]Department of Biochemistry, Institute of Biochemistry and Technical Biochemistry, University of Stuttgart, Stuttgart 70569, Germany. [4]Victor Phillip Dahdaleh Institute of Genomic Medicine, McGill University, Montréal H3A 0G1, Canada.
*Corresponding author. Email: jacques.drouin@ircm.qc.ca

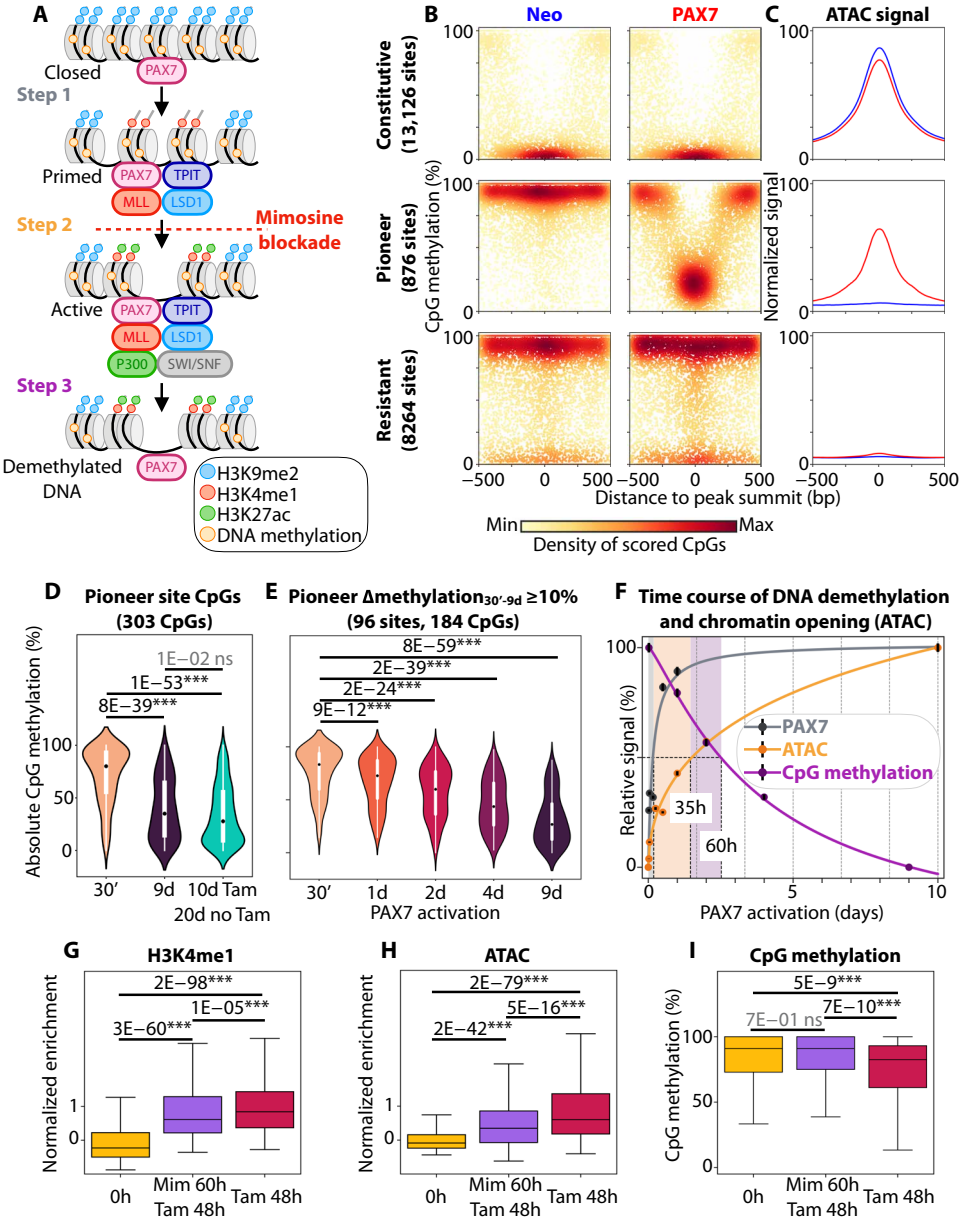

Fig. 1. PAX7 triggers progressive and localized DNA demethylation. (**A**) Current scheme for PAX7 pioneer action (*6*) indicating steps of enhancer chromatin opening (*20*). Step 2 is dependent on passage through replication and sensitive to mimosine blockade. (**B**) Scatter plots representing levels (%) of CpG dinucleotide methylation derived from WGBS performed on control AtT-20 (Neo) and PAX7-expressing AtT-20 (PAX7) cells. For representation, the 8450 CpGs present within the Pioneer enhancer subset ($n = 876$) were used together with subsets of the same number of randomly selected CpGs from the other types of enhancers centered on PAX7 peak summit. Data are shown for Constitutive, Pioneer, and Resistant enhancers, as defined in (*9*, *19*). (**C**) Average profiles of normalized ATAC-seq data for Neo (blue) and PAX7 cells (red) at indicated enhancer subsets. (**D**) Violin plots showing CpG methylation (%) at Pioneer enhancers derived from ChIP-bis-seq performed on AtT-20 expressing ER-PAX7 induced with Tam for 30 min (′) or 9 days (d) and for 10 days followed by 20 days without Tam. *P* values from the *t* test are shown. ns, not significant. (**E**) Violin plots showing CpG methylation (%) at Pioneer enhancers derived from ChIP-bis-seq performed on AtT-20 expressing ER-PAX7 induced with Tam for the indicated times. The data include CpGs within a ±150-bp window around PAX7 peaks (sequencing coverage ≥10 reads per CpG in all samples) that showed ≥10% of absolute demethylation between 30 min and 9 days of Tam treatments. (**F**) Time course of relative PAX7 recruitment (FlagM2 ChIP-seq), chromatin opening (ATAC-seq), and DNA demethylation [sites from (D)]. Standard errors of the means are shown as black bars. h, hours. (**G** to **I**) Box plots of normalized signals at Pioneer enhancers for H3K4me1 ChIP-seq (G) (±250 bp), ATAC-seq (H) (±50 bp), and 5mC determined by WGBS (I) (±150 bp) performed on ER-PAX7 cells treated or not with Tam and/or mimosine (Mim), as indicated. *P* values were calculated by the *t* test.

Fig. 1B). DNA demethylation accompanies opening of Pioneer enhancers, as revealed by the assay for transposon-accessible chromatin sequencing (ATAC-seq) (Fig. 1C). Whereas most Pioneer sites show global DNA demethylation, not all CpGs located within these enhancers demonstrate substantial demethylation (fig. S1A).

To further study the impact of PAX7 on DNA methylation, we used AtT-20 cells expressing ER-PAX7, a tamoxifen (Tam)–inducible chimera that allowed us to define two steps in chromatin opening (Fig. 1A), a first step that involves changes of histone marks and a second step that is replication dependent and that may involve changes in nucleoplasmic compartment (*20*). This system was used to show the long-term maintenance of DNA demethylation after removal of PAX7 (Fig. 1D) under conditions previously shown to preserve partial DNA accessibility (*9*). The system was also used to assess enhancer demethylation kinetics by chromatin immunoprecipitation followed by bisulfite sequencing (ChIP-bis-seq). We selected CpGs within a ±150–base pair (bp) window around PAX7 peak summits to assess CpG methylation (Fig. 1E and fig. S1B). These analyses show progressive loss of CpG methylation at Pioneer enhancers (Fig. 1, E and F) with no change at Constitutive and Resistant enhancers (Fig. 1B). Within this time frame, the half-time for CpG demethylation is ~60 hours compared to the ~35-hour half-time for chromatin opening revealed by ATAC-seq (Fig. 1F). Within a ±150-bp window, there is no correlation between the extent of individual CpG demethylation or the timing of demethylation and their distance from the summit of the PAX7 ChIP sequencing (ChIP-seq) peak (fig. S1C) or between the strength of PAX7 recruitment and extent of demethylation (fig. S1, D and E).

As we previously showed that step 1 of chromatin opening by PAX7 (Fig. 1A) results in enhancer priming, whereas full enhancer activation (step 2) requires passage through replication and cell division (*20*), we used mimosine to block DNA replication and the cell cycle in G_1 and determine whether DNA demethylation may proceed independently of replication. Consistent with enhancer priming in the presence of mimosine (Fig. 1A), we found partial gains of H3K4me1 and ATAC signals (Fig. 1, G and H) in mimosine-arrested cells. However, DNA demethylation is entirely blocked in these cells (Fig. 1I), indicating that it is fully dependent on passage through replication.

PAX7 interacts with UHRF1 and DNMT1

To investigate PAX7-dependent DNA demethylation, we performed rapid immunoprecipitation mass spectrometry of endogenous protein (RIME) assays (*21*) to identify potential protein partners for PAX7 action. We compared proteins immunoprecipitated with Flag-PAX7 in the RIME assay of PAX7 cells with those present in control (Neo) cells (Fig. 2A) or with those immunoprecipitated with T-box transcription factor 19 (TPIT) (Fig. 2B and table S1). Because TPIT is a nonpioneer factor acting with PAX7 at ~70% of the same enhancers (*18*, *22*), this comparison highlights pioneer-specific proteins and it notably identified three classes of PAX7-enriched proteins: In addition to expected chromatin remodeling proteins, many proteins of the replication complex, in particular lagging but not leading strand synthesis (fig. S2A), were enriched, together with DNMT1 and UHRF1, the DNA methylation maintenance complex (Fig. 2, A and B). It is noteworthy that putative interactions of UHRF1 and/or DNMT1 were documented for different PFs but not explored further (*21*, *23*–*25*).

While the RIME data indicated enrichment of DNMT1 and UHRF1 with PAX7 (Fig. 2, A and B), it does not test for direct interactions between these proteins. We used in vitro pull-down assays to assess whether UHRF1 and/or DNMT1 may interact directly with PAX7, and we show that both UHRF1 and DNMT1 interact with maltose binding protein (MBP)-PAX7 (Fig. 2C). The recruitment of DNMT1 by UHRF1 involves direct physical interaction between the DNMT1 replication foci domain (RFD) and/or catalytic domain and the plant homeodomain (PHD) and Set and RING finger–associated (SRA) domains of UHRF1 (*26*, *27*). We used an array of deletion mutants of PAX7, UHRF1, and DNMT1 to define the domains in each protein that are involved in their interactions.

PAX7 interacts in vitro with both UHRF1 and DNMT1 via its PAIRED domain (Fig. 2D), a DNA binding domain (DBD) highly conserved among PAX proteins (*28*). This DBD is necessary and sufficient to bind each partner, as shown with PAX7 constructs 1 to 170 amino acids and ΔPAIRED (Fig. 2D). We constructed a series of alanine replacement mutants within the PAX7 PAIRED domain to precisely define residues required for UHRF1 and/or DNMT1 interaction and in the hope of devising a mutant that may dissociate pioneering action from UHRF1-DNMT1 interaction. The mutants clearly defined a patch on the PAIRED domain that mediates interaction with UHRF1 (fig. S2, B and C, mutant MD11), but unfortunately, this mutant impaired PAX7 binding at its Pioneer sites and, thus, its pioneering ability.

UHRF1 requires both PHD and SRA domains to bind PAX7 strongly, as shown by deletions ΔPHD and ΔSRA and supported by the PHD-SRA construct (Fig. 2E). The DNMT1 RFD domain is sufficient for strong binding to PAX7 in vitro, as shown by the 291– to 575–amino acid construct (Fig. 2F), and the C-terminal domain of DNMT1 may also contribute to PAX7 interaction, but this was not further explored. The interacting domains of each protein are summarized (Fig. 2G) to highlight that the same domains of UHRF1 and DNMT1 are implicated in binding PAX7 and the other partner (*13*, *26*, *27*). This suggests a model of PAX7 competing the UHRF1-DNMT1 interaction. We tested this hypothesis in vitro using order-of-addition pull-down experiments (Fig. 2H). Under conditions where PAX7 is limiting, addition of UHRF1 first to MBP-PAX7 beads followed by DNMT1 prevents its recruitment to PAX7 (Fig. 2H), suggesting that DNMT1 cannot displace UHRF1 from PAX7. In contrast, when DNMT1 was added first followed by UHRF1, both proteins bound to PAX7, suggesting that prior occupancy by DNMT1 slows UHRF1 access or that the three proteins may form a complex when DNMT1 is at its core, given that DNMT1 contains two interacting domains for the other proteins (Fig. 2G). We therefore propose a model where PAX7 binding to UHRF1 blocks the binding of both to DNMT1.

PAX7 displaces UHRF1 from DNMT1 and prevents its activation

To directly assess in cells the model of PAX7 competition of UHRF1-dependent DNMT1 activation, we first used bimolecular fluorescence complementation (BiFC) (*29*). We set up a fluorescence complementation system that reveals UHRF1-DNMT1 interaction by fusing the C terminus of Cerulean (CC) to UHRF1 and the Cerulean N terminus (CN) to DNMT1. When transfected together in human embryonic kidney (HEK) 293 cells, blue fluorescence is observed in nuclei of transfected cells (Fig. 3, A and B). We then used the same CC-UHRF1 construct to assess fluorescence complementation with a PAX7 construct that contains the N terminus of Venus (VN). The VN fragment can complement with the CC fragment

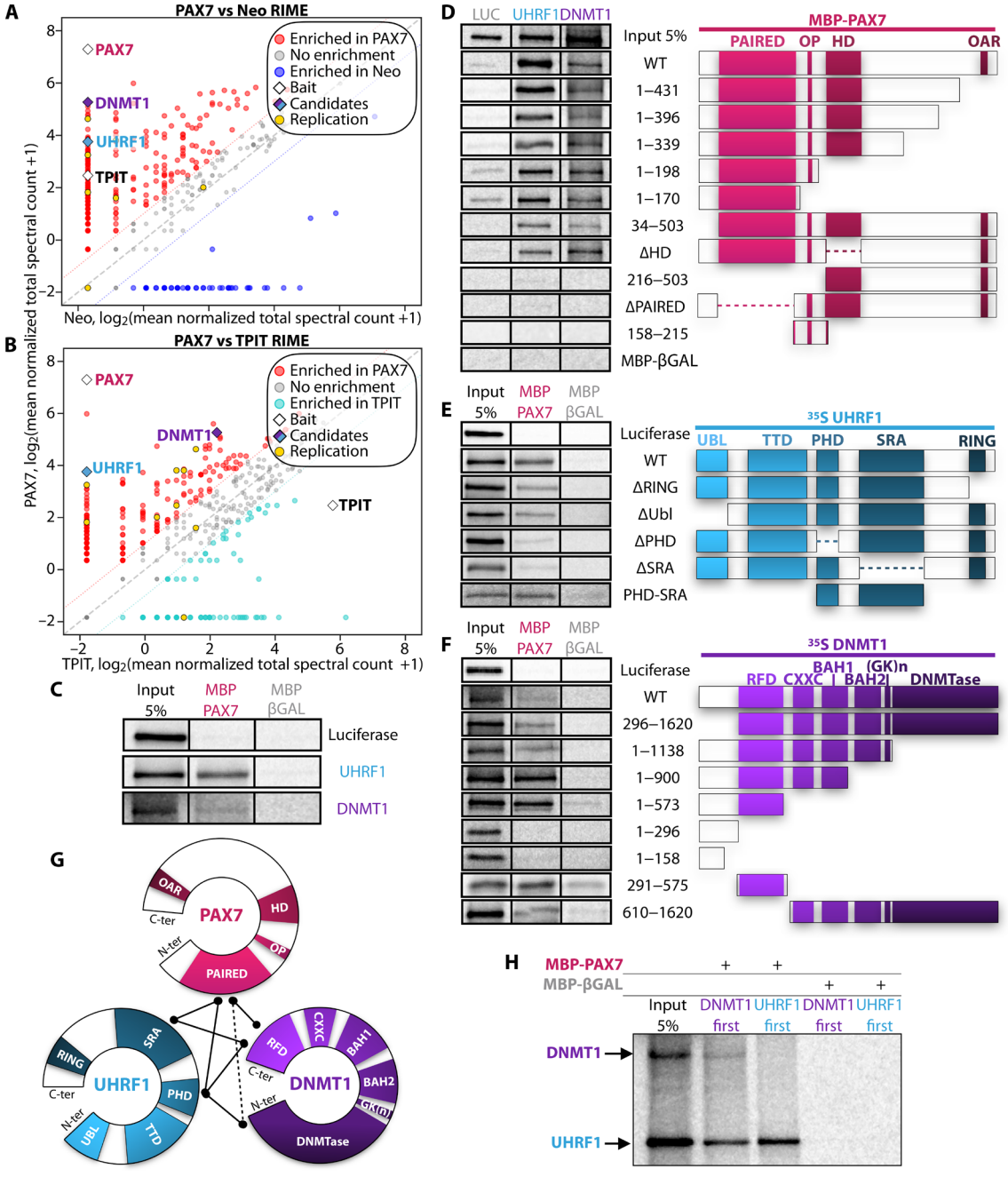

Fig. 2. Identification of candidates for DNA demethylation. (**A**) Scatter plot of protein enrichment determined by RIME (*21*) after immunoprecipitation using a Flag-M2 antibody in control Neo AtT-20 (blue) or PAX7 AtT-20 cells (red). (**B**) Same as (A), immunoprecipitation using a Flag-M2 antibody in AtT-20 cells expressing Flag-PAX7 (red) compared to TPIT immunoprecipitation in AtT-20 cells (green). (**C**) Pull-down assay of ^{35}S-labeled UHRF1 and DNMT1 binding to MBP-PAX7. MBP-βGAL is used as a control. (**D**) UHRF1 and DNMT1 binding to PAX7 was assessed by pull-down assays using indicated MBP-PAX7 chimeras. WT, wild type. (**E** and **F**) PAX7 binding to indicated deletions of UHRF1 (E) and DNMT1 (F) was assessed in pull-down assays. (**G**) Summary of interactions between PAX7, UHRF1, and DNMT1. (**H**) Order-of-addition pull-down experiment performed with either MBP-PAX7 or MBP-βGal. ^{35}S-labeled UHRF1 or DNMT1 was added first as indicated.

and produce a Venus-like signal (green) upon excitation (*30*). Because we could not use a VN construct containing full-length PAX7 (fig. S3A) for competitive complementation (Fig. 3A), we used a VN-PAX7$_{35–273}$ construct that is sufficient for interaction with UHRF1 (Figs. 2D and 3C) but not for pioneer activity (*31*). As observed in Fig. 3D and quantitated in Fig. 3E, the blue CC-UHRF1/CN-DNMT1 fluorescence signal is progressively displaced by the green CC-UHRF1/VN-PAX7$_{35–273}$ signal (Fig. 3A) with increasing amounts of PAX7$_{35–273}$ until a plateau of green fluorescence is observed (Fig. 3E). PAX7$_{35–273}$ thus displaces UHRF1 from DNMT1 in living cells.

We next tested the effect of PAX7 on UHRF1-dependent activation of DNMT1 by performing enzyme assays with purified DNMT1 and UHRF1 (*26*). For competition, we used the high-performance

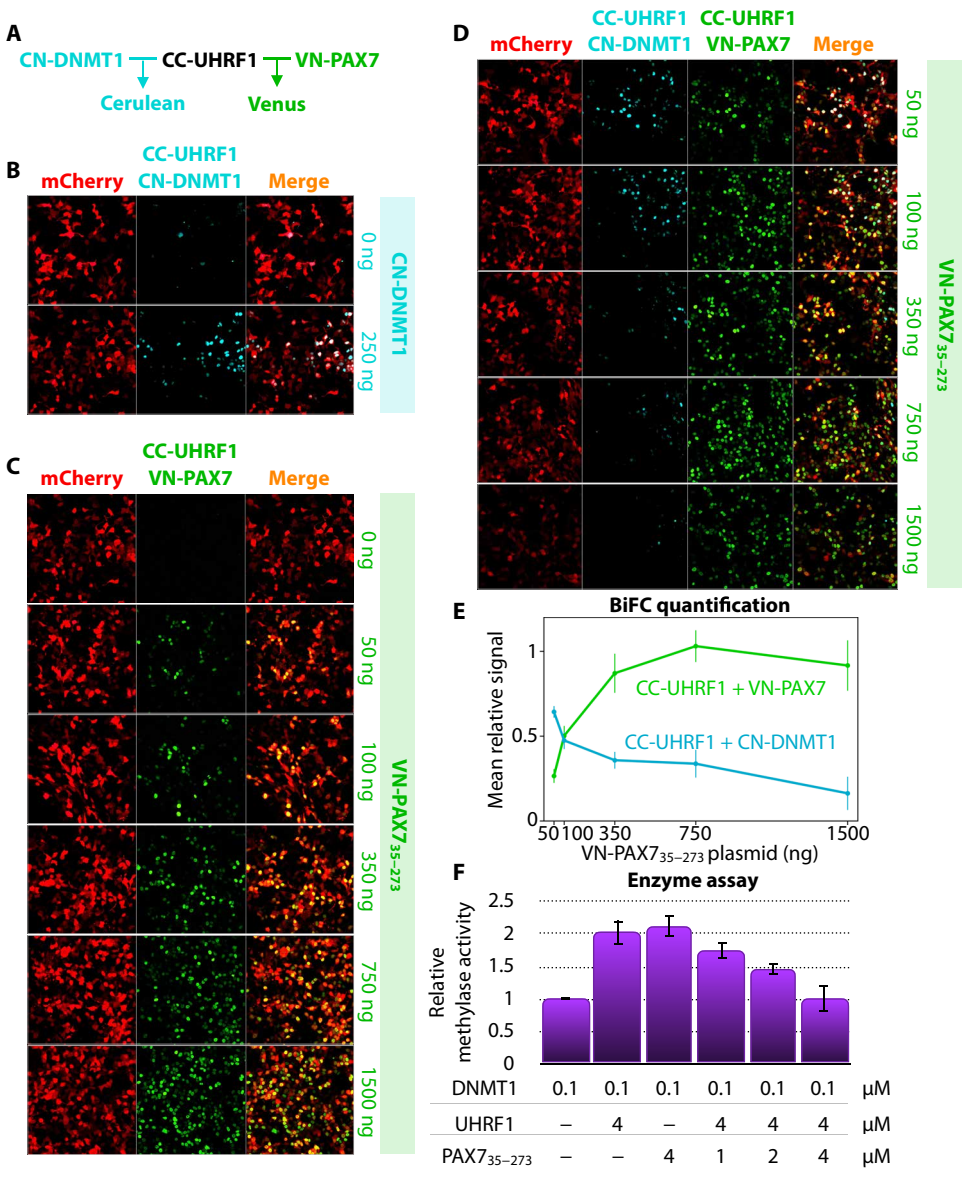

Fig. 3. PAX7 binding to UHRF1 displaces UHRF1:DNMT1 interaction and prevents DNMT1 activation. (**A**) Schema of fluorescence complementation between CC-UHRF1 and CN-DNMT1 or VN-PAX7$_{35–273}$. (**B**) BiFC between CC-UHRF1 and CN-DNMT1 (blue fluorescence) cotransfected with mCherry (red fluorescence) as a transfection control. (**C**) BiFC of CC-UHRF1 with increasing amounts (nanograms of plasmid per dish) of VN-PAX7$_{35–273}$ (green fluorescence), cotransfected with mCherry. (**D**) Bicolor BiFC between CC-UHRF1 and CN-DNMT1 (blue fluorescence) is competed by increasing concentrations of VN-PAX7$_{35–273}$ (green fluorescence). (**E**) Quantification of fluorescence signals (±SEM) in (D). (**F**) Relative DNA methylase activity of DNMT1 assessed with or without indicated concentrations of UHRF1, PAX7, or both.

liquid chromatography (HPLC)–purified PAX7$_{35–273}$ polypeptide that includes both DBDs and is sufficient for DNA binding (*19*). As previously shown (*26*), addition of UHRF1 increases the DNA methylation activity of DNMT1 (Fig. 3F). Addition of increasing amounts of PAX7$_{35–273}$ reversed this activation, with complete reversal observed at equimolar amounts of PAX7$_{35–273}$ and UHRF1 (Fig. 3F). These data support the hypothesis that PAX7 blocks DNMT1 activation by UHRF1 through its direct interaction with UHRF1. It is also noteworthy that PAX7 itself stimulates DNMT1 enzymatic activity: This could be explained by its binding to the DNMT1 RFD domain, because it was previously suggested that any protein binding to the RFD might cause activation of DNMT1 (*32*). This domain has an autoinhibitory effect on the DNMT1 catalytic domain as its binding by UHRF1 directly activates DNMT1 (*26*), so the mere binding of PAX7 to the RFD appears sufficient to deinhibit DNMT1 enzymatic activity as well. Of note, this effect is lost in the presence of UHRF1, suggesting that PAX7-UHRF1 binding is strong and that it prevents the interaction of the two individual proteins with DNMT1.

Involvement of TET demethylases

Active DNA demethylation is achieved by the TET enzymes (*33*, *34*) that catalyze the stepwise oxidation of 5mC (successively into 5hmC, 5fC, and finally, 5caC). 5fC and 5caC can be either replaced by the

base excision repair mechanism, resulting in an unmethylated cytosine, or diluted during replication because UHRF1 and DNMT1 are less efficient at recognizing these oxidized substrates (*14*, *35*, *36*). PF-induced DNA demethylation is not extensively studied, but some PFs involve TET enzymes (*8*, *37*–*39*). It is noteworthy that 5hmC may also be a stable epigenetic mark in specific contexts and is implicated in several processes such as gene regulation (*40*–*43*).

Although we did not identify any TET enzyme in our RIME analyses (Fig. 2, A and B, and table S1), it is possible that the assay is not sensitive enough to identify TET proteins: We thus compared the TET1/2/3 mRNA levels from AtT-20 cell RNA sequencing (RNA-seq) (*9*) to those of proteins detected in the assay (fig. S4A). This suggests borderline levels that may or not have been detectable in RIME. We next performed ChIP-seq against 5hmC in ER-PAX7 cells activated with Tam for 48 hours with or without mimosine to assess accumulation of 5hmC at Pioneer sites: We observed 5hmC accumulation, but it is replication dependent (Fig. 4A). This may be related to the greater affinity of TET enzymes for hemimethylated CpGs generated during replication and enriched after UHRF1-DNMT1 blockage when compared to fully methylated CpGs (*44*). Because TET3 has the highest mRNA level in AtT-20 cells (fig. S4B), we performed ChIP-quantitative polymerase chain reaction (qPCR) for this enzyme at a small subset of Pioneer sites that are rich in CpGs and where recruitment is detectable. These experiments showed weak TET3 recruitment at Pioneer sites but not control sites (Fig. 4B). To assess the time course of TET action, we measured accumulation of hmC by ChIP-seq. The time course shows accumulation of 5hmC over a 10-day period with the same half-time as chromatin opening measured by ATAC-seq (Fig. 4C and fig S4C). This stable accumulation is in contrast to the transient occurrence of 5hmC observed by Sardina *et al.* (*8*) during the TET2-dependent B cell reprogramming by the pioneers C/EBPa (CCAAT enhancer binding protein alpha) and KLF4 (Krueppel-like factor 4). It thus appears that TET enzymes may be active at PAX7 Pioneer sites but following replication, suggesting an indirect recruitment of these proteins. Under conditions of reduced UHRF1-DNMT1 activity, hemimethylated CpGs (5mC-C) may be methylated with a delay. Because hemimethylated CpGs are highly favored substrates of TET enzymes (*45*, *46*), and because conversion of 5mC-C to 5hmC-C strongly reduces DNMT1 activity (*14*), TET enzymes may serve to enhance DNA demethylation compared to simple dilution achieved by blockade of DNMT1 activity.

To ascertain this model, we used nanopore sequencing of genomic DNA at different times following Tam activation of ER-PAX7 (Fig. 4, D to G). The technology allows for direct sequencing of both strands of CpG dinucleotides within individual unamplified genomic DNA duplexes with the ability to identify C, 5mC, and 5hmC on each strand: We thus identified CpG dinucleotides bearing unique patterns of modifications (Fig. 4D). By real-time selection of sequenced duplexes from a panel of 4000 genomic loci containing 876 PAX7-pioneered sites (encompassing 2567 CpGs shown in Fig. 1C to undergo demethylation, >10% loss), we assessed changes in C methylation patterns over 48 hours following Tam activation of PAX7 (Fig. 4, F and G). The results show hydroxymethylation of CpG dinucleotides with particular accumulation of 5hmC-C dinucleotides that would derive from hemimethylated CpGs (5mC-C) produced at replication and of 5hmC-5mC dinucleotides generated by oxidation of methylated CpGs (5mC-5mC) (Fig. 4D); such accumulation is not seen at control CpGs that do not undergo demethylation (Fig. 4E). It is noteworthy that the presence of PAX7 is not sufficient to drive cytosine hydroxymethylation because PAX7 is recruited to Resistant sites that do not show changes in DNA methylation (Fig. 1B) and show no gain of 5hmC (fig. S4C). Analysis of sequences flanking CpGs (*14*, *45*, *46*) subject to PAX7-dependent demethylation indicates enrichment for sites of preferred TET action and of poor DNMT1 substrates, consistent with enhanced sensitivity to DNMT1 restriction (fig. S4D). It is interesting that a subset of already partially demethylated Pioneer enhancers ($n = 164$ of 876) never showed significant (P values $\geq 10^{-3}$) 5hmC ChIP-seq peaks during the 10-day test period but nonetheless exhibits similar CpG demethylation (Fig. 4H). These Pioneer sites may thus rely solely on blockade of methylation maintenance for PAX7-dependent demethylation. Collectively, the data support a model of efficient demethylation driven by both blockade of UHRF1-DNMT1 maintenance and recruitment of TET enzymes.

DISCUSSION

The present work uncovered a PF-specific mechanism for blockade of the DNA methylation maintenance system composed of UHRF1 and DNMT1. Previous work documented blockade of this mechanism on a broad genome-wide scale in early development by repression of UHRF1 expression (*47*). The involvement of a unique transcription factor in this mechanism (the UHRF1 association was detected with PAX7 but not with the associated non-PF TPIT; Fig. 2B) suggests that this may be an important evolutionary mechanism to implement lineage-specific long-term epigenetic memory. Epigenetic memory as implemented by DNA demethylation could contribute to establishment of transcriptional memory characterized by faster secondary transcriptional responses to transcription factors or signals or as we discussed previously (*6*).

The UHRF1-DNMT1 interaction leads to reversal of an autoinhibitory mechanism in DNMT1 (*26*). We showed that PAX7 interacts with the UHRF1 domains that overlap the UHRF1 domains involved in DNMT1 interaction and activation (Fig. 2, D to G) and that PAX7 competes the UHRF1-DNMT1 interaction in vivo (Fig. 3, D and E), thus blocking UHRF1-dependent activation of DNMT1 enzyme activity (Fig. 3F). Because the UHRF1-DNMT1 methylation maintenance system operates after DNA replication and PAX7-dependent enhancer DNA demethylation is dependent on replication, we may speculate that PAX7 is retained at the replication fork through its interaction with the replication machinery as our RIME data indicated PAX7 association with many proteins of the replication complex (Fig. 2, A and B, fig. S2A, and table S1). This interaction remains to be explored further, but it would ensure that PAX7 is present at the critical time when the maintenance function of UHRF1-DNMT1 is exerted. This blockade would result in dilution of CpG methylation during successive cell divisions. We found that PAX7-driven DNA demethylation at Pioneer sites also involves a replication-dependent oxidation of 5mC residues to 5hmC (Fig. 4A). Because hemimethylated CpGs, enriched after replication under UHRF1-DNMT1 inhibition, are the highest-affinity substrate for TET enzymes (*44*), we observed a significant increase in the 5hmC-C dinucleotide at Pioneer but not at other sites (Fig. 4, E and F). Hence, the postreplication action of TET enzymes downstream of UHRF1-DNMT1 blockage accelerates demethylation of enhancer CpGs beyond simple dilution (Fig. 4E) through replication cycles as oxidized cytosines are not efficiently recognized as methylated by the maintenance system (*14*, *17*). The implication of the TET system in

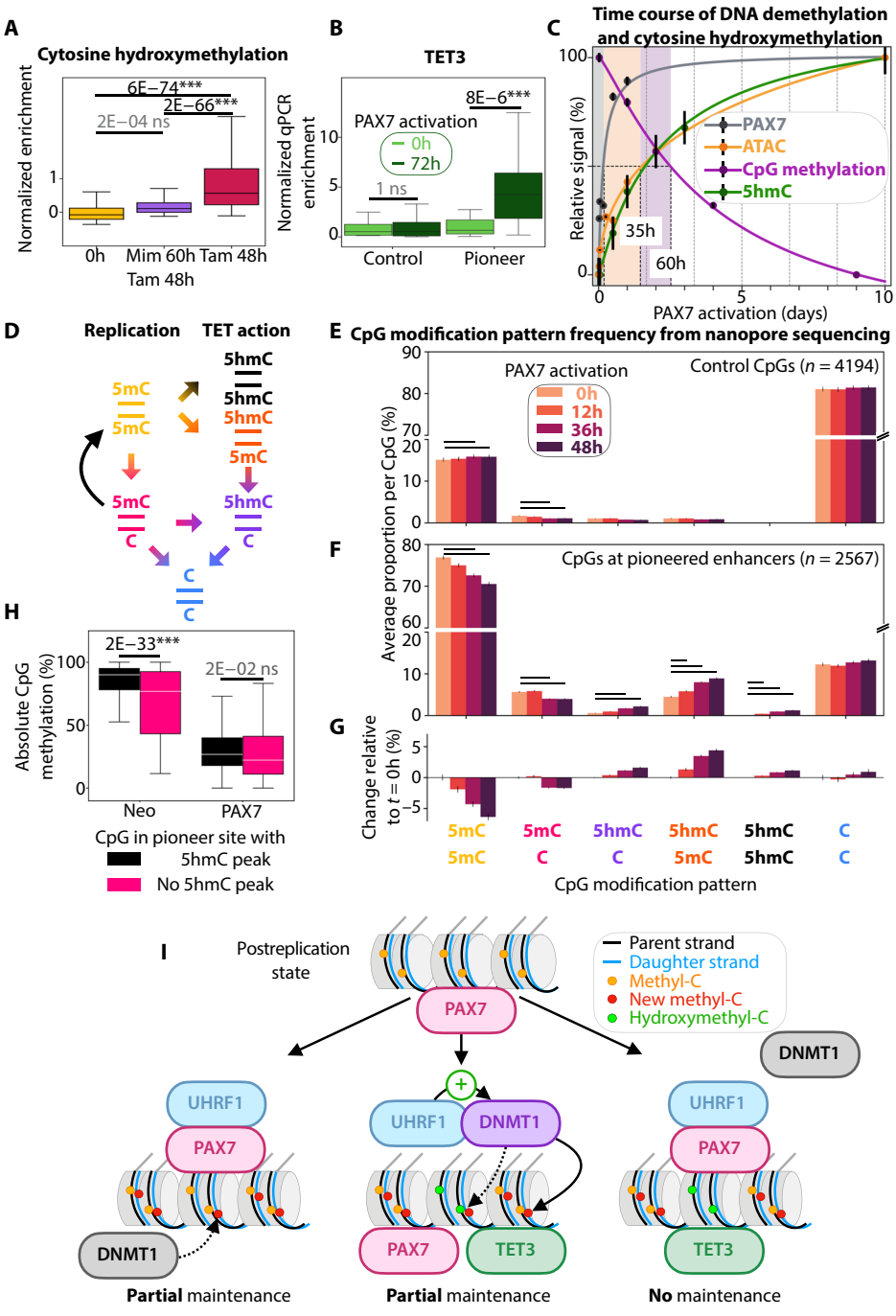

Fig. 4. Two mechanisms implement PAX7-driven DNA demethylation. (**A**) Box plot of normalized signals at PAX7 Pioneer enhancers for 5hmC ChIP-seq (±50 bp) after 0- or 48-hour Tam activation with/without mimosine cell cycle arrest. (**B**) TET3 recruitment (ChIP-qPCR) at control and PAX7 Pioneer enhancers [from (*20*)] before/after Tam activation for 72 hours. **(C)** Time course of 5hmC enrichment assessed by ChIP-seq (peak ±50 bp) at Pioneer enhancers compared to PAX7 recruitment (ChIP-seq), chromatin gain of accessibility (ATAC-seq), and DNA demethylation (WGBS). (**D**) Modified CpG dinucleotides predicted to occur after DNA replication (left) and following oxidation by TET enzymes (right). (**E**) Methylation patterns determined by nanopore sequencing of unamplified genomic DNA duplexes at control CpGs that are not demethylated (<1% change of absolute DNA methylation between Neo and PAX7 cells at PAX7 peak summits ±250 bp). (**F**) Methylation patterns at CpGs within Pioneer sites (PAX7 summits ±250 bp) that show demethylation (≥10%) after PAX7 activation. Significant changes compared to Tam at 0 hours (*P* values $\leq 10^{-5}$) are shown in (E) and (F) by black lines above relevant bars. *P* values are provided in table S2. (**G**) Relative change of occurrence for indicated CpG methylation patterns [(F) above] compared to Tam at 0 hours. (**H**) Box plots showing CpG methylation (%) before and after PAX7 activation at Pioneer enhancers that are positive (black) or not (pink) for 5hmC in ChIP-Seq at any time point assessed in (C). (**I**) Model for PAX7-driven enhancer DNA demethylation. Blockade of the UHRF1-DNMT1 DNA methylation maintenance system will lead to progressive CpG demethylation through dilution at each replication cycle (left). A postreplication TET action will lead to partial CpG methylation maintenance by hydroxymethylation of hemimethylated CpGs (center) that are poor substrates for UHRF1 and DNMT1 (*35*, *59*). The combined PAX7-dependent blockade of UHRF1 activation together with TET action would ensure efficient enhancer demethylation (right).

enhancer demethylation also involves oxidation of 5mC-5mC dinucleotides, but it appears nonetheless different compared to the transient and TET-dependent active mechanism reported for the action of C/EBPa and KLF4 (*8*). By acting at replication, the PAX7-dependent mechanism for DNA demethylation can operate as a safeguard to ensure the maintenance of the demethylated active state throughout cell generations of the PAX7-dependent lineage, such as pituitary melanotropes that maintain PAX7 expression throughout life.

MATERIALS AND METHODS

Plasmids

Plasmids were either constructed by enzymatic digestion and ligation or using the Gibson assembly method (*48*). Table S4 provides source plasmids and primers used for constructions.

Cell culture

AtT-20 cells were grown and selected as previously described (*18*) For AtT-20 ER-PAX7 cells, Tam (Sigma-Aldrich, H7904) was used at a final concentration of 400 nM with 0.1% ethanol, as described (*20*). For cells treated for more than 24 hours, the same quantity of Tam was added every 24 hours. Mimosine (Sigma-Aldrich, M0253) was added at a final concentration of 200 μM, as validated previously (*20*).

Assay for transposase-accessible chromatin (ATAC)

ATAC was performed as described (*9*, *20*, *22*). Nuclei (50,000) were isolated by incubation at 4°C first for 30 min in hypotonic cell lysis buffer 1 [0.1% (w/v) sodium citrate tribasic dihydrate and 0.1% (v/v) Triton X-100] followed by 30 min in buffer 2 [10 mM tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, and 0.1% (v/v) IGEPAL CA-630]. Nucleus pellets were tagmented for 30 min at 37°C (TDE1 Transposase, Illumina) using the Tagment DNA kit, following Illumina's recommendations. Tagmented DNA was purified on a DNA clean-5 column (Zymo Research). Illumina UDP Indexes were integrated by PCR enrichment (12 cycles), and final libraries were purified and size selected ($L = 1.1$; $R = 0.6$) with KAPA pure beads (Roche Diagnostics).

Chromatin immunoprecipitation (ChIP)

ChIP-qPCR and ChIP-seq were performed and analyzed as previously described (*20*). At least two independent ChIPs were pooled before library preparation.

On day 1, 20 μl of Dynabeads-Protein A (Invitrogen, no. 10001D) and 20 μl of Dynabeads-Protein G (Invitrogen, no. 10003D) were coupled with the indicated antibody overnight at 4°C in ChIP Dilution Buffer (1% Triton X-100, 0 mM tris, pH 8, 150 mM NaCl, and 2 mM EDTA). On day 2, cells were cross-linked at 1% formaldehyde at room temperature for 10 min, and the reaction was quenched for 5 min with 125 mM glycine in phosphate-buffered saline (PBS). Cells were then incubated for 5 min in buffer A (10 mM tris-HCl, pH 8, 0.25% Triton X-100, 10 mM EDTA, and 0.5 mM EGTA) followed by 30 min on ice in buffer B (10 mM tris-HCl, pH 8, 200 mM NaCl, 1 mM EDTA, and 0.5 mM EGTA). To generate fragments between 100 and 500 bp, sonication was performed using Covaris E220 in the sonication buffer (10 mM tris-HCl, pH 8, 140 mM NaCl, 0.5% Triton X-100, 0.5% SDS, 1 mM EDTA, and 0.5 mM EGTA). Chromatin was then incubated overnight at 4°C with the antibody-coupled magnetic beads.

On day 3, beads were washed four times (1: 1% Triton X-100, 0.1% SDS, 150 mM NaCl, 20 mM tris-HCl, pH 8, and 2 mM EDTA; 2: 1% Triton X-100, 0.1% SDS, 500 mM NaCl, 20 mM tris-HCl, pH 8, and 2 mM EDTA; 3: 1% NP-40, 250 mM LiCl, 10 mM tris-HCl, pH 8, and 1 mM EDTA; 4: 50 mM NaCl, 10 mM tris-HCl, pH 8, and 1 mM EDTA). Cross-linking was then reversed overnight at 65°C in elution buffer (50 mM tris-HCl, pH 8, 1% SDS, and 2 mM EDTA).

On day 4, DNA was incubated for 15 min at room temperature with ribonuclease A (0.1 mg/ml) followed by 1 hour at 65°C with proteinase K (0.8 mg/ml). DNA fragments were then purified using the QIAquick PCR Purification Kit (Qiagen, no. 28106). The libraries and flow cells were prepared by the IRCM Molecular Biology Core Facility according to Illumina's recommendations. Size distribution and the molarity of immunoprecipitated and input samples were evaluated on a 2100 Bioanalyzer (Agilent Technologies). ChIP-seq libraries were prepared using the KAPA HyperPrep library kit (Roche Diagnostics).

ChIP-seq and ATAC-seq analysis

Paired-end reads were first trimmed for adapter content and aligned to the mouse mm10/GRCm38 reference genome using Bowtie 2.4 (--no-unal --no-mixed --no-discordant) (*49*). Bam files were created, and duplicated reads were removed using view (-b) and markdup (-r) from SAMtools 1.12 (*50*). Reads from multiple biological replicates were merged (if applicable). Coverage tracks were created using bamCoverage (-bs 10 -e 150 --normalizeUsing BPM) from deepTools 3.5 (*51*). Peaks were called using callpeak (-f BAMPE -p 1e-3/1e-5 –g mm --min-length 100 -max-gap 50) from MACS 2.2.7.1 (*52*). Local signals were normalized relative to control loci (constitutively active enhancers and random genomic loci).

ChIP-bis-seq and WGBS

CHIP-bis-seq was performed like ChIP-seq. Six (24-hour ER-PAX7 Tam, 48-hour Tam, and 96-hour Tam) or 12 (9-day ER-PAX7 Tam, 30-min Tam, and 10-day Tam followed by no Tam for 20 days) independent ChIPs were pooled before bisulfite conversion and library preparation. For the 10-day Tam sample followed by the 20-day no-Tam sample, cells were treated with Tam for 30 min at the end of the experiment to allow ChIP. For WGBS, genomic DNA was extracted from 1,000,000 AtT-20Neo, AtT-20PAX7, and AtT-20 ER-PAX7 (untreated, 60-hour mimosine + 48-hour Tam, and 48-hour Tam) cells with the Qiagen DNeasy Blood & Tissue kit (Qiagen, no. 69504). Bisulfite conversion was done with the EZ DNA Methylation-Lightning kit (Zymo Research).

ChIP-bis-seq and WGBS analysis

Paired-end reads were first trimmed for adapter content and aligned to the mouse mm10/GRCm38 reference genome using Bismark 0.22 (--non_directional) (*53*). Bam files were created, and duplicated reads were removed using view (-b) and markdup (-r) from SAMtools 1.12 (*50*). Coverage tracks were created using bamCoverage (-bs 10 -e 150 –normalizeUsing BPM) from deepTools 3.5 (*51*). Peaks were called using callpeak (-f BAMPE -p 1e-3/1e-5 -g mm –min-length 100 –max-gap 50) from MACS 2.2.7.1 (*52*). Cytosine methylation was called using Bismark during alignment. CpG methylation was

extracted from bam files, and CpGs with a minimum coverage of 10 reads were used for analysis.

Library quality controls and sequencing (for all libraries)

The DNA libraries and sequencing flow cells were prepared by the IRCM Molecular Biology Core Facility. Library size distribution was assessed on a 2100 Bioanalyzer (Agilent Technologies). Normalization of the sample quantities was done after quantification of the ligation products by qPCR. Equimolar libraries were sequenced at Centre d'expertise et de services Génome Québec in paired-end reads (see table S3). Table S3 provides information on each library, including antibodies, sequencing depth, read length, and sequencing instrument.

Pull-down and order-of-addition assays

MBP proteins were expressed in *Escherichia coli* after IPTG induction, purified on amylose beads, and used as described (*22*). ^{35}S-labeled proteins were produced using either T7 TnT Coupled Reticulocyte Lysate Systems (L4610), T3 TnT Coupled Reticulocyte Lysate Systems (L4950), or TnT T7 Quick for PCR DNA (L5540) kits from Promega, following the manufacturer's protocols. Quantities of all ^{35}S-labeled and MBP proteins were assessed before pull-down assays to ensure that similar amounts were used. ^{35}S-labeled proteins were incubated for 4 hours with MBP proteins, followed by migration on 6 to 10% polyacrylamide gels. Gels were stained using Coomassie blue to confirm similar MBP protein loading before being dried and exposed to a low-energy autoradiography film. Order-of-addition pull-down assays were performed in the same way as pull-down assays, with a minor variation. ^{35}S-labeled proteins were incubated for either 4 hours (first protein introduced) or 2 hours (second protein introduced).

Bimolecular fluorescence complementation (BiFC)

Cell culture

HEK-293 cells were obtained from the American Type Culture Collection through LGC Standards Sarl (FR). Cells were cultured in Dulbecco's modified Eagle's medium (DMEM-GlutaMAX-I, Gibco by Life Technologies) supplemented with 10% (v/v) heat-inactivated fetal bovine serum, 1% (v/v) penicillin-streptomycin [5000 U penicillin and streptomycin (5 mg/ml)], and 5 μl of 100 mM sodium pyruvate (Gibco by Life Technologies), incubating at 37°C, in an atmosphere of 5% CO_2. In a six-well plate, 350,000 HEK-293 cells were seeded in each coverslip-containing well in 2 ml of culture media and incubated at 37°C in an atmosphere of 5% CO_2 for 16 to 24 hours or until cells reach a confluence of 60 to 80%.

Transfection

A four-plasmid transfection system is used in the BiFC assay for each PAX7 construct condition. In a 1.5-ml Eppendorf, 1 μl of each VN/CC/CN/mCherry-containing plasmid solution was added into 200 μl of jetPRIME buffer. The solution is then vortexed and spun down. Of note, an empty vector was used for compensation when needed. A ratio of 1:2 (plasmid DNA:jetPRIME reagent) was added to the mix followed by a brief vortex and spin-down. The transfection solution was incubated for 20 min at room temperature. Gently, 200 μl of transfection mix was added to a HEK-293–containing well and distributed evenly. The plates were incubated for 20 hours at 37°C and 5% CO_2, followed by fluorescence confocal imaging.

Microscopy and imaging

All the fluorescence microscopy images of cells were captured using confocal Zeiss LSM780. Images were recorded at a resolution of 1024 by 1024 pixels using a 20× objective and a pinhole of 1.0 Airy unit. All samples were imaged using identical settings, with argon as a laser source (488/561/633), using the 561-, 488-, and 633 nm laser lines for the excitation of mCherry, VN/CC, and CC/CN BiFCs, respectively. The detector gain was set at 700, and laser power was set at 2.0% (mCherry detection), 3.0% (VN/CC BiFC detection), or 10.0% (CC/CN Cerulean detection). A *z* axis was used to define the first and last sections. For each slide, at least three images were captured from different areas, ensuring that each image includes at least 100 cells. The experiment was performed in triplicates. Contrasts of all images were modified in the same way for figures.

Quantitative bicolor BiFC

All the images were quantified using Fiji software after several parameters were set up, choosing the Hyperstack and colorized in the Bio-formats Import options of the tool. Then, using the *z*-projection function, the maximum intensity projection was selected. For each condition, three fields were used, and in each field, different values for five cells were measured: the mCherry mean value, VN/CC mean value, and CC/CN mean value. The final normalized VN/CC value and final normalized CC/CN value were calculated by dividing with the mCherry fluorescence mean value.

Enzyme essays

The HPLC-purified $PAX7_{35–273}$ polypeptide was obtained from (*19*). Mouse DNMT1 and UHRF1 were purified as described (*26*). Radioactive DNA methylation assays of an oligonucleotide containing one hemimethylated CpG site were conducted as described.

Nanopore DNA sequencing

Genomic DNA was extracted from 1,000,000 AtT-20 ER-PAX7 cells treated with Tam for 0, 12, 36, or 48 hours with the Qiagen DNeasy Blood & Tissue kit. Long-read sequencing used the PromethION P24 with R10 chemistry flow cells together with the library preparation kit SQK-LSK-114 following the manufacturer's recommendations. The DNA was initially sheared to 8 kb to increase the sequencing yield using gTubes (Covaris), and the shearing parameters were 6600 rpm for 1 min with 60 ng/μl in 150 μl. Sequencing runs were set up with high-accuracy basecalling and adaptive sampling with a bed file of 4000 sites encompassing PAX7-bound regions ±16 kb (total size, 31 kb) with the mm10/GrCm38 reference. Each library was sequenced on one flow cell over 3 days. After the adaptive sampling runs, the pod5 raw files were rebasecalled with Dorado (0.5.3) in superaccuracy (SUP) and duplex mode. For basecalling, the pod files were split by the channel to increase the duplex calling processing speed (pod5 subset/see split_by_channel). The unaligned bam was then aligned with the Dorado aligner followed by modkit (0.2.6) to extract the double-stranded methylation calls in bed format. From the bed files, CpGs in a ±250-bp window of PAX7 peak summits with at least three duplex reads in all samples were used for analysis. Proportion of each CpG modification pattern was then assessed by locus, and these proportions were averaged to calculate the mean proportions and SEM.

RIME and analysis

RIME experiments were performed as described (*21*). Each sample was done in triplicates (Tpit on AtT-Neo) or quadruplicates (FlagM2 on AtT-PAX7 and AtT-Neo).

On day 1, 20 µl of Dynabeads-Protein A (Invitrogen, no. 10001D) and 20 µl of Dynabeads-Protein G (Invitrogen, no. 10003D) were incubated with either FlagM2 (Sigma-Aldrich, F3165) or Tpit (*54*) antibody in ChIP dilution buffer (1% Triton X-100, 10 mM tris, pH 8, 150 mM NaCl, and 2 mM EDTA) overnight at 4°C. On day 2, cells were cross-linked for 10 min at 1% formaldehyde at room temperature. Cross-linking was quenched with 200 mM glycine in PBS for 5 min. Cells were washed and harvested in cold PBS and then washed two additional times with PBS. Cells were then incubated for 10 min at 4°C in 10 ml of LB1 [50 mM Hepes-KOH (pH 7.5), 140 mM NaCl, 1 mM EDTA, 10% (v/v) glycerol, 0.5% (v/v) NP-40/IGEPAL CA-630, and 0.25% (v/v) Triton X-100]. Pellets were then incubated for 5 min at 4°C in 10 ml of LB2 [10 mM tris-HCl (pH 8.0), 200 mM NaCl, 1 mM EDTA, and 0.5 mM EGTA] and resuspended in 970 µl of LB3 [10 mM tris-HCl (pH 8.0), 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% (w/v) sodium deoxycholate, and 0.5% (v/v) *N*-lauroylsarcosine] for sonication. Sonication was performed using Covaris E220, and 1 ml of LB3 and 100 µl of 10% Triton X-100 were added. Lysates were centrifuged for 10 min at 10,000 rpm at 4°C and then incubated with magnetic beads overnight at 4°C.

On day 3, beads were washed 10 times with radioimmunoprecipitation assay buffer [50 mM Hepes (pH 7.6), 1 mM EDTA, 0.7% (w/v) sodium deoxycholate, 1% (v/v) NP-40, and 0.5 M LiCl] and then four times with ammonium hydrogen carbonate buffer (100 mM AMBIC, no. A6141, dissolved in HPLC-quality water). Last, beads were suspended in 50 µl of AMBIC buffer and sample preparation was performed by the IRCM Mass Spectrometry and Proteomics facility.

On-bead digestion

The on-bead proteins were first incubated for 10 min in 4 M urea/50 mM ammonium Bicarbonate and then diluted to 2 M urea/50 mM ammonium bicarbonate. On-bead trypsin digestion was performed overnight at 37°C. The samples were then reduced with 13 mM dithiothreitol at 37°C and, after cooling for 10 min, alkylated with 23 mM iodoacetamide at room temperature for 20 min in the dark. The supernatant was acidified with trifluoroacetic acid for desalting and removal of residual detergents by MCX (Waters Oasis MCX 96-well Elution Plate) following the manufacturer's instructions. After elution in 10% ammonium hydroxide/90% methanol, samples were dried and reconstituted under agitation for 15 min in 150 µl of 2% acetonitrile-1% formic acid.

LC-MS/MS analysis

The following procedures were previously described in (*55*). The liquid chromatography column was a PicoFrit fused silica capillary column (15-cm by 75-µm inside diameter; New Objective, Woburn, MA), self-packed with C-18 reverse-phase material (Jupiter 5-µm particles, 300-Å pore size; Phenomenex, Torrance, CA) using a high-pressure packing cell. This column was installed on the Easy-nLC II system (Proxeon Biosystems, Odense, Denmark) and coupled to the Q Exactive (Thermo Fisher Scientific, Bremen, Germany) equipped with a Proxeon Nanoelectrospray Flex ion source. The buffers used for chromatography were 0.2% formic acid (buffer A) and 100% acetonitrile/0.2% formic acid (buffer B). Peptides were loaded on-column at a flow rate of 600 nl/min and eluted with a two-slope gradient at a flow rate of 250 nl/min. Solvent B first increased from 1 to 40% in 52 min and then from 40 to 80% B in 18 min. Liquid chromatography–tandem mass spectrometry (LC-MS/MS) data were acquired using a data-dependent top 16 method combined with a dynamic exclusion window of 7 s. The mass resolution for full MS scan was set to 60,000 (at m/z 400), and lock masses were used to improve mass accuracy. The mass range was from 360 to 1800 m/z for MS scanning with a target value at 1×10^6, the maximum ion fill time at 100 ms, the intensity threshold at 1.0×10^4, and the underfill ratio at 0.5%. The data-dependent MS2 scan events were acquired at a resolution of 17,500 with the maximum ion fill time at 50 ms and the target value at 1×10^5. The normalized collision energy used was at 27, and the capillary temperature was 250°C. Nanospray and S-lens voltages were set to 1.3 to 1.7 kV and 50 V, respectively.

Protein identification

Protein database searches were performed using Mascot 2.3 (Matrix Science) against *Mus musculus*. The mass tolerance for precursor ions was set to 10 parts per million, and that for fragment ions was set to 0.5 Da. The enzyme specified was trypsin, and one missed cleavage was allowed. Cysteine carbamidomethylation was specified as fixed modification and methionine oxidation as variable modification. Scaffold (version Scaffold_5.3.1, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability by the Peptide Prophet algorithm (*56*) with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (*57*). Proteins that contained similar peptides and could not be differentiated on the basis of MS/MS analysis alone were grouped to satisfy the principles of parsimony. For each identified protein, a variation coefficient was calculated (standard deviation/mean) for the AtT-PAX7 samples. Proteins with a variation coefficient $\leq$0.75 were selected for further analysis and figures.

RNA-seq and analysis

For each biological RNA-seq duplicate, RNA was extracted from 1,000,000 cells in culture using the Qiagen RNeasy Plus Mini kit (Qiagen, no. 74134). Ribosomal depletion, library preparation, and flow cell preparation for sequencing were performed by the IRCM Molecular Biology Core Facility according to Illumina's recommendations. RNA libraries were prepared from total RNA. Ribosomal RNA was depleted using the KAPA RiboErase Kit (HMR), and libraries were prepared with the KAPA RNA HyperPrep Kit (Roche Diagnostics). Paired-end reads were first trimmed for adapter content and aligned to the mouse mm10/GRCm38 reference genome using Hisat 2.2 (--no-unal --no-mixed --no-discordant) (*58*). Bam files were created using view (-b) from SAMtools 1.12 (*50*). Coverage tracks were created using bamCoverage (-bs 10 -e 150 --normalizeUsing BPM) from deepTools 3.5 (*51*). Gene expression was assessed by integrating reads mapped to exons over transcript length and normalized across samples with the median-of-ratio method. Expression levels were then averaged between replicates and used for plotting.

Supplementary Materials

The PDF file includes:

Figs. S1 to S4
Legends for tables S1, S3, and S4
Table S2
References

Other Supplementary Material for this manuscript includes the following:
Tables S1, S3, and S4

REFERENCES AND NOTES

1. Z. D. Smith, A. Meissner, DNA methylation: Roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220 (2013).
2. A. P. Bird, CpG-rich islands and the function of DNA methylation. *Nature* **321**, 209–213 (1986).
3. A. Pacis, F. Mailhot-Léonard, L. Tailleux, H. E. Randolph, V. Yotova, A. Dumaine, J. C. Grenier, L. B. Barreiro, Gene activation precedes DNA demethylation in response to infection in human dendritic cells. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 6938–6943 (2019).
4. U. Jadhav, A. Cavazza, K. K. Banerjee, H. Xie, N. K. O'Neill, V. Saenz-Vash, Z. Herbert, S. Madha, S. H. Orkin, H. Zhai, R. A. Shivdasani, Extensive recovery of embryonic enhancer and gene memory stored in hypomethylated enhancer DNA. *Mol. Cell* **74**, 542–554.e5 (2019).
5. S. Orlanski, V. Labi, Y. Reizel, A. Spiro, M. Lichtenstein, R. Levin-Klein, S. B. Koralov, Y. Skversky, K. Rajewsky, H. Cedar, Y. Bergman, Tissue-specific DNA demethylation is required for proper B-cell differentiation and function. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5018–5023 (2016).
6. A. Balsalobre, J. Drouin, Pioneer factors as master regulators of the epigenome and cell fate. *Nat. Rev. Mol. Cell Biol.* **23**, 449–464 (2022).
7. J. Donaghey, S. Thakurela, J. Charlton, J. S. Chen, Z. D. Smith, H. Gu, R. Pop, K. Clement, E. K. Stamenova, R. Karnik, D. R. Kelley, C. A. Gifford, D. Cacchiarelli, J. L. Rinn, A. Gnirke, M. J. Ziller, A. Meissner, Genetic determinants and epigenetic effects of pioneer-factor occupancy. *Nat. Genet.* **50**, 250–258 (2018).
8. J. L. Sardina, S. Collombet, T. V. Tian, A. Gómez, B. Di Stefano, C. Berenguer, J. Brumbaugh, R. Stadhouders, C. Segura-Morales, M. Gut, I. G. Gut, S. Heath, S. Aranda, L. Di Croce, K. Hochedlinger, D. Thieffry, T. Graf, Transcription factors drive Tet2-mediated enhancer demethylation to reprogram cell fate. *Cell Stem Cell* **23**, 727–741.e9 (2018).
9. A. Mayran, K. Khetchoumian, F. Hariri, T. Pastinen, Y. Gauthier, A. Balsalobre, J. Drouin, Pioneer factor Pax7 deploys a stable enhancer repertoire for specification of cell fate. *Nat. Genet.* **50**, 259–269 (2018).
10. Y. Reizel, A. Morgan, L. Gao, J. Schug, S. Mukherjee, M. F. García, G. Donahue, J. A. Baur, K. S. Zaret, K. H. Kaestner, FoxA-dependent demethylation of DNA initiates epigenetic memory of cellular identity. *Dev. Cell* **56**, 602–612.e4 (2021).
11. T. M. Escobar, O. Oksuz, R. Saldaña-Meyer, N. Descostes, R. Bonasio, D. Reinberg, Active and repressed chromatin domains exhibit distinct nucleosome segregation during DNA replication. *Cell* **179**, 953–963.e11 (2019).
12. M. Bostick, J. K. Kim, P. O. Esteve, A. Clark, S. Pradhan, S. E. Jacobsen, UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760–1764 (2007).
13. J. Sharif, M. Muto, S. Takebayashi, I. Suetake, A. Iwamatsu, T. A. Endo, J. Shinga, Y. Mizutani-Koseki, T. Toyoda, K. Okamura, S. Tajima, K. Mitsuya, M. Okano, H. Koseki, The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**, 908–912 (2007).
14. S. Adam, V. Klingel, N. E. Radde, P. Bashtrykov, A. Jeltsch, On the accuracy of the epigenetic copy machine: Comprehensive specificity analysis of the DNMT1 DNA methyltransferase. *Nucleic Acids Res.* **51**, 6622–6633 (2023).
15. J. Arand, M. Wossidlo, K. Lepikhov, J. R. Peat, W. Reik, J. Walter, Selective impairment of methylation maintenance is the major cause of DNA methylation reprogramming in the early embryo. *Epigenetics Chromatin* **8**, 1 (2015).
16. Y. Zeng, T. Chen, DNA methylation reprogramming during mammalian development. *Genes* **10**, 257 (2019).
17. C.-J. Lio, X. Yue, I. F. Lopez-Moyado, M. Tahiliani, L. Aravind, A. Rao, TET methylcytosine oxidases: New insights from a decade of research. *J. Biosci.* **45**, 21 (2020).
18. L. Budry, A. Balsalobre, Y. Gauthier, K. Khetchoumian, A. L'Honore, S. Vallette, T. Brue, D. Figarella-Branger, B. Meij, J. Drouin, The selector gene Pax7 dictates alternate pituitary cell fates through its pioneer action on chromatin remodeling. *Genes Dev.* **26**, 2299–2310 (2012).
19. A. Pelletier, A. Mayran, A. Gouhier, J. G. Omichinski, A. Balsalobre, J. Drouin, Pax7 pioneer factor action requires both paired and homeo DNA binding domains. *Nucleic Acids Res.* **49**, 7424–7436 (2021).
20. A. Gouhier, J. Dumoulin-Gagnon, V. Lapointe-Roberge, J. Harris, A. Balsalobre, J. Drouin, Pioneer factor Pax7 initiates two-step cell-cycle-dependent chromatin opening. *Nat. Struct. Mol. Biol.* **31**, 92–101 (2024).
21. H. Mohammed, C. Taylor, G. D. Brown, E. K. Papachristou, J. S. Carroll, C. S. D'Santos, Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat. Protoc.* **11**, 316–326 (2016).
22. A. Mayran, K. Sochodolsky, K. Khetchoumian, J. Harris, Y. Gauthier, A. Bemmo, A. Balsalobre, J. Drouin, Pioneer and nonpioneer factor cooperation drives lineage specific chromatin opening. *Nat. Commun.* **10**, 3807 (2019).
23. S. Biswas, K. Kang, K. P. Ng, T. Radivoyevitch, K. Schalper, H. Zhang, D. J. Lindner, A. Thomas, D. MacPherson, B. Gastman, D. S. Schrump, K. K. Wong, V. Velcheti, Y. Saunthararajah, Neuroendocrine lineage commitment of small cell lung cancers can be leveraged into p53-independent non-cytotoxic therapy. *Cell Rep.* **42**, 113016 (2023).
24. D. C. Di Giammartino, A. Kloetgen, A. Polyzos, Y. Liu, D. Kim, D. Murphy, A. Abuhashem, P. Cavaliere, B. Aronson, V. Shah, N. Dephoure, M. Stadtfeld, A. Tsirigos, E. Apostolou, KLF4 is involved in the organization and regulation of pluripotency-associated three-dimensional enhancer networks. *Nat. Cell Biol.* **21**, 1179–1190 (2019).
25. M. R. Rafiee, C. Girardot, G. Sigismondo, J. Krijgsveld, Expanding the circuitry of pluripotency by selective isolation of chromatin-associated proteins. *Mol. Cell* **64**, 624–635 (2016).
26. P. Bashtrykov, G. Jankevicius, R. Z. Jurkowska, S. Ragozin, A. Jeltsch, The UHRF1 protein stimulates the activity and specificity of the maintenance DNA methyltransferase DNMT1 by an allosteric mechanism. *J. Biol. Chem.* **289**, 4106–4115 (2014).
27. M. Achour, X. Jacq, P. Ronde, M. Alhosin, C. Charlot, T. Chataigneau, M. Jeanblanc, M. Macaluso, A. Giordano, A. D. Hughes, V. B. Schini-Kerth, C. Bronner, The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression. *Oncogene* **27**, 2187–2197 (2008).
28. A. Mayran, A. Pelletier, J. Drouin, Pax factors in transcription and epigenetic remodelling. *Semin. Cell Dev. Biol.* **44**, 135–144 (2015).
29. M. Duffraisse, B. Hudry, S. Merabet, Bimolecular fluorescence complementation (BiFC) in live Drosophila embryos. *Methods Mol. Biol.* **1196**, 307–318 (2014).
30. C. D. Hu, T. K. Kerppola, Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat. Biotechnol.* **21**, 539–545 (2003).
31. V. Bascunana, A. Pelletier, A. Gouhier, A. Bemmo, A. Balsalobre, J. Drouin, Chromatin opening ability of pioneer factor Pax7 depends on unique isoform and C-terminal domain. *Nucleic Acids Res.* **51**, 7254–7268 (2023).
32. F. Syeda, R. L. Fagan, M. Wean, G. V. Avvakumov, J. R. Walker, S. Xue, S. Dhe-Paganon, C. Brenner, The replication focus targeting sequence (RFTS) domain is a DNA-competitive inhibitor of Dnmt1. *J. Biol. Chem.* **286**, 15344–15351 (2011).
33. J. A. Hackett, R. Sengupta, J. J. Zylicz, K. Murakami, C. Lee, T. A. Down, M. A. Surani, Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* **339**, 448–452 (2013).
34. S. Yamaguchi, L. Shen, Y. Liu, D. Sendler, Y. Zhang, Role of Tet1 in erasure of genomic imprinting. *Nature* **504**, 460–464 (2013).
35. M. Schneider, C. Trummer, A. Stengl, P. Zhang, A. Szwagierczak, M. C. Cardoso, H. Leonhardt, C. Bauer, I. Antes, Systematic analysis of the binding behaviour of UHRF1 towards different methyl- and carboxylcytosine modification patterns at CpG dyads. *PLOS ONE* **15**, e0229144 (2020).
36. T. Zhou, J. Xiong, M. Wang, N. Yang, J. Wong, B. Zhu, R. M. Xu, Structural basis for hydroxymethylcytosine recognition by the SRA domain of UHRF2. *Mol. Cell* **54**, 879–886 (2014).
37. Y. Costa, J. Ding, T. W. Theunissen, F. Faiola, T. A. Hore, P. V. Shliaha, M. Fidalgo, A. Saunders, M. Lawrence, S. Dietmann, S. Das, D. N. Levasseur, Z. Li, M. Xu, W. Reik, J. C. Silva, J. Wang, NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* **495**, 370–374 (2013).
38. J. Dubois-Chevalier, F. Oger, H. Dehondt, F. F. Firmin, C. Gheeraert, B. Staels, P. Lefebvre, J. Eeckhoute, A dynamic CTCF chromatin binding landscape promotes DNA hydroxymethylation and transcriptional induction of adipocyte differentiation. *Nucleic Acids Res.* **42**, 10943–10959 (2014).
39. L. de la Rica, J. Rodríguez-Ubreva, M. García, A. B. Islam, J. M. Urquiza, H. Hernando, J. Christensen, K. Helin, C. Gómez-Vaquero, E. Ballestar, PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. *Genome Biol.* **14**, R99 (2013).
40. H. Wu, A. C. D'Alessio, S. Ito, Z. Wang, K. Cui, K. Zhao, Y. E. Sun, Y. Zhang, Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev.* **25**, 679–684 (2011).
41. M. A. Hahn, R. Qiu, X. Wu, A. X. Li, H. Zhang, J. Wang, J. Jui, S. G. Jin, Y. Jiang, G. P. Pfeifer, Q. Lu, Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. *Cell Rep.* **3**, 291–300 (2013).
42. Y. Huang, L. Chavez, X. Chang, X. Wang, W. A. Pastor, J. Kang, J. A. Zepeda-Martínez, U. J. Pape, S. E. Jacobsen, B. Peters, A. Rao, Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1361–1366 (2014).
43. A. Tsagaratou, T. Äijö, C. W. Lio, X. Yue, Y. Huang, S. E. Jacobsen, H. Lähdesmäki, A. Rao, Dissecting the dynamic changes of 5-hydroxymethylcytosine in T-cell development and differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3306–E3315 (2014).
44. S. He, F. Wang, Y. Zhang, J. Chen, L. Liang, Y. Li, M. Zhang, X. Yang, H. Pang, Y. Li, X. Huang, D. Qin, D. Pei, H. Sun, H. Zheng, Hemi-methylated CpG sites connect Dnmt1-knockdown-induced and Tet1-induced DNA demethylation during somatic cell reprogramming. *Cell Discov.* **5**, 11 (2019).
45. S. Adam, H. Anteneh, M. Hornisch, V. Wagner, J. Lu, N. E. Radde, P. Bashtrykov, J. Song, A. Jeltsch, DNA sequence-dependent activity and base flipping mechanisms of DNMT1 regulate genome-wide DNA methylation. *Nat. Commun.* **11**, 3723 (2020).

46. S. Adam, J. Bräcker, V. Klingel, B. Osteresch, N. E. Radde, J. Brockmeyer, P. Bashtrykov, A. Jeltsch, Flanking sequences influence the activity of TET1 and TET2 methylcytosine dioxygenases and affect genomic 5hmC patterns. *Commun. Biol.* **5**, 92 (2022).
47. S. Kagiwada, K. Kurimoto, T. Hirota, M. Yamaji, M. Saitou, Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J.* **32**, 340–353 (2013).
48. D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison III, H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
49. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
50. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, 1000 Genome Project Data Processing Subgroup, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
51. F. Ramírez, D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar, T. Manke, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
52. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, X. S. Liu, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
53. F. Krueger, S. R. Andrews, Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).
54. A.-M. Pulichino, S. Vallette-Kasic, J. P. Tsai, C. Couture, Y. Gauthier, J. Drouin, Tpit determines alternate fates during pituitary cell differentiation. *Genes Dev.* **17**, 738–747 (2003).
55. Z. Butti, Y. E. Pan, J. Giacomotto, S. A. Patten, Reduced C9orf72 function leads to defective synaptic vesicle release and neuromuscular dysfunction in zebrafish. *Commun. Biol.* **4**, 792 (2021).
56. A. Keller, A. I. Nesvizhskii, E. Kolker, R. Aebersold, Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383–5392 (2002).
57. A. I. Nesvizhskii, A. Keller, E. Kolker, R. Aebersold, A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646–4658 (2003).
58. D. Kim, B. Langmead, S. L. Salzberg, HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
59. C. L. Seiler, J. Fernandez, Z. Koerperich, M. P. Andersen, D. Kotandeniya, M. E. Nguyen, Y. Y. Sham, N. Y. Tretyakova, Maintenance DNA methyltransferase activity in the presence of oxidized forms of 5-methylcytosine: Structural basis for ten eleven translocation-mediated DNA demethylation. *Biochemistry* **57**, 6061–6069 (2018).
60. H. E. Xu, M. A. Rould, W. Xu, J. A. Epstein, R. L. Maas, C. O. Pabo, Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. *Genes Dev.* **13**, 1263–1275 (1999).
61. M. Ravichandran, D. Rafalski, C. I. Davies, O. Ortega-Recalde, X. Nan, C. R. Glanfield, A. Kotter, K. Misztal, A. H. Wang, M. Wojciechowski, M. Rażew, I. M. Mayyas, O. Kardailsky, U. Schwartz, K. Zembrzycki, I. M. Morison, M. Helm, D. Weichenhan, R. Z. Jurkowska, F. Krueger, C. Plass, M. Zacharias, M. Bochtler, T. A. Hore, T. P. Jurkowski, Pronounced sequence specificity of the TET enzyme catalytic domain guides its cellular function. *Sci. Adv.* **8**, eabm2427 (2022).

Acknowledgments: We are grateful to J. P. Martinez Barbera (London) for mDNMT1 plasmids, to O. Neyret and S. Boissel for next-generation sequencing, and to V. Magoon for expert secretarial assistance. **Funding:** This work was supported by grants to J.D. from the Canadian Institutes of Health Research (FDN-154297) and the Digital Research Alliance of Canada (zmv-553). **Author contributions:** Conceptualization: J.H., A.M., A.J., A.B., and J.D. Data curation: J.D. Formal analysis: J.H., A.G., N.H.S., and S.M. Funding acquisition: J.D. Investigation: J.H., A.G., Y.G., S.M., M.D., P.B., S.-H.C., and J.D. Methodology: J.H., A.M., Y.G., N.H.S., S.M., A.J., S.-H.C., and J.D. Project administration: A.J. and J.D. Software: J.H. and H.D. Resources: A.J., H.D., S.-H.C., and J.D. Supervision: A.M., A.J., and J.D. Validation: J.H., A.M., S.M., and J.D. Visualization: J.H., N.H.S., and J.D. Writing—original draft: J.H., H.D., S.-H.C., and J.D. Writing—review and editing: J.H., A.M., S.M., A.J., and J.D. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Sequencing data are deposited with Gene Expression Omnibus (GEO) under accession numbers GSE87185, GSE212390, GSE125668, GSE280757, GSE280759, GSE280760, GSE280762, and GSE28063.

Submitted 15 November 2024
Accepted 11 April 2025
Published 16 May 2025
10.1126/sciadv.adu6632