Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive demethylation in mammals

Authors:

Christopher B. Mulholland[1], Atsuya Nishiyama[2†], Joel Ryan[1†], Ryohei Nakamura[3], Merve Yiğit[1], Ivo M. Glück[4], Carina Trummer[1], Weihua Qin[1], Michael D. Bartoschek[1], Franziska R. Traube[5], Edris Parsa[5], Enes Ugur[1,8], Miha Modic[7], Aishwarya Acharya[1], Paul Stolz[1], Christoph Ziegenhain[6], Michael Wierer[8], Wolfgang Enard[6], Thomas Carell[5], Don C. Lamb[4], Hiroyuki Takeda[3], Makoto Nakanashi[2], Sebastian Bultmann[*1§] and Heinrich Leonhardt[*1§]

Affiliations:

[1] Department of Biology II and Center for Integrated Protein Science Munich (CIPSM), Human Biology and BioImaging, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany

[2] Division of Cancer Cell Biology, The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan

[3] Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

[4] Physical Chemistry, Department of Chemistry, Center for Nanoscience, Nanosystems Initiative Munich and Center for Integrated Protein Science Munich, Ludwig-Maximilians-Universität München, Munich, Germany

[5] Center for Integrated Protein Science (CIPSM) at the Department of Chemistry, Ludwig-Maximilians-Universität München, Munich, Germany

[6] Department of Biology II, Anthropology and Human Genomics, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany

[7] The Francis Crick Institute and University College London, Institute of Neurology, London, United Kingdom

[8] Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Am Klopferspitz 18, 82152, Martinsried, Germany.

[†]These authors contributed equally

§These authors jointly supervised this work

*Correspondence and requests for material should be addressed to S.B. (bultmann@bio.lmu.de) or H.L. (h.leonhardt@lmu.de)

Abstract:

Genome-wide DNA demethylation is a unique feature of mammalian development and naïve pluripotent stem cells. So far, it was unclear how mammals specifically achieve global DNA hypomethylation, given the high conservation of the DNA (de-)methylation machinery among vertebrates. We found that DNA demethylation requires TET activity but mostly occurs at sites where TET proteins are not bound suggesting a rather indirect mechanism. Among the few specific genes bound and activated by TET proteins was the naïve pluripotency and germline marker *Dppa3* (*Pgc7*, *Stella*), which undergoes TDG dependent demethylation. The requirement of TET proteins for genome-wide DNA demethylation could be bypassed by ectopic expression of *Dppa3*. We show that DPPA3 binds and displaces UHRF1 from chromatin and thereby prevents the recruitment and activation of the maintenance DNA methyltransferase DNMT1. We demonstrate that DPPA3 alone can drive global DNA demethylation when transferred to amphibians (*Xenopus*) and fish (medaka), both species that naturally do not have a *Dppa3* gene and exhibit no post-fertilization DNA demethylation. Our results show that TET proteins are responsible for active and - indirectly also for - passive DNA demethylation; while TET proteins initiate local and gene-specific demethylation in vertebrates, the recent emergence of DPPA3 introduced a unique means of genome-wide passive demethylation in mammals and contributed to the evolution of epigenetic regulation during early mammalian development.

INTRODUCTION

During early embryonic development the epigenome undergoes massive changes. Upon fertilization the genomes of highly specialized cell types — sperm and oocyte — need to be reprogrammed in order to obtain totipotency. This process entails decompaction of the highly condensed gametic genomes and global resetting of chromatin states to confer the necessary epigenetic plasticity required for the development of a new organism[1]. At the same time the genome needs to be protected from the activation of transposable elements (TEs) abundantly present in vertebrate genomes[2]. Activation and subsequent transposition of TEs results in mutations that can have deleterious effects and are passed onto offspring if they occur in the germline during early development[2,3]. The defense against these genomic parasites has shaped genomes substantially[4,5].

DNA methylation in vertebrates refers to the addition of a methyl group at the C5 position of cytosine to form 5-methylcytosine (5mC). Besides its important role in gene regulation, the most basic function of DNA methylation is the repression and stabilization of TEs and other repetitive sequences[6]. Accordingly, the majority of genomic 5mC is located within these highly abundant repetitive elements. Global DNA methylation loss triggers the derepression of transposable and repetitive elements, which leads to genomic instability and cell death, highlighting the crucial function of vertebrate DNA methylation[7–12]. Hence, to ensure constant protection against TE reactivation, global DNA methylation levels remain constant throughout the lifetime of non-mammalian vertebrates[13–16]. Paradoxically, mammals specifically erase DNA methylation during preimplantation development[17,18], a process that would seemingly expose the developing organism to the risk of genomic instability through the activation of TEs. DNA methylation also acts as an epigenetic barrier to restrict and stabilize cell fate decisions and thus constitutes a form of epigenetic memory. The establishment of pluripotency in mammals requires the erasure of epigenetic memory and as such, global hypomethylation is a defining characteristic of pluripotent cell types including naïve embryonic stem cells (ESCs), primordial germ cells (PGCs), and induced pluripotent stem cells (iPSCs)[19].

In vertebrates, DNA methylation can be reversed to unmodified cytosine by two mechanisms; either actively by Ten-eleven translocation (TET) dioxygenase-mediated oxidation of 5mC in concert with the base excision repair machinery[20–23] or passively by a lack of functional DNA methylation maintenance during the DNA replication cycle[24,25]. In mammals, both active and passive demethylation pathways have been suggested to be involved in the establishment of global demethylation during preimplantation development[26–32]. Intriguingly however, whereas hypomethylated states during development are mammal-specific, TET-proteins as well as the DNA methylation machinery are highly conserved among all vertebrates[33,34]. This discrepancy implies the existence of so far unknown mammalian-specific pathways and factors controlling the establishment and maintenance of genomic hypomethylation.

Here, we use mouse embryonic stem cells (ESCs) cultured in conditions promoting naïve pluripotency[35–37] as a model to study global DNA demethylation in mammals. By dissecting the contribution of the catalytic activity of TET1 and TET2 we show that TET-mediated active demethylation drives the expression of the Developmental pluripotency-associated protein 3 (DPPA3/PGC7/STELLA). We show that DPPA3 directly binds UHRF1

prompting its release from chromatin, resulting in the inhibition of maintenance methylation and global passive demethylation. Although only found in mammals, DPPA3 can also induce global demethylation in non-mammalian vertebrates. In summary, we described a novel TET-controlled and DPPA3-driven pathway for passive demethylation in naïve pluripotency in mammals.

RESULTS

TET1 and TET2 indirectly protect the naïve genome from hypermethylation

Mammalian TET proteins, TET1, TET2, and TET3, share a conserved catalytic domain and the ability to oxidize 5mC but exhibit distinct expression profiles during development[38]. Naïve ESCs and the inner cell mass (ICM) of the blastocyst from which they are derived feature high expression of *Tet1* and *Tet2* but not *Tet3*[27,39–41]. To dissect the precise contribution of TET-mediated active DNA demethylation to global DNA hypomethylation in naïve pluripotency we generated isogenic *Tet1* (T1CM) and *Tet2* (T2CM) single as well as *Tet1*/*Tet2* (T12CM) double catalytic mutant mouse ESC lines using CRISPR/Cas-assisted gene editing (Supplementary Fig. 1). We confirmed the inactivation of TET1 and TET2 activity by measuring the levels of 5-hydroxymethylcytosine (5hmC), the product of TET-mediated oxidation of 5mC[20] (Supplementary Fig. 1i). While the loss of either *Tet1* or *Tet2* catalytic activity significantly reduced 5hmC levels, inactivation of both TET1 and TET2 resulted in the near total loss of 5hmC in naïve ESCs (Supplementary Fig. 1i) indicating that TET1 and TET2 account for the overwhelming majority of cytosine oxidation in naïve ESCs. We then used reduced representation bisulfite sequencing (RRBS) to determine the DNA methylation state of T1CM, T2CM, and T12CM ESCs as well as wild-type (wt) ESCs. All *Tet* catalytic mutant (T1CM, T2CM and T1CM) cell lines exhibited severe DNA hypermethylation throughout the genome including promoters, gene bodies, and repetitive elements (Fig. 1a,b). The increase in DNA methylation was particularly pronounced at LINE-1 (L1) elements of which 97%, 98%, and 99% were significantly hypermethylated in T1CM, T2CM, and T12CM ESCs, respectively (Supplementary Fig. 2a). This widespread DNA hypermethylation was reminiscent of the global increase in DNA methylation accompanying the transition of naïve ESCs to primed epiblast-like cells (EpiLCs)[40,42,43], which prompted us to investigate whether the hypermethylation present in T1CM, T2CM, and T12CM ESCs represents a premature acquisition of a more differentiated DNA methylation signature. In line with this hypothesis, *Tet* catalytic mutant ESCs displayed DNA methylation levels similar or higher than those of wt EpiLCs (Supplementary Fig. 2b). Moreover, hierarchical clustering and principal component analyses (PCA) of the RRBS data revealed that ESCs from *Tet* catalytic mutants clustered closer to wt EpiLCs than wt ESCs (Fig. 1c and Supplementary Fig. 2c). In fact, the vast majority of significantly hypermethylated CpGs in *Tet* catalytic mutant ESCs overlapped with those normally gaining DNA methylation during the exit from naïve pluripotency (Fig. 1d). However, T1CM, T2CM, and T12CM transcriptomes clearly clustered by differentiation stage indicating that the acquisition of an EpiLC-like methylome was not due to premature differentiation (Supplementary Fig. 2d). Intriguingly, we found that the majority of sites hypermethylated in *Tet* catalytic mutant ESCs are not bound by TET1 or TET2 (Fig. 1e,f) suggesting that the catalytic activity of TET1 and TET2 maintains the hypomethylated state of the naïve methylome by indirect means.

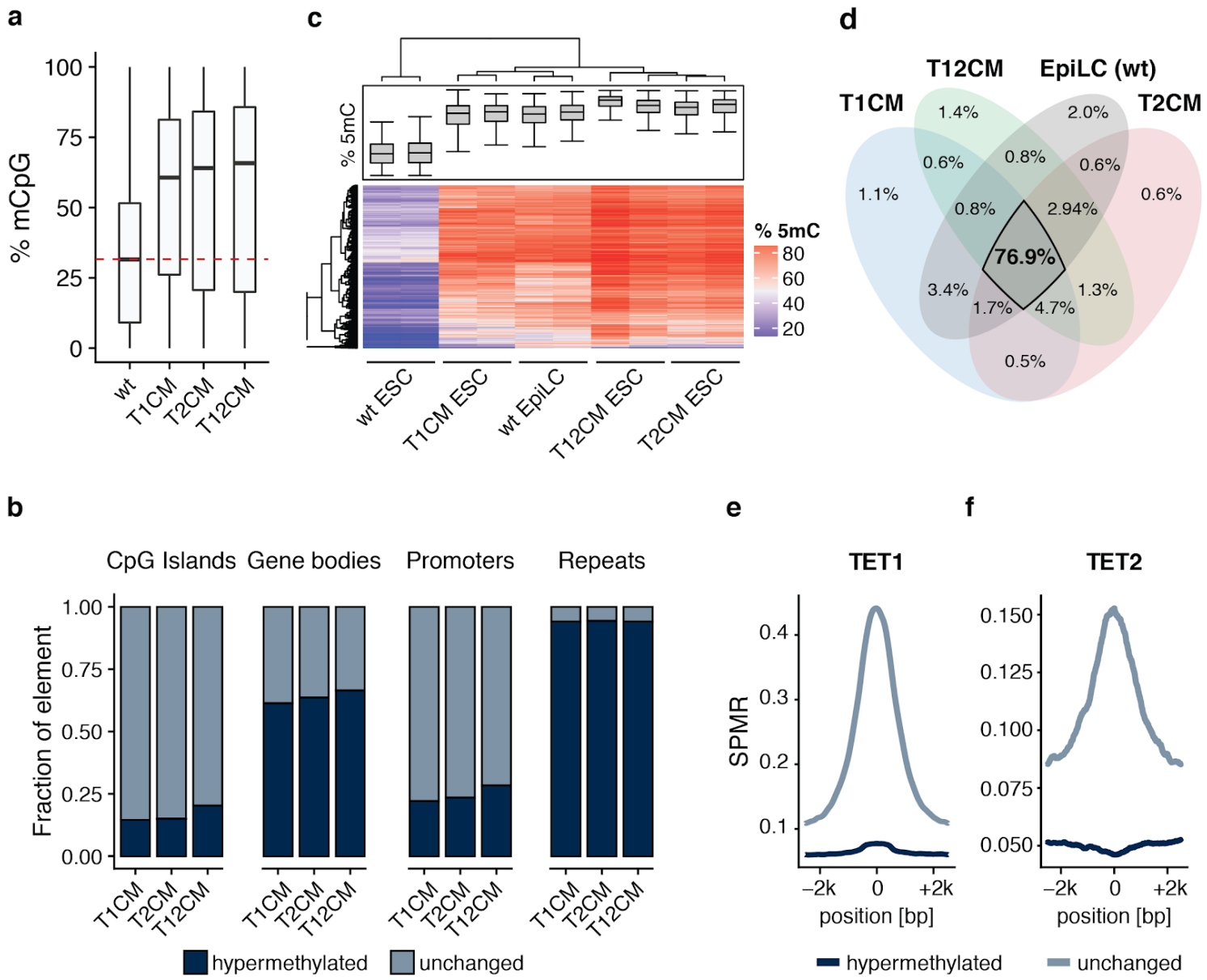


Fig. 1: TET1 and TET2 prevent hypermethylation of the naïve genome

a, Loss of TET catalytic activity leads to global DNA hypermethylation. Percentage of total 5mC as measured by RRBS. **b,** Loss of TET catalytic activity leads to widespread DNA hypermethylation especially at repetitive elements. Relative proportion of DNA hypermethylation (q value < 0.05; absolute methylation difference > 20%) at each genomic element in T1CM, T2CM, and T12CM ESCs compared to wt ESCs. **c,** Heat map of the hierarchical clustering of the RRBS data depicting the top 2000 most variable 1kb tiles during differentiation of wt ESCs to EpiLCs. **d,** Venn diagram depicting the overlap of hypermethylated (compared to wt ESCs; q value < 0.05; absolute methylation difference > 20%) sites among T1CM, T2CM, and T12CM ESCs and wt EpiLCs. **e,f,** TET binding is not associated with DNA hypermethylation in TET mutant ESCs. Occupancy of (**e**) TET1[44] and (**f**) TET2[45] over 1 kb tiles hypermethylated (dark blue) or unchanged (light blue) in T1CM and T2CM ESCs, respectively. In the boxplots in **a-c,** horizontal black lines within boxes represent median values, boxes indicate the upper and lower quartiles, and whiskers indicate the 1.5 interquartile range.

128 **TET1 and TET2 control *Dppa3* expression in a catalytically dependent manner**

129 To understand how TET1 and TET2 indirectly promote demethylation of the naïve genome, we examined the
130 expression of the enzymes involved in DNA methylation. Loss of TET catalytic activity was not associated with
131 changes in the expression of *Dnmt1*, *Uhrf1*, *Dnmt3a*, and *Dnmt3b*, indicating the hypermethylation in *Tet*
132 catalytic mutant ESCs is not a result of the DNA methylation machinery being upregulated (Fig. 2a). To identify
133 candidate factors involved in promoting global hypomethylation, we compared the transcriptome of
134 hypomethylated wild-type ESCs with those of hypermethylated cells, which included wt EpiLCs as well as
135 T1CM, T2CM, and T12CM ESCs (Fig. 2b). Among the 14 genes differentially expressed in hypermethylated cell
136 lines, the naïve pluripotency factor, *Dppa3* (also known as *Stella* and *Pgc7*), was an interesting candidate due
137 to its reported involvement in the regulation of global DNA methylation in germ cell development and oocyte
138 maturation[46–49]. In addition, *Dppa3* is also a direct target of PRDM14, a PR domain-containing transcriptional
139 regulator known to promote the DNA hypomethylation associated with naïve pluripotency[36,50–52] (Fig. 2e). While
140 normally highly expressed in naïve ESCs and only downregulated upon differentiation[53,54], *Dppa3* was
141 prematurely repressed in T1CM, T2CM, and T12CM ESCs (Fig. 2d). The significantly reduced expression of
142 *Dppa3* in TET mutant ESCs was accompanied by significant hypermethylation of the *Dppa3* promoter (Fig. 2e),
143 consistent with reports demonstrating *Dppa3* to be one of the few pluripotency factors regulated by promoter
144 methylation *in vitro* and *in vivo*[37,53–55]. In contrast to the majority of genomic sites gaining methylation in TET
145 mutant ESCs (Fig. 1e,f), hypermethylation at the *Dppa3* locus occurred at sites bound by both TET1 and TET2
146 (Fig. 2e)[44,45]. This hypermethylation overlapped with regions at which the TET oxidation product
147 5-carboxylcytosine (5caC) accumulates in Thymine DNA glycosylase (TDG)-knockdown ESCs (Fig. 2e)[56],
148 indicating that the Dppa3 locus is a direct target of TET/TDG-mediated active DNA demethylation in ESCs.

149 PRDM14 has been shown to recruit TET1 and TET2 to sites of active demethylation and establish global
150 hypomethylation in naïve pluripotency[36,40,51,52,57]. As the expression of *Prdm14* was not altered in *Tet* catalytic
151 mutant ESCs (Fig. 2a), we analysed PRDM14 occupancy at the *Dppa3* locus using publicly available ChIP-seq
152 data[51]. This analysis revealed that PRDM14 binds the same upstream region of *Dppa3* occupied by TET1 and
153 TET2 (Fig. 2e). Taken together, these data suggest that TET1 and TET2 are recruited by PRDM14 to maintain
154 the expression of *Dppa3* by active DNA demethylation. Strikingly, systematic analysis of public databases
155 revealed that while the DNA (de)methylation machinery (DNMTs, UHRF1, TETs) is conserved among
156 vertebrates, *Dppa3* is only present in mammals potentially representing a novel pathway that regulates
157 mammalian-specific global hypomethylation in naïve pluripotency (Fig. 2c).

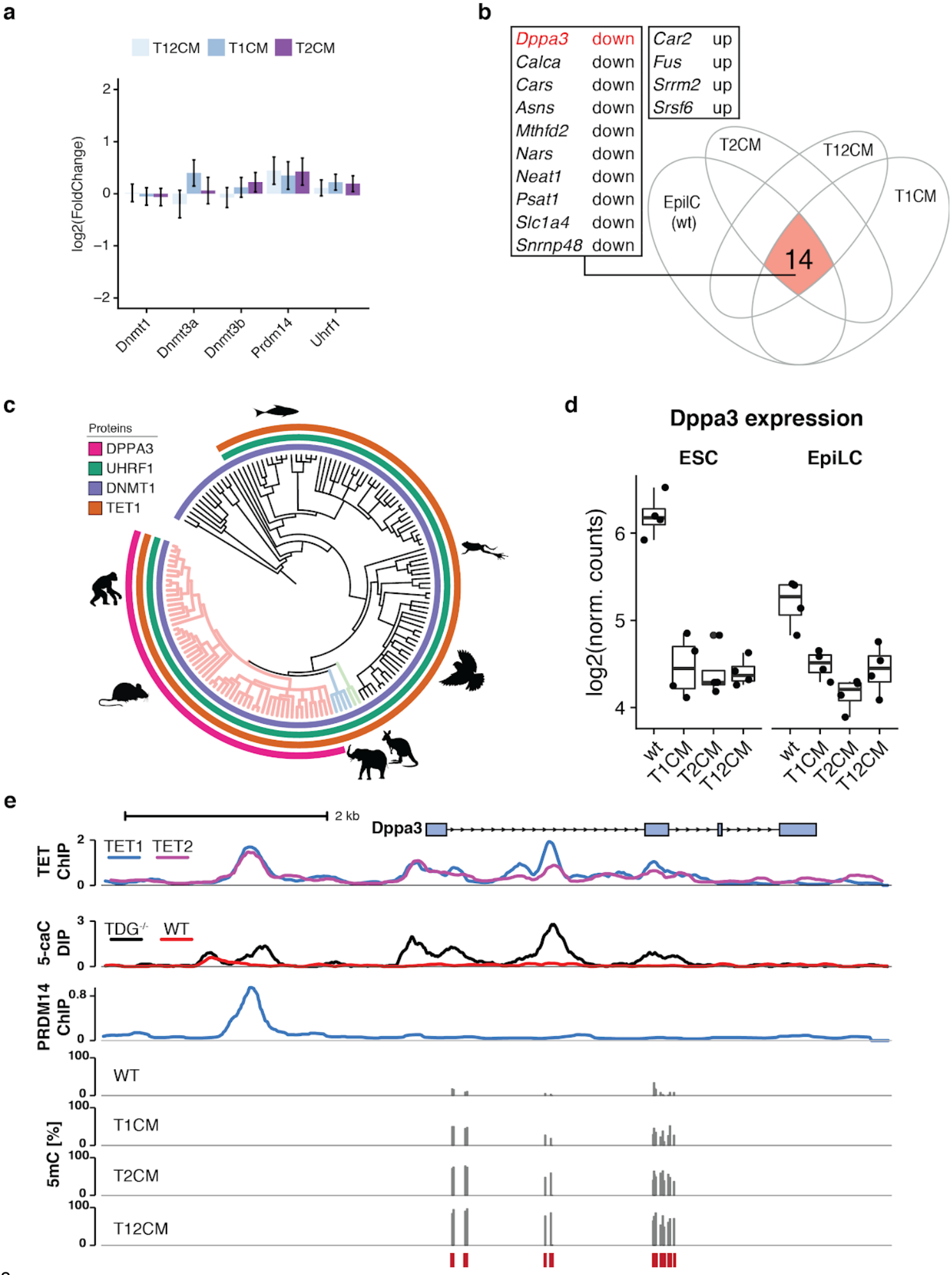

a
T12CM
T1CM
T2CM
log2(FoldChange)
2
1
0
−1
−2
Dnmt1
Dnmt3a
Dnmt3b
Prdm14
Uhrf1
b
Dppa3 down
Calca down
Cars down
Asns down
Mthfd2 down
Nars down
Neat1 down
Psat1 down
Slc1a4 down
Snrnp48 down
Car2 up
Fus up
Srrm2 up
Srsf6 up
T2CM
T12CM
T1CM
EpiLC
(wt)
14
c
Proteins
DPPA3
UHRF1
DNMT1
TET1
d
Dppa3 expression
ESC
EpiLC
log2(norm. counts)
6
5
4
wt
T1CM
T2CM
T12CM
wt
T1CM
T2CM
T12CM
e
2 kb
Dppa3
TET ChIP
TET1
TET2
5-caC DIP
TDG-/-
WT
PRDM14 ChIP
5mC [%]
WT
T1CM
T2CM
T12CM

Fig. 2: TET1 and TET2 catalytic activity is necessary for *Dppa3* expression

a, Expression of genes involved in regulating DNA methylation levels in T1CM, T2CM, and T12CM ESCs as assessed by RNA-seq. Expression is given as the $\log_2$ fold-change compared to wt ESCs. Error bars indicate mean ± SD, n = 4 biological replicates. No significant changes observable (Likelihood ratio test). **b,** *Dppa3* is downregulated upon loss of TET activity and during differentiation. Venn diagram depicting the overlap (red) of genes differentially expressed (compared to wt ESCs; adjusted $p < 0.05$) in T1CM, T2CM, T12CM ESCs, and wt EpiLCs. **c,** Phylogenetic tree of TET1, DNMT1, UHRF1, and DPPA3 in metazoa. **d,** *Dppa3* expression levels as determined by RNA-seq in the indicated ESC and EpiLC lines (n = 4 biological replicates). **e,** TET proteins bind and actively demethylate the *Dppa3* locus. Genome browser view of the *Dppa3* locus with tracks of the occupancy (Signal pileup per million reads; (SPMR)) of TET1[44], TET2[45], and PRDM14[51] in wt ESCs, 5caC enrichment in wt vs. TDG$^{-/-}$ ESCs[56], and 5mC (%) levels in wt, T1CM, T2CM, and T12CM ESCs (RRBS). Red bars indicate CpGs covered by RRBS. In **d,** boxplots horizontal black lines within boxes represent median values, boxes indicate the upper and lower quartiles, and whiskers indicate the 1.5 interquartile range. P-values were calculated using Welch's two-sided t-test comparing *Tet* catalytic mutants to their corresponding wt: ** $P < 0.01$; *** $P < 0.001$.

DPPA3 acts downstream of TET1 and TET2 and is required to safeguard the naïve methylome

DPPA3 has been reported to both prevent or promote DNA demethylation depending on the developmental time points[46,48,49,58–62]. However, the function of DPPA3 in the context of naïve pluripotency, for which it is a well-established marker gene[53], has yet to be explored. We first sought to characterize the relationship between the high expression of *Dppa3* and DNA hypomethylation both accompanying naïve pluripotency. To this end, we established isogenic *Dppa3* knockout (Dppa3KO) mouse ESCs using CRISPR/Cas (Supplementary Fig. 3a-c) and profiled their methylome by RRBS. Loss of DPPA3 led to severe global hypermethylation (Fig. 3a) with substantial increases in DNA methylation observed across all genomic elements (Supplementary Fig. 3d). Repetitive sequences and TEs, in particular, were severely hypermethylated including 98% of L1 elements (Supplementary Fig. 3d). A principal component analysis of the RRBS data revealed that Dppa3KO ESCs clustered closer to wt EpiLCs and *Tet* catalytic mutant ESCs rather than wt ESCs (Fig. 3b). Furthermore, we observed a striking overlap of hypermethylated CpGs between *Tet* catalytic mutant and Dppa3KO ESCs (Fig. 3c), suggesting that DPPA3 and TETs promote demethylation at largely the same targets. A closer examination of the genomic distribution of overlapping hypermethylation in *Tet* catalytic mutant and Dppa3KO ESCs revealed that the majority (~90%) of common targets reside within repetitive elements (Fig. 3d, Supplementary Fig. 3e,f) and are globally correlated with heterochromatic histone modifications (Supplementary Fig. 3h). In contrast, only half of the observed promoter hypermethylation among all cell lines was dependent on DPPA3 (classified as "common", Fig. 3d and Supplementary Fig. 3e-g). This allowed us to identify a set of strictly TET-dependent promoters (N = 1573) (Fig. 3d, Supplementary Fig. 3f and Supplementary Table 1), which were enriched for developmental genes (Fig 3e and Supplementary Table 2). Intriguingly, these TET-specific promoters contained genes (such as *Pax6*, *Foxa1* and *Otx2*) that have recently been shown to be conserved targets of TET-mediated demethylation during *Xenopus*, zebrafish and mouse development[63].

DPPA3 appeared to act downstream of TETs as the global increase in DNA methylation in Dppa3KO ESCs was

not associated with a reduction in 5hmC levels nor with a downregulation of TET family members (Fig. 3f and Supplementary Fig. 3i). In support of this notion, inducible overexpression of *Dppa3* (Supplementary Fig. 3j-l) completely rescued the observed hypermethylation phenotype at LINE-1 elements in T1CM as well as T2CM ESCs and resulted in a significant reduction of hypermethylation in T12CM cells (Fig. 3g). Strikingly, prolonged induction of *Dppa3* even resulted in hypomethylation in wild-type as well as T1CM ESCs (Fig. 3g). Collectively, these results show that TET1 and TET2 activity contributes to genomic hypomethylation in naïve pluripotency by direct and indirect pathways, the active demethylation of developmental promoters and the passive, DPPA3-mediated global demethylation.

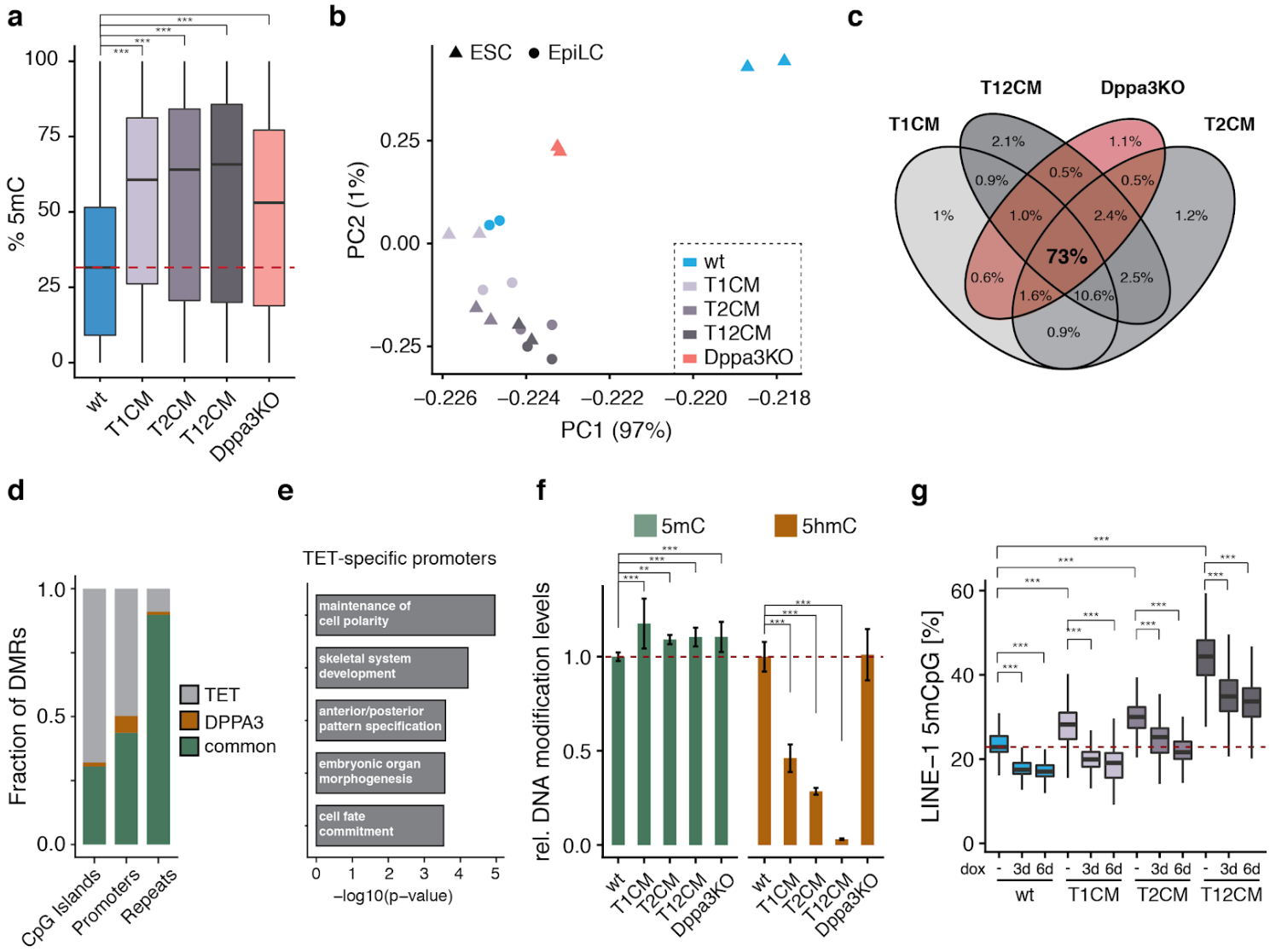


Fig. 3: DPPA3 acts downstream of TET1 and TET2 to establish and preserve global hypomethylation

a, *Dppa3* loss results in global hypermethylation. Percentage of total 5mC as measured by RRBS. **b,** *Dppa3* prevents the premature acquisition of a primed methylome. Principal component (PC) analysis of RRBS data from wt, T1CM, T2CM and T12CM ESCs, wt EpiLCs and Dppa3KO ESCs. **c,** DPPA3 and TET proteins promote demethylation of largely similar targets. Venn Diagrams depicting the overlap of hypermethylated sites among T1CM, T2CM, T12CM, and Dppa3KO ESCs. **d,** *Dppa3* protects mostly repeats from hypermethylation. Fraction of hypermethylated genomic elements classified as TET-specific (only hypermethylated in TET mutant ESCs), DPPA3-specific (only hypermethylated in Dppa3KO ESCs), or common (hypermethylated in TET mutant and Dppa3KO ESCs). **e,** Gene ontology (GO) terms associated with promoters specifically dependent on TET activity; adjusted p-values calculated using

Fisher's exact test followed by Benjamini-Hochberg correction for multiple testing. **f,** TET activity remains unaffected in Dppa3KO ESCs. Relative DNA modification levels for 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) as measured by mass spectrometry (LC-MS/MS). Error bars indicate mean ± SD calculated from $n > 3$ biological replicates. **g,** *Dppa3* expression can rescue the hypermethylation in TET mutant ESCs. DNA methylation levels at LINE-1 elements (%) as measured by bisulfite sequencing 0, 3, or 6 days after doxycycline (dox) induction of *Dppa3* expression. In the boxplots in **a, g,** horizontal black lines within boxes represent median values, boxes indicate the upper and lower quartiles, and whiskers indicate the 1.5 interquartile range. The dashed red line indicates the median methylation level of wt ESCs. In **a, f,** and **g,** P-values were calculated using Welch's two-sided t-test: ** $P < 0.01$; *** $P < 0.001$.

TET-dependent expression of DPPA3 regulates UHRF1 subcellular distribution and controls DNA methylation maintenance in embryonic stem cells

To investigate the mechanism underlying the regulation of global DNA methylation patterns by DPPA3 we first generated an endogenous DPPA3-HALO fusion ESC line to monitor the localization of DPPA3 throughout the cell cycle (Supplementary Fig. 4a,c). Recent studies have shown that DPPA3 binds H3K9me2[61] and that in oocytes its nuclear localization is critical to inhibit the activity of UHRF1[49], a key factor for maintaining methylation. Expecting a related mechanism to be present in ESCs, we were surprised to find a strong cytoplasmic localization of DPPA3 in ESCs (Fig. 4a). Furthermore, DPPA3 did not bind to mitotic chromosomes indicating a low or absent chromatin association of DPPA3 in ESCs (Fig. 4a). To further understand the mechanistic basis of DPPA3-dependent DNA demethylation in ESCs, we performed FLAG-DPPA3 pulldowns followed by liquid chromatography tandem mass spectrometry (LC MS-MS) to profile the DPPA3 interactome in naïve ESCs. Strikingly, among the 303 significantly enriched DPPA3 interaction partners identified by mass spectrometry, we found UHRF1 and DNMT1 (Fig. 4b and Supplementary Table 3), the core components of the DNA maintenance methylation machinery[64,65]. A reciprocal immunoprecipitation of UHRF1 confirmed its interaction with DPPA3 in ESCs (Supplementary Fig. 4f). Furthermore, GO analysis using the top 131 interactors of DPPA3 showed the two most enriched GO terms to be related to DNA methylation (Supplementary Table 4). These findings were consistent with previous studies implicating DPPA3 in the regulation of maintenance methylation[47,49]. In addition, we also detected multiple members of the nuclear transport machinery indicating that DPPA3 might undergo nuclear shuttling in ESCs (highlighted in purple, Fig. 4b and Supplementary Table 3), which prompted us to investigate whether DPPA3 influences the subcellular localization of UHRF1. Surprisingly, biochemical fractionation experiments revealed UHRF1 to be present in both the nucleus and cytoplasm of naïve wt ESCs (Supplementary Fig. 4e). Despite comparable total UHRF1 protein levels in wt and Dppa3KO ESCs (Supplementary Fig. 4g), loss of DPPA3 completely abolished the cytoplasmic fraction of UHRF1 (Supplementary Fig. 4e).

As maintenance DNA methylation critically depends on the correct targeting and localization of UHRF1 within the nucleus[66–69], we asked whether TET-dependent regulation of DPPA3 might affect the subnuclear distribution of UHRF1. To this end, we tagged endogenous UHRF1 with GFP in wild-type (U1G/wt) as well as Dppa3KO and T12CM ESCs (U1G/Dppa3KO and U1G/T12CM, respectively) enabling us to monitor UHRF1 localization dynamics in living cells (Supplementary Fig. 4b-d). Whereas UHRF1-GFP localized to both the nucleus and

223 cytoplasm of wt ESCs, UHRF1-GFP localization was solely nuclear in Dppa3KO and T12CM ESCs
224 (Supplementary Fig. 4h,i). In addition, UHRF1 appeared to display a more diffuse localization in wt ESCs
225 compared to Dppa3KO and T12CM ESCs, in which we observed more focal patterning of UHRF1 particularly
226 at heterochromatic foci (Supplementary Fig. 4h). To quantify this observation, we calculated the coefficient of
227 variation (CV) of nuclear UHRF1-GFP among wt, Dppa3KO, and T12CM ESCs. The CV of a fluorescent signal
228 correlates with its distribution, where low CV values reflect more homogenous distributions and high CV values
229 correspond to more heterogeneous distributions[70,71]. Indeed, the pronounced focal accumulation of
230 UHRF1-GFP observed in Dppa3KO and T12CM ESCs corresponded with a highly significant increase in the CV
231 values of nuclear UHRF1-GFP compared with wt ESCs (Supplementary Fig. 4h,i).

232 To assess whether these differences in nuclear UHRF1 distribution reflected altered chromatin binding, we
233 used fluorescence recovery after photobleaching (FRAP) to study the dynamics of nuclear UHRF1-GFP in wt,
234 Dppa3KO, and T12CM ESCs. Our FRAP analysis revealed markedly increased UHRF1 chromatin binding in
235 both Dppa3KO and T12CM ESCs as evidenced by the significantly slower recovery of UHRF1-GFP in these
236 cell lines compared to wt ESCs (Fig. 4c and Supplementary Fig. 4j,k). Additionally, these data demonstrated
237 increased UHRF1 chromatin binding to underlie the more heterogenous nuclear UHRF1 distributions in
238 Dppa3KO and T12CM ESCs. Interestingly, although strongly reduced compared to wt ESCs, UHRF1 mobility
239 was slightly higher in T12CM ESCs than Dppa3KO ESCs, consistent with a severe but not total loss of DPPA3
240 in the absence of TET activity (Supplementary Fig. 4l). Induction of *Dppa3* rescued the cytoplasmic fraction of
241 UHRF1 (N/C Ratio: Fig. 4d) as well as the diffuse localization of nuclear UHRF1 in Dppa3KO ESCs (CV: Fig.
242 4d), which reflected a striking increase in the mobility of residual nuclear UHRF1-GFP as assessed by FRAP
243 (Supplementary Figs. 4m and 5a,b). This analysis also revealed that UHRF1's nucleocytoplasmic translocation
244 and the inhibition of chromatin binding followed almost identical kinetics (N/C $t_{1/2}$ = 84.4 min; CV $t_{1/2}$ = 82.8)
245 (Fig. 4d). UHRF1 is required for the proper targeting of DNMT1 to DNA replication sites and therefore essential
246 for DNA methylation maintenance[64,72]. We observed a marked reduction of both UHRF1 and DNMT1 at
247 replication foci upon induction of *Dppa3*, indicating that DPPA3 promotes hypomethylation in naïve ESCs by
248 impairing DNA methylation maintenance (Supplementary Fig. 5c,d). Ectopic expression of DPPA3 not only
249 altered the subcellular distribution of endogenous UHRF1 in mouse ESCs (Fig. 4d and Supplementary Fig. 5e)
250 but also in human ESCs suggesting evolutionary conservation of this mechanism among mammals
251 (Supplementary Fig. 5f,g). Collectively our results demonstrate that TET-proteins control both the subcellular
252 localization and chromatin binding of UHRF1 via the regulation of DPPA3 levels in naïve ESCs. Furthermore,
253 these data show that DPPA3 is both necessary and sufficient for ensuring the nucleocytoplasmic translocation,
254 diffuse nuclear localization, and attenuated chromatin binding of UHRF1 in ESCs.

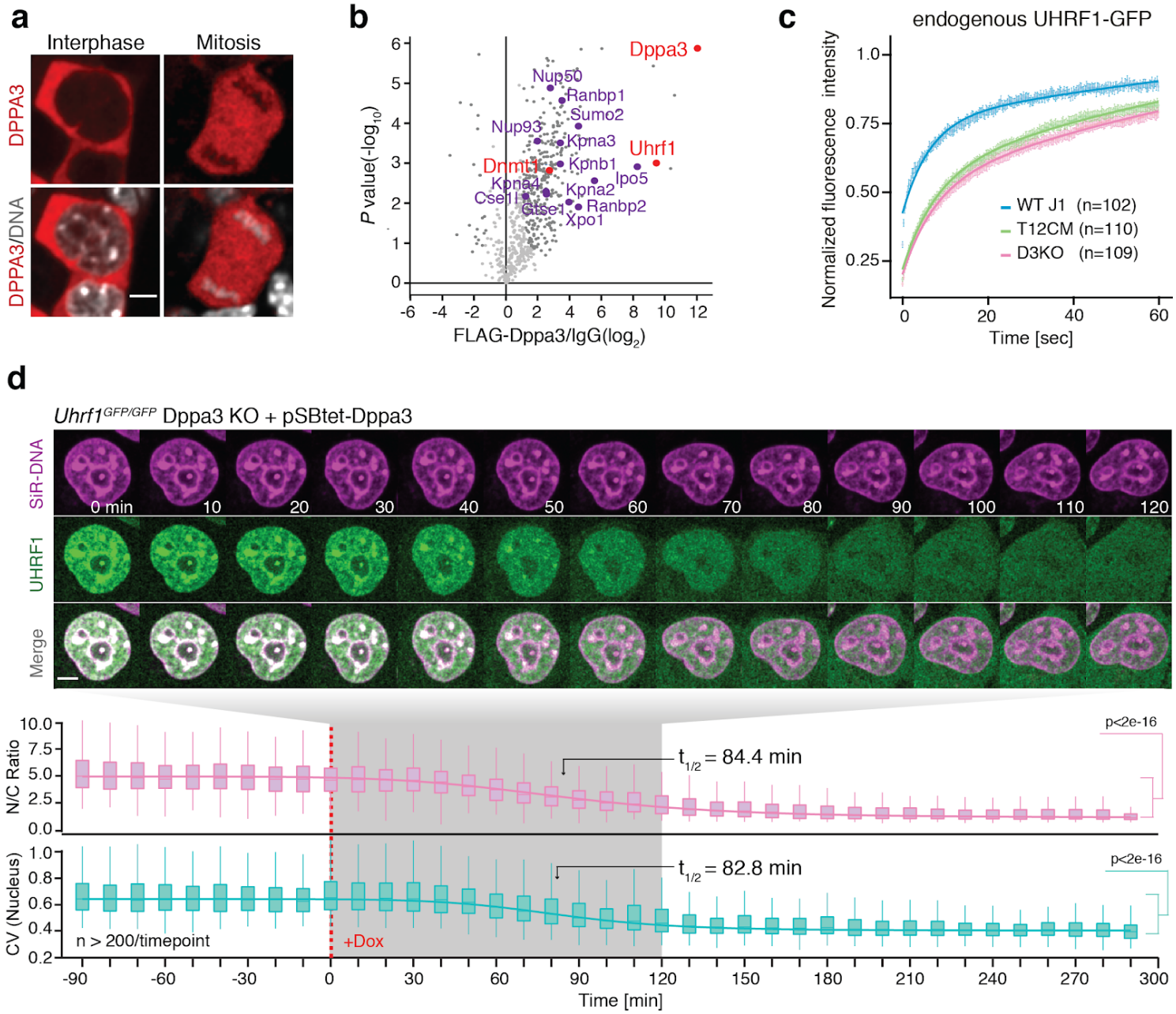


Fig. 4: DPPA3 prevents UHRF1 chromatin binding to impede maintenance methylation in embryonic stem cells

a, DPPA3 is primarily localized to the cytoplasm of ESCs. Images illustrating the localization of endogenous DPPA3-HALO in live ESCs counterstained with SiR-Hoechst (DNA). Scale bar: 5 µm. **b,** DPPA3 interacts with endogenous UHRF1 in ESCs. Volcano plot from DPPA3-FLAG pulldowns in ESCs. Dark grey dots: significantly enriched proteins (FDR < 0.05). Red dots: proteins involved in DNA methylation regulation. Purple dots: proteins involved in nuclear transport. anti-FLAG antibody: n = 3 biological replicates, IgG control antibody: n = 3 biological replicates. Statistical significance determined by performing a Student's t-test with a permutation-based FDR of 0.05 and an additional constant S0 = 1. **c,** Loss of DPPA3 leads to increased UHRF1 chromatin binding. FRAP analysis of endogenous UHRF1-GFP. Each genotype comprises the combined single-cell data from two independent clones acquired in two independent experiments. **d,** Normal UHRF1 localization can be rescued by ectopic *Dppa3* expression. Localization dynamics of endogenous UHRF1-GFP in response to *Dppa3* induction in U1G/D3KO + pSBtet-D3 ESCs with confocal time-lapse imaging over 8 h (10 min intervals). *t = 0* corresponds to start of *Dppa3* induction with doxycycline (+Dox). (*top panel*) Representative images of UHRF1-GFP and DNA (SiR-Hoechst stain) throughout confocal time-lapse imaging. Scale bar: 5 µm. (middle panel) Nucleus to cytoplasm ratio (N/C Ratio) of endogenous UHRF1-GFP signal. (bottom panel) Coefficient of variance (CV) of endogenous

UHRF1-GFP intensity in the nucleus. (*middle and bottom panel*) N/C Ratio and CV values: measurements in $n > 200$ single cells per time point, acquired at $n = 16$ separate positions. Curves represent fits of four parameter logistic (4PL) functions to the N/C Ratio (pink line) and CV (green line) data. Live-cell imaging was repeated 3 times with similar results. In **c,** the mean fluorescence intensity of n cells (indicated in the plots) at each timepoint are depicted as shaded dots. Error bars indicate mean ± SEM. Curves (solid lines) indicate double-exponential functions fitted to the FRAP data. In the boxplots in **d,** darker horizontal lines within boxes represent median values. The limits of the boxes indicate upper and lower quartiles, and whiskers indicate the 1.5-fold interquartile range. P-values based on Welch's two-sided t-test.

DPPA3-mediated inhibition of UHRF1 chromatin binding causes hypomethylation and is attenuated by nuclear export

Our results demonstrated cytoplasmic accumulation of UHRF1 and the disruption of its focal nuclear patterning to occur with almost identical kinetics upon induction of *Dppa3* expression (Fig. 4d). In principle, either a decrease in nuclear UHRF1 concentration or the impaired chromatin loading of UHRF1 would on their own be sufficient to impair maintenance DNA methylation[68,73]. To dissect these two modes and their contribution to the inhibition of maintenance methylation in naïve ESCs, we generated inducible *Dppa3*-mScarlet expression cassettes (Supplementary Fig. 6a) harboring mutations to residues described to be critical for its nuclear export (ΔNES)[48] and the interaction with UHRF1 (KRR and R107E)[49], as well as truncated forms of DPPA3 found in zygotes, 1-60 and 61-150[62] (Fig. 5a). After introducing these *Dppa3* expression cassettes into U1GFP/Dppa3KO ESCs, we used live-cell imaging to track each DPPA3 mutant's localization and ability to rescue the loss of DPPA3 (Fig. 5b). Whereas DPPA3-ΔNES and DPPA3 61-150, both lacking a functional nuclear export signal, were sequestered to the nucleus (Fig. 5b), DPPA3-WT, DPPA3-KRR, DPPA3-R107E, and DPPA3 1-60 mutants localized primarily to the cytoplasm (Fig. 5b) and recapitulated the localization of endogenous DPPA3 in naïve ESCs (Fig. 4a). Regardless, all tested DPPA3 mutants failed to efficiently reestablish nucleocytoplasmic translocation of UHRF1 (Fig. 5b and Supplementary Fig. 6b), indicating that DPPA3 requires both the capacity to interact with UHRF1 as well as a functional nuclear export signal to promote nucleocytoplasmic shuttling of UHRF1 in naïve ESCs.

Nevertheless, DPPA3-ΔNES and DPPA3 61-150 managed to significantly disrupt the focal patterning and association with chromocenters of UHRF1 within the nucleus itself, with DPPA3-ΔNES causing an even greater reduction in the CV of nuclear UHRF1 than DPPA3-WT (Fig. 5b and Supplementary Fig. 6c). In contrast, the loss or mutation of residues critical for its interaction with UHRF1 compromised DPPA3's ability to effectively restore the diffuse localization of nuclear UHRF1 (Fig. 5b and Supplementary Fig. 6c). On the one hand, FRAP analysis revealed that the disruption or deletion of the UHRF1 interaction interface (DPPA3-KRR, DPPA3-R107E, DPPA3 1-60) severely diminished the ability of DPPA3 to release UHRF1 from chromatin (Fig. 5c and Supplementary Fig. 6f-k). On the other hand, the C-terminal half of DPPA3 lacking nuclear export signal but retaining UHRF1 interaction came close to fully restoring the mobility of UHRF1 (Fig. 5c and Supplementary Fig. 6i-k). Remarkably, DPPA3-ΔNES mobilized UHRF1 to an even greater extent than DPPA3-WT (Fig. 5c and Supplementary Fig. 6d,e,j,k), suggesting that operative nuclear export might even antagonize DPPA3-mediated inhibition of UHRF1 chromatin binding. Supporting this notion, chemical inhibition of nuclear export using

leptomycin-B (LMB) significantly enhanced the inhibition of UHRF1 chromatin binding in U1G/D3KO ESCs expressing DPPA3-WT (Supplementary Fig. 5h-k). Taken together our data shows that the efficiency of DPPA3-dependent release of UHRF1 from chromatin requires its interaction with UHRF1 but not its nuclear export.

To further address the question whether the nucleocytoplasmic translocation of UHRF1 and impaired UHRF1 chromatin binding both contribute to DPPA3-mediated inhibition of DNA methylation maintenance, we assessed the ability of each DPPA3 mutant to rescue the hypermethylation of LINE-1 elements in Dppa3KO ESCs (Fig. 5d). Strikingly, DPPA3-ΔNES fully rescued the hypermethylation and achieved a greater loss of DNA methylation than DPPA3-WT, whereas DPPA3 mutants lacking the residues important for UHRF1 binding failed to restore low methylation levels (Fig. 5d). Overall, the ability of each DPPA3 mutant to reduce DNA methylation levels closely mirrored the extent to which each mutant impaired UHRF1 chromatin binding (Fig. 5c and Supplementary Fig. 6d-k). In line with the increased mobility of UHRF1 occuring in the absence of DPPA3 nuclear export (Fig. 5c, Supplementary Figs. 5h-k and 6d,e,j,k), the nucleocytoplasmic translocation of UHRF1 is not only dispensable but rather attenuates DPPA3-mediated inhibition of maintenance methylation (Fig. 5d). Collectively, our findings demonstrate the inhibition of UHRF1 chromatin binding, as opposed to nucleocytosolic translocation of UHRF1, to be the primary mechanism by which DPPA3 drives hypomethylation in naïve ESCs.

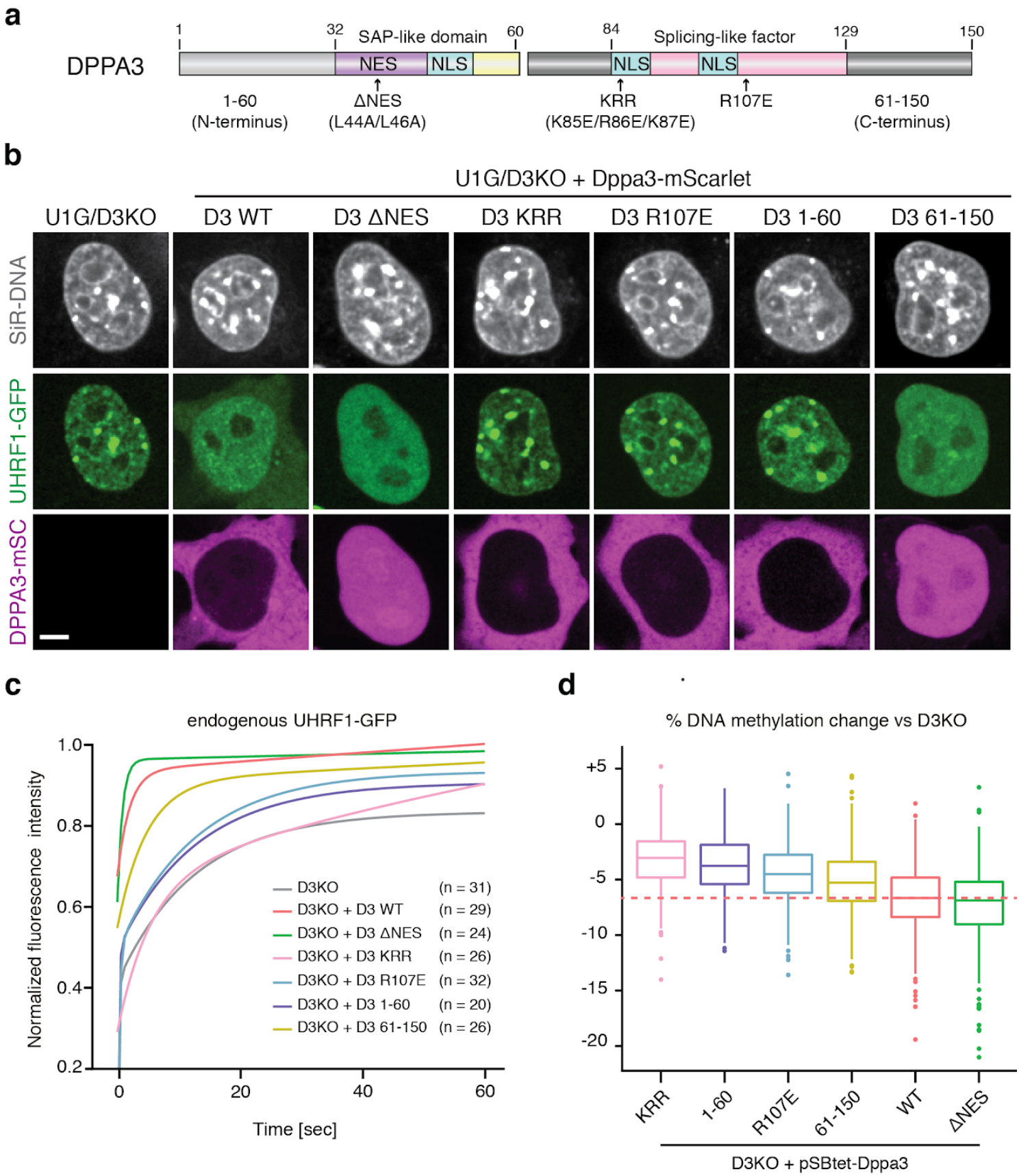


Fig. 5: DPPA3 inhibits maintenance DNA methylation by impairing UHRF1 chromatin binding

a, Schematic illustration of murine DPPA3 with the nuclear localization signals (NLS), nuclear export signal (NES), and predicted

domains (SAP-like and splicing factor-like[74]) annotated. For the DPPA3 mutant forms used in this study, point mutations are indicated with arrows (ΔNES, KRR, R107E) and the two truncations are denoted by the middle break (1-60, left half; 61-150, right half). **b,c,** Nuclear export and the C-terminus of DPPA3 are dispensable for disrupting focal UHRF1 patterning and chromatin binding in ESCs. **b,** Representative confocal images illustrating the localization of endogenous UHRF1-GFP and the indicated mDPPA3-mScarlet fusions in live U1G/D3KO + pSB-D3-mSC ESCs. DNA counterstain: SiR-Hoechst. Scale bar: 5 µm. **c,** FRAP analysis of endogenous UHRF1-GFP in U1G/D3KO ESCs expressing the indicated mutant forms of DPPA3. FRAP Curves (solid lines) indicate double-exponential functions fitted to the FRAP data acquired from *n* cells (shown in the plots). For single-cell FRAP data and additional quantification, see Supplementary Fig. 6d-k. **d,** DPPA3-mediated inhibition of UHRF1 chromatin binding is necessary and sufficient to promote DNA demethylation. Percentage of DNA methylation change at LINE-1 elements (%) in D3KO ESCs after induction of the indicated mutant forms of *Dppa3* as measured by bisulfite sequencing of $n = 4$ biological replicates. In the boxplot in **c,** horizontal black lines within boxes represent median values, boxes indicate the upper and lower quartiles, whiskers indicate the 1.5 interquartile range, and dots indicate outliers. P-values based on Welch's two-sided t-test

DPPA3 binds nuclear UHRF1 with high affinity prompting its release from chromatin in ESCs

We next set out to investigate the mechanistic basis of DPPA3's ability to inhibit UHRF1 chromatin binding in naïve ESCs. DPPA3 has been reported to specifically bind H3K9me2[61], a histone modification critical for UHRF1 targeting[68,75,76]. These prior findings led us to consider two possible mechanistic explanations for DPPA3-mediated UHRF1 inhibition in naïve ESCs: 1) DPPA3 blocks access of UHRF1 to chromatin by competing in binding to H3K9me2, 2) DPPA3 directly or indirectly binds to UHRF1 and thereby prevents it from accessing chromatin.

To simultaneously assess the dynamics of both UHRF1 and DPPA3 under physiological conditions in live ES cells, we employed raster image correlation spectroscopy with pulsed interleaved excitation (PIE-RICS) (Fig. 6a). RICS is a confocal imaging method which enables the measurement of binding and diffusive properties in living cells. Using images acquired on a laser scanning confocal microscope, spatiotemporal information of fluorescently labeled proteins can be extracted from the shape of the spatial autocorrelation function (ACF). A diffusive model is fitted to the ACF which yields the average diffusion coefficient, the concentration, and the fraction of bound molecules[77]. If two proteins are labeled with distinct fluorophores and imaged simultaneously with separate detectors, the extent of their interaction can be extracted from the cross-correlation of their fluctuations using cross-correlation RICS (ccRICS) (Fig. 6a)[78].

We first measured the mobility of DPPA3-mScarlet variants expressed in U1GFP/D3KOs (Supplementary Fig. 7a,b). RICS analysis revealed that over the timescale of the measurements, nuclear DPPA3-WT was predominantly unbound from chromatin and freely diffusing through the nucleus at a rate of 7.18 ± 1.87 $\mu m^2/s$ (Supplementary Fig. 7f). The fraction of mobile DPPA3-mScarlet molecules was measured to be 88.4 ± 5.2% (Fig. 6f), validating globally weak binding inferred from ChIP-Seq profiles[60]. These mobility parameters were largely unaffected by disruption of the UHRF1 interaction, with the DPPA3-KRR mutant behaving similarly to wild-type DPPA3 (Fig. 6f and Supplementary Fig. 7f). To rule out a potential competition between UHRF1 and DPPA3 for H3K9me2 binding, we next used RICS to determine if DPPA3 dynamics are altered in the absence

of UHRF1. For this purpose, we introduced the DPPA3-WT-mScarlet cassette into Uhrf1KO (U1KO) ESCs[79], in which free eGFP is expressed from the endogenous *Uhrf1* promoter (Supplementary Fig. 7c). However, neither the diffusion rate nor the mobile fraction of DPPA3 were appreciably altered in cells devoid of UHRF1, suggesting the high fraction of unbound DPPA3 to be unrelated to the presence of UHRF1 (Fig. 6f and Supplementary Fig. 7f). Overall, our RICS data demonstrate that, in contrast to zygotes [61], DPPA3 in ESCs lacks a strong capacity for chromatin binding, and as such is not engaged in competition with UHRF1 for chromatin binding.

We next used RICS to analyse the dynamics of UHRF1-GFP in response to DPPA3 induction (Fig. 6a). In cells expressing DPPA3-KRR, RICS measurements revealed that only 32.4 ± 10% of UHRF1 is mobile, indicating that the majority of UHRF1 is chromatin-bound (Fig. 6g). In contrast, expression of wild-type DPPA3 lead to a dramatic increase in the mobile fraction UHRF1 (60.6 ± 13.7% mobile fraction for UHRF1) (Fig. 6g and Supplementary Fig. 7g,h). Furthermore, the mobile fraction of UHRF1 increased as a function of the relative abundance of nuclear DPPA3 to UHRF1 (Supplementary Fig. 7i), thereby indicating a stoichiometric effect of DPPA3 on UHRF1 chromatin binding, consistent with a physical interaction. Thus, these results demonstrate that DPPA3 potently disrupts UHRF1 chromatin binding in live ESCs and suggest its interaction with UHRF1 to be critical to do so.

To determine whether such an interaction is indeed present in the nuclei of live ESCs, we performed cross-correlation RICS (ccRICS) (Fig. 6a). We first validated ccRICS in ESCs by analyzing live cells expressing a tandem eGFP-mScarlet fusion (Fig. 6e and Supplementary Fig. 7d), or expressing both freely diffusing eGFP and mScarlet (Fig. 6d and Supplementary Fig. 7e). For the tandem eGFP-mScarlet fusion, we observed a clear positive cross-correlation indicative of eGFP and mScarlet existing in the same complex (Fig. 6e,h), as would be expected for an eGFP-mScarlet fusion. On the other hand, freely diffusing eGFP and mScarlet yielded no visible cross-correlation (Fig. 6d,h), consistent with two independent proteins which do not interact. Upon applying ccRICS to nuclear UHRF1-GFP and DPPA3-mScarlet, we observed a prominent cross-correlation between wild-type DPPA3 and the primarily unbound fraction of UHRF1 (Fig. 6b,h), indicating that mobilized UHRF1 exists in a high affinity complex with DPPA3 in live ESCs. In marked contrast, DPPA3-KRR and UHRF1-GFP failed to exhibit detectable cross-correlation (Fig. 6c,h), consistent with the DPPA3-KRR mutant's diminished capacity to bind[49] and mobilize UHRF1 (Fig. 5c and Supplementary Fig. 6f,j,k). Overall, these findings demonstrate that nuclear DPPA3 interacts with UHRF1 to form a highly mobile complex in naïve ESCs which precludes UHRF1 chromatin binding.

To determine whether the DPPA3-UHRF1 complex identified *in vivo* (Fig. 6h) corresponds to a high affinity direct interaction, we performed microscale thermophoresis (MST) measurements using recombinant UHRF1-GFP and DPPA3 proteins. MST analysis revealed a direct and high affinity (K_D: 0.44 μM) interaction between the DPPA3 WT and UHRF1 (Fig. 6i). No binding was observed for DPPA3 1-60, lacking the residues essential for interaction with UHRF1 (Fig. 6i). In line with the results obtained by ccRICS, these data support the notion that DPPA3 directly binds UHRF1 *in vivo*. Interestingly, the affinity of the UHRF1-DPPA3 interaction

was comparable or even far greater than that reported for the binding of UHRF1 to H3K9me3 or unmodified H3 peptides, respectively[80,81].

To better understand how UHRF1 chromatin loading is impaired by its direct interaction with DPPA3, we applied a fluorescent-three-hybrid (F3H) assay to identify the UHRF1 domain bound by DPPA3 *in vivo* (Supplementary Fig. 7j,k). In short, this method relies on a cell line harboring an array of lac operator binding sites in the nucleus at which a GFP-tagged "bait" protein can be immobilized and visualized as a spot. Thus, the extent of recruitment of an mScarlet-tagged "prey" protein to the nuclear GFP-spot offers a quantifiable measure of the interaction propensity of the "bait" and "prey" proteins *in vivo* (Supplementary Fig. 7k)[82]. Using UHRF1-GFP domain deletions as the immobilized bait (Supplementary Fig. 7j,k), we assessed how the loss of each domain affected the recruitment of mDPPA3-mScarlet to the GFP spot. In contrast to the other UHRF1 domain deletions, removal of the PHD domain essentially abolished recruitment of DPPA3 to the lac spot, demonstrating DPPA3 binds UHRF1 via its PHD domain *in vivo* (Supplementary Fig. 7l,m). The PHD of UHRF1 is essential for its recruitment to chromatin[73,81,83], ubiquitination of H3 and recruitment of DNMT1 to replication foci[66,67]. Thus, our *in vivo* results suggest that the high affinity interaction of DPPA3 with UHRF1's PHD domain precludes UHRF1 from binding chromatin in ESCs, which is also supported by a recent report demonstrating that DPPA3 specifically binds the PHD domain of UHRF1 to competitively inhibit H3 tail binding *in vitro*[84].

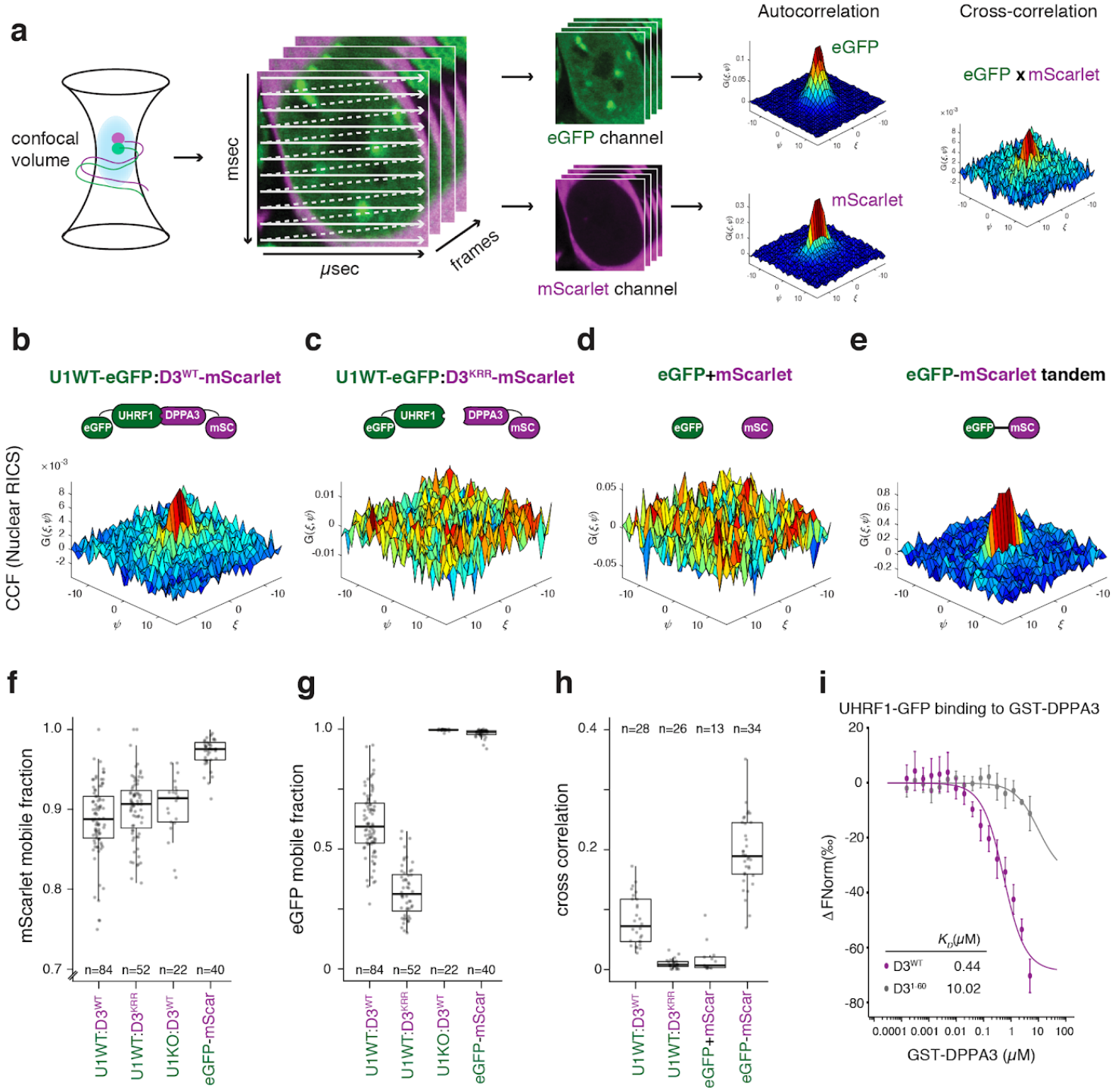


Fig. 6: DPPA3 binds nuclear UHRF1 with high affinity prompting its release from chromatin in ESCs

a, Overview of RICS and ccRICS. Confocal image series are acquired using a calibrated point-scanning laser, generating spatio-temporal fluorescence information on the microsecond and millisecond timescales. An autocorrelation function (ACF) is calculated from the fluorescence fluctuations and used to fit a diffusive model[77]. The cross-correlation of fluctuations between two channels is used to estimate the co-occurence of two fluorescent molecules within live cells[78]. The mean cross-correlation of fluctuations is calculated and shown in the 3D plot, color-coded according to the correlation value. **b,e,** Representative plots of the cross-correlation function (CCF) between the depicted fluorescent molecules in cells from each cell line measured, including U1G/D3KO + pSBtet-D3 ESCs expressing the following forms of DPPA3-mScalet: (**b**) wild-type (U1WT:D3^{WT}) and (**c**) K85E/R85E/K87E mutant (U1WT:D3^{KRR}), and control ESCs expressing (**d**) free eGFP and free mScarlet (eGFP + mScarlet) and (**e**) an eGFP-mScarlet tandem fusion (eGFP-mScarlet). See Supplementary Fig. 7 for the images and ACF plots of the cells used to make

the representative CCF plots. **f,g,** Mobile fraction of (**f**) mScarlet and (**g**) eGFP species in the cell lines depicted in (b, c, and e) as well as in Uhrf1KO ESCs expressing free eGFP and wild-type DPPA3-mScarlet (U1KO:D3^{WT}) (Supplementary Fig. 7c). The mobile fraction was derived from a two-component model fit of the autocorrelation function. Data are pooled from three (U1WT:D3^{WT}, U1WT:D3^{KRR}) or two (U1KO:D3^{WT}, eGFP-mScar) independent experiments. **h,** Mean cross-correlation values of mobile eGFP and mScarlet measured in the cell lines depicted in (**b-e**)**.** The fast timescale axis is indicated by ξ, and the slow timescale axis is indicated by ψ. Data are pooled from two independent experiments.**i,** Microscale thermophoresis measurements of UHRF1-eGFP binding to GST-DPPA3 WT (D3^{WT}) or GST-DPPA3 1-60 (D3$^{1-60}$). Error bars indicate mean ± SEM of n = 2 technical replicates from n = 4 independent experiments. In **f-h**, each data point represents the measured and fit values from a single cell where n = number of cells measured (indicated in the plots). In the box plots, darker horizontal lines within boxes represent median values. The limits of the boxes indicate upper and lower quartiles, and whiskers indicate the 1.5-fold interquartile range.

DPPA3 can inhibit UHRF1 function and drive global DNA demethylation in distantly related, non-mammalian species

Whereas UHRF1 and TET proteins are widely conserved throughout plants and vertebrates[34,85], both early embryonic global hypomethylation[86] and the *Dppa3* gene are unique to mammals. Consistent with UHRF1's conserved role in maintenance DNA methylation, a multiple sequence alignment of UHRF1's PHD domain showed that the residues critical for the recognition of histone H3 are completely conserved from mammals to invertebrates (Fig. 7a). This prompted us to consider the possibility that DPPA3 might be capable of modulating the function of distantly related UHRF1 homologs outside of mammals. To test this hypothesis, we used amphibian (*Xenopus laevis*) egg extracts to assess the ability of mouse DPPA3 (mDPPA3) to interact with a non-mammalian form of UHRF1. Despite the 360 million year evolutionary distance between mouse and *Xenopus*[87], mDPPA3 not only bound *Xenopus* UHRF1 (xUHRF1) with high affinity (Fig. 7b,c and Supplementary Fig. 8a,b) it also interacted with xUHRF1 specifically via its PHD domain (Supplementary Fig. 8c-e). Moreover, the first 60 amino acids of DPPA3 were dispensable for its interaction with UHRF1 (Supplementary Fig. 8a,b). Interestingly, mutation to R107, reported to be critical for DPPA3's binding with mouse UHRF1[49], diminished but did not fully disrupt the interaction (Supplementary Fig. 8b,e). The R107E mutant retained the ability to bind the xUHRF1-PHD domain but exhibited decreased binding to xUHRF1-PHD-SRA under high-salt conditions (Supplementary Fig. 8e), suggesting that R107E changes the binding mode of mDPPA3 to xUHRF1, rather than inhibiting the complex formation. Considering the remarkable similarity between DPPA3's interaction with mouse and *Xenopus* UHRF1, we reasoned that the ability of DPPA3 to inhibit UHRF1 chromatin binding and maintenance DNA methylation might be transferable to *Xenopus*. To address this, we took advantage of a cell-free system derived from interphase *Xenopus* egg extracts to reconstitute DNA maintenance methylation[66]. Remarkably, recombinant mDPPA3 completely disrupted chromatin binding of both *Xenopus* UHRF1 and DNMT1 without affecting the loading of replication factors such as xCDC45, xRPA2, and xPCNA (Fig. 7d). We determined that the inhibition of xUHRF1 and xDNMT1 chromatin loading only requires DPPA3's C-terminus (61-150 a.a.) and is no longer possible upon mutation of R107 (R107E) (Supplementary Fig. 8h), in line with our results in mouse ESCs (Fig. 5d). Moreover, DPPA3-mediated inhibition of xUHRF1 chromatin loading resulted in the severe perturbation of histone H3 dual-monoubiquitylation (H3Ub2), which is necessary for the recruitment

of DNMT1[66,67,88] (Supplementary Fig. 8f). To determine whether mDPPA3 can displace xUHRF1 already bound to chromatin, we first depleted *Xenopus* egg extracts of xDNMT1 to stimulate the hyper-accumulation of xUHRF1 on chromatin[66,89] and then added recombinant mDPPA3 after S-phase had commenced (Supplementary Fig. 8g). Under these conditions, both wild-type mDPPA3 and the 61-150 fragment potently displaced xUHRF1 from chromatin, leading to suppressed H3 ubiquitylation (Supplementary Fig. 8g).

We next assessed the effect of DPPA3 on *Xenopus* maintenance DNA methylation. Consistent with the severe disruption of xDNMT1 chromatin loading, both DPPA3 wild-type and 61-150 effectively abolished replication-dependent DNA methylation in *Xenopus* egg extracts (Fig. 7e). In contrast, DPPA3 1-60 and DPPA3 R107E, which both failed to suppress xUHRF1 and xDNMT1 binding, did not significantly alter maintenance DNA methylation activity (Fig. 7e and Supplementary Fig. 8d,e). Taken together, our data demonstrate DPPA3 to be capable to potently inhibit maintenance DNA methylation in a non-mammalian system.

These findings raised the question whether a single protein capable of inhibiting UHRF1 function like DPPA3 could establish a mammalian-like global hypomethylation during the early embryonic development of a non-mammalian organism. To explore this possibility we turned to the biomedical model fish, medaka (*Oryzias latipes*), which does not exhibit genome-wide erasure of DNA methylation[90] and diverged from mammals 450 million years ago[87]. We injected medaka embryos with *Dppa3* mRNA at the one-cell stage and then tracked their developmental progression. Remarkably, medaka embryos injected with *Dppa3* failed to develop beyond the blastula stage (Fig. 7f) and exhibited a near-complete elimination of global DNA methylation as assessed by immunofluorescence (Fig. 7g). DPPA3-mediated DNA methylation loss was both dose dependent and sensitive to the R107E mutation, which induced only partial demethylation (Supplementary Fig. 8i). Interestingly, medaka embryos injected with DPPA3 R107E showed far fewer developmental defects than those injected with wild-type DPPA3 (Fig. 7h), suggesting that the embryonic arrest resulting from DPPA3 expression is a consequence of the global loss of DNA methylation. Taken together, these results demonstrate that mammalian DPPA3 can inhibit UHRF1 to drive passive demethylation in distant, non-mammalian contexts.

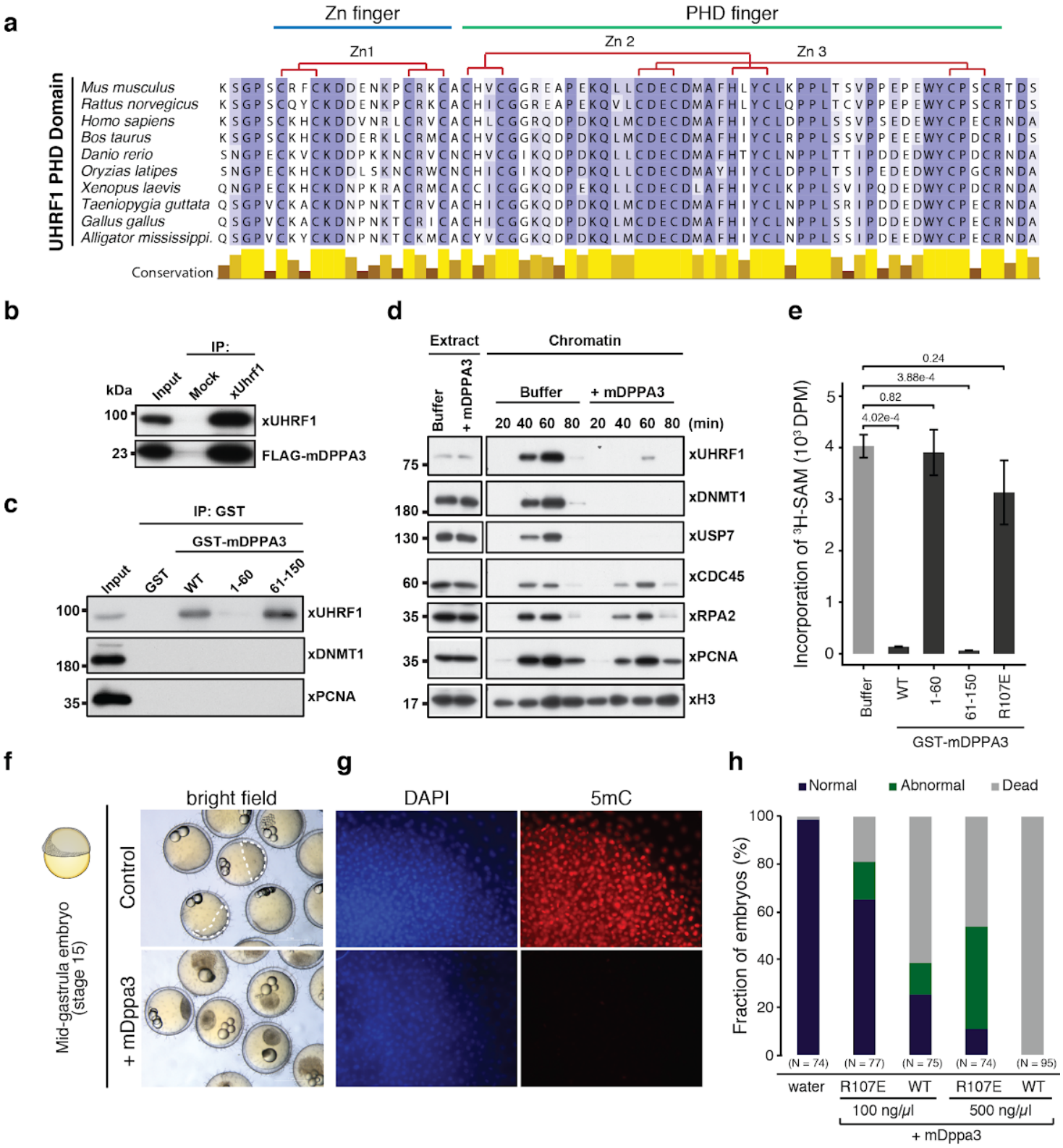

a
Zn finger
PHD finger
Zn1
Zn 2
Zn 3
UHRF1 PHD Domain
Mus musculus
Rattus norvegicus
Homo sapiens
Bos taurus
Danio rerio
Oryzias latipes
Xenopus laevis
Taeniopygia guttata
Gallus gallus
Alligator mississippi.
Conservation
b
IP:
kDa
Input
Mock
xUhrf1
100
23
xUHRF1
FLAG-mDPPA3
c
IP: GST
GST-mDPPA3
Input
GST
WT
1-60
61-150
100
180
35
xUHRF1
xDNMT1
xPCNA
d
Extract
Chromatin
Buffer
+ mDPPA3
Buffer
+ mDPPA3
20 40 60 80 20 40 60 80 (min)
75
180
130
60
35
35
17
xUHRF1
xDNMT1
xUSP7
xCDC45
xRPA2
xPCNA
xH3
e
Incorporation of ³H-SAM (10³ DPM)
0.24
3.88e-4
0.82
4.02e-4
Buffer
WT
1-60
61-150
R107E
GST-mDPPA3
f
Mid-gastrula embryo (stage 15)
Control
+ mDppa3
g
bright field
DAPI
5mC
h
Normal
Abnormal
Dead
Fraction of embryos (%)
(N = 74)
(N = 77)
(N = 75)
(N = 74)
(N = 95)
water
R107E
WT
R107E
WT
100 ng/µl
500 ng/µl
+ mDppa3
i
TET1/2
PRDM14
Dppa3
AAA
DNMT1/UHRF1
DPPA3
H3
Ub
UHRF1
PHD
DNMT1
CpG
mCpG
DPPA3
hemiCpG
j
Global
(mammal-specific)
Promoter-specific
(conserved)
DPPA3
TET1/2
passive demethylation
active demethylation
DNMT1/UHRF1
C
mC
hmC
fC
caC
base excision/TDG

Fig. 7: DPPA3 evolved in boreoeutherian mammals but also functions in lower vertebrates

a, Protein sequence alignment of the PHD domain of the UHRF1 family. The UHRF1 PHD domain shows high amino acid conservation throughout vertebrates, especially the residues involved in Zinc coordination (indicated above). **b,** Endogenous xUHRF1 binds mDPPA3. IPs were performed on *Xenopus* egg extracts incubated with FLAG-mDPPA3 using either a control (Mock) or anti-xUHRF antibody and then analyzed by immunoblotting using the indicated antibodies. **c,** xUHRF1 binds the C-terminus of mDPPA3. GST-tagged mDPPA3 wild-type (WT), point mutant R107E, and truncations (1-60 and 61-150) were immobilized on GSH beads and incubated with *Xenopus* egg extracts. Bound proteins were analyzed using the indicated antibodies. **d,** mDPPA3 inhibits xUHRF1 and xDNMT1 chromatin binding. Sperm chromatin was incubated with interphase *Xenopus* egg extracts supplemented with buffer (+buffer) or GST-mDPPA3 (+mDPPA3). Chromatin fractions were isolated and subjected to immunoblotting using the antibodies indicated. **e,** mDPPA3 inhibits maintenance DNA methylation in *Xenopus*. The efficiency of maintenance DNA methylation was assessed by the incorporation of radiolabelled methyl groups from S-[methyl-^{3}H]-adenosyl-L-methionine (^{3}H-SAM) into DNA purified from egg extracts. Disintegrations per minute (DPM). Error bars indicate mean ± SD calculated from n = 4 independent experiments. Depicted p-values based on a Student's two-sided t-test. **f,** mDPPA3 compromises the embryonic development of medaka. Representative images of developing mid-gastrula stage embryos (control injection) and arrested, blastula stage embryos injected with *mDppa3*. Injections were performed on one-cell stage embryos and images were acquired ~18 h after fertilization. **g,** mDPPA3 drives global DNA demethylation in medaka embryos. Representative 5mC immunostainings in control and *mDppa3*-injected medaka embryos at the late blastula stage (~8 h after fertilization). The assay was repeated > 3 times with similar results. Scale bars represent 50 µm. DNA counterstain: DAPI,4′,6-diamidino-2-phenylindole. **h,** Global hypomethylation causes developmental arrest in medaka. Percentage of normal, abnormal, or dead. Embryos were injected with wild-type *mDppa3* (WT) or *mDppa3* R107E (R107E) at two different concentrations (100 ng/µl or 500 ng/µl) or water at the one-cell stage and analyzed ~18 h after fertilization. N = number of embryos from n = 3 independent injection experiments. **i,** TET1 and TET2 are recruited by PRDM14 to the promoter of *Dppa3* where they promoter active DNA demethylation and transcription of *Dppa3*. DPPA3 is expressed and inhibits maintenance DNA methylation by directly binding UHRF1 and releasing it from chromatin. **j,** TET1 and TET2 control DNA methylation levels by two evolutionary and mechanistically distinct pathways.

440

DISCUSSION

While the appearance of genome-wide DNA demethylation in mammals represents a momentous change with far-reaching consequences to epigenetic gene regulation during early development, the key enzymes involved in DNA modification are highly conserved in vertebrates. As the role of TET enzymes in active demethylation is well documented[91], we investigated their contribution to the hypomethylated state of naïve ESCs. Mutation of the catalytic core of TET enzymes caused - as expected - a genome-wide increase in DNA methylation but mostly at sites where TET proteins do not bind suggesting a rather indirect mechanism. Among the few genes depending on TET activity for expression in naïve ESCs and downregulated at the transition to EpiLCs was *Dppa3.* Demethylation at the *Dppa3* locus coincides with TET1 and TET2 binding and TDG dependent removal of oxidized cytosine residues via base excision repair. DPPA3 in turn binds and displaces UHRF1 from chromatin and thereby prevents the recruitment of DNMT1 and the maintenance of DNA methylation in ESCs (see graphic summary in Fig. 7).

Despite long recognized as a marker of naïve ESCs resembling the inner cell mass[53,92], we provide, to our knowledge, the first evidence that DPPA3 directly promotes the genome-wide DNA hypomethylation characteristic of mammalian naïve pluripotency. This unique pathway, in which TET proteins indirectly cause passive demethylation, is based upon two uniquely mammalian innovations: the expression of TET genes in pluripotent cell types[39,63,93] and the evolution of the novel *Dppa3* gene, positioned within a pluripotency gene cluster and dependent on TET activity for expression. In support of this novel pathway for passive demethylation, we found that TET mutant ESCs show a similar phenotype as Dppa3KO cells with respect to UHRF1 inhibition and hypomethylation and can be rescued by ectopic expression of *Dppa3*.

Our findings also provide the missing link to reconcile previous, apparently conflicting reports. To date, three distinct mechanisms have been proposed for global hypomethylation accompanying naïve pluripotency: TET-mediated active demethylation[37,40,94], impaired maintenance DNA methylation[94], and PRDM14-dependent suppression of methylation[36,37,51]. As both a downstream target of TETs and PRDM14 as well as a direct inhibitor of maintenance DNA methylation, DPPA3 mechanistically links and integrates these three proposed pathways of demethylation (see graphic summary in Fig. 7).

Our mechanistic data showing DPPA3 to displace UHRF1 and DNMT1 from chromatin provide a conclusive explanation for the previous observation that global hypomethylation in naïve ESCs was accompanied by reduced levels of UHRF1 at replication foci[94]. The hypomethylated state of naïve ESCs has also been reported to be dependent on PRDM14[36,51], which has been suggested to promote demethylation by repressing *de novo* DNA methyltransferases[36,40,51,57]. However, recent studies have demonstrated that the loss of *de novo* methylation only marginally affects DNA methylation levels in mouse and human ESCs[94,95]. Interestingly, the loss of *Prdm14* lead to global hypermethylation and, to our surprise, also the downregulation of *Dppa3*[51,96,97]. Our results suggest that the reported ability of PRDM14 to promote hypomethylation in naïve ESCs largely relies on its activation of the *Dppa3* gene ultimately leading to an inhibition of maintenance methylation.

The comparison of TET catalytic mutants and Dppa3KO ESCs allows us to distinguish TET-dependent passive DNA demethylation mediated by DPPA3 from *bona fide* active demethylation. We show that TET activity is indispensable for the active demethylation of a subset of promoters in naïve ESCs, especially those of developmental genes. These findings uncover two evolutionary and mechanistically distinct functions of TET catalytic activity. Whereas TET-mediated active demethylation of developmental genes is evolutionarily conserved among vertebrates[63,98–100], the use of TET proteins to promote global demethylation appears to be specific to mammalian pluripotency[37,40,94] and mediated by the recently evolved *Dppa3* (Figs. 2c and 7j).

To date, our understanding of DPPA3's function in the regulation of DNA methylation had been clouded by seemingly conflicting reports from different developmental stages and cell types. DPPA3's ability to modulate DNA methylation was first described in the context of zygotes[48], where it was subsequently demonstrated to specifically protect the maternal genome from TET3-dependent demethylation[27,58,101]. In contrast, DPPA3 was later shown to facilitate DNA demethylation during PGC specification[46], iPSC reprogramming[59] and oocyte maturation[102,103]. Whereas DPPA3 was shown to disrupt UHRF1 function by sequestering it to the cytoplasm in oocytes[102], we demonstrate that DPPA3-mediated nucleocytoplasmic translocation of UHRF1 is not only dispensable but actually attenuates DPPA3's promotion of hypomethylation in ESCs. In light of our data from naïve ESCs, *Xenopus*, and medaka, DPPA3's capacity to directly bind UHRF1's PHD domain and thereby inhibit UHRF1 chromatin binding appears to be its most basal function. Considering that DPPA3 localization is highly dynamic during the different developmental time periods at which it is expressed[46,62,104], it stands to reason that its role in modulating DNA methylation might also be dynamically regulated by yet-to-be determined regulatory mechanisms. For example, immediately following fertilization, full length DPPA3 is cleaved and its C-terminal domain is specifically degraded[62]. Interestingly, we identified this exact C-terminal stretch of DPPA3 to be necessary and sufficient for DPPA3's inhibition of maintenance DNA methylation. Thus, the precisely timed destruction of this crucial domain might offer an explanation for the differing roles of DPPA3 in regulating DNA methylation between oocytes and zygotes[58,101–103].

As the most basic and evolutionarily conserved function of DNA methylation is the repression of transposable elements[105], the emergence of genome-wide DNA demethylation in mammals raises several fundamental questions. While the DPPA3 mediated erasure of parental DNA methylation might facilitate the establishment of new epigenetic patterns during development, it should be noted that non-mammalian vertebrates manage to undergo normal development without genome-wide demethylation. Moreover, the global loss of DNA methylation even poses a severe threat as excessive demethylation triggers derepression of TEs leading to genomic instability and ultimately cell death in most cell types[7–12]. Remarkably, mammalian naïve pluripotent cell types seem to have acquired the ability to tolerate global hypomethylation, suggesting that the evolutionary emergence of DPPA3 was likely accompanied by measures to control and productively integrate this new factor in epigenetic regulation in ESCs. In fact, many TEs in mammals are not only expressed during development but appear to have been co-opted to drive transcriptional networks critical for the establishment of pluripotency and progression through pre-implantation development[106,107].

512 A good example for the functional integration of TEs in regulatory networks is the reactivation of endogenous
513 retroviruses (ERVs) that is critical for the maternal-to-zygotic transition (MZT), the first major developmental
514 step when maternal mRNAs are degraded and zygotic transcription begins[108–111]. Strikingly, the activation of
515 ERVs is severely impaired in *Dppa3* knockout embryos resulting in MZT failure[60]. It is tempting to speculate that
516 mammal-specific demethylation originates from an arms race between TE and host. DPPA3 may have arisen as
517 a means to overcome the host defence system and was then co-opted by the host and gradually integrated
518 into regulatory networks during evolution. Such a scenario is compatible with the unique occurrence of *Dppa3*
519 in mammals and with our finding that DPPA3 alone is sufficient to inhibit DNA methylation maintenance in
520 *Xenopus* and medaka, species that harbor no *Dppa3* gene and exhibits constant DNA methylation levels at all
521 stages of development[13,14]. Follow-up studies that investigate the origin of *Dppa3* and whether a similar
522 rewiring of early development may have occurred in other, not yet studied branches of vertebrates, are needed
523 to further explore this evolutionary scenario.

REFERENCES

1. Ladstätter, S. & Tachibana, K. Genomic insights into chromatin reprogramming to totipotency in embryos. *J. Cell Biol.* **218**, 70–82 (2019).
2. Warren, I. A. *et al.* Evolutionary impact of transposable elements on genomic diversity and lineage-specific innovation in vertebrates. *Chromosome Res.* **23**, 505–531 (2015).
3. Arkhipova, I. R. Neutral Theory, Transposable Elements, and Eukaryotic Genome Evolution. *Mol. Biol. Evol.* **35**, 1332–1337 (2018).
4. Jacobs, F. M. J. *et al.* An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature* **516**, 242–245 (2014).
5. Friedli, M. & Trono, D. The developmental control of transposable elements and the evolution of higher species. *Annu. Rev. Cell Dev. Biol.* **31**, 429–451 (2015).
6. Rowe, H. M. & Trono, D. Dynamic control of endogenous retroviruses during development. *Virology* **411**, 273–287 (2011).
7. Walsh, C. P., Chaillet, J. R. & Bestor, T. H. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20**, 116–117 (1998).
8. Jackson-Grusby, L. *et al.* Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat. Genet.* **27**, 31–39 (2001).
9. Chernyavskaya, Y. *et al.* Loss of DNA methylation in zebrafish embryos activates retrotransposons to trigger antiviral signaling. *Development* **144**, 2925–2939 (2017).
10. Iida, A. *et al.* Targeted reduction of the DNA methylation level with 5-azacytidine promotes excision of the medaka fish Tol2 transposable element. *Genet. Res.* **87**, 187–193 (2006).
11. Chiappinelli, K. B. *et al.* Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. *Cell* **169**, 361 (2017).
12. Roulois, D. *et al.* DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. *Cell* **162**, 961–973 (2015).
13. Veenstra, G. J. & Wolffe, A. P. Constitutive genomic methylation during embryonic development of

551 Xenopus. *Biochim. Biophys. Acta* **1521**, 39–44 (2001).

552 14. Stancheva, I., El-Maarri, O., Walter, J., Niveleau, A. & Meehan, R. R. DNA methylation at promoter regions
553 regulates the timing of gene activation in Xenopus laevis embryos. *Dev. Biol.* **243**, 155–165 (2002).

554 15. Ortega-Recalde, O., Day, R. C., Gemmell, N. J. & Hore, T. A. Zebrafish preserve global germline DNA
555 methylation while sex-linked rDNA is amplified and demethylated during feminisation. *Nat. Commun.* **10**,
556 3053 (2019).

557 16. Skvortsova, K. *et al.* Retention of paternal DNA methylome in the developing zebrafish germline. *Nat.*
558 *Commun.* **10**, 3054 (2019).

559 17. Monk, M., Boubelik, M. & Lehnert, S. Temporal and regional changes in DNA methylation in the embryonic,
560 extraembryonic and germ cell lineages during mouse embryo development. *Development* **99**, 371–382
561 (1987).

562 18. Sanford, J. P., Clark, H. J., Chapman, V. M. & Rossant, J. Differences in DNA methylation during oogenesis
563 and spermatogenesis and their persistence during early embryogenesis in the mouse. *Genes Dev.* **1**,
564 1039–1046 (1987).

565 19. Lee, H. J., Hore, T. A. & Reik, W. Reprogramming the methylome: erasing memory and creating diversity.
566 *Cell Stem Cell* **14**, 710–719 (2014).

567 20. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL
568 partner TET1. *Science* **324**, 930–935 (2009).

569 21. He, Y.-F. *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA.
570 *Science* **333**, 1303–1307 (2011).

571 22. Cortellino, S. *et al.* Thymine DNA glycosylase is essential for active DNA demethylation by linked
572 deamination-base excision repair. *Cell* **146**, 67–79 (2011).

573 23. Ito, S. *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine.
574 *Science* **333**, 1300–1303 (2011).

575 24. Rougier, N. *et al.* Chromosome methylation patterns during mammalian preimplantation development.
576 *Genes Dev.* **12**, 2108–2113 (1998).

577 25. Howlett, S. K. & Reik, W. Methylation levels of maternal and paternal genomes during preimplantation
578 development. *Development* **113**, 119–127 (1991).

579 26. Carlson, L. L., Page, A. W. & Bestor, T. H. Properties and localization of DNA methyltransferase in
580 preimplantation mouse embryos: implications for genomic imprinting. *Genes Dev.* **6**, 2536–2541 (1992).
581 27. Wossidlo, M. *et al.* 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic
582 reprogramming. *Nat. Commun.* **2**, 241 (2011).
583 28. Gu, T.-P. *et al.* The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**,
584 606–610 (2011).
585 29. Wang, L. *et al.* Programming and inheritance of parental DNA methylomes in mammals. *Cell* **157**, 979–991
586 (2014).
587 30. Guo, F. *et al.* Active and passive demethylation of male and female pronuclear DNA in the mammalian
588 zygote. *Cell Stem Cell* **15**, 447–459 (2014).
589 31. Shen, L. *et al.* Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes
590 in mouse zygotes. *Cell Stem Cell* **15**, 459–471 (2014).
591 32. Amouroux, R. *et al.* De novo DNA methylation drives 5hmC accumulation in mouse zygotes. *Nat. Cell Biol.*
592 **18**, 225–233 (2016).
593 33. Zemach, A., McDaniel, I. E., Silva, P. & Zilberman, D. Genome-wide evolutionary analysis of eukaryotic
594 DNA methylation. *Science* **328**, 916–919 (2010).
595 34. Iyer, L. M., Tahiliani, M., Rao, A. & Aravind, L. Prediction of novel families of enzymes involved in oxidative
596 and other complex modifications of bases in nucleic acids. *Cell Cycle* **8**, 1698–1710 (2009).
597 35. Marks, H. *et al.* The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149**,
598 590–604 (2012).
599 36. Leitch, H. G. *et al.* Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol.*
600 *Biol.* **20**, 311–316 (2013).
601 37. Hackett, J. A. *et al.* Synergistic mechanisms of DNA demethylation during transition to ground-state
602 pluripotency. *Stem Cell Reports* **1**, 518–531 (2013).
603 38. Rasmussen, K. D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes*
604 *Dev.* **30**, 733–750 (2016).
605 39. Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass
606 specification. *Nature* **466**, 1129–1133 (2010).

607 40. Ficz, G. *et al.* FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic
608 ground state of pluripotency. *Cell Stem Cell* **13**, 351–359 (2013).
609 41. Boroviak, T. *et al.* Lineage-Specific Profiling Delineates the Emergence and Progression of Naive
610 Pluripotency in Mammalian Embryogenesis. *Dev. Cell* **35**, 366–382 (2015).
611 42. Habibi, E. *et al.* Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of
612 mouse embryonic stem cells. *Cell Stem Cell* **13**, 360–369 (2013).
613 43. Pfaffeneder, T. *et al.* Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA.
614 *Nat. Chem. Biol.* **10**, 574–581 (2014).
615 44. Khoueiry, R. *et al.* Lineage-specific functions of TET1 in the postimplantation mouse embryo. *Nat. Genet.*
616 (2017). doi:10.1038/ng.3868
617 45. Xiong, J. *et al.* Cooperative Action between SALL4A and TET Proteins in Stepwise Oxidation of
618 5-Methylcytosine. *Mol. Cell* **64**, 913–925 (2016).
619 46. Nakashima, H. *et al.* Effects of dppa3 on DNA methylation dynamics during primordial germ cell
620 development in mice. *Biol. Reprod.* **88**, 125 (2013).
621 47. Funaki, S. *et al.* Inhibition of maintenance DNA methylation by Stella. *Biochem. Biophys. Res. Commun.*
622 **453**, 455–460 (2014).
623 48. Nakamura, T. *et al.* PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat. Cell*
624 *Biol.* **9**, 64–71 (2007).
625 49. Li, Y. *et al.* Stella safeguards the oocyte methylome by preventing de novo methylation mediated by
626 DNMT1. *Nature* **564**, 136–140 (2018).
627 50. Magnúsdóttir, E. *et al.* A tripartite transcription factor network regulates primordial germ cell specification
628 in mice. *Nat. Cell Biol.* **15**, 905–915 (2013).
629 51. Yamaji, M. *et al.* PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic
630 pathways in mouse embryonic stem cells. *Cell Stem Cell* **12**, 368–382 (2013).
631 52. Okashita, N. *et al.* PRDM14 promotes active DNA demethylation through the ten-eleven translocation
632 (TET)-mediated base excision repair pathway in embryonic stem cells. *Development* **141**, 269–280 (2014).
633 53. Hayashi, K., Lopes, S. M. C. de S., Tang, F. & Surani, M. A. Dynamic equilibrium and heterogeneity of
634 mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* **3**, 391–401

635 (2008).
636 54. Kalkan, T. *et al.* Tracking the embryonic stem cell transition from ground state pluripotency. *Development*
637 **144**, 1221–1234 (2017).
638 55. Auclair, G., Guibert, S., Bender, A. & Weber, M. Ontogeny of CpG island methylation and specificity of
639 DNMT3 methyltransferases during embryonic development in the mouse. *Genome Biol.* **15**, 545 (2014).
640 56. Shen, L. *et al.* Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation
641 dynamics. *Cell* **153**, 692–706 (2013).
642 57. Grabole, N. *et al.* Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and
643 DNA methylation. *EMBO Rep.* **14**, 629–637 (2013).
644 58. Han, L. *et al.* Embryonic defects induced by maternal obesity in mice derive from Stella insufficiency in
645 oocytes. *Nat. Genet.* **50**, 432–442 (2018).
646 59. Xu, X. *et al.* Dppa3 expression is critical for generation of fully reprogrammed iPS cells and maintenance of
647 Dlk1-Dio3 imprinting. *Nat. Commun.* **6**, 6008 (2015).
648 60. Huang, Y. *et al.* Stella modulates transcriptional and endogenous retrovirus programs during
649 maternal-to-zygotic transition. *Elife* **6**, (2017).
650 61. Nakamura, T. *et al.* PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early
651 embryos. *Nature* **486**, 415–419 (2012).
652 62. Shin, S.-W., John Vogt, E., Jimenez-Movilla, M., Baibakov, B. & Dean, J. Cytoplasmic cleavage of DPPA3
653 is required for intracellular trafficking and cleavage-stage development in mice. *Nat. Commun.* **8**, 1643
654 (2017).
655 63. Bogdanović, O. *et al.* Active DNA demethylation at enhancers during the vertebrate phylotypic period. *Nat.*
656 *Genet.* **48**, 417–426 (2016).
657 64. Sharif, J. *et al.* The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated
658 DNA. *Nature* **450**, 908–912 (2007).
659 65. Hashimoto, H. *et al.* The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* **455**,
660 826–829 (2008).
661 66. Nishiyama, A. *et al.* Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and
662 replication. *Nature* **502**, 249–253 (2013).

663 67. Qin, W. *et al.* DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone
664 ubiquitination. *Cell Res.* **25**, 911–929 (2015).
665 68. Rothbart, S. B. *et al.* Association of UHRF1 with methylated H3K9 directs the maintenance of DNA
666 methylation. *Nat. Struct. Mol. Biol.* **19**, 1155–1160 (2012).
667 69. Zhao, Q. *et al.* Dissecting the precise role of H3K9 methylation in crosstalk with DNA maintenance
668 methylation in mammals. *Nat. Commun.* **7**, 12464 (2016).
669 70. Weihs, F. *et al.* Heterogeneous localisation of membrane proteins in Staphylococcus aureus. *Sci. Rep.* **8**,
670 3657 (2018).
671 71. Osswald, M., Santos, A. F. & Morais-de-Sá, E. Light-Induced Protein Clustering for Optogenetic
672 Interference and Protein Interaction Analysis in Drosophila S2 Cells. *Biomolecules* **9**, (2019).
673 72. Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**,
674 1760–1764 (2007).
675 73. Rothbart, S. B. *et al.* Multivalent histone engagement by the linked tandem Tudor and PHD domains of
676 UHRF1 is required for the epigenetic inheritance of DNA methylation. *Genes Dev.* **27**, 1288–1298 (2013).
677 74. Payer, B. *et al.* Stella is a maternal effect gene required for normal early development in mice. *Curr. Biol.*
678 **13**, 2110–2117 (2003).
679 75. Citterio, E. *et al.* Np95 is a histone-binding protein endowed with ubiquitin ligase activity. *Mol. Cell. Biol.*
680 **24**, 2526–2535 (2004).
681 76. Karagianni, P., Amazit, L., Qin, J. & Wong, J. ICBP90, a novel methyl K9 H3 binding protein linking protein
682 ubiquitination with heterochromatin formation. *Mol. Cell. Biol.* **28**, 705–717 (2008).
683 77. Digman, M. A. *et al.* Measuring fast dynamics in solutions and cells with a laser scanning microscope.
684 *Biophys. J.* **89**, 1317–1327 (2005).
685 78. Digman, M. A., Wiseman, P. W., Horwitz, A. R. & Gratton, E. Detecting protein complexes in living cells
686 from laser scanning confocal image sequences by the cross correlation raster image spectroscopy
687 method. *Biophys. J.* **96**, 707–716 (2009).
688 79. Karg, E. *et al.* Ubiquitome Analysis Reveals PCNA-Associated Factor 15 (PAF15) as a Specific
689 Ubiquitination Target of UHRF1 in Embryonic Stem Cells. *J. Mol. Biol.* **429**, 3814–3824 (2017).
690 80. Fang, J. *et al.* Hemi-methylated DNA opens a closed conformation of UHRF1 to facilitate its histone

691 recognition. *Nat. Commun.* **7**, 11197 (2016).

692 81. Harrison, J. S. *et al.* Hemi-methylated DNA regulates DNA methylation inheritance through allosteric
693 activation of H3 ubiquitylation by UHRF1. *Elife* **5**, (2016).

694 82. Herce, H. D., Deng, W., Helma, J., Leonhardt, H. & Cardoso, M. C. Visualization and targeted disruption of
695 protein interactions in living cells. *Nat. Commun.* **4**, 2660 (2013).

696 83. Arita, K. *et al.* Recognition of modification status on a histone H3 tail by linked histone reader modules of
697 the epigenetic regulator UHRF1. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 12950–12955 (2012).

698 84. Du, W. *et al.* Stella protein facilitates DNA demethylation by disrupting the chromatin association of the
699 RING finger-type E3 ubiquitin ligase UHRF1. *J. Biol. Chem.* **294**, 8907–8917 (2019).

700 85. Feng, S. *et al.* Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl.*
701 *Acad. Sci. U. S. A.* **107**, 8689–8694 (2010).

702 86. Wu, S. C. & Zhang, Y. Active DNA demethylation: many roads lead to Rome. *Nat. Rev. Mol. Cell Biol.* **11**,
703 607–620 (2010).

704 87. Kumar, S. & Hedges, S. B. A molecular timescale for vertebrate evolution. *Nature* **392**, 917–920 (1998).

705 88. Ishiyama, S. *et al.* Structure of the Dnmt1 Reader Module Complexed with a Unique Two-Mono-Ubiquitin
706 Mark on Histone H3 Reveals the Basis for DNA Methylation Maintenance. *Mol. Cell* **68**, 350–360.e7 (2017).

707 89. Yamaguchi, L. *et al.* Usp7-dependent histone H3 deubiquitylation regulates maintenance of DNA
708 methylation. *Sci. Rep.* **7**, 55 (2017).

709 90. Walter, R. B., Li, H.-Y., Intano, G. W., Kazianis, S. & Walter, C. A. Absence of global genomic cytosine
710 methylation pattern erasure during medaka (Oryzias latipes) early embryo development. *Comp. Biochem.*
711 *Physiol. B Biochem. Mol. Biol.* **133**, 597–607 (2002).

712 91. Wu, X. & Zhang, Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat. Rev.*
713 *Genet.* (2017). doi:10.1038/nrg.2017.33

714 92. Singer, Z. S. *et al.* Dynamic heterogeneity and DNA methylation in embryonic stem cells. *Mol. Cell* **55**,
715 319–331 (2014).

716 93. Almeida, R. D. *et al.* 5-hydroxymethyl-cytosine enrichment of non-committed cells is not a universal
717 feature of vertebrate development. *Epigenetics* **7**, 383–389 (2012).

718 94. von Meyenn, F. *et al.* Impairment of DNA Methylation Maintenance Is the Main Cause of Global

Demethylation in Naive Embryonic Stem Cells. *Mol. Cell* (2016). doi:10.1016/j.molcel.2016.04.025

95. Liao, J. *et al.* Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat. Genet.* **47**, 469–478 (2015).

96. Grabole, N. *et al.* Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. *EMBO Rep.* **14**, 629–637 (2013).

97. Ma, Z., Swigut, T., Valouev, A., Rada-Iglesias, A. & Wysocka, J. Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates. *Nat. Struct. Mol. Biol.* **18**, 120–127 (2011).

98. Li, X. *et al.* Tet proteins influence the balance between neuroectodermal and mesodermal fate choice by inhibiting Wnt signaling. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E8267–E8276 (2016).

99. Dai, H.-Q. *et al.* TET-mediated DNA demethylation controls gastrulation by regulating Lefty–Nodal signalling. *Nature* **538**, 528 (2016).

100. Verma, N. *et al.* TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. *Nat. Genet.* **50**, 83–95 (2018).

101. Nakamura, T. *et al.* PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486**, 415–419 (2012).

102. Li, Y. *et al.* Stella safeguards the oocyte methylome by preventing de novo methylation mediated by DNMT1. *Nature* **564**, 136–140 (2018).

103. Han, L., Ren, C., Zhang, J., Shu, W. & Wang, Q. Differential roles of Stella in the modulation of DNA methylation during oocyte and zygotic development. *Cell Discov* **5**, 9 (2019).

104. Payer, B. *et al.* Generation of stella-GFP transgenic mice: a novel tool to study germ cell development. *Genesis* **44**, 75–83 (2006).

105. Schmitz, R. J., Lewis, Z. A. & Goll, M. G. DNA Methylation: Shared and Divergent Features across Eukaryotes. *Trends Genet.* **35**, 818–827 (2019).

106. Cosby, R. L., Chang, N.-C. & Feschotte, C. Host-transposon interactions: conflict, cooperation, and cooption. *Genes Dev.* **33**, 1098–1116 (2019).

107. Robbez-Masson, L. & Rowe, H. M. Retrotransposons shape species-specific embryonic stem cell gene expression. *Retrovirology* **12**, 45 (2015).

747 108. Jachowicz, J. W. *et al.* LINE-1 activation after fertilization regulates global chromatin accessibility in the
748 early mouse embryo. *Nat. Genet.* **49**, 1502–1510 (2017).
749 109. Huang, Y. *et al.* Stella modulates transcriptional and endogenous retrovirus programs during
750 maternal-to-zygotic transition. *Elife* **6**, (2017).
751 110. Hendrickson, P. G. *et al.* Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage
752 genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925–934 (2017).
753 111. De Iaco, A. *et al.* DUX-family transcription factors regulate zygotic genome activation in placental
754 mammals. *Nat. Genet.* **49**, 941–945 (2017).
755 112. Hayashi, K. & Saitou, M. Generation of eggs from mouse embryonic stem cells and induced pluripotent
756 stem cells. *Nat. Protoc.* **8**, 1513–1524 (2013).
757 113. Mulholland, C. B. *et al.* A modular open platform for systematic functional studies under physiological
758 conditions. *Nucleic Acids Res.* **43**, e112 (2015).
759 114. Kowarz, E., Löscher, D. & Marschalek, R. Optimized Sleeping Beauty transposons rapidly generate stable
760 transgenic cell lines. *Biotechnol. J.* **10**, 647–653 (2015).
761 115. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*
762 **6**, 343–345 (2009).
763 116. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
764 117. Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G. F. & Chin, L. Post-translational Regulation of
765 Cas9 during G1 Enhances Homology-Directed Repair. *Cell Rep.* **14**, 1555–1566 (2016).
766 118. Mátés, L. *et al.* Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust
767 stable gene transfer in vertebrates. *Nat. Genet.* **41**, 753–761 (2009).
768 119. Bauer, C. *et al.* Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked
769 N-acetylglucosamine transferase (OGT). *J. Biol. Chem.* **290**, 4801–4812 (2015).
770 120. Méndez, J. & Stillman, B. Chromatin association of human origin recognition complex, cdc6, and
771 minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late
772 mitosis. *Mol. Cell. Biol.* **20**, 8602–8612 (2000).
773 121. Wagner, M. *et al.* Age-dependent levels of 5-methyl-, 5-hydroxymethyl-, and 5-formylcytosine in human
774 and mouse brain tissues. *Angew. Chem. Int. Ed Engl.* **54**, 12511–12514 (2015).

775 122. Ziegenhain, C. *et al.* Comparative Analysis of Single-Cell RNA Sequencing Methods. *Mol. Cell* **65**,
776 631–643.e4 (2017).
777 123. Boyle, P. *et al.* Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA
778 methylation profiling. *Genome Biol.* **13**, R92 (2012).
779 124. Parekh, S., Ziegenhain, C., Vieth, B., Enard, W. & Hellmann, I. zUMIs - A fast and flexible pipeline to
780 process RNA sequencing data with UMIs. *Gigascience* **7**, (2018).
781 125. Rau, A., Gallopin, M., Celeux, G. & Jaffrézic, F. Data-based filtering for replicated high-throughput
782 transcriptome sequencing experiments. *Bioinformatics* **29**, 2146–2152 (2013).
783 126. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data
784 with DESeq2. *Genome Biol.* **15**, 550 (2014).
785 127. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional
786 genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
787 128. Akalin, A. *et al.* methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation
788 profiles. *Genome Biol.* **13**, R87 (2012).
789 129. Illingworth, R. S. *et al.* Orphan CpG islands identify numerous conserved promoters in the mammalian
790 genome. *PLoS Genet.* **6**, e1001134 (2010).
791 130. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
792 131. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic*
793 *Acids Res.* **44**, W160–5 (2016).
794 132. Sporbert, A., Domaing, P., Leonhardt, H. & Cardoso, M. C. PCNA acts as a stationary loading platform for
795 transiently interacting Okazaki fragment maturation proteins. *Nucleic Acids Res.* **33**, 3521–3528 (2005).
796 133. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682
797 (2012).
798 134. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat.*
799 *Methods* **9**, 671–675 (2012).
800 135. McQuin, C. *et al.* CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biol.* **16**, e2005970
801 (2018).
802 136. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search

programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).

137. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128 (2007).

138. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).

139. Scheltema, R. A. & Mann, M. SprayQc: A Real-Time LC–MS/MS Quality Monitoring System To Maximize Uptime Using Off the Shelf Components. *J. Proteome Res.* **11**, 3458–3466 (2012).

140. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367 (2008).

141. Cox, J. *et al.* Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).

142. Cox, J., Hein, M. Y., Luber, C. A., Paron, I. & Nagaraj, N. MaxLFQ allows accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction. *Molecular & cellular* (2014).

143. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731 (2016).

144. Iwamatsu, T. Stages of normal development in the medaka Oryzias latipes. *Mech. Dev.* **121**, 605–618 (2004).

145. Hendrix, J. *et al.* Live-cell observation of cytosolic HIV-1 assembly onset reveals RNA-interacting Gag oligomers. *J. Cell Biol.* **210**, 629–646 (2015).

146. Müller, B. K., Zaychikov, E., Bräuchle, C. & Lamb, D. C. Pulsed interleaved excitation. *Biophys. J.* **89**, 3508–3522 (2005).

147. Schrimpf, W., Barth, A., Hendrix, J. & Lamb, D. C. PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. *Biophys. J.* **114**, 1518–1528 (2018).

148. Hendrix, J., Dekens, T., Schrimpf, W. & Lamb, D. C. Arbitrary-Region Raster Image Correlation Spectroscopy. *Biophys. J.* **111**, 1785–1796 (2016).

149. Borodovsky, A. *et al.* A novel active site-directed probe specific for deubiquitylating enzymes reveals

proteasome association of USP14. *EMBO J.* **20**, 5187–5196 (2001).

MATERIALS AND METHODS

Cell culture

Naïve J1 mouse ESCs were cultured and differentiated into EpiLCs as described previously[112,113]. In brief, for both naïve ESCs and EpiLCs defined media was used, consisting of N2B27: 50% neurobasal medium (Life Technologies), 50% DMEM/F12 (Life Technologies), 2 mM L-glutamine (Life Technologies), 0.1 mM β-mercaptoethanol (Life Technologies), N2 supplement (Life Technologies), B27 serum-free supplement (Life Technologies), 100 U/mL penicillin, and 100 µg/mL streptomycin (Sigma). Naïve ESCs were maintained on flasks treated with 0.2% gelatin in defined media containing 2i (1 µM PD032591 and 3 µM CHIR99021 (Axon Medchem, Netherlands)), 1000 U/mL recombinant leukemia inhibitory factor (LIF, Millipore), and 0.3% BSA (Gibco) for at least three passages before commencing differentiation. For reprogramming experiments, naïve media was supplemented with freshly prepared 100 µM Vitamin C (L-ascorbic acid 2-phosphate, Sigma). To differentiate naïve ESCs into Epiblast-like cells (EpiLCs), flasks were first pre-treated with Geltrex (Life Technologies) diluted 1:100 in DMEM/F12 (Life Technologies) and incubated at 37 °C overnight. Naïve ESCs were plated on Geltrex-treated flasks in defined medium containing 10 ng/mL Fgf2 (R&D Systems), 20 ng/mL Activin A (R&D Systems) and 0.1× Knockout Serum Replacement (KSR) (Life Technologies). Media was changed after 24 h and EpiLCs were harvested for RRBS and RNA-seq experiments after 48 h.

For CRISPR-assisted cell line generation, mouse ESCs were maintained on 0.2% gelatin-coated dishes in Dulbecco's modified Eagle's medium (Sigma) supplemented with 16% fetal bovine serum (FBS, Sigma), 0.1 mM ß-mercaptoethanol (Invitrogen), 2 mM L-glutamine (Sigma), 1× MEM Non-essential amino acids (Sigma), 100 U/mL penicillin, 100 µg/mL streptomycin (Sigma), homemade recombinant LIF tested for efficient self-renewal maintenance, and 2i (1 µM PD032591 and 3 µM CHIR99021 (Axon Medchem, Netherlands)).

For experiments in which cells were propagated in “serum LIF” conditions, the cells were maintained on 0.2% gelatin-coated dishes in Dulbecco's modified Eagle's medium (Sigma) supplemented with 16% fetal bovine serum (FBS, Sigma), 0.1 mM ß-mercaptoethanol (Invitrogen), 2 mM L-glutamine (Sigma), 1× MEM Non-essential amino acids (Sigma), 100 U/mL penicillin, 100 µg/mL streptomycin (Sigma), LIF (ESGRO, Millipore).

HESCs (line H9) were maintained in mTeSR1 medium (05850, STEMCELL Technologies) on Matrigel-coated plates (356234, Corning) prepared by 1:100 dilution, and 5 ml coating of 10 cm plates for 1 h at 37 °C. Colonies were passaged using the gentle cell dissociation reagent (07174, StemCell Technologies).

All cell lines were regularly tested for Mycoplasma contamination by PCR.

Sleeping Beauty Constructs

To generate the sleeping beauty donor vector with an N-terminal 3xFLAG tag and a fluorescent readout of doxycycline induction, we first used primers with overhangs harboring SfiI sites to amplify the IRES-DsRed-Express from pIRES2-DsRed-Express (Clontech). This fragment was then cloned into the NruI site in pUC57-GentR via cut-ligation to generate an intermediate cloning vector pUC57-SfiI-IRES-DsRed-Express-SfiI. A synthesized gBlock (IDT, Coralville, IA, USA) containing Kozak-BIO-3XFLAG-AsiSI-NotI-V5 was cloned into the Eco47III site of the intermediate cloning vector via cut-ligation. The luciferase insert from pSBtet-Pur[114] (Addgene plasmid #60507) was excised using SfiI. The SfiI-flanked Kozak-BIO-3XFLAG-AsiSI-NotI-V5-IRES-DsRed-Express cassette was digested out of the intermediate cloning vector using SfiI and ligated into the pSBtet-Pur vector backbone linearized by SfiI. The end result was the parental vector, pSBtet-3xFLAG-IRES-DsRed-Express-PuroR. The pSBtet-3x-FLAG-mScarlet-PuroR vector was constructed by inserting a synthesized gBlock (IDT, Coralville, IA, USA) containing the SfiI-BIO-3XFLAG-AsiSI-NotI-mScarlet sequence into the

871 SfiI-linearized pSBtet-Pur vector backbone using Gibson assembly[115]. For *Dppa3* expression constructs, the coding
872 sequence of wild-type and mutant forms of *Dppa3* were synthesized as gBlocks (IDT, Coralville, IA, USA) and inserted into
873 the pSBtet-3xFLAG-IRES-DsRed-Express-PuroR vector (linearized by AsiSI and NotI) using Gibson assembly. To produce
874 the *Dppa3*-mScarlet fusion expression constructs, wild-type and mutant forms of *Dppa3* were amplified from
875 pSBtet-3xFLAG-Dppa3-IRES-DsRed-Express-PuroR constructs using primers with overhangs homologous to the AsiSI
876 and NotI restriction sites of the pSBtet-3x-FLAG-mScarlet-PuroR vector. Wild-type and mutant *Dppa3* amplicons were
877 subcloned into the pSBtet-3x-FLAG-mScarlet-PuroR vector (linearized with AsiSI and NotI) using Gibson assembly.
878 For experiments involving the SBtet-3xFLAG-*Dppa3* cassette, all inductions were performed using 1 µg/mL doxycycline
879 (Sigma-Aldrich). The DPPA3-WT construct was able to rescue the cytoplasmic localization and chromatin association of
880 UHRF1 indicating that C-terminally tagged DPPA3 remains functional (Fig. 5b-d).

881 **CRISPR/Cas9 genome engineering**

882 For the generation of *Tet1*, *Tet2*, and *Tet1*/*Tet2* catalytic mutants, specific gRNAs targeting the catalytic center of *Tet1* and
883 *Tet2* were cloned into a modified version of the SpCas9-T2A-GFP/gRNA (px458[116], Addgene plasmid #48138), to which
884 we fused a truncated form of human Geminin (hGem) to SpCas9 in order to increase homology-directed repair
885 efficiency[117].
886 A 200 bp ssDNA oligonucleotide harboring the H1652Y and D1654A mutations and ~100 bp of homology to the genomic
887 locus was synthesized (IDT, Coralville, IA, USA). For targetings in wild-type J1 ESCs, cells were transfected with a 4:1 ratio
888 of donor oligo and Cas9/gRNA construct. Positively transfected cells were isolated based on GFP expression using
889 fluorescence-activated cell sorting (FACS) and plated at clonal density in ESC media 2 days after transfection. Cell lysis in
890 96-well plates, PCR on lysates, and restriction digests were performed as previously described[113]. *Tet1* catalytic mutation
891 was confirmed by Sanger sequencing.
892 As C-terminally tagged GFP labeled UHRF1 transgenes were shown to be able to rescue U1KO[67], the tagging of
893 endogenous *Uhrf1* was also performed at the C-terminus. For insertion of the HALO or eGFP coding sequence into the
894 endogenous *Dppa3* and *Uhrf1* loci, respectively, *Dppa3* and *Uhrf1* specific gRNAs were cloned into
895 SpCas9-hGem-T2A-Puromycin/gRNA vector, which is a modified version of SpCas9-T2A-Puromycin/gRNA vector
896 (px459;[116], Addgene plasmid #62988) similar to that described above. To construct the homology donors plasmids,
897 gBlocks (IDT, Coralville, IA, USA) were synthesized containing either the HALO or eGFP coding sequence flanked by
898 homology arms with ~200-400 bp homology upstream and downstream of the gRNA target sequence at the *Dppa3* or
899 *Uhrf1* locus, respectively, and then cloned into the NruI site of pUC57-GentR via cut-ligation. ESCs were transfected with
900 equimolar amounts of gRNA and homology donor vectors. Two days after transfection, cells were plated at clonal density
901 and subjected to a transient puromycin selection (1 ug/mL) for 40 h. After 5-6 days, ESCs positive for HALO or eGFP
902 integration were isolated via fluorescence-activated cell sorting (FACS) and plated again at clonal density in ESC media.
903 After 4-5 days, colonies were picked and plated on Optical bottom µClear 96-well plates and re-screened for the correct
904 expression and localization of eGFP or HALO using live-cell spinning-disk confocal imaging. Cells were subsequently
905 genotyped using the aforementioned cell lysis strategy and further validated by Sanger sequencing[113].
906 To generate *Dppa3* knockout cells, the targeting strategy entailed the use of two gRNAs with target sites flanking the
907 *Dppa3* locus to excise the entire locus on both alleles. gRNA oligos were cloned into the SpCas9-T2A-PuroR/gRNA vector
908 via cut-ligation. ESCs were transfected with an equimolar amount of each gRNA vector. Two days after transfection, cells
909 were plated at clonal density and subjected to a transient puromycin selection (1 ug/mL) for 40 h. Colonies were picked 6
910 days after transfection. The triple PCR strategy used for screening is depicted in Supplementary Fig. 3a. Briefly, PCR

primers 1F and 4R were used to identify clones in which the *Dppa3* locus had been removed, resulting in the appearance of a ~350 bp amplicon. To identify whether the *Dppa3* locus had been removed from both alleles, PCRs were performed with primers 1F and 2R or 3F and 4R to amplify upstream or downstream ends of the *Dppa3* locus, which would only be left intact in the event of mono-allelic locus excision. Removal of the *Dppa3* locus was confirmed with Sanger sequencing and loss of *Dppa3* expression was assessed by qRT-PCR.

For CRISPR/Cas gene editing, all transfections were performed using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's instructions. All DNA oligos used for gene editing and screening are listed in Supplementary Table 5.

Bxb1-mediated recombination and Sleeping Beauty Transposition

To generate stable mESC lines carrying doxycycline inducible forms of *Dppa3* or *Dppa3-mScarlet*, mES cells were first transfected with equimolar amounts of the pSBtet-3xFLAG-Dppa3-IRES-DsRed-PuroR or pSBtet-3xFLAG-Dppa3-mScarlet-PuroR and the Sleeping Beauty transposase, pCMV(CAT)T7-SB100[118] (Addgene plasmid #34879) vector using Lipofectamine 3000 (Thermo Fisher Scientific) according to manufacturer's instructions. Two days after transfection, cells were plated at clonal density and subjected to puromycin selection (1 ug/mL) for 5-6 days. To ensure comparable levels of *Dppa3* induction, cells were first treated for 18 h with doxycycline (1 µg/mL) and then sorted with FACS based on thresholded levels of DsRed or mScarlet expression, the fluorescent readouts of successful induction. Post sorting, cells were plated back into media without doxycycline for 7 days before commencing experiments.

To generate stable doxycycline-inducible *Dppa3* hESC lines, hES cells were first transfected with equimolar amounts of the pSBtet-3xFLAG-Dppa3-IRES-DsRed-PuroR and Sleeping Beauty transposase pCMV(CAT)T7-SB100[118] (Addgene plasmid #34879) vector using using the P3 Primary Cell 4D-NucleofectorTM Kit (V4XP-3012 Lonza) and the 4D-Nucleofector™ Platform (Lonza), program CB-156. Two days after nucleofection, cells were subjected to puromycin selection (1 ug/mL) for subsequent two days, followed by an outgrowth phase of 4days. At this stage, cells were sorted with FACS based on thresholded levels of DsRed expression to obtain two bulk populations of positive stable hESC lines with inducible *Dppa3*.

For the generation of the *Uhrf1*$^{GFP/GFP}$ cell line, we used our previously described ESC line with a C-terminal MIN-tag (*Uhrf1*$^{attP/attP}$*;* Bxb1 *attP* site) and inserted the GFP coding sequence as described previously[113]. Briefly, attB-GFP-Stop-PolyA (Addgene plasmid #65526) was inserted into the C-terminal of the endogenous *Uhrf1*$^{attP/attP}$ locus by transfection with equimolar amounts of Bxb1 and attB-GFP-Stop-PolyA construct, followed by collection of GFP-positive cells with FACS after 6 days.

Cellular fractionation and Western Blot

Western blot for T1CM ESCs were performed as described previously[113] using monoclonal antibody rat anti-TET1 5D6 (1:10)[119] and polyclonal rabbit anti-H3 (1:5,000; ab1791, Abcam) as loading control. Blots were probed with secondary antibodies goat anti-rat (1:5,000; 112-035-068, Jackson ImmunoResearch) and goat anti-rabbit (1:5,000; 170–6515, Bio-Rad) conjugated to horseradish peroxidase (HRP) and visualized using an ECL detection kit (Thermo Scientific Pierce). Cell fractionation was performed as described previously with minor modifications[120]. 1×10^7 ESCs were resuspended in 250 µL of buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol, 0.1% Triton X-100, 1 mM DTT, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1x mammalian protease inhibitor cocktail (PI; Roche)) and incubated for 5 min on ice. Nuclei were collected by centrifugation (4 min, 1,300 × g, 4 °C) and the cytoplasmic fraction (supernatant) was cleared again by centrifugation (15 min, 20,000 × g, 4 °C). Nuclei were washed once with buffer A, and then lysed in

buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT, 1 mM PMSF, 1x PI). Insoluble chromatin was collected by centrifugation (4 min, 1,700 × g, 4 °C) and washed once with buffer B. Chromatin fraction was lysed with 1x Laemmli buffer and boiled (10 min, 95 °C).

As markers of cytoplasmic and chromatin fractions, alpha-tubulin and histone H3 were detected using monoclonal antibody (mouse anti-alpha-Tubulin, Sigma T9026 or rat anti-Tubulin, Abcam ab6160) and polyclonal antibody (rabbit anti-H3, Abcam ab1791). UHRF1 was visualized by rabbit anti-UHRF1 antibody[75]. Western blots for DNMT1 were performed as described previously using a monoclonal antibody (rat anti-DNMT1, 14F6) or a polyclonal antibody (rabbit anti-DNMT1, Abcam ab87654)[113]. GFP and FLAG tagged proteins were visualized by mouse anti-GFP (Roche) and anti-FLAG M2 antibodies (Sigma, F3165), respectively.

Quantitative real-time PCR (qRT-PCR) Analysis

Total RNA was isolated using the NucleoSpin Triprep Kit (Macherey-Nagel) according to the manufacturer's instructions. cDNA synthesis was performed with the High-Capacity cDNA Reverse Transcription Kit (with RNase Inhibitor; Applied Biosystems) using 500 ng of total RNA as input. qRT-PCR assays with oligonucleotides listed in Supplementary Table 5 were performed in 8 µL reactions with 1.5 ng of cDNA used as input. FastStart Universal SYBR Green Master Mix (Roche) was used for SYBR green detection. The reactions were run on a LightCycler480 (Roche).

LC-MS/MS analysis of DNA samples

Isolation of genomic DNA was performed according to earlier published work[43]. 1.0–5 µg of genomic DNA in 35 µL H_2O were digested as follows: An aqueous solution (7.5 µL) of 480 µM $ZnSO_4$, containing 18.4 U nuclease S1 (Aspergillus oryzae, Sigma-Aldrich), 5 U Antarctic phosphatase (New England BioLabs) and labeled internal standards were added ([$^{15}N_2$]-cadC 0.04301 pmol, [$^{15}N_2$,D_2]-hmdC 7.7 pmol, [D_3]-mdC 51.0 pmol, [$^{15}N_5$]-8-oxo-dG 0.109 pmol, [$^{15}N_2$]-fdC 0.04557 pmol) and the mixture was incubated at 37 °C for 3 h. After addition of 7.5 µl of a 520 µM $[Na]_2$-EDTA solution, containing 0.2 U snake venom phosphodiesterase I (Crotalus adamanteus, USB corporation), the sample was incubated for 3 h at 37 °C and then stored at –20 °C. Prior to LC/MS/MS analysis, samples were filtered by using an AcroPrep Advance 96 filter plate 0.2 µm Supor (Pall Life Sciences).

Quantitative UHPLC-MS/MS analysis of digested DNA samples was performed using an Agilent 1290 UHPLC system equipped with a UV detector and an Agilent 6490 triple quadrupole mass spectrometer. Natural nucleosides were quantified with the stable isotope dilution technique. An improved method, based on earlier published work [43,121] was developed, which allowed the concurrent analysis of all nucleosides in one single analytical run. The source-dependent parameters were as follows: gas temperature 80 °C, gas flow 15 L/min (N_2), nebulizer 30 psi, sheath gas heater 275 °C, sheath gas flow 15 L/min (N_2), capillary voltage 2,500 V in the positive ion mode, capillary voltage –2,250 V in the negative ion mode and nozzle voltage 500 V. The fragmentor voltage was 380 V/ 250 V. Delta EMV was set to 500 V for the positive mode. Chromatography was performed by a Poroshell 120 SB-C8 column (Agilent, 2.7 µm, 2.1 mm × 150 mm) at 35 °C using a gradient of water and MeCN, each containing 0.0085% (v/v) formic acid, at a flow rate of 0.35 mL/min: 0 →4 min; 0 →3.5% (v/v) MeCN; 4 →6.9 min; 3.5 →5% MeCN; 6.9 →7.2 min; 5 →80% MeCN; 7.2 →10.5 min; 80% MeCN; 10.5 →11.3 min; 80 →0% MeCN; 11.3 →14 min; 0% MeCN. The effluent up to 1.5 min and after 9 min was diverted to waste by a Valco valve. The autosampler was cooled to 4 °C. The injection volume was amounted to 39 µL. Data were processed according to earlier published work[43].

RNA-seq and Reduced Representation Bisulfite Sequencing (RRBS)

989 For RNA-seq, RNA was isolated using the NucleoSpin Triprep Kit (Machery-Nagel) according to the manufacturer's
990 instructions. Digital gene expression libraries for RNA-seq were produced using a modified version of single-cell RNA
991 barcoding sequencing (SCRB-seq) optimized to accommodate bulk cells[122] in which a total of 70 ng of input RNA was
992 used for the reverse-transcription of individual samples. For RRBS, genomic DNA was isolated using the QIAamp DNA
993 Mini Kit (QIAGEN), after an overnight lysis and proteinase K treatment. RRBS library preparation was performed as
994 described previously[123], with the following modifications: bisulfite treatment was performed using the EZ DNA
995 Methylation-Gold™ Kit (Zymo Research Corporation) according to the manufacturer's protocol except libraries were
996 eluted in 2 x 20 µL M-elution buffer. RNA-seq and RRBS libraries were sequenced on an Illumina HiSeq 1500.

997 **Targeted Bisulfite Amplicon (TaBA) Sequencing**

998 Genomic DNA was isolated from 10^6 cells using the PureLink Genomic DNA Mini Kit (Thermo Fisher Scientific) according
999 to the manufacturer's instructions. The EZ DNA Methylation-Gold Kit (Zymo Research) was used for bisulfite conversion
1000 according to the manufacturer's instructions with 500 ng of genomic DNA used as input and the modification that bisulfite
1001 converted DNA was eluted in 2 x 20 µL Elution Buffer.
1002 The sequences of the locus specific primers (Supplementary Table 5) were appended with Illumina TruSeq and Nextera
1003 compatible overhangs. The amplification of bisulfite converted DNA was performed in 25 µL PCR reaction volumes
1004 containing 0.4 µM each of forward and reverse primers, 2 mM Betaiinitialne (Sigma-Aldrich, B0300-1VL), 10 mM
1005 Tetramethylammonium chloride solution (Sigma-Aldrich T3411-500ML), 1x MyTaq Reaction Buffer, 0.5 units of MyTaq HS
1006 (Bioline, BIO-21112), and 1 µL of the eluted bisulfite converted DNA (~12.5 ng). The following cycling parameters were
1007 used: 5 min for 95 °C for initial denaturation and activation of the polymerase, 40 cycles (95 °C for 20 s, 58 °C for 30 s, 72
1008 °C for 25 s) and a final elongation at 72 °C for 3 min. Agarose gel electrophoresis was used to determine the quality and
1009 yield of the PCR.
1010 For purifying amplicon DNA, PCR reactions were incubated with 1.8x volume of CleanPCR beads (CleanNA, CPCR-0005)
1011 for 10 min. Beads were immobilized on a DynaMag™-96 Side Magnet (Thermo Fisher, 12331D) for 5 min, the supernatant
1012 was removed, and the beads washed 2x with 150 µL 70% ethanol. After air drying the beads for 5 min, DNA was eluted in
1013 15 µL of 10 mM Tris-HCl pH 8.0. Amplicon DNA concentration was determined using the Quant-iT™ PicoGreen™ dsDNA
1014 Assay Kit (Thermo Fisher, P7589) and then diluted to 0.7 ng/µL. Thereafter, indexing PCRs were performed in 25 µL PCR
1015 reaction volumes containing 0.08 µM (1 µL of a 2 µM stock) each of i5 and i7 Indexing Primers, 1x MyTaq Reaction Buffer,
1016 0.5 units of MyTaq HS (Bioline, BIO-21112), and 1 µL of the purified PCR product from the previous step. The following
1017 cycling parameters were used: 5 min for 95 °C for initial denaturation and activation of the polymerase, 40 cycles (95 °C
1018 for 10 s, 55 °C for 30 s, 72 °C for 40 s) and a final elongation at 72 °C for 5 min. Agarose gel electrophoresis was used to
1019 determine the quality and yield of the PCR. An aliquot from each indexing reaction (5 µL of each reaction) was then pooled
1020 and purified with CleanPCR magnetic beads as described above and eluted in 1 µL x Number of pooled reactions.
1021 Concentration of the final library was determined using PicoGreen and the quality and size distribution of the library was
1022 assessed with a Bioanalyzer. Dual indexed TaBA-seq libraries were sequenced on an Illumina MiSeq in 2x300 bp output
1023 mode.

1024 **RNA-seq processing and analysis**

1025 RNA-seq libraries were processed and mapped to the mouse genome (mm10) using the zUMIs pipeline[124]. UMI count
1026 tables were filtered for low counts using HTSFilter[125]. Differential expression analysis was performed in R using DESeq2[126]
1027 and genes with an adjusted $P < 0.05$ were considered to be differentially expressed. Hierarchical clustering was performed
1028 on genes differentially expressed in TET mutant ESCs respectively, using k-means clustering ($k = 4$) in combination with

the ComplexHeatmap R-package[127]. Principal component analysis was restricted to genes differentially expressed during wild-type differentiation and performed using all replicates of wild-type, TET mutant, and Dppa3KO ESCs.

RRBS alignment and analysis

Raw RRBS reads were first trimmed using Trim Galore (v.0.3.1) with the '--rrbs' parameter. Alignments were carried out to the mouse genome (mm10) using bsmap (v.2.90) using the parameters '-s 12 -v 10 -r 2 -l 1'. CpG-methylation calls were extracted from the mapping output using bsmaps methratio.py. Analysis was restricted to CpG with a coverage > 10. methylKit [128] was used to identify differentially methylated regions between the respective contrasts for the following genomic features: 1) all 1-kb tiles (containing a minimum of three CpGs) detected by RRBS; 2) Repeats (defined by Repbase); 3) gene promoters (defined as gene start sites –2kb/+2kb); and 4) gene bodies (defined as longest isoform per gene) and CpG islands (as defined by [129]). Differentially methylated regions were identified as regions with $P < 0.05$ and a difference in methylation means between two groups greater than 20%. Principal component analysis of global DNA methylation profiles was performed on single CpGs using all replicates of wild-type, T1KO and T1CM ESCs and EpiLCs.

Chromatin immunoprecipitation (ChIP) and Hydroxymethylated-DNA immunoprecipitation (hMeDIP) alignment and analysis.

ChIP–seq reads for TET1 binding in ESCs and EpiLCs were downloaded from GSE57700[45] and PRJEB19897[44], respectively. hMeDIP reads for wild-type ESCs and T1KO ESCs were download from PRJEB13096[44]. Reads were aligned to the mouse genome (mm10) with Bowtie (v.1.2.2) with parameters '-a -m 3 -n 3 --best --strata'. Subsequent ChIP–seq analysis was carried out on data of merged replicates. Peak calling and signal pile up was performed using MACS2 callpeak[130] with the parameters '--extsize 150' for ChIP, '--extsize 220' for hMeDIP, and '--nomodel -B --nolambda' for all samples. Tag densities for promoters and 1kb Tiles were calculated using the deepTools2 computeMatrix module[131]. TET1 bound genes were defined by harboring a TET1 peak in the promoter region (defined as gene start sites –2kb/+2kb).

Immunofluorescence staining

For immunostaining, naïve ESCs were grown on coverslips coated with Geltrex (Life Technologies) diluted 1:100 in DMEM/F12 (Life Technologies), thereby allowing better visualization of the cytoplasm during microscopic analysis. All steps during immunostaining were performed at room temperature. Coverslips were rinsed two times with PBS (pH 7.4; 140 mM NaCl, 2.7 mM KCl, 6.5 mM Na_2HPO_4, 1.5 mM KH_2PO_4) prewarmed to 37 °C, cells fixed for 10 min with 4% paraformaldehyde (pH 7.0; prepared from paraformaldehyde powder (Merck) by heating in PBS up to 60 °C; store at -20 °C), washed three times for 10 min with PBST (PBS, 0.01% Tween20), permeabilized for 5 min in PBS supplemented with 0.5% Triton X-100, and washed two times for 10 min with PBS. Primary and secondary antibodies were diluted in blocking solution (PBST, 4% BSA). Coverslips were incubated with primary and secondary antibody solutions in dark humid chambers for 1 h and washed three times for 10 min with PBST after primary and secondary antibodies. For DNA counterstaining, coverslips were incubated 6 min in PBST containing a final concentration of 2 µg/mL DAPI (Sigma-Aldrich) and washed three times for 10 min with PBST. Coverslips were mounted in antifade medium (Vectashield, Vector Laboratories) and sealed with colorless nail polish.

Following primary antibodies were used: polyclonal rabbit anti-DPPA3 (1:200; ab19878, Abcam) and monoclonal rabbit anti-UHRF1 (1:250; D6G8E, Cell Signaling). Following secondary antibodies were used: polyclonal donkey anti-rabbit conjugated to Alexa 488 (1:500; 711-547-003, Dianova), polyclonal donkey anti-rabbit conjugated to Alexa 555 (1:500; A31572, Invitrogen), and polyclonal donkey anti-rabbit conjugated to DyLight fluorophore 594 (1:500; 711-516-152, Dianova).

1068

Immunofluorescence and Live-cell imaging

1069 For immunofluorescence, stacks of optical sections were collected on a Nikon TiE microscope equipped with a Yokogawa
1070 CSU-W1 spinning disk confocal unit (50 µm pinhole size), an Andor Borealis illumination unit, Andor ALC600 laser beam
1071 combiner (405nm/488nm/561nm/640nm), Andor IXON 888 Ultra EMCCD camera, and a Nikon 100x/1.45 NA oil immersion
1072 objective. The microscope was controlled by software from Nikon (NIS Elements, ver. 5.02.00). DAPI or fluorophores were
1073 excited with 405 nm, 488 nm, or 561 nm laser lines and bright-field images acquired using Nikon differential interference
1074 contrast optics. Confocal image z-stacks were recorded with a step size of 200 nm, 16-bit image depth, 1 x 1 binning, a
1075 frame size of 1024 × 1024 pixels, and a pixel size of 130 nm. Within each experiment, cells were imaged using the same
1076 settings on the microscope (camera exposure time, laser power, and gain) to compare signal intensities between cell lines.
1077 For live-cell imaging, cells were plated on Geltrex-coated glass bottom 2-well imaging slides (Ibidi). Both still and
1078 timelapse images were acquired on the Nikon spinning disk system described above equipped with an environmental
1079 chamber maintained at 37 °C with 5% CO_2 (Oko Labs), using a Nikon 100x/1.45 NA oil immersion objective and a Perfect
1080 Focus System (Nikon). Images were acquired with the 488, 561, and 640 nm laser lines, full-frame (1024 x 1024) with 1 x 1
1081 binning, and with a pixel size of 130 nm. Transfection of a RFP-PCNA vector[132] was used to identify cells in S-phase. For
1082 DNA staining in live cells, cells were exposed to media containing 200 nM SiR-DNA (Spirochrome) for at least 1 h before
1083 imaging. For imaging endogenous DPPA3-HALO in live cells, cells were treated with media containing 50 nM
1084 HaloTag-TMR fluorescent ligand (Promega) for 1 h. After incubation, cells were washed 3x with PBS before adding back
1085 normal media. Nuclear export inhibition was carried out using media containing 20 nM leptomycin-B (Sigma-Aldrich).

1086

Image analysis

1087 For immunofluorescence images, Fiji software (ImageJ 1.51j)[133,134] was used to analyze images and create RGB stacks.
1088 For analysis of live-cell imaging data, CellProfiler Software (version 3.0)[135] was used to quantify fluorescence intensity in
1089 cells stained with SiR-DNA. CellProfiler pipelines used in this study are available upon request. In brief, the SiR-DNA signal
1090 was used to segment ESC nuclei. Mean fluorescence intensity of GFP was measured both inside the segmented area
1091 (nucleus) and in the area extending 4-5 pixels beyond the segmented nucleus (cytoplasm). GFP fluorescence intensity was
1092 normalized by subtracting the experimentally-determined mean background intensity and background-subtracted GFP
1093 intensities were then used for all subsequent quantifications shown in Fig. 4 and Supplementary Figs. 4h, 5h, and 6b,c.

1094

Fluorescence Recovery After Photobleaching (FRAP)

1095 For FRAP assays, cells cultivated on Geltrex-coated glass bottom 2-well imaging slides (Ibidi) were imaged in an
1096 environmental chamber maintained at 37 °C with 5% CO_2 either using the Nikon system mentioned above equipped with
1097 a FRAPPA photobleaching module (Andor) or on an Ultraview-Vox spinning disk system (Perkin-Elmer) including a FRAP
1098 Photokinesis device mounted to an inverted Axio Observer D1 microscope (Zeiss) equipped with an EMCCD camera
1099 (Hamamatsu) and a 63x/1.4 NA oil immersion objective, as well as 405, 488 and 561 nm laser lines.
1100 For endogenous UHRF1-GFP FRAP, eight pre-bleach images were acquired with the 488 nm laser, after which an area of 4
1101 x 4 pixels was irradiated for a total of 16 ms with a focused 488 nm laser (leading to a bleached spot of ~1 µm) to bleach a
1102 fraction of GFP-tagged molecules within cells, and then recovery images were acquired every 250 ms for 1-2 min.
1103 Recovery analysis was performed in Fiji. Briefly, fluorescence intensity at the bleached spot was measured in
1104 background-subtracted images, then normalized to pre-bleach intensity of the bleached spot, and normalized again to the
1105 total nuclear intensity in order to account for acquisition photobleaching. Images of cells with visible drift were discarded.
1106

Xenopus egg extracts

The interphase extracts (low-speed supernatants (LSS)) were prepared as described previously[66]. After thawing, LSS were supplemented with an energy regeneration system (5 μg/ml creatine kinase, 20 mM creatine phosphate, 2 mM ATP) and incubated with sperm nuclei at 3,000-4,000 nuclei per μl. Extracts were diluted 5-fold with ice-cold CPB (50 mM KCl, 2.5 mM MgCl2, 20 mM HEPES-KOH, pH 7.7) containing 2% sucrose, 0.1% NP-40 and 2 mM NEM, overlaid onto a 30% sucrose/CPB cushion, and centrifuged at 15,000g for 10min. The chromatin pellet was resuspended in SDS sample buffer and analyzed by SDS-PAGE. GST-mDPPA3 was added to egg extracts at 50 ng/μl at final concentration.

Monitoring DNA methylation in *Xenopus* egg extracts

DNA methylation was monitored by the incorporation of S-[methyl-3H]-adenosyl-L-methionine, incubated at room temperature, and the reaction was stopped by the addition of CPB containing 2% sucrose up to 300 μl. Genomic DNA was purified using a Wizard Genomic DNA purification kit (Promega) according to the manufacturer's instructions. Incorporation of radioactivity was quantified by liquid synchillation counter.

Plasmid construction for Recombinant mDPPA3

To generate GST-tagged mDPPA3 expression plasmids, mDPPA3 fragment corresponding to full-length protein was amplified by PCR using mouse DPPA3 cDNA and primers
5'- TTAGCAGCCGGATCCCTAATTCTTCCCGATTTTCGCA-3' and
5'-CGTGGATCCCCGAATTCCATGGAGGAACCATCAGAGAAAGTC'. The resulting DNA fragment was cloned into pGEX4T-3 vector digested with EcoRI and SalI using an In-Fusion HD Cloning Kit.

Protein expression and purification

For protein expression in Escherichia coli (BL21-CodonPlus), the mDPPA3 genes were transferred to pGEX4T-3 vector as described above. Protein expression was induced by the addition of 0.1 mM Isopropyl β–D-1-thiogalactopyranoside (IPTG) to media followed by incubation for 12 hour at 20˚C. For purification of Glutathione S transferase (GST) tagged proteins, cells were collected and re-suspended in Lysis buffer (20 mM HEPES-KOH (pH 7.6), 0.5M NaCl, 0.5 mM EDTA, 10% glycerol, 1 mM DTT) supplemented with 0.5% NP40 and protease inhibitors, and were then disrupted by sonication on ice. After centrifugation, the supernatant was applied to Glutathione Sepharose (GSH) beads (GE Healthcare) and rotated at 4 ˚C for 2 hour. Beads were then washed three times with Wash buffer 1 (20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1% TritionX-100, 1 mM DTT) three times and with Wash buffer 2 (100 mM Tris-HCl (pH 7.5), 100 mM NaCl) once. Bound proteins were eluted in Elution buffer (100 mM Tris-HCl (pH 7.5), 100 mM NaCl, 5% glycerol, 1 mM DTT) containing 42 mM reduced Glutathione and purified protein was loaded on PD10 desalting column equilibrated with EB buffer (10 mM HEPES/KOH at pH 7.7, 100 mM KCl, 0.1 mM CaCl2, 1 mM MgCl2) containing 1 mM DTT, and then concentrated by Vivaspin (Millipore).

Data collection for the presence of TET1, UHRF1, DNMT1 and DPPA3 throughout metazoa

Megabat (*Pteropus vampyrus*) protein sequences for TET1 (ENSVPAP00000010999), UHRF1 (ENSPVAP00000002809), DNMT1 (ENSPVAP00000003477) and DPPA3 (ENSVPAP00000004109) were subjected to three iterations of PSI-blast [136], respectively. Subsequently, result tables were filtered to contain correct gene names using a custom python script. Presence of the proteins throughout metazoa was visualized using iTOL[137].

Chromatin immunoprecipitation coupled to Mass Spectrometry and Proteomics data analysis

For Chromatin immunoprecipitation coupled to Mass Spectrometry (ChIP-MS), whole cell lysates of the doxycycline inducible *Dppa3*-FLAG mES cells were used by performing three separate immunoprecipitations with an anti-FLAG

antibody and three samples with a control IgG. Proteins were digested on the beads after the pulldown and desalted subsequently on StageTips with three layers of C18[138]. Here, peptides were separated by liquid chromatography on an Easy-nLC 1200 (Thermo Fisher Scientific) on in-house packed 50 cm columns of ReproSil-Pur C18-AQ 1.9-µm resin (Dr. Maisch GmbH). Peptides were then eluted successively in an ACN gradient for 120 min at a flow rate of around 300 nL/min and were injected through a nanoelectrospray source into a Q Exactive HF-X Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher Scientific). After measuring triplicates of a certain condition, an additional washing step was scheduled. During the measurements, the column temperature was constantly kept at 60 °C while after each measurement, the column was washed with 95% buffer B and subsequently with buffer A. Real time monitoring of the operational parameters was established by SprayQc[139] software. Data acquisition was based on a top10 shotgun proteomics method and data-dependent MS/MS scans. Within a range of 400-1650 m/z and a max. injection time of 20 ms, the target value for the full scan MS spectra was 3×10^6 and the resolution at 60,000.

The raw MS data was then analyzed with the MaxQuant software package (version 1.6.0.7)[140]. The underlying FASTA files for peak list searches were derived from Uniprot (UP000000589_10090.fasta and UP000000589_10090 additional.fasta, version June 2015) and an additional modified FASTA file for the FLAG-tagged *Dppa3* in combination with a contaminants database provided by the Andromeda search engine[141] with 245 entries. During the MaxQuant-based analysis the "Match between runs" option was enabled and the false discovery rate was set to 1% for both peptides (minimum length of 7 amino acids) and proteins. Relative protein amounts were determined by the MaxLFQ algorithm[142], with a minimum ratio count of two peptides.

For the downstream analysis of the MaxQuant output, the software Perseus[143] (version 1.6.0.9) was used to perform Student's *t*-test with a permutation-based FDR of 0.05 and an additional constant S0 = 1 in order to calculate fold enrichments of proteins between triplicate chromatin immunoprecipitations of anti-FLAG antibody and control IgG. The result was visualized in a scatter plot.

For GO analysis of biological processes the Panther classification system was used[143]. For the analysis, 131 interactors of DPPA3 were considered after filtering the whole amount of 303 significant interactors for a p-value of at least 0.0015 and 3 or more identified peptides. The resulting GO groups were additionally filtered for a fold enrichment of observed over expected amounts of proteins of at least 4 and a p-value of 5.30 E-08.

Dppa3 overexpression in medaka embryos and immunostaining

Medaka d-rR strain was used. Medaka fish were maintained and raised according to standard protocols. Developmental stages were determined based on a previous study[144]. *Dppa3* and mutant *Dppa3* (R107E) mRNA were synthesized using HiScribe T7 ARCA mRNA kit (NEB, E2060S), and purified using RNeasy mini kit (QIAGEN, 74104). *Dppa3* or mutant *Dppa3* (R107E) mRNA was injected into the one-cell stage (stage 2) medaka embryos. After 7 hours of incubation at 28 ˚C, the late blastula (stage 11) embryos were fixed with 4% PFA in PBS for 2 hours at room temperature, and then at 4 ˚C overnight. Embryos were dechorionated, washed with PBS, and permeabilized with 0.5% Triton X-100 in PBS for 30 minutes at room temperature. DNA was denatured in 4 M HCl for 15 minutes at room temperature, followed by neutralization in 100 mM Tris-HCl (pH 8.0) for 20 minutes. After washing with PBS, embryos were blocked in blocking solution (2% BSA, 1%DMSO, 0.2% Triton X-100 in PBS) for 1 hour at room temperature, and then incubated with 5-methylcytosine antibody (1:200; Active Motif #39649) at 4 °C overnight. The embryos were washed with PBSDT (1% DMSO, 0.1% Triton X-100 in PBS), blocked in blocking solution for 1 hours at room temperature, and incubated with Alexa Fluor 555 goat anti-mouse 2nd antibody (1:500; ThermoFisher Scientific #A21422) at 4 °C overnight. After washing with PBSDT, cells were mounted on slides and examined under a fluorescence microscope.

1186

Fluorescence Three Hybrid (F3H) assay

1187 The F3H assay was performed as described previously[82]. In brief, BHK cells containing multiple lac operator repeats were
1188 transiently transfected with the respective GFP- and dsRed-constructs on coverslips using PEI and fixed with 3.0%
1189 formaldehyde 24 hrs after transfection. For DNA counterstaining, coverslips were incubated in a solution of DAPI (200
1190 ng/ml) in PBS-T and mounted in Vectashield. Images were collected using a Leica TCS SP5 confocal microscope. To
1191 quantify the interactions within the lac spot, the following intensity ratio was calculated for each cell in order to account for
1192 different expression levels: $(mCherry_{spot}–mCherry_{background})/(GFP_{spot}–GFP_{background})$.

1193

Microscale Thermophoresis (MST)

1194 For MST measurements, mUHRF1 C-terminally tagged with GFP- and 6xHis-tag was expressed in HEK 293T cells and
1195 then purified using Qiagen Ni-NTA beads (Qiagen #30230). Recombinant mDPPA3 WT and 1-60 were purified as
1196 described above. Purified UHRF1 (200 nM) was mixed with different concentrations of purified DPPA3 (0.15 nM to 5 μM)
1197 followed by a 30 min incubation on ice. The samples were then aspirated into NT.115 Standard Treated Capillaries
1198 (NanoTemper Technologies) and placed into the Monolith NT.115 instrument (NanoTemper Technologies). Experiments
1199 were conducted with 80% LED and 80% MST power. Obtained fluorescence signals were normalized (F_{norm}) and the
1200 change in F_{norm} was plotted as a function of the concentration of the titrated binding partner using the MO. Affinity Analysis
1201 software version 2.1 (NanoTemper Technologies). For fluorescence normalization ($F_{norm} = F_{hot}/F_{cold}$), the manual analysis
1202 mode was selected and cursors were set as follows: F_{cold} = -1 to 0 s, F_{hot} = 10 to 15 s. The Kd was obtained by fitting the
1203 mean F_{norm} of eight data points (four independent replicates, each measured as a technical duplicate).

1204

RICS

1205 Data for Raster Image Correlation Spectroscopy (RICS) was acquired on a home-built laser scanning confocal setup
1206 equipped with a pulsed interleaved excitation (PIE) system as used elsewhere[145]. Samples were excited using pulsed
1207 lasers at 470 (Picoquant) and 561 nm (Toptica Photonics), synchronized to a master clock, and then delayed ~20ns
1208 relative to one another to achieve PIE. Laser excitation was separated from descanned fluorescence emission by a
1209 Di01-R405/488/561/635 polychroic mirror (Semrock, AHF Analysentechnik) and eGFP and mScarlet fluorescence emission
1210 was separated by a 565 DCXR dichroic mirror (AHF Analysentechnik) and collected on avalanche photodiodes, a Count
1211 Blue (Laser Components) and a SPCM-AQR-14 (Perkin-Elmer) with 520/40 and a 630/75 emission filters (Chroma, AHF
1212 Analysentechnik). Detected photons were recorded by time-correlated single-photon counting.

1213 The alignment of the system was verified prior to each measurement session by performing FCS with PIE of a mixture of
1214 Atto-488 and Atto565 dyes, excited with pulsed 470 and 561 nm lasers set to 10 μW (measured in the collimated space
1215 before entering the galvo-scanning mirror system), 1 μm above the surface of the coverslip[146]. Cells were plated on Ibidi
1216 two-well glass bottom slides, and induced with doxycycline overnight prior to measurements. Scanning was performed in
1217 cells at maintained at 37 ˚C with a stage top incubator, with a total field of view of 12 x 12 μm, composed of 300 pixels x
1218 300 lines a pixel size of 40 nm, a pixel dwell time of 11 μs, a line time 3.33 ms, at one frame per second, for 100-200
1219 seconds. Pulsed 470 and 561 nm lasers were adjusted to 4 and 5 μW respectively.

1220 Image analysis was done using the Pulsed Interleaved Excitation Analysis with Matlab (PAM) software[147]. Briefly, time
1221 gating of the raw photon stream was performed by selecting only photons emitted on the appropriate detector after
1222 pulsed excitation, thereby allowing to generate cross-talk free images for each channel. Then, using Microtime Image
1223 Analysis (MIA), slow fluctuations were removed by subtracting a moving average of 3 frames, and a region of interest

corresponding to the nucleus was selected, excluding nucleoli and dense aggregates. The spatial autocorrelation and cross-correlation functions (SACF and SCCF) were calculated as done previously[148] using arbitrary region RICS:

$$G(\xi,\psi) = \frac{\langle I_{RICS,1}(x,y) I_{RICS,2}(x+\xi, y+\psi)\rangle_{XY}}{\langle I_{RICS,1}\rangle_{XY}\langle I_{RICS,2}\rangle_{XY}}$$

where ξ and ψ are the correlation lags in pixel units along the x- and y-axis scan directions. The correlation function was then fitted to a two-component model (one mobile and one immobile component) in MIAfit:

$$G_{fit}(\xi,\psi) = A_{mob} G_{fit,\,mob}(\xi,\psi) + A_{imm} \exp\left(-\delta r^2 \omega_{imm}^{-2}\left(\xi^2+\psi^2\right)\right) + y_0 ,$$

where

$$G_{fit,\,mob}(\xi,\psi) = \left(1+\frac{4D(\tau_p\xi+\tau_l\psi)}{\omega^2{}_r}\right)^{-1}\left(1+\frac{4D(\tau_p\xi+\tau_l\psi)}{\omega^2{}_z}\right)^{-1/2}\cdot\exp\left(-\frac{\delta r^2(\xi^2+\psi^2)}{\omega^2{}_r+4D(\tau_p\xi+\tau_l\psi)}\right) ,$$

which yields parameters such as the diffusion coefficient (D) and the amplitudes of the mobile and immobile fractions (A_{mob} and A_{imm}). The average number of mobile molecules per excitation volume on the RICS timescale was determined by:

$$N_{mob} = \frac{\gamma}{A_{mob}}\frac{2\Delta F}{2\Delta F+1}$$

where γ is a factor pertaining to the 3D Gaussian shape of the PSF, and 2ΔF/(2ΔF+1) is a correction factor when using a moving average subtraction prior to calculating the SACF. The bound fraction is the contribution of particles which remain visible during the acquisition of 5-10 lines of the raster scan, corresponding to ~30 ms. The cross-correlation model was fitted to the cross-correlation function, and the extent of cross-correlation was calculated from the amplitude of the mobile fraction of the cross-correlation fit divided by the amplitude of the mobile fraction of the autocorrelation fit of DPPA3-mScarlet.

Data and Code Availability

Sequencing data reported in this paper are available at ArrayExpress (EMBL-EBI) under accessions E-MTAB-6785 (wild-type and Tet catalytic mutants RRBS), E-MTAB-6797 (RNA-seq), E-MTAB-6800 (Dppa3KO RRBS). The raw mass spectrometry proteomics data from the FLAG-DPPA3 pulldown will be deposited at the ProteomeXchange Consortium via the PRIDE partner repository. Significant interactors of FLAG-DPPA3 in ESCs can be found in Supplementary Table 3.

Statistics and reproducibility

No statistical methods were used to predetermine sample size, the experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment. Blinding was not implemented in this study as analysis was inherently objective in the overwhelming majority of experiments. For microscopy analysis,

where possible, experimenter bias was avoided by selecting fields of view (or individual cells) for acquisition of UHRF1-GFP or DNMT1-GFP signal using the DNA stain (or another marker not being directly assessed in the experiment e.g. DsRed/mScarlet as a readout of Dppa3 induction or RFP-PCNA). To further reduce bias, imaging analysis was subsequently performed indiscriminately on all acquired images using semi-automated analysis pipelines (either with CellProfiler or Fiji scripts). All the experimental findings were reliably reproduced in independent experiments as indicated in the Figure legends. The number of replicates used in each experiment are described in the Figure legends and/or in the Methods section, as are the Statistical tests used. P values values or adjusted P values are given where possible. Unless otherwise indicated, all statistical calculations were performed using R Studio 1.2.1335. Next-generation sequencing experiments include at least two independent biological replicates. RNA-seq experiments include n= 4 biological replicates comprised of n=2 independently cultured samples from 2 clones (for T1CM, T2CM, T12CM ESCs and EpiLCs) or 4 independently cultured samples (for wild-type ESCs and EpiLCs). For RRBS experiments, data are derived from n= 2 biological replicates. For bisulfite sequencing of LINE-1 elements n= 2 biological replicates were analyzed from 2 independent clones for T1CM, T2CM, T12CM, and Dppa3KO ESCs or 2 independent cultures for wt ESCs. LC-MS/MS quantification was performed on at least 4 biological replicates comprising at least 2 independently cultured samples (usually even more) from n = 2 independent clones (T1CM, T2CM, T12CM, and Dppa3KO ESCs) or 4 independently cultured samples (wild-type ESCs and cell lines shown in Fig. 5d).

Acknowledgements:

We would like to thank Dr. H. Blum, Dr. S. Krebs (Laboratory for Functional Genome Analysis, LMU Munich), and Dr. A. Brachmann (Faculty of Biology, Department of Genetics, LMU Munich) for next generation sequencing services. We further thank Dr. Ralf Heermann (Bioanalytics core facility, LMU Munich) for help with the MST measurements. We thank V. Laban, C. Kirschner, T.J. Fischer, D. Nixdorf and V.B. Ozan for help with experiments, J. Koch for technical assistance, Dr. M. Mann and Dr. I. Paron for mass spectrometry and Dr. I. Hellmann for advice on data analysis and providing computational infrastructure. We would like to thank Dr. S. Stricker and Dr. I. Solovei for helpful discussions and constructive criticism on the manuscript and E. Ntouliou for critical reading of the manuscript. We thank Dr. Feng Zhang for providing the pSpCas9(BB)-2A-Puro (PX459) and SpCas9-T2A-GFP (PX458) (Addgene plasmids #62988 and #48138) plasmids, Dr. Eric Kowarz for the gift of the pSB-tetPur (Addgene plasmid #60507) construct, and Dr. Zsuzsanna Izsvak for the gift of the pCMV(CAT)T7-SB100 (Addgene plasmid #34879). J.R. and M.D.B. are fellows of the International Max Planck Research School for Molecular Life Sciences (IMPRS-LS). C.B.M gratefully acknowledges the support of the Fulbright Commission and the late Dr. Glenn Cuomo. F.R.T. thanks the Boehringer Ingelheim Fonds and J.R. the Fonds de Recherche du Québec en Santé for a PhD fellowship.

Funding:

The work was funded by the Deutsche Forschungsgemeinschaft (DFG grants SFB1064/A22 to S.B, SFB1064/A17 and SFB1243/A01 to H.L., and SFB1243/A14 to W.E.).

Authors Contributions:

C.B.M. and S.B. designed and conceived the study. S.B. and H.L. supervised the study. C.M., S.B., and H.L. prepared the manuscript with the help of M.D.B. C.B.M. performed cellular and molecular experiments. C.B.M. generated cell lines with help from M.Y.. C.B.M. performed RRBS and RNA-Seq with help and supervision from C.Z., S.B., and W.E.. J.R. and C.B.M. performed live-cell microscopy and photobleaching analyses. I.G., J.R. and C.B.M. performed RICS experiments under the supervision of D.L.. C.T. performed MST and F3H assays. A.N. performed *Xenopus* experiments under the supervision of M.N.. R.N. performed the experiments in medaka embryos under the supervision of H.T.. M.D.B. and P.S. helped with cell line validation and performed fluorescence microscopy analysis. W.Q. performed the biochemical analyses with assistance from A.A. M.M performed hESC experiments. E.U. conducted proteomics experiments and analyses under the guidance of M.W. F.R.T. and E.P. quantified modified cytosines by LC-MS/MS with the supervision by T.C. S.B. performed data analysis. All authors read, discussed, and approved the manuscript.

Competing interests:

The authors declare no competing interests.

SUPPLEMENTARY FIGURES for:

Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive demethylation in mammals

Authors:

Christopher B. Mulholland[1], Atsuya Nishiyama[2†], Joel Ryan[1†], Ryohei Nakamura[3], Merve Yiğit[1], Ivo M. Glück[4], Carina Trummer[1], Weihua Qin[1], Michael D. Bartoschek[1], Franziska R. Traube[5], Edris Parsa[5], Enes Ugur[1,8], Miha Modic[7], Aishwarya Acharya[1], Paul Stolz[1], Christoph Ziegenhain[6], Michael Wierer[8], Wolfgang Enard[6], Thomas Carell[5], Don C. Lamb[4], Hiroyuki Takeda[3], Makoto Nakanashi[2], Sebastian Bultmann[*1§] and Heinrich Leonhardt[*1§]

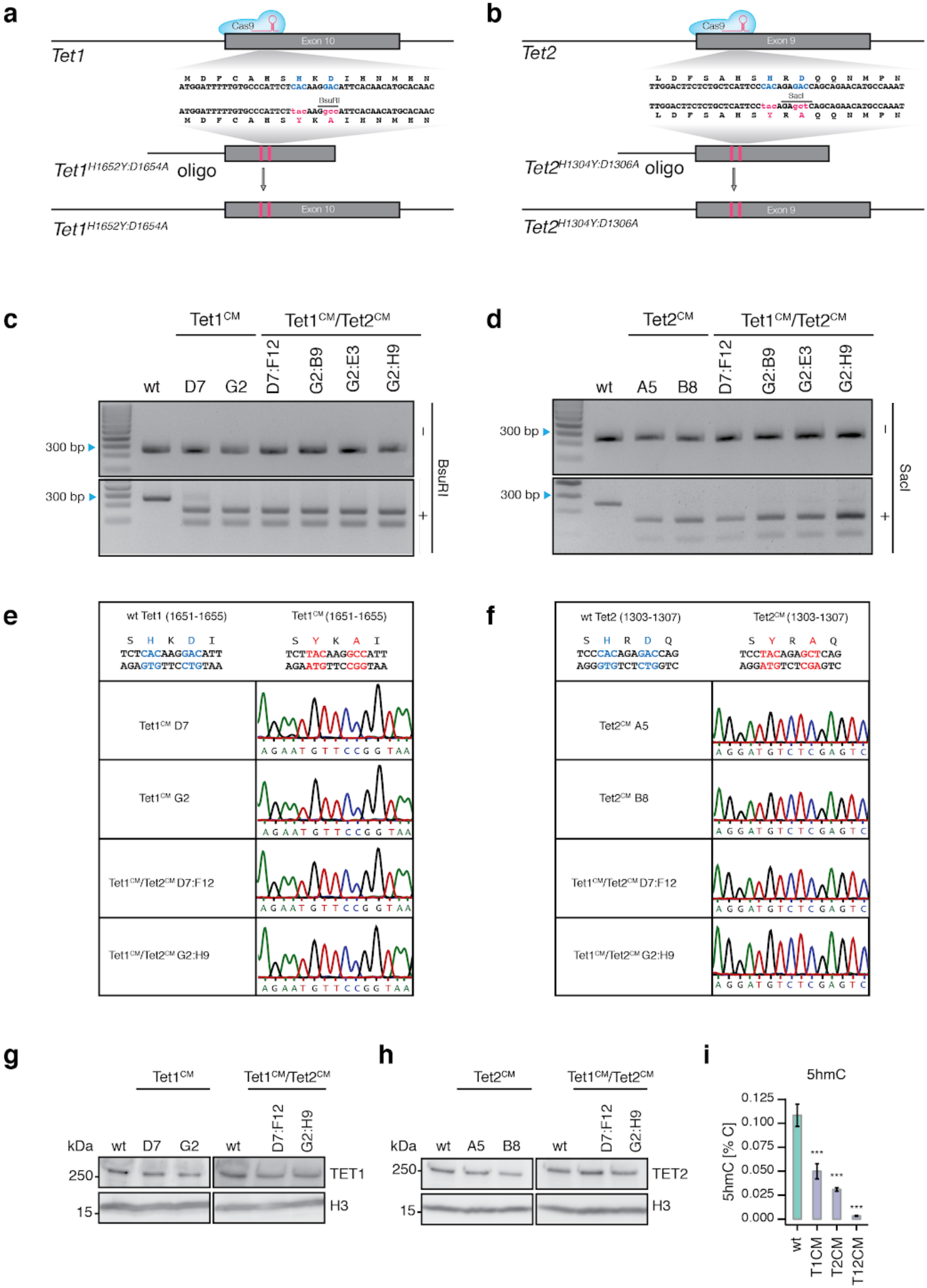
Correspondence: bultmann@bio.lmu.de or h.leonhardt@lmu.de

This PDF includes:

Supplementary Fig. 1
Supplementary Fig. 2
Supplementary Fig. 3
Supplementary Fig. 4
Supplementary Fig. 5
Supplementary Fig. 6
Supplementary Fig. 7
Supplementary Fig. 8

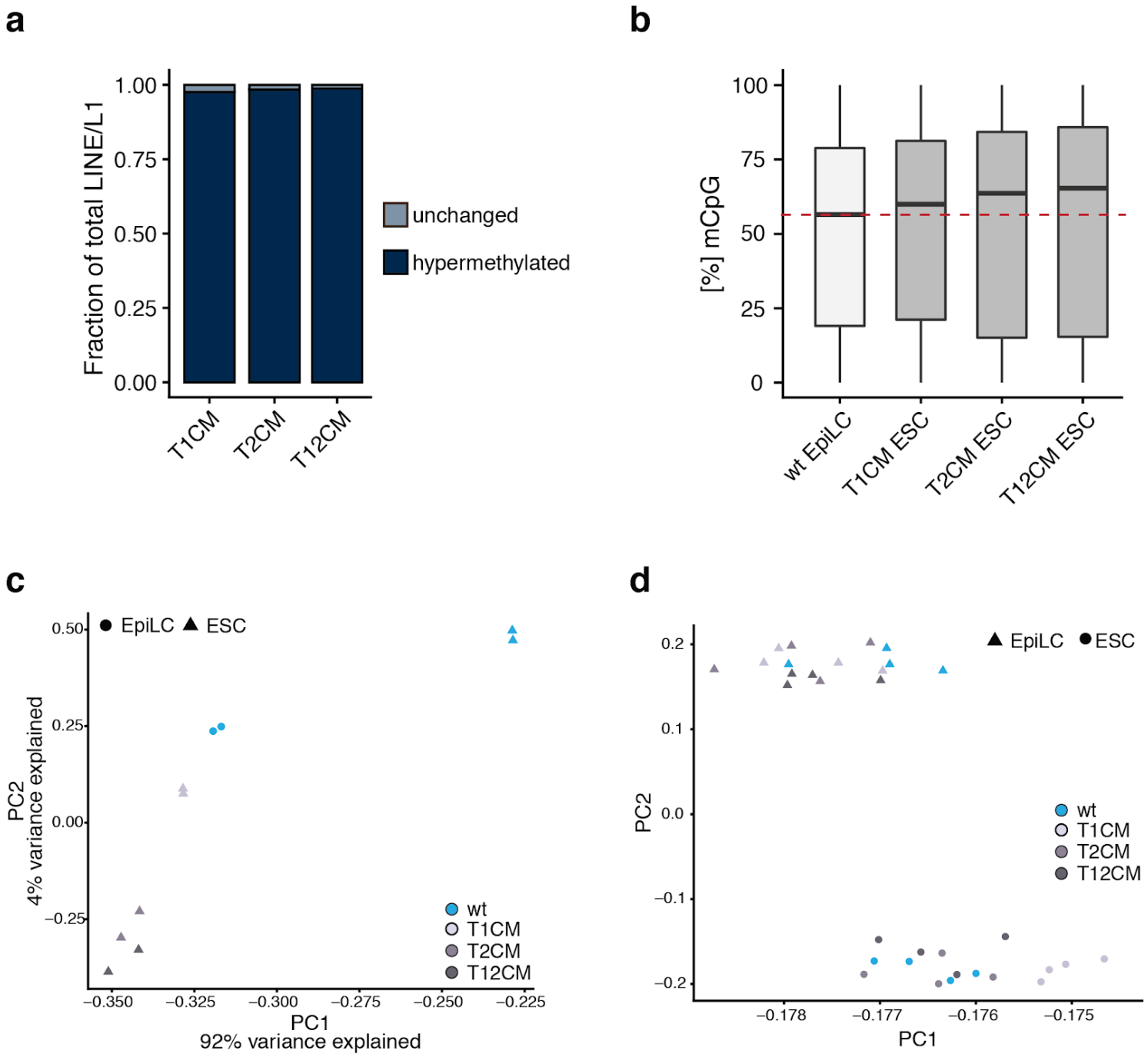
Other Supplementary Information for this manuscript not in this PDF:

Supplementary Table 1
Supplementary Table 2
Supplementary Table 3
Supplementary Table 4
Supplementary Table 5


a
Cas9
Exon 10
Tet1
M D F C A H S H K D I H N M H N
ATGGATTTTTGTGCCCATTCTCACAAGGACATTCACAACATGCACAAC
BsuFI
ATGGATTTTTGTGCCCATTCTtacAAGgccATTCACAACATGCACAAC
M D F C A H S Y K A I H N M H N
Tet1H1652Y:D1654A oligo
Exon 10
Tet1H1652Y:D1654A
b
Cas9
Exon 9
Tet2
L D F S A H S H R D Q Q N M P N
TTGGACTTCTCTGCTCATTCCCACAGAGACCAGCAGAACATGCCAAAT
SacI
TTGGACTTCTCTGCTCATTCCtacAGAgctCAGCAGAACATGCCAAAT
L D F S A H S Y R A Q Q N M P N
Tet2H1304Y:D1306A oligo
Exon 9
Tet2H1304Y:D1306A
c
Tet1CM
Tet1CM/Tet2CM
wt D7 G2 D7:F12 G2:B9 G2:E3 G2:H9
300 bp
300 bp
BsuRI
d
Tet2CM
Tet1CM/Tet2CM
wt A5 B8 D7:F12 G2:B9 G2:E3 G2:H9
300 bp
300 bp
SacI
e
wt Tet1 (1651-1655)
S H K D I
TCTCACAAGGACATT
AGAGTGTTCCTGTAA
Tet1CM (1651-1655)
S Y K A I
TCTTACAAGGCCATT
AGAATGTTCCGGTAA
Tet1CM D7
Tet1CM G2
Tet1CM/Tet2CM D7:F12
Tet1CM/Tet2CM G2:H9
A G A A T G T T C C G G T A A
f
wt Tet2 (1303-1307)
S H R D Q
TCCCACAGAGACCAG
AGGGTGTCTCTGGTC
Tet2CM (1303-1307)
S Y R A Q
TCCTACAGAGCTCAG
AGGATGTCTCGAGTC
Tet2CM A5
Tet2CM B8
Tet1CM/Tet2CM D7:F12
Tet1CM/Tet2CM G2:H9
A G G A T G T C T C G A G T C
g
Tet1CM
Tet1CM/Tet2CM
kDa wt D7 G2 wt D7:F12 G2:H9
250
15
TET1
H3
h
Tet2CM
Tet1CM/Tet2CM
kDa wt A5 B8 wt D7:F12 G2:H9
250
15
TET2
H3
i
5hmC
5hmC [% C]
0.125
0.100
0.075
0.050
0.025
0.000
***
***
***
wt
T1CM
T2CM
T12CM

Supplementary Figure 1: Generation and characterization of T1CM, T2CM, and T12CM mESCs

a,b, Schematic representation of CRISPR/Cas9 gene editing strategy used to mutate the catalytic center (HxD) of *Tet1* (**a**) and *Tet2* (**b**). gRNA target sequences and restriction enzyme recognition sites for restriction fragment length polymorphism (RFLP) screening are shown (See Supplementary Table 5). **c,d,** Genotyping using RFLP analysis of the *Tet1* (**c**) and *Tet2* (**d**) locus. **e,f,** Sanger sequencing results confirming the successful insertion of point mutations, YxA, at the *Tet1* (**g**) and *Tet2* (**h**) locus, HxD. **g,h,** Immunoblot detection of endogenous TET1 (**g**) and TET2 (**h**) protein levels in T1CM, T2CM, and T12CM mESCs. **i,** DNA modification levels as percentage of total cytosines in wt (n = 24), T1CM (n = 8), T2CM (n = 12), and T12CM (n = 11) mESCs. Depicted are mean values ± standard deviation; *P < 0.05, **P < 0.01, ***P < 0.005 to wt as determined using a one-way ANOVA followed by a post-hoc Tukey HSD test.

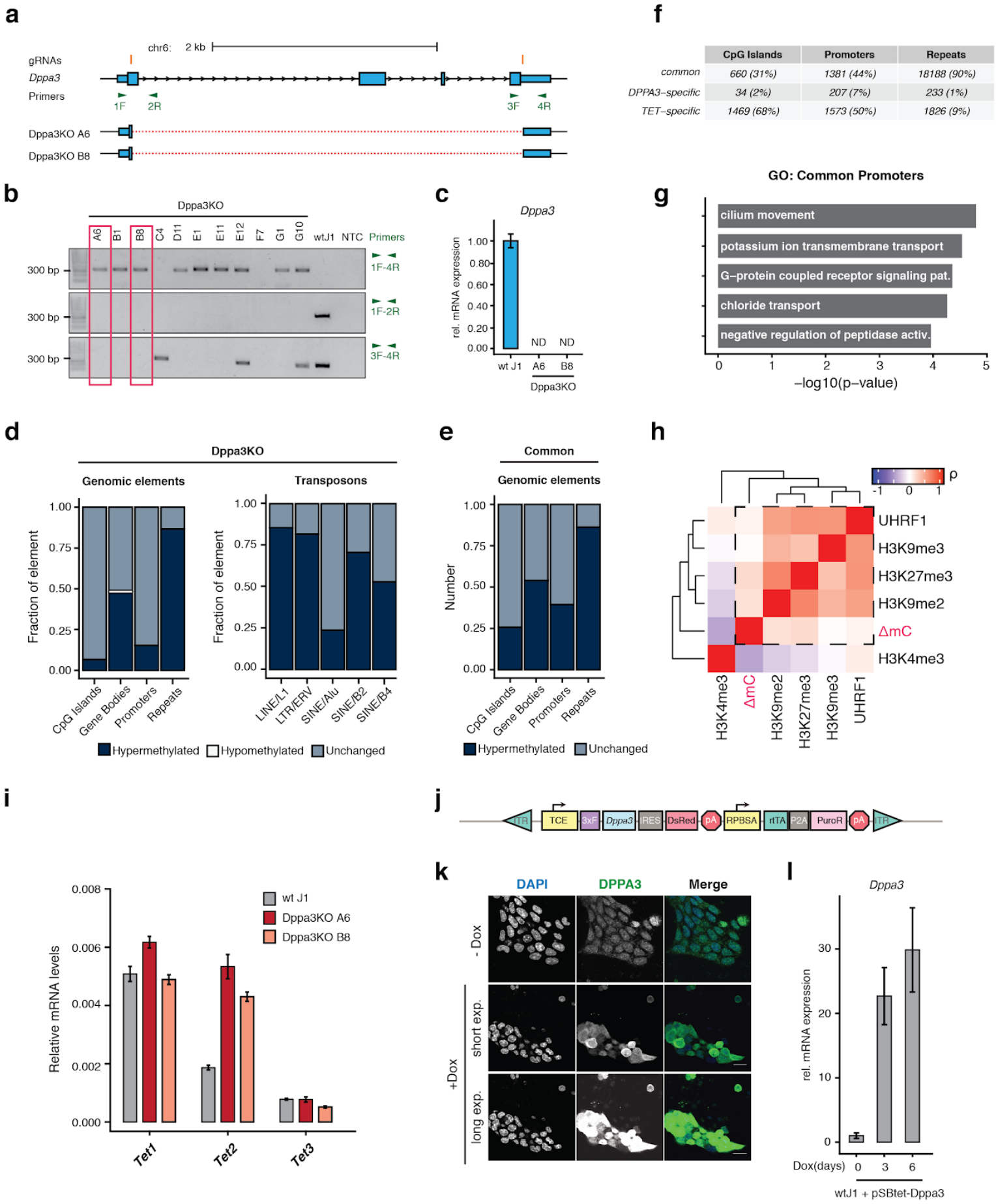


Supplementary Figure 2: Methylome and transcriptome analysis of T1CM, T2CM, and T12CM ESCs

a, Relative proportion of DNA hypermethylation (q value < 0.05; absolute methylation difference > 20%) at LINE1/L1 elements in T1CM, T2CM, and T12CM ESCs compared to wt ESCs. **b,** Loss of TET catalytic activity in ESCs results in similar or higher DNA methylation levels than in wt EpiLCs. Percentage of total CpG DNA methylation (5mCpG) as measured by RRBS. Horizontal black lines within boxes represent median values, boxes indicate the upper and lower quartiles, and whiskers indicate the 1.5 interquartile rangeDashed red line indicates the median mCpG methylation in wt EpiLCs. **c,** Principal component (PC) analysis of RRBS data from wt, T1CM, T2CM, and T12CM ESCs and wt EpiLCs. **d,** PC analysis of RNA-seq data from wt, T1CM, T2CM, and T12CM ESCs during EpiLC differentiation.

1337

1338


a
chr6: 2 kb
gRNAs
Dppa3
Primers
1F
2R
3F
4R
Dppa3KO A6
Dppa3KO B8
b
Dppa3KO
A6
B1
B8
C4
D11
E1
E11
E12
F7
G1
G10
wtJ1
NTC
Primers
300 bp
1F-4R
1F-2R
3F-4R
c
Dppa3
rel. mRNA expression
ND
ND
wt J1
A6
B8
Dppa3KO
f
CpG Islands
Promoters
Repeats
common
660 (31%)
1381 (44%)
18188 (90%)
DPPA3–specific
34 (2%)
207 (7%)
233 (1%)
TET–specific
1469 (68%)
1573 (50%)
1826 (9%)
g
GO: Common Promoters
cilium movement
potassium ion transmembrane transport
G–protein coupled receptor signaling pat.
chloride transport
negative regulation of peptidase activ.
–log10(p–value)
d
Dppa3KO
Genomic elements
Transposons
Fraction of element
CpG Islands
Gene Bodies
Promoters
Repeats
LINE/L1
LTR/ERV
SINE/Alu
SINE/B2
SINE/B4
Hypermethylated
Hypomethylated
Unchanged
e
Common
Genomic elements
Number
CpG Islands
Gene Bodies
Promoters
Repeats
Hypermethylated
Unchanged
h
ρ
UHRF1
H3K9me3
H3K27me3
H3K9me2
ΔmC
H3K4me3
i
Relative mRNA levels
wt J1
Dppa3KO A6
Dppa3KO B8
Tet1
Tet2
Tet3
j
ITR
TCE
3xF
Dppa3
IRES
DsRed
pA
RPBSA
rtTA
P2A
PuroR
k
DAPI
DPPA3
Merge
- Dox
+Dox
short exp.
long exp.
l
Dppa3
rel. mRNA expression
Dox(days)
wtJ1 + pSBtet-Dppa3

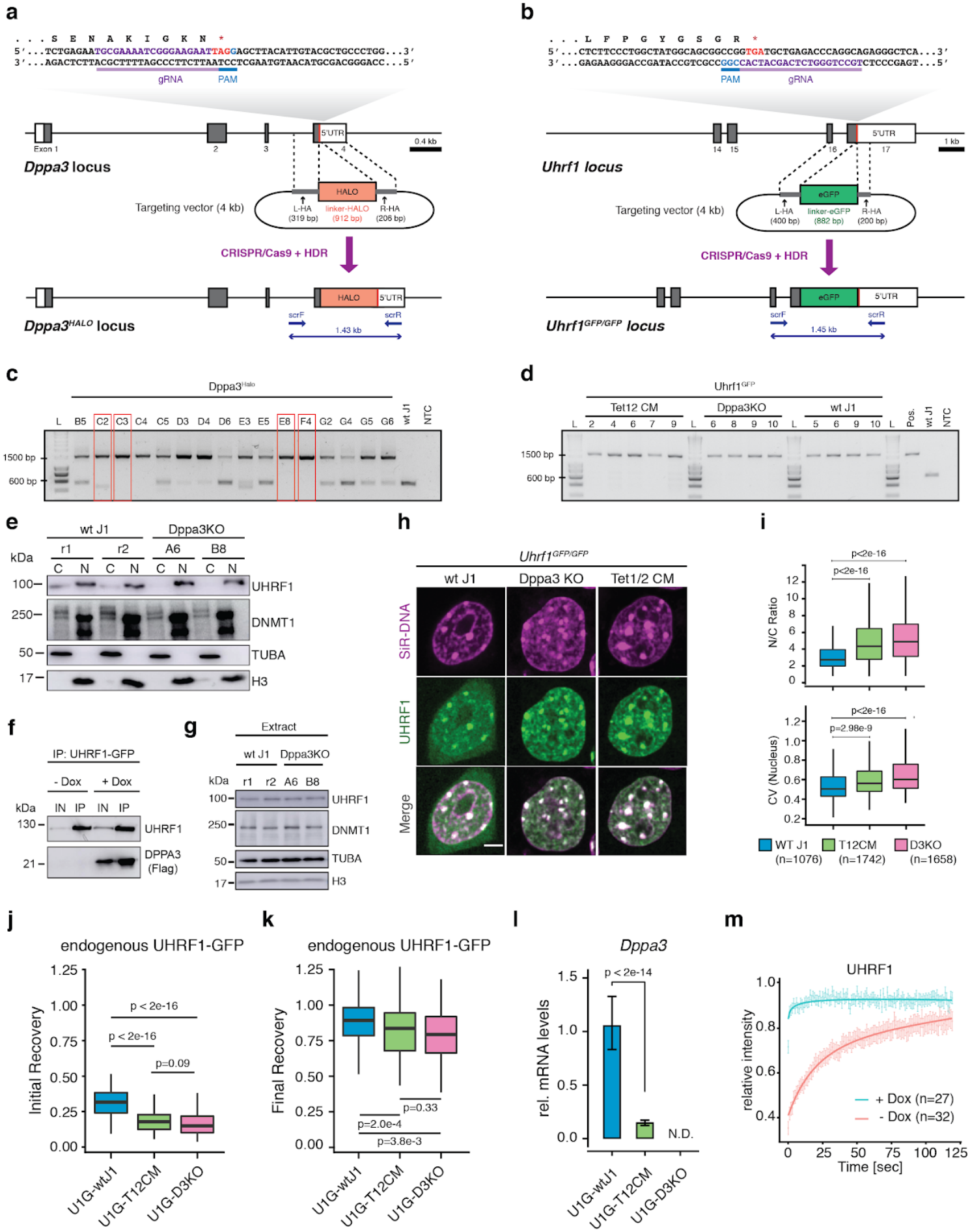
Supplementary Figure 3: Generation, characterization, and methylome profiling of Dppa3KO ESCs

a, Schematic representation of the CRISPR/Cas9 gene editing strategy used to excise the entire *Dppa3* locus. The position of the two locus-flanking gRNAs are shown in orange. PCR primers for determination of locus removal and zygosity are indicated in green. **b,** Results of PCR using the primers indicated in (**a**). *Dppa3* knockout (Dppa3KO) clones, A6 and B8, chosen for further experiments are indicated by the red boxes. The wild-type (wt) and no template control (NTC) reactions are depicted on the right. The sequence alignment of the amplicon generated using primers 1 and 4 for Dppa3KO clones A6 and B8 is provided in the lower portion of (**a**) with solid boxes and the dashed lines representing successfully aligned sequences and genomic sequences not found in the sequenced amplicons respectively. **c,** *Dppa3* transcript levels of the two Dppa3KO clones A6 and B8 are depicted relative to mRNA levels in wt J1 ESCs after normalization to *Gapdh*. Error bars indicate mean ± SD calculated from technical triplicate reactions from $n = 4$ biological replicates. N.D., expression not detectable. **d,** Relative proportion of DNA hypermethylation occurring at each genomic element or retrotransposon class in Dppa3KO ESCs. **e,** Relative proportion of DNA hypermethylation common to T1CM, T2CM, and T12CM ESCs and Dppa3KO ESCs at each genomic element. **f,** Summary of differentially methylated regions either unique to TET mutants (TET-specific) and Dppa3KO ESCs (DPPA3-specific) or shared among TET mutants and Dppa3KO ESCs (common). **g,** Gene ontology (GO) terms associated with promoters specifically dependent on TET activity; adjusted p-values calculated using Fisher's exact test. **h,** Cross-correlation analysis between DNA hypermethylation (ΔmC) in Dppa3KO ESCs and genomic occupancy of histone modifications and UHRF1 binding. Correlations are calculated over 1kb tiles. Positive correlations with ΔmC are bounded by a dashed-line square. **i,** Hypermethylation in Dppa3 KO ESCs is not associated with *Tet* downregulation. Expression of *Tet* genes in wt Dppa3KO ESC clones depicted as mRNA levels relative to *Gapdh*. Error bars indicate mean ± SD calculated from technical triplicate reactions from $n = 2$ biological replicates of each genotype or clone depicted. **j,** Schematic representation of the pSBtet-D3 (pSBtet-3xFLAG-Dppa3-IRES-DsRed) cassette for the Sleeping Beauty transposition-mediated generation of doxycycline (Dox) inducible *Dppa3* ESC lines. Abbreviations: inverted terminal repeat (ITR), tetracycline response element plus minimal CMV (TCE), 3xFLAG tag (3xF), internal ribosomal entry site (IRES), polyA signal (pA), constitutive RPBSA promoter (RPBSA), reverse tetracycline-controlled transactivator (rtTA), self-cleaving peptide P2A (P2A), puromycin resistance (PuroR). **k,** Confirmation of DPPA3 protein induction as assessed by immunofluorescence in uninduced ESCs (-Dox) or after 24 h of doxycycline treatment (+Dox). To illustrate the increase in DPPA3 protein levels, the same acquisition settings used to detect DPPA3, including a long exposure, in uninduced cells were applied for detection of DPPA3 after induction (+Dox (long exp.); bottom panel) leading to a saturated signal. Shorter exposure settings were also applied to the induced cells (+Dox (short exp.); middle panel) to better resolve the localization of DPPA3 after induction. The staining was repeated in 3 biological replicates with similar results. Scale bar: 10 μm. **l,** *Dppa3* expression before induction and after 3 or 6 days of doxycycline treatment in wt J1 + pSBtet-D3 ESCs. mRNA levels of *Dppa3* are shown relative to those in uninduced wt J1 ESCs after normalization to *Gapdh*. Error bars indicate mean ± SDcalculated from technical triplicate reactions from $n = 2$ biological replicates.

In **d-h,** hypermethylation is a gain in 5mC compared to wt ESCs (q value < 0.05; absolute methylation difference > 20%).

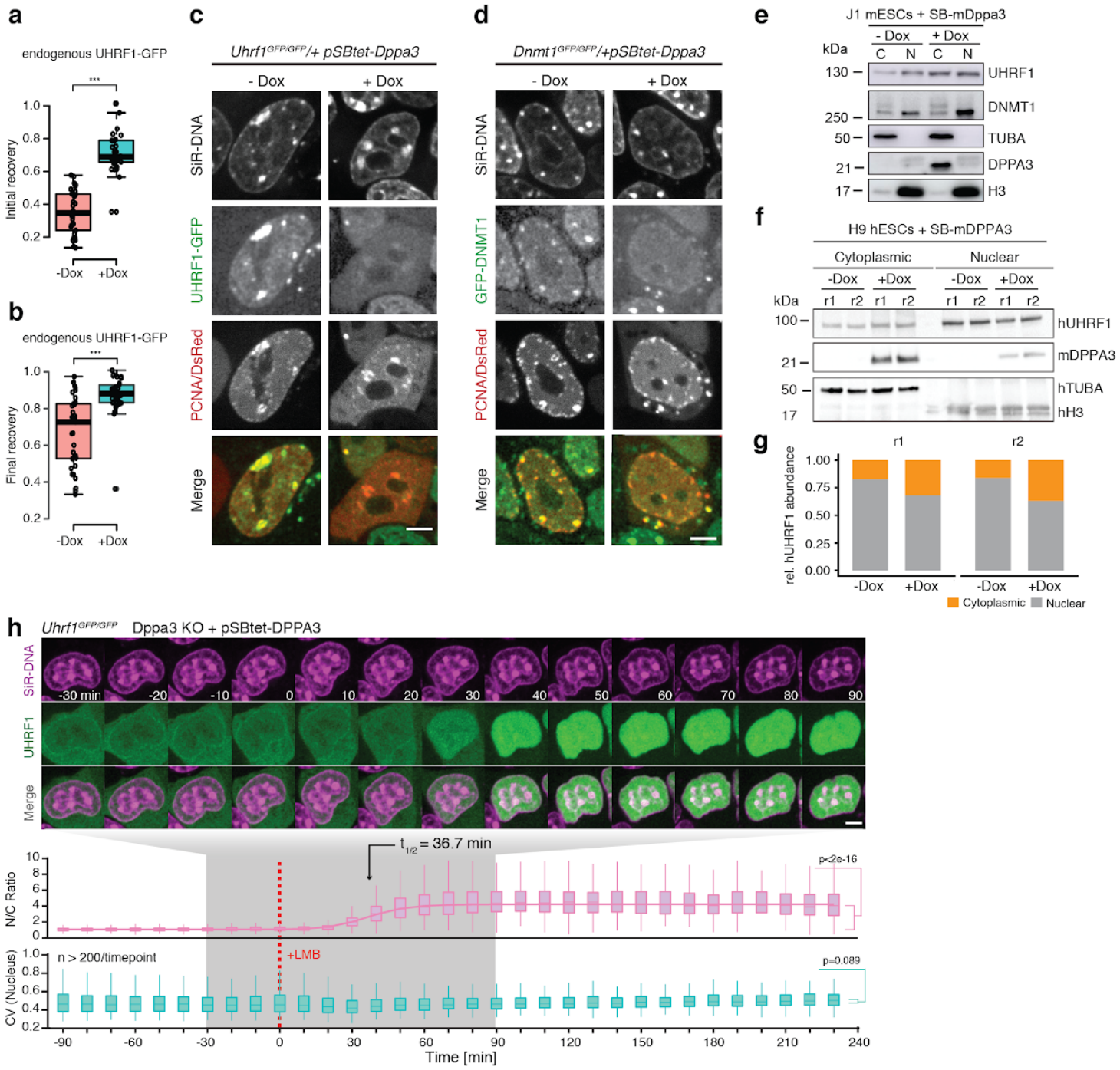
1339

1340


a
. . . S E N A K I G K N *
5'...TCTGAGAATGCGAAAATCGGGAAGAATTAGGAGCTTACATTGTACGCTGCCCTGG...3'
3'...AGACTCTTACGCTTTTAGCCCTTCTTAATCCTCGAATGTAACATGCGACGGGACC...5'
gRNA
PAM
Exon 1
2
3
4
5'UTR
0.4 kb
Dppa3 locus
Targeting vector (4 kb)
HALO
L-HA (319 bp)
linker-HALO (912 bp)
R-HA (206 bp)
CRISPR/Cas9 + HDR
Dppa3HALO locus
scrF
scrR
1.43 kb
b
. . . L F P G Y G S G R *
5'...CTCTTCCCTGGCTATGGCAGCGGCCGGTGATGCTGAGACCCAGGCAGAGGGCTCA...3'
3'...GAGAAGGGACCGATACCGTCGCCGGCCACTACGACTCTGGGTCCGTCTCCCGAGT...5'
PAM
gRNA
14
15
16
17
5'UTR
1 kb
Uhrf1 locus
Targeting vector (4 kb)
eGFP
L-HA (400 bp)
linker-eGFP (882 bp)
R-HA (200 bp)
CRISPR/Cas9 + HDR
Uhrf1GFP/GFP locus
1.45 kb
c
Dppa3Halo
L B5 C2 C3 C4 C5 D3 D4 D6 E3 E5 E8 F4 G2 G4 G5 G6 wt J1 NTC
1500 bp
600 bp
d
Uhrf1GFP
Tet12 CM
Dppa3KO
wt J1
L 2 4 6 7 9 L 6 8 9 10 L 5 6 9 10 L Pos. wt J1 NTC
1500 bp
600 bp
e
wt J1
Dppa3KO
r1
r2
A6
B8
kDa
C N C N C N C N
100
250
50
17
UHRF1
DNMT1
TUBA
H3
f
IP: UHRF1-GFP
- Dox
+ Dox
kDa
IN IP IN IP
130
21
UHRF1
DPPA3 (Flag)
g
Extract
wt J1
Dppa3KO
kDa
r1 r2 A6 B8
100
250
50
17
UHRF1
DNMT1
TUBA
H3
h
Uhrf1GFP/GFP
wt J1
Dppa3 KO
Tet1/2 CM
SiR-DNA
UHRF1
Merge
i
p<2e-16
p<2e-16
N/C Ratio
p<2e-16
p=2.98e-9
CV (Nucleus)
WT J1 (n=1076)
T12CM (n=1742)
D3KO (n=1658)
j
endogenous UHRF1-GFP
Initial Recovery
p < 2e-16
p < 2e-16
p=0.09
U1G-wtJ1
U1G-T12CM
U1G-D3KO
k
endogenous UHRF1-GFP
Final Recovery
p=0.33
p=2.0e-4
p=3.8e-3
U1G-wtJ1
U1G-T12CM
U1G-D3KO
l
Dppa3
rel. mRNA levels
p < 2e-14
N.D.
U1G-wtJ1
U1G-T12CM
U1G-D3KO
m
UHRF1
relative intensity
+ Dox (n=27)
- Dox (n=32)
Time [sec]

Supplementary Figure 4: DPPA3 hinders UHRF1 chromatin association in mouse ESCs

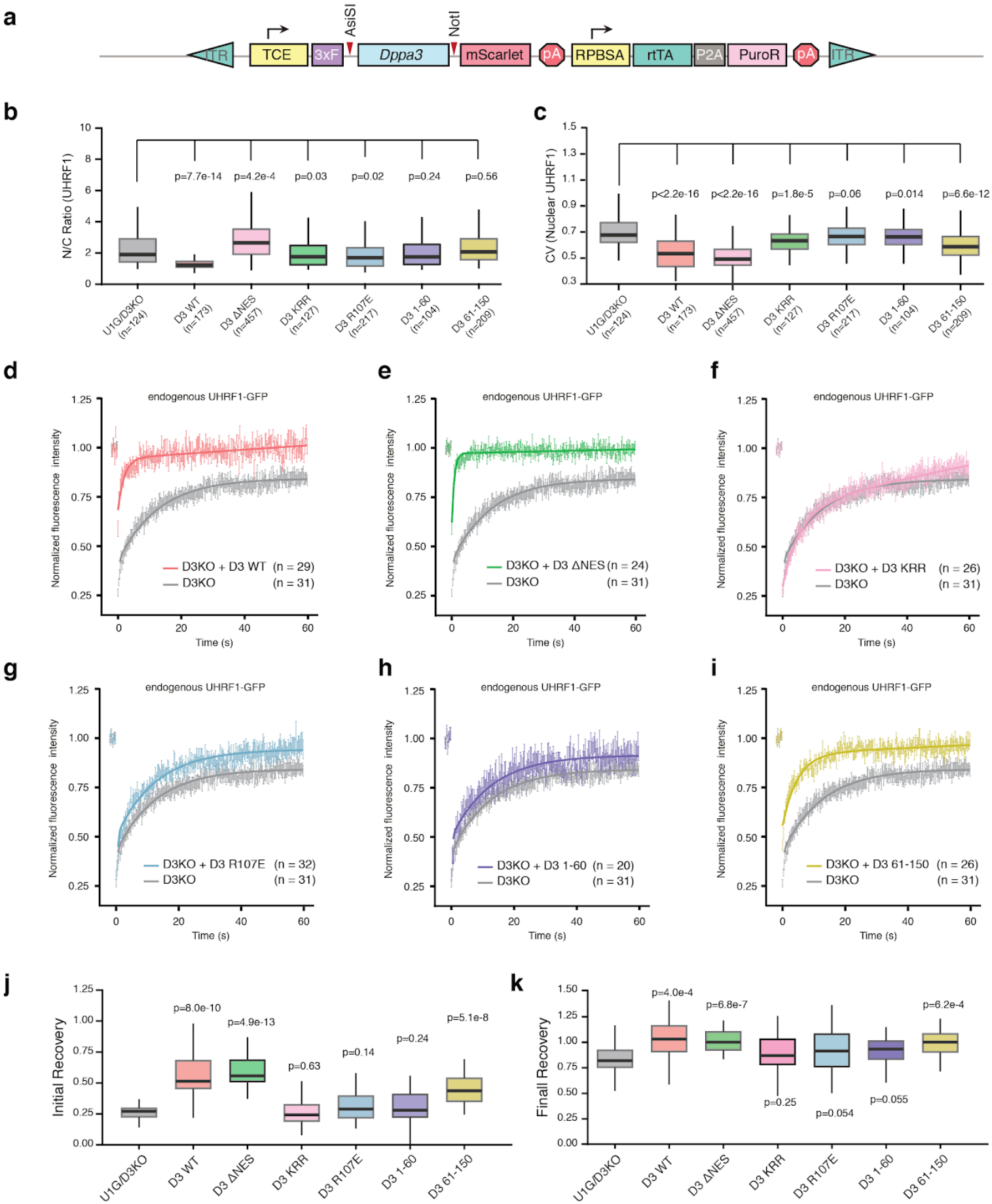
a,b, Schematic of the CRISPR/Cas9 targeting strategy used to insert fluorescent protein tags at the (**a**) *Dppa3* and (**b**) *Uhrf1* loci. Exons are depicted as boxes, with the coding sequences shaded in grey. Stop codons are indicated with a red line. The nucleotide and amino acid sequence encompassing the gRNA target site are enlarged for detail. gRNAs target sequences are colored purple and their respective PAMs blue. Donor constructs harboring the coding sequence for (**a**) HALO or (**b**) eGFP are flanked by homology arms (L-HA, left homology arm and R-HA, right homology arm) and indicated as grey rectangles. Genotyping PCR primers used in (**c,d**) are shown as arrows along with the size of the amplicon confirming successful integration of the donor sequence. **c,d,** Results of genotyping PCRs for confirming the successful integration of (**c**) HALO into the *Dppa3* locus of wt J1 ESCs and (**d**) eGFP into the *Uhrf1* locus of wt J1, Dppa3KO, and T12CM ESCs. Wild-type (wt) and negative control without DNA template (NTC) are depicted on the right. **c,** Clones with Sanger-sequence validated homozygous insertion of the HALO coding sequence are indicated with red boxes. **d,** PCR results after pre-screening via microscopy. All depicted clones have a correct homozygous insertion of eGFP as validated by Sanger-sequencing. PCRs were performed using the primers depicted in (**a,b**). **e,** DPPA3 regulates subcellular UHRF1 distribution. Cytoplasmic, “C”, and nuclear, “N”, fractions from wt and Dppa3KO ESCs analyzed by immunoblot detection using the indicated antibodies. The fractionation and immunoblot were repeated 2 times with similar results obtained. **f,** Immunoblot detection of anti-GFP immunoprecipitated material (IP) and the corresponding input (IN) extracts from U1G/wt + pSBtet-D3 ESCs treated before (-Dox) and after (+Dox) 24 h induction of *Dppa3* with doxycycline. Immuoblot was repeated 2 times with similar results obtained. **g,** UHRF1 and DNMT1 protein levels are unaffected by DPPA3 loss. Whole-cell extracts from wt and Dppa3KO ESCs analyzed by immunoblot detection using the indicated antibodies. Immunoblot depicts $n = 2$ biological replicates. **h,i,** DPPA3 alters the localization and nuclear patterning of endogenous UHRF1-GFP. **h,** Representative confocal images of UHRF1-GFP in live U1G-wt, U1G-T12CM, U1G-Dppa3KO ESCs with DNA counterstain (SiR-Hoechst). Scale bar: 5 µm. **i,** Quantification of endogenous UHRF1-GFP (top panel) Nucleus to cytoplasm ratio (N/C Ratio) and (bottom panel) coefficient of variance (CV) within the nucleus of cells (number indicated in the plot) from $n = 3$ biological replicates per genotype. The assay was repeated 2 times with similar results. **j,k,** Endogenous DPPA3 prevents excessive UHRF1 chromatin binding in ESCs. Further analysis of the single cell FRAP data presented in Fig. 4c. showing the (**j**) initial and (**k**) final relative recovery of endogenous UHRF1-GFP. Intensity measurements 1 s and 60 s after photobleaching were used for calculating (**j**) initial recovery and (**k**) final recovery, respectively. **l,** *Dppa3* is significantly downregulated in U1G-T12CM ESCs. *Dppa3* transcript levels in U1G-wt, U1G-T12CM, U1G-Dppa3KO ESCs are depicted relative to mRNA levels in U1G-wt J1 ESCs after normalization to Gapdh. Error bars indicate mean ± SD calculated from technical triplicate reactions from $n = 4$ biological replicates (independent clones). N.D., expression not detectable. **m,** DPPA3 expression abolishes UHRF1 chromatin binding. FRAP analysis of endogenous UHRF1-GFP before (-Dox) and after 48 h of *Dppa3* induction (+Dox) in U1G-wt + pSBtet-D3 ESCs. The FRAP experiment was repeated at least 3 times with similar results obtained. In the boxplots in **i-k**, darker horizontal lines within boxes represent median values. The limits of the boxes indicate upper and lower quartiles, and whiskers indicate the 1.5-fold interquartile range. P-values in **i-l**; Welch’s two-sided t-test. In **m,** the mean fluorescence intensity of n cells (indicated in the plots) at each timepoint depicted as a shaded dot. Error bars indicate mean ± SEM. Curves (solid lines) indicate double-exponential functions fitted to the FRAP data.


a
endogenous UHRF1-GFP
Initial recovery
-Dox
+Dox
b
endogenous UHRF1-GFP
Final recovery
c
Uhrf1GFP/GFP/+ pSBtet-Dppa3
- Dox
+ Dox
SiR-DNA
UHRF1-GFP
PCNA/DsRed
Merge
d
Dnmt1GFP/GFP/+pSBtet-Dppa3
GFP-DNMT1
e
J1 mESCs + SB-mDppa3
kDa
C N C N
UHRF1
DNMT1
TUBA
DPPA3
H3
f
H9 hESCs + SB-mDPPA3
Cytoplasmic
Nuclear
r1 r2 r1 r2 r1 r2 r1 r2
hUHRF1
mDPPA3
hTUBA
hH3
g
rel. hUHRF1 abundance
Cytoplasmic
Nuclear
h
Uhrf1GFP/GFP Dppa3 KO + pSBtet-DPPA3
SiR-DNA
UHRF1
Merge
t1/2 = 36.7 min
N/C Ratio
p<2e-16
n > 200/timepoint
+LMB
CV (Nucleus)
p=0.089
Time [min]
i
endogenous UHRF1-GFP
Fluorescence Intensity
+DOX + LMB (n = 39)
+DOX - LMB (n = 36)
- DOX + LMB (n = 33)
- DOX - LMB (n = 34)
Time (s)
j
endogenous UHRF1-GFP
Initial Recovery (t =1 s)
p<2.2e-16
p=2.2e-4
p=0.04
+DOX + LMB
+DOX - LMB
- DOX + LMB
- DOX - LMB
k
endogenous UHRF1-GFP
Final Recovery (t = 60 s)
p=3.9e-4
p=0.13
p=0.26
p=1.6e-6
p=3.6e-7
p=8.9e-11

Supplementary Figure 5: DPPA3 alters the localization and chromatin binding of endogenous UHRF1 and DNMT1 in ESCs

a,b, DPPA3 releases UHRF1 from chromatin. Initial (**a**) and final (**b**) relative recovery of UHRF1-GFP in *Uhrf1*$^{GFP/GFP}$*/SBtet-3xFLAG-Dppa3* ESCs before (-Dox) and after 48 h of *Dppa3* induction (+Dox) (+Dox: n = 27; -Dox: n = 32) (data from FRAP experiment in Supplementary Fig. 4m). **c,d,** DPPA3 disrupts the recruitment of UHRF1 and DNMT1 to replication foci. Live cell imaging illustrating the late S-phase localization of (**c**) UHRF1-GFP in U1G/wt + pSBtet-D3 ESCs or (**d**) GFP-DNMT1 in live *Dnmt1*$^{GFP/GFP}$ + pSBtet-D3 ESCs before (-Dox) and after 48 h of *Dppa3* induction. Transfected RFP-PCNA marks sites of active DNA replication within the nucleus. The assay was performed 2 times with similar results. Expressed free DsRed is a marker of doxycycline induction (see cytoplasmic signal in +Dox PCNA/DsRed panels and Supplementary Fig. 3j). DNA was stained in live cells using a 30 min treatment of SiR-DNA (SiR-Hoechst). Scale bar: 5 µm. **e-g,** DPPA3 alters the subcellular distribution of endogenous UHRF1 in mouse ESCs and human ESCs. **e,f,** Immunoblots of nuclear, "N", and cytoplasmic, "C", fractions from (**e**) U1G/wt + pSBtet-D3 mouse ESCs and (**f**) Human H9 ESCs + SBtet-3xFLAG-Dppa3 before (-Dox) and after 24 hours of *Dppa3* induction (+Dox) using the indicated antibodies. An anti-FLAG antibody was used for detection of FLAG-DPPA3. **g,** Quantification of the relative abundance of hUHRF1 in the cytoplasmic and nuclear fractions shown in (**f**) with the results of two independent biological replicates (r1 and r2) displayed. **h,** Nuclear export is dispensable for DPPA3-mediated inhibition of UHRF1 chromatin association. Localization dynamics of endogenous UHRF1-GFP in response to inhibition of nuclear export using leptomycin-B (LMB) after *Dppa3* induction in U1G/D3KO + pSBtet-D3 ESCs with confocal time-lapse imaging over 5.5 h (10 min intervals). $t = 0$ corresponds to start of nuclear export inhibition (+LMB). (*top panel*) Representative images of UHRF1-GFP and DNA (SiR-Hoechst stain) throughout confocal time-lapse imaging. Scale bar: 5 µm. (middle panel) Nucleus to cytoplasm ratio (N/C Ratio) of endogenous UHRF1-GFP signal. (bottom panel) Coefficient of variance (CV) of endogenous UHRF1-GFP intensity in the nucleus. (*middle and bottom panel*) N/C Ratio and CV values: measurements in $n > 200$ single cells per time point, acquired at $n = 20$ separate positions. Curves represent fits of four parameter logistic (4PL) functions to the N/C Ratio (pink line) and CV (green line) data. The experiment was repeated 2 times with similar results. **i-k,** Nuclear export is not only dispensable for but attenuates DPPA3-mediated inhibition of UHRF1-GFP chromatin binding. **i,** FRAP analysis of endogenous UHRF1-GFP within the nucleus of U1G/D3KO + pSBtet-D3 ESCs before (-Dox) and after 48 h of *Dppa3* induction and before (-LMB) and after (+LMB) 3 h of leptomycin-B (LMB) treatment. The mean fluorescence intensity of n cells (indicated in the plots) at each timepoint depicted as a shaded dot. Error bars indicate mean ± SEM. The FRAP experiment was repeated 3 times with similar results. **j,k,** Further analysis of the data from (**i**) showing the (**j**) initial and (**k**) final relative recovery of endogenous UHRF1-GFP.

For **a,b,j,k,** initial and final recoveries were calculated using the intensities measured 1 s or 60 s after photobleaching, respectively. In the boxplots, darker horizontal lines within boxes represent median values. The limits of the boxes indicate upper and lower quartiles, and whiskers indicate the 1.5-fold interquartile range. Welch's two-sided t-test. *** $P < 0.001$; ** $P = 0.002$; * $P = 0.03$; n.s.: not significant.

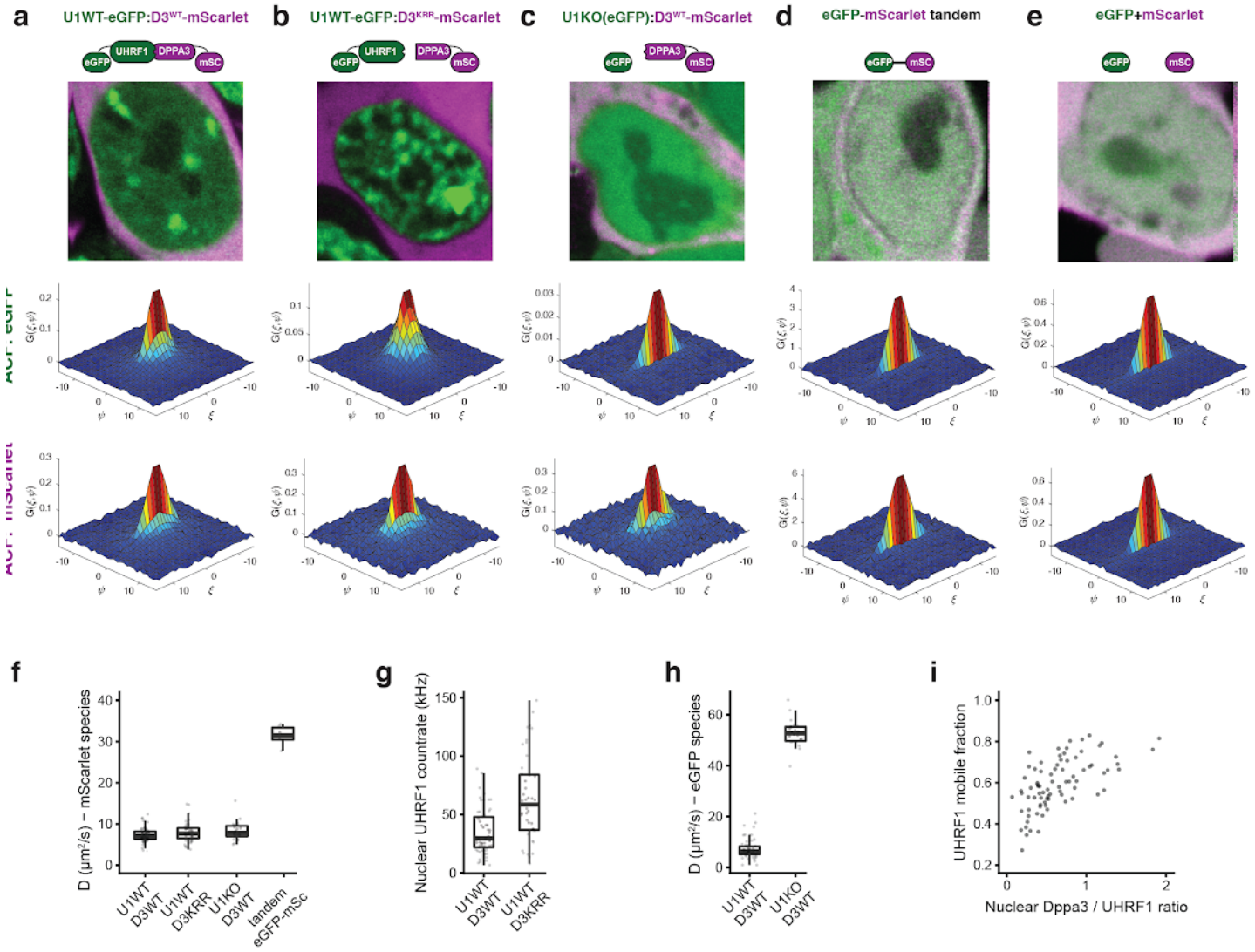

a
ITR
TCE
3xF
AsiSI
Dppa3
NotI
mScarlet
pA
RPBSA
rtTA
P2A
PuroR
pA
ITR
b
N/C Ratio (UHRF1)
p=7.7e-14
p=4.2e-4
p=0.03
p=0.02
p=0.24
p=0.56
U1G/D3KO (n=124)
D3 WT (n=173)
D3 ΔNES (n=457)
D3 KRR (n=127)
D3 R107E (n=217)
D3 1-60 (n=104)
D3 61-150 (n=209)
c
CV (Nuclear UHRF1)
p<2.2e-16
p<2.2e-16
p=1.8e-5
p=0.06
p=0.014
p=6.6e-12
d
endogenous UHRF1-GFP
Normalized fluorescence intensity
Time (s)
D3KO + D3 WT (n = 29)
D3KO (n = 31)
e
D3KO + D3 ΔNES (n = 24)
D3KO (n = 31)
f
D3KO + D3 KRR (n = 26)
D3KO (n = 31)
g
D3KO + D3 R107E (n = 32)
D3KO (n = 31)
h
D3KO + D3 1-60 (n = 20)
D3KO (n = 31)
i
D3KO + D3 61-150 (n = 26)
D3KO (n = 31)
j
Initial Recovery
p=8.0e-10
p=4.9e-13
p=0.63
p=0.14
p=0.24
p=5.1e-8
U1G/D3KO
D3 WT
D3 ΔNES
D3 KRR
D3 R107E
D3 1-60
D3 61-150
k
Finall Recovery
p=4.0e-4
p=6.8e-7
p=0.25
p=0.054
p=0.055
p=6.2e-4

Supplementary Figure 6: Further analysis of endogenous UHRF1-GFP localization and chromatin binding in response to mutant forms of mDPPA3

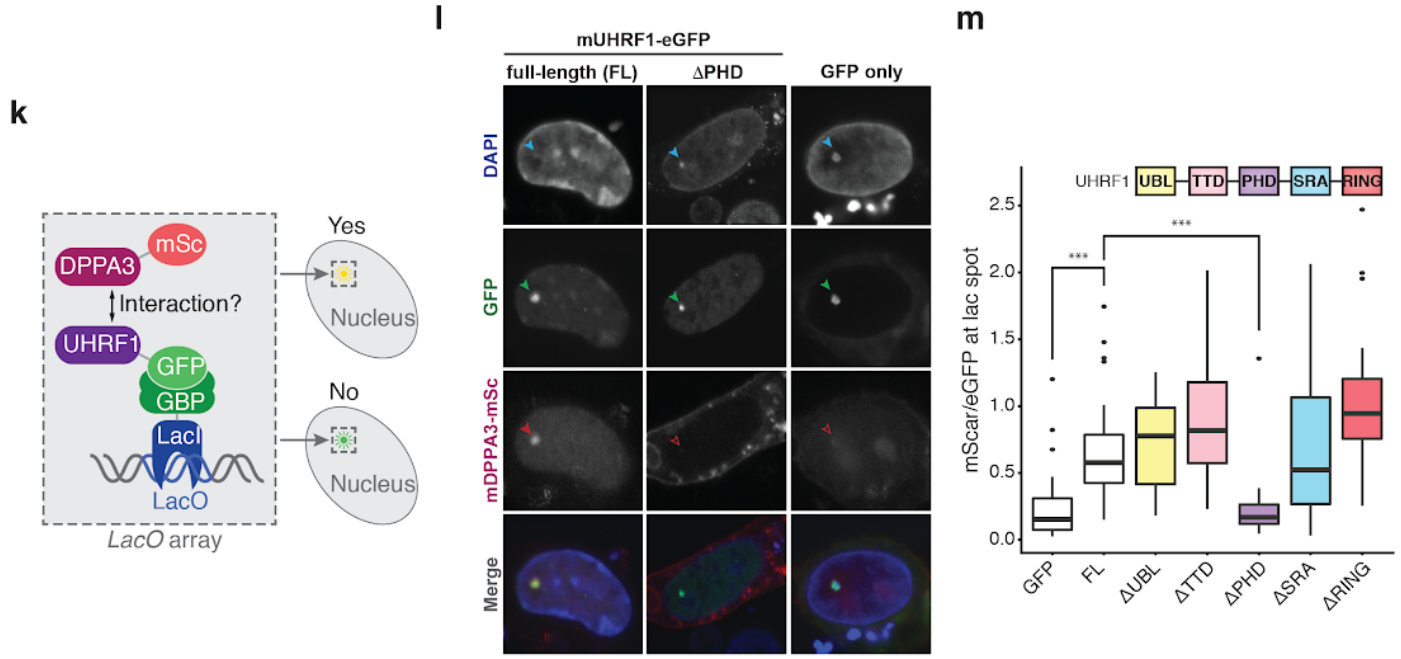
a, Schematic representation of the pSBtet-D3-mSC (pSBtet-3xFLAG-Dppa3-mScarlet) cassette for the Sleeping Beauty transposition-mediated generation of ESC lines with doxycycline (Dox) inducible expression of DPPA3-mScarlet fusions. Abbreviations: inverted terminal repeat (ITR), tetracycline response element plus minimal CMV (TCE), 3xFLAG tag (3xF), internal ribosomal entry site (IRES), polyA signal (pA), constitutive RPBSA promoter (RPBSA), reverse tetracycline-controlled transactivator (rtTA), self-cleaving peptide P2A (P2A), puromycin resistance (PuroR). Restriction sites used for exchanging the *Dppa3* coding sequences are indicated with red arrows. **b,c,** DPPA3 alters the nuclear localization of UHRF1 independently of promoting UHRF1 nucleocytoplasmic translocation. Quantification of endogenous UHRF1-GFP (**b**) Nucleus to cytoplasm ratio (N/C Ratio) and (**c**) coefficient of variance (CV) within the nucleus of U1G/D3KO + pSBtet-D3 ESCs expressing the indicated mutant forms of *Dppa3*. *n* = number of cells analyzed (indicated in the plot). **d-i** Nuclear export and the N-terminus of DPPA3 are dispensable for its inhibition of UHRF1 chromatin binding. FRAP analysis of endogenous UHRF1-GFP within the nucleus of U1G/D3KO + pSBtet-D3 ESCs expressing the following forms of DPPA3: (**d**) wild-type (D3 WT), (**e**) nuclear export mutant L44A/L46A (D3-ΔNES), (**f**) K85E/R85E/K87E mutant (D3-KRR), (**g**) R107E mutant (D3-R107E), (**h**) N-terminal 1-60 fragment (D3 1-60), (**i**) C-terminal 61-150 fragment (D3 61-150). In **d-i,** the mean fluorescence intensity of *n* cells (indicated in the plots) at each timepoint is depicted as a shaded dot. Error bars indicate mean ± SEM. Curves (solid lines) indicate double-exponential functions fitted to the FRAP data. Individual same fits correspond to those in Fig. 5c. FRAP of UHRF1 in cells expressing *Dppa3* mutants was repeated 2 times. **j,k,** Further analysis of the data from (**d-i**) showing the (**j**) initial and (**k**) final relative recovery of endogenous UHRF1-GFP. Intensity measurements 1 s and 60 s after photobleaching were used for calculating (**j**) initial recovery and (**k**) final recovery, respectively.

In the boxplots in **b, c, j,** and **k,** darker horizontal lines within boxes represent median values. The limits of the boxes indicate upper and lower quartiles, and whiskers indicate the 1.5-fold interquartile range. P-values based on Welch's two-sided t-test of the difference between U1/D3KO and each DPPA3 mutant.

1344

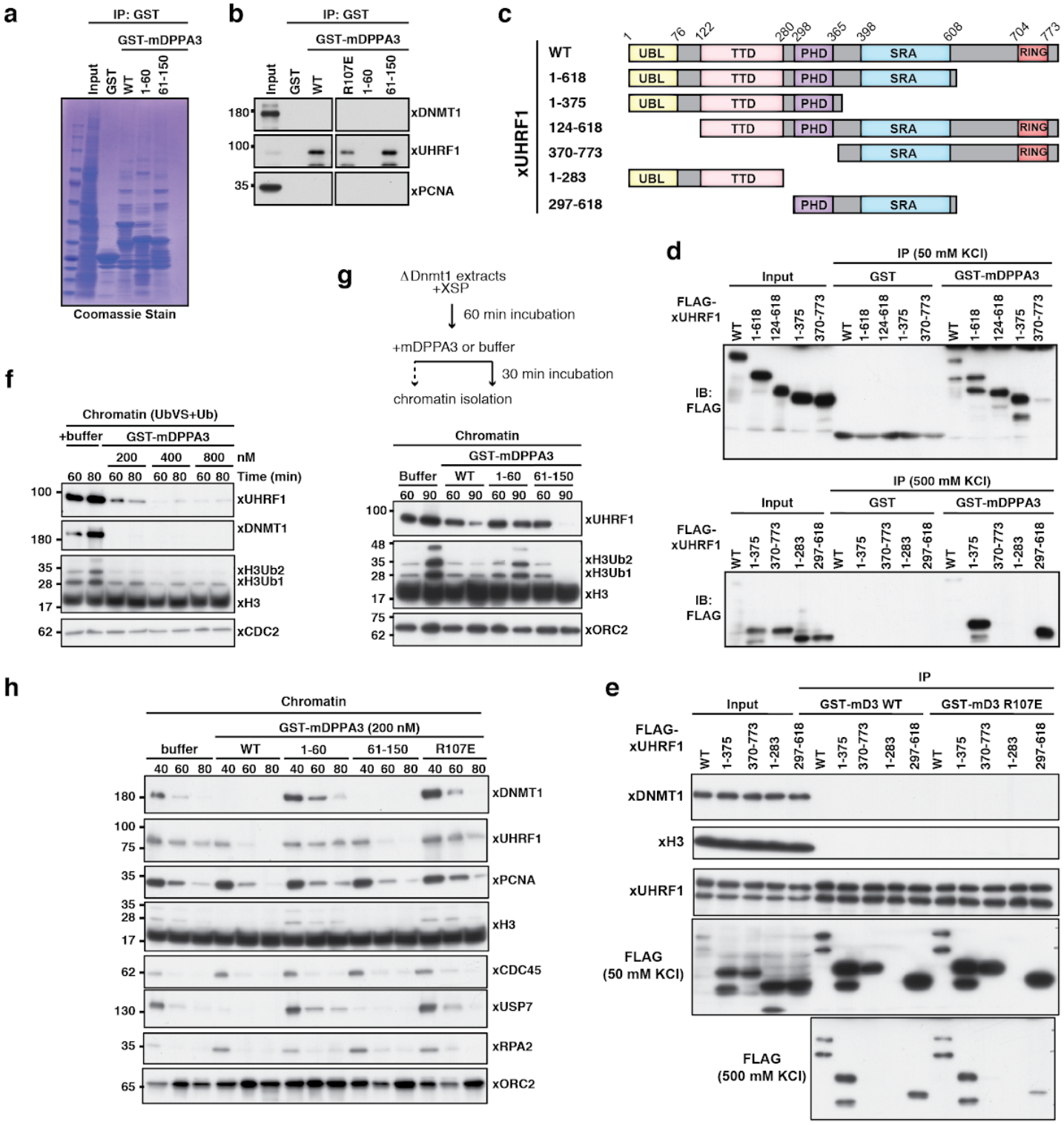

a
U1WT-eGFP:D3WT-mScarlet
b
U1WT-eGFP:D3KRR-mScarlet
c
U1KO(eGFP):D3WT-mScarlet
d
eGFP-mScarlet tandem
e
eGFP+mScarlet
eGFP
UHRF1
DPPA3
mSc
f
D (μm²/s) – mScarlet species
U1WT D3WT
U1WT D3KRR
U1KO D3WT
tandem eGFP-mSc
g
Nuclear UHRF1 countrate (kHz)
U1WT D3WT
U1WT D3KRR
h
D (μm²/s) – eGFP species
U1WT D3WT
U1KO D3WT
i
UHRF1 mobile fraction
Nuclear Dppa3 / UHRF1 ratio

j
mUHRF1
1
78
129
280
304
371
424
586
713
752
782
UBL
TTD
PHD
SRA
RING


k
DPPA3
mSc
Interaction?
UHRF1
GFP
GBP
LacI
LacO
LacO array
Yes
Nucleus
No
Nucleus
l
mUHRF1-eGFP
full-length (FL)
ΔPHD
GFP only
DAPI
GFP
mDPPA3-mSc
Merge
m
UHRF1
UBL
TTD
PHD
SRA
RING
mScar/eGFP at lac spot
***
***
GFP
FL
ΔUBL
ΔTTD
ΔPHD
ΔSRA
ΔRING

Supplementary Fig. 7: DPPA3 binds UHRF1 via its PHD domain to form a mobile complex in ESCs

a-e, Representative dual-color live confocal images (top row) and associated 3D plots of mean autocorrelation (ACF) values of eGFP (middle row) and mScarlet (bottom row) species in U1G/D3KO + pSBtet-D3 ESCs expressing the following forms of DPPA3-mScalet: (**a**) wild-type (U1WT:D3^{WT}) and (**b**) K85E/R85E/K87E mutant (U1WT:D3^{KRR}), in (**c**) Uhrf1KO ESCs expressing free eGFP and wild-type *Dppa3*-mScarlet (U1KO:D3^{WT}), and in control ESCs expressing (**d**) an eGFP-mScarlet tandem fusion (eGFP-mScarlet) and (**e**) free eGFP and free mScarlet (eGFP + mScarlet). In **a-e**, each image is the result of merging the sum projections of 250-frame image series acquired simultaneously from eGFP (green) and mScarlet (magenta) channels. Autocorrelation plots are color-coded to indicate mean correlation value. The fast timescale axis is indicated by ξ, and the slow timescale axis is indicated by ψ. **f,** Calculated diffusion coefficient of mScarlet species in different cell types (those shown in (**a-d**)), derived from a two-component fit of mScarlet autocorrelation functions. Data are pooled from three independent experiments, except tandem eGFP-mScarlet which was from two independent experiments. **g,** Photon count rates of UHRF1-eGFP in nuclei of U1G/D3KOs expressing either DPPA3-WT (U1WT:D3^{WT}) or DPPA3-KRR mutant (U1WT:D3^{KRR}). Data are pooled from three independent experiments. **h,** Calculated diffusion coefficients of UHRF1-GFP (in U1WT:D3^{WT}) or free eGFP (in U1KO:D3^{WT}), derived as those in C. Data are pooled from three independent experiments **i,** Scatter plot showing the relationship between the mobile fraction of UHRF1-eGFP and the ratio of DPPA3-mScarlet:UHRF1-eGFP photon count rates in the nucleus. Data are pooled from three independent experiments. **j,** A schematic diagram illustrating the domains of mUHRF1: ubiquitin-like (UBL), tandem tudor (TTD), plant homeodomain (PHD), SET-and-RING-associated (SRA), and really interesting new gene (RING). **k,** Overview of the F3H assay used to find the domain of UHRF1 which binds to DPPA3. **l,m,** The PHD domain of UHRF1 is necessary for mediating the interaction with DPPA3. **l,** Representative confocal images of free GFP or the ΔPHD UHRF1-GFP constructs (**j**) immobilized at the lacO array (indicated with green arrows). Efficient or failed recruitment of DPPA3-mScarlet to the lacO spot are indicated by solid or unfilled red arrows respectively. **m,** The efficiency of DPPA3-mScarlet recruitment to different UHRF1-eGFP deletion constructs immobilized at the lacO array is given as the fluorescence intensity ratio of mScarlet (DPPA3) to eGFP (UHRF1 constructs) at the nuclear LacO spot. Data represents the results of $n > 10$ cells analyzed from two independent experiments. In **f-h**, each data point indicates the measured and fit values from a single cell. In the box plots **f-h** and **m**, darker horizontal lines within boxes represent median values. The limits of the boxes indicate upper and lower quartiles, and whiskers indicate the 1.5-fold interquartile range. In **m,** P-values were calculated using Welch's two-sided t-test: ** $P < 0.01$; *** $P < 0.001$.


a
IP: GST
GST-mDPPA3
Input
GST
WT
1-60
61-150
Coomassie Stain
b
IP: GST
GST-mDPPA3
Input
GST
WT
R107E
1-60
61-150
180
100
35
xDNMT1
xUHRF1
xPCNA
c
xUHRF1
WT
1-618
1-375
124-618
370-773
1-283
297-618
UBL
TTD
PHD
SRA
RING
1
76
122
280
298
365
398
608
704
773
d
IP (50 mM KCl)
Input
GST
GST-mDPPA3
FLAG-xUHRF1
WT
1-618
124-618
1-375
370-773
IB: FLAG
IP (500 mM KCl)
1-283
297-618
e
IP
Input
GST-mD3 WT
GST-mD3 R107E
xDNMT1
xH3
xUHRF1
FLAG (50 mM KCl)
FLAG (500 mM KCl)
f
Chromatin (UbVS+Ub)
+buffer
GST-mDPPA3
200
400
800
nM
60 80 60 80 60 80 60 80 Time (min)
100
180
35
28
17
62
xUHRF1
xDNMT1
xH3Ub2
xH3Ub1
xH3
xCDC2
g
ΔDnmt1 extracts
+XSP
60 min incubation
+mDPPA3 or buffer
30 min incubation
chromatin isolation
Chromatin
GST-mDPPA3
Buffer
WT
1-60
61-150
60 90 60 90 60 90 60 90
100
48
35
28
17
75
62
xUHRF1
xH3Ub2
xH3Ub1
xH3
xORC2
h
Chromatin
GST-mDPPA3 (200 nM)
buffer
WT
1-60
61-150
R107E
40 60 80 40 60 80 40 60 80 40 60 80 40 60 80
180
100
75
35
35
28
17
62
130
35
65
xDNMT1
xUHRF1
xPCNA
xH3
xCDC45
xUSP7
xRPA2
xORC2
i
Control
No injection
+ mDppa3 WT
+ mDppa3 R107E
100 ng/μl
200 ng/μl
500 ng/μl
5mC
DAPI

Supplementary Figure 8: DPPA3 evolved in boreoeutherian mammals but its function is conserved in lower vertebrates

a, mDPPA3 C-terminus is sufficient for xUHRF1 binding. GST-tagged mDPPA3 wild-type (WT) and truncations (1-60 and 61-150) were immobilized on GSH beads and incubated with *Xenopus* egg extracts. The samples shown in Fig. 7c were subjected to SDS-PAGE and stained with Coomassie Blue. **b,** High affinity interaction between mDPPA3 and xUHRF1 is weakened by R107E. GST-tagged mDPPA3 wild-type (WT), the point mutant R107E, and truncations (1-60 and 61-150) were immobilized on GSH beads and incubated with *Xenopus* egg extracts . Pull-downs were subjected to stringent 500 mM KCl washing. Bound proteins were analyzed using the indicated antibodies. **c,** Schematic diagram illustrating the xUHRF1 deletion constructs used in the pull-downs in (**d,e**). **d,e,** xUHRF1 binds mDPPA3 via its PHD domain. *In vitro* translated xUHRF1 fragments were added to interphase *Xenopus* egg extracts. (**d**) GST and GST-mDPPA3 wild-type (WT) or (**e**) GST-tagged mDPPA3 wild-type (WT) and GST-tagged mDPPA3 R107E were immobilized on GSH beads and then used for GST-pulldowns on egg extracts containing the indicated recombinant xUHRF1 fragments. Bound proteins were analyzed using the denoted antibodies. Pull-downs were subjected to either 50 mM KCl or more stringent 500 mM KCl washing as indicated. **f,** mDPPA3 disrupts xUHRF1-dependent ubiquitylation of H3. As dual-monoubiquitylation of H3 (H3Ub2) is hard to detect given its quick turnover[89], we specifically enhanced ubiquitylation by simultaneous treatment of extracts with ubiquitin vinyl sulfone (UbVS), a pan-deubiquitylation enzyme inhibitor[149] and free ubiquitin (+Ub) as described previously[89]. Buffer or the displayed concentration of recombinant mDPPA3 were then added to the extracts. After the indicated times of incubation, chromatin fractions were isolated and subjected to immunoblotting using the antibodies indicated. **g,** mDPPA3 displaces chromatin-bound xUHRF1. To stimulate xUHRF1 accumulation on hemi-methylated chromatin, *Xenopus* extracts were first immuno-depleted of xDNMT1 as described previously[66]. After addition of sperm chromatin, extracts were incubated for 60 min to allow the accumulation of xUHRF1 on chromatin during S-phase and then the indicated form of recombinant mDPPA3 (or buffer) was added. Chromatin fractions were isolated either immediately (60 min) or after an additional 30 min incubation (90 min) and subjected to immunoblotting using the antibodies indicated. **h,** The region 61-150 of mDPPA3 is sufficient but requires R107 to inhibit xUHRF1 chromatin binding. Sperm chromatin was incubated with interphase *Xenopus* egg extracts supplemented with buffer (+buffer) and GST-mDPPA3 wild-type (WT), the point mutants R107E, truncations (indicated). Chromatin fractions were isolated after the indicated incubation time and subjected to immunoblotting using the antibodies indicated. **i,** mDPPA3-mediated DNA demethylation in medaka. Representative 5mC immunostainings in late blastula stage (~8 h after fertilization) medaka embryos injected with wild-type *mDppa3* (WT) or *mDppa3* R107E (R107E) at three different concentrations (100 ng/µl, 200 ng/µl, or 500 ng/µl). Scale bars represent 50 µm. DNA counterstain: DAPI,4′,6-diamidino-2-phenylindole.

1347