HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2018 July 05.

Published in final edited form as:

Nat Genet. 2017 July ; 49(7): 1061–1072. doi:10.1038/ng.3868.

Lineage-specific functions of TET1 in the postimplantation mouse embryo

Rita Khoueiry[1,7], **Abhishek Sohni**[1,7], **Bernard Thienpont**[2,3], **Xinlong Luo**[1], **Joris Vande Velde**[1], **Michela Bartoccetti**[1], **Bram Boeckx**[2,3], **An Zwijsen**[4,5], **Anjana Rao**[6], **Diether Lambrechts**[2,3], and **Kian Peng Koh**[1]

[1]KU Leuven Department of Development and Regeneration, Stem Cell Institute Leuven, Leuven, Belgium

[2]VIB Center for Cancer Biology, Laboratory for Translational Genetics, Leuven, Belgium

[3]KU Leuven Department of Human Genetics, Laboratory for Translational Genetics, Leuven, Belgium

[4]VIB Center for the Biology of Disease, Leuven, Belgium

[5]KU Leuven Department of Human Genetics, Leuven, Belgium

[6]La Jolla Institute for Allergy and Immunology, La Jolla, California, USA

Abstract

The mammalian TET enzymes catalyze DNA demethylation. While they have been intensely studied as major epigenetic regulators, little is known about their physiological roles and the extent of functional redundancy following embryo implantation. Here we define non-redundant roles for TET1 at an early postimplantation stage of the mouse embryo, when its paralogs *Tet2* and *Tet3* are not detectably expressed. TET1 regulates numerous genes defining differentiation programs in the epiblast and extraembryonic ectoderm. In epiblast cells, TET1 demethylates gene promoters via hydroxymethylation and maintains telomere stability. Surprisingly, TET1 represses a majority of epiblast target genes independently of methylation changes, in part through regulation of the gene encoding the transcriptional repressor JMJD8. Dysregulated gene expression in the absence of TET1 causes embryonic defects, which are partially penetrant in an inbred strain but fully lethal in

Correspondence should be addressed to K.P.K. (kian.koh@kuleuven.be).

[7]These authors contributed equally to this work.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

AUTHOR CONTRIBUTIONS

R.K. performed all experiments involving mouse embryo dissection and manipulation, derivation of ESC lines, cell culture, cellular flux assays and amplicon sequencing analysis. A.S. performed RNA–seq analysis, Flow-FISH and luciferase reporter assays and generated TET1 and JMJD8 rescue cell lines. B.T. performed MeDIP, hMeDIP, mass spectrometry and integrative ChIP–seq with WGBS analysis. X.L. performed TET1 and ΔN-JMJD8 ChIP–qPCR and ChIP–seq analysis. J.V.V. performed dot blot analysis, ELISA and confocal microscopy. M.B. contributed X-gal staining of embryos and assisted in derivation of ESC lines. B.B. contributed IPA analysis. A.Z. contributed expertise on mouse dissection and WISH assays. A.R. contributed the *Tet1* GT mouse strain. D.L. provided next-generation sequencing expertise and support. R.K. and A.S. wrote the methods and prepared all figures. K.P.K. conceived the study, directed the research and wrote the manuscript with assistance from all co-authors.

non-inbred mice. Collectively, our study highlights an interplay between the catalytic and non-catalytic activities of TET1 that is essential for normal development.

Dynamic changes in cytosine methylation are critical components of genome-wide epigenetic events that program the early mammalian embryo for development[1,2]. DNA methylation of CpG dinucleotides at gene promoters often elicits transcriptional repression critical for cell fate commitment[3,4]. The role of DNA methylation erasure mechanisms during early peri- and postimplantation stages is much less understood.

Mouse embryonic stem cells (ESCs), which are derived from the inner cell mass (ICM) of embryonic day (E) 3.5 blastocysts, are *in vitro* models of 'naive' pluripotency[5]. However, ESCs in conventional serum-containing cultures deviate from the ICM by gaining *de novo* DNA methylation[6–8]. To reach a 'ground state' that recapitulates the ICM, ESCs are re-adapted to serum-free medium containing leukemia inhibitory factor (LIF) and inhibitors of MEK and GSK3 (2iL)[6]. ESCs in 2iL can be differentiated *in vitro* into 'primed' epiblast-like cells (EpiLCs) that mimic the E5.5–E6.0 pre-gastrulation epiblast[9]. In comparison, epiblast stem cells (EpiSCs) derived from postimplantation E5.5–E7.5 epiblasts progress in culture to resemble the E7.5 anterior primitive streak of the mouse gastrula[10–12]. An adequate choice of *in vitro* models is therefore crucial to capture the rapidly changing cellular states of peri- and postimplantation development.

The ten–eleven translocation (TET) DNA dioxygenases (TET1, TET2 and TET3) erase DNA methylation by reiterative oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC)[13–15]. The contributions of TET1 and TET2 in naive pluripotency, where both are highly expressed, have been intensely studied in ESCs[16–19]. Because *Tet1* and *Tet2* expression levels are low in EpiSCs, both genes were presumed to be silenced soon after implantation[7,20]. However, we recently showed a persistence of *Tet1* expression and silencing of only *Tet2* in the transition from ESCs to EpiLCs[21].

The embryonic lethality of *Tet1*-deficient mice has been controversial[22–25] (Supplementary Note). On the basis of functional redundancies between the TET enzymes, recent studies have examined embryonic defects in combined TET double and triple knockouts[25–27]. However, the precise roles of individual TET enzymes during development remain unresolved. Here we report non-redundant functions of TET1 in pre-streak mouse embryos and EpiLCs.

RESULTS

Tet1 is detected with genomic 5hmC in the postimplantation epiblast

To profile the temporal and spatial expression of *Tet1*, *Tet2* and *Tet3* in mice at early postimplantation stages, we performed whole-mount *in situ* hybridization (WISH) on C57BL/6 (B6) embryos. *Tet1* expression was prominent in the epiblast (Epi) at E6.5 but was diminished at E7.5, with persistence of a faint signal in the anterior half of the gastrula (Fig. 1a). During these stages, *Tet2* and *Tet3* were undetectable by WISH and qPCR (Fig. 1a,b). At E8.5, *Tet1* and *Tet3* were weakly expressed in the head folds and neural tube, whereas

Tet2 was not detectable. At E9.5–E10.5, all three TET genes were expressed in the developing brain (Fig. 1a,b).

Given that TET proteins catalyze the formation of 5hmC, we next analyzed genomic 5hmC distribution in the early postimplantation embryo. On the basis of immunohistochemistry and dot blot detection, 5hmC was enriched in the E6.5 Epi in comparison to the extraembryonic ectoderm (ExE) and surrounding decidual tissues (Fig. 1c,d). The 5hmC patterns faded out at E7.5–E8.5 and reappeared in tissue-specific regions at E9.5–E10.5 (Fig. 1d, Supplementary Fig. 1a and data not shown), in line with the expression patterns of the TET genes. Thus, whereas all three TETs may collectively contribute 5hmC in the E9.5–E10.5 embryo, TET1 appears to have a unique role in the E6.5 epiblast.

Loss of TET1 causes early embryonic defects

We asked whether the solitary expression of *Tet1* in the E6.5–E7.5 epiblast had any functional implication. In previous studies, knockouts in which the *Tet1* gene was targeted were generated by deletions encompassing exon 11, which encodes the catalytic site; however, deletions of only 3′ exons may allow expression of an N-terminal fragment (~70 kDa encoded by exons 2–4). To delineate the implications of more complete loss of TET1, we first examined the *Tet1* GT(RRG140) strain (donor 129P2/OlaHsd), in which transcription is disrupted after the first coding exon (exon 2) on the mutant allele (Supplementary Fig. 1b–h).

X-gal staining in heterozygous $Tet1^{GT/wt}$ embryos recapitulated the developmental patterns of *Tet1* expression observed by WISH (Fig. 1e,f, Supplementary Fig. 2a,b and Supplementary Note). Interestingly, $Tet1^{GT/GT}$ embryos obtained from early mixed-strain intercrosses (≤3 backcrosses to B6) were developmentally abnormal in comparison to wild-type and heterozygous littermate embryos at E8.0, and none survived beyond E9.5 (Supplementary Fig. 2c,d). Further backcrossing (>5 crosses) produced viable *Tet1*-deficient mice, but these were born below the expected Mendelian ratio and were developmentally stunted (Fig. 1g–j, Supplementary Fig. 2e,f and Supplementary Note). These observations indicate embryonic defects at late gastrulation in both inbred and non-inbred *Tet1*- deficient strains *in vivo*; the degree of penetrance is surprisingly reduced by inbreeding.

To validate these gene-trap (GT) phenotypes by an independent strategy, we generated a B6 line containing a *Tet1*-targeted mutation ($Tet1^{tm1Koh}$), in which the *lacZ* reporter is inserted in frame immediately downstream of the ATG start codon (Supplementary Fig. 2g). This approach effectively ablates expression of the full coding sequence of *Tet1*. In this mutant, we observed embryonic defects that recapitulated observations in the incipient congenic B6 $Tet1^{GT/GT}$ mice. By E9.5, several *Tet1*-null embryos showed deformities in fore-brain development associated with incomplete closure of the anterior neuropore (Fig. 1k). Collectively, our analysis in two alternative *Tet1*-mutant strains may suggest additional non-catalytic functions of TET1 requiring its N-terminal domain.

Intriguingly, $Tet1^{GT/GT}$ embryos were stained more intensely by X-gal and in broader regions than their $Tet1^{GT/wt}$ littermates (Fig. 1f,g and Supplementary Fig. 2c–f). We reasoned that the presence of biallelic *lacZ* reporters in $Tet1^{GT/GT}$ embryos has allowed

detection of the GT fusion product in somatic and extraembryonic cells, which would otherwise be below the detection threshold when only a single allele is expressed in *Tet1*$^{GT/wt}$ embryos (Supplementary Note). If this premise were valid, we would expect *Tet1*$^{lacZ/wt}$ embryos from the *Tet1*^{tm1Koh} strain, which expresses β-galactosidase directly from the *Tet1* promoter without gene fusion, to have a lower threshold for X-gal detection and stain like *Tet1*$^{GT/GT}$ embryos. Indeed, E6.5 *Tet1*$^{lacZ/wt}$ embryos showed very intense staining in the Epi, with lighter staining also observed in the ExE (Supplementary Fig. 2h), indicating *Tet1* expression at different levels in both embryonic and extraembryonic lineages.

TET1 regulates lineage genes in the pre-streak mouse epiblast

On the basis of TET1 expression, we examined the impact of TET1 loss on gene expression in the pre-gastrula embryo. Incipient congenic N5-6 B6 embryos were collected from *Tet1*$^{GT/wt}$ (heterozygous) intercrosses at E6.25—staged by visible absence of the primitive streak, the first morphological hallmark of gastrulation—and dissected into Epi and ExE tissues for RNA–seq (Fig. 2a, Supplementary Fig. 3a–c and Supplementary Note).

Among the differentially expressed genes in *Tet1*-deficient Epi were several lineage markers (Fig. 2b, Supplementary Fig. 3d and Supplementary Tables 1 and 2). *Tet1* was not detected by these criteria because of reads mapping to exon 2 (1.9 kb) in *Tet1*$^{GT/GT}$ samples; otherwise, reads downstream of the GT cassette were absent (Fig. 2c). Interestingly, heat map clustering of the differentially expressed genes showed that most were upregulated rather than downregulated in *Tet1*$^{GT/GT}$ Epi (Fig. 2d); several were primitive streak markers, which were expressed at low levels in wild-type and heterozygous Epi, as expected at the pre-streak (E6.25) stage. They included *Twist1* and *Cd44*, both associated with epithelial-to-mesenchymal transition[28]—a trait of primitive streak formation—and the mesodermal transcription factors *Barx1* and *Bin1* (Fig. 2e and Supplementary Fig. 3d). These *Tet1*$^{GT/GT}$ Epi profiles are suggestive of precocious entry into mesendodermal fate that precedes morphological manifestation of the primitive streak, providing further evidence that TET1 regulates the onset of gastrulation.

Notably, the few genes downregulated in *Tet1*$^{GT/GT}$ Epi included *Nnat* (neuronatin; also known as *Peg5*) and *Lefty2* (Fig. 2f). *Nnat* has been implicated in promoting neuronal fate in ESCs[29]. *Lefty2* blocks TGF-β–Nodal signaling by graded spatial expression during gastrulation to regulate lineage segregation[30]. Loss of both *Nnat* and *Lefty2* expression in *Tet1*-deficient Epi is consistent with observations of hyperactive Nodal signals and skewing toward the mesendoderm when *Tet1*-deficient ESCs differentiate into embryoid bodies *in vitro*[16]. In wild-type E7.5 embryos, *Lefty2* was prominently detected in the embryonic mesoderm in the posterior lateral region by WISH (Fig. 2g). In *Tet1*$^{GT/GT}$ littermate embryos, *Lefty2* was posteriorly restricted and reduced in expression, both in inbred and outbred mice (Fig. 2g), confirming *Lefty2* as a gene regulated by TET1 *in vivo*.

To examine whether loss of TET1 affected DNA methylation at the differentially expressed genes identified *in vivo*, we performed locus-specific bisulfite sequencing. *Tet1*$^{GT/GT}$ Epi exhibited increased methylation (5mC and 5hmC) in an intragenic CpG island (CGI) of *Lefty2* (Fig. 2h and Supplementary Fig. 3e), consistent with TET1 occupancy at this locus as

previously shown in ESCs[18]. Similarly, *Tet1* CGI promoter and enhancer sites showed substantial increases in methylation in *Tet1*$^{GT/GT}$ Epi cells (Fig. 2i and Supplementary Fig. 3f), in agreement with autoregulation by TET1 (ref. 31). However, the *Nnat* promoter was fully methylated in wild-type and *Tet1*$^{GT/GT}$ Epi (data not shown). On the other hand, the *Cd44* transcription start site (TSS) gained methylation upon loss of TET1, although gene expression increased (Fig. 2j), suggesting that gain in promoter methylation upon loss of TET1 does not strictly repress gene expression.

TET1 suppresses metabolic genes in the extraembryonic ectoderm

Next, we examined TET1 loss of function in the ExE. Consistent with *Tet1* expression, RNA–seq analysis produced reads mapping to *Tet1* exons in wild-type ExE samples, albeit at lower levels than in the Epi (compare Fig. 2c and Fig. 3a). Reads were reduced in heterozygous ExE samples, compatible with subthreshold X-gal detection of *Tet1* reporter activity in heterozygous ExE (Supplementary Fig. 3c). In contrast, few reads were found for *Tet2* and *Tet3* (<2 RPKM), which was validated by qPCR (Fig. 3b). These observations confirm TET1 as the dominant TET protein also in the ExE, without substantial compensation by TET2 and TET3 in its absence.

Pairwise comparisons between *Tet1*$^{GT/GT}$ and either heterozygous or wild-type ExE identified 15 differentially expressed genes (false discovery rate (FDR) < 0.05) (Fig. 3c and Supplementary Tables 2 and 3). Heat map clustering showed that 12 of these were upregulated while only 3 (including *Tet1*) were downregulated in *Tet1*$^{GT/GT}$ samples (Fig. 3d). Although five genes appeared to be upregulated in wild-type relative to heterozygous ExE (FDR < 0.05), they were expressed at very low levels (RPKM < 5) and likely reflect transcriptional noise. Of the additional differentially expressed genes that were identified using a less stringent cutoff ($P < 0.001$), several were highly expressed metabolic genes (*Atp5l*, *Pfdn4*, *Polr2k*, *Rpa3*, *Rpl22l1*, *Rps27l* and *Uqcrb*) that were upregulated by 1.5- to 2-fold in *Tet1*$^{GT/GT}$ ExE (Supplementary Fig. 4a,b). On the other hand, *Dmkn*, *Efcc1* and *Xaf1* were silent in wild-type and heterozygous ExE but activated upon loss of TET1 (Fig. 3e and Supplementary Fig. 4b).

When the differentially expressed genes ($P < 0.05$) were subjected to Ingenuity Pathway Analysis (IPA), the top pathways were related to mitochondrial dysfunction and oxidative phosphorylation (Fig. 3f). Indeed, several components of the mitochondrial electron transport chain were upregulated by 1.5- to 2-fold in *Tet1*$^{GT/GT}$ ExE (Supplementary Fig. 4c and Supplementary Table 4). Although these fold changes in expression are moderate, the collective upregulation of multiple pathway components may amplify into a biologically significant effect.

A shift in metabolic dependence from glycolysis to oxidative respiration is a hallmark of epiblast differentiation[32]. The metabolic profiles of the ExE were until now unknown. To address whether the upregulated genes in *Tet1*$^{GT/GT}$ ExE might represent an altered metabolic and differentiation status, we used trophoblast stem cell (TSC) lines derived from E6.5 ExE[33]. The epigenetic state of TSCs differs markedly from that of the ExE[34], including low levels of expression of all TET genes upon adaptation to culture (Supplementary Fig. 4d); nonetheless, these cells are the closest available *in vitro* correlates for the ExE. We

observed that the differentially expressed genes upregulated in *Tet1*$^{GT/GT}$ ExE were also upregulated during TSC differentiation by growth factor withdrawal (Fig. 3g). Moreover, differentiation was associated with increased mitochondrial respiratory capacity (Fig. 3h). Collectively, these results suggest that subsistence levels of *Tet1* in the ExE regulate the balance between the glycolytic and oxidative metabolic pathways, the loss of which triggers a metabolic shift favoring precocious differentiation.

We then examined the methylation status at the promoters of 13 differentially expressed genes by high-throughput bisulfite sequencing. At four loci (*Xaf1*, *Efcc1*, *Rpa3* and *Pet100*), loss of methylation was associated with gain of gene expression in *Tet1*$^{GT/GT}$ ExE (Fig. 3i,j and Supplementary Fig. 4e,f). A striking example was *Xaf1*, where promoter methylation was drastically reduced from 60.3% to 1.7% by loss of TET1. Interestingly, both *Xaf1* and *Efcc1* were silent in wild-type but activated in *Tet1*$^{GT/GT}$ ExE, suggesting that derepression of genes in TET1-deficient ExE may involve erasure of methylation. Moreover, the *Tet1* enhancer site also lost methylation in *Tet1*$^{GT/GT}$ ExE (Fig. 3k), whereas the same locus was 100% methylated in *Tet1*$^{GT/GT}$ Epi isolated from the same pool of embryos (Fig. 2i). On the other hand, loss of TET1 increased the gene expression levels of *Aph1b* and *Uqcrb* along with a gain in promoter methylation, but had no effect on the remaining six loci (Supplementary Fig. 4g,h). The lack of detectable 5hmC in the ExE (Fig. 1d) implies that methylation counts can be largely attributed to 5mC. Thus, our results suggest a contribution of TET1 in basal 5mC maintenance at specific sites in the ExE, which invokes a mechanism of action opposite to that more prevalent in the Epi, that is, erasure of 5mC.

TET1 regulates the DNA methylome in epiblast-like cells

The highly detectable *Tet1* and genomic 5hmC in the postimplantation epiblast prompted us to examine further global TET1 loss-of-function effects on the transcriptome and methylome in the epiblast lineage. To facilitate these studies, we used EpiLCs *in vitro* by first deriving wild-type, heterozygous and *Tet1*$^{GT/GT}$ (referred to hereafter as ‘knockout’; Supplementary Fig. 5a,b) ESC lines from individual blastocysts of incipient congenic (N5 or >95% B6) *Tet1* GT intercrosses. All lines (in triplicate per genotype) efficiently differentiated to EpiLCs. As previously shown, the omission of activin A did not affect naive-to-primed conversion[35]. However, loss of TET led to a notable reduction in total 5hmC content in EpiLCs, as measured by dot blots and ELISA (Fig. 4a,b). These differences between knockouts and controls (wild type and heterozygous) were less prominent when ESCs were maintained in serum and LIF (SL). The loss of 5hmC in knockout EpiLCs had minor effects on global 5mC content during naive-to-primed transition (Supplementary Fig. 5c,d). These observations suggest that loss of TET1 may affect the methylome landscape in primed pluripotent cells.

To profile the methylome of EpiLC lines, we first performed 5hmC and 5mC methylated DNA immunoprecipitation (hMeDIP and MeDIP, respectively) (Supplementary Note). By this approach, we identified 11,659 regions with differential 5hmC signals in knockout as compared to wild-type (or heterozygous) EpiLCs, of which 71% exhibited loss in knockouts, as well as 829 of 976 regions (85%) with enriched 5mC signals in knockouts, of which 32% resided within CGI regions (Supplementary Fig. 5e–i). Notably, these changes in *Tet1*-

knockout EpiLCs were largely not inherited from pre-existing differences between knockout and wild-type ESCs cultured in 2iL (Supplementary Fig. 5j–l), suggesting that TET1 continues to regulate the DNA methylome after the transition to primed pluripotency.

Therefore, we performed whole-genome bisulfite sequencing (WGBS) in parallel bisulfite and oxidative bisulfite[36] converted libraries to obtain base-resolution profiles of 5mC and 5hmC in wild-type and knockout EpiLCs (two independent lines each) (Supplementary Fig. 6a and Supplementary Note). More than 80% of methylated cytosines were found at CpGs (Supplementary Fig. 6b). In knockouts, methylated CpGs appeared to be redistributed from highly methylated regions with 50–100% methylation to lowly methylated regions with 10–50% methylation; the latter are known to be enriched for CpG-poor distal regulatory regions[37] (Fig. 4c and Supplementary Fig. 6c,d).

Although recent studies have shown a globally hypomethylated genome in female as compared to male ESCs[8,20], these differences are less pronounced in EpiLCs[38]. Because two of our knockout lines were female while all control lines were male, our dot blot and MeDIP analyses likely underestimate gains of methylation in knockouts. Supporting this deduction, we identified several differentially methylated regions (DMRs), >90% of which had elevated DNA methylation levels, in the male knockout line (line 12) in comparison to a male wild-type line (line 15) (Fig. 4d, Supplementary Table 5 and Supplementary Data 1–4). DMRs spanned 1–2 kb and predominantly overlapped gene body (exons and introns) and promoter regions (Fig. 4e). Even in female knockout EpiLCs, 5mCpG levels were marginally elevated at CGIs despite global loss at other genomic regions (Supplementary Fig. 6e), consistent with our MeDIP results. Therefore, TET1 regulates 5mC patterns at specific genomic features in EpiLCs, in conformity with its DNA demethylation function.

We then identified sites with 5hmC in EpiLCs and estimated that 80–90% of these resided within CpGs (Supplementary Fig. 6f), in agreement with previous base-resolution 5hmC profiling in ESCs[39]. An average of 502,874 CpGs in wild-type and 384,633 in knockout EpiLCs were found to contain 5hmC; the majority exhibited 5hmC levels of 20–60% occupancy and coincided with fully methylated (90–100%) sites profiled by bisulfite sequencing (Supplementary Fig. 6g,h). On a global scale, these 5hmC levels may correspond to 5hmC at only ~0.04% of all cytosines, barely above the detection limit of mass spectrometry; nonetheless, these levels can be biologically critical when localized at regulatory sites. When global 5hmC levels were distributed according to annotated genomic feature, depressed levels were discernible in knockout EpiLCs at CGIs (Supplementary Fig. 6i).

To determine whether differentially hydroxymethylated CpGs in knockout EpiLCs were concentrated at TET1-bound regions, we performed ChIP–seq using an anti-TET1 antibody. Interestingly, we observed a concentration of TET1 occupancy sites within promoter regions in EpiLCs, to a greater extent than previously determined in serum-cultured mouse ESCs[18] (Fig. 4f). Globally, TET1-bound regions are centered at unmethylated sites, in line with demethylating activity (Fig. 4g and Supplementary Fig. 6j). These regions exhibited elevated 5mC and reduced 5hmC levels in knockouts, where 5mC-to-5hmC conversion at peak centers were evident in wild-type but absent in knockout EpiLCs (Fig. 4h–j). Collectively,

our results suggest that TET1 regulates DNA methylation via hydroxymethylation at gene-proximal promoters in the postimplantation epiblast.

Loss of TET1 affects gene expression independently of methylation changes

We next asked how the impact of TET1 on the EpiLC methylome influences gene expression. By RNA–seq, we identified 613 differentially expressed genes (FDR < 0.05) in knockout versus wild type (DE1), 133 in knockout versus heterozygous (DE2) and only 3 in heterozygous versus wild type (DE3) pairwise comparisons (Fig. 5a, Supplementary Tables 6 and 7, and Supplementary Note). When the 613 genes in DE1 were overlapped with the TET1-bound regions defined by ChIP–seq, we observed that ~40% were bound by TET1 (Fig. 5b). The majority (~72%; 441 genes) gained expression in *Tet1*-knockout EpiLCs (Fig. 5c). Gene ontology (GO) analysis of DE1 upregulated genes identified the MHC class I protein complex and structural molecule activity as the only significantly enriched terms; for downregulated genes, the only significant term was nucleic acid binding (Supplementary Fig. 7a). Over 60% of the differentially expressed genes were not associated with oxWGBS-defined DMRs (Fig. 5c).

GO and KEGG enrichment analyses of WGBS DMRs showed highly significant enrichment of terms for diverse biological processes and diseases (Supplementary Fig. 7b,c), in stark contrast to the limited GO terms from differential gene expression analysis. These results suggest that, while loss of TET1 affects the DNA methylome in EpiLCs, these effects do not yet affect expression of over 90% of the associated genes (Fig. 5c).

TET1 regulates telomere stability in epiblast-like cells

We focused on 95 differentially expressed genes commonly dysregulated in knockout versus wild-type or heterozygous EpiLCs. Almost 70% of these gained expression in *Tet1*-knockout EpiLCs (Fig. 5a,d and Supplementary Tables 6 and 7). The smaller set of genes downregulated by *Tet1* knockout included the entire *Zscan4a–Zscan4f* cluster. ZSCAN4 is known to be critical for telomere maintenance, as a means to maintain karyotype integrity in ESCs[40,41]. Because *Zscan4c*, *Zscan4d* and *Zscan4f* were previously shown to be upregulated in TET triple-knockout serum-cultured ESCs[42], we measured their expression in cells cultured under serum and LIF (SL), 2iL and EpiLC conditions to reconcile this discrepancy. We observed that ZSCAN4 gene expression was indeed increased in knockout ESCs from the SL condition but was decreased under the 2iL condition and in EpiLCs (Supplementary Fig. 8a), suggesting that this gene cluster can be dynamically regulated by TET1 in different cell states.

To investigate a potential defect in telomere homeostasis in *Tet1*-knockout cells, we performed flow cytometry FISH (flow-FISH), which uses a fluorescent probe to measure telomere length in cells in G1 phase of the cell cycle. This method found a significant increase in the proportion of cells harboring shortened telomeres in all three knockout lines upon conversion to EpiLCs (Fig. 5e,f). The defect was specific to primed cells, as naive knockout ESCs in 2iL showed no detectable fraction with shortened telomeres (Supplementary Fig. 8b). We note that others have recently shown a role for TET enzymes in regulating telomere length, but these effects were only noticeable in TET double- or

triple-knockout ESCs[42,43]. Together, these results suggest that fluctuations in ZSCAN4 levels caused by loss of TET1 in naive pluripotent states may have no measurable impact on telomere stability but can be deleterious in primed pluripotency.

To detect the presence of 5hmC in the ZSCAN4 cluster, we performed amplicon sequencing following TET-assisted bisulfite sequencing (TAB)[39] (Supplementary Note). At a unique TET1-bound site downstream of *Zscan4f*, we identified CpG sites that had 10–30% 5hmC levels in wild-type EpiLCs but significantly reduced levels in knockout EpiLCs (Fig. 5g). These results provide evidence for TET1 catalytic activity at the ZSCAN4 locus, together with a role in telomere maintenance, during the transition to primed pluripotency.

TET1 exerts catalysis-independent repression of target genes in epiblast-like cells

While TET1 is expected to promote DNA demethylation, we were surprised that the majority of differentially expressed genes were upregulated in knockout Epi, EpiLCs and ExE. This may imply a gene regulatory mode of TET1 independent of 5mC oxidation, in line with a previous study suggesting gene regulation by TET1 even in methylation-deficient ESCs[18]. We therefore focused on the most biologically relevant differentially expressed genes identified *in vivo*, which can be further analyzed *in vitro* using EpiLCs, by overlapping the differentially expressed data sets of EpiLCs and Epi.

Of the ten genes in this shortlist, eight were upregulated, including *Bin1*, *Cd44*, *Tnfrsf1a* and two RAS family oncogenes (Fig. 6a). TET1 ChIP–qPCR validated that they were bound by TET1 in EpiLCs (Fig. 6b and Supplementary Fig. 9a). Methylation changes at CpG sites flanking the promoters of these genes suggest that TET1 is indeed exerting DNA demethylating activity at its targets (Supplementary Fig. 9a). To resolve the paradox of why loss of TET1 would increase gene expression despite gain of 5mC at promoters, we performed TET1 rescue in a knockout line to demonstrate whether its repressive function involves 5mC oxidation. Re-expression of TET1, either as a full-length wild-type transcript or a catalytically dead mutant (Supplementary Fig. 9b–d), potently abrogated upregulation of *Bin1*, *Tnfrsf1a*, *Rab20*, *Ppp1r1a*, *Kirrel* and *Rasl10b* (Fig. 6c) in *Tet1*-knockout EpiLCs. These results are strongly suggestive of TET1 catalysis-independent activity that represses target genes in the primed epiblast.

Strain-specific modifiers of the *Tet1*-knockout phenotype

To determine why mixed-strain and outbred *Tet1* knockout showed full embryonic lethality in comparison to partial lethality in B6 knockout (Supplementary Note), we examined wild-type and knockout CD1 EpiLCs converted from ESC lines (two males and one female per group), derived from *Tet1* GT mice after three generations of outbreeding. In the genetically heterogeneous background, RNA–seq identified 48 differentially expressed genes (FDR < 0.05). Heat map cluster showed again that the majority (~75%) were upregulated, independently of sex differences (Supplementary Fig. 10a). *Bin1*, *Cd44*, *Kirrel* and *Tnfrsf1a* were differentially expressed genes common in CD1 and B6 knockouts (Supplementary Fig. 10b), which showed enhanced differential expression in CD1 relative to B6 (Fig. 6d). Moreover, GO analysis of differentially expressed genes in CD1 knockouts revealed several terms (FDR < 1×10^{-4}) related to regulation of cell differentiation and development, none of

which were enriched in the analysis of B6 lines (Supplementary Fig. 10c). Together, these results provide molecular evidence that loss of TET1 affected genes involved in differentiation more potently in non-inbred than in inbred mice (Supplementary Note).

Jmjd8 is a TET1 target gene that encodes a transcriptional co-repressor

Interestingly, the two genes downregulated both in B6 *Tet1*-knockout Epi and EpiLCs included *Tet1* itself and *Jmjd8*, a relatively unknown Jumonji family member (Fig. 6a). We identified two TSS-proximal sites in *Jmjd8* that are bound by TET1 in EpiLCs (Fig. 6b). Expression of *Jmjd8* in *Tet1*$^{GT/GT}$ Epi was diminished to basal levels in comparison to wild-type and heterozygous controls (Fig. 7a). In addition, *Jmjd8* downregulation *in vivo* was associated with significant gains of methylation at two sites upstream of the TSS (Fig. 7b). These results implicate *Jmjd8* as a direct target of TET1 activity in the epiblast, prompting us to investigate its biological function further.

Jmjd8 belongs to the family of iron- and 2-oxoglutarate-dependent chromatin-modifying enzymes that includes TET DNA oxygenases and histone lysine demethylases. However, like the founder *Jumonji* member, JARID2, the JmjC domain of JMJD8 contains catalysis-mutating substitutions[44]. Supporting the possibility that JMJD8 may instead be a transcriptional repressor like JARID2 (ref. 45), we observed that JMJD8 repressed both the mammalian *thymidine kinase* and viral SV40 promoters in dual-reporter assays (data not shown). This repressive activity was validated either by using a promoter-less reporter to normalize for transfection (Fig. 7c) or sorting transfected cells (Fig. 7d). Surprisingly, immunofluorescence detection of exogenous tagged protein showed full-length JMJD8 to be expressed predominantly in the cytosol (Fig. 7e). However, removal of a predicted N-terminal signal peptide (Fig. 7f) allowed expression of a truncated isoform (ΔN-JMJD8) in both cytosol and nuclei (Fig. 7e and Supplementary Fig. 11a).

To assess whether ΔN-JMJD8 can bind target genes, we performed ChIP–seq using wild-type EpiLCs overexpressing tandem-tagged ΔN-JMJD8 (Supplementary Fig. 11b). This analysis showed colocalization of nuclear JMJD8 at TET1-bound sites (Fig. 7g,h and Supplementary Fig. 11c). Binding to the *Bin1* and *Tnfrsf1a* loci is likely to be functionally relevant, as these two genes were down-regulated in knockout EpiLCs upon overexpression of ΔN-JMJD8 (Fig. 7i and Supplementary Fig. 11d). This indicates that the repressive effects of ΔN-JMJD8 override the demethylating activity of TET1 at these genes, providing a functional link between the catalytic and non-catalytic activities of TET1.

DISCUSSION

Here we demonstrate non-redundant functions of *Tet1* at an early postimplantation stage and explain why previous *Tet1*-mutant mice, which ablated TET1 catalytic function, had little or no defects *in utero*[24,25,46]. In two alternative *Tet1*-null strains in which the 5′ coding sequence is also ablated, we observed severe post-gastrulation defects, suggesting that, while the catalytic domain of TET1 can be dispensable, further regulation by its N terminus is critical for gastrulation. In agreement, TET1 represses several target genes in the epiblast without requiring its catalytic activity. Others have shown that TET1 interacts with SIN3A[18] or Polycomb co-repressors in ESCs[31]. Here we identify JMJD8 as a new TET1 co-repressor.

Surprisingly, few TET1-bound genes rely on DNA demethylation for expression. We identified two of these as *Jmjd8* and *Lefty2*; both coincidentally encode repressors of biological processes, in antithesis to the expected role for DNA demethylation to activate gene expression programs. Rather, our results implicate a role for TET1 as a safeguard against precocious differentiation of embryos at pre-gastrulation. Modest dysregulation of genes (*Bin1*, *Cd44* and *Kirrel*) driving progression of differentiation may be tolerated in inbred embryos but more pronounced dysregulation becomes deleterious in non-inbred ones, which may account for the strain dependence of knockout phenotypes.

The maintenance of DNA methylation is essential for the viability of primed pluripotent cells but dispensable in naive cells[47]. We found TET1 to be critical for preventing telomere shortening in EpiLCs, likely by regulating *Zscan4*, which raises the possibility that telomere maintenance requires regulation of DNA methylation in primed pluripotency. This impact on genome stability may be pertinent to a wide spectrum of diseases[24,46,48] (Supplementary Note). In conclusion, our study highlights critical roles for TET1 in non-catalytic gene regulation during primed pluripotency, whereas its catalytic effects on the DNA methylome may serve independently as an epigenetic safeguard against dysregulation later in postnatal development. The involvement of TET1 in the extraembryonic methylome opens new avenues to understand the epigenetic regulation of placental development.

URLs

qPCR protocol to assess gene dosage in GT mice, http://www.genetrap.org/info/protocols/baygenomics/QPCR_neo_Gene_dosage.pdf; WISH protocol, http://goodrich.med.harvard.edu/uploads/3/7/7/1/37718659/whole_mount_in_situ.pdf; Venn diagram generator, http://genevenn.sourceforge.net/; calculation of methylation density, http://www.bioconductor.org/packages/2.13/bioc/html/bsseq.html.

ONLINE METHODS

Characterization of the *Tet1* gene-trap allele

The *Tet1*$^{GT/wt}$ mouse line was generated by blastocyst injection of gene-trap ESC clone RRG140 (BayGenomics). Southern blot analysis was performed as previously described[49]. To map the gene-trap (GT) insertion precisely, genomic DNA was sequenced by targeted locus amplification (Cergentis) as previously described[50]. SYBR Green–based quantitative measurements of *lacZ* gene copies were normalized to those of *Ngf* and calibrated to a targeted cell line harboring one allelic copy of *lacZ* (see URLs for the qPCR protocol).

C57BL/6 *Tet1*^{tm1Koh} mouse creation

Vector design and construction, gene targeting in B6 ESCs, blastocyst injection and breeding of chimeras to B6 mice using goGermline technology[51] were performed at Ozgene. In this mouse model, a *lacZ* reporter gene is introduced immediately after the ATG start codon, flanked by two *loxP* sites, in a cassette that allows reporter exchange by Cre-mediated recombination (Supplementary Fig. 2g).

Isolation of embryos

To collect preimplantation embryos, 4- to 5-week-old wild-type female mice were superovulated by intraperitoneal injection with 5 IU pregnant mare serum gonadotropin (Sigma) followed by 7.5 IU human chorionic gonadotropin (Sigma) 46–48 h later and were then mated with *Tet1*$^{GT/wt}$ males. Embryos were collected at daily intervals at 1.5 d (eight-cell stage), 2.5 d (morula stage) and 3.5 d (blastocyst stage) post coitum by flushing the oviduct (for the eight-cell and morula stages) or the uterus (for blastocysts). Postimplantation embryos from E5.5 to E10.5 were isolated from naturally mated timed-pregnant mice. Upon isolation of intact postimplantation conceptuses from decidua, the outer Reichert's membranes with attached parietal endoderm and trophoblast giant cells were removed with the aid of fine tungsten needles. To prepare E6.25 Epi and ExE tissues for RNA–seq or bisulfite sequencing, embryos free of the Reichert's membrane were incubated for 5–10 min at 4 °C in 2.5% pancreatin/0.5% trypsin in Ca^{2+}/Mg^{2+}-free Tyrode Ringer's saline solution and then washed in PBS to aid removal of the VE by mechanical dissection. Epi and ExE tissues were subsequently separated; interface tissues were removed to minimize cross-contaminating cell types. To genotype embryos, cells of the ectoplacental cone, visceral endoderm layers or yolk sac were lysed at 55 °C overnight in lysis buffer (10 mM Tris-HCl pH 8.3, 50 mM KCl, 2 mM $MgCl_2$, 0.45% NP40 and 0.45% Tween-20) supplemented with 1 μg/ml proteinase K. Digestion was stopped by incubation for 10 min at 98 °C before PCR. The primers used for genotyping are listed in Supplementary Table 8. All experimental procedures on mice have been reviewed and approved by the KU Leuven Ethical Committee for Animal Experimentation (P177/2011) in compliance with European Directive 2010/63/EU.

Whole-mount *in situ* hybridization

Embryos were fixed overnight at 4 °C in 4% paraformaldehyde (PFA) in PBS, dehydrated in increasing methanol concentrations in PBST (PBS + 0.1% Tween-20) and stored in 100% methanol at −20 °C until further processing, according to a standard online protocol with minor modifications during the staining steps (see URLs for the protocol). Briefly, to detect the hybridized signal, embryos were washed with MABT (100 mM maleic acid, 150 mM NaCl, pH 7.5, containing 0.1% Tween-20), pre-blocked in 20% serum, 2% Boehringer Blocking Reagent (Roche) at room temperature for 90–120 min and then incubated overnight at 4 °C with a 1:2,000 dilution of an alkaline phosphatase (AP)-conjugated anti-digoxigenin antibody (Roche, 11093274910) in blocking solution. After several washes in MABT of 1 h each, an additional overnight wash at 4 °C was performed. To limit the background signal and internal phosphatase activity of E9.5–E10.5 embryos, the embryos were washed three times for 15 min each with 2 mM levamisole (Sigma) in MABT. Detection of staining was performed using the BM purple AP substrate (Roche, 11442074001) for 6–24 h, depending on the probe. The embryos were post-fixed with 4% PFA overnight at 4 °C and stored in 70% ethanol. Imaging was performed using an S8 APO stereomicroscope (Leica).

Sense (negative-control) and antisense RNA probes for *Tet1*, *Tet2* and *Tet3* were generated by PCR using primers flanked with SP6 or T7 sequences. Digoxigenin labeling was performed using the DIG RNA Labeling kit and T3 RNA polymerase (Roche). Primer

sequences are listed in Supplementary Table 8. The *Lefty2* probe was kindly offered by the laboratory of H. Hamada (Osaka University)[52]. More than five embryos from at least three different litters were stained per embryonic stage, per genotype (wild type or *Tet1*$^{Gt/Gt}$) and per probe.

X-gal staining

Preimplantation embryos and whole-mount E5.5–E10.5 embryos were fixed in 0.2% glutaraldehyde (Sigma), 2 mM $MgCl_2$, 5 mM EGTA and 0.02% NP-40 in PBS at room temperature for 15–30 min (http://www.sanger.ac.uk/genetrap/). Fixation was followed by three washes in 2 mM $MgCl_2$, 0.02% NP-40 and 0.01% sodium deoxycholate in PBS. Staining for β-galactosidase was performed in 1 mg/ml X-gal (5-bromo-4-chloro-3-indolyl-β-_-galactopyranoside) (Fermentas), 5 mM potassium ferrocyanide, 5 mM potassium ferricyanide, 2 mM $MgCl_2$, 0.02% NP-40 and 0.01% sodium deoxycholate in PBS at 37 °C for 6 h (*Tet1* GT preimplantation and E6.5 *Tet1*$^{lacZ/wt}$ embryos) or overnight (*Tet1* GT postimplantation embryos). Postimplantation embryos were post-fixed in 4% PFA overnight at 4 °C and stored in 70% ethanol at 4 °C until further use in paraffin sectioning. Preimplantation embryos were imaged using a CKX41 Olympus microscope. For optimal whole-mount images, postimplantation embryos were further cleared in increasing concentrations of glycerol (25% to 100%) and imaged using an S8 APO stereomicroscope (Leica). For each embryonic stage, at least six embryos from more than four different litters were analyzed.

Paraffin embedding and sectioning

X-gal-stained embryos were positioned properly in 1.5% agarose equilibrated at 60 °C using Dumont 5 forceps (Roboz Surgical Instruments). The embryos were then further dehydrated and embedded in paraffin using an automated Excelsior ES system (Thermo Scientific). 5-μm sections were cut using a Microm HM360 microtome and were mounted and counterstained with Mayer's hematoxylin and eosin using standard protocols. Stained sections were imaged using an upright CTR5500B microscope (Leica).

5hmC immunohistochemistry

E6.5–E7.5 (within intact decidua) and E8.5 embryos were fixed in 4% PFA and paraffin embedded. At least 3–5 wild-type embryos per stage were sectioned and further stained for 5hmC. Briefly, tissue sections were dewaxed in xylene two times for 10 min each and rehydrated in a graded series of decreasing ethanol concentration (100% to 50%). To expose the 5hmC epitope, sections were treated with 2 N HCl for 30 min and then neutralized in 100 mM Tris-HCl (pH 8.5) for 10 min. Immunohistochemical detection was performed using tyramide signal amplification (TSA) according to kit instructions (PerkinElmer, NEL 700). We used as primary antibody an anti-5hmC rabbit polyclonal antibody (Active Motif, 39769; 1:5,000 dilution) and as secondary antibody a donkey anti-rabbit biotin-SP antibody (Lucron Bioproducts, 711-065-152; 1:300 dilution). Diaminobenzoate chromagen (Dako) was used for signal visualization. Sections were finally counterstained with Harris hematoxylin (Sigma), dehydrated, mounted in DPX mounting medium (Sigma) and imaged using a Leica CTR5500B microscope.

Genomic 5mC and 5hmC quantification

For dot blot assays, 250 ng of DNA was serially diluted twofold in TE buffer, denatured in 0.4 M NaOH/10 mM EDTA at 95 °C for 10 min, neutralized in an equal volume of ice-cold 2 M ammonium acetate and kept on ice for 10 min. Samples were subsequently spotted onto a Hybond+ Zeta-Probe nylon membrane (Bio-Rad) using a 96-well Bio-Dot apparatus (Bio-Rad). The spotted membrane was rinsed in 2× SSC, air-dried for 5–10 min and UV cross-linked two times at 120,000 μJ/cm^2 and was then blocked in 5% nonfat milk in TBST for 1 h before incubation overnight at 4 °C with anti-5hmC (Active Motif, 39769; 1:10,000 dilution) or anti-5mC (Active Motif, 39649; 1:5,000 dilution) antibody. Detection was performed using an HRP-conjugated anti-rabbit (Dako, P0217; 1:5,000 dilution) or anti-mouse (Dako, P0447; 1:5,000 dilution) secondary antibody, and signal was developed using an enhanced chemiluminescence reagent (Bio-Rad). On every membrane, a serial dilution of DNA containing known amounts of 5hmC or 5mC was spotted in parallel as standards. Signal intensities were measured using ImageJ software and calibrated against the linear range of the standard curves to estimate the quantities of 5hmC and 5mC in each sample.

5hmC ELISA was performed using the Global 5-hmC Quantification kit (Active Motif). For mass spectrometry, genomic DNA was extracted using the Promega Wizard kit, treated with RNase A (NEB) and purified using the DNeasy Blood and Tissue kit (Qiagen). Purified DNA was hydrolyzed using the DNA degradase plus kit (Zymo). Analysis of DNA hydrolysates was performed according to Bachman *et al.*[53].

Derivation of ESC lines

E3.5 blastocysts were collected from *Tet1*$^{GT/wt}$ superovulated females mated with *Tet1*$^{GT/wt}$ males (>95% B6 background or after three outcrosses to CD1 mice). ESC lines were derived from blastocyst in defined medium containing 2iL as previously described[54]. From the B6 incipient congenic embryos, we cultured 20 blastocysts separately and generated a total of 16 independent cell lines: 5 wild type, 8 heterozygous and 3 *Tet1*$^{GT/GT}$ (knockout). From CD1 embryos, we derived 37 independent cell lines from 41 blastocysts: 7 wild type, 22 heterozygous and 7 *Tet1*$^{GT/GT}$ (knockout). Immunoblot analysis was performed on whole-cell lysates using an anti-TET1 antibody (Millipore, 09-872; 1:1,000 dilution). Karyotyping was performed by mitotic chromosome counting as previously described[54], using Wright's staining instead of DAPI.

Cell culture and differentiation

All mouse ESC lines were cultured in standard ESC culture medium as previously described[16]. When collecting feeder-cultured cells, cells were depleted of feeder cells by plating them twice for 30–45 min on tissue culture plates to remove adherent cells. Differentiation toward EpiLCs was performed as previously described with bFGF (Peprotech) but without activin A supplementation[35]. TSCs, derived from E6.5 embryos and generously provided by J. Rossant (University of Toronto), were cultured on mitomycin-inactivated fibroblast feeders in TSC medium containing RPMI 1640, 20% ES-qualified FBS, 25 ng/ml recombinant FGF4 (Peprotech), 1 ng/ml heparin (Sigma-Aldrich), 2 mM L-glutamine, 0.1 mM βmercaptoethanol, 1 mM sodium pyruvate and 50 units/ml penicillin-streptomycin (Invitrogen). Afterward, cells were passaged 1:10 or 1:20 on feeder-free

gelatin-coated plates and cultured for 4 d with 70% feeder-conditioned medium in the presence or absence of FGF4 and heparin.

Gene expression analysis3

Total RNA was extracted using the RNeasy Plus Mini kit (Qiagen) for cell lines or the RNeasy micro kit (Qiagen) for embryos and converted to cDNA using the SuperScript III First-Strand Synthesis SuperMix kit (Life Technologies). SYBR Green–based qPCR (Life Technologies) was performed on either a ViiA 7 Real-Time PCR System or a Step One Plus Real-Time PCR System (Applied Biosystems). Primer sequences are listed in Supplementary Table 8.

MeDIP–seq and hMeDIP–seq

Genomic DNA was extracted from ESCs after five passages in 2iL and upon conversion to EpiLCs as previously described[55]. Three independent biological replicates were analyzed. For MeDIP and hMeDIP, anti-5mC (Eurogentec, BI-MECY-0100) and anti-5hmC (Active Motif, 39791) antibodies were used, respectively. Single ends from these libraries were sequenced on an Illumina HiSeq 2500 instrument with a v4 short-read flow cell (Illumina), yielding ~55 million reads of 50 bp in length per sample (range 40–90 million). Reads were mapped to mouse genome build GCRm38 using Bowtie 0.12.8 (-v 2 -m 1 --best -S). Mapping efficiencies were ~70% for input and hMeDIP and ~40% for MeDIP, thus matching previously reported values[17].

5mC and 5hmC DNA immunoprecipitation enrichment peak calling

Reads were extended *in silico* to 220 bp, to match the average insert sizes determined by Bioanalyzer HiSensitivity chips (Illumina). Areas of enriched MeDIP or hMeDIP signal (5mC and 5hmC peaks) were called in SeqMonk 1.0 by applying MACS[56], using as input store a group of input data (combining the reads from all the individual input data sets) and as ChIP store a group of MeDIP or hMeDIP data (combining reads from all individual MeDIP and hMeDIP data sets, respectively). At a *P*-value cutoff of 1×10^{-5}, 241,137 MeDIP and 247,769 hMeDIP peaks were called from EpiLC data sets. Peak areas were filtered for having a relatively equivalent number of reads mapping to the top and bottom strands (maximum of threefold difference) to remove potential mapping and DNA immunoprecipitation artifacts[57,58]. Unsupervised hierarchical clustering was performed using the Pearson distance and Ward method. Principal-component analysis (PCA) plots and heat maps were generated for the top 1,000 most differentially (hydroxy)methylated regions. Differential (hydroxy)methylation within the filtered peaks was analyzed using EdgeR[59] and called at FDRs below 5% (FDR < 0.05) or 10% (FDR < 0.1).

RNA sequencing

RNA from the EpiLCs of three independent lines per genotype was isolated using the TRIzol method (Invitrogen). 4 μg of total RNA was used for library preparation with the KAPA stranded mRNA Library prep kit for Illumina platform (KAPA Biosystems). Library concentrations were normalized on the basis of qPCR, and equimolar amounts were pooled for single-end sequencing on an Illumina HiSeq 2500 instrument with a v4 short-read flow

cell (Illumina) to yield ~44 million reads (range 41–45 million) of 50 bp in length per sample. PCR adaptors were trimmed using FastX (0.013).

In the low-input protocol, RNA was purified using the RNeasy micro kit (Qiagen) to obtain 1–10 ng per pool. First-strand cDNA synthesis and amplification were performed using SMARTer Ultra-Low-Input RNA amplification kit v3 (Clontech). Single-end libraries were prepared using NEBNext Ultra DNA library prep for Illumina and sequenced on a NextSeq 500 instrument, yielding ~11 million reads (range 8–13 million) of 75 bp in length. Low-quality ends and adaptor sequences were trimmed from the Illumina reads with FastX 0.0.13 and cutadapt 1.7.1. Using FastX and ShortRead 1.16.3, we subsequently filtered out short reads (length <35 bp), poly(A) reads (where >90% of the bases are adenine), ambiguous reads (containing Ns) and low-quality reads (where >50% of the bases have quality <Q25).

Differential expression analysis

Reads from all data sets (EpiLC, Epi and ExE) were aligned using TopHat v2.0 with default parameter options (--library-type unstranded --min-intron-length 70 --maximum-intron-length 500000) and mapped to mouse genome build GRCm38/mm10. The mapping efficiencies of the EpiLC data sets were >98%, and those of the *in vivo* Epi and ExE data sets were >85%. SAMtools 1.1 were used to generate aligned BAM files that served as input for read counting using the summarizeOverlaps count function in the R Bioconductor package GenomicAlignments. Genes with no reads in all samples were filtered out. The features were counted for single-end reads using the union method in the function, yielding an average of 34 million (78.07% of processed reads), 7.7 million (69.32%) and 7.9 million (69.61%) reads mapping as unique or primary in the EpiLC, Epi and ExE data sets, respectively. A read count matrix of these mapped reads was used as input for differential expression analysis using the DESeq2 package in R. The resulting matrix was used for PCA with the top 1,000 differentially expressed genes. The P values obtained from the DESeq2 analysis were corrected for multiple testing with the Benjamini–Hochberg method to control the FDR. Genes were considered significantly differentially regulated on the basis of either FDR < 0.05 or $P < 0.001$. Venn diagrams were generated using an online tool (see URLs). Differential expression was visualized as heat maps generated in R using the heatmap.2 function. Hierarchical clustering was performed using a Pearson correlation matrix. RPKM values were obtained by aligning mapped reads to mm10 using Cufflinks (v2.2.1).

Bisulfite amplicon sequencing analysis

Epi and ExE tissues from E6.25 embryos were pooled by genotype (3–5 per pool) and lysed according to the EpiTect Fast Bisulfite Conversion kit instructions (Qiagen). For *Zscan4* locus methylation and hydroxymethylation analysis, genomic DNA extracted from EpiLCs was converted using the 5hmC TAB-Seq kit (WiseGene). Bisulfite conversion was performed according to the EpiTect Fast Bisulfite Conversion kit instructions. Nested PCR products were generated using HiFi Taq polymerase (Invitrogen). Amplicons of 300–450 bp were subsequently purified for library preparation using the NEBNext Ultra DNA Library prep kit (Illumina), and libraries were sequenced on an Illumina MiSeq 600 instrument. Methylation quantification, for sequences with a depth of ≥1,000×, was performed using the CLC genomics workbench (Qiagen) accordingly to previously published pipelines[60] with an

additional step to remove duplicate sequences. For TAB–seq, 5mC and 5hmC controls were added as spiked-in controls in the samples as previously described[39]. The 5hmC protection by glucosylation was >80%, and the oxidation rate was >95%. For clonal sequencing (GATC biotech), purified PCR products were subcloned into the pGEM-T Easy vector (Promega). The CpG methylation profiles of sequenced clones were analyzed using QUMA[61]. Sequences with a conversion rate lower than 95%, sequence identity lower than 90% or identical bisulfite sequences were excluded. Primer sequences are listed in Supplementary Table 8.

Oxygen consumption rate measurement

4×10^4 TSCs cultured with and without FGF4 (25 ng/ml) and heparin (1 ng/ml) were seeded onto XFp Seahorse gelatine-coated wells 24 h before measurement for 90% confluency. Culture medium was replaced with XFp Seahorse base medium supplemented with 2 mM glutamine, 1 mM sodium pyruvate and 10 mM glucose 1 h before the assay, and this medium was kept for the duration of the measurement. Selective inhibitors were injected during measurement to achieve final concentrations of oligomycin (2.5 mM), FCCP (300 nM) and antimycin (2 mM), following the manufacturer's instructions. Measurement was performed using the Seahorse XFp cellular flux analyzer (Agilent). Measurements were normalized to total protein content per sample. The maximal respiration rate was calculated by subtracting oxygen consumption rate (OCR) levels measured at the time of oligomycin injection from the levels recorded at FCCP injection.

Whole-genome bisulfite and oxidative bisulfite sequencing

WGBS and oxWGBS were performed by Novogene. Briefly, 5 μg of genomic DNA per sample from two wild-type (lines 15 and 19) and two knockout (lines 2 and 12) EpiLC lines was fragmented by sonication to 200–300 bp in length and subjected to bisulfite or oxidative bisulfite treatment using the CEGX TrueMethyl Whole-Genome (TMWG) kit (Epigenetix) according to the manufacturer's instructions. The resulting single-stranded DNA fragments were PCR amplified using KAPA HiFi HotStart Uracil + ReadyMix (2×). Libraries were sequenced on an Illumina HiSeq X-10 platform to generate 150-bp paired-end reads. Sequencing of the bisulfite sequencing libraries generated 582, 525, 546 and 586 million reads for the wt19, wt15, KO2 and KO12 samples, respectively; the oxidative bisulfite sequencing libraries generated 508, 890, 514 and 604 million reads for the wt19, wt15, KO2 and KO12 samples, respectively.

Whole-genome bisulfite and oxidative bisulfite sequencing analysis

Adaptor sequences were trimmed from the Illumina reads. Subsequently, reads with ≥10% Ns were removed. In addition, reads with low quality (Phred score ≤5 and percentage of low-quality bases ≥50%) were discarded. The remaining reads were mapped to mouse genome build GRCm38/mm10 using Bismark software (version 0.12.5)[62] with default parameters. Reads aligned to the same regions of the genome were regarded as duplicates and counted once. The results from the methylation extractor were transformed into bigWig format for visualization using the IGV browser. A set of six control duplexes containing cytosine, 5hmC, 5mC and 5fC bases at known positions, provided in the TMWG kit, was used to estimate the bisulfite and oxidation conversion rates. In all oxWGBS libraries,

oxidation efficiency based on 5hmC-to-thymidine conversion was >99%. In all WGBS and oxWGBS libraries, bisulfite conversion rate based on cytosine-to-thymidine conversion was >99%.

Estimation of the methylation level

To identify the sites of methylation, the sum $\mathbf{s}^{+}_{i,j}$ of the methylated counts was modeled as a binomial random variable with methylation rate $\mathbf{r}_{i,j}$.

$$\mathbf{s}^{+}_{i,j} \sim Bin(\mathbf{s}^{+}_{i,j} + \mathbf{s}^{-}_{i,j}, \mathbf{r}_{i,j})$$

To calculate the methylation density, a sliding window approach was employed. This is conceptually similar to approaches that have been used for bulk bisulfite sequencing (see URLs). With window size w = 3,000 bp and step size = 600 bp[63], the sum of the methylated and unmethylated read counts in each window was calculated. The methylation level (ML) for each cytosine site corresponds to the fraction of methylated cytosines and is defined as

$$\mathrm{ML(corrected)} = \frac{\mathrm{ML} - r}{1 - r}$$

where r is the bisulfite non-conversion rate. The final 5hmC level is calculated and corrected for the 5hmC conversion rate.

$$\mathrm{level} = \frac{\left(\mathrm{cytosine}_{\mathrm{ratio}_{\mathrm{BS}}} - \mathrm{cytosine}_{\mathrm{ration}_{\mathrm{oxBS}}}\right)}{\mathrm{conversion\ rate}}$$

DMRs were identified using swDMR software with default parameters, setting the sliding window length to 1,000 bp and the step length to 100 bp. Fisher's test was implemented to detect DMRs[64]. The distribution of DMRs among genomic features was generated using the ChIPseeker package in R[65].

ChIP–qPCR and ChIP–seq

Cells were trypsinized and fixed in freshly prepared 1% PFA at room temperature for 10 min and quenched in 0.125 M glycine. Subsequently, cells were lysed with buffer I (0.25% Triton X-100, 0.5% NP-40, 10% glycerol, 10 mM EDTA, 140 mM NaCl, 10 mM HEPES-KOH, pH 7.5) and washed with buffer II (200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 10 mM Tris-HCl, pH 8.0) to remove detergents. Nuclei were then lysed in buffer III (100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% sodium deoxycholate, 0.5% *N*-lauroylsarcosine, 10 mM Tris-HCl pH 8.0, 0.1% SDS). Nuclear lysates were sheared to 200–500 bp using a Bioruptor Plus sonicator (Diagenode) set at high power for 20 cycles (30 s on, 30 s off). Sheared chromatin was immunoprecipitated using 3 μg (for ChIP–qPCR) or 10 μg (for ChIP–seq) of the appropriate antibody (anti-TET1: Millipore, 09-872; anti-6×His tag: Abcam, ab18184; anti-V5 tag: Abcam, ab9116) or the same amount of control IgG (Santa

Cruz Biotechnology, sc-2027) and incubated overnight with Protein G Dynabeads (Thermo Scientific, 10004D). Immunocomplexes were washed for 5 min each sequentially with the following buffers: low-salt buffer (20 mM Tris pH 8.1, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), high-salt buffer (20 mM Tris pH 8.1, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), LiCl buffer (10 mM Tris pH 8.1, 250 mM LiCl, 1 mM EDTA, 1% deoxycholate, 1% NP-40) and twice in 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 50 mM NaCl. Chromatin was eluted with elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS), and cross-linking was reversed by incubation overnight at 65 °C in 5 M NaCl buffer. The de-cross-linked DNA samples were incubated with DNase-free RNase at 37 °C for 30 min and afterward with 10 mg/ml proteinase K at 55 °C for 1–2 h. DNA was purified using the QIAquick PCR Purification kit (Qiagen, 28106) for qPCR or the Zymo ChIP DNA Clean & Concentrator kit (Zymo, D5201) for ChIP–seq library preparation. TET1 ChIP–seq was performed on two biological replicates (two wild-type and two *Tet1*-knockout EpiLC lines). V5 and His ChIP assays were performed on tandem-tagged ΔN-JMJD8-V5/His expressed in wild-type EpiLCs. Libraries were prepared using the NEBNext Ultra DNA Library Prep kit for Illumina and sequenced on a HiSeq 2500 instrument to generate 15–25 million 50-bp reads. FASTQ files were adaptor trimmed and mapped using Bowtie2 (ref. 66). Binding peaks relative to background (IgG) were called using MACS2 with default parameters[56]. BED files were further generated for peak annotation using the ChIPSeeker package in R[65]. ChIP peaks in specific target genes were visualized using the IGV browser. ChIP heat maps were generated using custom R script. ChIP–qPCR primer sequences are listed in Supplementary Table 8.

Telomere length assays

A telomere PNA kit with FITC-labeled probes for flow cytometry (Dako, K5327) was used following the manufacturer's protocol.

Cloning of JMJD8 overexpression vectors

The *Jmjd8* coding sequence was amplified using HiFi Taq polymerase (Invitrogen) and subcloned into the pOZ-FH-N-IRES-CD25 vector, including a 3′-end stop codon, between XhoI and NotI restriction sites. Subsequently, the FH-JMJD8-IRES-CD25 cassette was digested with BglII and BamHI and subcloned into the pEF1/V5-His vector between KpnI and BamHI sites to yield pEF1-FH-JMJD8-IRES-CD25. To add a C-terminal tag, the *Jmjd8* coding sequence was subcloned in frame with a sequence encoding a C-terminal V5-His tag within the pEF1 vector using the KpnI and NotI sites to generate pEF1-JMJD8-V5-His. We further generated a vector expressing truncated JMJD8 lacking the predicted N-terminal signal peptide by subcloning at the SpeI and NotI sites, to obtain pEF1-JMJD8-(27-272)-V5-His.

Luciferase reporter assays

A dual-reporter luciferase assay (Promega) was performed in HEK293T cells in which thymidine-kinase-promoter-driven *Renilla* luciferase vector (pRL-TK) was used as the reporter and the pro-moterless firefly luciferase vector (pGL3-Basic) was used as the normalizing control at a 1:5 ratio. To test the activity of JMJD8 and ΔN-JMJD8, 1 μg of the corresponding overexpression construct (pEF1-JMJD8-V5/His or pEF1-ΔN-JMJD8-V5/His)

or control (pEF1-mock) was cotransfected into the cells using Mirus TransIT-293T transfection reagent (Mirus Bio). In a second approach, we performed a single-reporter assay using SV40-promoter-driven luciferase plasmid (pGL3-SV40P) transfected together with pEF1 vector coexpressing JMJD8 with a human CD25 surface antigen (pEF1-FH-JMJD8-IRES-CD25) or mock (pEF1-IRES-CD25) into HEK293T cells. Forty-eight hours after transfection, cells were sorted using microbeads conjugated to mono-clonal anti-human CD25 antibodies (CD25 Microbeads II, Miltenyi Biotec) on an MS column (Myltenyl Bio). Luciferase assays were performed on the CD25-enriched populations; readings were further normalized to the total protein content of the cells determined by Bradford assay.

Immunofluorescence

HEK293T cells were transiently transfected with 1 μg of expression constructs for V5/His-tagged full-length (amino acids 1–271) or N-terminally truncated (amino acids 27–271; ΔN-JMJD8) JMJD8. JMJD8 was detected in fixed cells using anti-6×His antibody (Abcam, ab18184) and Alexa Fluor 488–conjugated secondary antibody (Thermo Fisher, a-11001). Images were captured using a Leica SP8x confocal microscope. Nuclear and cytoplasmic fractions were isolated from cells with standard protocols using a Dounce homogenizer. Standard immunoblotting was performed to detect JMJD8 using the anti-6×His antibody (Abcam, ab18184). Antibodies to nuclear lamin A (Santa Cruz Biotechnology, sc-20680) and tubulin (Cell Signaling Technology, 2144) were used to mark the nuclear and cytoplasmic fractions, respectively.

TET1 and JMJD8 overexpression cell lines

The ΔN-JMJD8 coding sequence was subcloned into the pPyCAGIP vector containing a puromycin resistance gene, and the resulting vector was electroporated into mouse ESCs. Puromycin-resistant clones were picked and characterized for expression of the His/V5 tag. For TET1 and JMJD8 rescue in knockout cell lines, the coding sequence for wild-type *Tet1*, catalytic-domain-mutant *Tet1* (mHxD) or ΔN-JMJD8 was subcloned into the piggyBac (PB) transposon vector under the control of the doxycycline-inducible *tetO* promoter. The PB expression plasmids were cotransfected with a tetracycline transactivator (rtTA)-IRES-hygromycin resistance gene vector and a vector encoding the transposase enzyme at a molar ratio of ~3:1:1 (PB-TET1 or PB-JMJD8:PB-rtTA:PBase-transposase). Hygromycin-resistant clones were selected, and the expression of the TET1 and ΔN-JMJD8 coding sequences was assayed by immunoblotting (in the presence or absence of 1 μg/ml doxycycline).

Statistical tests

Two-tailed paired or unpaired *t* tests were used with 95% confidence intervals. We applied Tukey's HSD (honest significant difference) test for multiple comparisons when ANOVA was used.

Data availability

RNA–seq data for Epi and ExE samples are accessible from the Gene Expression Omnibus (GEO) under accession GSE71744. RNA–seq, MeDIP–seq and hMeDIP-seq, and ChIP–seq data for EpiLC samples are available from ArrayExpress under accessions E-MTAB-4578,

E-MTAB-4581 and E-MTAB-5562, respectively. WGBS and oxWGBS data are accessible from ArrayExpress under accession E-MTAB-5568.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank G. Daley and A. Yabuuchi (Harvard Medical School) for generating the *Tet1* GT mouse. SMARTer RNA–seq and amplicon sequencing were performed by K. Coeck at the VIB Nucleomics Core (KU Leuven) under the expert guidance of P. Verhasselt, R. Janky, W. Van Delm and S. Derveaux. WGBS and oxWGBS were performed at Novogene, and data were analyzed by Y.S. Li. We also thank L. Vermeire and L. Umans for technical guidance in mouse embryo manipulation and protocols for WISH, and H. Zhao for help with bioinformatic analyses. We are grateful for suggestions from M. Wilkinson and C. Verfaillie in the critique of this manuscript. This work was supported by Fonds voor Wetenschappelijk Onderzoek (FWO) Research Foundation–Flanders grants G.0C56.13N and G.0632.13N, Ministerie van de Vlaamse Gemeenschap and Marie Curie Career Integration grant PCIG-GA-2012-321658 (K.P.K.), US NIH grant R35 CA210043 (A.R.) and ERC Consolidator Grant award CHAMELEON 617595 (D.L.).

References

1. Cantone I, Fisher AG. Epigenetic programming and reprogramming during development. Nat Struct Mol Biol. 2013; 20:282–289. [PubMed: 23463313]
2. Smith ZD, et al. A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature. 2012; 484:339–344. [PubMed: 22456710]
3. Bird AP, Wolffe AP. Methylation-induced repression—belts, braces, and chromatin. Cell. 1999; 99:451–454. [PubMed: 10589672]
4. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. Nat Rev Genet. 2013; 14:204–220. [PubMed: 23400093]
5. Rossant J. Stem cells and early lineage development. Cell. 2008; 132:527–531. [PubMed: 18295568]
6. Ying QL, et al. The ground state of embryonic stem cell self-renewal. Nature. 2008; 453:519–523. [PubMed: 18497825]
7. Ficz G, et al. FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. Cell Stem Cell. 2013; 13:351–359. [PubMed: 23850245]
8. Habibi E, et al. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell Stem Cell. 2013; 13:360–369. [PubMed: 23850244]
9. Hayashi K, Ohta H, Kurimoto K, Aramaki S, Saitou M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell. 2011; 146:519–532. [PubMed: 21820164]
10. Brons IG, et al. Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature. 2007; 448:191–195. [PubMed: 17597762]
11. Tesar PJ, et al. New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature. 2007; 448:196–199. [PubMed: 17597760]
12. Kojima Y, et al. The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. Cell Stem Cell. 2014; 14:107–120. [PubMed: 24139757]
13. Tahiliani M, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009; 324:930–935. [PubMed: 19372391]
14. Ito S, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science. 2011; 333:1300–1303. [PubMed: 21778364]
15. He YF, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science. 2011; 333:1303–1307. [PubMed: 21817016]

16. Koh KP, et al. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell. 2011; 8:200–213. [PubMed: 21295276]
17. Ficz G, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature. 2011; 473:398–402. [PubMed: 21460836]
18. Williams K, et al. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature. 2011; 473:343–348. [PubMed: 21490601]
19. Wu H, Zhang Y. Reversing DNA methylation: mechanisms, genomics, and biological functions. Cell. 2014; 156:45–68. [PubMed: 24439369]
20. Hackett JA, et al. Synergistic mechanisms of DNA demethylation during transition to ground-state pluripotency. Stem Cell Rep. 2013; 1:518–531.
21. Sohni A, et al. Dynamic switching of active promoter and enhancer domains regulates Tet1 and Tet2 expression during cell state transitions between pluripotency and differentiation. Mol Cell Biol. 2015; 35:1026–1042. [PubMed: 25582196]
22. Dawlaty MM, et al. Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell. 2011; 9:166–175. [PubMed: 21816367]
23. Yamaguchi S, et al. Tet1 controls meiosis by regulating meiotic gene expression. Nature. 2012; 492:443–447. [PubMed: 23151479]
24. Zhang RR, et al. Tet1 regulates adult hippocampal neurogenesis and cognition. Cell Stem Cell. 2013; 13:237–245. [PubMed: 23770080]
25. Kang J, et al. Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis. Proc Natl Acad Sci USA. 2015; 112:E4236–E4245. [PubMed: 26199412]
26. Dawlaty MM, et al. Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. Dev Cell. 2013; 24:310–323. [PubMed: 23352810]
27. Dai HQ, et al. TET-mediated DNA demethylation controls gastrulation by regulating Lefty–Nodal signalling. Nature. 2016; 538:528–532. [PubMed: 27760115]
28. Li J, Zhou BP. Activation of β-catenin and Akt pathways by Twist are critical for the maintenance of EMT associated cancer stem cell–like characters. BMC Cancer. 2011; 11:49. [PubMed: 21284870]
29. Lin HH, et al. Neuronatin promotes neural lineage in ESCs via Ca^{2+} signaling. Stem Cells. 2010; 28:1950–1960. [PubMed: 20872847]
30. Shen MM. Nodal signaling: developmental roles and regulation. Development. 2007; 134:1023–1034. [PubMed: 17287255]
31. Wu H, et al. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature. 2011; 473:389–393. [PubMed: 21451524]
32. Zhou W, et al. HIF1α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. EMBO J. 2012; 31:2103–2116. [PubMed: 22446391]
33. Tanaka S, Kunath T, Hadjantonakis AK, Nagy A, Rossant J. Promotion of trophoblast stem cell proliferation by FGF4. Science. 1998; 282:2072–2075. [PubMed: 9851926]
34. Rugg-Gunn PJ, Cox BJ, Ralston A, Rossant J. Distinct histone modifications in stem cell lines and tissue lineages from the early mouse embryo. Proc Natl Acad Sci USA. 2010; 107:10783–10790. [PubMed: 20479220]
35. Buecker C, et al. Reorganization of enhancer patterns in transition from naive to primed pluripotency. Cell Stem Cell. 2014; 14:838–853. [PubMed: 24905168]
36. Booth MJ, et al. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science. 2012; 336:934–937. [PubMed: 22539555]
37. Stadler MB, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature. 2011; 480:490–495. [PubMed: 22170606]
38. Shirane K, et al. Global landscape and regulatory principles of DNA methylation reprogramming for germ cell specification by mouse pluripotent stem cells. Dev Cell. 2016; 39:87–103. [PubMed: 27642137]

39. Yu M, et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell. 2012; 149:1368–1380. [PubMed: 22608086]
40. Zalzman M, et al. Zscan4 regulates telomere elongation and genomic stability in ES cells. Nature. 2010; 464:858–863. [PubMed: 20336070]
41. Nakai-Futatsugi Y, Niwa H. Zscan4 is activated after telomere shortening in mouse embryonic stem cells. Stem Cell Rep. 2016; 6:483–495.
42. Lu F, Liu Y, Jiang L, Yamaguchi S, Zhang Y. Role of Tet proteins in enhancer activity and telomere elongation. Genes Dev. 2014; 28:2103–2119. [PubMed: 25223896]
43. Yang J, et al. Tet enzymes regulate telomere maintenance and chromosomal stability of mouse ESCs. Cell Rep. 2016; 15:1809–1821. [PubMed: 27184841]
44. Johansson C, et al. The roles of Jumonji-type oxygenases in human disease. Epigenomics. 2014; 6:89–120. [PubMed: 24579949]
45. Kim TG, Kraus JC, Chen J, Lee Y. JUMONJI, a critical factor for cardiac development, functions as a transcriptional repressor. J Biol Chem. 2003; 278:42247–42255. [PubMed: 12890668]
46. Rudenko A, et al. Tet1 is critical for neuronal activity-regulated gene expression and memory extinction. Neuron. 2013; 79:1109–1122. [PubMed: 24050401]
47. Liao J, et al. Targeted disruption of *DNMT1*, *DNMT3A* and *DNMT3B* in human embryonic stem cells. Nat Genet. 2015; 47:469–478. [PubMed: 25822089]
48. Cimmino L, et al. TET1 is a tumor suppressor of hematopoietic malignancy. Nat Immunol. 2015; 16:653–662. [PubMed: 25867473]
49. Vermeire L, et al. Essential validation of gene trap mouse ES cell lines: a test case with the gene *Ttrap*. Int J Dev Biol. 2009; 53:1045–1051. [PubMed: 19598121]
50. de Vree PJ, et al. Targeted sequencing by proximity ligation for comprehensive variant detection and local haplotyping. Nat Biotechnol. 2014; 32:1019–1025. [PubMed: 25129690]
51. Koentgen F, et al. Exclusive transmission of the embryonic stem cell–derived genome through the mouse germline. Genesis. 2016; 54:326–333. [PubMed: 27012318]
52. Meno C, et al. Two closely-related left–right asymmetrically expressed genes, *lefty-1* and *lefty-2*: their distinct expression domains, chromosomal linkage and direct neuralizing activity in *Xenopus* embryos. Genes Cells. 1997; 2:513–524. [PubMed: 9348041]
53. Bachman M, et al. 5-Hydroxymethylcytosine is a predominantly stable DNA modification. Nat Chem. 2014; 6:1049–1055. [PubMed: 25411882]
54. Czechanski A, et al. Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains. Nat Protoc. 2014; 9:559–574. [PubMed: 24504480]
55. Taiwo O, et al. Methylome analysis using MeDIP–seq with low DNA concentrations. Nat Protoc. 2012; 7:617–636. [PubMed: 22402632]
56. Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008; 9:R137. [PubMed: 18798982]
57. Thomson JP, et al. Comparative analysis of affinity-based 5-hydroxymethylation enrichment techniques. Nucleic Acids Res. 2013; 41:e206. [PubMed: 24214958]
58. Matarese F, Carrillo-de Santa Pau E, Stunnenberg HG. 5-Hydroxymethylcytosine: a new kid on the epigenetic block? Mol Syst Biol. 2011; 7:562. [PubMed: 22186736]
59. Robinson MD, et al. Evaluation of affinity-based genome-wide DNA methylation data: effects of CpG density, amplification bias, and copy number variation. Genome Res. 2010; 20:1719–1729. [PubMed: 21045081]
60. Masser DR, Stanford DR, Freeman WM. Targeted DNA methylation analysis by next-generation sequencing. J Vis Exp. 2015; 96:e52488.
61. Kumaki Y, Oda M, Okano M. QUMA: quantification tool for methylation analysis. Nucleic Acids Res. 2008; 36:W170–W175. [PubMed: 18487274]
62. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite–Seq applications. Bioinformatics. 2011; 27:1571–1572. [PubMed: 21493656]
63. Smallwood SA, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat Methods. 2014; 11:817–820. [PubMed: 25042786]

64. Wang Z, et al. swDMR: a sliding window approach to identify differentially methylated regions based on whole genome bisulfite sequencing. PLoS One. 2015; 10:e0132866. [PubMed: 26176536]

65. Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics. 2015; 31:2382–2383. [PubMed: 25765347]

66. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012; 9:357–359. [PubMed: 22388286]

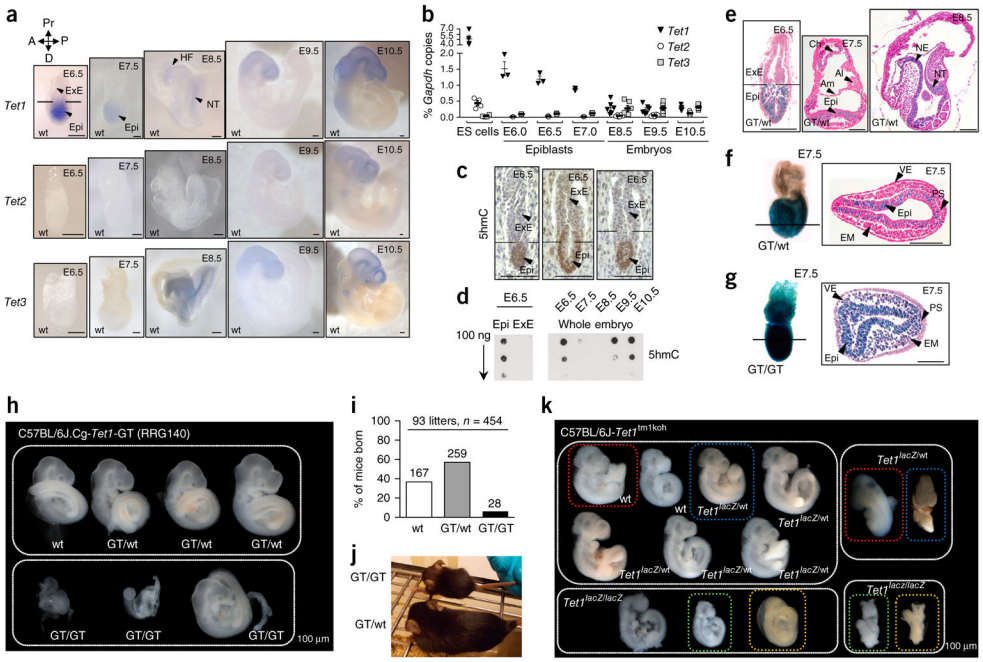

Figure 1.

Tet1 expression and loss of function in the postimplantation mouse embryo. (**a**) WISH analysis of E6.5–E10.5 B6 wild-type (wt) embryos. Proximal (Pr), distal (D), anterior (A) and posterior (P) axes refer to E6.5–E8.5. (**b**) qRT–PCR of TET transcript copies of TET genes normalized to *Gapdh*. Error bars, s.e.m. of biological replicates of ESC lines ($n = 3$), pooled E6.0–E7.0 epiblasts ($n = 3$) and entire E8.5–E10.5 embryos ($n = 5$ or 6). Values for *Tet1* were obtained from Sohni *et al.*[21]. (**c**) 5hmC immunostaining of serial E6.5 embryo sections. (**d**) 5hmC dot blots of E6.5 Epi and ExE (left) and whole embryos (right). (**e**) Sagittal sections of X-gal-stained $Tet1^{GT/wt}$ embryos. (**f**,**g**) Transverse section of E7.5 $Tet1^{GT/wt}$ (**f**) and $Tet1^{GT/GT}$ (**g**) X-gal-stained embryos. Whole embryos with plane of section are indicated on the left. (**h**,**i**) Representative litters of E9.5 embryos from intercrosses of $Tet1^{GT/wt}$ mice (**h**) and numbers of mice born at late backcross generations to B6 ($N \geq 6$). (**j**) A 4-week-old $Tet1^{GT/GT}$ mouse next to a $Tet1^{GT/wt}$ littermate. (**k**) Representative litters of E9.5 embryos from intercrosses of B6 $Tet1^{tm1koh}$ heterozygous mice. Colored dotted lines delimit the same embryos shown in frontal or dorsal view on the right, and lateral on the left. Scale bars, 100 µm. Epi, epiblast; ExE, extraembryonic ectoderm; HF, head folds; NT, neural tube; Ch, chorion; Am, amnion; Al, allantois; NE, neuroepithelium; EM, embryonic mesoderm; PS, primitive streak; VE, visceral endoderm.

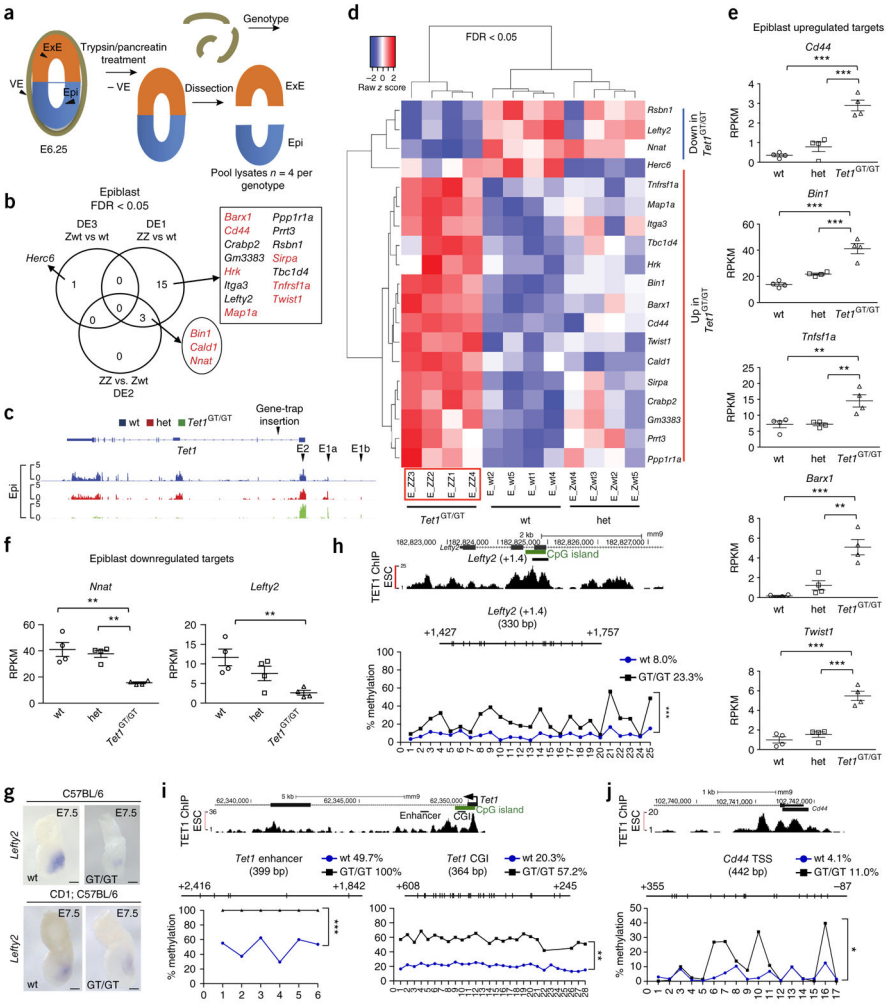

Figure 2.

Genes regulated by TET1 in the epiblast. (**a**) Embryo dissection procedure for RNA–seq. (**b**) Venn diagram showing differentially expressed genes. Gene names highlighted in red are common to DE1 and DE2 using $P < 0.001$ (supplementary Fig. 3d). Zwt, heterozygous (het); ZZ, *Tet1*$^{GT/GT}$. (**c**) Representative read coverage over *Tet1*. Positions of exons (E; E1a, E1b are alternative exons 1) and GT insertion are indicated by black arrowheads. The *y* axis indicates reads per million. (**d**) Heat map cluster of the differentially expressed genes shown in **b**. (**e**,**f**) RPKM of selected upregulated (**e**) and downregulated (**f**) genes in the epiblast. (**g**) WISH detection of *Lefty2* in E7.5 embryos from B6 and outbred CD1 mice. (**h**–**j**) Bisulfite amplicon sequencing analysis at *Lefty2* (**h**), *Tet1* (**i**) and *Cd44* (**j**) in E6.25 Epi. Percentages in legends indicate average methylation levels of all CpGs over each locus. Data shown in **h** and **i** are representative of two independent biological replicate pools. Above, ChIP–seq tracks of TET1 occupancy in mouse ESCs[18]. Middle, relative CpG positions in the amplicon, delineated relative to TSSs. Paired Student's *t* test, $*P < 0.05$.

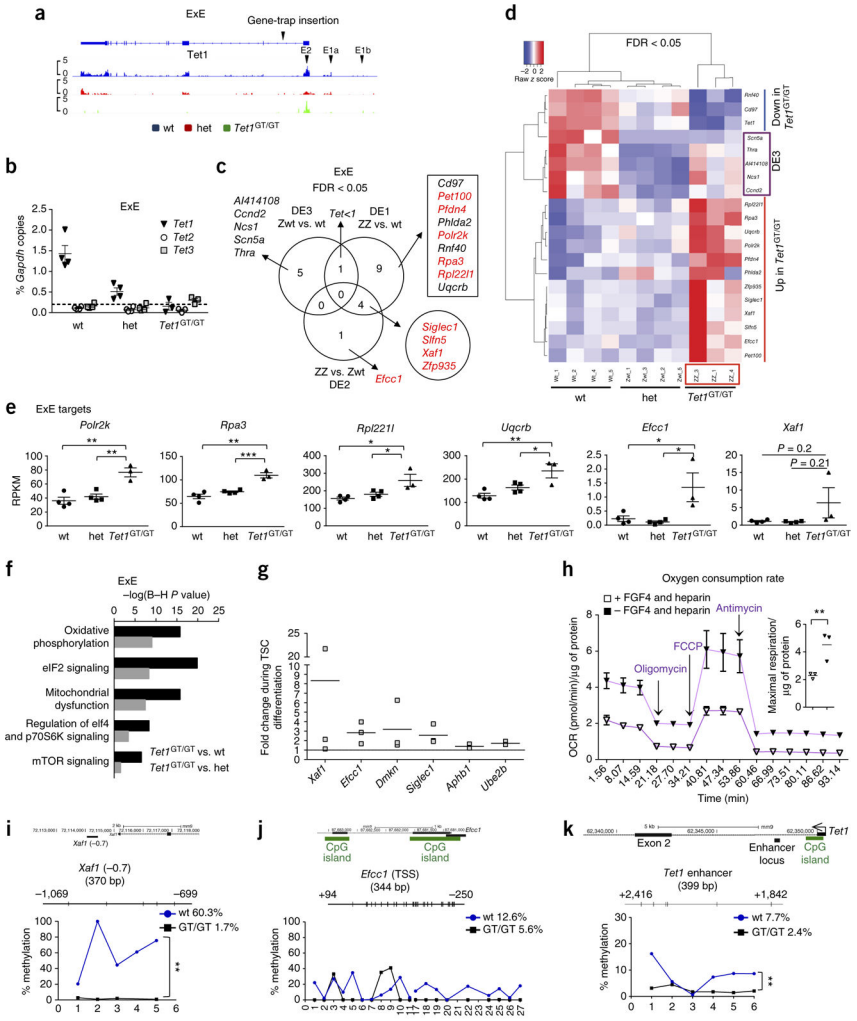

Figure 3.

Genes regulated by TET1 in the extraembryonic ectoderm. (**a**) Read coverage over *Tet1*, as described in Figure 2c. (**b**) TET gene expression, as described in Figure 1b. Error bars, s.e.m. of biological replicates ($n = 3$ or 4). The dotted line indicates background detection. (**c**) Venn diagram of differentially expressed genes. Gene names highlighted in red are common to DE1 and DE2 using $P < 0.001$ (supplementary Fig. 4a). (**d**) Heat map clustering of all differentially expressed genes shown in **c**. (**e**) RPKM values of selected differentially expressed genes. ANOVA, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. (**f**) Top five pathways identified by Ingenuity using Benjamini-Hochberg (BH)-corrected $P < 0.05$. (**g**) Fold change in expression of selected genes during differentiation of TSCs by withdrawal of fibroblast growth factor 4 (FGF4) and heparin. Error bars, s.e.m. of biological replicates ($n = 3$ or 4 cell cultures). Unpaired Student's *t* test, $*P < 0.05$. (**h**) Oxygen consumption rate normalized to protein in TSCs. Error bars, s.d. of technical replicates ($n = 3$). The inset shows a significant increase in normalized maximal respiration rates during differentiation. (**i**–**k**) Bisulfite amplicon sequencing analysis at *Xaf1* (**i**), *Efcc1* (**j**) and *Tet1* (**k**) in E6.25 ExE, as described in Figure 2h–j. Data shown here are representative of two independent biological replicate pools.

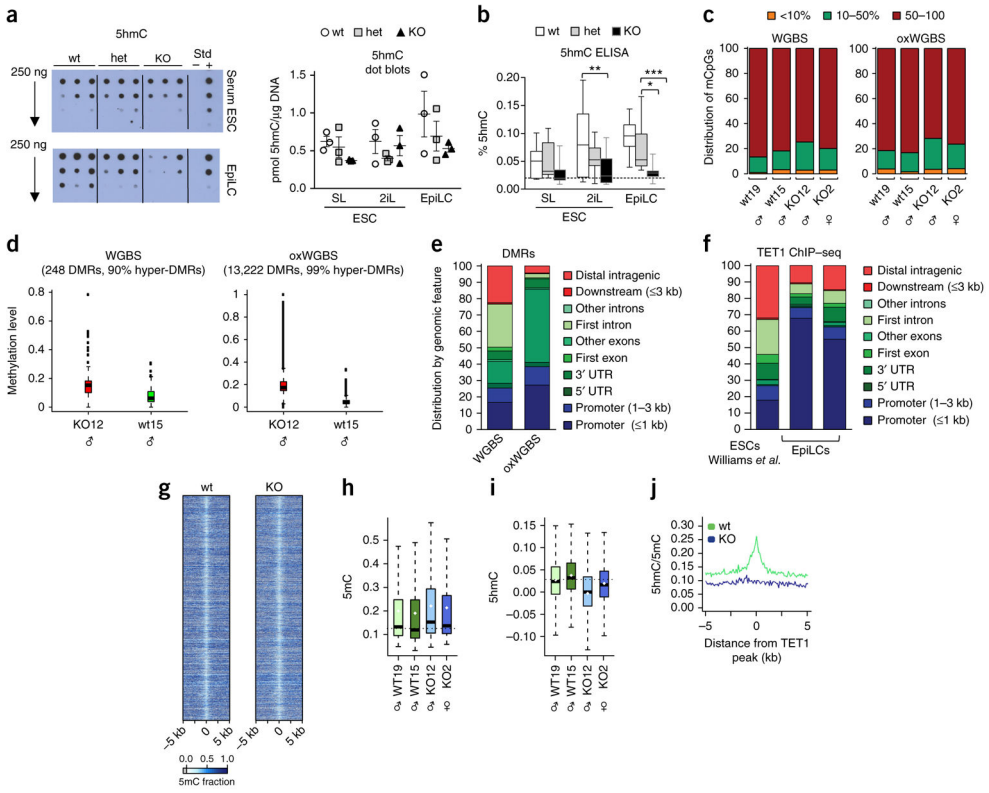

Figure 4.

Effects of TET1 on the DNA methylome in epiblast-like cells. (**a**) Representative 5hmC dot blots (left) and quantification (right). Serial dilution of an oligonucleotide containing 5hmC (+) or lacking 5hmC (−) serve as standards (Std). Error bars, s.e.m. of biological replicates ($n = 3$ cell lines). (**b**) ELISA-based 5hmC quantification of data from biological triplicate lines analyzed in 2–3 independent assays. The dotted horizontal line represent the background level. Nested ANOVA, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. (**c**) Distribution of methylated CpGs (mCpGs) (coverage ≥ 10) among unmethylated (<10% where 0% is not included), lowly methylated (10–50%) and highly methylated (50–100%) fractions. (**d**) Methylation levels of DMRs in male wild-type (15) and knockout (12) EpiLC lines. (**e**) Distribution of DMRs identified by WGBS and oxWGBS among genomic features. (**f**) Distribution of TET1 ChIP–seq peak regions classified by genomic features in EpiLCs (two replicate lines) and a published ESC data set[18]. (**g**) Heat map showing 5mC levels (mean of replicate oxWGBS profiles) within 100-bp bins spanning ±5-kb regions centered at TET1 ChIP–seq peaks. Heat maps for individual samples are shown in supplementary Figure 6j. (**h**,**i**) Average 5mC (**h**) and 5hmC (**i**) levels at TET1 binding peaks (summit peak ± 200 bp). The dashed horizontal line represents the median level of wild-type samples. In box plots, boxes show the interquartile range (IQR) around the median and whiskers extend from the minimum to the maximum value unless the distance to the first or third quartile was >1.5× the IQR. (**j**) Line plot of average hmC over mC levels at ±5-kb regions centered at TET1 ChIP–seq peaks.

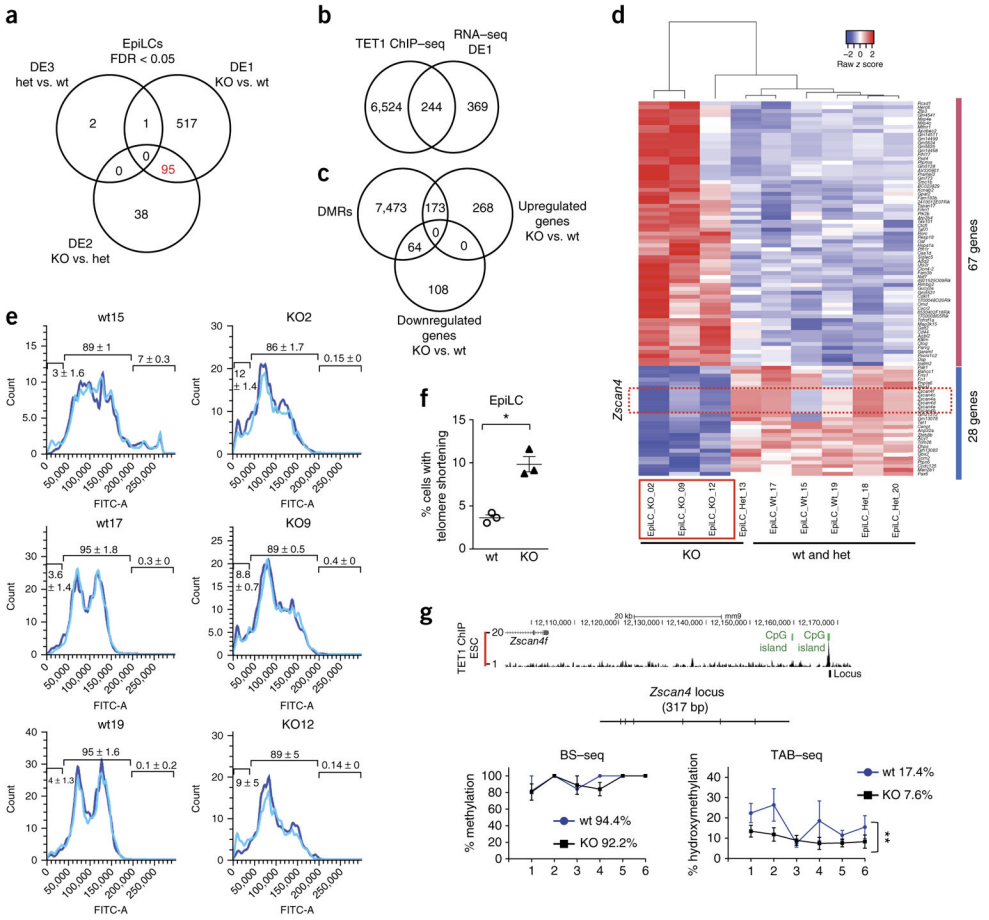

Figure 5.

Genes regulated by TET1 in epiblast-like cells. (**a**) Venn diagram showing differentially expressed genes at FDR < 0.05. het, *Tet1*$^{GT/wt}$; KO, *Tet1*$^{GT/GT}$. (**b**) Overlap of DE1 genes (KO versus wt) with genes associated with TET1 binding. (**c**) Overlap of up- and downregulated genes in DE1 with DMRs from oxWGBS analysis shown in Figure 4d. (**d**) Heat map cluster of 95 differentially expressed genes in the overlap of DE1 and DE2 shown in **a**. (**e**,**f**) Flow-FISH analysis of B6 EpiLC lines. Each line was analyzed in technical duplicates (shown as light and dark blue line profiles), with percentage values of fractions shown as mean ± s.d. in **e**. The values of subtelomere length fractions are shown as mean ± s.e.m. from the biological triplicates in **f**. (**g**) Bisulfite (BS) and TET-assisted bisulfite (TAB) amplicon sequencing analysis at a *Zscan4f* locus. Error bars, s.e.m. of biological triplicate EpiLC lines. Above, ChIP–seq tracks of TET1 occupancy in mouse ESCs[18]. Paired Student's *t* test, $*P < 0.05$.

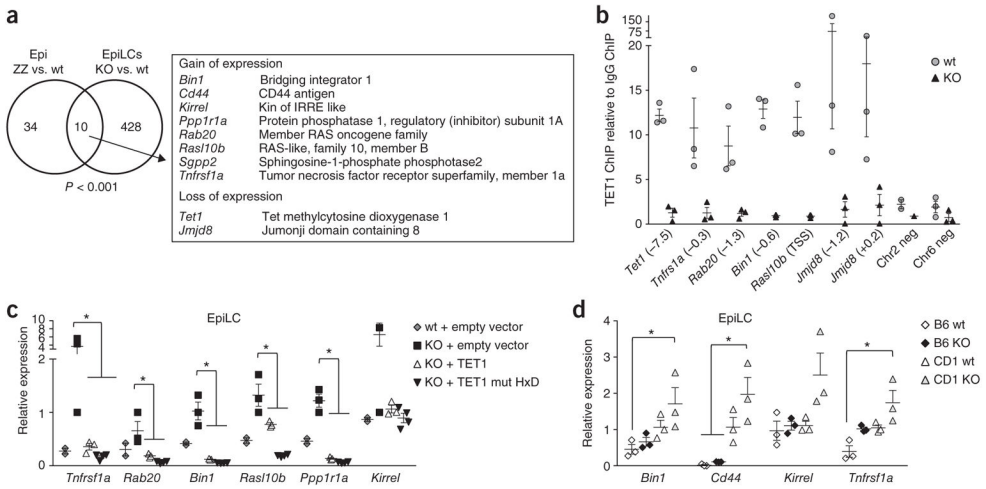

Figure 6.

Target genes of TET1 in the epiblast. (**a**) Venn diagram overlap of TET1-regulated genes common in Epi and EpiLCs. ZZ and KO, *Tet1*$^{GT/GT}$. (**b**) TET1 ChIP–qPCR at differentially expressed genes in EpiLCs. Error bars, s.e.m. of biological triplicate lines. (**c**) Repression of differentially expressed genes upon rescue overexpression of full-length TET1 wild type and catalytic mutant (mut HxD) transcripts in *Tet1*-knockout EpiLCs. Error bars, s.e.m. of two (wt + empty vector), three (KO + empty and KO + TET1) and four (KO + TET1 mut HxD) transfected clones; expression levels are relative to one mock-transfected knockout clone. (**d**) Expression of differentially expressed genes in B6 versus CD1 wild-type and knockout EpiLCs. Error bars, s.e.m. of biological triplicate lines; expression is relative to one CD1 wild-type EpiLC line. ANOVA, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

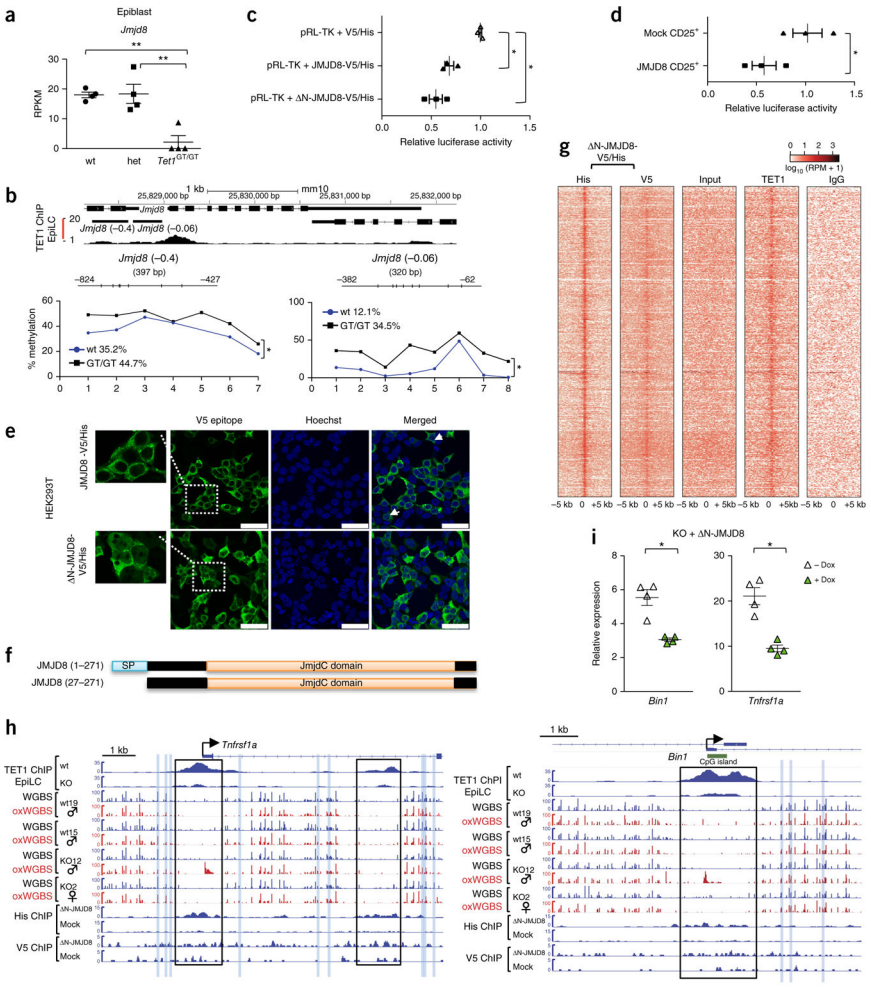

Figure 7.

Transcriptional repression by JMJD8. (**a**) RPKM of *Jmjd8* in the Epi. (**b**) Bisulfite amplicon sequencing analysis in E6.25 Epi. Above, TET1 ChIP–seq peaks in wild-type EpiLCs. Paired Student's *t* test, $*P < 0.05$. (**c**) Dual-reporter luciferase assay in HEK293T cells indicating thymidine kinase promoter-driven *Renilla* activity (normalized to cotransfected promoterless pGL3-Basic) upon overexpression of JMJD8, a truncated JMJD8 nuclear isoform (ΔN-JMJD8) or empty vector. Error bars, s.e.m. of three independent experiments. (**d**) Single-reporter luciferase assay in HEK293T cells with activity normalized to protein content in cells sorted for transfection. Error bars, s.e.m. of three independent experiments. ANOVA, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. (**e**) Confocal images of HEK293T cells expressing full-length or truncated C-terminally tagged JMJD8 stained with an anti-6× His antibody. White arrows indicate two cells showing nuclear staining. Scale bars, 50 μm. (**f**) Schematic representation of JMJD8 indicating the predicted signal peptide (SP) and JmjdC domain. The SP-deleted isoform is ΔN-JMJD8. (**g**) Heat map depiction of 1,295 ΔN-JMJD8-His ChIP peaks (>4 RPKM at the summit peak ± 200 bp) alongside ΔN-JMJD8-V5 and TET1 ChIP peak intensities. (**h**) TET1 and JMJD8 ChIP–seq peaks and mCpG levels at *Bin1* and *Tnfr1s1a* in EpiLCs. Regions showing co-binding of TET1 and JMJD8 are boxed. Statistically validated 5hmCpGs in wild-type EpiLCs are highlighted in blue. (**i**) Expression

of *Bin1* and *Tnfrsf1a* in *Tet1*-knockout EpiLCs expressing doxycycline (DOX)-inducible ΔN-JMJD8, relative to wild type. Error bars, s.e.m. of four clonal transfected lines. Unpaired Student's *t* test, $*P < 0.05$.