Coordination of germ layer lineage choice by TET1 during primed pluripotency

Xinlong Luo,[1] Bernard K. van der Veer,[1] Lei Sun,[1] Michela Bartoccetti,[1] Matteo Boretto,[2] Hugo Vankelecom,[2] Rita Khoueiry,[1,3] and Kian Peng Koh[1]

[1]Laboratory for Stem Cell and Developmental Epigenetics, [2]Laboratory of Tissue Plasticity in Health and Disease, Department of Development and Regeneration, Katholieke Universiteit Leuven, 3000 Leuven, Belgium

Gastrulation in the early postimplantation stage mammalian embryo begins when epiblast cells ingress to form the primitive streak or develop as the embryonic ectoderm. The DNA dioxygenase Tet1 is highly expressed in the epiblast and yet continues to regulate lineage specification during gastrulation when its expression is diminished. Here, we show how Tet1 plays a pivotal role upstream of germ layer lineage bifurcation. During the transition from naive pluripotency to lineage priming, a global reconfiguration redistributes Tet1 from Oct4-cobound promoters to distal regulatory elements at lineage differentiation genes, which are distinct from high-affinity sites engaged by Oct4. An altered chromatin landscape in *Tet1*-deficient primed epiblast-like cells is associated with enhanced Oct4 expression and binding to Nodal and Wnt target genes, resulting in collaborative signals that enhance mesendo-dermal and inhibit neuroectodermal gene expression during lineage segregation. A permissive role for Tet1 in neural fate induction involves Zic2-dependent engagement at neural target genes at lineage priming, is dependent on the signaling environment during gastrulation, and impacts neural tube closure after gastrulation. Our findings provide mechanistic information for epigenetic integration of pluripotency and signal-induced differentiation cues.

Embryonic development in mammals involves an early phase of gastrulation, when cell movements result in a massive reorganization of the embryo from a hollow sphere of cells, or blastocyst, into a multilayered gastrula following implantation (Tam and Loebel 2007). The gastrula contains three germ layers known as ectoderm, mesoderm, and endoderm that contain progenitors of all tissues in the adult body. In mice, the onset of gastrulation occurs at embryonic day E6.5 with the appearance of a primitive streak in the posterior region of the embryo, through which epiblast cells ingress by an epithelial-mesenchymal transition to form progenitor cells of the mesoderm and definitive endoderm, also known as "mesendoderm" (ME). Descendents of epiblast cells that do not pass through the primitive streak form the surface and neural ectoderm (NE). These early lineages are specified temporally and spatially through the coordinated activities of various signaling pathways downstream from wingless-related integrated site (Wnt), Nodal and bone morphogenetic protein (BMP) family receptors (Robertson 2014; Muñoz-Descalzo et al. 2015; Zinski et al. 2018). However, recent studies suggest that cells acquire compe-tence to differentiate in response to inductive signaling cues at earlier stages of epiblast transitions (Smith 2017). The molecular mechanisms acting upstream of germ layer lineage decisions remain unclear.

The developmental progression of the epiblast through lineage priming and germ layer fate allocation has been largely recapitulated in vitro using embryonic stem cell (ESC) cultures (Niwa 2010). Murine ESCs derived from the preimplantation E3.5 blastocyst can be maintained in a homogenous "ground state" with unrestricted "naive" pluripotency in chemically defined medium containing leukemia inhibitory factor (LIF) and inhibitors of two differentiation-inducing kinase pathways, commonly known as 2iL (Ying et al. 2008). A 2-d differentiation protocol of 2iL-adapted ESCs to transient epiblast-like cells (EpiLCs) in vitro captures the gastrulation-poised "primed" pluripotent state of the postimplantation epiblast (Hayashi et al. 2011; Buecker et al. 2014). Earlier differentiation models utilizing murine ESCs relied on cellular aggregation in nonadherent conditions to form embryoid bodies (EBs) containing germ layer progenitors

[3]Present address: Epigenetics Group, International Agency for Research on Cancer, 69008 Lyon, France.
Corresponding author: kian.koh@kuleuven.be

mimicking the gastrula (ten Berge et al. 2008). Under differentiation-permissive conditions, a synergistic interdependency between canonical Wnt/β-catenin and Nodal signaling pathways drives ME differentiation (Funa et al. 2015; Wang et al. 2017). Further lineage specification depends on modulation of collaborative signal inputs; strong Nodal signals in EBs favor a definitive endoderm fate while low levels or Wnt favor a mesoderm fate (Gadue et al. 2006), in agreement with how graded Nodal morphogen gradients in the primitive streak drive patterning of definitive endoderm anteriorly and mesoderm posteriorly in vivo (Vincent et al. 2003; Schier 2009; Robertson 2014).

The use of cell culture mimics of early embryonic differentiation has facilitated recent exploration of genomic regulatory mechanisms underlying early developmental cell state transitions. For example, the transitions from naive to primed pluripotency during conversion of 2iL ESCs to EpiLCs have revealed global reorganization of the master pluripotency transcription factor Oct4 at cell state-specific enhancer sites (Buecker et al. 2014). In human and murine ESCs, Oct4 co-occupies distal regulatory elements with Smad2 and Smad3 (Smad2/3), closely related downstream intracellular effectors of TGF-β/Activin/Nodal receptors (Shi and Massagué 2003), to regulate target genes including *Oct4* (also known as *Pou5f1*), *Nodal*, and its antagonist, *Lefty1* (Mullen et al. 2011), in an autoregulatory feedback loop. Upon exit from pluripotency, Oct4 continues to determine germ layer fate selection by promoting ME and repressing NE fate (Thomson et al. 2011), through new cooperative interactions with tissue-specific and signal-dependent factors at lineage-priming genomic sites (Funa et al. 2015; Simandi et al. 2016). Collectively, these studies suggest that the capacity of transcription factors (TFs) to access enhancer sites to drive cell fate choice is highly context-dependent. The complex circuitry that decides which factor(s) dominates in the process of ESC commitment to ME or NE fate, whether stochastic or predetermined by regulators acting upstream of lineage bifurcation, remains to be worked out in further detail.

Dynamic changes in DNA methylation constitute a fundamental epigenetic mechanism in cell fate commitment in the early embryo (Santos et al. 2002). The Ten–Eleven-Translocation (TET) DNA dioxygenase (Tet1, Tet2, and Tet3) family proteins mediate DNA demethylation by iterative oxidation of 5-methylcytosine (5mC) (Tahiliani et al. 2009; He et al. 2011; Ito et al. 2011). As important regulators of epigenetic reprogramming, they are intensively studied for their roles in development and disease (Pastor et al. 2013; Wu and Zhang 2014). In the early postimplantation mouse embryo, *Tet1* is sustained in expression without detectable *Tet2* and *Tet3* (Khoueiry et al. 2017). In agreement with a nonredundant biological role of *Tet1* in lineage priming, loss of *Tet1* in ESCs results in hyperactive Nodal signaling in EBs and increased competence to differentiate toward ME fate (Koh et al. 2011). Furthermore, disruption of *Tet1* function can severely compromise neurulation in the mouse embryo, manifested as failed cranial neuropore closure after gastrulation (Fong et al. 2016; Khoueiry et al. 2017). However, expres-

sion of all three TET genes is low during gastrulation (Khoueiry et al. 2017), suggesting that an epigenetic regulatory mechanism allows TET1 to influence lineage segregation even when expression is down-regulated.

Here, we ask how Tet1 controls developmental outcomes by coordinating pluripotency and differentiation signals at entry points in the molecular circuitries of early embryonic lineage segregation. Our findings clarify how Tet1 regulates the primed pluripotency epigenetic landscape at the crossroads of ME and NE fate entry, providing a cellular memory that regulates engagement of Oct4, Wnt/β-catenin, and Nodal signaling effectors at developmental loci during gastrulation.

Results

Tet1 is repatterned to developmental fate loci during pluripotency transition

To examine whether Tet1 genomic occupancy changes during the transition between naive and primed pluripotency, we first performed Tet1 chromatin immunoprecipitation sequencing (ChIP-seq) in ESCs cultured in 2iL (naive state) and following in vitro conversion to EpiLCs (primed state). Based on differential binding intensities determined from biological replicates ($n = 2$) of wild-type (WT) cells and normalized to *Tet1*-deficient (knockout [KO]) negative controls (Supplemental Fig. S1A), we defined Tet1-bound peaks preferentially enriched in 2iL-treated ESCs (ESC-specific, 1256) or EpiLCs (EpiLC-specific, 5557) and those where Tet1 binding did not show significant differences between the two states (common, 5741) (Fig. 1A; Supplemental Table S1). The sizable numbers of EpiLC-specific binding peaks suggest the presence of de novo Tet1 binding, in line with detection of EpiLC-specific and Tet1-dependent differentially hydroxymethylated regions previously observed during ESC-to-EpiLC differentiation (Khoueiry et al. 2017). Moreover, even though Tet1 binding profiles were preferentially enriched at gene promoter regions in both pluripotency states, the state-specific binding peaks showed higher proportions located at gene body and intergenic regions compared with common peaks (Fig. 1B). To demonstrate that EpiLC-specific peaks are primarily located at promoter-distal regulatory sites, we examined the peak regions for enhancer features marked by the presence of p300 histone acetyltransferase occupancy, histone 3 Lys4 monomethylation (H3K4me1) and histone 3 Lys27 acetylation (H3K27ac) using a published data set profiled in EpiLCs (Buecker et al. 2014). In comparison with strong p300 and H3K27ac and weaker H3K4me1 signals detected at common and ESC-specific Tet1 peaks, which are consistent with a prevalence at active promoters, EpiLC-specific Tet1 peaks displayed stronger p300 and H3K4me1 co-occupancies and much weaker enrichment for the active H3K27ac mark (Fig. 1C). These results indicated that de novo Tet1 genomic occupancy in EpiLCs occurs mainly at primed or poised enhancers.

Gene ontology (GO) analysis showed that EpiLC-specific binding sites were highly enriched in GO terms for
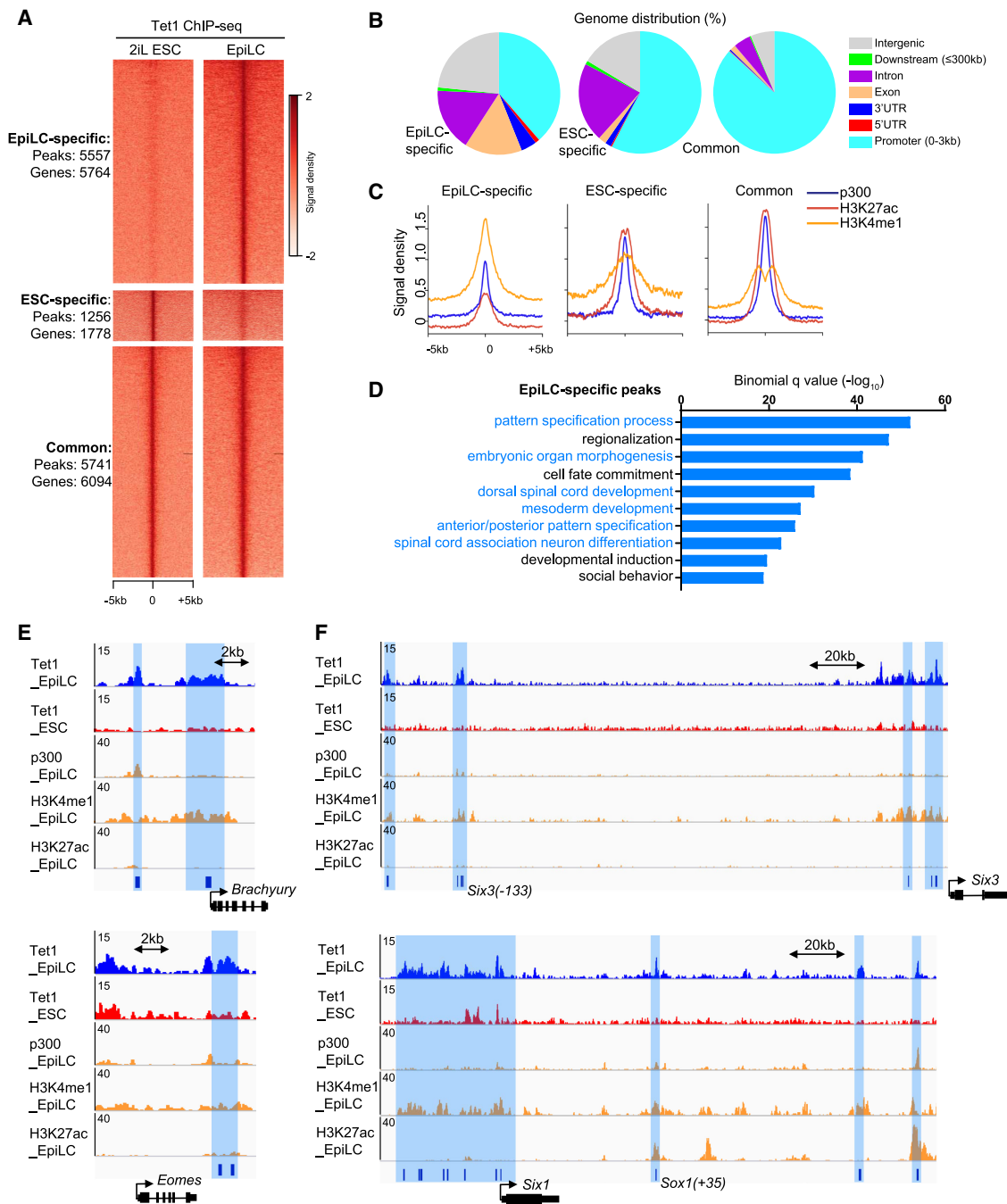
Figure 1. Tet1 is repatterned to developmental fate loci during pluripotency transition . (*A*) Heat map of Tet1 ChIP-seq signal densities in 2iL-cultured ESCs and EpiLCs at defined EpiLC-specific, ESC-specific, and common peaks (*n* = 2 biological replicates). See Supplemental Table S1 for an annotated list of all peak regions. (kb) Kilobases. (*B*) Pie charts indicating genomic feature distribution of Tet1-bound EpiLC-specific, ESC-specific, and common peaks. (*C*) Density aggregation plots of p300, H3K4me1, and H3K27ac ChIP-seq signal densities (Buecker et al. 2014) in EpiLCs at Tet1-bound EpiLC-specific, ESC-specific, and common peaks. (*D*) Functional annotation of GO terms enriched in EpiLC-specific peaks analyzed by GREAT (McLean et al. 2010). The text highlighted in blue indicates terms associated with lineage development. See Supplemental Table S2 for an annotated list of ranked GO terms. (*E,F*) Integrative Genomics Viewer (IGV) snapshots of Tet1 ChIP-seq tracks in EpiLCs and 2iL-cultured ESCs, and p300, H3K4me1, H3K27ac ChIP-seq tracks in EpiLCs, illustrating EpiLC-specific loci associated with ME genes (*Brachyury* and *Eomes* in *E*) or NE genes (*Six3* and *Sox1* in *F*). Gene structure annotations are shown *below* each track set. Blue highlights indicate regions containing statistically significant differential peaks of Tet1 binding between ESCs and EpiLCs identified by MACS2 from replicates (BED annotations are shown as blue bars *below* tracks). Gene loci *Six3(−133)* and *Sox1(+35)* are neural lineage-associated enhancers identified from a previous study (Cruz-Molina et al. 2017).

pattern specification, cell fate commitment, mesoderm development, and spinal cord-associated neuron differentiation (Fig. 1D; Supplemental Table S2). These regions were associated with several genes that are key determinants of lineage fate, including *Brachyury*, *Eomes*, *Six3*, and *Sox1* (Fig. 1E,F; Arnold et al. 2008; Gouti et al. 2014; Cruz-Molina et al. 2017; Koch et al. 2017). Moreover, pathway analysis from the Molecular Signatures Database (MSigDB) showed these Tet1-bound EpiLC-specific peaks to be highly enriched in Wnt-mediated signal transduction and regulation of SMAD2/3 signaling (Supplemental Fig. S1B). In stark contrast, these developmental-related terms were missing among ESC-specific and common peaks, which were instead enriched in general transcriptional and translational processes (Supplemental Fig. S1C,D; Supplemental Table S2). As expected for a functional relevance in blastocyst development, ESC-specific binding sites were also enriched in GO terms related to placenta development and stem cell maintenance (Supplemental Fig. S1C) and can be found at distal sites of *Esrrb* and *Klf4* (Supplemental Fig. S1E), genes encoding naive pluripotency factors known to have pioneer activities in mouse ESCs (Adachi et al. 2018).

Tet1 influences chromatin accessibility at primed developmental loci

We previously detected widespread hypermethylated genomic regions in *Tet1*-deficient EpiLCs enriched at genes associated with neurological functions (Khoueiry et al. 2017). Therefore, we asked whether the presence of Tet1 in EpiLCs also regulates chromatin accessibility, by profiling B6.129P2-*Tet1*$^{Gt (RRG140)}$ (abbreviated to *Tet1*^{GT}) WT and *Tet1*-deficient EpiLCs using the assay for transposase-accessible chromatin sequencing (ATAC-seq). Based on analysis of biological replicates ($n = 2$), we identified 5661 regulatory elements that were differentially accessible in *Tet1*$^{GT/GT}$ (KO) EpiLCs compared with WT EpiLCs, of which 4280 sites showed decreased chromatin accessibility (referred to as loss-in-KO sites) and 1381 sites showed increased accessibility (referred to as gain-in-KO sites) (Fig. 2A; Supplemental Fig. S2A; Supplemental Table S3). The higher incidences and extents of losing chromatin accessibility compared with gaining in *Tet1*-deficient EpiLCs are in line with the role of Tet1 in promoting an open chromatin state at regulatory elements through DNA demethylation.

Genomic distribution of all loci exhibiting Tet1-dependent differential chromatin accessibility showed locations within promoters, intronic and intergenic regions (Supplemental Fig. S2B). Promoter regions (0–3 kb upstream of transcription start sites or TSS) were relatively more prevalent in loss-in-KO and depleted in gain-in-KO loci, whereas intergenic and intronic regions were reciprocally more prevalent in gain-in-KO loci (Supplemental Fig. S2B). The apparent enrichment of promoter regions in loss-in-KO loci relative to gain in KO may be attributed to the presence of common peaks bound by Tet1 in both naive and primed states, which mostly occurred at promoters (Fig. 1B). We asked whether the fraction

(∼50%) located within introns and intergenic regions may nonetheless be distal enhancers. Indeed, loss-in-KO loci showed enrichment for markers of active distal regulatory elements (p300, H3K27ac, and H3K4me1) in WT EpiLCs, which were comparably much weaker in gain-in-KO loci (Fig. 2B), suggesting that Tet1 promotes chromatin accessibility predominantly at active enhancers, whereas sites that gained accessibility in the absence of Tet1 are normally silent in EpiLCs.

When centered at TET1-bound regions in EpiLCs, normalized ATAC-seq signal intensities were lower in *Tet1*$^{GT/GT}$ EpiLCs compared with WT (Fig. 2C; Supplemental Fig. S2C). Of all Tet1-bound regions, 84.6% were located within accessible genomic regions in EpiLCs, overlapping considerably with loss-in-KO but not with gain-in-KO sites (Fig. 2D; Supplemental Fig. S2D). Thus, the latter were not direct targets of Tet1. In our previous study, we had observed that Tet1-bound regions in EpiLCs generally lost 5hmC and gained 5mC upon loss of TET1, without a corresponding reduction in expression of a majority of genes associated with these loci (Khoueiry et al. 2017). By overlapping those whole-genome differentially methylated regions (DMRs) identified by "true 5mC" oxidative bisulfite sequencing with ATAC-seq loss-in-KO loci, we found 2749 regions where loss of chromatin accessibility were associated with gains in DNA methylation (Fig. 2E). Interestingly, these regions are associated with neurodevelopmental lineage genes, including *Sox* family genes *Nestin* and *Sim2*.

In addition to developmental processes, GO terms for genes associated with loss-in-KO sites include protein localization to telomeric region and G1 DNA damage checkpoint (Supplemental Fig. S2E), consistent with a role for Tet1 in maintaining genomic stability (Cimmino et al. 2015; Yang et al. 2016; Khoueiry et al. 2017). On the other hand, gain-in-KO sites enriched for RNA processing, translation, and protein targeting as top GO terms and only weakly for a developmental term for mesenchymal stem cell differentiation (Supplemental Fig. S2E). The latter would be consistent with differentiation skewing toward ME at the expense of NE (Koh et al. 2011). However, gain-in-KO sites within ME-related genes (*Pdgfra*, *Rest*, *Sox5*, and *Sox9*) included in this GO term were minor peaks located within broader genomic regions where most peaks exhibited reduced accessibility in KO (data not shown). When we further examined a subset of 359 loss-in-KO sites that overlapped with EpiLC-specific Tet1 ChIP-seq peaks, we observed that these regions, associated with 505 genes, were significantly enriched for GO terms for pattern specification and spinal cord development (Fig. 2F). These data suggest that Tet1 occupies new enhancer regions during naive-to-primed pluripotency transitions, primarily at neurodevelopmental loci where it promotes an open chromatin state.

To determine whether Tet1-dependent chromatin accessibility in EpiLCs were state-specific or pre-existing in ESCs, we performed further ATAC-seq analysis in WT and KO 2iL-treated ESCs. At loss-in-KO sites defined in EpiLCs, ATAC-seq signals in WT and KO ESCs were detected at similar levels, as in KO EpiLCs, suggesting that
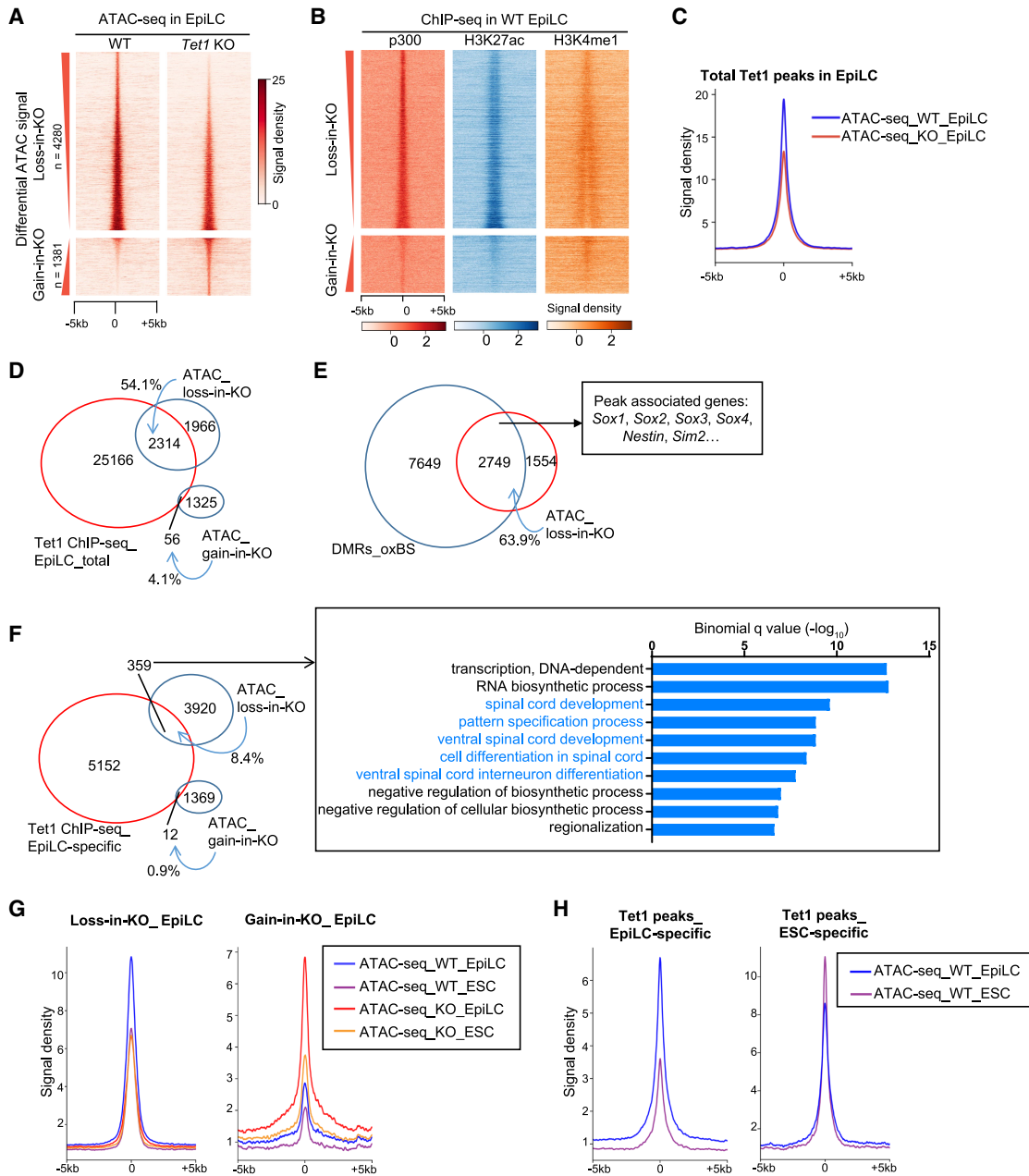
Figure 2. Tet1 influences chromatin accessibility at primed developmental loci. (*A*) Heat map of ATAC-seq signal densities in EpiLCs at regions showing differential chromatin accessibility between WT and $Tet1^{\mathrm{GT/GT}}$ (KO) ($n = 2$ biological replicates) of strain B6.129P2-$Tet1^{\mathrm{Gt(RRG140)}}$ (Khoueiry et al. 2017). Peak regions with decreased (4280; referred to as "loss in KO"; *top*) and increased chromatin accessibility (1381; referred to as "gain in KO"; *bottom*) in KO EpiLCs are ranked by decreasing and increasing ATAC signal change, respectively. See Supplemental Table S3 for an annotated list of all peak regions. (*B*) Heat map of p300, H3K27ac, and H3K4me1 ChIP-seq signal densities in WT EpiLCs at loss-in-KO and gain-in-KO ATAC-seq peaks in EpiLCs as defined in *A*. (*C*) Density aggregation plot of ATAC-seq signal densities in WT and *Tet1* KO EpiLCs at total Tet1-bound peaks in EpiLCs. (*D*) Venn overlap of total Tet1 ChIP-seq peaks in EpiLCs with the subsets of loss-in-KO and gain-in-KO ATAC-seq peaks. (*E*) Venn overlap of loss-in-KO ATAC-seq peaks in EpiLCs with DMRs previously defined by oxidative whole-genome bisulfite sequencing data (Khoueiry et al. 2017). (*F*) Venn overlap of Tet1 EpiLC-specific ChIP-seq peaks with the subsets of loss-in-KO and gain-in-KO ATAC-seq peaks and GO analysis of the former overlapping subset. GO terms associated with neural lineage development are highlighted in blue. (*G*) Density aggregation plots of ATAC-seq signal densities in WT and *Tet1* KO 2iL-cultured ESCs and EpiLCs centered at loss-in-KO (*left*) and gain-in-KO (*right*) ATAC-seq peaks in EpiLCs. (*H*) Density aggregation plots of ATAC-seq signal densities in WT 2iL-cultured ESCs and EpiLCs centered at Tet1-bound EpiLC-specific (*left*) and ESC-specific (*right*) peaks.

these sites were already accessible independently of TET1 in ESCs, but that further gain of accessibility in EpiLCs required Tet1 (Fig. 2G, left panel). On the other hand, gain-in-KO EpiLCs sites showed low signal intensities in ESCs but were already differentially more accessible in KO ESCs compared with WT; since there were minimal overlap with ESC-specific and common Tet1-bound regions, Tet1 may be countering chromatin opening at these loci indirectly (Fig. 2G, right panel). Comparing accessibility of state-specific Tet1 bound regions in the two cell states showed increased chromatin opening at EpiLC-specific sites but only slight reduction at ESC-specific sites during ESC-to-EpiLC conversion (Fig. 2H). These observations suggest passive engagement of Tet1 at ESC-specific regions where it has minimal impact on chromatin accessibility. Upon conversion to EpiLCs, Tet1 engages new primed loci where it promotes increased chromatin opening; however, these latter sites already possess an open chromatin state in ESCs.

We then asked whether rescue of Tet1 expression in KO EpiLCs can restore the chromatin landscape to WT. B6-$Tet1^{GT/GT}$ ESC clonal lines harboring doxycycline-inducible $Tet1$ WT or catalytic mutant cDNA constructs (Khoueiry et al. 2017) were generated (Supplemental Fig. S2F). Two replicate lines with rescued expression of either WT (Rwt) or mutant Tet1 (Rmt) were converted to EpiLCs for ATAC-seq analysis. Unsupervised hierarchical clustering of ATAC-seq signals showed that Tet1 rescued (WT or mutant) lines clustered away from 2iL ESC and EpiLC samples (Supplemental Fig. S2G). Reverse transcription quantitative PCR (RT-qPCR) analysis of naive and primed pluripotency markers indicated efficient conversion of both WT and KO ESCs to EpiLCs, as previously observed (Khoueiry et al. 2017), as shown by complete silencing of naive genes ($Esrrb$ and $Tbx3$) and induction of primed markers ($Fgf5$, $Otx2$, and $Dnmt3b$). While Tet1 rescued lines showed similar changes, the silencing of naive markers appeared incomplete (Supplemental Fig. S2H). Both WT and mutant rescue in EpiLCs failed to restore chromatin accessibility at differential sites to WT levels; on the contrary, further reduction in accessibility was observed (Supplemental Fig. S2I). Thus, the altered accessible chromatin landscape in $Tet1$-deficient EpiLCs could no longer be reversed by subsequent restoration of $Tet1$ expression, arguing against a pioneer role for Tet1 in chromatin remodeling.

Tet1 engagement at primed loci depends on Zic2

To determine whether the state specificity of Tet1 genomic localization involves partner cofactors, we performed motif enrichment analysis of bound regions. EpiLC-specific Tet1 peaks were highly enriched for motifs of developmental regulator families, including zinc finger protein of the cerebellum (Zic), Wilms tumor 1 homolog (Wt-1), high-mobility group (HMG) sex-determining region Y (SRY)-box transcription factors Sox3 and Sox4, and basic helix–loop–helix (bHLH) transcription factors Tcf12 and NeuroD1 (Fig. 3A). In contrast, ESC-specific peaks were enriched for motifs occupied by Kruffel-like factors

(KLF) (Klf3, Klf4, Klf5, Klf6, and Klf14), estrogen-related nuclear receptor (NR)-β (Esrrb), and the Wnt/β-catenin target transcription factors Sp5 and Tcfcp2l1, many of which are naive-specific factors (Fig. 3A). Similar EpiLC- and ESC-specific motifs were found in loss-in-KO- and gain-in-KO-accessible sites, respectively (Fig. 3B). Loss-in-KO-enriched motifs included Sox3, Sox4, and Zic, whereas gain-in-KO sites were enriched in motifs for the KLF family, Sp2 and Sp5.

Using published RNA sequencing (RNA-seq) gene expression data sets of 2iL-treated ESCs and EpiLCs (Buecker et al. 2014), we verified that cognate transcription factors for several EpiLC-specific motifs were indeed expressed at higher levels in EpiLCs than in ESCs (Supplemental Fig. S3A,B). Western blot analysis of Zic2, Wt1, and Sox3 showed strikingly an abundance of Zic2 protein in EpiLCs but undetectable levels in ESCs (Fig. 3C). Using a published Zic2 ChIP-seq data set obtained in postimplantation embryo-derived epiblast stem cells (EpiSCs) (Matsuda et al. 2017), we observed that regions where EpiLC-specific Tet1 occupancy overlapped with loss-in-KO-accessible sites also showed Zic2 occupancy (Fig. 3D). To determine whether Zic2 could be a partner factor that recruits Tet1 to neural fate genes in ESCs (Frank et al. 2015; Luo et al. 2015), we performed Tet1 ChIP at specific target loci upon Zic2 depletion during ESC-to-EpiLC conversion. Using siRNA transfection, we observed >80% depletion of $Zic2$ transcripts (Fig. 3E). We selected distal regulatory elements associated with four lineage genes—$Lhx5$, $Foxp1$, $Sox1$, and $Pax3$—where we found overlap with EpiLC-specific Tet1 occupancies, loss-in-KO chromatin accessibility in EpiLCs, and Zic2 occupancies in EpiSCs (Fig. 3F; Supplemental Fig. S3C). In particular, $Lhx5$ is critical for anterior fate of neuronal progenitor cells (Gouti et al. 2014; Cruz-Molina et al. 2017). In three of four biological replicates, we detected loss of Tet1 binding at several sites—$Lhx5$(–91.6), $Foxp1$(+42.5), $Sox1$ (+6.2), and $Sox1$(+82.6)—upon Zic2 depletion (Fig. 3F; Supplemental Fig. S3C). These results suggest that Zic2 may recruit Tet1 to new regulatory loci in the primed epiblast and regulate germ layer fate.

Distinct engagement of Tet1 and Oct4 at primed enhancers

As previously reported, Oct4-bound genomic sites are globally reorganized at enhancers during ESC-to-EpiLC conversion (Buecker et al. 2014). We asked how our observed repatterning of Tet1 genomic occupancy sites compared with those of Buecker et al.'s (2014) Oct4 ChIP-seq data sets. The global overview of all ChIP-seq peaks in both data sets suggests that Tet1 may undergo an even more pronounced genome-scale redistribution compared with Oct4 during the cell state transitions (Supplemental Fig. S4A). While total numbers of Oct4-bound sites stayed relatively similar between the two states, in line with sustained expression of Oct4 protein, over twofold more Tet1 occupancy sites were detected in EpiLCs than in 2iL ESCs, with the gains mostly within distal intergenic and exonic regions (Supplemental Fig. S4A).
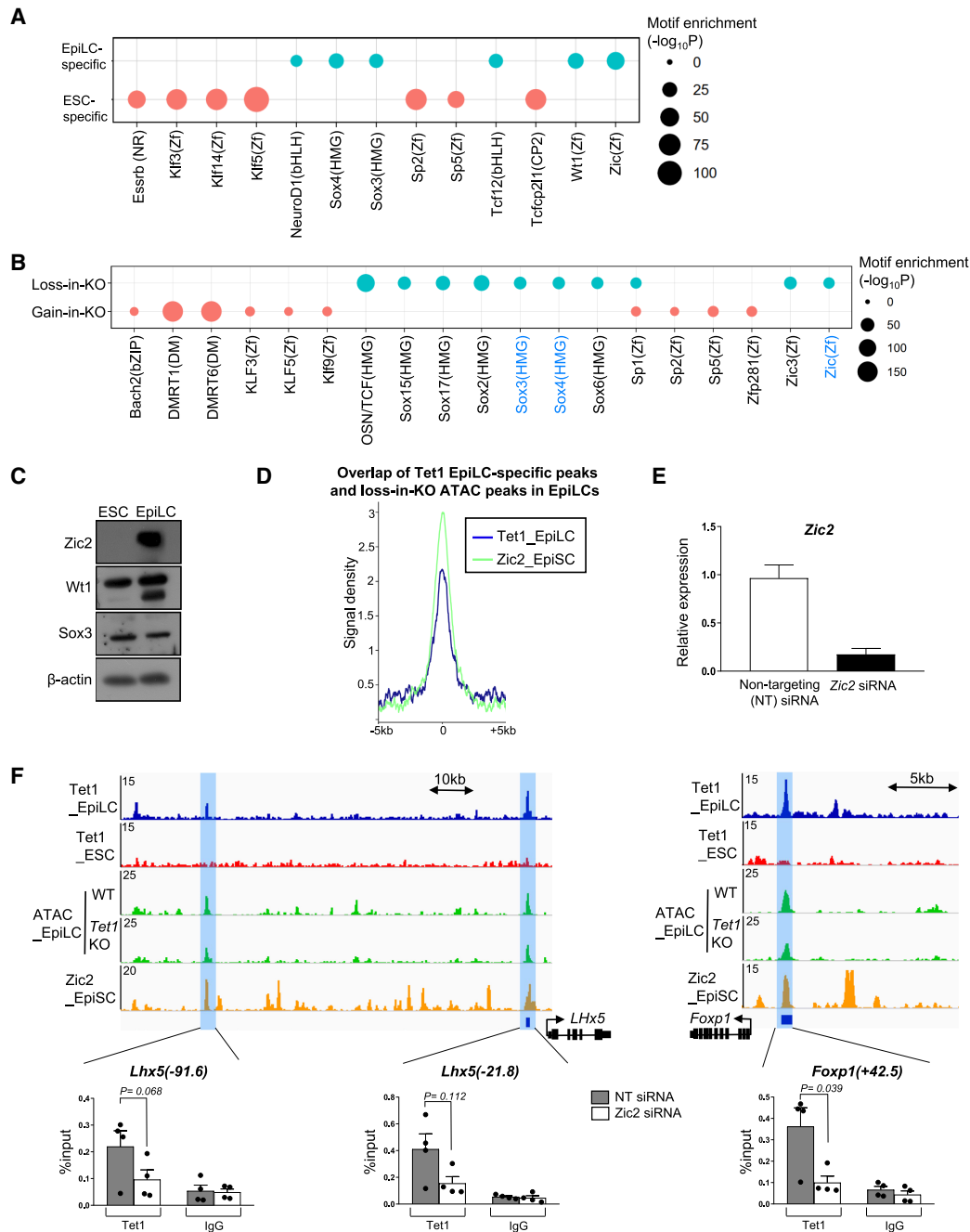
Figure 3. Tet1 engagement at primed loci depends on Zic2. (*A*) Bubble plots indicating transcription factors associated with top enriched motifs in Tet1-bound EpiLC-specific and ESC-specific peaks analyzed by HOMER (Heinz et al. 2010). (*B*) Bubble plots indicating transcription factors associated with top enriched motifs in loss-in-KO and gain-in-KO ATAC-seq peaks in EpiLCs. Motif names highlighted in blue are enriched in both Tet1-bound EpiLC-specific peaks and loss-in-KO ATAC-seq peaks in EpiLCs. (*C*) Western blot analysis of Zic2, Wt1, and Sox3 expression in 2iL-cultured ESCs and EpiLCs. (*D*) Density aggregation plot of Tet1 ChIP-seq signal densities in EpiLCs and Zic2 ChIP-seq signal densities in EpiSCs (Matsuda et al. 2017) at the overlap of Tet1 EpiLC-specific peaks and loss-in-KO ATAC-seq peaks in EpiLCs. (*E*) Knockdown efficiency in nontargeting (NT) siRNA or Zic2 siRNA treated EpiLCs. Data shown are mean values ± SEM (*n* = 3) from three independent cell lines. (*F*) ChIP-qPCR analysis of Tet1 binding at the *Lhx5(-91.6)*, *Lhx5(-21.8)* (*left* panel), and *Foxp1(+42.5)* (*right* panel) in *Zic2* knockdown EpiLCs. (*Top*) IGV snapshots of Tet1 ChIP-seq tracks in EpiLCs and 2iL-cultured ESCs, ATAC-seq in WT and *Tet1*$^{GT/GT}$ (KO) EpiLCs, and Zic2 ChIP-seq tracks in EpiSCs at *Lhx5* (*left* panel) and *Foxp1* (*right* panel). Gene structure annotations are shown *below* each track set. Regions where EpiLC-specific Tet1 peaks (BED annotations of statistically significant peaks shown as blue bars below tracks) coincide with loss-in-KO ATAC-seq peaks in EpiLCs are highlighted in blue. (*Bottom*) ChIP-qPCR data shown are mean ± SEM (*n* = 4) from three independent cell lines. Position coordinates in parenthesis are with respect to TSS in kilobases. The enrichments of ChIP samples are normalized to the values of the corresponding input.

When comparing normalized binding intensities centered at Tet1-bound peaks, we observed similarly strong Oct4 co-occupancy at ESC-specific Tet1-bound peaks in 2iL ESCs; at these loci, both Tet1 and Oct4 binding intensities diminished in EpiLCs (Fig. 4A). Venn analysis incorporating ESC-specific Oct4-bound peaks showed an overlap with 47% of ESC-specific Tet1-bound peaks (Fig. 4B). At EpiLC-specific Tet1 peaks, Oct4-binding intensities in EpiLCs were similar to pre-existing Oct4 levels in ESCs, but comparatively weaker relative to those of Tet1 in EpiLCs (Fig. 4C). Only 8% of these EpiLC-specific Tet1-bound peaks overlapped with EpiLC-specific Oct4-bound peaks (Fig. 4D). Because of the larger size of Buecker et al.'s (2014) Oct4 ChIP-seq peaks, we reclassified those by their genome distribution either at promoters or intergenic regions. When centered on ESC-specific Oct4-bound peaks, colocalization of Tet1 and Oct4 binding was evident only at promoter but not intergenic regions in ESCs (Supplemental Fig. S4B). At EpiLC-specific Oct4-bound peaks, Tet1-binding intensities were relatively weak at both promoters and intergenic regions. Collectively, these data suggest that Tet1 and Oct4 coregulate genes at promoter regions in the naive state but diverge to engage distinct sets of promoters and distal enhancers in the primed state.

In further support that Tet1 and Oct4 regulate disparate EpiLC loci, we examined our peak regions displaying differential ATAC-seq signals in $Tet1^{GT/GT}$ EpiLCs with Buecker et al.'s (2014) Oct4 ChIP-seq data and observed undetectable signals for Oct4 binding at these loci in WT EpiLCs (Fig. 4E). In another study that examined WT and Smad2/3-double mutant EpiLCs by ATAC-seq, Smad2/3-dependent sites that display differential accessibility were shown to colocalize with Oct4 binding (Senft et al. 2018). Examining ATAC-seq signals from that data set at our Tet1-dependent regions showed that chromatin accessibility at Tet1-regulated sites were not affected by combined loss of Smad2 and Smad3 (Fig. 4E). Collectively, these analyses suggest that the effects of Tet1 on chromatin accessibility in EpiLCs are independent of Nodal signal effectors and Oct4.

Based on these observations, we hypothesize that Tet1 and Oct4 may function antagonistically of each other in EpiLCs. Therefore, we performed Oct4 ChIP-seq analysis in Tet1-deficient EpiLCs to determine whether Oct4 genomic occupancies may be altered in the absence of Tet1. To facilitate Oct4 ChIP-seq, we derived WT and Tet1 KO ESCs from a new B6129S6F1 hybrid strain harboring a $Tet1^{tm1Koh}$ allele by which knock-in/knockout gene targeting disrupted transcription of the entire Tet1 coding region (Khoueiry et al. 2017). This hybrid strain, on a background similar to the v6.5 strain used by Buecker et al. (2014), also facilitated subsequent ESC differentiation assays as EBs or monolayer cultures (see Figs. 4, 5). In ESCs and EpiLCs, absence of Tet1 did not affect Oct4 expression (Supplemental Fig. S4C,D). Using MACS2 bdgdiff module (Zhang et al. 2008) to compare Oct4 ChIP-seq peaks collated from WT and Tet1 KO EpiLC biological replicates ($n = 2$), we identified 1244 Oct4 ChIP-seq peaks, of which 1151 (92.5%) were differentially

higher in Tet1 KO, 93 (7.5%) similar in WT and KO, but none lower in KO. Normalized read densities of all 1244 Oct4 ChIP-seq peaks showed significantly increased Oct4 occupancy upon loss of Tet1 (Fig. 4F). These peak regions were lowly accessible sites engaged by Oct4 in WT cells and associated with 1602 genes, many of which are involved in lineage development (Supplemental Table S4). The most striking examples of enhanced Oct4-binding sites include a proximal promoter element of Oct4 itself (Fig. 4G), Nodal targets including Foxh1, Lefty1, Skil, and Agpat3 (Fig. 4H), and distal sites of Wnt signaling components Axin2, Lef1, and Fzd7 (Supplemental Fig. S4E). Similar enhanced Oct4 genomic occupancies were recapitulated in Oct4 ChIP-seq analysis of B6.129P2-$Tet1^{GT/GT}$ EpiLCs (Fig. 4G,H; Supplemental Fig. S4E).

Focusing on the 1151 up-in-KO Oct4 peaks in EpiLCs, we found minimal overlap with EpiLC-specific (only 49 or 4.3%) or with common Tet1-bound regions (72 or 6.3%) (Fig. 4I). Therefore, at the majority of these up-in-KO Oct4 ChIP-seq peak regions, Oct4 is not replacing Tet1. GO analysis of these up-in-KO Oct4 ChIP-seq peaks showed enrichment for somatic stem cell population maintenance, but region and gene set coverages were <5% (Supplemental Table S5). Using an alternative differential binding analysis tool DiffBind (Ross-Innes et al. 2012), we defined a smaller subset of 388 up-in-KO peaks (Supplemental Table S4) at FDR < 0.1, of which 337 (86.9%) overlapped with the 1151 differential peaks defined using MACS2 bdgdiff. These 388 up-in-KO Oct4 peaks were more distinctly enriched for GO biological processes for morphogenesis of the embryonic epithelium and Mouse Genome Informatics (MGI) phenotype ontology related to abnormal neural tube morphology (highest-ranked term with region and gene set coverages of 13.66% and 8.07%, respectively) (Fig. 4J,K; Supplemental Table S5). Collectively, these observations suggest that the presence of Tet1 at primed enhancers may indirectly regulate chromatin affinity of Oct4 at distant sites, preventing precocious engagement at Nodal and Wnt targets during lineage priming and safeguarding proper development of the neuroepithelium.

Tet1 coordinates ME and NE fate choice via regulation of the Oct4–Nodal–Wnt axis

To follow up on the induction of key developmental genes during germ layer differentiation, we adapted WT and Tet1 KO ESC lines of B6129S6F1-$Tet1^{tm1Koh}$ strain to serum- and feeder-free conditions followed by 4 d of cellular aggregation as EBs (Fig. 5A). Gene expression of key lineage genes were measured over a 4-d time course by RT-qPCR. In defined culture media used in this assay, differentiation in the absence of exogenously added growth factors (nondirected) favored induction of NE markers (Pax6 and Sox1) in WT cells (Ying et al. 2003); addition of Activin A (25 ng/mL) activated primitive streak ME and repressed NE markers by day 4 (Gadue et al. 2006). In Tet1 KO cells, however, ME marker expression (Brachyury, Eomes, Mixl, Foxa2, and Goosecoid) were elevated by day 3 without requiring exogenous Activin A
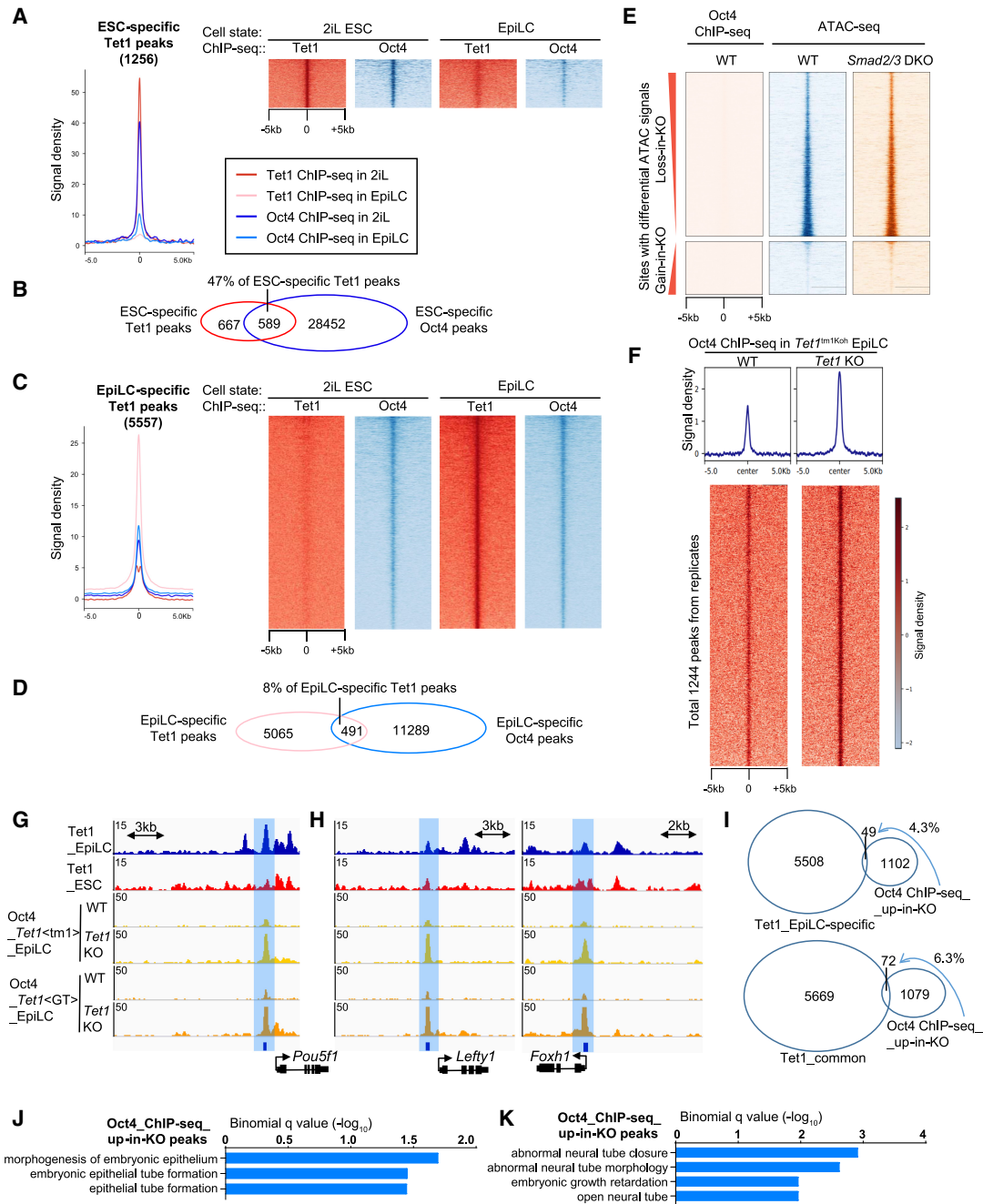
Figure 4. Distinct engagement of Tet1 and Oct4 at primed enhancers. (*A*) Heat map and density aggregation plots of Tet1 and Oct4 ChIP-seq signal densities in 2iL-cultured ESCs and EpiLCs centered at Tet1-bound ESC-specific peaks. (*B*) Venn diagram indicating the overlap of Tet1-bound ESC-specific peaks and Oct4-bound ESC-specific peaks. (*C*) Same as in 3A, but centered at Tet1-bound EpiLC-specific peaks. (*D*) Same as in 3B for Tet1-bound and Oct4-bound EpiLC-specific peaks. (*E*) Heat map of Oct4 ChIP-seq signal densities in EpiLCs (Buecker et al. 2014) and ATAC-seq signal densities in WT and *Smad2;Smad3* KO EpiLCs (Senft et al. 2018) at differential accessibility peaks defined in Figure 2A. (*F*) Heat map and density aggregation plot of Oct4 ChIP-seq signal densities in WT and *Tet1* KO EpiLCs of B6129S6F1-*Tet1*^tm1Koh strain at total 1244 Oct4-bound peaks defined by MACS2 bdgdiff module. See Supplemental Table S4 for an annotated list of all peaks. (*G,H*) IGV snapshots of Tet1 ChIP-seq tracks in EpiLCs and 2iL-cultured ESCs, and Oct4 ChIP-seq tracks WT and *Tet1* KO EpiLCs, at indicated loci located at proximal promoters of *Oct4* (*G*) and Nodal targets *Lefty1* and *Foxh1* (*H*). Gene annotations are shown *below* each track set. Blue highlights indicate statistically differential peaks of Oct4 binding identified by MACS2 in two strains of *Tet1*-deficient EpiLCs (BED annotations shown as blue bars *below* tracks). (*I*) Venn diagram indicating the overlap of Oct4-bound up-in-KO ChIP-seq peaks in EpiLCs with Tet1-bound EpiLC-specific (*top* panel) or Tet1-bound common (*bottom* panel) ChIP-seq peaks. (*J,K*) Functional annotation of GO terms (*J*) and phenotype annotation of mouse genome informatics (MGI) ontology (*K*) enriched in 388 Oct4-bound up-in-KO ChIP-seq peaks in EpiLCs defined by R/Bioconductor package DiffBind at FDR < 0.1. See Supplemental Table S5 for an annotated list of ranked GO terms.
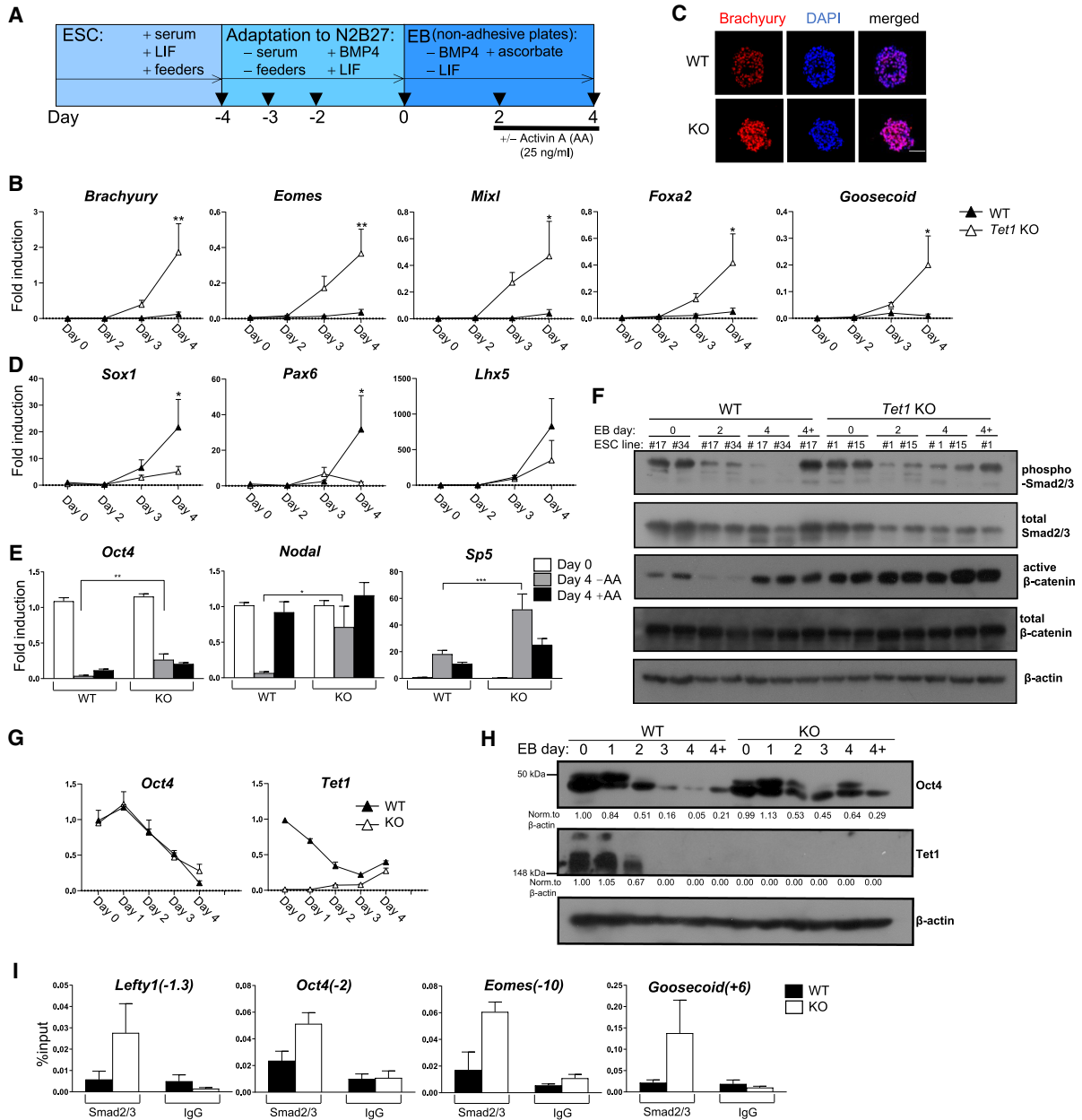
Figure 5. Tet1 coordinates ME and NE fate choice via regulation of the Oct4–Nodal–Wnt axis (*A*) Schematics of serum-free embryoid bodies (EB) differentiation assay. Activin A (AA) was supplied at day 2 as a positive control to stimulate ME differentiation. Arrowheads denote cell splitting or EB reaggregations. (*B*) Time-course expression analysis of ME markers (*Brachyury, Eomes, Mixl, Foxa2,* and *Goosecoid*) in WT and *Tet1*[tm1Koh] (KO) EBs. Data shown are mean ± SEM of *n* = 3 biological replicate ESC lines per genotype. Normalized expression values are relative to values of WT EBs at day 4 without AA stimulation for ME markers and relative to values of WT EBs at day 0 for NE markers. Statistical test: two-way ANOVA. (*) $P < 0.05$; (**) $P < 0.01$; (***) $P < 0.001$. (*C*) Immunofluorescence staining of Brachyury (red) in sectioned WT and *Tet1* KO EBs at day 4. DAPI (blue) indicates the nuclei. Scale bar, 50 µm. (*D*) Same as in 5B for NE markers (*Sox1, Pax6,* and *Lhx5*) in WT and *Tet1* KO EBs. (*E*) Expression of *Oct4, Nodal,* and Wnt signaling target *Sp5* in WT and *Tet1* KO EBs. Normalized expression values are relative to values of WT EBs at day 0. Data shown are mean ± SEM (*n* = 4) from three biological replicate ESC lines per genotype. (*F*) Western blot analysis of phospho-Smad2, total Smad2/3, active-β-catenin, and total β-catenin in whole-cell lysates of day 0, 2, and 4 WT and *Tet1* KO EBs. (4+) WT EBs at day 4 treated with exogenous AA (25 ng/mL). Biological replicate lines per genotype are presented. (*G*) Time-course expression analysis of *Oct4* and *Tet1* in WT and *Tet1* KO EBs. Data shown are mean ± SEM (*n* = 3) from three biological replicate ESC lines per genotype. Normalized expression values are relative to the values of WT EBs at day 0. (*H*) Western blot time-course analysis of Oct4 and Tet1 in whole-cell lysates from WT and *Tet1* KO EBs. (4+) WT EBs at day 4 stimulated with AA (25 ng/mL). Relative values of Oct4 and Tet1 band intensities normalized to β-actin levels are shown below the respective signal bands. (*I*) ChIP-qPCR analysis of Smad2/3 binding to the *Lefty1* (−1.3) and *Oct4* (−2) proximal promoters, and *Eomes* (−10), and *Goosecoid* (+6) distal enhancers in day 4 WT and *Tet1* KO EBs. Position coordinates in parenthesis are with respect to TSS in kilobases. The enrichments of ChIP samples are normalized to the values of the corresponding input.

(Fig. 5B; Supplemental Fig S5A). A homogeneous pattern of ME induction in KO EBs was verified by enhanced Brachyury immunofluorescence staining in almost all cells in EB sections (Fig. 5C). This skewing phenotype toward ME differentiation recapitulates our previous observations following acute *Tet1* depletion by siRNA in v6.5 ESCs (Koh et al. 2011). Furthermore, nondirected differentiation was associated with decreased expression of NE (*Pax6*, *Sox1*, and *Lhx5*) genes by day 4 (Fig. 5D).

Naive pluripotency markers, including *Klf4*, *Esrrb*, and *Rex1*, were fully silenced by day 4 in WT and KO EBs (data not shown), verifying that exit from pluripotency was not inhibited by loss of *Tet1*. However, we consistently observed sustained expression of *Oct4* in KO cells compared with its fully repressed levels in WT EBs by day 4, along with elevated expression of *Nodal* and the Wnt target *Sp5* (Fig. 5E). By performing Western blot analysis during the differentiation time course, we verified that following removal of BMP4 and LIF at day 0, active phosphorylated Smad2/3 proteins were extinguished by day 4 in WT EBs unless Activin A was added; in contrast, Smad2/3 proteins sustained their phosphorylation status in KO EBs between days 2 and 4 (Fig. 5F). In concert with these differences, active β-catenin proteins were detectable at higher levels in KO compared with WT EBs throughout the differentiation time courses. *Oct4* transcripts were down-regulated by day 3 in WT EBs, following down-regulation of *Tet1* by day 2 (Fig. 5G). In contrast, *Tet1* KO EBs sustained Oct4 protein expression even when *Oct4* transcripts were down-regulated by day 4, suggesting that Tet1 may regulate Oct4 levels during germ layer differentiation by a posttranslational mechanism (Fig. 5H). Similar observations were made using incipient congenic B6-*Tet1*[GT/GT] ESCs (Supplemental Fig. S5B,C). Moreover, we often observed Oct4 signals to appear as doublet bands of about 50 kDa, higher than the expected molecular mass of 37 kDa, which may suggest the presence of posttranslational modifications such as O-GlcNAcylation (Jang et al. 2012). To further validate elevated Nodal signals at key ME lineage genes in *Tet1* KO EBs, we performed Smad2 ChIP and interrogated Smad2 recruitment to known Smad2/3 target loci (Mullen et al. 2011; Wang et al. 2017). At proximal promoter sites of *Oct4* and *Lefty1* and distal sites of *Eomes* and *Goosecoid*, increased Smad2 occupancy were detected in *Tet1* KO compared with WT EBs (Fig. 5I).

Previous studies have shown collaborative interactions between Wnt/β-catenin, Nodal–Smad2/3, and Oct4 in the regulation of ME differentiation of mouse and human ESCs (Funa et al. 2015; Wang et al. 2017). To clarify how these effectors modulate differentiation phenotypes downstream from Tet1, we inhibited each of these three components in *Tet1* KO EBs (Fig. 6A). First, we silenced Oct4 expression in EBs by applying two rounds of *Oct4* SMARTpool siRNA transfection during serum-free EB differentiation at days 0 and 2 (Fig. 6B). In line with a contribution of Oct4 toward ME skewing in *Tet1* KO EBs, depletion of *Oct4* expression in *Tet1* KO EBs reduced expression of *Brachyury*, *Mixl1*, and *Eomes* to near WT levels by day 4 (Fig. 6C). However, *Oct4* depletion can enhance expression of the neural lineage marker *Sox1* and

Lhx5 in WT EBs, but cannot restore expression in *Tet1* KO EBs to WT levels (Fig. 6D). Oct4 depletion in ESCs results in spontaneous loss of self-renewal and differentiation into extraembryonic lineages (Niwa et al. 2000; Radzisheuskaya et al. 2013). In our approach, in which we depleted *Oct4* after induction of EB differentiation, silencing of naive markers (*Klf4*, *Esrrb*, and *Rex1*) was not affected in WT EBs (Supplemental Fig. S6A); among trophoblast markers (*Cdx2*, *Hand1*, and *Mash2*) tested, only *Cdx2* expression was increased by *Oct4* depletion. In *Tet1* KO EBs, further loss of Oct4 resulted in small increases in *Klf4* and *Rex1* expression, and more pronounced differential up-regulation of *Hand1* and *Mash2* by day 4 (Supplemental Fig. S6B). Collectively, these results reveal the contribution of Oct4 in embryonic and extraembryonic differentiation in the context of Tet1 loss of function.

Next, we added chemical inhibitors during EB reaggregation at day 2 to block signaling components, namely, SB431542 to inhibit Nodal–Smad2/3 pathway, IWP-2 to inhibit Wnt palmitoylation, and XAV939 to stabilize the Wnt negative regulator Axin2 (Anastas and Moon 2013). To access whether these treatments provided effective blockade of relevant signaling, we examined gene expression of *Nodal* and the Wnt targets *Axin2* and *Sp5*. As expected, hyperactivated *Nodal* expression levels in *Tet1* KO EBs were fully abrogated by SB431542, but also by the Wnt inhibitors, consistent with an action of Wnt upstream of Nodal-Smad2/3 (Fig. 6E). On the other hand, the Wnt target genes were effectively repressed by IWP-2 and XAV939, but variably enhanced in WT and KO EBs by SB431542, suggesting that Nodal signals may provide negative feedback regulation of Wnt (Fig. 6E). The effects of inhibitor treatment on *Brachyury*, Eomes, and *Mixl* expression phenocopied the effects on *Nodal*, in agreement with these ME genes being downstream targets of collaborative Wnt and Nodal inputs (Fig. 6F). In contrast, expression of *Pax6*, *Sox1* and *Lhx5* in KO EBs can be rescued to nearly WT levels by SB431542, but not by the Wnt inhibitors, suggesting that hyperactivated Nodal signals in *Tet1* KO EBs cross-inhibits these NE markers while promoting ME fate (Fig. 6G). An additional treatment with inhibitors at day 0 in this assay produced no further changes (data not shown), in agreement with the action of extracellular cues occurring between days 2 to 4 of differentiation (Fig. 5F; Gadue et al. 2006).

Loss of Tet1 impairs differentiation into anterior neuroectoderm lineage

Because expression of NE marker genes can be variable in the nondirected differentiation assay described above, we tested our ESC lines upon inductive neuronal differentiation. In the first differentiation protocol based on supplementation with retinoic acid (RA) (Bibel et al. 2007), we collected EBs in serum media after a 4-d treatment with RA at 5 μM and replated the cells for terminal neuronal differentiation (Fig. 7A). *Tet1*[GT/GT] ESCs generated EBs with severely impaired expression of NE markers *Pax6*, *Sox1*, *NeuroD1*, and *Nestin*, concomitantly with elevated *Oct4* levels (Supplemental Fig. S7A,B). Because these
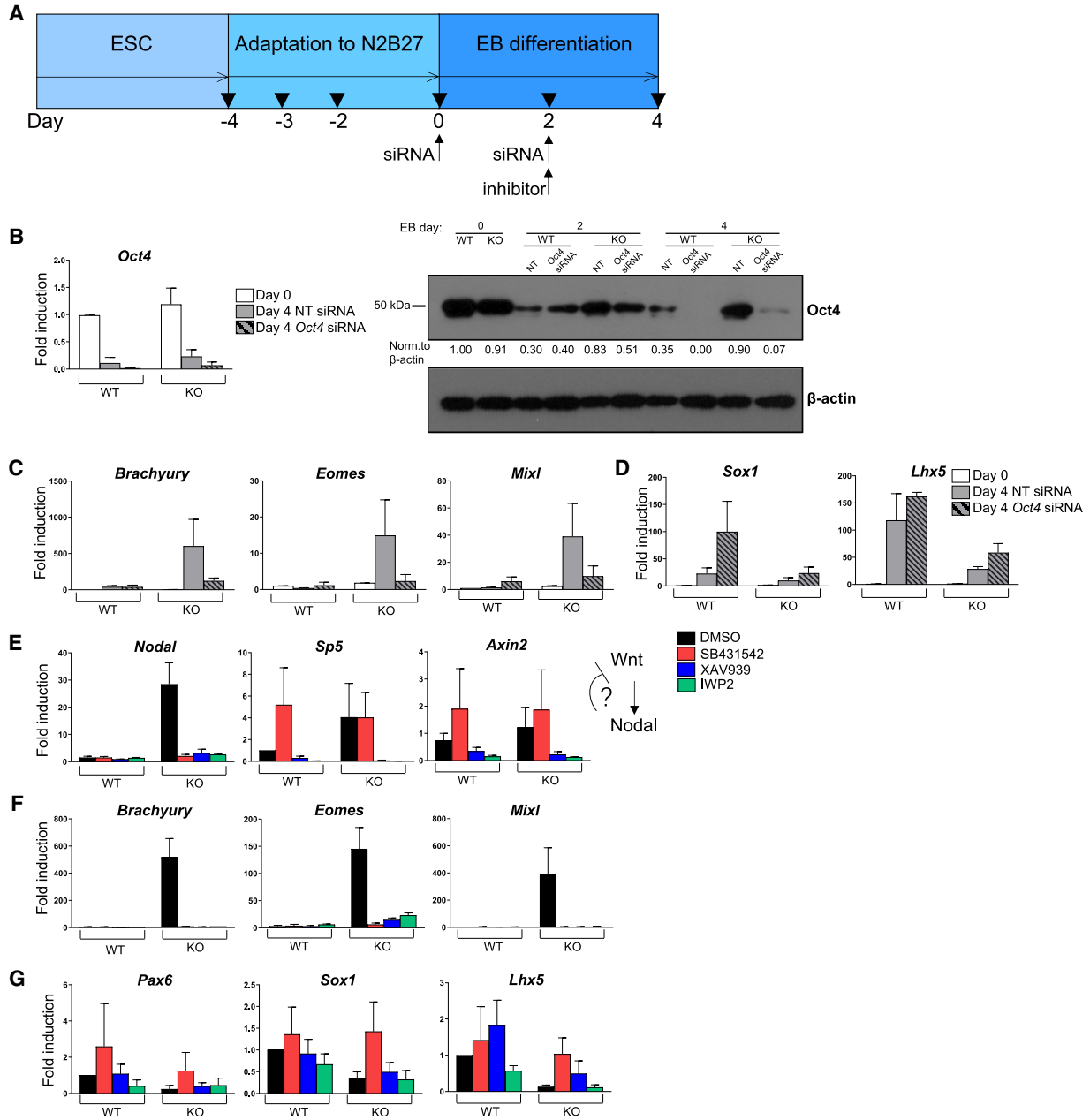
Figure 6. *Oct4* knockdown or Wnt/Nodal signaling blockade rescue ME skewing in *Tet1* KO. (*A*) Schematics of serum-free EB differentiation assays including treatment with *Oct4* siRNA or signaling inhibitors. To silence the expression of *Oct4* completely, two rounds of Oct4 siRNA were added at day 0 and day 2. Signaling inhibitors, including SB431542, XAV939, and IWP2, or DMSO as control were added at day 2 only. (*B*) RT-qPCR (*left* panel) and Western blot (*right* panel) analysis of *Oct4* expression in days 0, 2, and 4 WT and *Tet1*[tm1Koh] KO EBs treated with *Oct4* siRNA. RT-qPCR data shown are mean ± SEM of $n = 3$ independent experiments from two lines per genotype. Relative values of Oct4 band intensities normalized to β-actin levels are shown *below* the signal bands. (*C,D*) Expression of ME markers (*C*) and NE markers (*D*) in WT and *Tet1*[tm1Koh] KO EBs treated with *Oct4* siRNA. Data shown are mean values ± SEM of $n = 3$ independent experiments from two lines per genotype. (*E–G*) Expression of Nodal and Wnt target genes (*E*), ME markers (*F*), and NE markers (*G*) in WT and *Tet1*[tm1Koh] (KO) EBs treated with signaling inhibitors. Schematic at the *right* in *E* indicates the inferred Wnt and Nodal signaling cascade. Data shown are the mean ± SEM of $n = 3$ independent experiments from three biological replicate lines per genotype. Normalized expression values are relative to values of WT EBs at day 0.

genes are normally activated to high levels in WT EBs after treatment with RA at day 4, when *Tet1* expression is no longer detected, we asked how the chromatin status at their gene promoters or enhancers were altered

by absence of Tet1. We performed ChIP-qPCR analysis for the presence of histone H3 Lys27 trimethylation (H3K27me3) and acetylation (H3K27ac), which are mutually exclusive marks for poised/silenced and active
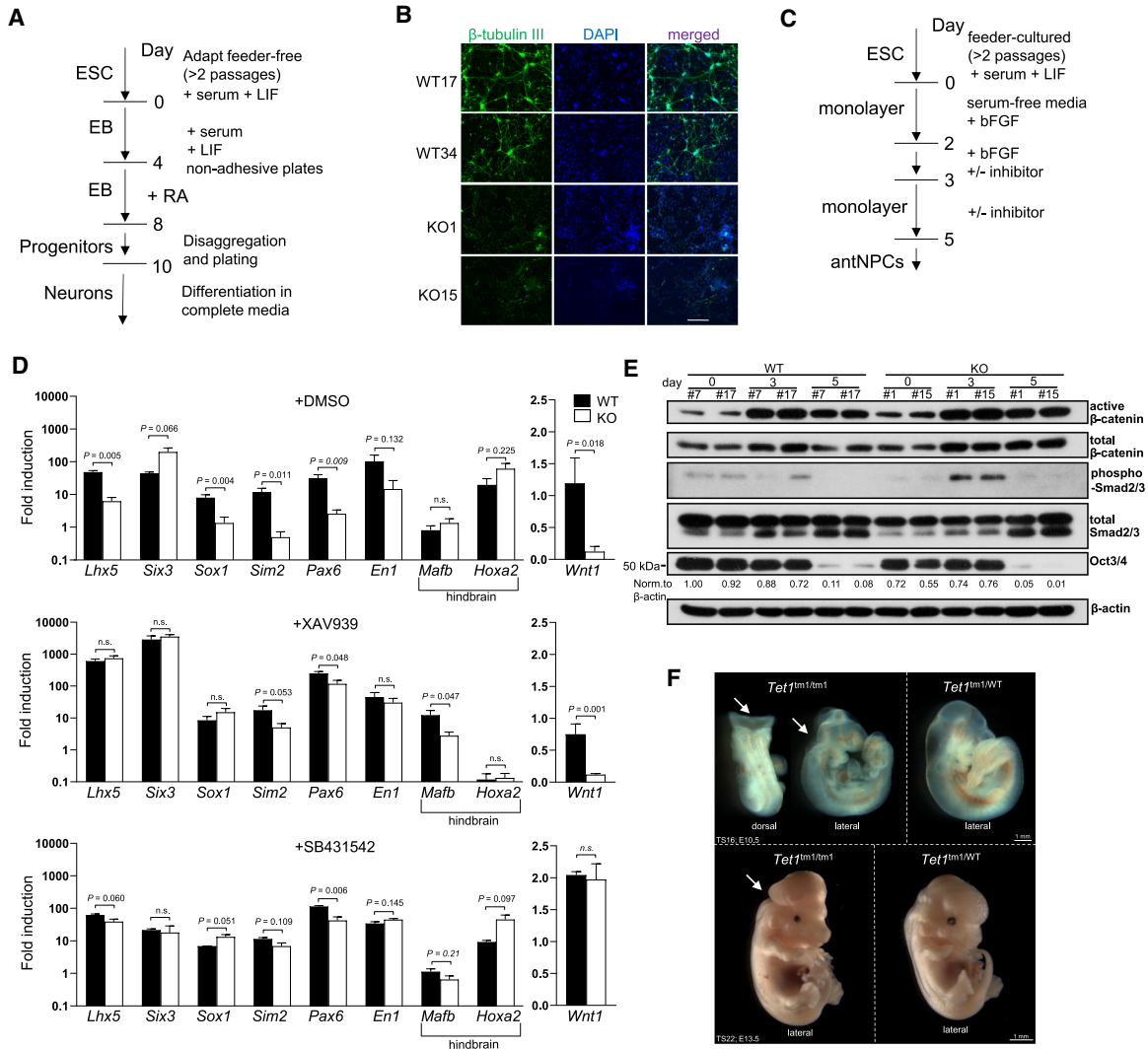
Figure 7. Loss of *Tet1* impairs differentiation into anterior neuroectoderm lineage. (*A*) Schematics of retinoic acid (RA)–directed EB differentiation assay. (*B*) Immunofluorescence staining of the neural marker β-tubulin III (green) in terminal neurons differentiated from WT (WT17 and WT34) and KO (KO1 and KO15) ESCs of strain B6129S6F1*Tet1*[tm1Koh]. DAPI (blue) indicates the nuclei. Scale bars, 100 μm. (*C*) Schematics of anterior neural progenitor cell (antNPC) monolayer differentiation assay. ESCs are cultured in serum-free medium supplied with bFGF for 3 d. Inhibitors or DMSO control were added from day 2. (*D*) Gene expression of neural markers in WT and *Tet1*[tm1Koh] KO ESCs following differentiation into antNPCs treated with the Wnt signaling inhibitor XAV939, Nodal signaling inhibitor SB431542, or DMSO (vehicle control). Except for hindbrain markers *Mafb* and *Hoxa2*, all other genes are markers of anterior fate. Data shown are the mean ± SEM of two independent differentiation experiments in which a total of $n = 5$ or 6 biological replicate lines per genotype were analyzed. Fold induction is expressed relative to values of a WT cell line at day 0 on a $\log_{10}$ scale, except *Wnt1* for which values are relative to a WT line at day 5 on a linear scale. *P*-values are obtained by multiple *t*-tests. (*E*) Western blot analysis of phospho-Smad2, total Smad2/3, active-β-catenin, total β-catenin, and Oct4 in whole-cell lysates of days 0, 3, and 5 WT and *Tet1*[tm1Koh] (KO) antNPCs. Biological replicate lines per genotype are presented. Relative values of Oct4 band intensities normalized to β-actin levels are shown below the signal bands. (*F*) Embryos at E10.5 (Theiler stage 16) and E13.5 (Theiler stage 22) that are homozygous or heterozygous for the *Tet1*[tm1Koh] allele on a C57BL/6 background. Failed neuropore closure is indicated with a white arrow. Scale bar, 1 mm.

regulatory regions, respectively (Creyghton et al. 2010; Rada-Iglesias et al. 2011), at distal regulatory loci of *Pax6*, *Sox1*, and *Nestin* (Supplemental Fig. S7C). At day 0 of EB differentiation onset, we did not observe any differences in the presence of these marks at all three loci between WT and *Tet1*[GT/GT] ESCs, arguing against any effect of Tet1 on pre-existing histone status prior to activation by differentiation cues. By day 6 of EB aggregation, RA

induced gains in H3K27ac at the activated NE gene loci in WT cells; these chromatin changes were completely abolished in *Tet1*[GT/GT] EBs, indicating a severe defective cellular response to RA stimulation. Because both WT and *Tet1*[GT/GT] ESCs of the B6.129P2 strain cannot adhere well to substrates during replating, we switched to comparing WT and KO ESCs of strain B6129S6F1-*Tet1*[tm1Koh] to test further monolayer differentiation into terminal

neurons expressing β-tubulin III (also known as Tuj1). We observed extensive loss of Tuj1-expressing neurons in KO lines compared with WT lines (Fig. 7B). These phenotypes were reproduced during terminal neuronal differentiation of E14Tg2a ESC lines (129 strain) in which *Tet1* is stably depleted (Supplemental Fig. S7D,E) and also during differentiation of induced pluripotent stem cells (iPSCs) reprogrammed from KO mouse embryonic fibroblasts of mixed B6;129S6-*Tet1*^tm1Koh strain (Supplemental Fig. S7F; Bartoccetti et al. 2020).

To further dissect the effects of *Tet1* loss of function in NE lineage development, we used a more recently established serum-free protocol that directed mouse ESCs to differentiate into anterior neural progenitor cells (antNPCs) on monolayer cultures by supplementation with fibroblast growth factor-basic (bFGF) (Fig. 7C; Gouti et al. 2014; Cruz-Molina et al. 2017). In this system, *Tet1* transcripts and protein expression were lost by day 3, but *Oct4* expression persisted until day 4 (Supplemental Fig. S7G,H). In the absence of small molecule inhibitors, several anterior NPC fate markers *Lhx5*, *Sim2*, and *Wnt1* were significantly compromised in *Tet1* KO compared with WT cells (Fig. 7D). Upon blockade of posteriorization activities provided by Wnt signaling, or Nodal inhibition, expression of *Lhx5*, *Sim2*, *Sox1*, *Pax6*, and *En1* were rescued in *Tet1* KO ESCs to WT levels (Fig. 7D). In agreement with hyperactivated Nodal and canonical Wnt signaling, we observed increased expression of active β-catenin and phosphorylated Smad2/3 in KO cells by day 3 of differentiation (Fig. 7E). These results suggest that impaired neuronal differentiation potential in *Tet1* KO ESCs is not the result of inherent NE gene defects, but largely a response to altered extrinsic signaling inputs.

Finally, we examined mouse embryos in midgestation using our B6-*Tet1*^tm1Koh strain. From heterozygous-intercross timed pregnancies, we consistently observed failure of midbrain–hindbrain neuropore closure in E10.5 *Tet1*-null embryos, which manifested as exencephaly by E13.5 (Fig. 7F). We previously described that the first morphological defects in *Tet1*-deficient embryos were detected as deformities at E8.5 after gastrulation, coincident with initiation of neuropore closure, and failed anterior neuropore closure at E9.5 (Khoueiry et al. 2017). Therefore, regulatory control of anterior neuroectodermal lineage potential by Tet1 is critical for proper neural tube closure.

Discussion

Our results provide a molecular framework for understanding how a pre-existing epigenetic landscape shaped by Tet1 in the epiblast influences the integration of cell-intrinsic pluripotency factor expression and extrinsic signaling cues during subsequent germ layer specification. While Tet1 is highly expressed in both naive and primed pluripotent states, the genomic occupancy patterns are dynamic and related to distinct biological functions relevant to each cell state. Moreover, de novo Tet1 occupancy during lineage priming affects chromatin accessibility at developmental loci comprising neural fate genes. These

Tet1-dependent primed loci appear distinct from those engaged by the essential pluripotency factor Oct4 and independent of Smad2/3 regulation. However, absence of Tet1 enhances Oct4 binding to a subset of its genomic targets, including *Oct4* itself and Nodal and Wnt targets, unleashing a feed-forward loop that amplifies Smad2/3 and canonical Wnt signaling recruitment to ME genes during differentiation when *Tet1* expression is no longer sustained. We propose a model in which Tet1 safeguards both DNA methylation fidelity and chromatin accessibility at neural fate developmental loci and indirectly modulates occupancy by Oct4 at key ME-fate loci in postimplantation epiblast cells. Entry into the primitive streak then requires rapid loss of *Tet1* expression to allow Oct4 to collaborate with Wnt and Nodal signals to induce ME while suppressing NE differentiation.

While Oct4 exhibits enhancer repatterning during conversion of naive ESCs to primed EpiLCs, its dynamic genomic occupancy appears to be dependent on other cell state-specific TFs, including Zic2/3 and Otx2 (Buecker et al. 2014). Similarly, we observed that EpiLC-specific Tet1 bound regions are enriched for consensus motifs for Wt-1, Zic, and Sox family factors, but not Otx2, suggesting that a distinct set of partner TFs may facilitate engagement of Tet1 at primed enhancers separately from those engaged by Oct4. Accessibility of these loci are lost in the absence of Tet1 but cannot be restored by rescued re-expression of Tet1, arguing against a pioneer function for Tet1 as a "de novo" DNA demethylase that can open up closed chromatin. Consistent with having a highly generic DNA sequence specificity, the N-terminal CXXC zinc finger domain in Tet1 confers an affinity for unmethylated CpG motifs, enriched in the majority of mammalian gene promoters (Zhang et al. 2010; Frauer et al. 2011), such that Tet1 by itself binds predominantly to CpG-rich regions in ESCs where it preserves methylation fidelity by preventing aberrant de novo DNA methylation activity (Williams et al. 2011; Wu et al. 2011; Xu et al. 2011). Here, we showed that Zic2 may be a key cofactor recruiting Tet1 to its primed enhancer sites, where Tet1 can then promote an open chromatin state.

Defined basal media conditions in ESC differentiation models favor development along the neural lineage in the absence of extrinsic signals, supporting a "neural default" fate in epiblast cells (Ozair et al. 2013). In support of this model, a recent study demonstrated that Polycomb proteins facilitate a permissive regulatory topology in ESCs at a set of poised enhancers that are necessary for the induction of anterior neural genes during differentiation (Cruz-Molina et al. 2017). Our study adds *Tet1* as another important component of the "neural default" wiring in primed epiblast cells by regulating the chromatin states at neurodevelopmental loci, possibly as a "maintenance" DNA demethylase. A recent study showed that *TET1*, *TET2*, and *TET3* triple-KO in human ESCs results in DMNT3B-mediated DNA hypermethylation predominantly at bivalent promoters and a neural differentiation defect associated with failure of *PAX6* induction during differentiation, which can be partially rescued by re-expression of *TET1* (Verma et al. 2018). These studies,

much in agreement with our observations in this study and previously in *Tet1* KO mice (Khoueiry et al. 2017), suggest that disruption of the epigenetic landscape in primed pluripotency can occur without corresponding effects on gene expression at pregastrulation, but affect activation of neural genes later during development. Possible effects of *Tet1* KO on recruitment of Polycomb-repressive components to primed enhancers remain to be investigated, although we had not observed any changes in H3K27me3 status at distal enhancers of *Pax6*, *Sox1*, and *Nestin* when comparing WT and KO ESCs at the initiation of EB differentiation.

In *Tet1* KO EpiLCs, the altered chromatin landscape is accompanied by increased genomic occupancy by Oct4 at several low-affinity sites at developmental gene loci. These observations are consistent with a previous study that described Oct4 binding at a large set of low-accessible and differentiation-related genomic regions, potentially priming cell fate enhancers (Simandi et al. 2016). Because Tet1 occupancy sites and Tet1-dependent chromatin accessibility in EpiLCs are largely disparate from Oct4-bound regions, functional interactions between Tet1 and Oct4 at the genome-scale are likely to be indirect. However, we observed an exceptional coincidence of an EpiLC-specific Tet1-bound site that showed enhanced Oct4 binding in the absence of Tet1 at the proximal primed enhancer of *Oct4* (Fig. 4G), suggesting that Tet1 may evict Oct4 from an autoregulatory element at the *Oct4* promoter. Using a *POU5f1*-luciferase reporter construct transfected in KO EpiLCs, we did not observe any competitive effect of rescued *Tet1* expression on *OCT4* promoter-reporter activity (data not shown); however, we cannot preclude direct functional antagonism between Tet1 and Oct4 in the native chromatin environment of the *Oct4* autoregulatory site or posttranslational regulation of Oct4 by Tet1. Nodal and Wnt targets with enhanced Oct4 binding also include genes encoding negative feedback regulators (*Lefty1*, *Lefty2*, *Skil*, and *Axin2*); however, these effects are likely downstream from collaborative Oct4 and extracellular signals. In naive ESCs in vitro and the epiblast in vivo, loss of TET proteins causes loss of expression of neighboring *Lefty1* and *Lefty2* gene loci in association with hypermethylation of their promoters and enhancers (Koh et al. 2011; Dai et al. 2016; Khoueiry et al. 2017). Collectively, these studies suggest that a fundamental role of TET catalytic activity at pregastrulation stages is to keep epiblast cells in "neural default" mode by promoting accessibility of TFs to neural loci, modulating Oct4 expression and genomic occupancy, and sustaining expression of Nodal and Wnt antagonists to prevent precocious signaling activation of primitive streak entry.

Stochastic differences in expression levels of pluripotency factors may influence germ layer fate: Higher ratios of Oct4 to Sox2 suppresses NE while promoting ME differentiation and vice versa (Thomson et al. 2011). We observed that elevated *Oct4* expression in *Tet1* KO EBs increased ME gene expression, but did not directly inhibit NE potential, in line with a positive role of Oct4 in integrating ESC response to external cues (Simandi et al.

2016). Moreover, we observed loss of *Sox2* in both WT and KO EBs by similar extents and kinetics during EB differentiation (data not shown); thus, dysregulation of *Sox2* expression is unlikely to account for loss of NE potential in *Tet1* KO EBs. However, inhibition of Nodal signals rescued differentiation skewing toward ME and restored NE gene expression. The latter results are consistent with previous observations in human ESCs, where Wnt/β-catenin induction of neural crest markers are sustained only if Nodal signaling is inhibited (Funa et al. 2015). Thus, the impaired NE potential of *Tet1* KO cells is primarily a response to aberrant extrinsic cues that repress TF gene expression.

NE genes most affected in *Tet1* KO EBs are anterior neuronal progenitor markers (*Lhx5*, *Sim2*, and *Wnt1*), which are coincidently targets of repression by the posteriorization activity of Wnt/β-catenin signaling during in vitro generation of neuromesodermal progenitors (Gouti et al. 2014). Consistent with evidence of hyperactive β-catenin in *Tet1* KO EBs, inhibiting Wnt restores expression of anterior neuronal genes in KO cells during neuronal differentiation. It remains to be investigated whether modulating Wnt/β-catenin and graded Nodal signals may produce differential effects in WT and *Tet1* KO EBs on the anterior versus posterior identities of differentiating neuronal progenitors. Since loss-in-KO accessibility regions are enriched in the GO term for paraxial mesoderm development (Supplemental Fig. S2E), Tet1 may potentially also regulate the bipotency of Sox2$^+$/T$^+$ neuro-mesoderm progenitors (Gouti et al. 2014; Li et al. 2016; Koch et al. 2017). Nonetheless, the dysregulation of anterior markers including *Lhx5*, *Sim2*, and *Wnt1* by loss of *Tet1* is consistent with failed anterior neuropore closure observed in *Tet1*-null embryos (Khoueiry et al. 2017) and in *tuft* mutant mice in which a nonsense mutation in the *Tet1* gene has been detected (Fong et al. 2016), suggesting that Tet1 primarily regulates anterior neural identity.

Our results position Tet1 as an upstream negative regulator of Wnt/β-catenin signaling. Possible modes of action include gene activation of Wnt repressors by DNA demethylation or noncatalytic repression of signal effectors. As precedents for the former mechanism, promoter hypomethylation and derepression of Wnt pathway gene inhibitors by TET enzymes have been previously reported—*Sfrp4* by Tet3 during serum-free neural EB differentiation (Li et al. 2016); *DKK3* and *DKK4* by TET1 in human colon cancer cells to suppress tumor growth (Neri et al. 2015). Here, we observed that the axis of regulatory control involves modulation of Oct4 by Tet1 in the primed epiblast, since several gene targets of elevated Oct4 binding in *Tet1* KO EpiLCs encode canonical Wnt components such as *Fzd1*, *Fzd7*, and *Fzd8*. Moreover, de novo occupancy sites of Tet1 in EpiLCs are enriched in motifs for candidate TFs that are potentially negative regulators of Wnt/β-catenin, such as Wt-1 and Zic2/3 (Kim et al. 2009; Pourebrahim et al. 2011). Future investigations will identify Tet1 partners with potential pioneering activities in orchestrating the primed state chromatin, which in turn controls the tripartite axes of differentiation cues involving Oct4, Nodal and Wnt.

Materials and methods

Derivation of ESC lines

Mouse ESC lines were derived from blastocysts as previously described (Czechanski et al. 2014). Briefly, 4- to 5-wk-old females are injected intraperitoneally (i.p.) with 5 IU of pregnant mare serum gonadotropin (Calbiochem) and again 46–48 h later with 7.5 IU of human chorionic gonadotopin (Sigma) and paired with male studs. The morning after, the females were checked for the presence of copulation plugs (E0.5). E3.5 blastocysts were collected by flushing from the uterine horn and adapted in defined medium containing 2iL for 7–9 d. Following attachment to feeder cell layer and expansion, blastocysts outgrowth was dissociated with 0.05% Trypin-EDTA and passaged at least five to seven times in standard ESC culture medium (described below) supplemented with 1 µM MEK inhibitor PD0325901 (Axon Medchem BV Axon 1386), 3 µM GSK3 β inhibitor CHIR99021 (Axon Medchem BV Axon 1408), and 1000 units/mL ESGRO LIF (Millipore ESG1107) to obtain stable ESC lines.

B6.129P2-$Tet1^{Gt(RRG140)}$ ESC lines were derived as described previously (Khoueiry et al. 2017). To obtain C57BL/6J (abbreviated as B6)×129S6 F1 hybrid ESCs, we first backcrossed B6-$Tet1^{tm1.1Koh}$ to 129S6 (129/SvEvTac) for three generations using marker-assisted accelerated backcrossing (Charles River) to obtain heterozygous N3 males that are 91% 129S6 based on 192 informative SNPs on the 384 mouse SNP panel (Charles River). Heterozygous intercrosses between 129S6;B6 males and B6 superovulated females were performed to obtain the hybrid strain ESC lines. Genotyping of the $Tet1^{tm1Koh}$ allele was performed by PCR amplification of genomic DNA using primers: Tet1_1F (5′-TTAG ACCCCAAACTCAGGTGAC-3′), Tet1_2R (5′-TTTTCCGGGG TTCACTGCCTT-3′), and LacZ_6R (5′-CGGATTGACCGTAA TGGGATAG-3′) to distinguish WT and mutant alleles (Khoueiry et al. 2017). Gender determination of all ESC lines were performed by PCR using the sex-specific primers SX_F (5′-GATG ATTTGAGTGGAAATGTGAGGTA-3′) and SX_R (5′-CTTATG TTTATAGGCATGCACCATGTA-3′) (McFarlane et al. 2013). All comparisons between WT and KO groups were between male lines in this study.

Cell culture

ESC lines were cultured on feeder layers of mitotically inactivated mouse embryonic fibroblasts (MEFs) in standard ESC culture medium composed of knockout DMEM (Thermo Fisher 10829-018), 15% ESC-qualified FBS (Invitrogen 10270-106), 2 mM L-glutamine (Invitrogen 25030-024), 0.1 mM each nonessential amino acids (Invitrogen 11440-035), 0.1 mM β-mercaptoethanol (Invitrogen 31350-010), 100:100 units:µg/mL penicillin:streptomycin (Invitrogen 15140), supplemented with in-house leukemia inhibitory factor (LIF) culture supernatant. MEFs were cultured in the MEF medium consisting of DMEM GlutaMAX (61965-026), 10% FBS (Sigma-Aldrich F7524), 2 mM L-glutamine, 1 mM sodium pyruvate, 0.1 mM each of nonessential amino acids, 0.1 mM β-mercaptoethanol and 100:100 U:µg/mL penicillin:streptomycin. When harvesting feeder-cultured ESCs, cells were replated twice for 30 min on tissue culture plates to remove adherent feeder cells.

EpiLC differentiation

Differentiation toward EpiLCs was performed as described previously but without activin A supplementation (Buecker et al. 2014). Briefly, feeder-depleted ESCs were first adapted at least five passages in defined media referred as 2iL media, which is composed of N2B27 basal medium supplemented with 1 µM MEK inhibitor PD0325901 (Axon Medchem BV Axon 1386), 3 µM GSK3 β inhibitor CHIR99021 (Axon Medchem BV Axon 1408), and 1000 units/mL ESGRO LIF (Millipore ESG1107). The N2B27 basal medium was prepared as a 1:1 mixture of DMEM-F12 (Invitrogen 11320-074) and Neurobasal medium (Invitrogen 21103-049) supplemented with 1× N2 (Invitrogen 17502-048), 1× B27 (Invitrogen 17504-044), 1 mM L-glutamine, 0.1 mM 2-mercaptoethanol, and 100:100 units:µg/mL penicillin:streptomycin. 2iL-adapted ESCs were seeded at 2×10^5 to 3×10^5 cells per 10 cm^2 on dishes coated with 5 µg/10 cm^2 fibronectin and differentiated for 2 d in N2B27 basal medium supplemented with 1% knockout serum replacement KSR (Invitrogen 10828-028) and 12 ng/mL bFGF (Peprotech 100-18C). To knock down Zic2 expression using siRNA during EpiLC conversion, siGENOME SMARTpool targeting Zic2 (Dharmacon) was supplemented twice: once in 2iL cultured ESCs 2 d before the conversion and once in EpiLCs 12 h after the conversion.

Serum-free embryoid body (EB) differentiation

Serum-free EB differentiation was performed as previously described (Gadue et al. 2006). Briefly, feeder-depleted ESCs were adapted for at least two passages on feeder-free gelatin-coated plates in defined medium comprised of knockout DMEM/F12 (Invitrogen 12660-012), 0.5× N2 (Invitrogen 17502-048), 0.5× B27 (Invitrogen 17504-044), 2 mM L-glutamine, 0.05% BSA (Invitrogen 15260-370) with fresh supplementation of 1000 units/mL ESGRO LIF (Millipore ESG1107), 0.15 µM 1-thioglycerol (Sigma-Aldrich M6145), and 10 ng/mL BMP4 (Peprotech 120-05ET). At day 0 of EB differentiation, adapted ESCs were dissociated using TrypLE Express (Invitrogen 12605010) and then plated at the density of 1×10^5 cells/mL onto nonadherent bacteriological Petri six-well plates (Greiner 657102) with EB differentiation medium composed of 75% IMDM (Invitrogen 21980-032), 25% Ham's F12 medium (Invitrogen 21765-029), 0.5× N2, 0.5× B27 without retinoic acid (Invitrogen 12587-010), 0.05% BSA, supplemented freshly with 0.45 µM 1-thioglycerol, and 50 µg/mL L-ascorbic acid (Sigma-Aldrich A4544). At day 2, EBs were dissociated with TyrpLE Express and reaggregated with or without addition of 25 ng/mL Activin A (PeproTech 120-14E); the latter served as a positive control for induction of definitive endoderm markers. To knockdown Oct4 expression using siRNA during EB differentiation, siGENOME SMARTpool targeting Oct4 (Dharmacon) was supplemented at day 0 and day 2 to dissociated cells according to manufacturer's instructions. To inhibit the signaling pathways, the appropriate chemical inhibitors were added at the final concentration of 2.5 µM (in 0.05% DMSO vehicle) at day 2. Inhibitors used in this study were TGF-β receptor/Nodal signaling inhibitor SB431542 (Sigma-Aldrich S4317), and Wnt inhibitors IWP2 (Sigma-Aldrich X10536) and XAV939 (Sigma-Aldrich X3004).

Terminal neuronal differentiation

Neural differentiation was performed as described previously (Bibel et al. 2007). Feeder-depleted ESCs were adapted in gelatin-coated dishes in standard ESC culture medium for two passages. At day 0, 4×10^6 cells were plated onto nonadherent bacteriological 10-cm Petri dishes (Greiner 633102) in 15 mL of aggregation medium, which is the same as MEF medium but using 10% ES-qualified FBS. Media were refreshed every other day of EB aggregation. At day 4, the media were supplemented with 5 µM retinoic acid (Sigma-Aldrich R2625). At day 8, EBs were fully dissociated into single cells with 0.5% trypsin-EDTA and seeded at 2×10^6 to 3×10^6 per 12-well culture dish containing coverglass precoated with 50 µg/mL poly L-ornithine (Sigma-

Aldrich P3655) and 5 ng/mL laminin (Sigma-Aldrich L2020) in DMEM/F-12 (Invitrogen 31330-038), supplemented with 1× N2, 25 µg/mL insulin (Roche 1376497), 50 µg/mL BSA, and 2 mM L-glutamine. At day 10, media were replaced with DMEM/F-12 (Invitrogen 31330-038) supplemented with 1× B27 and 2 mM L-glutamine (Invitrogen 17504-044). Neural cells were fixed at day 14 for immunofluorescent staining.

Anterior neural progenitor cell differentiation

Anterior neural progenitor differentiation was performed as described previously (Arnold et al. 2008; Gouti et al. 2014; Cruz-Molina et al. 2017; Koch et al. 2017). Feeder-depleted serum-cultured ESCs were plated on gelatin-coated plates at 10,000 cells/cm^2 (90,000 cells per well in a six-well plate) in N2B27-defined medium compromising a 1:1 mix of 1:1 DMEM/F12 (Invitrogen 12634-010) and neurobasal medium (10888-022) supplemented with 0.5× B27 without vitamin A (12587-010), 0.5× N2 (17502-048), 2 mM L-glutamine, 40 mg/mL BSA fraction V, 0.1 mM 2-mercaptoethanol, and 100:100 U:µg/mL penicillin:streptomycin. Cells were supplemented with 10 ng/mL bFGF (Peprotech 100-18C) from day 0 until day 3. At day 3, the medium was changed to medium without bFGF. Next, 5 µM Wnt inhibitor XAV939 (Sigma-Aldrich X3004), 2.5 µM TGF-β receptor inhibitor SB431542 (Sigma-Aldrich S4317), or DMSO (vehicle control) was added from day 2 until day 5. The medium was changed every day during differentiation. The cells were considered as anterior neural progenitor cells at day 5.

Timed pregnancies

To obtain *Tet1* KO embryos, B6-*Tet1*[tm1Koh] mice were naturally mated by heterozygote intercrossing. The day on which a copulation plug was detected in the females was timed as E0.5 of gestation. E10.5–E13.5 embryos were dissected and genotyped as previously described (Khoueiry et al. 2017). Images were taken using a Leica S8APO microscope equipped with a Leica MC170 HD camera. All experimental procedures on mice have been reviewed and approved by the Katholieke Universiteit Leuven Ethical Committee for Animal Experimentation (P101/2016) in compliance with the European Directive 2010/63/EU.

Immunostaining

Cells were fixed on coverglass in 4% paraformaldehyde for 10 min at room temperature and permeabilized with 0.5% Triton X-100 in PBS for 5–10 min. Fixed cells were then incubated with primary antibodies for anti-TUJ-1/β3TUB antibody (1:500; Synaptic System 302302) overnight at 4°C in PBS containing 5% BSA and 0.02% Tween20. Following incubation with the appropriate fluorophore-labeled secondary antibodies in blocking solution for 30 min in the dark and counterstaining with DAPI (Sigma-Aldrich D9542) with intervening washes with PBST (PBS + 0.1% TritonX-100), slides were mounted in ProLong Gold Antifade reagent (Thermo Fisher ab P36930) and imaged using a Zeiss Axio Imager Z1 fluorescent microscope.

Embryoid body embedding, section, and immunostaining

Embryoid bodies were collected and fixed in 4% PFA for 1 h at room temperature and washed three times with DPBS (PBS without Ca^{2+} and Mg^{2+}) (Thermo Fisher 14190136) for 15 min at room temperature thereafter. Embryoid bodies were then resuspended in 70% ethanol and stored at 4°C. For paraffin embedding, the embryoid bodies were sequentially washed three times with 100%

ethanol for 30 min at room temperature, three times with 1-buthanol for 30 min at room temperature, and three times with paraffin for 30 min at 60°C before being manually embedded in paraffin blocks. Sections 5 µm thick were prepared with a microtome and subjected to immunostaining. In brief, sections were deparaffinized by incubation for 1 h at 57°C and dehydrated in descending ethanol gradient twice for 3 min each. Sections were subsequently rehydrated in DPBS and antigen was unmasked by 30-min incubation in 10 mM Na-citrate buffer (PH 6) at 95°C. Sections were allowed to cool down for 20 min at room temperature and blocked in blocking buffer (DPBS + 1% BSA, 0.15% glycine, 0.1% Triton, 10% donkey serum) for 1 h at room temperature. Sections were finally incubated overnight at 4°C with anti-Brachyury antibody (1:100; Santa Cruz Biotechnology sc17743) diluted in blocking buffer. The following day, the sections were washed three times for 15 min in PBST (PBS + 0.1% Triton X-100) and incubated with the appropriate conjugated secondary antibody (donkey antigoat IgG [H + L] cross-adsorbed secondary antibody, Alexa fluor 555; Thermo Fisher A-21432) in dark for 30 min at room temperature. A coverslip was finally mounted with VectaShield Antifade mounting medium with DAPI (Vector Laboratories H-1200) and the stain was visualized by Zeiss AxioImager Z1 fluorescent microscope.

Western blots

Cells were lysed in RIPA buffer (50 mM Tris at pH 8.0, 150 mM NaCl, 0.2 mM EDTA, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate) containing 1× protease inhibitor cocktail (Roche 11836153001) for 30 min on ice. Protein extracts were run on a 8% or 10% SDS–polyacrylamide gel and then transferred to a PVDF membrane. Membranes were blocked with 5% nonfat milk or 5% BSA for 1 h and then incubated with primary antibodies overnight at 4°C, followed by incubation with secondary antibodies for 1 h at room temperature. The signal was detected using Clarity Western ECL substrate (Bio-Rad 1705060) on AGFA Curix 60 Film Processor. The quantification of signals in Western blot analysis were performed using ImageJ software. Primary antibodies used in this study were anti-Oct4 (1:1000; Santa Cruz Biotechnology sc-5279), anti-phospho-Smad2 (Ser465/467)/Smad3 (Ser423/425) (D27F4) (1:1000; Cell Signaling Technology 8828), anti-Smad2/3 (1:1000; Cell Signaling Technology 3102), antinonphospho (active) β-Catenin (Ser33/37/Thr41) (D13A1) (1:1000; Cell Signaling Technology 8814), anti-β-Catenin (Biosciences 610153), anti-Tet1 (Millipore 09-872), anti-β-Actin (D6A8) (rabbit polyclonal, 1:2000; Cell Signaling Technology 8457), and anti-β-Actin (mouse monoclonal, 1:2000; Sigma-Aldrich A1978).

RNA extraction and reverse transcription-quantitative PCR (RT-qPCR)

Total RNA was extracted using the RNeasy minikit (Qiagen 74004) according to the manufacturer's instructions, and then converted to cDNA using SuperScript III first-strand synthesis supermix kit (Thermo Fisher 11752-050). Quantitative real-time PCR (qPCR) was performed using cDNA at 1:100 in SYBR Green PCR master mix (Thermo Fisher 11733-046) on either a StepOnePlus or ViiA7 real-time PCR system (Applied Biosystems). Relative expression was normalized to *Gapdh* expression using ΔΔCt method. Primer sequences are listed in Supplemental Table S6.

ChIP-qPCR and ChIP-seq

We chemically cross-linked 5 × 10^6 cells per ChIP-qPCR or 2 × 10^7 to 3 × 10^7 cells per ChIP-seq sample with 1% methanol-free

formaldehyde (Polysciences 04018) for 10 min at room temperature and then quenched with 0.125 M glycine. Fixed cells were lysed sequentially in lysis buffer I (50 mM HEPES-KOH at pH 7.5, 140 mM NaCl, 0.25% Triton X-100, 0.5% NP-40, 10% glycerol, 1 mM EDTA) and lysis buffer II (10 mM Tris-HCl at pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA), followed by lysis buffer III (10 mM Tris-HCl at pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% sodium deoxycholate, 0.5% N-lauroylsarcosine).

Chromatin fractions were sheared to 200–500 bp using Bioruptor Plus sonicator under a high power setting for 20 cycles (30 sec on, 30 sec off). The sheared DNA was measured on an Agilent 2100 Bioanalyzer using High Sensitivity DNA analysis kits (Agilent 5067-4626). The high-size fragments (>500 bp), if detected, was then removed with size-selection using Agencourt AMPure XP beads (Beckman Coulter A63881). The sheared DNA then was incubated using appropriate antibodies (and the same amount of control IgG for ChIP-qPCR) overnight and then precipitated with Protein G Dynabeads (Thermo Scientific 10004D). Precipitates were washed sequentially using the following washes for 5 min each: low-salt buffer (20 mM Tris-HCl at pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), high-salt buffer (20 mM Tris-HCl at pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), 1× LiCl buffer (10 mM Tris-HCl at pH 8.0, 250 mM LiCl, 1 mM EDTA, 1% deoxychlate, 1% NP-40), and twice in 1× TE buffer + 50 mM NaCl. Chromatin antibody beads were eluted in 50 mM Tris-HCl (pH 8.0), 10 mM EDTA, and 1% SDS and decross-linked in 5 M NaCl solution overnight at 65°C. Chromatin extracts were incubated with DNase-free RNase (Roche 04716728001) for 30 min at 37°C and afterward with 10 mg/mL proteinase K (Roche 03115879001) for 2 h at 55°C. Sheared 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.

ChIP-qPCR primer sequences are listed in Supplemental Table S6. ChIP-seq libraries were prepared using the NEBNext Ultra DNA library preparation kit for Illumina according to manufacturer's instructions and then sequenced on HiSeq 2500 to generate 15–25 million 50-bp single-end reads. TET1 ChIP–seq assays of 2iL ESCs and EpiLCs were performed in biological duplicates using two WT and two $Tet1^{GT/GT}$ (KO) C57BL/6 incipient congenic cell lines. Oct4 ChIP-seq in WT and $Tet1$ KO EpiLCs was performed using two biological replicate cell lines of the $Tet1^{tm1Koh}$ B6129S6F1 strain and repeated using one pair of WT and $Tet1^{GT/GT}$ (KO) B6 incipient congenic cell lines. Antibodies used in this study were anti-Tet1 (Millipore 09-872) for Tet1 ChIP-seq, anti-Tet1 (Diagenode GTX125888) for Tet1 ChIP-qPCR, anti-Oct4 (Diagenode C15410305) for Oct4 ChIP-seq in WT and KO $Tet1^{tm1Koh}$ B6129S6F1 EpiLCs, and anti-Oct3/4 (Santa Cruz Biotechnology sc-8286) for Oct4 ChIP-seq in WT and $Tet1^{GT/GT}$ C57BL/6 EpiLCs.

ChIP-seq analysis

Demultiplexed fastq files were aligned to UCSC mm10 genome reference using bowtie2 (v2.26) with default arguments. The generated BAM files were then analyzed using MACS2 (v2.1.0) with the following settings and the corresponding controls to identify significant binding peaks enriched in the target proteins: For Tet1 ChIP-seq in WT and $Tet1^{GT/GT}$ 2iL-cultured ESCs, the MACS2 settings were "–nomodel –extsize 200 -q 0.01," in which replicates in WT 2iL-cultured ESCs were set as the treatment group and the replicates in $Tet1^{GT/GT}$ 2iL-cultured ESCs were set as the background group for normalization. For Tet1 ChIP-seq in WT and $Tet1^{GT/GT}$ EpiLCs (E-MTAB-5562) (Khoueiry et al.

2017), the MACS2 settings were "–nomodel –extsize 200 -q 0.01," in which replicates in WT EpiLCs were set as the treatment group and the replicates in $Tet1^{GT/GT}$ EpiLCs were set as the background group for normalization. For H3K27ac, H3K4me1, and p300 ChIP-seq in EpiLCs (GSE56098) (Buecker et al. 2014), the MACS2 settings were "–broad –broad-cutoff 0.1 –nomodel –extsize 146." For Oct4 ChIP-seq in 2iL ESCs and EpiLCs (GSE56098) (Buecker et al. 2014), the MACS2 settings were "—nomodel –extsize 200 -q 0.01." For Oct4 ChIP-seq in WT and KO B6129S6F1-$Tet1^{tm1Koh}$ EpiLCs, the MACS2 settings were "-q 0.05," in which the replicates of Oct4 ChIP samples in WT (or KO) EpiLCs were set as the treatment group and the replicates of WT (or KO) total genomic input DNA were set as the background group for normalization. For Oct4 ChIP-seq in WT and $Tet1^{GT/GT}$ EpiLCs, the MACS2 settings were "-q 0.05," in which the Oct4 ChIP samples in WT (or $Tet1^{GT/GT}$) EpiLCs were set as the treatment group and WT (or $Tet1^{GT/GT}$) total genomic input DNA were set as the background group for normalization. For Zic2 ChIP-seq in EpiLCs (GSE74636) (Matsuda et al. 2017), the MACS2 settings were "-q 0.05."

To define the state-specific Tet1-bound sites, the MACS2 subcommand bdgdiff was used with default arguments to statistically compare the MACS2-generated Tet1 peak files in 2iL ESCs with those in EpiLCs. To define differential Oct4-bound peaks between WT and KO B6129S6F1-$Tet1^{tm1Koh}$ EpiLCs, two methods were used: (1) the MACS2 subcommand bdgdiff with default arguments the same as above and (2) R/Bioconductor package DiffBind (v2.10.0) were used with FDR < 0.1 and other default arguments.

To annotate genomic feature distribution of these peak sets, R/Bioconductor package ChIPseeker (v1.18.0) was used. To annotate the ChIP-seq peaks to their nearby specific genes (at proximal religions or enhancers) followed by functional interpretation analysis (gene ontology of biological process, gene ontology of molecular function, and mouse genome informatics phenotype ontology), the online bioinformatic tool GREAT (genomic regions enrichment of annotations tool; v3.0.0 for Tet1 ChIP-seq and v4.0.4 for Oct4 ChIP-seq) (McLean et al. 2010) with the default "basal plus extension" settings was used. To generate density heat maps and density profiling plots, the BAM files were first summarized and transformed into BigWig files with bamCompare command from deepTools2 (v3.3.1) using the following settings: –scaleFactorsMethod None –normalizeUsing RPKM –ignoreDuplicates ; plots were generated with plotHeatmap and plotProfile commands. The visualization of peak tracks was generated with Integrative Genomics Viewer (IGV) using BigWig files. To find overlapped peaks from different peak files, findOverlapsOfPeaks function from R/Bioconductor package ChIPpeakAnno (v3.16.1) was used with default arguments.

DNA motif enrichment analysis was performed using findMotifsGenome.pl from the HOMER package (v4.10) with the following arguments: findMotifsGenome.pl mm10 -size given. The enriched motifs were then filtered with the following criteria: FDR < 0.0001, foreground present in >10% of peaks, and foreground/background ratio >1.2 as described previously (Knaupp et al. 2017).

ATAC-seq

ATAC-seq was performed as described previously (Buenrostro et al. 2013; Corces et al. 2017). Fifty-thousand freshly isolated cells were washed once with 100 μL PBS and resuspended in ice-cold 50 μL lysis buffer (10 mM Tris-HCl at pH 7.4, 10 mM NaCl, 3 mM MgCl$_2$, 0.1% NP40, 0.1% Tween-20, 0.01% digitonin) for 3 min. The suspension of nuclei was then centrifuged at

500g for 10 min at 4°C, followed by one wash with lysis buffer without Tween-20 and digitonin. The visible pellets were incubated with 50 μL of transposition reaction mix (25 μL of 2× TD buffer, 2.5 μL [or 5 μL] of Tn5 Transposase, 16.5 μL of PBS, 0.5 μL of 1% digitonin, 0.5 μL of 10% Tween-20, nuclease-free H$_2$O) (Illumina FC-121-1030) at 500g for 30 min at 37°C with mixing. DNA was isolated using Zymo DNA Clean & Concentrator-5 kit (Zymo D4014). Library amplification was done by two sequential PCR reactions (five and eight cycles, respectively) using NEBNext high-fidelity 2× PCR master mix (New England Biolabs M0541S). Library quality was measured on an Agilent 2100 Bioanalyzer using high-sensitivity DNA analysis kits (Agilent 5067-4626). The fragmented DNA was then size-selected using Agencourt AMPure XP beads (Beckman Coulter A63881) to generate 200- to 500-bp DNA fragments, followed by sequencing on an Illumina HiSeq 4000 to generate 40 million to 50 million 50-bp single-end reads per library. Samples for ATAC-seq were generated from two independent ESC-to-EpiLC conversions using a pair of wild-type (WT17) and $Tet1^{GT/GT}$ (KO12) ESC lines with addition of either 2.5 or 5 μL of Tn5 transposase.

ATAC-seq analysis

Demultiplexed fastq files were analyzed using automated EN-CODE ATAC-seq pipeline (http://doi.org/10.5281/zenodo.156534) developed by Anshul Kundaje's laboratory with default arguments.

To call the statistically differential accessible regions between WT and $Tet1^{GT/GT}$ EpiLCs, two resulting ATAC peak files and BAM files from each replicate per genotype were then loaded into R/Bioconductor package DiffBind (v2.10.0) with default arguments.

The genomic feature annotation, gene ontology analysis and motif enrichment analysis were performed the same as ChIP-seq analysis. To generate density heat maps and density profiling plots, the BAM files were first summarized and transformed into BigWig files with bamCoverage command from deepTools2 using the following settings: –normalizeUsing RPKM –ignoreDuplicates.

Data availability

The accession number for the data reported here is GEO GSE144869, composed of SubSeries GSE144867 (ATAC-seq) and GSE144868 (ChIP-seq). All source data and code supporting the findings of this study are available from the corresponding author on reasonable request.

References

Adachi K, Kopp W, Wu G, Heising S, Greber B, Stehling M, Araúzo-Bravo MJ, Boerno ST, Timmermann B, Vingron M, et al. 2018. Esrrb unlocks silenced enhancers for reprogramming to naive pluripotency. *Cell Stem Cell* 23: 266–275.e6. doi:10.1016/j.stem.2018.05.020

Anastas JN, Moon RT. 2013. WNT signalling pathways as therapeutic targets in cancer. *Nat Rev Cancer* 13: 11–26. doi:10.1038/nrc3419

Arnold SJ, Hofmann UK, Bikoff EK, Robertson EJ. 2008. Pivotal roles for eomesodermin during axis formation, epithelium-to-mesenchyme transition and endoderm specification in the mouse. *Development* 135: 501–511. doi:10.1242/dev.014357

Bartoccetti M, van der Veer BK, Luo X, Khoueiry R, She P, Bajaj M, Xu J, Janiszewski A, Thienpont B, Pasque V, et al. 2020. Regulatory dynamics of Tet1 and Oct4 resolve stages of global DNA demethylation and transcriptomic changes in reprogramming. *Cell Rep* 30: 2150–2169.e9. doi:10.1016/j.celrep.2020.01.065

Bibel M, Richter J, Lacroix E, Barde YA. 2007. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. *Nat Protoc* 2: 1034–1043. doi:10.1038/nprot.2007.147

Buecker C, Srinivasan R, Wu Z, Calo E, Acampora D, Faial T, Simeone A, Tan M, Swigut T, Wysocka J. 2014. Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* 14: 838–853. doi:10.1016/j.stem.2014.04.003

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10: 1213–1218. doi:10.1038/nmeth.2688

Cimmino L, Dawlaty MM, Ndiaye-Lobry D, Yap YS, Bakogianni S, Yu Y, Bhattacharyya S, Shaknovich R, Geng H, Lobry C, et al. 2015. TET1 is a tumor suppressor of hematopoietic malignancy. *Nat Immunol* 16: 653–662. doi:10.1038/ni.3148

Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B, et al. 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* 14: 959–962. doi:10.1038/nmeth.4396

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, et al. 2010. Histone H3K27ac separates active from poised

enhancers and predicts developmental state. *PNAS* **107:** 21931–21936. doi:10.1073/pnas.1016071107

Cruz-Molina S, Respuela P, Tebartz C, Kolovos P, Nikolic M, Fueyo R, van Ijcken WFJ, Grosveld F, Frommolt P, Bazzi H, et al. 2017. PRC2 facilitates the regulatory topology required for poised enhancer function during pluripotent stem cell differentiation. *Cell Stem Cell* **20:** 689–705.e9. doi:10.1016/j.stem.2017.02.004

Czechanski A, Byers C, Greenstein I, Schrode N, Donahue LR, Hadjantonakis AK, Reinholdt LG. 2014. Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains. *Nat Protoc* **9:** 559–574. doi:10.1038/nprot.2014.030

Dai HQ, Wang BA, Yang L, Chen JJ, Zhu GC, Sun ML, Ge H, Wang R, Chapman DL, Tang F, et al. 2016. TET-mediated DNA demethylation controls gastrulation by regulating Lefty–Nodal signalling. *Nature* **538:** 528–532. doi:10.1038/nature20095

Fong KS, Hufnagel RB, Khadka VS, Corley MJ, Maunakea AK, Fogelgren B, Ahmed ZM, Lozanoff S. 2016. A mutation in the *tuft* mouse disrupts TET1 activity and alters the expression of genes that are crucial for neural tube closure. *Dis Model Mech* **9:** 585–596. doi:10.1242/dmm.024109

Frank CL, Liu F, Wijayatunge R, Song L, Biegler MT, Yang MG, Vockley CM, Safi A, Gersbach CA, Crawford GE, et al. 2015. Regulation of chromatin accessibility and Zic binding at enhancers in the developing cerebellum. *Nat Neurosci* **18:** 647–656. doi:10.1038/nn.3995

Frauer C, Rottach A, Meilinger D, Bultmann S, Fellinger K, Hasenöder S, Wang M, Qin W, Söding J, Spada F, et al. 2011. Different binding properties and function of CXXC zinc finger domains in Dnmt1 and Tet1. *PLoS One* **6:** e16627. doi:10.1371/journal.pone.0016627

Funa Nina S, Schachter Karen A, Lerdrup M, Ekberg J, Hess K, Dietrich N, Honoré C, Hansen K, Semb H. 2015. β-Catenin regulates primitive streak induction through collaborative interactions with SMAD2/SMAD3 and OCT4. *Cell Stem Cell* **16:** 639–652. doi:10.1016/j.stem.2015.03.008

Gadue P, Huber TL, Paddison PJ, Keller GM. 2006. Wnt and TGF-β signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. *PNAS* **103:** 16806–16811. doi:10.1073/pnas.0603916103

Gouti M, Tsakiridis A, Wymeersch FJ, Huang Y, Kleinjung J, Wilson V, Briscoe J. 2014. In vitro generation of neuromesodermal progenitors reveals distinct roles for wnt signalling in the specification of spinal cord and paraxial mesoderm identity. *PLoS Biol* **12:** e1001937. doi:10.1371/journal.pbio.1001937

Hayashi K, Ohta H, Kurimoto K, Aramaki S, Saitou M. 2011. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146:** 519–532. doi:10.1016/j.cell.2011.06.052

He Y-F, Li B-Z, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, et al. 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333:** 1303–1307. doi:10.1126/science.1210944

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38:** 576–589. doi:10.1016/j.molcel.2010.05.004

Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333:** 1300–1303. doi:10.1126/science.1210597

Jang H, Kim TW, Yoon S, Choi SY, Kang TW, Kim SY, Kwon YW, Cho EJ, Youn HD. 2012. O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network. *Cell Stem Cell* **11:** 62–74. doi:10.1016/j.stem.2012.03.001

Khoueiry R, Sohni A, Thienpont B, Luo X, Velde JV, Bartoccetti M, Boeckx B, Zwijsen A, Rao A, Lambrechts D, et al. 2017. Lineage-specific functions of TET1 in the postimplantation mouse embryo. *Nat Genet* **49:** 1061–1072. doi:10.1038/ng.3868

Kim MK-H, McGarry TJ, Broin P Ó, Flatow JM, Golden AA-J, Licht JD. 2009. An integrated genome screen identifies the Wnt signaling pathway as a major target of WT1. *PNAS* **106:** 11154–11159. doi:10.1073/pnas.0901591106

Knaupp AS, Buckberry S, Pflueger J, Lim SM, Ford E, Larcombe MR, Rossello FJ, de Mendoza A, Alaei S, Firas J, et al. 2017. Transient and permanent reconfiguration of chromatin and transcription factor occupancy drive reprogramming. *Cell Stem Cell* **21:** 834–845.e6. doi:10.1016/j.stem.2017.11.007

Koch F, Scholze M, Wittler L, Schifferl D, Sudheer S, Grote P, Timmermann B, Macura K, Herrmann BG. 2017. Antagonistic activities of Sox2 and Brachyury control the fate choice of neuro-mesodermal progenitors. *Dev Cell* **42:** 514–526.e7. doi:10.1016/j.devcel.2017.07.021

Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, Laiho A, Tahiliani M, Sommer CA, Mostoslavsky G, et al. 2011. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8:** 200–213. doi:10.1016/j.stem.2011.01.008

Li X, Yue X, Pastor WA, Lin L, Georges R, Chavez L, Evans SM, Rao A. 2016. Tet proteins influence the balance between neuroectodermal and mesodermal fate choice by inhibiting Wnt signaling. *PNAS* **113:** E8267–E8276. doi:10.1073/pnas.1617802113

Luo Z, Gao X, Lin C, Smith ER, Marshall SA, Swanson SK, Florens L, Washburn MP, Shilatifard A. 2015. Zic2 is an enhancer-binding factor required for embryonic stem cell specification. *Mol Cell* **57:** 685–694. doi:10.1016/j.molcel.2015.01.007

Matsuda K, Mikami T, Oki S, Iida H, Andrabi M, Boss JM, Yamaguchi K, Shigenobu S, Kondoh H. 2017. ChIP-seq analysis of genomic binding regions of five major transcription factors highlights a central role for ZIC2 in the mouse epiblast stem cell gene regulatory network. *Development* **144:** 1948–1958. doi:10.1242/dev.143479

McFarlane L, Truong V, Palmer JS, Wilhelm D. 2013. Novel PCR assay for determining the genetic sex of mice. *Sex Dev* **7:** 207–211. doi:10.1159/000348677

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. 2010. GREAT improves functional interpretation of *cis*-regulatory regions. *Nat Biotechnol* **28:** 495–501. doi:10.1038/nbt.1630

Mullen AC, Orlando DA, Newman JJ, Lovén J, Kumar RM, Bilodeau S, Reddy J, Guenther MG, DeKoter RP, Young RA. 2011. Master transcription factors determine cell-type-specific responses to TGF-β signaling. *Cell* **147:** 565–576. doi:10.1016/j.cell.2011.08.050

Muñoz-Descalzo S, Hadjantonakis A-K, Arias AM. 2015. Wnt/β-catenin signalling and the dynamics of fate decisions in early mouse embryos and embryonic stem (ES) cells. *Semin Cell Dev Biol* **47-48:** 101–109. doi:10.1016/j.semcdb.2015.08.011

Neri F, Dettori D, Incarnato D, Krepelova A, Rapelli S, Maldotti M, Parlato C, Paliogiannis P, Oliviero S. 2015. TET1 is a tumour suppressor that inhibits colon cancer growth by

derepressing inhibitors of the WNT pathway. *Oncogene* **34:** 4168–4176. doi:10.1038/onc.2014.356

Niwa H. 2010. Mouse ES cell culture system as a model of development. *Dev Growth Differ* **52:** 275–283. doi:10.1111/j.1440-169X.2009.01166.x

Niwa H, Miyazaki J, Smith AG. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24:** 372–376. doi:10.1038/74199

Ozair MZ, Kintner C, Brivanlou AH. 2013. Neural induction and early patterning in vertebrates. *Wiley Interdiscip Rev Dev Biol* **2:** 479–498. doi:10.1002/wdev.90

Pastor WA, Aravind L, Rao A. 2013. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol* **14:** 341–356. doi:10.1038/nrm3589

Pourebrahim R, Houtmeyers R, Ghogomu S, Janssens S, Thelie A, Tran HT, Langenberg T, Vleminckx K, Bellefroid E, Cassiman J-J, et al. 2011. Transcription factor Zic2 inhibits Wnt/β-catenin protein signaling. *J Biol Chem* **286:** 37732–37740. doi:10.1074/jbc.M111.242826

Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. 2011. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470:** 279–283. doi:10.1038/nature09692

Radzisheuskaya A, Le Bin Chia G, dos Santos RL, Theunissen TW, Castro LFC, Nichols J, Silva JCR. 2013. A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all embryonic lineages. *Nat Cell Biol* **15:** 579–590. doi:10.1038/ncb2742

Robertson EJ. 2014. Dose-dependent Nodal/Smad signals pattern the early mouse embryo. *Semin Cell Dev Biol* **32:** 73–79. doi:10.1016/j.semcdb.2014.03.028

Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, et al. 2012. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481:** 389–393. doi:10.1038/nature10730

Santos F, Hendrich B, Reik W, Dean W. 2002. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* **241:** 172–182. doi:10.1006/dbio.2001.0501

Schier AF. 2009. Nodal morphogens. *Cold Spring Harb Perspect Biol* **1:** a003459. doi:10.1101/cshperspect.a003459

Senft AD, Costello I, King HW, Mould AW, Bikoff EK, Robertson EJ. 2018. Combinatorial Smad2/3 activities downstream of nodal signaling maintain embryonic/extra-embryonic cell identities during lineage priming. *Cell Rep* **24:** 1977–1985.e7. doi:10.1016/j.celrep.2018.07.077

Shi Y, Massagué J. 2003. Mechanisms of TGF-β signaling from cell membrane to the nucleus. *Cell* **113:** 685–700. doi:10.1016/S0092-8674(03)00432-X

Simandi Z, Horvath A, Wright Lyndsey C, Cuaranta-Monroy I, De Luca I, Karolyi K, Sauer S, Deleuze J-F, Gudas Lorraine J, Cowley Shaun M, et al. 2016. OCT4 acts as an integrator of pluripotency and signal-induced differentiation. *Mol Cell* **63:** 647–661. doi:10.1016/j.molcel.2016.06.039

Smith A. 2017. Formative pluripotency: the executive phase in a developmental continuum. *Development* **144:** 365–373. doi:10.1242/dev.142679

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, et al. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324:** 930–935. doi:10.1126/science.1170116

Tam PP, Loebel DA. 2007. Gene function in mouse embryogenesis: get set for gastrulation. *Nat Rev Genet* **8:** 368–381. doi:10.1038/nrg2084

ten Berge D, Koole W, Fuerer C, Fish M, Eroglu E, Nusse R. 2008. Wnt signaling mediates self-organization and axis formation in embryoid bodies. *Cell Stem Cell* **3:** 508–518. doi:10.1016/j.stem.2008.09.013

Thomson M, Liu Siyuan J, Zou L-N, Smith Z, Meissner A, Ramanathan S. 2011. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* **145:** 875–889. doi:10.1016/j.cell.2011.05.017

Verma N, Pan H, Doré LC, Shukla A, Li QV, Pelham-Webb B, Teijeiro V, González F, Krivtsov A, Chang C-J, et al. 2018. TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. *Nat Genet* **50:** 83–95. doi:10.1038/s41588-017-0002-y

Vincent SD, Dunn NR, Hayashi S, Norris DP, Robertson EJ. 2003. Cell fate decisions within the mouse organizer are governed by graded Nodal signals. *Genes Dev* **17:** 1646–1662. doi:10.1101/gad.1100503

Wang Q, Zou Y, Nowotschin S, Kim SY, Li QV, Soh C-L, Su J, Zhang C, Shu W, Xi Q, et al. 2017. The p53 family coordinates Wnt and Nodal inputs in mesendodermal differentiation of embryonic stem cells. *Cell Stem Cell* **20:** 70–86. doi:10.1016/j.stem.2016.10.002

Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PAC, Rappsilber J, Helin K. 2011. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473:** 343–348. doi:10.1038/nature10066

Wu H, Zhang Y. 2014. Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell* **156:** 45–68. doi:10.1016/j.cell.2013.12.019

Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, Cui K, Zhao K, Eve Sun Y, Zhang Y. 2011. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473:** 389–393. doi:10.1038/nature09934

Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, Barbera AJ, Zheng L, Zhang H, Huang S, et al. 2011. Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell* **42:** 451–464. doi:10.1016/j.molcel.2011.04.005

Yang J, Guo R, Wang H, Ye X, Zhou Z, Dan J, Wang H, Gong P, Deng W, Yin Y, et al. 2016. Tet enzymes regulate telomere maintenance and chromosomal stability of mouse ESCs. *Cell Rep* **15:** 1809–1821. doi:10.1016/j.celrep.2016.04.058

Ying Q-L, Nichols J, Chambers I, Smith A. 2003. BMP induction of id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **115:** 281–292. doi:10.1016/S0092-8674(03)00847-X

Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A. 2008. The ground state of embryonic stem cell self-renewal. *Nature* **453:** 519–523. doi:10.1038/nature06968

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9:** R137. doi:10.1186/gb-2008-9-9-r137

Zhang H, Zhang X, Clark E, Mulcahey M, Huang S, Shi YG. 2010. TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res* **20:** 1390–1393. doi:10.1038/cr.2010.156

Zinski J, Tajer B, Mullins MC. 2018. TGF-β family signaling in early vertebrate development. *Cold Spring Harb Perspect Biol* **10:** a033274. doi:10.1101/cshperspect.a033274