Deletion of Tet proteins results in quantitative disparities during ESC differentiation partially attributable to differences in gene expression

Michael J Reimer  
Versiti
Kirthi Pulakanti  
Versiti
Linzheng Shi  
Vanderbilt University
Alex Abel  
Versity
Mingyu Liang  
Medical College of Wisconsin
Subramaniam Malarkannan  
Medical College of Wisconsin
Sridhar Rao (✉ sridhar.rao@versiti.org)  
Versiti  https://orcid.org/0000-0003-2688-9541

Research article

Keywords: Embryonic stem cells, DNA methylation, Ten eleven translocation (TET proteins), Differentiation

Posted Date: April 24th, 2019

DOI: https://doi.org/10.21203/rs.2.9320/v1

License: ☕️ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License

Version of Record: A version of this preprint was published on July 8th, 2019. See the published version at https://doi.org/10.1186/s12861-019-0196-6.
Abstract

Background: The Tet protein family (Tet1, Tet2, and Tet3) regulate DNA methylation through conversion of 5-methylcytosine to 5-hydroxymethylcytosine which can ultimately result in DNA demethylation and play a critical role during early mammalian development and pluripotency. While multiple groups have generated knockouts combining loss of different Tet proteins in murine embryonic stem cells (ESCs), differences in genetic background and approaches has made it difficult to directly compare results and discern the direct mechanism by which Tet proteins regulate the transcriptome. To address this concern, we utilized genomic editing in an isogenic pluripotent background which permitted a quantitative, flow-cytometry based measurement of pluripotency in combination with genome-wide assessment of gene expression and DNA methylation changes. Our ultimate goal was to generate a resource of large-scale datasets to permit hypothesis-generating experiments. Results: We demonstrate a quantitative disparity in the differentiation ability among Tet protein deletions, with Tet2 single knockout exhibiting the most severe defect, while loss of Tet1 alone or combinations of Tet genes showed a quantitatively intermediate phenotype. Using a combination of transcriptomic and epigenomic approaches we demonstrate an increase in DNA hypermethylation and a divergence of transcriptional profiles in pluripotency among Tet deletions, with loss of Tet2 having the most profound effect in undifferentiated ESCs. Conclusions: We conclude that loss of Tet2 has the most dramatic effect both on the phenotype of ESCs and the transcriptome compared to other genotypes. While loss of Tet proteins increased DNA hypermethylation, especially in gene promoters, these changes in DNA methylation did not correlate with gene expression changes. Thus, while loss of different Tet proteins alters DNA methylation, this change does not appear to be directly responsible for transcriptome changes. Thus, loss of Tet proteins likely regulates the transcriptome epigenetically both through altering 5mC but also through additional mechanisms. Nonetheless, the transcriptome changes in pluripotent Tet2-/- ESCs compared to wild-type implies that the disparities in differentiation can be partially attributed to baseline alterations in gene expression.

Background

DNA methylation plays a critical role in regulating gene expression during development and is maintained in mammals through a complex interplay between DNA methyltransferases (DNMT) and a family of proteins termed Ten Eleven Translocation (Tet). DNMTs are the “writer” of DNA methylation, whereas Tet proteins are the “erasers” [1-4]. All Tet members are capable of oxidizing methylated cytosines (5mC) to 5-hydroxymethylated cytosine (5hmC). Further oxidation of the 5hmC mark by Tet proteins yields 5-carboxy cytosine (5caC) and 5-formyl cytosine (5fC; [5]) which are unstable and removed by base excision repair (BER; [6-8]), resulting in conversion to an unmodified cytosine and is referred to as active DNA demethylation. Tet proteins can also cause DNA demethylation through a passive mechanism. The 5hmC mark is not recognized by maintenance DNMTs during DNA replication and will be absent in the daughter strand, resulting in passive DNA demethylation in a cell-cycle dependent manner. Regulation of DNA
methylation is especially important during early embryogenesis where dynamic, rapid changes in the epigenome are required for early differentiation steps.

Tet1 and Tet2 are highly expressed in the inner cell mass (ICM; [9]) from which murine embryonic stem cells (ESCs) are derived, but studies in ESCs on the role of Tet proteins have resulted in conflicting reports. *In vivo* studies of Tet deletion have established that loss of either Tet1 [10] or Tet2 [11] alone had no effect on early development. The combination deletion of Tet1 and Tet2 (double knockout, DKO) exhibited a partially penetrant phenotype in which a fraction of embryos died perinatally, but others were overtly normal [12]. Even though Tet3 is only expressed at low levels in ESCs and early embryos, combined loss of Tet1, Tet2, and Tet3 in ESCs prevents them from contributing to chimeric embryos during blastocyst complementation [13], implying that some amount of Tet protein activity is required for normal embryonic development. Importantly, because of the complex cell:cell interactions within the developing embryo, as well as differences in mouse strain background, it has not been possible to translate these *in vivo* results into how Tet proteins regulate ESCs and pluripotency. There remains debate about which Tet genes are required for pluripotency and/or differentiation, and how they may interact in combination remains unclear. Part of this likely relates to the differences in how the ESC were isolated, differences between ESC genetic background, as well as the use of classical homologous recombination approaches which induce large genomic deletions [10, 12, 13]. Genomic editing, which can generate null alleles by inducing small frame-shifts to prevent functional protein production rather than large genomic deletions, has been used at all the Tet genes [14, 15], but the resulting ESCs have not been carefully quantitatively characterized in terms of their differentiation and pluripotency. To address this, we utilized a GFP reporter line to permit quantitative tracking of pluripotency in combination with combinatorial genomic-editing to ablate all three Tet genes in an isogenic background. By utilizing both transcriptome and genome-wide DNA methylation analysis we establish that the differences we observe are at least partially related to baseline differences in the epigenome and transcriptome of these cells.

**Results**

**Deletion of Tet proteins Does Not Disrupt Pluripotency**

We utilized a previously generated ESC line in which eGFP was “knocked-in” the Pou5f1 locus, permitting use of eGFP expression as a quantitative marker of pluripotency (Figures 1A-C). A single clone was used as a parental line for all further experiments and hereafter referred to as wild-type (WT). Published gRNAs [14] to all three Tet genes were transfected into the parental line to generate the following lines: Tet1-/-, Tet2-/-, DKO (Tet1-/-:Tet2-/-), and TKO (Tet1-/-:Tet2-/-:Tet3-/-; Figure 1A). Unless otherwise noted, all experiments were performed with 3 individual clones of WT, Tet1-/- and 2 clones of Tet2-/-, DKO, TKO. Generation of indels by genomic editing was confirmed by direct sequencing (data not shown) and loss of protein by Western blot (Figure 1D). All Tet deletion lines expressed similar levels of eGFP (Figure 1E) and common pluripotency markers (Figure 1D). Expression of three lineage markers, Gata6 (endoderm), Cdx2 (trophectoderm), and Brachyury (mesoderm) were variable across lines (Figure 1F) when measured by RT-qPCR, likely due to clone-to-clone variation given the very low-level expression of these markers.
These data are consistent with published reports that deletion of Tet proteins in any combination does not alter pluripotency per se.

**Deletion of Tet Proteins Impairs Differentiation**

To determine the effect of Tet deletion on differentiation we performed a leukemia inhibitory factor (LIF) withdrawal assay. LIF maintains ESC pluripotency and its withdrawal promotes differentiation down all three germ layers. All lines were grown without LIF for six days and differentiation assayed by flow cytometry for eGFP expression. All Tet deletion lines, including independent clones for each genotype, were resistant to differentiation as measured by retention of eGFP expression, consistent with prior publications (Figure 2A, B; [10, 12, 13, 16]). Surprisingly, Tet2-/- clones while quantitatively more resistant to differentiation as expected [17] than the other lines, the DKO and TKO lines more closely resembled the Tet1-/-. This suggests that Tet1 deletion is the “dominant” phenotype since the DKO and TKO differentiation mirrors Tet1-/- and not Tet2-/-.

Given the agreement among clones, we used multiple clones for all subsequent analyses and grouped them as biological replicates, thereby allowing interclonal variations to drive our statistical tests.

To confirm the loss of pluripotency and resistance to differentiation using an independent approach, we performed Western blots for the pluripotency factors Nanog and Oct3/4 and the mesoderm marker Brachyury/T (Figure 2C). All Tet genotypes retained higher levels of both Nanog and Oct3/4 when compared to the WT (Figure 2C). Importantly, Oct3/4 protein levels as measured by Western blot closely resemble the eGFP expression, indicating our flow cytometry approach is a quantitative measurement of pluripotency. To identify morphological differences during differentiation, cells were stained for alkaline phosphatase (AP) at day 0 and 6 of LIF withdrawal (Figure 2E). At D0 no gross morphological changes were observed in any Tet deletion lines (Figure 2E, top panel). Following six days of LIF withdrawal, WT colonies at D6 appear flatter and are negative for AP, Tet2-/- colonies were less spread out than WT and remained AP positive whereas the other lines displayed an intermediate phenotype (Figure 2E, bottom panel). In agreement with several published studies, we found no defects in pluripotency following deletion of Tet proteins but uncovered a quantitative discrepancy between Tet2-/- and other genotypes which phenocopied Tet1-/-.

Transcriptional Profiling of Tet Deletions in Pluripotency and LIF Withdrawal

To delineate the disparities in differentiation among the various genotypes, RNA-seq was performed at D6 of LIF withdrawal. We chose LIF withdrawal because it permits ESCs to differentiate down multiple lineages through loss of a signaling cascade rather than a strong positive differentiation signal which may promote its own set of transcriptome changes. We performed RNA sequencing (RNA-seq) on differentiated (D6) samples and read counts were normalized using ERCC spike-ins to enhance quantification [18]. D6 analysis of the top 1000 most variable genes across genotypes revealed a unique transcriptional profile of Tet2-/- cells that was not shared by other Tet deletions or the WT based upon
unsupervised hierarchical clustering, consistent with the Tet2-/−-cells exhibiting a very different cell state after LIF withdrawal compared with the other genotypes (Figure 3A). Consistent with our eGFP analysis, the Tet1-/−, DKO, and TKO cells most closely resembled each other at the transcriptome level, again indicating that the Tet1-/−-intermediate differentiation phenotype was dominant over Tet2-/−. To further refine which genes were differentially expressed, we utilized both a statistical (adjusted p-val <0.05) and fold change (>2-fold) criteria to generate refined gene lists which likely represented the most differentially expressed as compared to WT cells differentiated in parallel. In terms of the genes which were underexpressed (Figure 3B) or overexpressed (Figure 3C) we found that the Tet2-/−-cells exhibited the most uniquely dysregulated genes compared to WT, consistent with Tet2-/−-being in a distinct cell state compared to differentiated WT cells. Our results are consistent with the Tet2-/−-cells at D6 of differentiation being very different from WT, with Tet1-/−, DKO, and TKO cells representing an intermediate defect in differentiation following LIF withdrawal.

Given our results, we wondered whether the loss of Tet genes induced baseline transcriptome changes in the ESCs which were compatible with pluripotency, or alternatively Tet proteins were only required during differentiation as posited in [17]. We performed RNA-seq on undifferentiated ESCs (D0) using the same approach as above. The top 1000 most variable genes are displayed as a heat map in Figure 4A and indicate that Tet2-/−-cells are transcriptionally distinct from the other genotypes at D0. In contrast to the D6 results, the TKO cells appear to be distinct at D0 than the Tet1-/−, WT, and DKO cells. This was surprising, since Tet3 is generally expressed at low levels in ESCs, although other groups have shown that Tet3 expression is upregulated in DKO cells with minimal effects on global levels of 5hmC [12]. Nonetheless, our results are consistent with the Tet2-/−-having a very different gene expression pattern compared to the other genotypes. Given the role of Tet proteins in DNA demethylation, it was not surprising that compared to WT cells at D0 our differential expression analysis showed virtually all genes were downregulated (Figure 4A). Not surprisingly, genes expressed at lower levels in Tet2-/−-versus WT cells were predominantly unique to this genotype and not shared with the others, again indicating that loss of Tet2 induced a unique change on the transcriptome in ESCs compared to the other genotypes at D0.

Prior work from others has demonstrated that loss of Tet2 in ESCs delays differentiation because of a failure to activate enhancers critical to differentiation programs [17]. Thus, loss of Tet proteins could prevent proper differentiation either through baseline transcriptome differences, an inability to activate differentiation-critical gene expression programs, or perhaps both. To address this, gene lists for significantly downregulated genes at D0 were compared for overlap between different knockouts (Figure 4B). Tet2-/− had the most downregulated genes (299) as compared to the WT and showed little overlap with any other Tet genotype besides TKO (60 shared). Taken together, we conclude that Tet proteins predominantly function to activate gene expression and Tet2-/−-induces unique transcriptome changes not shared by other Tet deletions at D0. This would be consistent with baseline transcriptome differences in the different genotypes at D0. Collectively, our data along with the literature [17] would suggest loss of Tet proteins prevents differentiation through multiple mechanisms.
To obtain a more global view of transcriptome differences between the various genotypes at both D0 and D6, we performed a principal component analysis (PCA; Figure 4C). Consistent with our prior analysis, differentiated (D6) WT cells segregated into a unique quadrant, indicating they were highly distinct from the other cells. D6 DKO, TKO, and Tet1-/- clustered together and were distinct from the D6 Tet2-/-, again confirming, using an alternative analytic approach, that D6 Tet2-/- cells were different than other genotypes. Surprisingly, the D0 cells, all of which display a pluripotent phenotype, were less clustered than anticipated. The D0 WT cells were the most distinct, and again the D0 DKO, TKO, and Tet1-/- formed a cluster. The D0 Tet2-/- cells were distinct from the other cells again, indicating that at baseline their transcriptome is distinct from the other genotypes.

To better delineate the pathways which may be altered in Tet1-/- and Tet2-/- as compared to WT cells we first performed Gene Ontology analysis using PANTHER [19] to assign differentially expressed down-regulated transcripts to their respective biological process (Figure 5A). Because Tet1-/- had far fewer down-regulated transcripts (24) then Tet2-/- (421), fewer overall pathways were identified overall. Importantly, there was rough agreement in general between the processes effected by loss of either Tet1 or Tet2. Thus, while loss of Tet2 had a more dramatic effect on the overall transcriptome then loss of Tet1, there was general agreement between the biological processes impacted by loss of either Tet1 or Tet2.

**Tet Deletion Alters DNA Methylation which does not correlate with gene expression changes.**

We next hypothesized that changes in DNA methylation in pluripotent ESCs explains their transcriptome differences, at least partially. To quantitate DNA methylation differences, we performed reduced representation bisulfite sequencing (RRBS) on D0 samples. RRBS is highly quantitative and provides robust coverage of CpG islands which are abundant in mammalian gene promoters. Differentially methylated regions (DMRs) were defined as having a change of at least 25% compared to WT. Overall, we observed more hypermethylated than hypomethylated regions in all Tet genotypes, consistent with their role in DNA demethylation (Figure 5B; [3, 4]). It should be noted that the Tet2-/- cells displayed the most hypermethylation overall compared to wild-type, whereas the Tet1-/-, DKO, and TKO all showed a similar phenotype. DMRs were found at consistent ratios in all genotypes across promoters, introns, exons, and intergenic elements (data not shown), consistent with a genome-wide change in DNA methylation. Importantly, when we looked at a small window around TSS (+/- 2kb) corresponding to gene promoters, we observed a larger predominance of DNA hypermethylation rather than hypomethylation (Figure 5C), consistent with prior literature that Tet proteins bind to a large fraction of promoters [20, 21].

To determine if changes in DNA methylation correlated with changes in mRNA, we identified DMRs within promoters for each genotype and then mapped the fold change in the gene (Figure 6). Surprisingly, changes in DNA methylation at gene promoters did not correlate with transcriptional changes in the D0 RNA-seq. This is consistent with prior literature [10, 12, 13, 20, 21]. Thus, even though we utilized genomic editing in an isogenic background with multiple clones, we were unable to mechanistically link changes in DNA methylation to altered gene expression.
Discussion

In this study we used CRISPR/Cas9 to generate Tet1-/-, Tet2-/-, DKO, and TKO in an isogenic ESC line. In accordance with published studies, we did not observe any perturbation of pluripotency following Tet deletion. We did however uncover a quantitative disparity across genotypes during differentiation. All genotypes were resistant to differentiation, consistent with published reports [10, 12, 13, 16]. Importantly, these experiments revealed three key findings. First, while loss of Tet proteins caused a block in differentiation, the loss of Tet2 induced a near complete block with a >90% retention of eGFP expression. Second, loss of Tet1 caused only a partial reduction in differentiation. Third, that in both DKO and TKO cells their differentiation potential phenocopied loss of Tet1 rather than loss of Tet2, implying that the Tet1-/phenotype is dominant. This quantitative difference has not previously been described within the literature and was most likely difficult to conclusively observe because of the lack of isogenic backgrounds and our use of at least two, independent clones for all genotypes minimizes the chance that differences are simply related to clone-to-clone variation. In addition, by using LIF withdrawal rather than a potent differentiation agent such as retinoic acid we were able to observe a more subtle but nonetheless quantitative difference among genotypes. Importantly, many labs have published using TKO embryonic cells representing the most dramatic phenotype [13, 22]. Our data would argue this may not be the case, since the Tet2-/ESCs exhibit a substantial, quantitative reduction in their ability to differentiate.

Importantly, while our study focused on the role of Tet proteins in vitro, these findings are inconsistent with the observed in vivo phenotypes. As mentioned earlier, loss of either Tet1 or Tet2 appears to be well-tolerated by embryos, and in fact double knockouts (Tet1-/Tet2-/Tet3-/Tet3-/-) are able to generate healthy adults, although not at the expected Mendelian ratios [10-12]. Only triple knock-out (Tet1-/Tet2-/Tet3-/-Tet3-/-) animals exhibit early developmental defects, illustrating that the in vivo and in vitro results of ESCs are not fully compatible [13]. Given the complex environment in vivo where both external signals including cell-cell interactions and diffusible signaling factors play critical roles in organismal development, it is perhaps not surprising there is discordance between our in vitro results and in vivo studies. Nonetheless, while our studies may not lend insights into the in vivo role of Tet proteins, they do provide some key insights into how they operate to regulate the transcriptome of pluripotent cells.

A second area of disparity with the literature surrounds seminal studies by Meelad Dawlaty and Rudy Jaenisch [10, 12, 13], in which ESCs deficient in different combinations of Tet proteins were derived. In these studies, classical homologous recombination was used to generate animals with combinatorial loss of Tets, and then ESCs were derived from blastocysts. While in vivo these cells contributed poorly to chimeras during blastocyst complementation, they were able to form embryoid bodies in vitro, consistent with our findings that TKO cells have an impaired but not complete block in differentiation. This disparity between our groups results and the Jaenisch findings could be for two different reasons. The first may be technical, given the significant differences between genomic editing in a single, parental ESC line as opposed to classical homologous recombination, breeding adult animals, followed by derivation of ESCs. The second is our use of an Oct4-EGFP reporter, originally developed by the Jaenisch group [15, 23], which permits quantitative measurement of pluripotency and a direct comparison among the different
genotypes. Irrespective of the differences, it is clear from our work and others [17] that loss of Tet2 has the most dramatic phenotype on ESC differentiation, but further studies will be required to explain the discordance between our work and the studies from others.

The second part of this study used genome-wide approaches to quantitate transcriptome and epigenomic differences. To our surprise given their phenotypic similarity at D0, RNA-seq showed substantial gene expression differences between Tet2/- and other genotypes. These changes did not affect pluripotency but imply the differences during LIF withdrawal are at least partly attributable to baseline disparities in the transcriptomes between genotypes. Importantly, we cannot exclude that Tet2 is also required for remodeling chromatin during differentiation [17], which warrants further investigation to tease apart these possibilities. From the D0 results we conclude that Tet2/- cells induce a unique transcriptional program as compared to loss of the other Tets. Given the profound block in differentiation we saw in the Tet2/-ESCs, we were surprised to find that additional deletion of Tet1/- in DKO or TKO cells more closely phenocopied Tet1/- cells, both in terms of their differentiation potential and their baseline transcriptome. This phenotype requires further investigation, since in terms of baseline transcriptome changes loss of Tet1 alone had modest effects. This implies that the combinatorial loss of both Tet1 and Tet2 that is critical to the intermediate phenotype. Importantly, given the close correlation between the baseline D0 transcriptome of Tet1/-, DKO, and TKO cells, it is likely that the intermediate differentiation phenotype again is at least partially attributable to the role of different Tet proteins in chromatin remodeling during differentiation. Nonetheless, this pattern of differentiation block among the various genotypes could only have been uncovered using a single, parental, isogenic line which easily permitted the quantitative measurement of pluripotency.

Since Tet proteins canonically function in DNA demethylation, we hypothesized that changes in D0 RNA-seq would be secondary to alterations in DNA methylation. Overall, we observed a majority of DMRs in the Tet deletion genotypes were hypermethylated compared to WT. Tet2/- had the largest number of hypermethylated DMRs even though this genotype had the second fewest of total DMRs. Nonetheless, our data is consistent with the work of others, in that there was little correlation between the changes in DNA methylation and transcriptome changes we observed [9, 10, 12, 13, 20]. This would imply that the observed transcriptome changes are separate from the role of Tet proteins in active DNA demethylation. It may be that our transcriptome differences are secondary to loss of 5hmC, which we did not directly measure. Alternatively, it may be the result of Tet proteins modulating gene expression independent of their oxidative capacity. In either case, this important question remains unresolved and requires further study.

**CONCLUSIONS**

Tet proteins are required for proper embryonic development [10, 12, 13] and in ESC by interacting with the pluripotency-associated transcription factor Nanog [24]. While the in vivo roles during early embryogenesis has been well characterized, there remains conflicting literature about the role either individually or combinatorially for the Tet proteins in ESCs. To address this gap, we utilized genomic
editing to delete different \textit{Tet} genes individually and in combination. Key findings are that loss of \textit{Tet}2 by itself displayed the most significant, quantitative block in differentiation during LIF withdrawal. What is most surprising is that \textit{Tet1-/-}, DKO, and TKO cells phenocopied each other during differentiation and displayed very similar transcriptome changes both prior and following LIF withdrawal. Consistent with other publications however, we were unable to correlate changes in gene expression with altered DNA methylation. Thus, while transcriptome differences are at best only partially attributable to the DNA demethylase activity of Tet proteins. Nonetheless, the baseline changes in the transcriptome induced by loss of \textit{Tet}2 in particular are likely at least partially responsible for the inability of these cells to differentiate. Given the isogenic background we utilized and the use of multiple clones for each genotype, our work represents a valuable resource for investigators interested in pluripotency, epigenetics, and the role of Tet proteins in regulating gene expression.

\section*{Methods}

\textbf{Generation of Tet knockout and Oct4:IRES:EGFP ESC lines}

ESC line used has been described previously [25-27]. Briefly, they are a 129/SVj derived in house through blastocyst outgrowth, and then adapted to grow under feeder free conditions on gelatin. All lines ESCs were grown on gelatin-coated tissue culture plates in DMEM supplemented with 15\% FBS, 2\% pen/strep, 1\% nucleoside mix, 1\% L-glutamine, 1\% non-essential amino acids, 10-4 M 2ME, and 103 U LIF.

Generation of the Oct4:IRES:EGFP was previously described [27]. To generate the Tet knockout lines, published gRNAs (Wang et al., 2013) were cloned into pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene #62988) and transiently transfected into the Oct4:IRES:EGFP line. Briefly, 1-2 x 106 cells were transfected with 2 \( \mu \)g plasmid DNA in the case of DKO and TKO. 24 hours later cells were selected with 2 \( \mu \)g/mL puromycin for 2 days and individual clones isolated. Clones were screened by restriction digest using primers listed in Supplemental Table 1 [14]. Restriction testing was done with the following pairs: \textit{Tet1-Sacl}, \textit{Tet2-EcoRV}, \textit{Tet3-Xhol}. Indels were confirmed by sequencing and loss of protein was confirmed by Western blot. Two (\textit{Tet}2-/-, DKO, TKO) or three (\textit{Tet}1-/-, WT) individual clones were used for each experiment.

\textbf{RNA Isolation and RT-qPCR}

RNA was isolated using TRIzol (Thermo Fisher 15596026) according to the manufacturer’s instructions. Further purification of RNA was done using the Qiagen RNeasy Mini Kit (Qiagen 7404) and converted to cDNA with the iScript cDNA Synthesis Kit (Bio-Rad 4106228). Approximately 20ng of cDNA was used for each reaction. Primers used are listed in Table 1. All primers, unless previously published, were designed using mm9.

\textbf{Western Blot}

Total protein was extracted following lysis in RIPA buffer (25mM Trizma pH 7.4, 150mM NaCl, .1\% SDS, .5\% sodium deoxycholate, and 1\% NP-40 substitute (Sigma, 74385)). Protease inhibitors were added to
RIPA as follows: 1:1000 DTT (Sigma, 646563), 5:1000 PMSF (Sigma, 93482-50ML-F), and 1:1000 Protease Inhibitor Cocktail (Sigma, P8340). 10μg of protein was run for each sample on a 4-20% Criterion Tris-HCl protein gel (Bio-Rad 3450033) and processed using standard western blot technique. All primary incubations were performed overnight in 5% BSA/TBST. Secondary incubations were done in 5% milk/TBST for approximately 1 hour. Imaging was performed on a GE Amersham Imager 600. Analysis was done in ImageJ and values were normalized to GAPDH. Western Blot: Sall4 (Abcam, ab29112), GAPDH (Santa Cruz, sc25778), Nanog (Santa Cruz sc8822), Oct3/4 (Santa Cruz, sc9081), Tet1 (Abcam, ab191698), Tet2 (Abcam, ab94580), donkey anti-rabbit IgG-HRP (Santa Cruz, sc2313), m-IgG κ BP-HRP (Santa Cruz, sc516102), Brachyury/T (Santa Cruz, N-19 sc17743)

**LIF Withdrawal and Flow Cytometry**

Approximately 5,000 ESC were plated on one well of a six-well gelatin-coated plate in ESC media with LIF and replaced the following day with ESC media lacking LIF. Media was changed every day and on day 6 the cells were analyzed using the BD LSR II flow cytometer and analyzed with FlowJo.

**Statistical Analysis**

Statistically significant differences were measured using a two-tailed Student’s t-test and a p-value <.05. Bar graphs represent the mean of all experiments and errors bars are standard error of the mean (SEM).

**Next-Generation Sequencing Library Preparation**

**RNA-seq:** 1 μg total RNA was obtained, as described above; ERCC RNA (Thermo 4456740) was then added to each sample prior to Poly-A mRNA selection (NEB E7490). RNA-seq libraries were made using the NEBNext Ultra RNA Library Prep Kit for Illumina (E7530). The following numbers of clones were used for both RNA-seq and RRBS analysis: 3 WT, 3 Tet1-/-, 2 Tet2-/-, 2 DKO, and 2 TKO.

**RRBS:** Libraries were made from 100 ng gDNA using Premium Reduced Representation Bisulfite Sequencing Kit (Diagenode C02030033). All libraries were run as paired-end (76 total cycles) on an Illumina NextSeq 500.

**RRBS Analysis**

Raw sequence reads had the first 6 base pairs clipped off the 5’ end and were also trimmed to remove both poor quality calls using Trim Galore v0.50 (--clip_r1 6) (www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Adapter sequences were removed with Cutadapt v1.16 [28]. The refined sequences were mapped to mouse reference genome(mm9) using Bismark v0.19.1 [29] with Bowtie2 v2.1.0 [30] at default parameters.

The R package methylKit v1.4.1 [31] was used for further analysis. The aligned files from Bismark were utilized to extract the methylation calls occurring at only CpG dinucleotides with minimum of 10 read coverage. Differentially Methylated Regions (DMRs) were identified for each comparison between
treatment and wild type. DMRs were selected based on q-value < 0.05 and those that meet the minimum percent methylation difference cut-off of 25%. Reported DMRs were then annotated using HOMER v4.10 software [32] and the distribution of DMRs in different genomic elements were plotted. Promoters were defined as -1kb to +100bp around TSS of RefSeq gene. Intergenic partitions were defined as genomic regions that did not overlap with promoters, exons and introns.

RNA-seq Analysis

The raw RNA sequence reads were mapped to mouse reference genome build mm9 using STAR v2.5.1 [33] using default parameters and including normalization using the ERCC spike-ins. Quality control matrices were confirmed with a FastQC program. Differential Expression analysis was done using DESeq package in R [34]. Differentially expressed genes were called as significant at Benjamin-Hochberg adjusted p-value < 0.05 and fold change of 2. Principal Component Analysis (PCA) was done in R to see the variance between the samples. Heatmaps were done in R using heatmap.2 in R with Pearson method for distance and average agglomeration for clustering. Top 1000 variable genes with log2 read counts were plotted in heatmap after removing duplicates and miRNA genes.

List Of Abbreviations

Tet: Ten Eleven Translocation

5mC: 5-MethylCytosine

5hmC: 5-HydroxymethylCytosine

5caC: 5-CarboxyCytosine

5fC: 5-FormylCytosine

ESC: Embryonic Stem Cells

DNMT: DNA MethylTransferase

BER: Base Excision Repair

ICM: Inner Cell Mass

KO: KnockOut of a single gene such as Tet1/- or Tet2/-

DKO: Double KnockOut- Tet1-/-;Tet2-/-

TKO: Triple KnockOut- Tet1-/-;Tet2-/-;Tet3-/-

Indel: Insertion or Deletion
eGFP: Enhanced Green Fluorescent Protein

WT: Wild-Type

gRNA: Guide RNA

RT-qPCR: Reverse Transcription Quantitative PCR

LIF: Leukemia Inhibitory Factor

AP: Alkaline Phosphatase

D0: Day 0, prior to LIF withdrawal

D6: Day 6 following LIF withdrawal

RNA-seq: Next Generation sequencing of RNA for transcriptomics

PCA: Principal Component Analysis

RRBS: Reduced Representation Bisulfite Sequencing

CpG: Cytosine-Guanine dinucleotide

DMR: Differentially Methylated Regions

TSS: Transcriptional Start Site

DMEM: Dulbecco's Modified Eagle's Media

FBS: Fetal Bovine Serum

Pen/Strep: Penicillin and Streptomycin

2ME: 2-MercaptoEthanol

cDNA: Complementary DNA

mm9: Mus Musculus genomic build 9

RIPA: RadioImmunoPrecipitation Assay Buffer

DTT: DiThioThreitol

BSA: Bovine Serum Albumin

TBST: Tris-Buffered Saline with 0.1% Triton X-100
Declarations

Ethics approval and consent to participate: Not Applicable

Consent for publication: Not Applicable.

Availability of data and material: All data sets generated in this publication have been submitted to the Gene Expression Omnibus (GEO) under SuperSeries GSE122814. This data is currently private but can be accessed by reviewers using the token: uxkteuqildujniz.

Competing interests: The authors declare that they have no competing interests.

Funding: This study was supported by the Versiti Blood Research Foundation and the Midwest Athletes against Childhood Cancer to SR. SR is also supported by NCI R01 (CA204231). The funding agencies did not participate in the study design, data collection, data analysis, data interpretation, nor manuscript preparation.

Authors' contributions: SR and MR designed the experiments. MR performed the bulk of the experiments. KP performed the computational analysis. LS assisted with experiments. AA and SM assisted with flow cytometry. ML provided assistance with the DNA methylation analysis.

Acknowledgements: We would like to thank Rudolf Jaenisch and Meelad Dawlaty for providing WT and DKO cells from a prior publication [12]. These cells provided a critical foundation to our initial experiments and we remain grateful to their generous contribution.

References

REFERENCES:

1. Guo JU, Su Y, Zhong C, Ming G-I, Song H: Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. cc 2011, 10(16):2662-2668.

2. Ito K, Suda T: Metabolic requirements for the maintenance of self-renewing stem cells. Nat Rev Mol Cell Biol 2014, 15(4):243-256.

3. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L et al: Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science 2009, 324(5929):930-935.
4. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y: Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 2011, 333(6047):1300-1303.

5. Ito S, D’ Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y: Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 2010, 466(7310):1129-1133.

6. Cortazar D, Kunz C, Selfridge J, Lettieri T, Saito Y, MacDougall E, Wirz A, Schuermann D, Jacobs AL, Siegrist F et al: Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* 2011, 470(7334):419-423.

7. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D et al: Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* 2011, 146(1):67-79.

8. Muller U, Bauer C, Siegl M, Rottach A, Leonhardt H: TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation. *Nucleic Acids Res* 2014, 42(13):8592-8604.

9. Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, Laiho A, Tahiliani M, Sommer CA, Mostoslavsky G et al: Tet1 and Tet2 Regulate 5-Hydroxymethylcytosine Production and Cell Lineage Specification in Mouse Embryonic Stem Cells. *Cell Stem Cell* 2011, 8(2):200-213.

10. Dawlaty MM, Ganz K, Powell BE, Hu Y-C, Markoulaki S, Cheng AW, Gao Q, Kim J, Choi S-W, Page DC et al: Tet1 Is Dispensable for Maintaining Pluripotency and Its Loss Is Compatible with Embryonic and Postnatal Development. *Cell Stem Cell* 2011, 9(2):166-175.

11. Li Z, Cai X, Cai CL, Wang J, Zhang W, Petersen BE, Yang FC, Xu M: Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood* 2011, 118(17):4509-4518.

12. Dawlaty MM, Breiling A, Le T, Raddatz G, Barrasa MI, Cheng AW, Gao Q, Powell BE, Li Z, Xu M et al: Combined Deficiency of Tet1 and Tet2 Causes Epigenetic Abnormalities but Is Compatible with Postnatal Development. *Developmental Cell* 2013, 24(3):310-323.

13. Dawlaty MM, Breiling A, Le T, Barrasa MI, Raddatz G, Gao Q, Powell BE, Cheng AW, Faull KF, Lyko F et al: Loss of Tet Enzymes Compromises Proper Differentiation of Embryonic Stem Cells. *Developmental Cell* 2014, 29(1):102-111.

14. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R: One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell* 2013, 153(4):910-918.

15. Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R: One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering. *Cell* 2013, 154(6):1370-1379.
16. Hu X, Zhang L, Mao S-Q, Li Z, Chen J, Zhang R-R, Wu H-P, Gao J, Guo F, Liu W et al: Tet and TDG Mediate DNA Demethylation Essential for Mesenchymal-to-Epithelial Transition in Somatic Cell Reprogramming. Stem Cell 2014, 14(4):512-522.

17. Hon GC, Song C-X, Du T, Jin F, Selvaraj S, Lee AY, Yen C-a, Ye Z, Mao S-Q, Wang B-A et al: 5mC Oxidation by Tet2 Modulates Enhancer Activity and Timing of Transcriptome Reprogramming during Differentiation. Mol Cell 2014:1-12.

18. Risso D, Ngai J, Speed TP, Dudoit S: Normalization of RNA-seq data using factor analysis of control genes or samples. Nature Biotechnology 2014, 32(9):896-902.

19. Muruganujan A, Ebert D, Mi H, Thomas PD, Huang X: PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Research 2018, 47(D1):D419-D426.

20. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PAC, Rappsilber J, Helin K: TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature 2011, 473(7347):343-348.

21. Wu H, D’alessio AC, Ito S, Xia K, Wang Z, Cui K, Zhao K, Eve Sun Y, Zhang Y: Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature 2011, 473(7347):389-393.

22. Lu F, Liu Y, Jiang L, Yamaguchi S, Zhang Y: Role of Tet proteins in enhancer activity and telomere elongation. Genes Dev 2014, 28(19):2103-2119.

23. Lengner CJ, Camargo FD, Hochedlinger K, Welstead GG, Zaidi S, Gokhale S, Scholer HR, Tomilin A, Jaenisch R: Oct4 expression is not required for mouse somatic stem cell self-renewal. Cell Stem Cell 2007, 1(4):403-415.

24. Costa Y, Ding J, Theunissen TW, Faiola F, Hore TA, Shliaha PV, Fidalgo M, Saunders A, Lawrence M, Dietmann S et al: NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. Nature 2013:1-5.

25. Pulakanti K, Pinello L, Stelloh C, Blinka S, Allred J, Milanovich S, Kiblawi S, Peterson J, Wang A, Yuan G-C et al: Enhancer transcribed RNAs arise from hypomethylated, Tet-occupied genomic regions. Epigenetics 2013, 8(12):1303-1320.

26. Blinka S, Reimer MH, Jr., Pulakanti K, Rao S: Super-Enhancers at the Nanog Locus Differentially Regulate Neighboring Pluripotency-Associated Genes. Cell Rep 2016, 17(1):19-28.

27. Stelloh C, Reimer MH, Pulakanti K, Blinka S, Peterson J, Pinello L, Jia S, Roumiantsev S, Hessner MJ, Milanovich S et al: The cohesin-associated protein Wapal is required for proper Polycomb-mediated gene silencing. Epigenetics & chromatin 2016, 9:14.
28. M M: Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetJournal* 2011, 17(1).

29. Krueger F, Andrews SR: Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 2011, 27(11):1571-1572.

30. Langmead B, Salzberg SL: Fast gapped-read alignment with Bowtie 2. *Nat Meth* 2012, 9(4):357-359.

31. Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE: methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology* 2012, 13(10):R87.

32. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK: Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010, 38(4):576-589.

33. Bioinformatics AD: RNA-STAR: ultrafast universal spliced sequences aligner: Supplementary materials.

34. Anders S, Huber W: Differential expression analysis for sequence count data. *Genome Biol* 2010, 11(10):R106.

**Supplementary Material Legends**

**Supplemental Table 1.** Primers and gRNAs.

**Supplemental Table 2.** RNA-seq readcounts at D0 and D6 for all samples.

**Supplemental Table 3.** RRBS DMRs.

**Figures**
Figure 1

Loss of Tet proteins does not impair pluripotency A) Schematic showing the generation of the Oct4:IRES:eGFP parental cell line followed by deletion of each TET protein individually or combinatorially (DKO=Tet1-/-::Tet2-/-, TKO=Tet1-/-::Tet2-/-::Tet3-/-). B) Representative histogram showing eGFP expression in a WT clone. Parental (GFP-) mESC was used as a negative control. C) Brightfield and fluorescent micrographs (10x) of WT Oct4:IRES:eGFP before and after exposure to 5µM retinoic acid to induce differentiation. D) Representative western blot of Tet1, Tet2, and common pluripotency factors in each line. GAPDH was used as a loading control. E) Representative histogram of eGFP expression in WT, Tet1-/-, Tet2-/-, DKO, and TKO clone. F) RT-qPCR of endoderm (Gata6), trophectoderm (Cdx2), and mesoderm (Brachyury/T) associated transcription factors. *=p<.05 compared to WT.
Figure 2

Loss of Tet proteins impairs differentiation. A) Column scatter plot showing the ratio of D6 MFI to D0 eGFP MFI of each cell line. Individual clones are shown to demonstrate minimal inter-clone variability. B) Similar to (A), but clones are grouped together for the purpose of statistical tests. *=p <.05 as compared to WT, #=p <.05 compared to Tet2-/- . WT n=10, Tet1-/- n=8, Tet2-/- n=8, DKO n=4, TKO n=4. C) Representative western blot of the pluripotency factors Nanog, Oct3/4 and the mesoderm marker Brachyury (T). D) Quantification of western blots for Nanog and Oct3/4. n=2. *=p<.05, compared to WT.
E) Brightfield micrographs (top panel) of alkaline phosphatase staining at day 0 (4x) and following 6 days (bottom panel) of LIF withdrawal (10x).

**Figure 3**

RNA-seq analysis of D6 LIF withdrawal samples. A) Heatmap of top 1000 most variable genes at D6. B) Venn diagram of significantly Up (p<.05, |log2 fold change| >1). C) Venn diagram of significantly Up (p<.05, |log2 fold change| >1).
Figure 4

RNA-seq analysis following D0 A) Heatmap of top 1000 variable genes at D0. B) Venn diagram of significantly down (p<.05, log2 fold change <-1) genes at D0, as compared to WT. C) PCA plot of RNA-seq data from D0 (red) and D6 (green) samples. % variability encompassed within each principle components are shown in parentheses.
Figure 5

Gene ontology for altered gene expression and DNA methylation. A) Gene Ontology terms for Tet1/- and Tet2/- down regulated genes compared to WT were identified using Panther. # to the right of each bar represents the number of downregulated genes identified in each pathway. The total number of genes used for each knock-out is shown in parentheses within the legend. B) Percentage of hyper and
hypomethylated DMRs as compared to WT. C) Number of hyper and hypomethylated DMRs located in proximity to gene promoter/TSS.

Figure 6

DNA methylation changes do not necessarily correlate with changes in gene expression. For each genotype (Tet1 KO, Tet2 KO, DKO, TKO), DMRs were identified in promoters as compared to WT. The corresponding Log2 Fold change in the gene's expression was then graphed. Because the dynamic range
across the different genotypes was so different, color schemes are different for each genotype and are shown at the top of each column. Total number of DMRs used for each genotype is shown below each column.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- supplement1.xlsx
- supplement2.xlsx
- supplement3.xlsx