TET2 binding to enhancers facilitates transcription factor activity in hematopoietic cells

Kasper D. Rasmussen,1,2,7,8 Ivan Berest,3,7 Sandra Keßler,1,2 Koutarou Nishimura,1,4 Lucía Simón-Carrasco,1,2 George S. Vassiliou,5,6 Marianne T. Pedersen,1,2 Jesper Christensen,1,2 Judith B. Zaugg,3 and Kristian Helin1,2,4

1Biotech Research and Innovation Centre (BRIC), Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark; 2The Novo Nordisk Foundation Center for Stem Cell Biology (Danstem), Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark; 3European Molecular Biology Institute, Structural and Computational Unit, 69115 Heidelberg, Germany; 4Cell Biology Program, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA; 5Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge CB2 0XY, United Kingdom; 6Department of Haematology, Cambridge University Hospitals NHS Trust, Cambridge CB2 0PT, United Kingdom

The epigenetic regulator TET2 is frequently mutated in hematological diseases. Mutations have been shown to arise in hematopoietic stem cells early in disease development and lead to altered DNA methylation landscapes and an increased risk of hematopoietic malignancy. Here, we show by genome-wide mapping of TET2 binding sites in different cell types that TET2 localizes to regions of open chromatin and cell-type–specific enhancers. We find that deletion of T2 in native hematopoiesis as well as fully transformed acute myeloid leukemia (AML) results in changes in transcription factor (TF) activity within these regions, and we provide evidence that loss of TET2 leads to attenuation of chromatin binding of members of the basic helix–loop–helix (bHLH) TF family. Together, these findings demonstrate that TET2 activity shapes the local chromatin environment at enhancers to facilitate TF binding and provides an example of how epigenetic dysregulation can affect gene expression patterns and drive disease development.

[Supplemental material is available for this article.]

The tet methylcytosine dioxygenase (also known as ten-eleven translocation [TET]) enzymes (TET1-3) mediate active DNA demethylation of cytosines in CG dinucleotides. This occurs by processive TET-mediated oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). The presence of 5hmC may lead to passive replication-dependent loss of DNA methylation, whereas 5fC and 5caC can be excised by thymine DNA glycosylase (TDG) and be replaced by unmodified cytosine via the base-excision repair (BER) pathway. Although targeting of TET1 to chromatin has been investigated (Williams et al. 2011; Wu et al. 2011), the mechanisms of TET2 recruitment to chromatin remain poorly understood (for review, see Rasmussen and Helin 2016).

Loss-of-function mutations of TET2 have been found in patients with a wide range of hematological diseases, including acute myeloid leukemia (AML) (Scourzic et al. 2015). More recently, high frequencies of TET2 mutations have also been observed in aging-associated clonal hematopoiesis (Genovese et al. 2014; Jaiswal et al. 2014; Xie et al. 2014) and in the poorly studied disorder clonal cytopenia of unknown significance (Kwok et al. 2015; Hansen et al. 2016). In previous studies, we and others identified a role of TET2 in protecting enhancer elements from aberrant DNA methylation (Hon et al. 2014; Lu et al. 2014; An et al. 2015; Rasmussen et al. 2015; Yamazaki et al. 2015). In addition, inhibition of TET proteins was shown to perturb chromatin architecture at enhancers in an embryonal carcinoma cell line undergoing neuronal differentiation (Mahé et al. 2017). Despite these results, direct TET2 binding at enhancers in hematopoietic cells has not been reported. In fact, previous studies mapping TET2 genome-wide occupancy in embryonic stem (ES) cells noted significant TET2 binding at CpG islands and promoters (Chen et al. 2013; Deplus et al. 2013; Peng et al. 2016) or at promoter-distal SALL4A binding sites located at enhancers (Xiong et al. 2016). These seemingly contradictory observations as well as the impact of aberrant DNA methylation at enhancer elements in hematopoietic cells remains to be resolved.

Gene expression is regulated by transcription factors (TFs) that bind DNA in a sequence-specific manner. Activation of a specific gene locus depends both on concentration of individual TFs as well as their ability to access the regulatory genomic elements that control gene expression. TF binding outside gene promoters is associated with low- or intermediate DNA methylation, enrichment of specific histone marks (e.g., monomethylation at histone H3 lysine 4 and acetylation of H3 lysine 27), as well as the presence of a nucleosome-depleted region (Stadler et al. 2011; Thurman et al. 2012). Although much work has focused on understanding the role of aberrant TF expression in leukemia, less is known about the role of the chromatin environment, and hence DNA methylation (Blattler and Farnham 2013), in modulating TF access to their cognate binding sites.
The occurrence of enhancer DNA hypermethylation and hematological malignancies upon TET2 mutation suggests that DNA methylation may pose a challenge for TF binding (Thurman et al. 2012). Although several TFs have been demonstrated to bind methylated DNA and induce DNA hypomethylation (e.g., CTCF and REST) (Lienert et al. 2011; Stadler et al. 2011), many TFs show an inherent binding preference in vitro for motifs with either methylated or unmethylated CpG sites (Hashimoto et al. 2014; Wang et al. 2017; Yin et al. 2017). As an illustration of this, global loss of DNA methylation in ES cells was shown to unmask previously inaccessible genomic binding sites for the TF NRF1 (Domcke et al. 2015). Moreover, DNA methylation and TET-mediated oxidation products potentially influence binding of epigenetic “readers” (e.g., Methyl-Binding Domain [MBD] proteins) (Song and Pfeifer 2016), histone variants and nucleosome remodeling enzymes (Conerly et al. 2010; Brunelle et al. 2015), and modify the physical properties of chromatin and the shape of DNA itself (Ngo et al. 2016).

Here, we generate and analyze genomic profiles from ES cells and several hematopoietic cell types to understand the mechanisms by which loss of TET2, and the resulting altered DNA methylation landscape, affects gene expression patterns and promotes the onset of hematological malignancies.

## Results

Depletion of TET2 in ES and hematopoietic cells results in widespread changes in the DNA methylation landscape (Ko et al. 2010; Asmar et al. 2013; Hon et al. 2014; An et al. 2015; Rasmussen et al. 2015; Yamazaki et al. 2015). However, it still remains unclear which regions are directly targeted by TET2 as opposed to being altered due to secondary effects of its depletion. Previous studies, in our laboratory and others, have been hampered by chromatin immunoprecipitation (ChIP) artifacts—recently coined as “Phantom peaks” (Jain et al. 2015)—and the lack of ChIP-grade antibodies against TET2. To circumvent this, we used CRISPR homology-directed repair to introduce ChIP-verified epitope tags into the endogenous Tet2 locus in murine ES cells. We identified, upon clonal expansion and characterization, two independent lines with in-frame integration of a V5-tag or a 2xFLAG (2xFL)-tag at the C terminus of Tet2 (Supplemental Fig. S1A,B). Analysis of ChIP coupled with high-throughput sequencing (ChIP-seq) using V5 or FLAG-specific antibodies revealed that the presence of the protein–protein crosslinker disuccinimidyl glutarate (DSG) in combination with formaldehyde (FA) greatly increased the signal-to-noise ratio of the ChIP signal (Fig. 1A) and increased the number of called peaks (Fig. 1B). In contrast,
crosslinking conditions with formaldehyde alone, or the combina-
tion of the longer ethylene glycol bis(succinimidy1 succinate) (EGS)
crosslinker plus formaldehyde, were insufficient to capture these
regions.

Encouraged by these results, we tested if untagged endoge-
nous Tet2 could be mapped to chromatin using these crosslinking
conditions. ES cell chromatin was crosslinked with DSG + FA and
immunoprecipitation was performed using an antibody raised
against the N terminus of Tet2 (TET2-N) or a commercially avail-
able antibody directed against the TET2 C terminus (TET2-C). We
observed a strong and reproducible enrichment of TET2-bound re-
gions using the TET2-N antibody, and this signal was absent in
knockout cells or at a negative control region (Fig. 1C). Thus, the
combination of specific antibodies and optimized workflows al-
lows for the precise mapping of TET2-bound genomic regions.

TET2 binds regions of open chromatin with enhancer features in
embryonic stem cells

Next, we generated a high-quality ChIP-seq data set to determine
genome-wide TET2 occupancy in ES cells. Quality control of bio-
logical duplicate experiments using antibodies to endogenous
TET2 (TET2-N) or the FLAG-tagged TET2 expressed from the endo-
genous locus (FLAG M2) showed excellent correlation between rep-
licate samples (Supplemental Fig. S1C–F). To avoid the false
discovery of “phantom peaks,” TET2 binding sites were defined by
enrichment over nonspecific ChIP-enriched regions identified in
Tet2 knockout cells or parental cells without FLAG-tagged
TET2. In total, this analysis revealed 26,512 TET2-bound regions
of which approximately one-third, referred to as “high-confi-
dence” TET2 binding sites, were identified with both antibodies
and had a stronger ChIP-seq signal (Fig. 2A; Supplemental Fig.
S1G,H). The vast majority (93.4%) of these 8262 sites were localized
in regions of open chromatin defined by DNase I hypersensitivity
(DHS) (Fig. 2B,C).

To understand whether TET2 is enriched in specific regions of
open chromatin in the genome, we compared TET2 binding sites
to sets of random control regions (∼5) matched for size, orienta-
tion, and distance to DHS sites. This showed a strong enrichment
of TET2 high-confidence binding sites at DHS with enhancer fea-
tures such as EP300 binding (fourfold), H3K27ac (2.5-fold), and
H3K4me1 (1.6-fold), whereas TET2 binding at promoters, CpG is-
lands, and CTCF binding sites was depleted (−0.3-fold each) (Fig
2D). In agreement with this, we found that TET2 binding sites
were located in CpG-spars regions as compared to CpG islands
(Supplemental Fig. S1I). This binding pattern is furthermore illus-
trated by a representative TET2-bound enhancer region 10 kb up-
stream of the master ES cell regulator Zfp281 (Fig. 2E). In total,
up to 90% of TET2 binding sites are associated with features of
distal regulatory elements (EP300, H3K27ac, H3K4me1, or DHS), and
nearly half of TET2-bound regions overlap with promoter-distal
EP300 binding sites (Fig. 2F). Chromatin occupancy of EP300 is a
hallmark of active enhancers, and a biochemical interaction be-
 tween EP300 and TET2 has recently been reported (Heintzman
et al. 2007; Zhang et al. 2017). Thus, our data are in agreement
with a model in which TET2 can be directly recruited to a subset
of its chromatin targets via a direct interaction with EP300
(Zhang et al. 2017).

A previous study examined the genome-wide changes in
5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC)
that occur upon Tet2 knockout in ES cells (Hon et al. 2014). We
used this data set to identify epigenetic changes around (±250
bp) high-confidence TET2 binding sites. Genetic knockout of
Tet2 is likely to result in gain of 5mC (due to absence of DNA deme-
ylation), loss of 5hmC (stable product of methylcytosine oxida-
tion by TET2), or both, at sites of TET2 recruitment. We did not
observe a linear correlation between TET2 binding and changes
in DNA methylation in Tet2 knockout (Supplemental Fig S1J,K).
However, we found DNA methylation changes consistent with
loss of TET2 catalytic activity (hyper-5mC, hypo-5hmC, or both)
at the majority of high-confidence TET2 binding sites (−93%) (Fig
2G), thereby confirming that TET2 modifies the DNA methyl-
ation state. The analysis was repeated with the full set of 26,512
sites with evidence of TET2 binding (Supplemental Fig. S2A).
Within this larger and less stringent set of regions, we observed a
modest enrichment of TET2 binding at CpG islands and gene pro-
moters (Supplemental Fig. S2B). However, in contrast to a pro-
ounced DNA hypermethylation at sites cobound by TET2 and
EP300, DNA methylation did not change at these CpG islands
and promoters in Tet2-deficient ES cells (Supplemental Fig. S2C,
D). This is consistent with the previously identified role of TET2
in maintaining a low level of DNA methylation at distal regulatory
elements, whereas CpG islands and promoters are protected from
DNA hypermethylation by additional mechanisms (Rasmussen
and Helin 2016). It should be noted that most differentially methy-
lated regions reported in Tet2 knockout ES cells (over 60,000 hy-
per-DMRs and 130,000 hypo-DMRs) (Hon et al. 2014) did not
show detectable binding of TET2 in our data set (Supplemental
Fig. S2E). This suggests that TET2 functions in the absence of
robust and persistent binding at these sites and that some of these
methylation changes (especially hypo-DMRs) can occur as a con-
sequence of secondary events to TET2 depletion, or as a result of cell
passaging (misincorporation during DNA replication) (Rasmussen
et al. 2015). Taken together, our data show that TET2 binds to re-
gions of open chromatin with enhancer features that undergo
TET2-dependent DNA demethylation.

Cell-type–specific binding pattern of TET2 in myeloid
hematopoietic versus ES cells

Next, we determined the genome-wide binding sites of TET2 in he-
matopoietic cells. To this end, we used myeloid cells immortalized
by AML1-ETO with the potential for inducible deletion of Tet2
(Tet2fl/fl;AE;Rosa26+/-CreERT2) (Rasmussen et al. 2015). Analysis
of biological duplicate ChIP experiments using the TET2-N antibody
(ChIP using a FLAG-tagged version of TET2 was not performed
because this would require generating a novel FLAG-tagged Tet2
mouse line) identified 19,706 regions significantly enriched
over knockout control. In total, only 7.4% of these regions were
shared between ES cells and myeloid cells (Fig. 3A; Supplemental
Fig. S2F,G).

We then overlapped the TET2 myeloid binding sites with vari-
ous genomic regions. Although binding of TET2 could be detect-
ed in a subset of promoters and CpG islands, the majority of
binding sites was promoter-distal (67.4%), and nearly half
(44.4%) were associated with enhancer features (Fig. 3B). An illus-
trative example of this binding pattern is shown at enhancers up-
stream of the hematopoietic master regulator Gata2 (Fig. 3C).
Since promoter DNA methylation patterns are largely unaffected
by depletion of Tet2 (Rasmussen et al. 2015), we decided to focus
our analysis on promoter-distal TET2 binding sites. Identification
of enriched Gene Ontology (GO) terms by Genomic Regions
Enrichment of Annotations Tool (GREAT) (McLean et al. 2010) re-
vealed a striking separation by cell type. For example, the

Rasmussen et al.
GO terms “stem cell maintenance” and “blastocyst formation” were enriched for TET2 binding sites in ES cells, whereas “immune system process” was highly significant for sites in myeloid cells (Supplemental Fig. S2H,I). Next, we performed motif enrichment analysis of known DNA-binding TFs. The top enriched motifs belonged to master TFs of the respective cell types such as SOX2, KLF4, ESRG, POU5F1, and NANOG in ES cells, and ERG, RUNX1, CEBPA, and GATA1 in myeloid cells (Fig. 3D,E). Together, these results indicate that TET2 is recruited to chromatin in a highly cell-type-specific manner and that TET2 binding at promoter-distal regulatory elements colocalizes with a wide range of predicted TF binding sites.

Differential analysis of chromatin accessibility reveals widespread changes in TF activity upon TET2 loss

Analysis of chromatin accessibility in gene regulatory regions can be used as a genome-wide and locus-specific measure of nonhistone protein binding to DNA (Hesselberth et al. 2009). It therefore offers an opportunity to systematically assess whether particular sets of TFs (as defined by chromatin accessibility changes at their DNA-binding motifs) are affected by loss of TET2 and potentially drive the genomic and phenotypic changes observed in TET2-mutated hematopoietic cells. Therefore, we decided to perform genomic profiling by Assay for Transposase Accessible Chromatin

Figure 2. TET2 binds regions of open chromatin with enhancer features in ES cells. (A) Venn diagram showing overlap of called peaks in TET2-N or FLAG M2 ChIP-seq experiments (left) as well as average ChIP signal from replicate samples (right). High- and low-confidence TET2 binding sites are defined, respectively, as regions showing evidence of TET2 binding in both peak sets (high) or only supported by a called positive region in one peak set (low). (B) Heat maps of ChIP-seq signal from wild-type TET2 or TET2 with two copies of a FLAG tag (2xFL). Tracks of H3K27ac and EP300 enrichment as well as regions of DHS in ES cells are also shown. The vertical axis contains all high-confidence TET2 binding sites defined in A, sorted by decreasing EP300 read counts. The horizontal axis is centered on TET2 peaks. (C) Histogram showing overlap of high-confidence TET2 binding sites with regions of DHS in ES cells. Random matched control regions (Control) were generated with same size, orientation, and distance relative to gene bodies. (D) Histogram showing overlap as C but with random matched control regions relative to DHS. The percentage of CpG islands, proximal (±250 bp) or distal (±1500 bp) transcription start sites (TSS), EP300 binding sites, H3K27ac, and H3K4me1 enriched domains, Gene body, and CTCF binding sites, overlapping TET2 high-confidence binding sites are shown. Data indicate the number of TET2 BS overlapping a given region set divided by the total number of loci in the region set (e.g., 175,237 DHS in ES cells) and fold enrichment over matched control are stated (red). (E) Representative ChIP-seq tracks showing a TET2-bound region upstream of the Zfp281 gene in ES cells. (F) Pie chart showing the distribution of high-confidence TET2 binding sites with respect to the indicated genomic regions. Each TET2 binding site is counted only once and excluded when moving clockwise from proximal promoters (e.g., 8% of TET2 high-confidence BS overlap with DHS but does not overlap with any of the preceding genomic elements). Blue hues indicate the fraction of TET2-bound regions with regulatory potential at promoter-distal sites (~90% of all TET2-bound regions). (G) Pie chart as in F showing average change in DNA methylation (5mC and 5hmC) upon TET2 loss in ES cells (whole-genome bisulfite sequencing data [WGBS] from Hon et al. 2014) in high-confidence TET2 binding sites and flanking regions (±250 bp). Only CpG sites covered by more than 10 reads were included in the analysis. Each TET2 binding site is counted only once and excluded when moving clockwise. Red hues indicate the fraction (~93%) of TET2-bound regions showing DNA methylation changes consistent with loss of TET2 catalytic activity (gain or 5mC or loss of 5hmC, or both). Regions in B–F enriched for DHS, EP300, H3K27ac, H3K4me1, and CTCF sites in ES cells were experimentally determined in the ENCODE Project (The ENCODE Project Consortium 2011).
AML cells (30%)

Finally, a considerable fraction of these peaks in both MPP and AML cells revealed that the majority of differential regions with increased DNA methylation in hematopoietic stem cells with loss of DNA accessibility in MPP cells were associated with increased DNA methylation in myeloid hematopoietic cells. We found significantly higher normalized TET2 read counts in regions with reduced accessibility upon Tet2 knockout compared to those with increased accessibility (Supplemental Fig. S3D). This suggests that TET2 is likely to have a direct role in maintaining open chromatin (and TF binding), whereas the increases in chromatin accessibility at other regions may be due to accumulated epigenetic dysregulation or associated secondary events in response to TET2 loss (e.g., imbalance of differentiation-associated cytokines and growth factors).

To identify TF binding events that are affected by TET2 loss in native chromatin, we developed a novel computational method, difftF (Berest et al. 2018), to assess “TF activity” on a genome-wide scale using profiles of chromatin accessibility by ATAC-seq. In this method, we compare the accessibility changes at putative binding sites within ATAC-seq peaks for each TF and compare this distribution to the background distribution of accessibility fold changes for all other TFs. If the putative binding sites of a TF are overall less open in the Tet2 mutant cells, we define this TF to be less “active” in Tet2 knockout cells and vice versa. Thus “TF activity” is here defined as the TF being associated with increased chromatin accessibility at its target sites. We predicted differential TF activity using a comprehensive collection of DNA-binding motifs for mouse TFs (Kulakovskiy et al. 2016). However, although this analysis identified numerous TFs with altered activity in one or several cell types, the high similarity of DNA-binding motifs in groups of related TFs (e.g., GATA family) makes it challenging to pinpoint an observed effect to a single group member. Therefore, we clustered TFs based on similarity of DNA-binding motifs (Supplemental Table S1) and compared the overall differential activity of clusters in wild-type and Tet2 knockout (Fig. 5A,B).

Loss of TET2 resulted in widespread changes in TF activity in all observed cell types. In GMP cells, we detected a strong signature of aberrant lineage differentiation characterized by increased activity of the IRF family of TFs (Cluster 18) and decreased activity of GATA family members (Cluster 3). Conversely, MPP cells showed a pronounced increase in chromatin accessibility in predicted binding sites of the GATA and CCAAT/Enhancer binding protein family (Clusters 3 and 23) and AML, GMP, and MPP cells were

Figure 3. Cell-type–specific binding pattern of TET2 in myeloid hematopoietic cells. (A) Venn diagram showing overlap of TET2 binding sites in ES cells and myeloid hematopoietic cells. Binding sites are defined from biological replicate samples by enrichment over nonspecific ChIP-enriched regions in Tet2 knockout cells using the TET2-N antibody. (B) Pie chart showing distribution of TET2-bound regions in myeloid hematopoietic cells. Each TET2 binding site is counted once only and excluded when moving clockwise from proximal promoters (e.g., 10.1% of TET2 BS overlap with DHS but does not overlap with any of the preceding genomic elements). Blue hues indicate fraction (~44%) of TET2-bound regions with regulatory potential at promoter-distal sites based on H3K27 acetylation and H3K4 methylation in myeloid cells as well as DHS in the related mouse erythroleukemia (MEL) cell line. (C) Representative ChIP-seq tracks showing TET2-bound enhancer regions upstream of the Gata2 gene in myeloid hematopoietic cells. (D) List of top enriched logos and their associated DNA-binding TFs identified in promoter-distal (~1.5 kb/+500 bp from TSS) TET2-bound regions in ES cells. (E) As in D, but for myeloid hematopoietic cells. Regions in B and C enriched for H3K27ac, H3K4me3, and H3K4me1 were experimentally determined in Rasmussen et al. (2015), and DHS in MEL cells was downloaded from the ENCODE Project (The ENCODE Project Consortium 2011).
**TET2 and TF binding in enhancers**

characterized by a loss of activity of members of the HOX protein family (Cluster 8). Finally, Tet2 knockout ES cells exhibited a prominent decrease in activity of a large TF cluster containing steroid hormone receptors including estrogen receptor alpha (Cluster 6). Several recent studies are consistent with a direct role of DNA methylation and TET2 catalytic function to facilitate activity and/or expression of members of these TF clusters: (1) CEBPB has been reported to have a preference for binding to motifs with a methylated cytosine, thus potentially implicating TET2 as a factor that restricts C/EBP family TF binding to chromatin (Mann et al. 2013; Zib et al. 2016); (2) consistent with loss of HOX gene activity in AML cells, TET2 catalytic activity has been implicated in maintenance of expression of the HOXA cluster in various cell types (Bocker et al. 2012); and (3) TET2 has recently been shown to mediate recruitment of estrogen receptor alpha by demethylation of estrogen responsive enhancers in the breast cancer cell line MCF-7 (Wang et al. 2018). Together, these results suggest that loss of TET2 catalytic activity, and the resulting alterations in DNA methylation and hydroxymethylation, has pleiotropic effects on chromatin accessibility and TF binding, most likely related to cell-type-specific chromatin structure and TF expression patterns.

**Basic helix-loop-helix (bHLH) TFs are adversely affected by loss of TET2 in hematopoietic cells**

To understand whether the observed TF activity changes could potentially explain the altered hematopoietic differentiation program observed in Tet2 knockout hematopoietic cells, we focused on a cluster of basic helix-loop-helix (bHLH) TFs (Cluster 12), whose activity was consistently decreased upon Tet2 loss in all cell types and this decrease was observed at sites of TET2 binding (e.g., “AML.ChIP” and “ES.ChIP”) (Fig. 5A, B). Cluster 12 consists of 22 TFs that preferentially bind enhancer-box (E-box) elements (5′-CANNTG-3′) in the genome. E-box motifs often contain a central CpG site, and 5mC, as well as the TET-mediated oxidation products 5hmC and 5fC, in this motif have been shown to impair in vitro DNA binding (Wang et al. 2017; Yin et al. 2017). Analysis of mRNA expression of Cluster 12 TFs in MPP cells revealed that the most abundantly expressed TFs were the MYC/MAX, MAX/MXI1, and HIF1A/ARNT heterodimers, as well as the E proteins ITE2 and ZEB1. Therefore, it is likely that the observed changes in chromatin accessibility originate from altered binding activity of these TFs or other Cluster 12 TF members are lowly expressed or absent (Fig. 5C). None of the abundant TFs were found to be differentially expressed in Tet2 knockout MPP cells, thus suggesting that loss of TET2 results in decreased TF binding activity rather than inducing changes in TF expression (Fig. 5D).

We reasoned that a potential loss of TF binding may result in decreased expression of TF target genes. To analyze the effect of TF
we also observed a decrease in expression of the validated ITF2 target genes, Ccr9 and Dnmt3b, in Tet2 knockout MPPs (Fig. 6B), and analysis of DNA methylation changes in predicted ITF2 binding sites revealed an overall increase in 5mC levels in hematopoietic stem cells upon ablation of TET activity (Fig. 6C).

Mice with hematopoietic-specific knockout of a single allele of If2 display defects in differentiation of plasmacytoid dendritic cells (pDCs) (Cisse et al. 2008). Therefore, we reasoned that a change in ITF2 binding activity should recapitulate this. Phenotypic analysis of aged Tet2 knockout mice revealed a partial impairment of pDC development that is highly similar to previously published observations in If2 heterozygous knockout bone marrow (Fig 6D,E; Cisse et al. 2008). Overall, these results suggest that the knockout of Tet2 phenocopies the haploinsufficiency of If2 in hematopoietic stem cells by decreasing the binding activity of ITF2 rather than changing its expression. This would derail normal pDC differentiation in the bone marrow. Frequent mutations of TET2 have been found in patients with Blastic plasmacytoid dendritic cell neoplasm (BPDCN), a rare myeloid neoplasm characterized by proliferation of aberrant pDC precursor cells (Scourzic et al. 2015). Thus, TET2 may prevent BPDCN by keeping the binding sites of ITF2 accessible and thus maintaining the expression of ITF2-regulated genes.

Within Cluster 12, MYC was by far the most abundantly expressed TF in hematopoietic stem cells. It has previously been reported that DNA methylation inhibits MYCN chromatin binding in Tet1-3 triple knockout ES cells (Yin et al. 2017), and a recent study has uncovered that Tet2 is recruited to MYC binding sites by SNIP1 in cancer cell lines (Chen et al. 2018). However, analysis of chromatin accessibility by diffTF at MYC binding sites (predicted individually from all available MYC DNA-binding motifs) did not show a significant difference in wild-type and Tet2 knockout cells, possibly due to a complex MYC binding pattern at E-boxes throughout the genome. To further analyze if MYC activity could be affected by TET2, we performed gene set enrichment analysis (GSEA) on RNA-seq data from wild-type and Tet2 knockout MPP cells. This analysis revealed a pronounced down-regulation of MYC target gene signatures (Fig. 6F; Supplemental Table S2; Menssen and Hermeking 2002; Zeller et al. 2003). Furthermore, we also found a general down-regulation of gene signatures of translation,
ribozyme biogenesis and metabolism, which are processes known to be controlled by MYC (Fig. 6G; Supplemental Table S2). Thus, transcriptome analysis suggests that MYC activity is indeed inversely affected in Tet2 knockout MPP cells, likely through altered DNA methylation at MYC target sites and decreased MYC TF chromatin binding. In support of this, conditional knockout of Myc has been shown to result in hematopoietic stem cell expansion (Wilson et al. 2004) similar to what has been observed in aged Tet2 knockout animals (Moran-Crusio et al. 2011; Quivoron et al. 2011). Together, these results suggest that impairment of MYC activity may confer a competitive advantage to Tet2 knockout hematopoietic stem cells, and that impairment of other bHLH TFs, such as ITF2, renders terminal differentiation inefficient.

Discussion

Although DNA methylation landscapes in TET2 mutant cells have been extensively characterized (Ko et al. 2010; Asmar et al. 2013; Hon et al. 2014; Rasmussen et al. 2015; Yamazaki et al. 2015), much less is known about the functional consequences that ultimately drive hematopoietic stem cell expansion and malignancy. Here, we resolve the inconsistency between regions reported to be bound by TET2 (Chen et al. 2013; Dephus et al. 2013; Peng et al. 2016) and those affected by TET2 catalytic function. We show that TET2 is predominantly recruited to promoter-distal regions of open chromatin, including enhancers. Therefore, it is likely that it is epigenetic perturbation of these elements that results in dysregulated hematopoeisis of TET2-mutated stem cells, possibly through pleiotropic effects on many genes that together derail homeostasis and differentiation. A recent study has reported a direct biochemical interaction between the histone acetyltransferase EP300 and TET2 (Zhang et al. 2017). In ES cells, we find that approximately half of high-confidence TET2 binding sites colocalize with promoter-distal EP300-enriched regions (Fig. 2B,F), and nearly half of all EP300-enriched regions show evidence of TET2 binding (Supplemental Fig. S2B). Taken together these results may suggest that EP300 is involved in recruiting TET2 to chromatin through direct protein–protein interactions. However, a direct validation of this and a comprehensive analysis of the role of additional potential recruiters, such as WT1 (Rampal et al. 2014; Wang et al. 2015), remain an area of active investigation.

To dissect the impact of TET2 loss in hematopoiesis, we investigated genome-wide changes in enhancer function using chromatin accessibility as a measure of activity in a native chromatin context. We report widespread changes that ultimately drive hematopoiesis of TET2-mutated stem cells, possibly through pleiotropic effects on many genes that together derail homeostasis and differentiation. A recent study has reported a direct biochemical interaction between the histone acetyltransferase EP300 and TET2 (Zhang et al. 2017). In ES cells, we find that approximately half of high-confidence TET2 binding sites colocalize with promoter-distal EP300-enriched regions (Fig. 2B,F), and nearly half of all EP300-enriched regions show evidence of TET2 binding (Supplemental Fig. S2B). Taken together these results may suggest that EP300 is involved in recruiting TET2 to chromatin through direct protein–protein interactions. However, a direct validation of this and a comprehensive analysis of the role of additional potential recruiters, such as WT1 (Rampal et al. 2014; Wang et al. 2015), remain an area of active investigation.

To dissect the impact of TET2 loss in hematopoiesis, we investigated genome-wide changes in enhancer function using chromatin accessibility as a measure of activity in a native chromatin context. We report widespread changes, of which a majority of differentially open regions in all tested cell types, except GMP, were less accessible upon ablation of TET2. This is consistent with a recent study showing that most DNA methylation-sensitive regulatory elements (~86%) are inhibited
by DNA methylation (Lea et al. 2018). To perform an unbiased analysis of TF signatures that were globally enriched or depleted, we developed a novel computational tool, diffTF, to assess differential TF binding from ATAC-seq profiles (Berest et al. 2018). We could detect signatures of aberrant activity of several clusters of hematopoietic TFs such as IRF, C/EBP, GATA, and HOX protein families in different hematopoietic cell types. In addition to these changes, we also found a strong and consistent loss of activity for a cluster of bHLH TFs (including MYC and ITF2) that preferentially bind enhancer-box (E-box) elements in the genome. To examine the molecular and phenotypic consequences of this, we uncovered that (1) MYC and ITF2 target genes were down-regulated in Tet2 knockout hematopoietic stem cells; (2) ITF2 binding sites were DNA hypermethylated upon ablation of TET activity; and (3) Tet2 knockout mice show hallmarks of Mhc-deficient and Itf2-deficient hematopoiesis. Thus, our results support a model by which Tet2 recruitment and catalytic functions are necessary to shape the chromatin structure for permissive TF binding.

It should be noted that although inferring differential TF occupancy from ATAC-seq data provides a global view of the transcriptional network of a cell, it does have limitations: The use is dependent on collections of TF binding motifs of high quality (not applicable to TFs with poorly defined or unknown DNA-binding motifs) as well as the inherent impact of binding of specific TFs on nucleosome density. For instance, it is possible that additional TF binding events that were not identified in this study may also require TET2 activity; however, such TFs may not show strong global differences or are not detected from TF motif analysis of chromatin accessibility alone. Vice versa, due to the high similarity of DNA-binding motifs within certain TF families, the role of a TF identified by this approach should subsequently be characterized to confirm the identity as well as the biological significance.

Frequent mutations in the DNA methylation machinery have been found in patients suffering from hematological diseases as well as solid cancers (Rasmussen and Helin 2016). Yet, whether DNA methylation is causally involved in shaping gene expression patterns, rather than passively mirroring transcriptional states, is still a matter of debate (Schübeler 2015). Here, we present data in support of a role of TET2-mediated cytosine modifications at enhancers to facilitate TF binding and fidelity of target gene expression in hematopoietic cells. Rather than being an obligate activator, TET2 functions to reinforce binding of some TFs and contributes in this manner to enhancer activity and gene expression. Analysis of chromatin maturation after replication fork passage has indicated that open chromatin at cell-type-specific enhancers is only slowly reestablished through competition of TFs with newly deposited nucleosomes (Ramachandran and Henikoff 2016). Thus, cells that naturally divide rapidly and must undergo coordinated cell state transitions (e.g., hematopoietic stem and progenitor cells) may be particularly sensitive to epigenetic disturbances of TF binding kinetics. Comprehensive analysis of enhancer function in TET2-mutated hematopoiesis will foster a greater understanding of the role of epigenetic dysregulation in disease and may lead to discoveries of potential clinical significance.

Methods

ES cell lines

Mouse ES cell lines were derived from blastocysts harvested from Tet2fl/fl animals (Moran-Crusio et al. 2011). To generate a clean Tet2 knockout ES cell line (to use as ChIP control), cells were transiently transfected with a Cre recombinase-expressing plasmid and subjected to identification of a constitutive Tet2 knockout ES cell line. Endogenous tagging of TET2 was performed using CRISPR homology-directed repair with a single-stranded DNA oligonucleotide repair template. Briefly, a sgRNA targeting the insertion site was cloned into pSpCas9(BB)-2A-GFP (PX458) (Addgene #48138). This plasmid was cotransfected in ES cells with oligonucleotides encoding either two copies of Flag (DYKDDDDDKDYKDIDDDKKDKDDDK) or a single copy of V5 tag (GKPIPNL LGLDST) as well as homology arms (60 bp each). Transfected cells were single-cell sorted, and the resulting clones were screened for stable expression of epitope-tagged TET2 by Western blot. All mouse ES cell lines were cultured in feeder-free, gelatinized plates in “Serum/2i/LIF” conditions: Dulbecco’s Modified Eagle Medium (DMEM) supplemented with 15% fetal bovine serum, 2 mM Glutamax (Gibco), 0.1 mM 2-Mercaptoethanol (Gibco), 1× nonessential amino acids (Gibco), 1× Pen/Strep (Gibco), 3 µM GSK3 inhibitor (CHIR99021), 1 µM MEK1 inhibitor (PD0325901), and leukemia inhibitory factor (LIF).

Mice

For analysis of aging-related Tet2-deficient hematopoiesis, cohorts of age-matched litter mates (8-wk old) of Tet2fl/fl or Tet2fl/fl;Mx1-Cre+ mice (Quivoron et al. 2011) were injected three times intraperitoneally with 250 µg polynosinic-polycytidylic acid (PolyI:C, LMW, InvivoGen) at experimental days 0, 2, and 4. The mice were subsequently allowed to age and sacrificed at 10 mo of age. To generate a Tet2 knockout AML model, the following genetically modified mouse lines Npm1cA-Flox (Vassiliou et al. 2011), Flt3-TTD (Lee et al. 2007), and Tet2fl/fl;Mx1-Cre+ (Quivoron et al. 2011) were intercrossed and Npm1cA-Flox;Flt3-TTD;Mx1-Cre+ or Npm1cA-Flox;Flt3-TTD;Tet2fl/fl;Mx1-Cre+ mice were monitored for disease development. AML cells were harvested and 2.5 × 104 c-Kit+Gr1+Mac1− AML splenocytes transplanted into lethally irradiated (650Rad) Ly5.1 recipient animals by tail vein injection. Ly5.1 mice were maintained on medicated water (Ciprofloxacin 100 µg/ml) for 3 wk following the irradiation procedure. To establish in vitro culture of AML cells, c-Kit+Gr1+Mac1− splenocytes harvested from moribund mice were purified by FACS and cultured in suspension in nontissue culture treated plasticware in StemPro-34 SFM media (Thermo Fisher Scientific) supplemented with 2 mM Glutamax (Gibco), 1× Pen/Strep (Gibco), 0.1 mM 2-mercaptoethanol (Sigma-Aldrich), as well as the cytokines SCF (50 ng/ml), IL3 (10 ng/ml), and IL6 (10 ng/ml) (Peprotech). All animal work was carried out in compliance with ethical regulation under license by the Danish regulatory authority.

Chromatin immunoprecipitation and sequencing

ChIP-seq experiments were carried out using a two-step crosslinking procedure with 2 mM disuccinimidyl glutarate (DSG) and 1% formaldehyde (FA). The genomic coordinates of all region sets derived from TET2 ChIP analysis in various cell types can be found in Supplemental Table S3. See Supplemental Methods for a full description of experimental protocol, processing, and analysis of ChIP data sets.

ATAC-seq library generation and analysis

ATAC-seq libraries were generated and sequenced as described previously (Buenrostro et al. 2013; Lara-Astiaso et al. 2014). See Supplemental Methods for a description of experimental protocol and data preprocessing steps.
diffTF analysis

In order to define potential difference in TF activity based on the chromatin accessibility (ATAC-seq) data, we developed a computational pipeline called diffTF (deposited at https://git.embl.de/grp-zaugg/diffTF; documented at https://diffTF.readthedocs.io) (Berest et al. 2018). This method takes as input predicted TF binding sites defined by genome-wide PWM scanner (PWMscan) (Ambrosini et al. 2018) overlapped with the open regions from ATAC-seq. For the PWM database source, we used HOCOMOCO v10 database (Kulakovskiy et al. 2016), in which we replaced CDX4 and EVX1 motifs (that were retracted from this database) with PWM models from the recent methyl-SELEX study (Yin et al. 2017). We then extended TF binding sites by 100 bp in each direction and calculated a fold change of chromatin accessibility between two conditions. Finally, we removed potential CG-content bias from the data and used an analytical approach for significance assessment (Supplemental Methods). The final weighted mean difference value for each TF corresponds to the genome-wide difference in accessible DNA. This tool was used in the recent methyl-SELEX study (Yin et al. 2017). We then extended TF binding sites by 100 bp in each direction and calculated a fold change of chromatin accessibility between two conditions. Finally, we removed potential CG-content bias from the data and used an analytical approach for significance assessment (Supplemental Methods). The final weighted mean difference value for each TF corresponds to the genome-wide difference in accessible DNA.

External data sets

The following experimentally determined data sets were downloaded from the ENCODE Project Database (ENCODE; https://www.encodeproject.org/): DNase-seq “ENCFF001YOK” and “ENCFF001UPU”, EP300 binding sites “ENCFF001YAD”, H3K27 acetylation “ENCFF001XWY”, H3K4 monomethylation “ENCFF001XWY”, and CTCF binding sites “ENCFF001YAC”. Active and poised enhancers in myeloid hematopoietic cells were obtained from the NCBI Gene Expression Omnibus (GEO) under accession number GSE59591. DNA methylation data from wild-type and Tet2 knockout cells were obtained from GEO: GSE48519 and GSE72630.

Data access

The ChIP-seq, ATAC-seq, and RNA-seq data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE115972. The source code for the diffTF software used in this study is available as Supplemental Code and at https://git.embl.de/grp-zaugg/diffTF.

Acknowledgments

We thank members of the Helin laboratory for advice and discussion. K.D.R. was supported by a postdoctoral fellowship from the Danish Medical Research Council (FSS 1333-00120B), K.N. was supported by Program for Advancing Strategic International Networks to Accelerate the Circulation of Talented Researchers, JSPS (S2704), and L.S.-C. was supported by a Marie Skłodowska-Curie individual fellowship (Horizon 2020 Framework Programme, grant agreement no. H2020-MSCA-IF-2017-796341). The work was supported by grants to K.H. from The European Research Council (294666, DNA MET), the Danish Cancer Society, the Danish National Research Foundation (DNRF82), and through a center grant from the Novo Nordisk Foundation (NNF17CC0027852). This work was also supported by the National Institutes of Health (NIH, P30 CA008748).

Author contributions: K.D.R. designed the study, performed the majority of experiments, analyzed data, and wrote the first draft of the manuscript. J.B. and J.B.Z. provided genomic data analysis and interpretation (ATAC-seq, RNA-seq, ChIP-seq) and designed the diffTF software. S.K., K.N., L.S.-C., and M.T.P. performed experiments. G.S.V. and J.C. provided essential reagents and intellectual input. K.H. and J.B.Z. designed and supervised the study, analyzed data, secured funding, and wrote the manuscript with K.D.R.

References


Received May 7, 2018; accepted in revised form February 19, 2019.
TET2 binding to enhancers facilitates transcription factor recruitment in hematopoietic cells

Kasper D. Rasmussen, Ivan Berest, Sandra Kebler, et al.

Genome Res. 2019 29: 564-575 originally published online February 22, 2019
Access the most recent version at doi:10.1101/gr.239277.118

Supplemental Material
http://genome.cshlp.org/content/suppl/2019/03/15/gr.239277.118.DC1

References
http://genome.cshlp.org/content/29/4/564.full.html#ref-list-1
This article cites 67 articles, 19 of which can be accessed free at:

Creative Commons License
This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see http://genome.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.