Supplementary Information

TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS

Rachel Deplus[1,#] Benjamin Delatte[1,#], Marie K. Schwinn[2,#], Matthieu Defrance[1], Jacqui Méndez[2], Nancy Murphy[2], Mark A. Dawson[3,4], Michael Volkmar[1], Pascale Putmans[1], Emilie Calonne[1], Olivier Bernard[5], Thomas Mercher[5], Eric Solary[6], Marjeta Urh[2], Danette L. Daniels[2,*] & François Fuks[1,*]

Supplementary Information is divided into two parts (excluding Supplementary Tables and References): Supplementary Methods and Supplementary Figures. The later is listed according to the Main Text and the Main Figures to which they are related.

Contents

1. Supplementary Methods

1.1 Expression plasmids. Clones expressing N-Terminal HaloTag fusions of human TET2 (NM_00127208), N-terminal HaloTag fusion of human TET2ΔCD (NM_017628.4), a transcript variant lacking the catalytic domain, TET3 (NM_144993), OGT (NM_181673), SETD1A (NM_014712)), and WDR82 (NM_025222) were obtained from Kazusa DNA Research Institute (Kisarazu, Japan) as pFN21A HaloTag CMV Flexi Vectors (Promega). HaloTag Control (ADN27525.1) vector (Promega) was used for expression of the HaloTag protein alone. For BRET experiments, Histone H3.3 (NM_002107) was synthesized by IDT and cloned into pFC14K HaloTag CMV Flexi vector (Promega) and SETD1A (NM_014712) was subcloned into an N-terminal NanoLuc expression vector pFCMV_N-NLuc_G418 (Promega). For expression in *E. coli*, human full-length OGT and the TET2 catalytic domain (amino acids 1362-1905 of full-length human TET2) were sub-cloned into the pFN18K HaloTag. pcDNA3.1 TET2 CD Flag and pcDNA3.1 TET3 CD Flag were already described (Ito et al, 2010).

1.2 Cell culture, transfections and drug treatments. HEK293T cells (ATCC #CRL-11268) were maintained in Dubelcco's modified Eagle's medium (DMEM, Gibco BRL) supplemented with 10% foetal bovine serum and grown at 37°C under 5% CO_2. Transient transfections were performed as described (Vire et al, 2006). The cells were treated with Alloxan 5mM (A6316, Sigma) or PUGNAc 150µM (132489-69-1, Carbosynth) for 24h.

1.3 Mass spectrometry analysis. Purified complexes from the HT mammalian pull-down assays were analyzed and processed by MS Bioworks, LLC (Ann Arbor, Michigan). The samples were separated on a SDS-PAGE gel which was subsequently

Coomassie stained and cut into 10 fragments. Each gel piece was processed with the Progest Protein Digestion Station (Digilab). Briefly, gel slices were washed using 25 mM ammonium bicarbonate and acetonitrile, followed by reduction with 10 mM dithiothreitol, and alkylation with 50 mM iodoacetamide. Proteins were digested with trypsin (Promega) for 4h and digestion was quenched with formic acid. Gel digests were analyzed directly by nano LC/MS/MS with a NanoAcquity HPLC (Waters) interfaced with an Orbitrap Velos Pro (Thermo Scientific) tandem mass spectrometer. Digested peptides were loaded on a trapping column and eluted over a 75 µm analytical column packed with Jupiter Proteo Resin (Phenomenex) at 350nl/min. The mass spectrometer was operated in data-dependent mode, with MS performed in the Orbitrap at 60.000 full width at half maximum (FWHM) resolution, and MS/MS performed in the LTQ. The 15 most abundant ions were selected for MS/MS. The data were searched with Mascot (Matrix Science) against the concatenated forward/decoy UniProt Human Database, and Mascot DAT files were visualized and filtered by Scaffold (Proteome Software). Data were filtered using a minimum protein value of 90%, a minimum peptide value of 50% (Protein and Peptide Prophet scores), and required at least two unique peptides per protein. Spectral counting was performed and normalized spectral abundance factors determined (Sardiu et al, 2008). Data were reported at less than 1% false discovery rate (FDR) at the protein level based on counting the number of forward and decoy matches. GlcNAcylated peptides were identified by an unbiased search of mass spectrometry data for HexNAc as a variable modification. Any spectra arising from HexNAC modified peptides were shown as a match in the search and associated peptides determined. These spectral are dominated by neutral loss, helping to determine if the match is correct.

1.4 Immunoprecipitations and western blot analysis. Standard procedures were used for coimmunoprecipitations and Western blotting (Vire et al, 2006). For endogenous

immunoprecipitations, antibodies were incubated with HeLa nuclear extracts (CILBiotech) in IPH buffer at 4°C overnight. Antibodies against OGT (ab50270; Abcam), and IgG (sc-2027; Santa Cruz) were used for immunoprecipitation and anti TET2 (ab94580; Abcam) anti TET3 (ab94577; Abcam), anti Flag (F3165, Sigma), anti HaloTag (G9211, Promega), anti H3 (Abcam, ab 1791), anti H3K4me3 (39159, Active Motif), anti H3K36me3 (Diagenode: pAb-058-050), anti H3K9me3 (Diagenode: pAb-056-050), anti Tubulin (ab7291, Abcam), anti beta-actin (Abcam ab8226), and anti HDAC1 (pAb-053-050, Diagenode) for western blot. The global level of GlcNAcylation was evaluated by western blotting using an anti *O*-GlcNAc antibody (MMS-248R; Covance).

1.5 Purification of OGT in *E. coli.* KRX competent Cells (Promega) were transformed with the HT-TET2CD and HT0OGT vectors and grown in 5 ml TB medium supplemented with 25 μg/μl kanamycin at 37^{o}C overnight. The overnight culture was diluted 1:100 in 50ml TB and induced with 0.05% glucose and 0.05% rhamnose for 20 h at 25^{o}C. Cells were harvested by centrifugation at 4,000 x g for 20 min at 4^{o}C and then resuspended in purification buffer (50 mM Tris pH 7.5, 150 mM NaCl, 10% glycerol, and 1μM DTT) containing 10 μg/ml lysozyme, RQ1 RNase-Free DNase, and Protease Inhibitor Cocktail (Promega). Following a 10 min incubation on ice, cells were sonicated with the Model 705 Sonic Dismembrator (Fisher Scientific) equipped with a 1/16" microtip and set at an amplitude of 12 for 60s. Lysate was clarified by centrifugation at 10,000 x g for 30 min at 4^{o}C and then incubated with HaloLink Resin (Promega) that had been pre-equilibrated in purification buffer for 1 h at 22^{o}C with rotation. Following three washes in purification buffer, OGT was cleaved from the resin with TEV Protease (Promega) diluted in purification buffer for 1 h at 22^{o}C with rotation. Resin was centrifuged at 2000 x g for 5 min, and the supernatant containing the cleaved OGT was incubated with HisLink Resin (Promega) for 20 min at room

temp to remove the TEV protease. Resin was centrifuged at 1,000 x g for 5 min, and the supernatant containing the purified OGT was collected.

1.6 *In Vitro* Complex Isolation. Bacterial lysates containing HT-TET2CD or HT alone (prepared as described for OGT) were incubated with HaloLink Resin that had been pre-equilibrated in resin wash buffer (TBS and 0.05% IGEPAL CA-640 (Sigma)) for 1 h at 22^{o}C. Resin was then washed 5 times in wash buffer and then incubated with purified OGT (prepared as described above) for 1 h. Resin was then washed 10 times with wash buffer and incubated with SDS elution buffer (50 mM Tris-HCl, pH 7.5, and 1% SDS) at 70^{o}C for 5 min. Samples were analyzed for OGT by Western blotting using anti-OGT antibody.

1.7 Dot Blot for 5-hydroxymethylcytosine quantification. Isolated genomic DNA (50ng/µl) was incubated at 95°C for 15min and immediately chilled on ice to obtain single stranded DNA. 75ng (1.5µl) of DNA was spotted in duplicate onto Nylon membrane (GE Healthcare Hybond-N+). After drying the membrane, cross-linking was performed 2 times with 200 000µJ/cm^2 UV. The membrane was blocked in 10 %(w/v) non-fat dry milk + 1%(w/v) BSA in PBS+0.1% tween-20 overnight at 4°C. The membrane was then transferred into blocking solution supplemented with 1:1000 of anti-5hmC polyclonal antibody from rabbit (Active Motif) validated for specificity(Ito et al, 2010; Nestor et al) (Supplementary Fig.4C) or 1/100 of anti-ssDNA antibody from human (LSBio) and incubated 1h at room temperature. Membrane was washed 3 times with PBS+0.1% tween-20 for 10min and then transferred into blocking)solution supplemented with 1:5000 of anti-rabbit isotype linked to HRP (GE Healthcare) or 1:10 000 anti-human isotype linked to HRP (Jackson ImmunoResearch) and incubated 1h at room temperature, followed by 3 washing steps with PBS+0.1% tween-20. Peroxidase activity was detected by ECL (Western Lightning Plus-

ECL, Perkin-Elmer) and signal detection was done with the LAS-3000 imager (Fujifilm). For spot signal quantification, AIDA image analyzer was used. Membrane stripping was performed in PBS supplemented of 20% SDS and 1.5% β-mercaptoethanol, 30 minutes at 50°C, then washed extensively with PBS+0.1% tween-20.

1.8 RNA Interference, retroviral infection and RT-PCR analysis. Interfering RNA for TET2 and TET3 were generated as previously described (Vire et al, 2006). Briefly, the target sequence used to silence TET2 and TET3 was inserted as a short hairpin into the pRetroSuper (pRS) retroviral vector according to the manufacturer's recommendations (OligoEngine) to form RNAi TET2 and RNAi TET3. Retrovirus production by HEK293GP cells and infection of target cells were performed as described(Vire et al, 2006). Infected cells were selected with 1.5µg/ml puromycin (Sigma). For double knockdown TET2/3 cells, a second round of infection was performed. For this, we used the pRetroSuper neoGFP (pRSneoGFP) according to the manufacturer's recommendations (OligoEngine). Infected cells were then reselected with G418 at the concentration of 750µg/ml (Sigma). RNA purification and RT-PCR analysis were performed as described previously (Vire et al, 2006).

For Figure 1H, quantitative RT-PCR analysis was performed in the double TET2/3 kd (RNAi TET2/3) and the double RNAi Control (pRS empty and pRS neoGFP empty) or OGT kd cells (RNAi OGT) and its related RNAi Control (pRS empty). The results were normalized relative to the following 5 housekeeping genes: *SDHA, GAPDH, ACTIN, HPRT and rRNA 18S*. Expression of each RNAi control was set to a value of 1. Error bars represent standard deviations of two independent biological experiments. Primer sequences are listed in Table S1.

1.9 Radioactive GlcNAc assays. OGT activity measurements from nuclear lysate were performed as described previously with minor modifications (Song et al, 2008). 120µg of nuclear extracts were added into a 40µl reaction mixture containing 1mM HPLC purified CKII peptide (PGGSTPVSSANMM), 1µCi of UDP-[^{3}H]GlcNac (NET 434250UC, Perkin Elmer), 50mM Tris-HCl pH 7.4, 1mM dithiothreitol and 12.5mM MgCl2. The reaction was incubed at 30°C for 90min at 220rpm and stopped by the addition of 450µl of 50mM formic acid to the supernatant. The resulting reaction mixture was loaded onto a Sephedax columns (PD midi trap G10 28-9180-11, GE healthcare) equilibrated in 50mM formic acid. The columns were washed with 9ml of formic acid and eluted with 2ml of NaCl 0.5M. Incorporation of [^{3}H]GlcNAc into the peptides was quantitated using scintillation counting.

1.10 HaloCHIP. 8 x 10^5 HEK293T cells were plated in each well of a 6-well plate. At 70-80% confluency, cells were transfected for 24h with 2µg of HaloTag fusion constructs (experimental samples) or HaloTag control (which was used as the reference to discard non-specific peaks. This "HaloTag only" control gave essentially background results; data not shown). Both experimental and control cells were crosslinked with 0.75% formaldehyde (Sigma) for 10min at 22^{o}C, quenched with 0.125M glycine for 5min, and harvested with a cell scraper. Lysates were generated by incubating cells in mammalian lysis buffer on ice for 15min followed by homogenization with a syringe. To obtain DNA fragments 500-1500bp in size, lysates were sonicated using the Model 705 Sonic Dismembrator (Fisher Scientific) equipped with a 1/16" microtip and set at an amplitude of 12 with a total run time of 120s. Sonicated lysate was clarified by centrifugation at 14.000 x g for 5min and then incubated with HaloLink Resin for 2h at 22^{o}C with rotation. Resin was then washed once with mammalian lysis buffer, twice with nuclease free water, once with high salt wash buffer (50mM Tris-HCl, pH 7.5, 700mM NaCl, 1% Triton X-100, 0.1% sodium deoxycholate, and

5mM EDTA), and then three more times with nuclease free water. To reverse the crosslinks and release the DNA, the resin was incubated at 65°C for 16hrs in reversal buffer (10mM Tris-HCl, pH 8.0, 1mM EDTA, and 300mM NaCl). Isolated DNA fragments were purified using the QIAquick PCR Purification Kit (Qiagen) and eluted in 100µl nuclease free water.

1.11 Methyl-DNA Immunoprecipitation (MeDIP). 5µg of genomic DNA was diluted in 150µl of 1X TE buffer, then sonicated using bioruptor (Diagenode) in cold water with the following settings: 2 cycles of 10 minutes, sonication strength set on high, with intervals of 30 seconds ON/OFF. DNA was then quantified by fluorometry using the Qubit 2.0 (Invitrogen) and 4µg of sonicated DNA was diluted in 450µl of 1X TE buffer. In order to increase pull down efficiency, DNA was denatured 10 minutes at 95°C, then chilled 10 minutes on ice + water. 50µl of 10X IP buffer (100mM sodium phosphate, pH 7.0; 1,4M NaCl ; 0,5% Triton X-100) was added and 10% of DNA was used as input and store at 4°C. 8,5µg of mouse monoclonal antibody for 5-methylcytosine (MAb-081-100, Diagenode) or mouse monoclonal antibody for GAL4 (sc-510, Santa Cruz) were added to DNA and incubated 2h at 4°C on a rotating wheel. 40µl of anti-mouse IgG magnetic beads (Dynabeads, Invitrogen) were blocked twice over 5 minutes with 800µl of 1X PBS-BSA 0,1%, then were resuspended in 40µl of 1X IP buffer and added to the immune complexes overnight at 4°C. Beads were washed 3 times 10 minutes with 700µl of 1X IP buffer and resuspended in 250µl of digestion buffer (50mM Tris-HCl pH 8.0 ; 10mM EDTA ; 0,5% SDS). 7µl of proteinase K (10mg/ml) were added and beads were incubated 3h at 50°C under vigorous agitation. Proteinase K was inactivated 10 minutes at 80°C, then eluted DNA was purified using QIAquick PCR purification kit (Qiagen). 3µl of enriched fragmented DNA or 3µl of input was subjected to 40 cycles of quantitative PCR as described above. Primer sequences are listed in Table S1.

1.12 Hydroxymethyl-DNA Immunoprecipitation (hMeDIP). 5µg of genomic DNA was diluted in 150µl of 1X TE buffer, then sonicated using bioruptor (Diagenode) in cold water with the following settings: 2 cycles of 10 minutes, sonication strength set on high, with intervals of 30 seconds ON/OFF. DNA was then quantified by fluorometry using the Qubit 2.0 (Invitrogen) and 1,1µg sonicated DNA was engaged in hMedIP reaction following the manufacturer's protocol (hMedIP kit, Active Motif). After immunoprecipitation and washing steps, DNA was purified using QIAquick PCR purification kit (Qiagen). 3µl of enriched fragmented DNA or 3µl of input was subjected to 40 cycles of quantitative PCR as described above. Primer sequences are described in Table S1.

1.13 Library preparation and deep sequencing workflow. The library preparation was performed using ChIP-Seq DNA sample prep kit (Illumina). Briefly, 10ng of qPCR verified ChIPed DNA was subjected to 5' and 3' protruding ends repair, then, non-templated adenines were added to the 3' ends of the blunted DNA fragments. This last step allows SR (Single Read) Illumina's adapters ligation. DNA fragments were then size selected with a 2% agarose gel electrophoresis in order to remove all unligated adapters and to sequence 300bp fragments. 18 cycles of PCR were done to amplify the library. DNA was then quantified by fluorometry using the Qubit 2.0 and its integrity was assessed with 2100 bioanalyzer (Agilent). 3,5pM DNA library spiked with 1% PhiX viral DNA was clustered on cBot (Illumina), and then sequenced on HiScanSQ module (Illumina).

1.14 ChIP-Seq data analysis. **Reads mapping**. Sequencing reads were mapped to the human genome (NCBI Build 36 / UCSC hg18) for the human samples and to the mouse genome (NCBI Build 37 / UCSC mm9) for the mice samples using the BWA software(Li & Durbin, 2009). Reads not uniquely mapped to the reference genome were discarded [see

Supplementary Fig. 7 for details about sequencing statistics]. Raw data were submitted to the Gene Expression Omnibus [http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=ljypfsuuwwiqmvk&acc=GSE36620] database. **Peak calling**. Enriched regions (peaks) were identified using the MACS software (version 1.4)(Zhang et al, 2008) with the default parameters and a P-value cut-off $<10^{-10}$. Peaks were then filtered to avoid significant overlaps (< 80%) with repetitive regions using the UCSC repeat marker tracks [see Table S2 and Table S3 for the full peak data]. For HaloTag ChIP-Seqs (HT-TET2, HT-TET3 and HT-OGT) peaks were filtered by discarding regions showing a statistically significant enrichment (P-value < 10^{-5}) in the HaloTag control ChIP-Seq. **Peak annotation**. Genomic location classes were determined by computing the relative peak density in RefSeq promoter regions (2,000 bp upstream the TSS), in gene body regions and intergenic regions. CpG island relations were evaluated by computing the relative peak density in UCSC annotated CpG islands (CGI), in 2,000 bp regions surrounding CpG islands (Shores) and in the rest of the genome (Distant). Promoters were categorized according to their CpG content and GC ratio in three categories (high, intermediate and low CpG promoters) as defined in Weber *et al*. (Weber et al, 2007). RefSeq promoters (2,000 bp upstream the TSS) were annotated as HCPs (high-CpG promoters) if at least one 500bp window with CpG ratio > 0.75 and a GC content > 0.55 was found or as LCPs (Low CpG promoters) if no 500bp window reached a CpG ratio of 0.48 and as ICPs otherwise. Distance to the TSS was evaluated by computing the distance between peak summits and RefSeq annotated TSSs. **Venn diagrams.** Peak occurring at overlapping genomic location were considered as overlapping and reported in the intersection of the Venn diagrams. For the case of enriched promoters, peaks were considered as overlapping when a co-occurrence was detected in the 1kb region around the TSS. Statistical significance of the overlap was estimated by computing the P-value as $P(X \geq k) = \text{Sum}(k = x \text{ to } m)\ C(m, k)\ C(N - m, n - k) /$

C(N, n) where x is the number of overlapping peaks, n the number of peaks in the first set, m the peaks in the second set and N the total number of allowed peaks that was lower bound by the total number of RefSeq genes (N ≥ Count(RefSeq genes)) (Plaisier et al, ; Robinson et al, 2002). **Heatmaps**. Heatmaps were generated by plotting for all the annotated RefSeq genes the peak significance ($-\log_{10}$ P-value) as a colour gradient at their relative location using the annotated gene TSS as the reference. For all heatmaps, genes were first ordered using the average peak significance in the 1kb region around the TSS for one selected ChIP-Seq (OGT in the case of HEK293T cells and Tet2 in the case of mouse bone marrow experiments). **Encode Data**. ChIP-Seq data for Pol II, H3K4me1 and CTCF were retrieved from the ENCODE project. **Differential analysis**. Decrease in ChIP-Seq signal for the RNAi TET2 in HEK293T cells and Tet2 Ko in mouse bone marrow tissue were evaluated by comparing the statistical significance (peak P-value) and the normalized read count at enriched region in the control for both samples (RNAi TET2 versus RNAi Control and Tet2 Wt versus Tet2 Ko). Decrease in H3K4me3 signal in RNAi TET2 sample vs RNAi control sample was evaluated by computing in each TET2 TET3 OGT enriched regions the enrichment of H3K4me3 (normalized read count) for both samples. Statistical significance was assessed using a Fisher's exact test with a FDR cut-off < 5%. Decrease in H3K4me3 signal in Bone Marrow Tet2 Ko sample versus Bone Marrow Wt was evaluated by comparing the enrichment (peak P-value) at Tet2 *O*-GlcNAc H3K4me3 enriched regions found in the Wt sample. Statistical significance of the decrease was computed using a Student's t-test. All statistics and plotting were performed using the R statistical software.

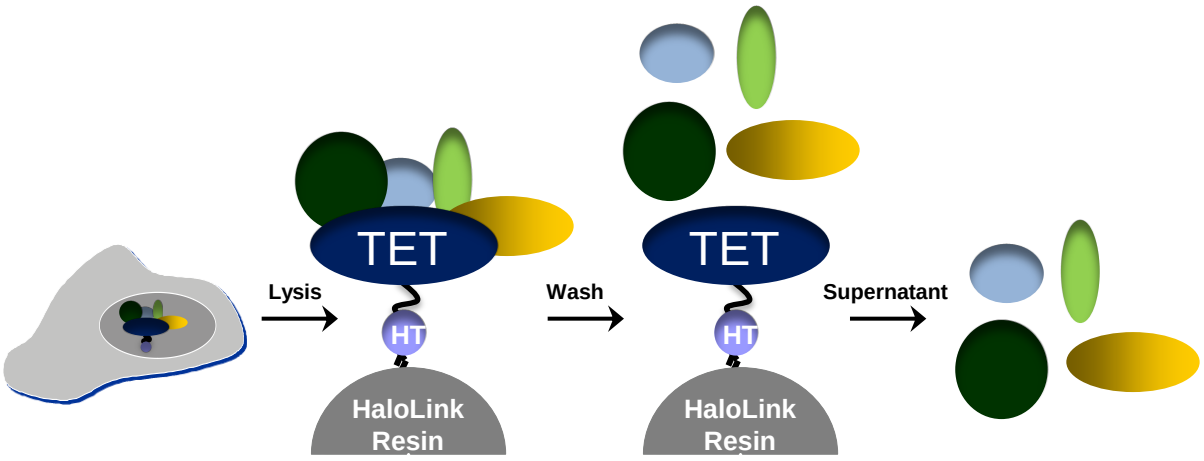
1.15 Mouse tissue preparation. Whole bone marrow cells from wild-type 8-10 weeks-old C57BL/6 mice were collected in PBS 1x supplemented with 2% FBS and subjected to red blood cell lysis prior to molecular analyses. Approval for the use of animals in this

study was granted by the Institut Gustave Roussy Institutional Animal Care and Use Committee. Tet2 Ko mice tissues were received from Ross L. Levine (Memorial Sloan-Kettering Cancer Center, New York) and the generation of Tet2-deficient mice was previously described (Moran-Crusio et al, 2011).

1.16 Bioluminescence Resonance Energy Transfer (BRET assays). HEK293T or RNAi TET2/3 cells (8×10^5) were plated in each well of a 6-well plate and co-transfected with 1.98 μg of Histone H3.3-HaloTag and 0.02 μg of NanoLuc-SETD1A vectors. 24h post-transfection, 1×10^4 cells were re-plated in a 96-well assay plate (Corning Costar #3917) and media exchanged into phenol-red free DMEM, 10% FBS, in the absence (control BRET) or the presence of 100nM fluorescent TOM HaloTag ligand (Promega) for experimental samples. Cells are incubated for 18-24 h at 37^{o}C and 5% CO2 and then NanoLuc furimazine substrate (Promega) was added to both control and experimental samples to a final concentration of 20 μM. Readings were performed within 5 minutes using the Varioskan Flash spectral scanning multimode reader (Thermo Scientific) equipped with 450/80 nm bandpass and 610 nm longpass filters. A corrected BRET ratio was calculated and is defined as the ratio of the emission at 610 nm/450 nm for experimental samples (i.e. those treated with fluorescent TOM ligand) subtracted by and the emission at 610 nm/450 nm for control samples (not treated with TOM ligand). BRET ratios are expressed as milliBRET units (mBU), where 1 mBU corresponds to the corrected BRET ratio multiplied by 1000. (Pfleger et al, 2006; Xu et al, 1999)

1.17 RNA Pol II ChIP-Seq and RNA-Seq analysis. RNA Pol II ChIP-Seq and RNA-Seq data for the HEK293T studies were retrieved from the GEO database (GSM301566 and GSM301568)[23] and mapped to the human reference genome (NCBI Build 36 / UCSC hg18).

Normalized read counts (expressed in reads per kilobase per million reads mapped) were computed for all 1kb regions around RefSeq annotated TSSs for the RNA Pol II data and in RefSeq coding exons for the RNA-Seq data. Read counts are reported for genes with and without TET2-TET3-OGT-H3K4me3 peaks around the TSSs (+/-1 kb regions). RNA Pol II ChIP-Seq and RNA-Seq data for the bone marrow tissue studies were retrieved from the UCSC ENCODE project database (wgEncodeEM001688 and wgEncodeEM001706)[26], mapped to the mouse reference genome mm9 and analyzed using the previously described protocol.



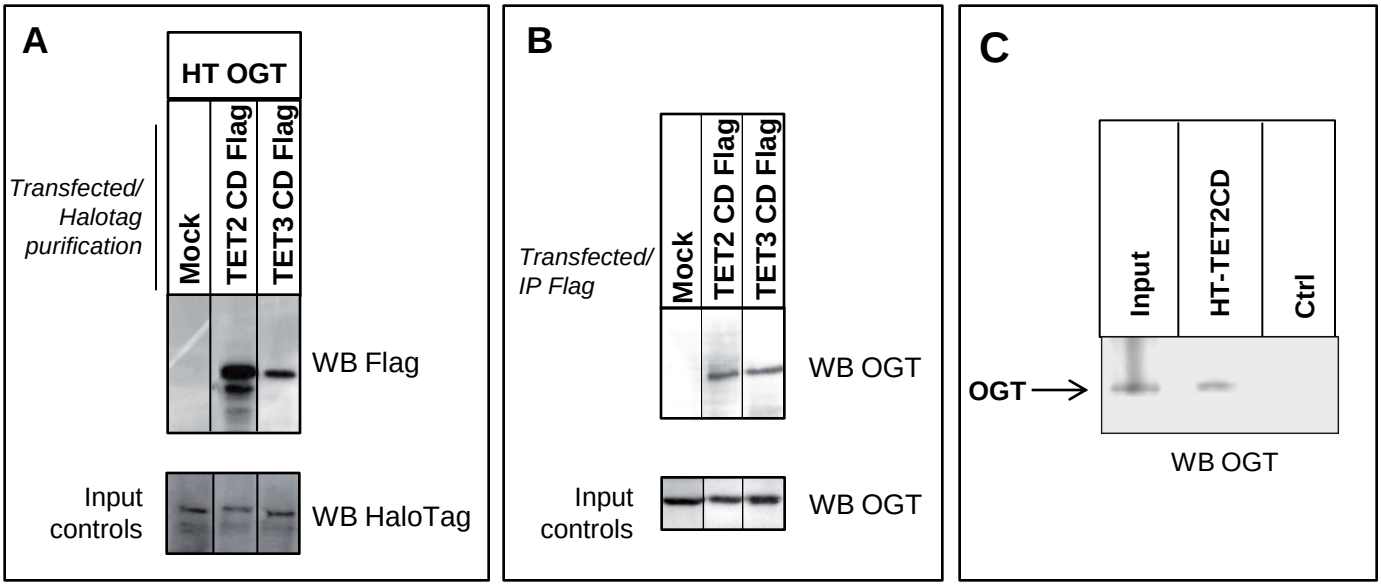
Supplementary Fig. 1 (related to Figs. 1-3): Schematic of protein complex isolation using HaloTag fusions. HaloTag fusion proteins, shown here with TET as an example, are either transiently or stably expressed in mammalian cells. Cells are lysed and fusion proteins are covalently captured directly from a lysate on HaloTag affinity resin, HaloLink. Captured complexes are washed gently on the resin, and interacting partners are eluted with SDS buffer. Due to the covalent binding of HaloTag fusion proteins to its resin, they remain attached to the resin after elution. Interacting partners can be analyzed in a variety of methods, including SDS-PAGE, Western blotting, or mass spectromentry.

Protein Fusion:

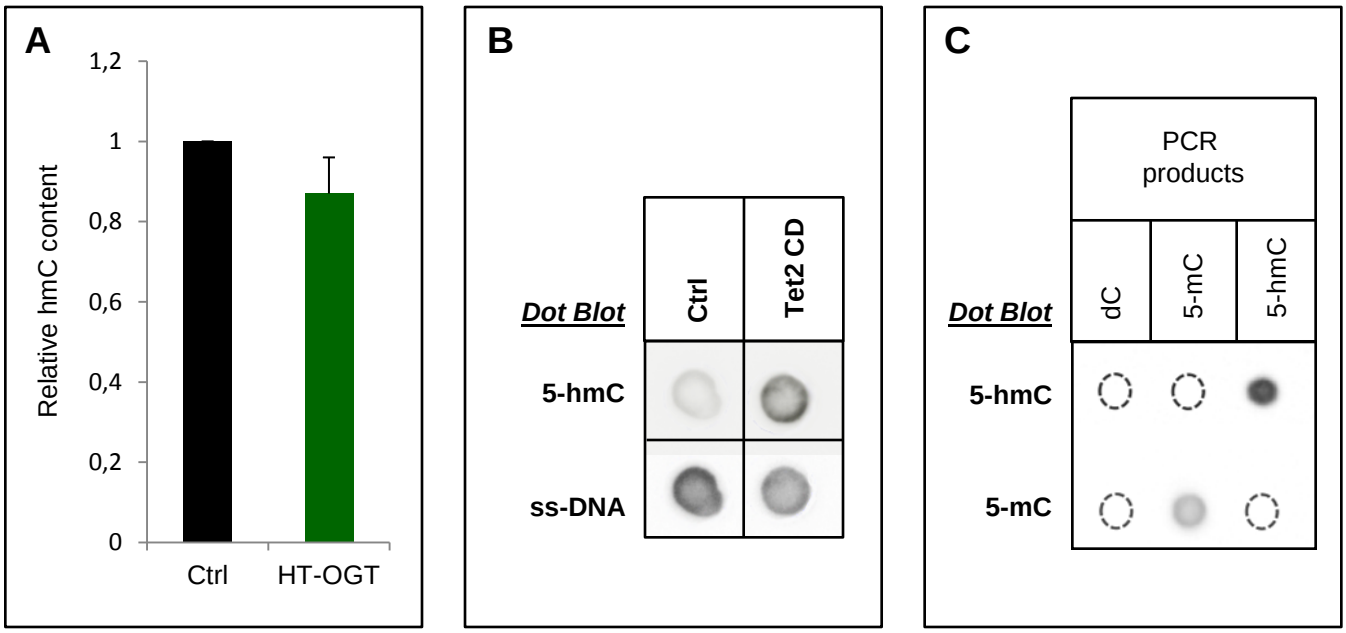
Interactors	TET2ΔCD		Ctrl	
	#1	#2	#1	#2
OGT	3	3	0	0
HCF1	2	0	0	0
YWHAH	18	9	0	0
YWHAG	20	12	0	0
YWHAQ	13	8	0	0
BRG1	6	2	0	2
SMARCC2	0	0	0	0

HT-TET2ΔCD partners purification (Spectral counts)

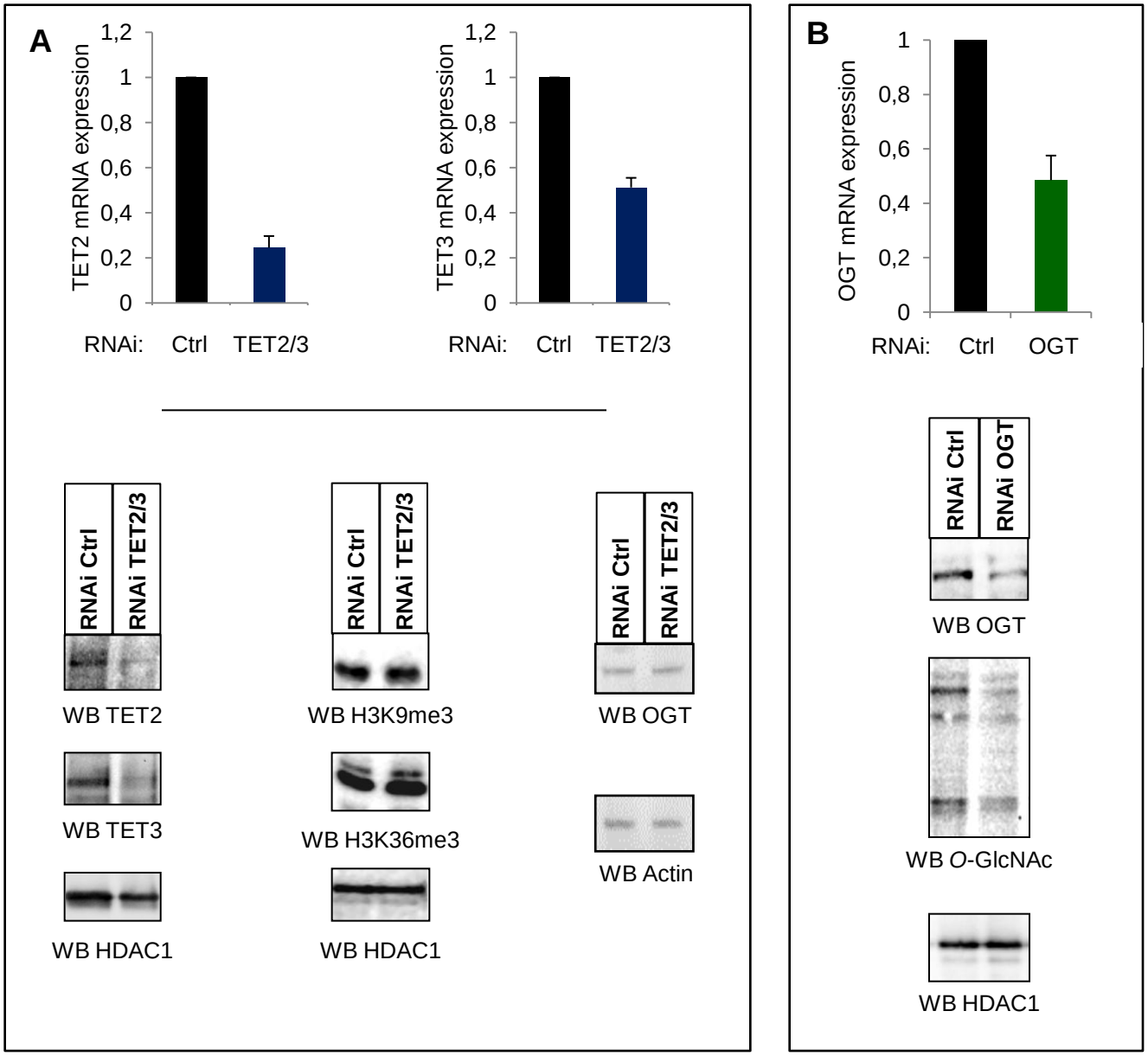
Supplementary Fig. 2 (related to Fig. 1B): HT-TET2ΔCD pulldown mass spectrometry analysis showing signficiant loss of interaction with OGT as compared to full-length TET2. A variant of TET2 lacking the catalytic domain (TET2ΔCD) shows minimal interaction with OGT. Reduced interactions are also seen with HCF1, BRG1, and SMARCC2, but other interactors remain relatively unchanged. Biological duplicates are shown and spectral counts for each protein are listed.



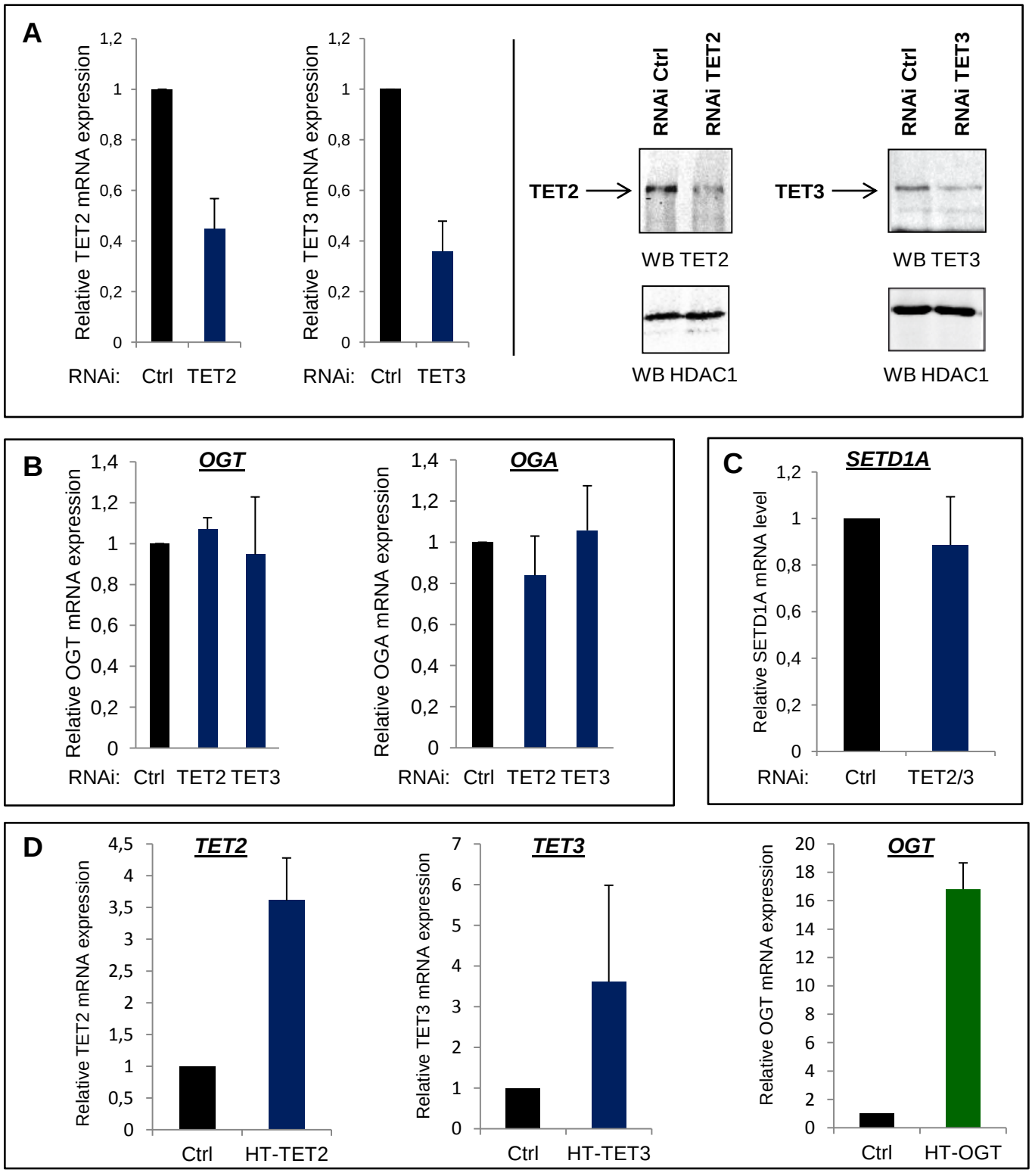
Supplementary Fig. 3 (related to Fig. 1C): TET2 and TET3 catalytic domains alone interact with OGT domains and TET2 directly associates with OGT. (A) HEK293T cells were transfected with the indicated plasmids and the HaloLink resin was used to isolate HaloTag-OGT and interacting proteins. To detect the presence of TET2CD-FLAG or TET3CD-FLAG, an anti-Flag antibody was used in Western blot analysis. Anti-HaloTag western blots were performed as input loading controls to measure HaloTag-OGT. **(B)** Following transfection of the indicated constructs, nuclear cell extracts were precipitated with the anti-Flag antibody and the presence of endogenous OGT was detected with anti-OGT antibody. Endogenous OGT was used as a input loading control. **(C)** Direct interaction of the TET2CD with OGT. *In vitro* interaction of *E. coli* purified OGT and HT-TET2CD. Full-length human OGT was expressed and purified from *E. coli,* then subjected to resin containing covalently bound HT-TET2CD or HT alone (HT-Ctrl) also expressed in *E. coli*. After extensive washing and SDS elution, Western blotting of the resulting supernatant was performed with an anti-OGT antibody.



Supplementary Fig. 4 (related to Fig. 1F): OGT does not affect hmC levels, while TET2CD does in Dot Blot assay with a specific 5-hmC antibody. (A) Overexpression of OGT followed by Dot blot to detect 5-hmC activity does not impact hmC, similar to what was observed with either alloxan or PUGNAc treatment (Fig 1d). **(B), (C),** The lack of effect of Alloxan or PUGNAc on 5-hmC activity is not due to a technical failure of Dot blot assays: **(B**) increased 5-hmC activity is observed in cells overexpressing TET2 catalytic domains (CD) and **(C),** the 5-hmC antibody (from Active Motif) is specific for 5-hmC.



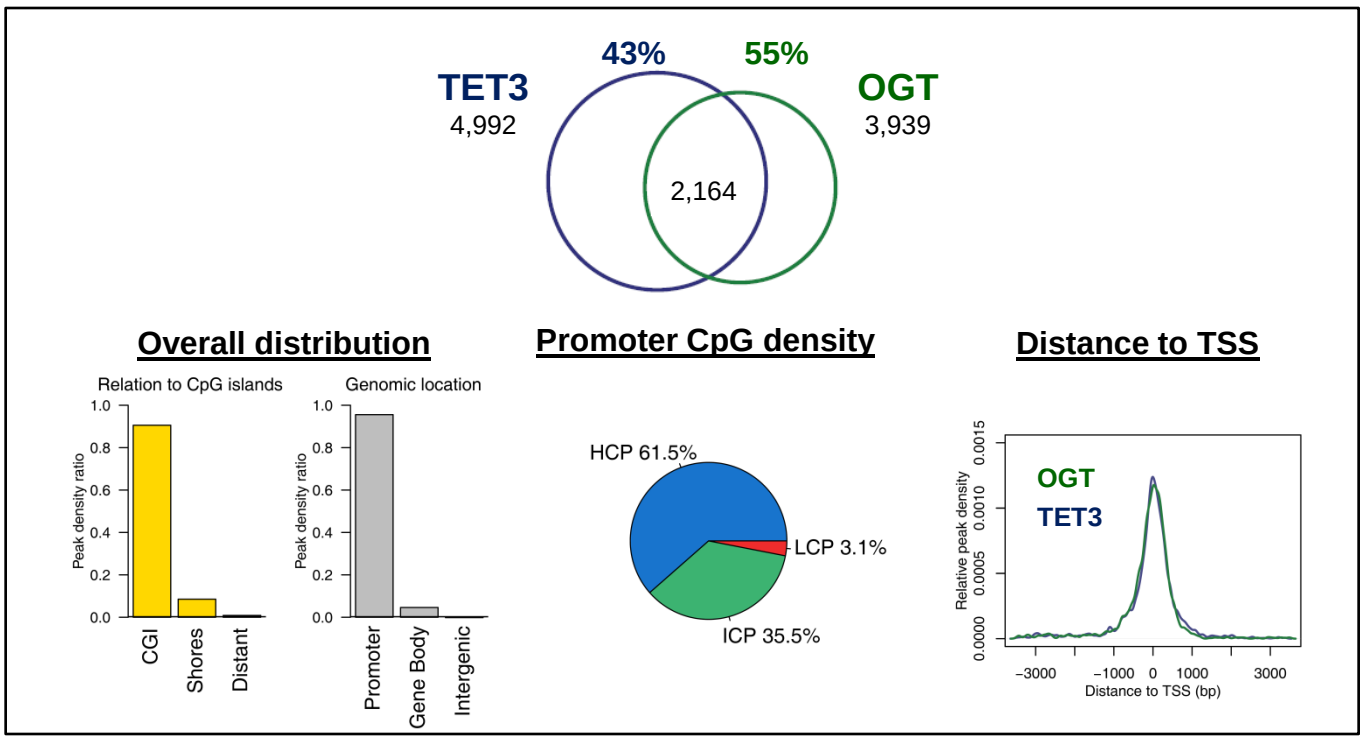
Supplementary Fig. 5 (related to Figs. 1-3): Control experiments showing levels of RNAi knockdowns of TET2/3 and OGT. (A) Validation of RNAi TET2/3 by quantitative PCR analysis (top) and Western blots (bottom) for the TET2, TET3, as well as for several loading controls. For RT-qPCR, each qPCR analysis was normalized with the expression of housekeeping gene SDHA and expression of each mRNA with RNAi control was set to a value of 1. Error bars represent SD of 3 independent experiments. **(B)** Validation of RNAi OGT. Upper part, quantitative PCR analysis of mRNA expression. Normalisation was done as in (A). Error bars represent SD of 3 independent experiments. Lower part, Western blot for OGT. Control antibodies are shown .



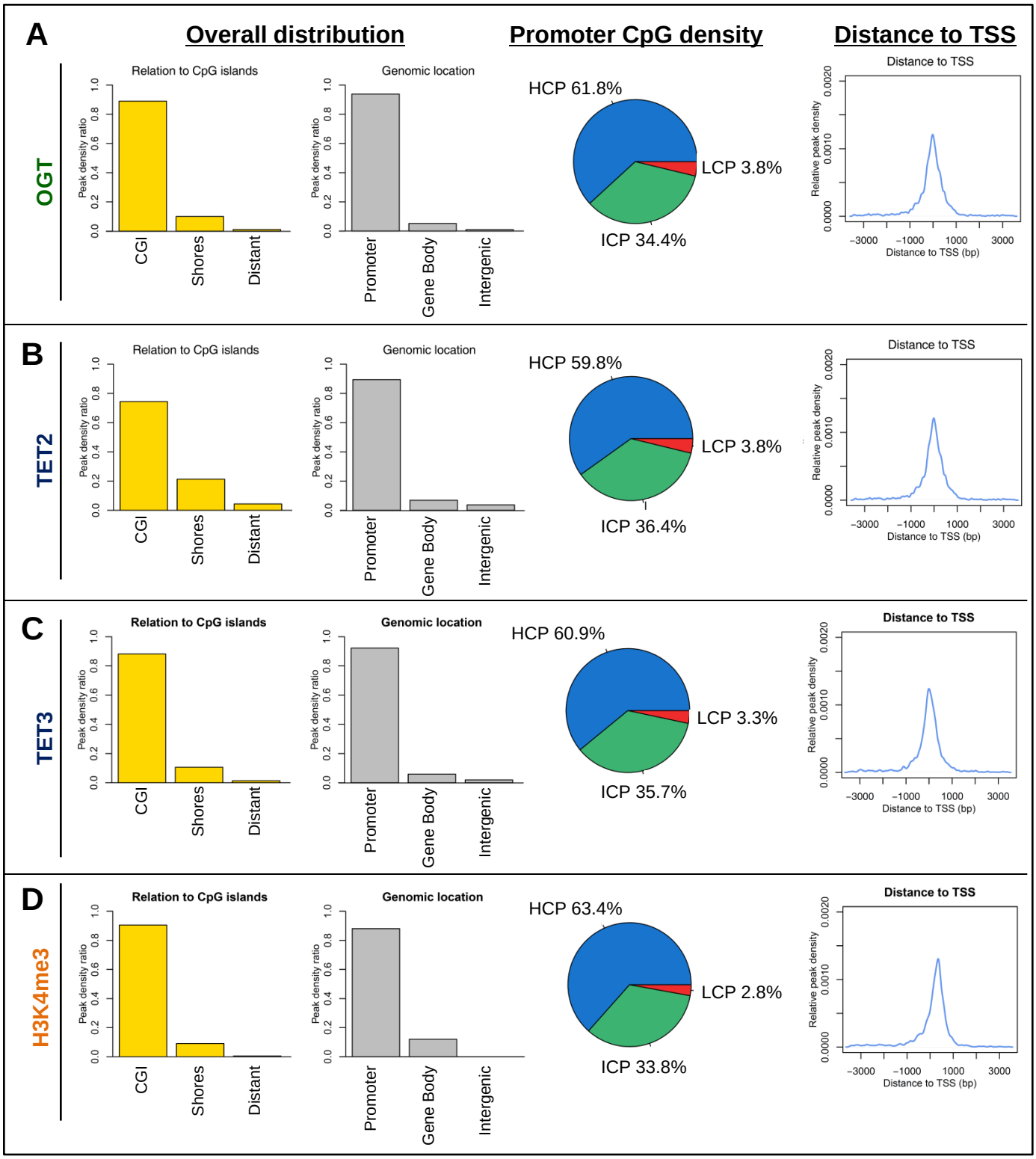
Supplementary Fig. 6 (related to Figs. 1-3): Control experiments showing levels of RNAi knockdown of TET2 and TET3, their influence on OGT, OGA, and SETD1A expression, as well as HT TET2/3-OGT expression levels. **(A)** Validation of RNAi TET2 or RNAi TET3 by RTqPCR (left) and Western blotting (right). **(B), (C),** OGT, OGA and SETD1A expression levels are not affected in TET2/3 knockdown cells as assessed by RT-qPCR and normalized with the expression of housekeeping gene SDHA. Expression of each mRNA with RNAi control was set to a value of 1. Validation at protein and mRNA levels of RNAi for TET2 and TET3. **(D)** Overexpression of HaloTag TET2, TET3, and OGT proteins. Quantitative PCR analysis was performed on control cells (HT-Ctrl) and HT-OGT, HT-TET2, or HT-TET3 overexpressed HEK293T cells. Primers were designed to amplify endogenous and overexpressed mRNA. The results were normalized with the expression of housekeeping gene SDHA. Expression of each mRNA with empty vector (HT-Ctrl) was set to a value of 1.

Sample		Read Count	Percentage Mapped	Mapped Read Count
HEK293T	H3K4me3	47346891	75.84	35906475
HEK293T	HT-TET2	45435735	81.96	37240410
HEK293T	HT-TET3	44808539	79.80	35757243
HEK293T	HT-OGT	42423294	81.23	34460216
HEK293T	*O*-GlcNAc	37323865	81.02	30240447
HEK293T	HCF-1	57937349	86.68	50217828
HEK293T RNAi Ctrl	H3K4me3	34509438	81.16	28009414
HEK293T RNAi TET2	H3K4me3	32691456	76.61	25046354
Bone Marrow	Tet2	73255638	74.57	54625683
Bone Marrow	*O*-GlcNAc	45974200	79.16	36393342
Bone Marrow	H3K4me3	65400033	80.77	52824542
Bone Marrow Wt	H3K4me3	79295188	69.19	54864174
Bone Marrow Tet2 Ko	H3K4me3	27284025	77.41	21120193

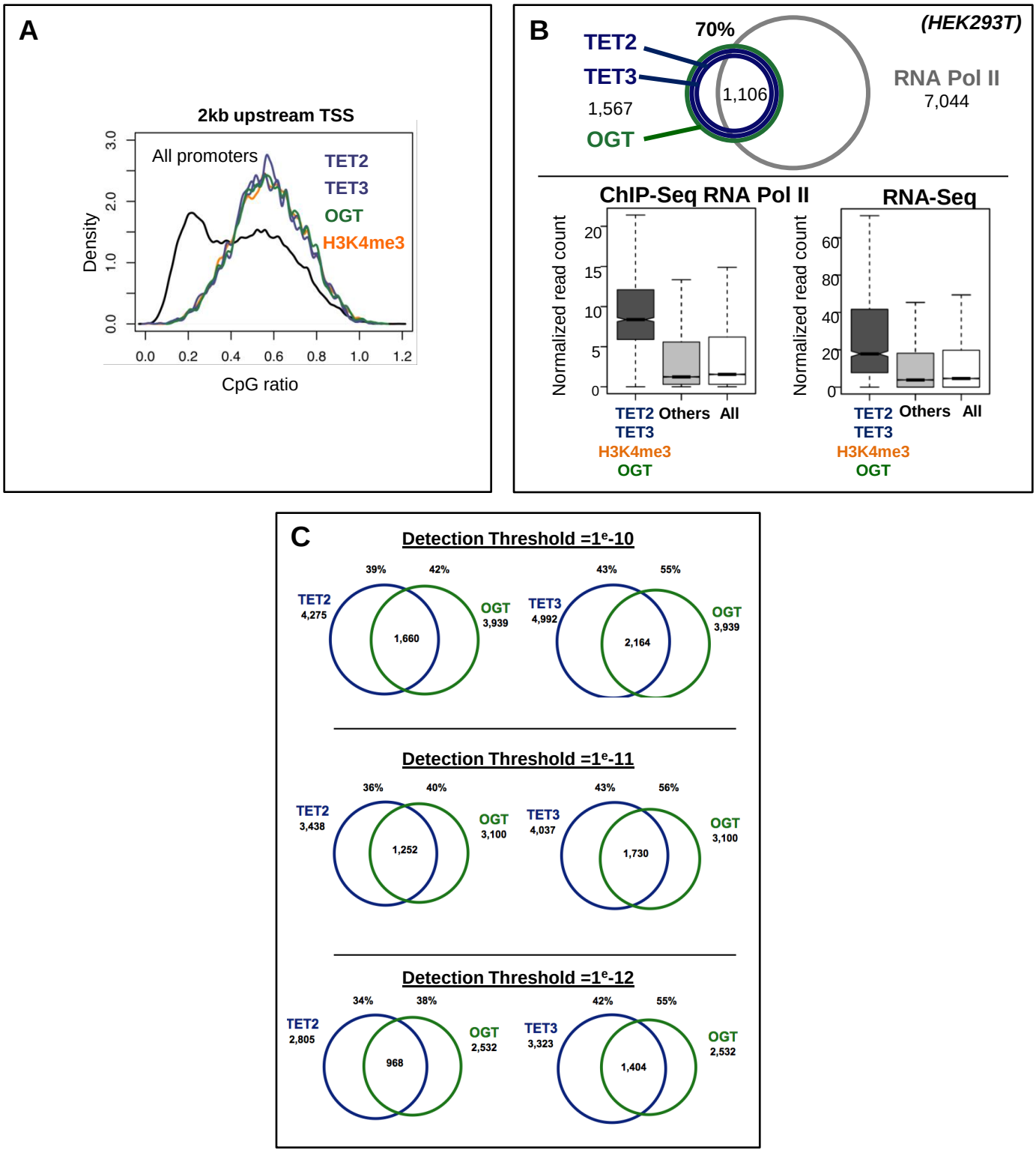
Supplementary Fig. 7 (related to Figs. 2-4): Illumina sequencing run statistics. For each sample the raw number of reads, the number of reads mapped to the reference genome (hg18 for H293T and mm9 for Bone Marrow samples), and the coverage achieved are shown.



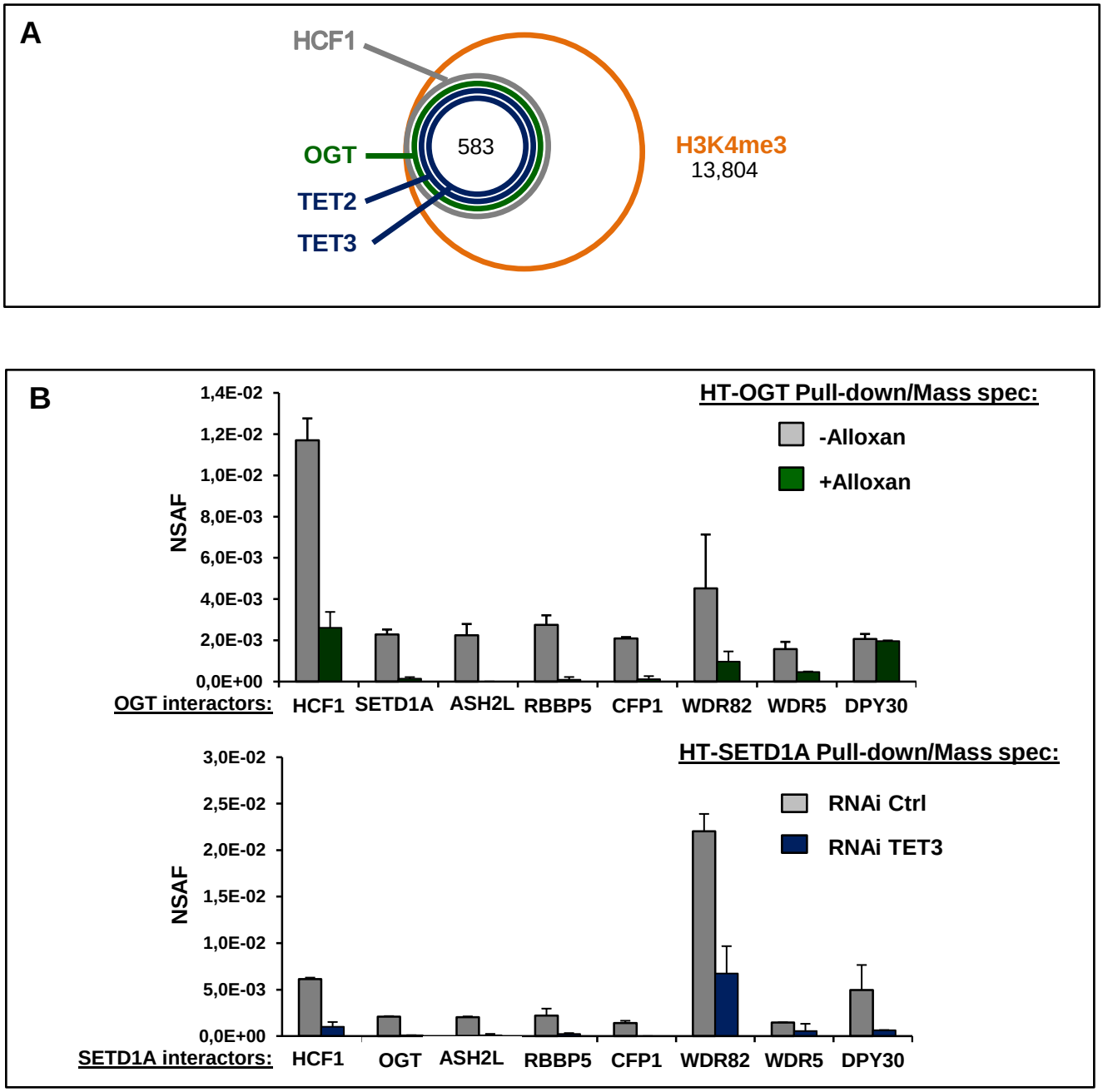
Supplementary Fig. 8 (related to Fig. 2A): Genome-wide distribution of TET3 and OGT enriched regions. Upper-part depicting the overlap (P-value overlap < 10^{-10}) between TET3-OGT in HEK293T. Bottom: Left part, relation to CpG islands and genomic location of the enriched regions are shown as barplot of the relative peak density inside CpG islands (CGI), close to CpG islands (Shores) or distant and inside promoters (2kb upstream the TSS), inside gene body regions or intergenic regions. Center part, pie chart showing the fraction of enriched promoters belonging to CpG-rich promoters (HCP), CpG poor promoters (LCP) or intermediate CpG promoters (ICP). Right part, density plots showing the distribution of enriched region location around the TSS.



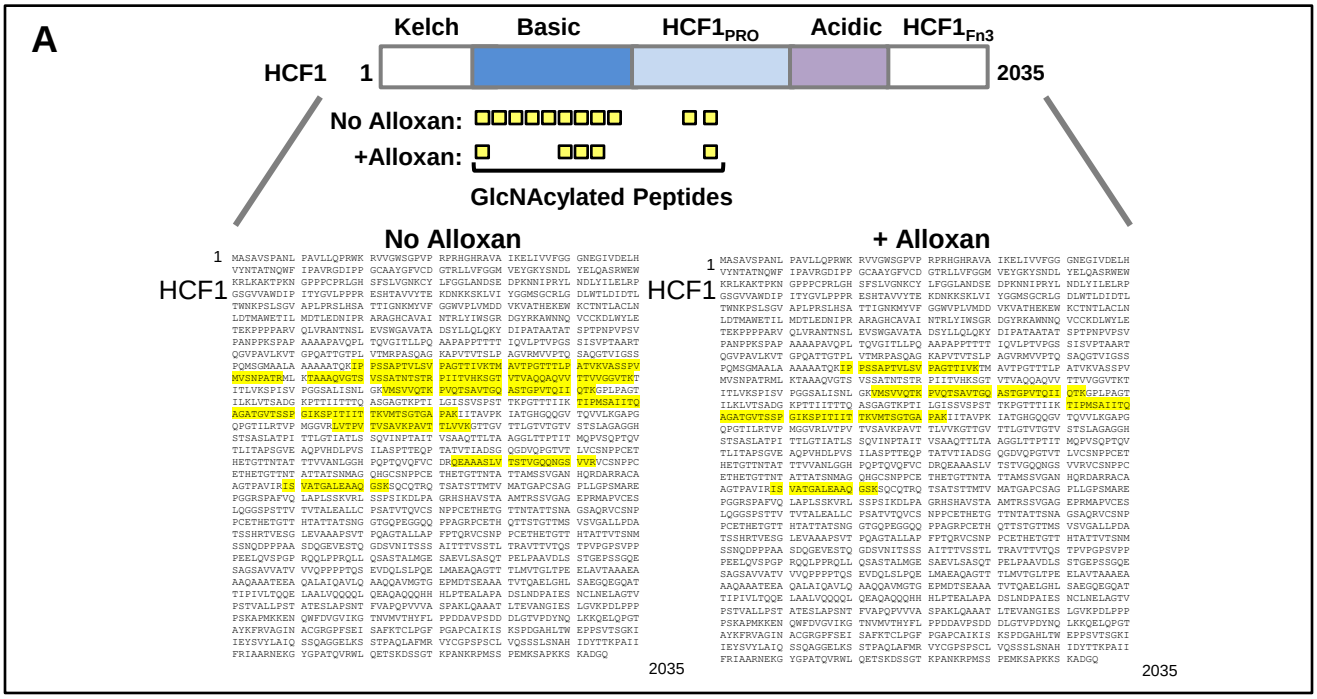
Supplementary Fig. 9 (related to Fig. 2a): Genome-wide distribution of the OGT, TET2, TET3 and H3K4me3 enriched regions. (A) OGT, **(B)** TET2, **(C)** TET3, **(D)** H3K4me3. Left-part, relation to CpG islands and genomic location of the enriched regions are shown as barplot of the relative peak density inside CpG islands (CGI), close to CpG islands (Shores) or distant and inside promoters (2kb upstream the TSS), inside gene body regions or intergenic regions. Center-part, pie chart showing the fraction of enriched promoters belonging to CpG-rich promoters (HCP), CpG poor promoters (LCP) or intermediate CpG promoters (ICP). Lower-right, density plots showing the distribution of enriched region location around the TSS.

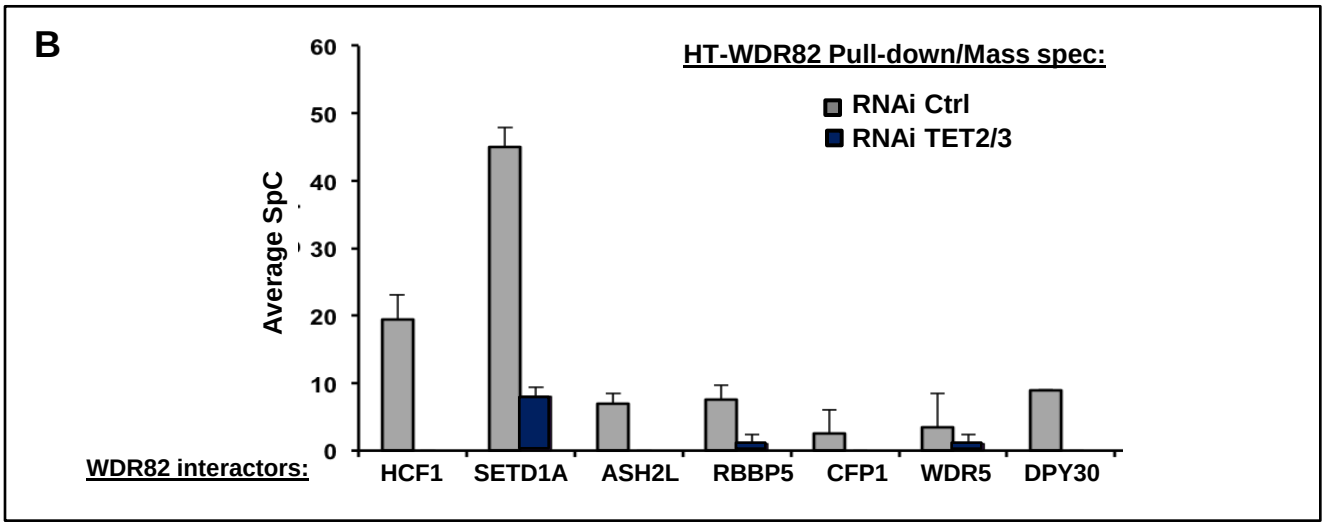


Supplementary Fig. 10 (related to Fig. 2): Overlap profiles of TET2, TET3, OGT, H3K4me3, RNA Pol II along with HaloCHIP controls in HEK293T cells. (A) Density plot of CpG ratio for all promoters (black) and for promoters enriched in HT-TET2, HT-TET3 (blue), HT-OGT (green) and H3K4me3 (orange). **(B)** Upper panel: Venn diagrams depicting the overlap (P-value overlap < 10^{-10}) between TET2-TET3-OGT-H3K4me3 and RNA Pol. Lower panel: Box plots showing the correlation between TET2-TET3-OGT-H3K4me3 targets and RNA Pol II or expression levels. **(C)** Percentage of overlapping TET2-OGT or TET3-OGT targets (Venn diagrams) minimally affected by the peak selection and remains in the same range for different thresholds. Overlap is computed for the indicated different peaks calling thresholds.



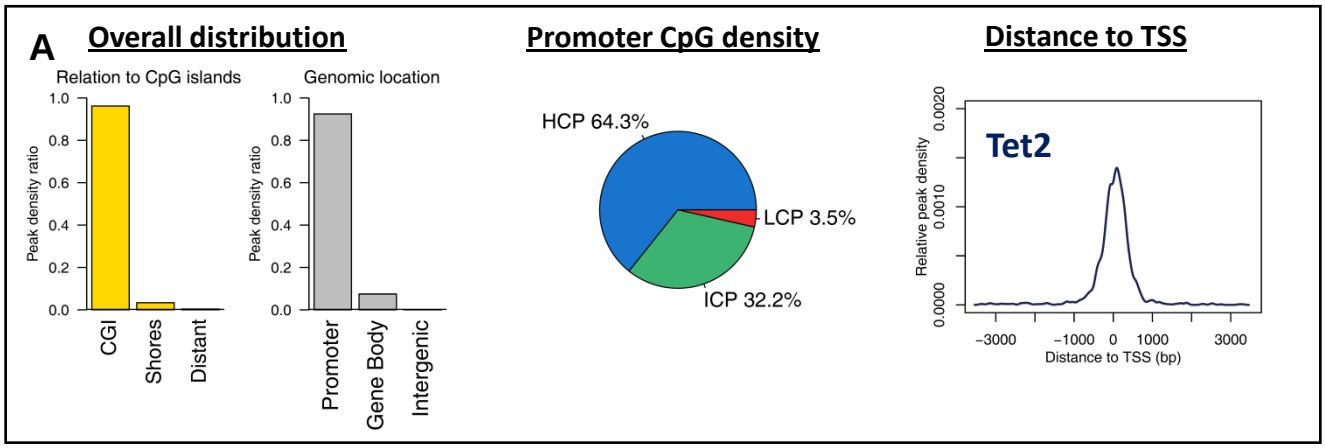
Supplementary Fig. 11 (related to Fig. 3): Genomic overlap of endogenous HCF1 in HEK293T cells with TET2, TET3, OGT, and H3K4me and NSAF values for OGT and SETD1A isolations (A) Venn diagram shows full overlap of HCF1 target genes, as determined by ChIP-Seq against endogenous HCF1, with those of Halo-TET2, TET3, OGT, and H3K4me3. **(B)** Relative abundance of SET1/COMPASS components in HT-OGT isolations (NSAF values). Plot of normalized spectral abundance factors (NSAFs) for SET1/COMPASS components identified by mass spectrometry from HT-OGT complex isolations in the absence (grey bars) or presence (green bars) of Alloxan (top part**) or** from HT-SETD1A complex isolations from control (grey bars) or TET3 RNAi kd (blue bars) cells **(**bottom part). Values represent average NSAF of biological replicates +/- SD.



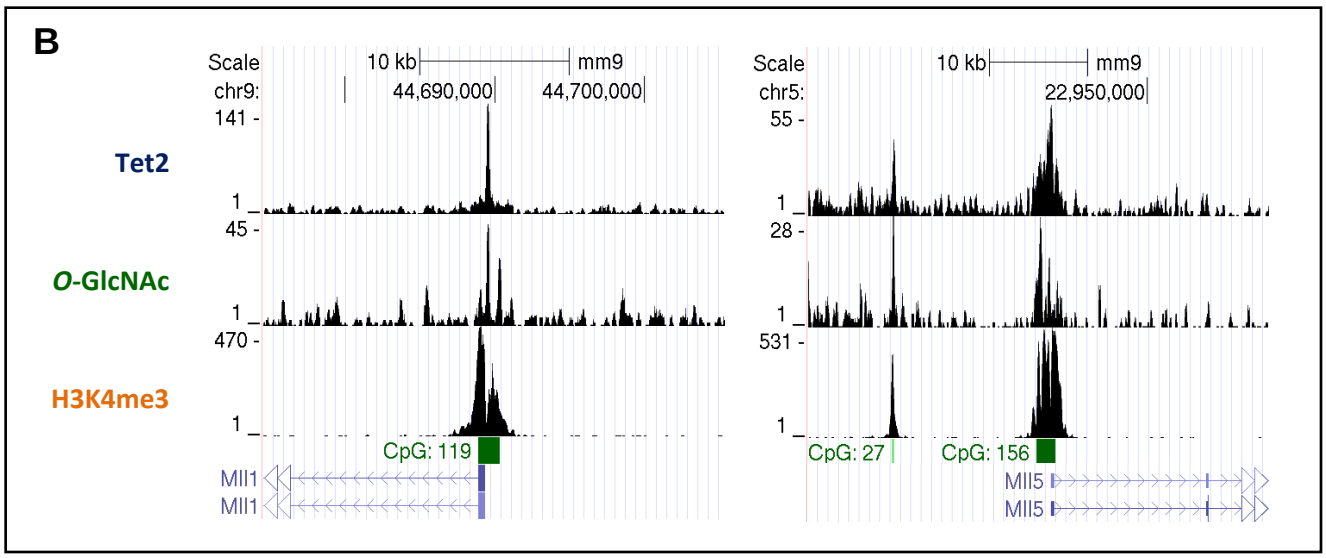


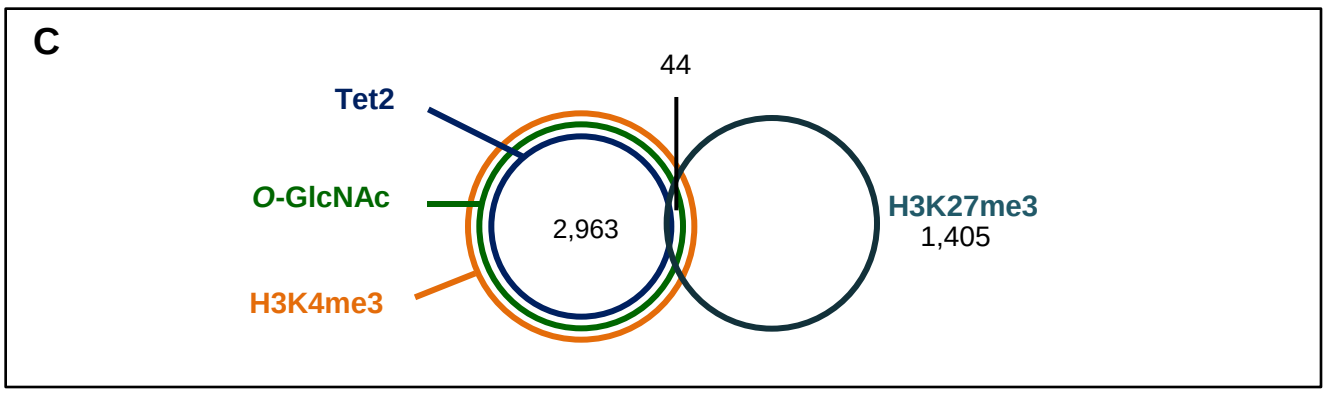
Supplementary Fig. 12 (related to Fig. 3): GlcNAcylation of HCF1 in OGT complex isolations +/- Alloxan treatment and proteomics of WDR82 (a unique component of SET1/COMPASS) shows disruption of SET1/COMPASS complex in TET2/3 RNAi knockdown cells. (A) Upper part: schematic representation of full-length HCF1. Indicated below are the GlcNAcylated peptides identified by mass spectrometry after HT-OGT isolation in HEK293T cells grown in the absence or presence of Alloxan. Lower part: HCF1 amino acid sequence (NP_005325.2) shows the corresponding GlcNAcylated peptides highlighted in yellow. These results match with previously published data[15]. **(B)** Relative abundance of SET1/COMPASS complex components in HT-WDR82 isolations in HEK293T cells. RNAi control (grey bars) or RNAi TET2/3 (blue bars) are shown. Values represent average SpC of biological replicates +/- SD. These data with WDR82 support the disruption of the SET1/COMPASS complex in TET2/3 knockdown shown in Fig. 3d.

Mouse bone marrow:
ChIP-Seq Tet2

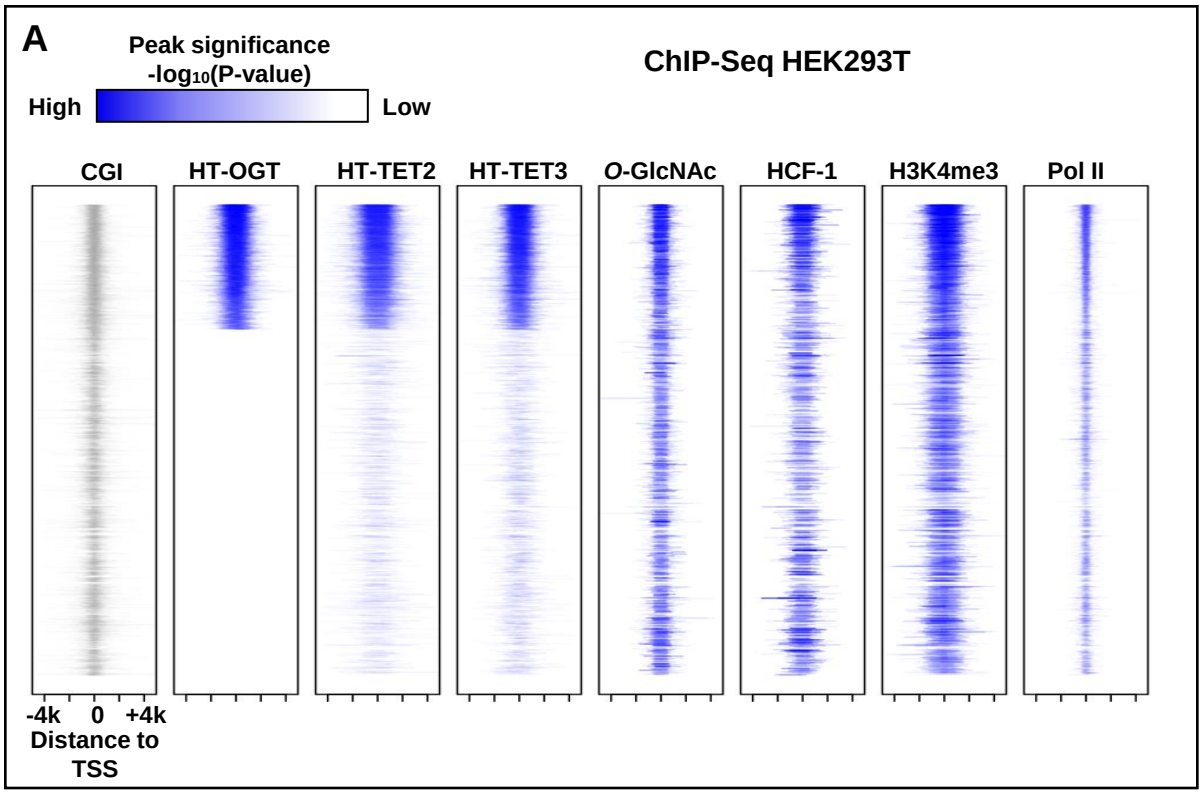


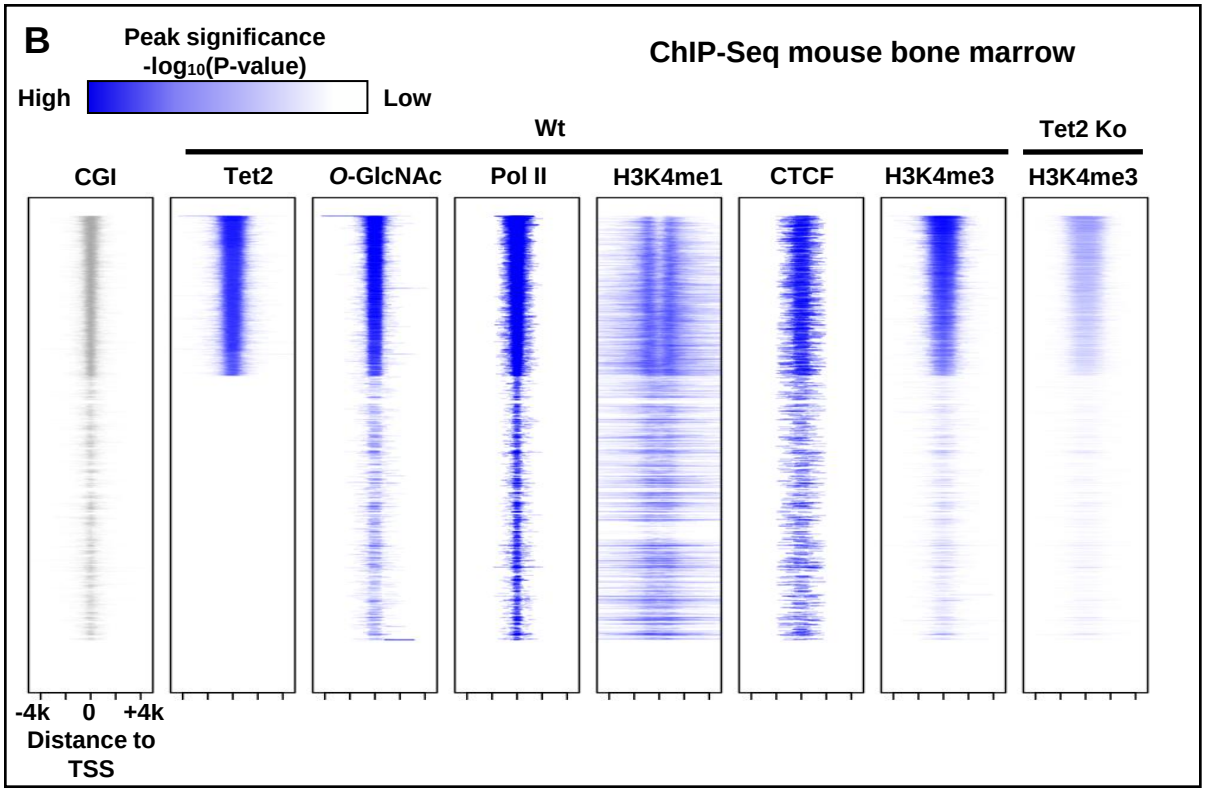
Mouse bone marrow:
ChIP-Seq Tet2, O-GlcNac, H3K4me3





Supplementary Fig. 13 (related to Figs. 4a, b): Genome-wide binding profile in mouse bone marrow tissue of Tet2, which is mostly enriched at CpG-rich regions and TSSs. (A) From left to right: overall genomic distribution (barplots of the relative peak density in relation to CGI or genomic location); promoter distribution of HCP, ICP, and LCP; relative peak density for bound regions with respect to the distance to TSSs. **(B) Examples of ChIP-seq profiles of Tet2, *O*-GlcNAc, and H3K4me3 binding:** *Mll1* and *Mll5* promoter regions are shown (UCSC tracks). **(C)** H3K27me3 does not show overlaps with Tet2, O-GlcNAc, and H3K4me3 in mouse bone marrow tissue. Venn diagrams as determined by ChIPSeq in mouse bone marrow tissue.





Supplementary Fig. 14 (related to Figs. 2, 4): Heatmap showing ChIP-Seq in HEK293T and mouse bone marrow tissue. **(A)** Shown are enriched region around RefSeq genes TSS identified by ChIP-Seq for TET2, TET3, *O*-GlcNAc, HCF1, H3K4me3 and, Pol II in HEK293T. UCSC annotated CpG islands (CGI) are also reported. Genes are ordered according their Tet2 enrichment (average peak significance in 1kb region surrounding the TSS). **(B)** shown are enriched region around RefSeq genes TSS identified by ChIP-Seq for Tet2, *O*-GlcNAc, H3K4me3, Pol II, H3K4me1, CTCF in mouse Wt bone marrow tissue and H3K4me3 for Tet2 Ko. UCSC annotated CpG islands (CGI) are also reported. Genes are ordered according their Tet2 enrichment (average peak significance in 1kb region surrounding the TSS).

Supplementary Table 1: Primer lists and sequences.

Name	Sequence 5'-3'	Target	Application
TET1 F	CCGAATCAAGCGGAAGAATA	human TET1	qPCR expression
TET1 R	ACTTCAGGTTGCACGGTCTC	human TET1	qPCR expression
TET2 F	AGCCCCATCACGTACAAAAC	human TET2	qPCR expression
TET2 R	TGTGGTGGCTGCTTCTGTAG	human TET2	qPCR expression
TET3 F	CAGCAGCCGAGAAGAAGAAG	human TET3	qPCR expression
TET3 R	GGACAATCCACCCTTCAGAG	human TET3	qPCR expression
OGT F	GCAACCTAGCCAATGCTCTC	Human OGT	qPCR expression
OGT R	GCAGCAAACTCTGGGAAGAC	Human OGT	qPCR expression
POLR3C F	CATGCTCCGAATGAGTGAGA	POLR3C	qPCR expression
POLR3C R	GCCAACAGGTAGGGATCTGA	POLR3C	qPCR expression
TMEM116 F	TGGCCCTTTATGTTCTCCAG	TMEM116	qPCR expression
TMEM116 R	GCCTCCTGCTTTAGTTGGTG	TMEM116	qPCR expression
CYB5D2 F	GCTTGGAGGCCAACAAACTA	CYB5D2	qPCR expression
CYB5D2 R	CTCCACTCTTCTGGGAGCAC	CYB5D2	qPCR expression
NSUN4 F	AACTGGAGGGGGACACCTAT	NSUN4	qPCR expression
NSUN4 R	CTGTCGCTCCTTCTTCCTTG	NSUN4	qPCR expression
SLC3A2 F	CCAGGTTCGGGACATAGAGA	SLC3A2	qPCR expression
SLC3A2 R	GAGTTAGTCCCCGCAATCAA	SLC3A2	qPCR expression
ACTA2 F	CCGGGAGAAAATGACTCAAA	ACTA2	qPCR expression
ACTA2 R	GCGTCCAGAGGCATAGAGAG	ACTA2	qPCR expression
CENPO F	AACCCAACCAAACAGTGGAG	CENPO	qPCR expression
CENPO R	CGGCTGGTCAGTTTACCACT	CENPO	qPCR expression
STYXL1 F	TAACGAGCAGACCTTGCAGA	STYXL1	qPCR expression
STYXL1 R	GTCTTCTCCCATTCCAGCAG	STYXL1	qPCR expression
CDKL4 F	TGTTGAAGCAATTAAAACATCCA	CDKL4	qPCR expression
CDKL4 R	TCCTTTTTCTCCTGAACACCTC	CDKL4	qPCR expression
MITF F	AAGCAAGAGCACTGGCCAAA	MITF	qPCR expression
MITF R	CCTTGTTCCAGCGCATGTCT	MITF	qPCR expression
ACTINE F	CTGGAACGGTGAAGGTGACA	ACTINE	qPCR expression
ACTINE R	AAGGGACTTCCTGTAACAATGCA	ACTINE	qPCR expression
GAPDH F	TGCACCACCAACTGCTTAGC	GAPDH	qPCR expression
GAPDH R	GGCATGGACTGTGGTCATGAG	GAPDH	qPCR expression
SDHA F	TGGGAACAAGAGGGCATCTG	SDHA	qPCR expression
SDHA R	CCACCACTGCATCAAATTCATG	SDHA	qPCR expression
HPRT1 F	TGACACTGGCAAAACAATGCA	HPRT1	qPCR expression
HPRT1 R	GGTCCTTTTCACCAGCAAGCT	HPRT1	qPCR expression
rRNA 18S F	CAGCCACCCGAGATTGAGCA	rRNA 18S	qPCR expression
rRNA 18S R	TAGTAGCGACGGGCGGTGTG	rRNA 18S	qPCR expression

Supplementary Table 1 (continued) : Primer lists and sequences.

Name	Sequence 5'-3'	Target	Application
shRNA TET1 F	gatccccAGATTAAGTTTGAGGCTAAttcaagagaTTAGCCTCAAACTTAATCTttttta	human TET1	RNAi
shRNA TET1 R	aggttaaaaaAGATTAAGTTTGAGGCTAAtctcttgaaTTAGCCTCAAACTTAATCTggg	human TET1	RNAi
shRNA TET2 F	gatccccGGATCATTCTTTGGCCAGAttcaagagaTCTGGCCAAAGAATGATCCttttta	human TET2	RNAi
shRNA TET2 R	agcttaaaaaGGATCATTCTTTGGCCAGAtctcttgaaTCTGGCCAAAGAATGATCCggg	human TET2	RNAi
shRNA TET3 F	gatccccGGAGAAAGATGAAGGTCCAttcaagagaTGGACCTTCATCTTTCTCCttttta	human TET3	RNAi
shRNA TET3 R	aggttaaaaaGGAGAAAGATGAAGGTCCAtctcttgaaTGGACCTTCATCTTTCTCCggg	human TET3	RNAi
shRNA OGT F	gatccccGCTGAGCAGTATTCCGAGAAAttcaagagaTTTCTCGGAATACTGCTCAGCttttta	human OGT	RNAi
shRNA OGT R	agcttaaaaaGCTGAGCAGTATTCCGAGAAAtctcttgaaTTTCTCGGAATACTGCTCAGCggg	human OGT	RNAi
EFCAB7-F	CCGGATCAGCGCTTCCTAGA	EFCAB7	MeDIP/hMeDIP/ChIP/Halo-ChIP (#1)
EFCAB7-R	CATGGCAACTGCACCTCACC	EFCAB7	MeDIP/hMeDIP/ChIP/Halo-ChIP (#1)
LRIG2-F	CTGCGTCTGCTCCTTGCATG	LRIG2	MeDIP/hMeDIP/ChIP
LRIG2-R	GGCGGAGAAGCTCAGCAGAG	LRIG2	MeDIP/hMeDIP/ChIP
TROVE2-F	GCAAACGCCGCGAAACTATC	TROVE2/UCHL5	MeDIP/hMeDIP/ChIP
TROVE2-R	GCGATGCTGGTTCCGTTACC	TROVE2/UCHL5	MeDIP/hMeDIP/ChIP
MTR-F	GCGCAAGGAGGAGACTCGAC	MTR	MeDIP/hMeDIP/ChIP
MTR-R	ACACAACCCAACCACGCAGA	MTR	MeDIP/hMeDIP/ChIP
Intergenic-F	GGAGCCATCATCATCAACACA	Intergenic	MeDIP/hMeDIP/ChIP /Halo-ChIP (#10)
Intergenic-R	CGCTGAATCCCAAAGCATGA	Intergenic	MeDIP/hMeDIP/ChIP /Halo-ChIP (#10)
PPIH-F	GGGAACGAGGAAGTGCCTGA	PPIH	MeDIP/hMeDIP/ChIP/Halo-ChIP (#11)
PPIH-R	GGGCCTGTGGGAACATTCTA	PPIH	MeDIP/hMeDIP/ChIP/Halo-ChIP (#11)
NBR2-F	CGCACAGGTCTCCAATCTATC	NBR2	MeDIP
NBR2-R	AATGAAAGGTCTTCGCCACA	NBR2	MeDIP
MAN1A1-F	TCCCTCTCCCTCTCCCTCTCTG	MAN1A1	hMeDIP
MAN1A1-R	AGTGCCCATCCAGATTTCCTGA	MAN1A1	hMeDIP
USP32-F	TGAAGAAGACGGAGAGAAGGAAGG	USP32	hMeDIP
USP32-R	AGCCAGGCATGGAGAGGAAA	USP32	hMeDIP
TAF12-F	AAGACTCACAGGCAGCAGCGTCTATC	TAF12	ChIP /Halo-ChIP (#2)
TAF12-R	GCTGCTGAGACGAACGCTTCACTG	TAF12	ChIP /Halo-ChIP (#2)
ZNF335-F	ACTTTATACGTCGCCCGGTGTTTG	ZNF335	ChIP /Halo-ChIP (#7)
ZNF335-R	CCTACGTAACACTCCGCCGAGATAAAG	ZNF335	ChIP / Halo-ChIP (#7)
EFCAB7-F	TTCTAGGCCCAGGACAGCACCTATTTC	EFCAB7	Halo-ChIP (#3)
EFCAB7-R	GGTGGTATGGCCAGAGTAGTGAGAGAAG	EFCAB7	Halo-ChIP (#3)
RPS7-F	CCTGCACTGAAGACGGGAACCTTG	RPS7	Halo-ChIP (#4)
RPS7-R	TTCCGCCCTAGGATTCGTACTTTGG	RPS7	Halo-ChIP (#4)
ABT1-F	CGTCAGTTGGCAAGGCACTTTACG	ABT1	Halo-ChIP (#5)
ABT1-R	CTTCTCCGATTCCTCTGCCTCCATG	ABT1	Halo-ChIP(#5)
SSBP1-F	TTCCTCTGGCGAGCTTTGCGTTC	SSBP1	Halo-ChIP (#6)
SSBP1-R	CACCTGACGATCTAACCCGAGCAG	SSBP1	Halo-ChIP(#6)
C21ORF59-F	CCACAACCGCTCCACATTCCAG	C21ORF59	Halo-ChIP (#8)
C21ORF59-R	GCCGGAAACAGCGTGGAAAC	C21ORF59	Halo-ChIP (#8)
HDAC8-F	GGTTCCAGTTCCTGCTCCTCTGATC	HDAC8	Halo-ChIP (#9)
HDAC8-R	CTCCGGCGATCTCAAGAATAGGAGG	HDAC8	Halo-ChIP (#9)

Supplementary Table 2: Enriched regions for TET2, TET3, OGT, OGT-TET2, OGT-TET3, H3K4me3 in RNAi Ctrl and in RNAi TET2 HEK293T cells. ChIP-Seq peaks with absolute genomic coordinates on human reference genome hg18 for TET2, TET3, OGT, OGT-TET2, OGT-TET3, H3K4me3, H3K4me3 RNAi Ctrl, H3K4me3 RNAi TET2, HCF1, O-GlcNAc.

Supplementary Table 3: Enriched regions for *O*-GlcNAc, Tet2 and H3K4me3 ChIP-Seq in mouse bone marrow. Peaks with absolute genomic coordinates on mouse reference genome mm9 for *O*-GlcNAc, Tet2, H3K4me3, H3K4me3 Wt and H3K4me Tet2Ko.

3. Supplementary References

The R Project for Statistical Computing. www.r-project.org.

Capotosti F, Guernier S, Lammers F, Waridel P, Cai Y, Jin J, Conaway JW, Conaway RC, Herr W (2011) O-GlcNAc transferase catalyzes site-specific proteolysis of HCF-1. *Cell* **144**(3)**:** 376-388

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

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**(14)**:** 1754-1760

Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C, Figueroa ME, Vasanthakumar A, Patel J, Zhao X, Perna F, Pandey S, Madzo J, Song C, Dai Q, He C, Ibrahim S, Beran M, Zavadil J, Nimer SD, Melnick A, Godley LA, Aifantis I, Levine RL (2011) Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**(1)**:** 11-24

Nestor CE, Ottaviano R, Reddington J, Sproul D, Reinhardt D, Dunican D, Katz E, Dixon JM, Harrison DJ, Meehan RR Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res* **22**(3)**:** 467-477

Pfleger KD, Seeber RM, Eidne KA (2006) Bioluminescence resonance energy transfer (BRET) for the real-time detection of protein-protein interactions. *Nat Protoc* **1**(1)**:** 337-345

Plaisier SB, Taschereau R, Wong JA, Graeber TG Rank-rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures. *Nucleic Acids Res* **38**(17)**:** e169

Robinson MD, Grigull J, Mohammad N, Hughes TR (2002) FunSpec: a web-based cluster interpreter for yeast. *BMC Bioinformatics* **3:** 35

Sardiu ME, Cai Y, Jin J, Swanson SK, Conaway RC, Conaway JW, Florens L, Washburn MP (2008) Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. *Proc Natl Acad Sci U S A* **105**(5)**:** 1454-1459

Song M, Kim HS, Park JM, Kim SH, Kim IH, Ryu SH, Suh PG (2008) o-GlcNAc transferase is activated by CaMKIV-dependent phosphorylation under potassium chloride-induced depolarization in NG-108-15 cells. *Cell Signal* **20**(1)**:** 94-104

Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM, Bollen M, Esteller M, Di Croce L, de Launoit Y, Fuks F (2006) The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**(7078)**:** 871-874

Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* **39**(4)**:** 457-466

Xu Y, Piston DW, Johnson CH (1999) A bioluminescence resonance energy transfer (BRET) system: application to interacting circadian clock proteins. *Proc Natl Acad Sci U S A* **96**(1)**:** 151-156

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**(9)**:** R137