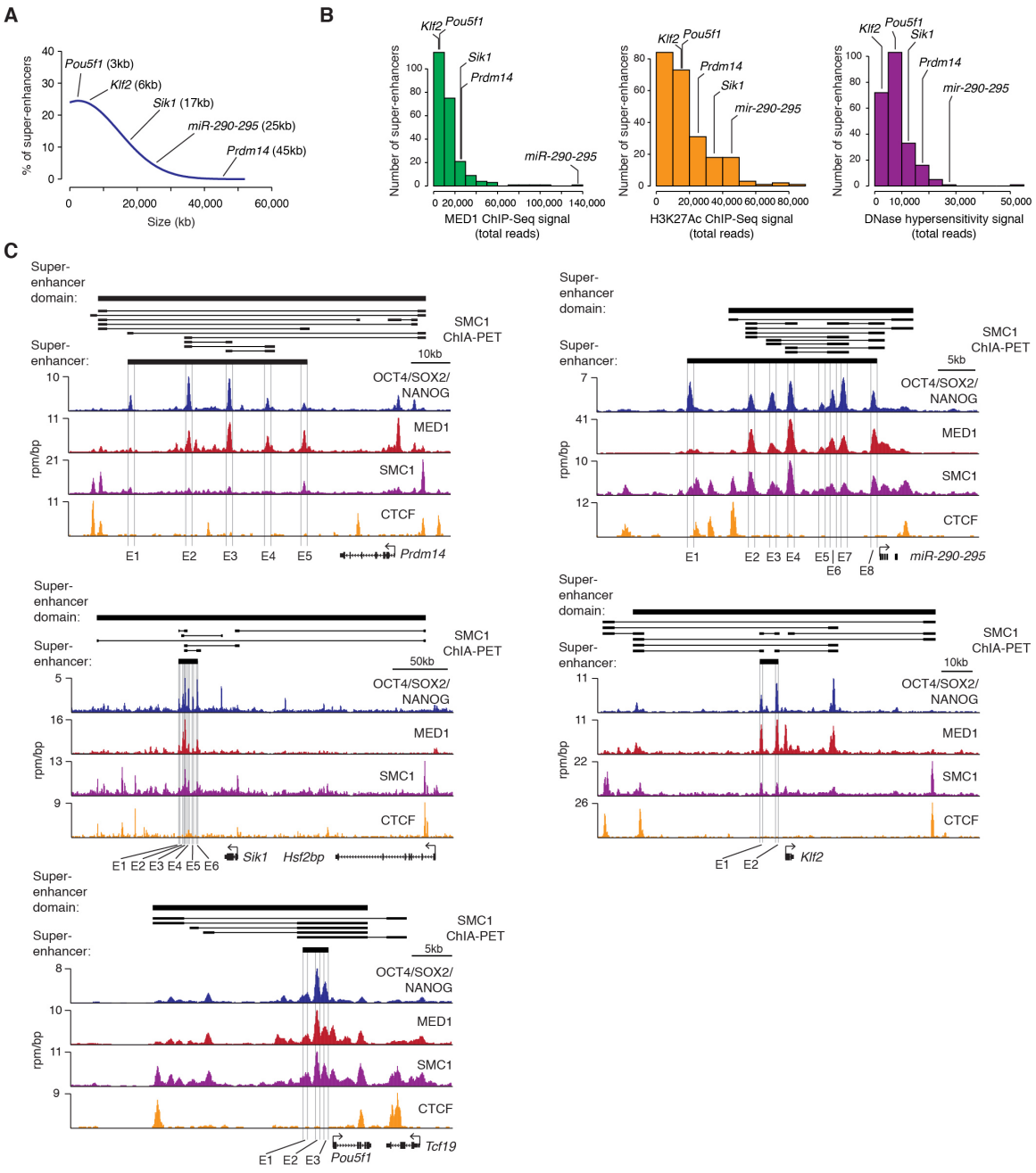
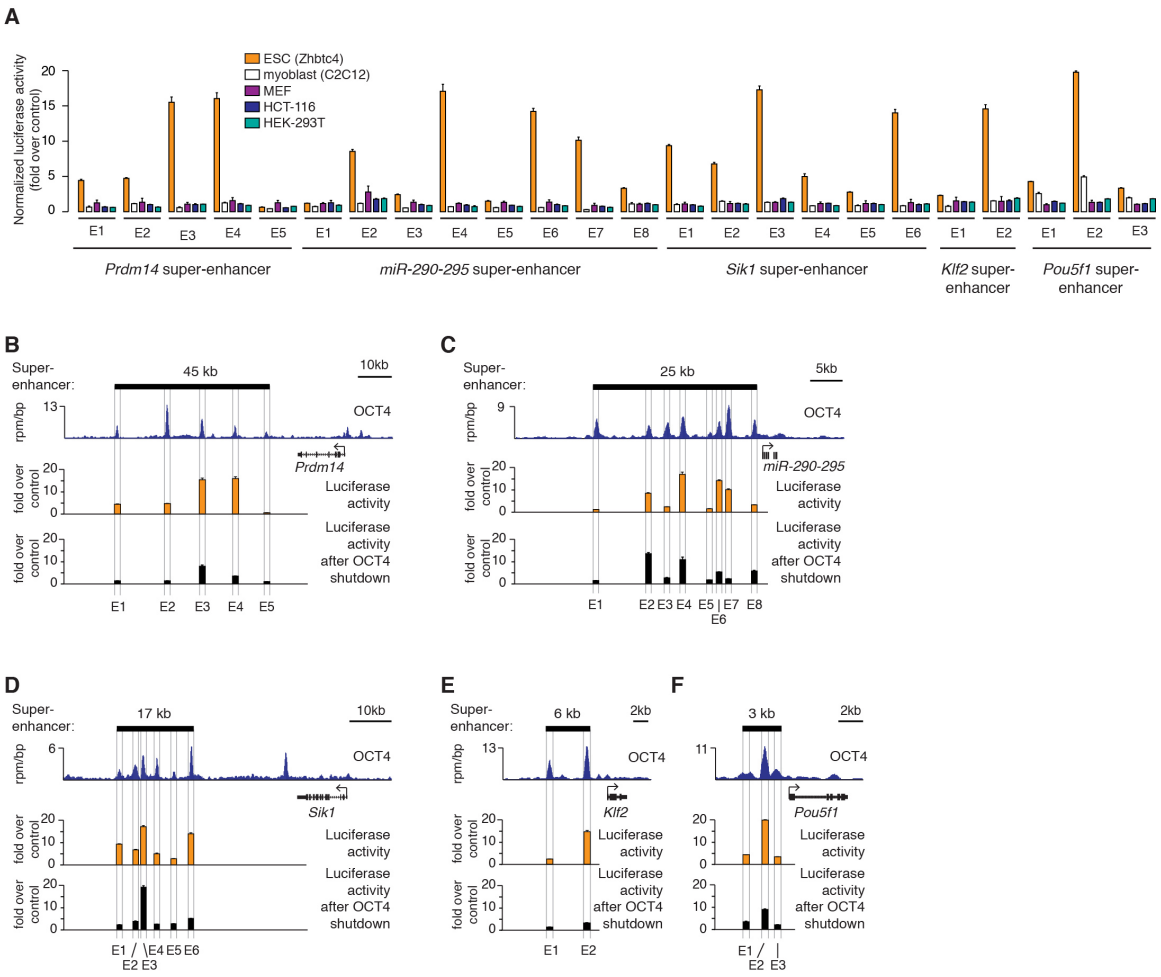


Supplemental Figure 1. Features of super-enhancers in ESCs. Related to Figure 1. (**A**) Histogram of the genomic length of the 231 super-enhancers in murine ESCs. The super-enhancers associated with *Pou5f1*, *Klf2*, *Sik1*, *miR-290-295* and *Prdm14* are highlighted. (**B**) Histograms of Mediator (MED1) and H3K27Ac ChIP-Seq and DNase-hypersensitivity signal at the 231 super-enhancers in murine ESCs. (**C**) Super-enhancer domains at the *Prdm14*, *miR-290-295*, *Sik1, Klf2* and *Pou5f1* loci in ESC. High-

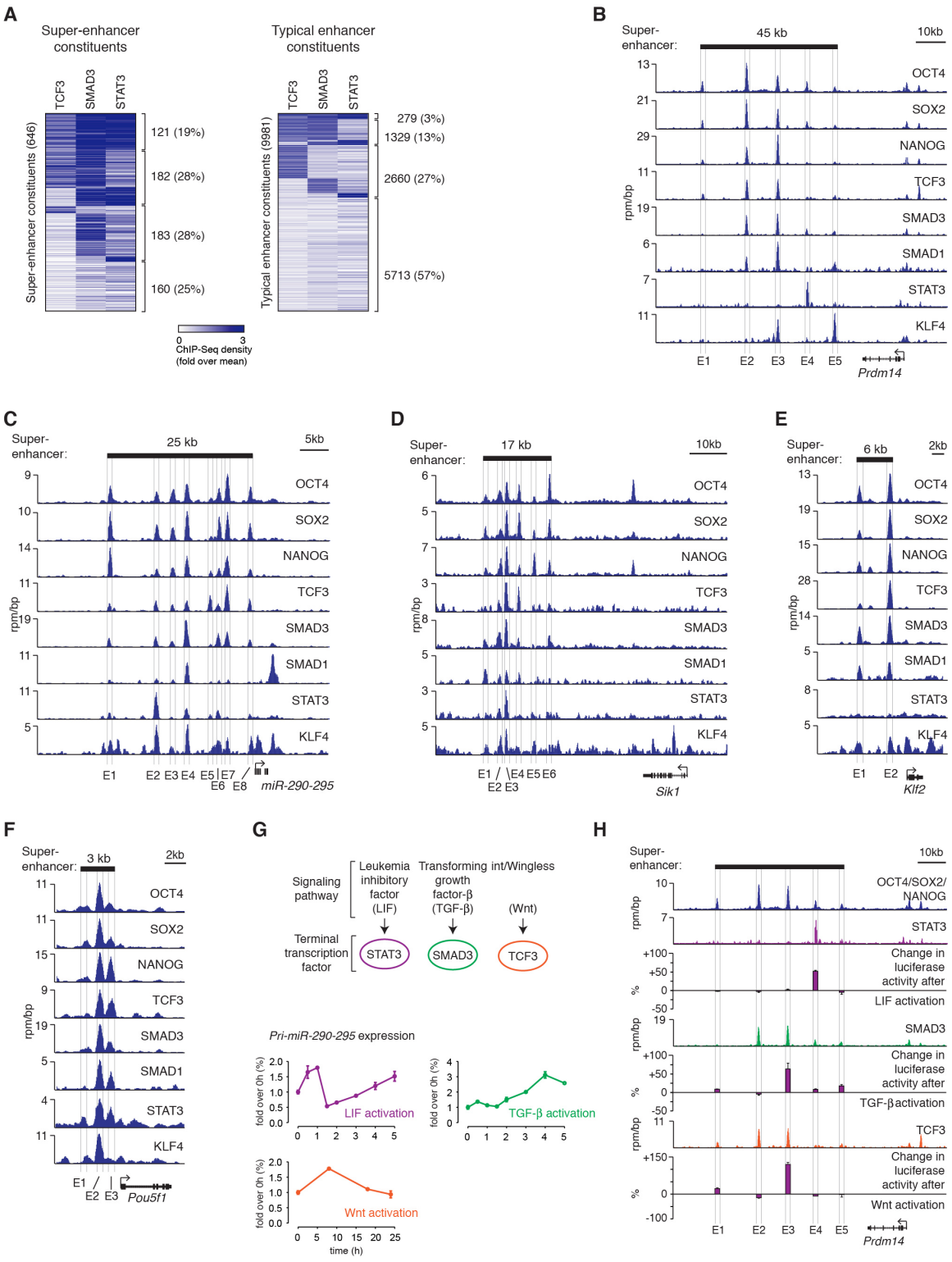
confidence cohesin (SMC1) ChIA-PET interactions are from (Dowen et al., 2014). ChIP-Seq binding profiles for OCT4, SOX2 and NANOG (merged), Mediator (MED1), Cohesion (SMC1) and CTCF are displayed below the high-confidence interactions.



Supplemental Figure 2. Activities of SE-constituents. Related to Figure 1.

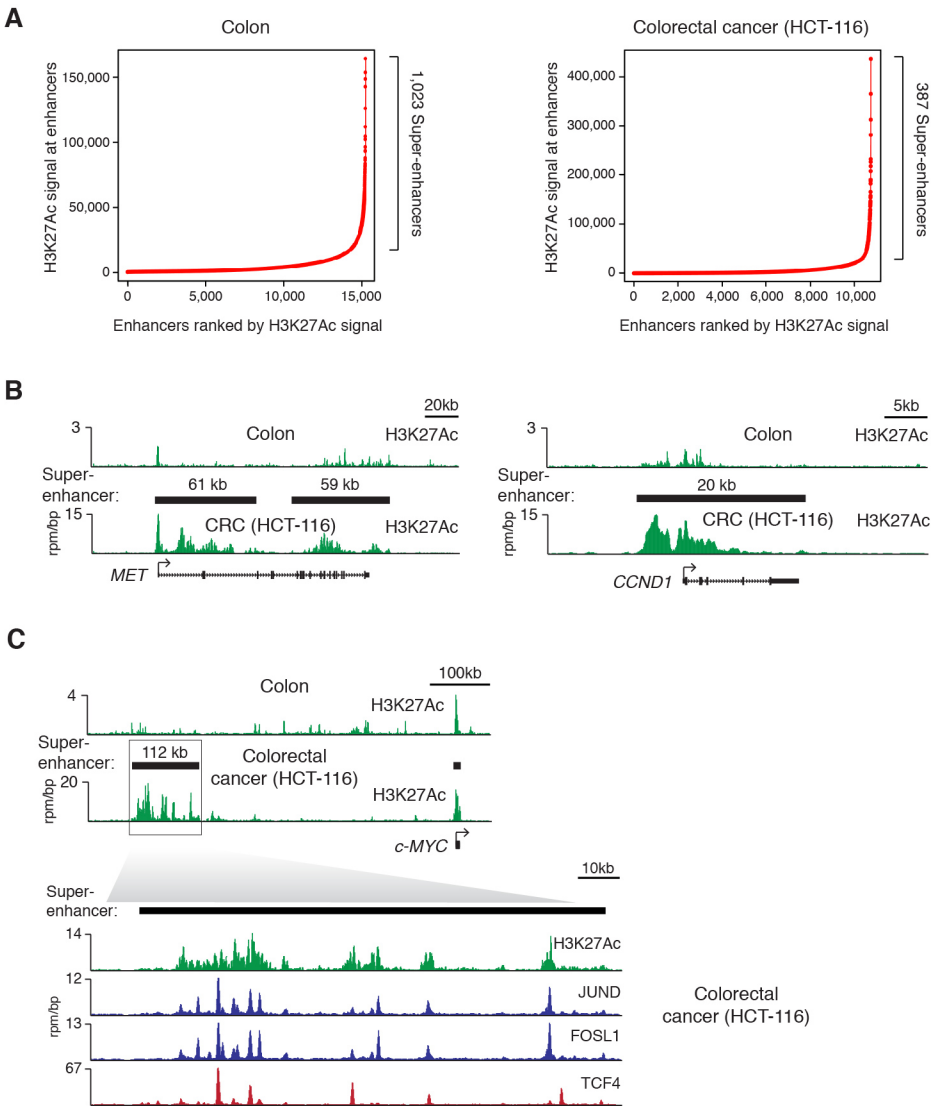
(**A**) Enhancer activity of the mESC SE constituents measured in mESCs, murine myoblasts (C212), primary murine embryonic fibroblasts (MEF), human embryonic kidney cells (HEK-293T) and the colorectal cancer cell line HCT-116. Values correspond to mean +SD from three biological replicate experiments. (**B**) Relationship between OCT4 binding and loss of enhancer activity after OCT4 shutdown. Enhancer activity measured in luciferase reporter assays in wild type cells and the enhancer activity after OCT4 shutdown is plotted for each constituent enhancer within the super-enhancer. Values correspond to mean +SD from three biological replicate experiments. The difference in values after OCT4 shutdown is statistically significant for all constituents, except from *miR-290-295* M1, M3 and M5 ($P<0.05$, Student´s t-test). OCT4 ChIP-Seq

density at each constituent correlates with the change in luciferase activity after OCT4 shutdown (Pearson´s r=0.58).



Supplemental Figure 3. Convergence of signaling at super-enhancers in ESCs. Related to Figure 3.

(**A**) Heat map of ChIP-Seq densities of TCF3, SMAD3 and STAT3 at constituent enhancers of super-enhancers (left) and typical enhancers (right). Read density values were normalized against the mean value calculated for the union of all constituent enhancers. (**B-F**) ChIP-Seq binding profiles for OCT4, SOX2 and NANOG (merged) and the indicated transcription factors at the (**B**) *Prdm14,* (**C**) *miR-290-295*, (**D**) *Sik1,* (**E**) *Klf2* and (**F**) *Pou5f1* loci in ESCs. (**G**) (top) schematic representation of the Wnt, TGF-β and LIF signaling pathways and their terminal transcription factors in ESCs. (bottom) Activation of the Wnt, TGF-β and LIF signaling pathways induces transcription of the SE-associated *miR-290-295* cluster. Gene expression was assayed by qPCR. Values are normalized against the expression level at the start of pathway activation, and correspond to mean +SD from three qPCR replicates. P-value<0.01 for LIF (30min, 1h, 90min, 2h), TGF-β (2h, 3h, 4h, 5h), Wnt (8h) (Student´s t-test). (**H**) Manipulation of the Wnt, TGF-β and LIF signaling pathways tend to affect the activities of SE-constituents occupied by the terminal TFs of the respective pathways. STAT3, SMAD3, TCF binding profiles in ESCs are displayed, along with the change in enhancer activities of the indicated SE-constituents after the stimulation of the LIF, TGF-β and Wnt signaling pathways. Values correspond to mean +SD from three biological replicate experiments.



Supplemental Figure 4. Super-enhancers acquired in colorectal cancer. Related to Figure 4.

(**A**) Distribution of background normalized H3K27Ac ChIP-seq signal at enhancers in normal human colon tissue and in the colorectal cancer cell (CRC) line HCT-116 used to identify super-enhancers (Hnisz et al., 2013). (**B**) ChIP-Seq binding profiles for H3K27Ac at the *MET* and *CCND1* loci in colon and colorectal cancer cells (HCT-116). (**C**) ChIP-Seq binding profiles for H3K27Ac and the indicated transcription factors at the *c-MYC* locus.

Supplemental Table 1. Primer sequences used for the luciferase reporter constructs. Related to Figure 1 and Figure 4.

Supplemental Table 2. Information on sgRNA sequences for CRISPR/Cas9-mediated genome editing, and allele sequences of reference and mutant alleles.
Related to Figure 2.

Supplemental Table 3. Datasets used in this study. Related to Figure 3 and Figure 4.

EXTENDED EXPERIMENTAL PROCEDURES

Cell culture

V6.5 murine embryonic stem cells (ESCs) were cultured on irradiated murine embryonic fibroblasts (MEF) under standard ESC conditions as previously described (Whyte et al., 2013). Briefly, cells were cultured on 0.2% gelatinized (Sigma, G1890) tissue culture plates in ESC media containing DMEM-KO (Invitrogen, 10829-018) supplemented with 15% fetal bovine serum, 1000 U/ml LIF (ESGRO, ESG1106), 100 μM nonessential amino acids (Invitrogen, 11140-050), 2 mM L-glutamine (Invitrogen, 25030-081), 100 U/ml penicillin, 100 μg/ml streptomycin (Invitrogen, 15140-122), and 8 nL/ml of 2-mercaptoethanol (Sigma, M7522).

Zhbtc4 murine ESCs were cultured on 0.2% gelatinized (Sigma, G1890) tissue culture plates in ESC media without a MEF feeder layer. Zhbtc4 cells harbor a Tetracycline-activated transactivator protein (tTA) and Tetracycline-repressible *Pou5f1 (Oct4)* alleles (Niwa et al., 2000). OCT4 expression was shut down by the addition of 2 μg/ml doxycycline to the culture media. After 24 hours of shutdown, the decrease in OCT4 protein in the cells is over 90% (Niwa et al., 2000; Rahl et al., 2010; Whyte et al., 2012).

C2C12 murine myoblasts (ATCC, CRL-1772) and murine primary embryonic fibroblasts (MEF) were cultured on 0.2% gelatinized (Sigma, G1890) tissue culture plates media containing DMEM (Invitrogen, 11965) supplemented with 10% fetal bovine serum, 100 μM nonessential amino acids (Invitrogen, 11140-050), 2 mM L-glutamine (Invitrogen, 25030-081), 100 U/ml penicillin, 100 μg/ml streptomycin (Invitrogen, 15140-122), and 8 nl/ml of 2-mercaptoethanol (Sigma, M7522).

HCT-116 human colorectal cancer cells (ATCC, CCL-247) and HEK-293T cells were cultured in DMEM (Invitrogen, 11965) supplemented with 10% fetal bovine serum, glutaMAX, 100 U/ml penicillin and 100 μg/ml streptomycin (Invitrogen, 15140-122).

Luciferase reporter assays

Murine ESC super-enhancer constituent enhancers (~400bp) were cloned into a pGL3 (Promega) reporter vector (BamHI-SalI sites) that contains a Firefly luciferase gene driven by a minimal *Oct4* promoter (Whyte et al., 2013). V6.5 genomic DNA was used as template DNA, and primer sequences are listed in Supplemental Table 1. $2*10^5$ Zhbtc4

cells or V6.5 ESCs were transfected with 490 ng of the Firefly reporter plasmid in 24-well plates using Lipofectamine 2000 (Invitrogen) according to the manufacturer´s instructions. The amount of transfected plasmid was adjusted according to its size whenever necessary (Figure 1G). 10ng of a Renilla luciferase control plasmid (pRL-SV40; Promega) was co-transfected as a normalization control. In Zhbtc4 cells OCT4-expression was shut down by addition of 2 μg/ml doxycycline to the culture media during transfection. The data displayed on Figure 1G and Figure S3H was generated in V6.5 mESCs. C2C12 cells and MEFs were transfected exactly as mESCs, except Lipofectamine 3000 (Invitrogen) was used. After 24 hours of incubation luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega).

HCT-116 super-enhancer regions (~400bp) were cloned into a pGL3 (Promega) reporter vector (BamHI-SalI sites) that contains a Firefly luciferase gene driven by a minimal *c-Myc* promoter (Loven et al., 2013). HCT-116 genomic DNA was used as template DNA, and primer sequences are listed in Supplemental Table 1. $1*10^5$ HCT-116 cells were transfected with 50ng of the reporters using pPEI (Sigma), in the presence or absence of 20ng of an S33Y-β-catenin expression vector (Addgene: 19286) (Wnt activation on Figure 4B) or 100ng of a ΔNTCF4 expression vector (kind gift from Hans Clevers) (Wnt blockage on Figure 4B). 10ng of a Renilla luciferase control plasmid (pRL-SV40; Promega) was co-transfected as a normalization control. After 15 hours of incubation luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega). All luciferase reporter assays were performed in triplicates. Luciferase activity was normalized to the activity measured in cells transfected with a construct containing only the promoter (empty vector). HEK-293T cells were transfected exactly as the HCT-116 cells using pPEI (Sigma).

Pathway manipulation

In the enhancer reporters assays displayed on Figure S3H, the LIF, TGF-β and Wnt signaling pathways were manipulated as follows. *LIF:* V6.5 mESC cells were plated immediately prior to transfection onto gelatin-coated plates in media without LIF. Following transfection, cells were incubated for 24 hours. After 24 hours, media was replaced either with fresh media without LIF, or with fresh media containing 1000 U/ml LIF (ESGRO, ESG1106). Cells were incubated for further 6 hours, and luciferase measurements were performed as described above. *TGF-β*: V6.5 mESC cells were

plated immediately prior to transfection onto gelatin-coated plates in media supplemented with 10μM SB431542 (STEMGENT) to suppress TGF-β signaling. Following transfection, cells were incubated for 24 hours. After 24 hours, media was replaced either with fresh media containing 10μM SB431542, or with fresh media containing 10ng/μl Activin A (R&D) to stimulate TGF-β signaling. Cells were incubated for further 6 hours, and luciferase measurements were performed as described above. *Wnt*: V6.5 mESCs were cultured in media containing 3μM IWP-2 (STEMGENT) for 24 hours prior to transfection to suppress Wnt signaling. Cells were then transfected either in media containing 3μM IWP-2 or in media containing 50ng/μl recombinant Wnt3a (R&D). Transfected cells were incubated for 24 hours, and luciferase measurements were performed as described above. The pathway manipulation for the gene expression assays displayed on Figure S3G was performed exactly as described above, except the Wnt pathway was stimulated by the addition of 3μM CHIR99021 (Selleckchem) instead of the Wnt3a ligand.

Note on the activity of the Wnt pathway in murine ESCs: Murine ES cells have low endogenous Wnt activity in our standard culture conditions. Under these conditions TCF3 recruits co-repressors and functions as a transcriptional repressor (Brantjes et al., 2001; Roose et al., 1998). When the Wnt pathway is further stimulated in these cells, a portion of the TCF3-occupied genes is activated (Cole et al., 2008). This activation occurs through accumulation of β-catenin in the nucleus where it displaces TLE/Groucho to activate transcription (Daniels and Weis, 2005). This type of experiment, where mESC cells were treated with Wnt3a conditioned medium, was used for the GSEA analysis in Figure 3D. Here, the β-catenin/TCF3 complexes that form as a result of exogenous Wnt-stimulation act as activators of transcription.

Genome editing

Genome editing was performed using CRISPR/Cas9 essentially as described (Wang et al., 2013). Briefly, target-specific oligonucleotides were cloned into a plasmid carrying a codon-optimized version of Cas9 (pX330, Addgene: 42230). The genomic sequences complementary to guide RNAs are listed in Supplemental Table 2. V6.5 murine ESCs were transfected with two plasmids expressing Cas9 and sgRNA targeting regions around 200 basepairs up- and down-stream of the center of the targeted SE-constituent

(as defined by OCT4/SOX2/NANOG co-binding; see below), respectively. A plasmid expressing PGK-puroR was also co-transfected. Transfection was carried out with the X-fect reagent (Clontech) according to the manufacturer's instructions. One day after transfection, cells were re-plated on DR4 MEF feeder layers. One day after re-plating puromycin (2µg/ml) was added for three days. Subsequently, puromycin was withdrawn for three to four days. Individual colonies were picked, and genotyped by PCR, and the edited alleles were verified by Sanger sequencing. All cell lines used in subsequent experiments were homozygous deletion lines. Reference and deletion allele sequences are listed in Supplemental Table 2.

RNA isolation and quantitative RT-PCR

ESC lines were split off MEFs for two passages. RNA was isolated using Trizol reagent (Invitrogen) or RNeasy purification kit (Promega), and reverse transcribed using oligo-dT primers and SuperScript III reverse transcriptase (Invitrogen) according to the manufacturers´ instructions. Quantitative real-time PCR was performed on a 7000 AB Detection System using the following Taqman probes, according to the manufacturer´s instructions (Applied Biosystems):
Gapdh: Mm99999915_g1
Prdm14: Mm01237814_m1
Mmu-mir-292b: Mm03307733_pri
Sik1: Mm00440317_m1

ChIP-Seq

ChIP was performed as in (Lee et al., 2006) with a few adaptations. Cells were crosslinked for 10 min at room temperature by the addition of one-tenth of the volume of 11% formaldehyde solution (11% formaldehyde, 50 mM HEPES pH 7.3, 100 mM NaCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0) to the growth media followed by 5 min quenching with 100 mM glycine. Cells were washed twice with PBS, then the supernatant was aspirated and the cell pellet was flash frozen in liquid nitrogen. Frozen crosslinked cells were stored at −80°C. 100ul of Dynal magnetic beads (Sigma) were blocked with 0.5% BSA (w/v) in PBS. Magnetic beads were bound with 10 µg of anti-H4K24Ac antibody (Abcam Ab-4729). Crosslinked cells were resuspended and

sonicated in lysis buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA pH 8.0, 0.1% SDS, and 1% Triton X-100). Cells were sonicated for 12 cycles at 30s each on ice (18-21 W) with 60 s on ice between cycles. Sonicated lysates were cleared and incubated overnight at 4°C with magnetic beads bound with antibody to enrich for DNA fragments bound by the indicated factor. Beads were washed with wash buffer A (50 mM HEPES-KOH pH7.9, 140 mM NaCl, 1 mM EDTA pH 8.0, 0.1% Na-Deoxycholate, 1% Triton X-100, 0.1% SDS), B (50 mM HEPES-KOH pH7.9, 500 mM NaCl, 1 mM EDTA pH 8.0, 0.1% Na-Deoxycholate, 1% Triton X-100, 0.1% SDS), C (20 mM Tris-HCl pH8.0, 250 mM LiCl, 1 mM EDTA pH 8.0, 0.5% Na-Deoxycholate, 0.5% IGEPAL C-630 0.1% SDS) and D(TE with 50 mM NaCl) sequentially. DNA was eluted in elution buffer (50 mM Tris-HCL pH 8.0, 10 mM EDTA, 1% SDS). Cross-links were reversed overnight. RNA and protein were digested using RNase A and Proteinase K, respectively and DNA was purified with phenol chloroform extraction and ethanol precipitation.

Illumina Sequencing and Library Generation

Purified ChIP DNA was used to prepare Illumina multiplexed sequencing libraries. Libraries for Illumina sequencing were prepared following the Illumina TruSeq DNA Sample Preparation v2 kit protocol with the following exceptions. After end-repair and A-tailing, Immunoprecipitated DNA (~10-50 ng) or Whole Cell Extract DNA (50 ng) was ligated to a 1:50 dilution of Illumina Adaptor Oligo Mix assigning one of 24 unique indexes in the kit to each sample. Following ligation, libraries were amplified by 18 cycles of PCR using the HiFi NGS Library Amplification kit from KAPA Biosystems. Amplified libraries were then size-selected using a 2% gel cassette in the Pippin Prep system from Sage Science set to capture fragments between 200 and 400 bp. Libraries were quantified by qPCR using the KAPA Biosystems Illumina Library Quantification kit according to kit protocols. Libraries with distinct TruSeq indexes were multiplexed by mixing at equimolar ratios and running together in a lane on the Illumina HiSeq 2000 for 40 bases in single read mode.

ChIP-Seq Data analysis

Accessing data used in this study. All ChIP-Seq, RNA-Seq and microarray data used in this study can be found online at the Gene Expression Omnibus (GEO)

(www.ncbi.nlm.nih.gov/geo/) or Arrayexpress (http://www.ebi.ac.uk/arrayexpress/). Accession numbers are listed in Supplemental Table 3.

Gene sets and annotations. All analysis was performed using RefSeq (NCBI37/MM9) mouse gene annotations, and RefSeq (GRCh37/hg19) human gene annotations.

Super-enhancers and typical enhancers. Super-enhancers and typical enhancers in mouse embryonic stem cells (ESCs) were previously identified (Whyte et al., 2013). Briefly, the intersection of the binding peaks for OCT4, SOX2, and NANOG was taken to identify constituent enhancers. Constituent enhancers were stitched together if they were within 12.5 kb, and these stitched enhancers were ranked by their ChIP-Seq read signal of MED1, using ROSE (https://bitbucket.org/young_computation/rose). Genomic co-ordinates of super-enhancers and typical enhancers in human sigmoid colon and HCT-116 colorectal cancer cells were downloaded from a previous study (Hnisz et al., 2013). Briefly, in these cells ChIP-Seq binding peaks of H3K27Ac were stitched together if they were within 12.5 kb, and these stitched enhancers were ranked by their signal of H3K27Ac using ROSE. H3K27Ac data was previously shown to be a satisfactory surrogate to identify SEs (Hnisz et al., 2013).

Calculating read density. The normalized read densities of ChIP-Seq datasets in mouse ESCs for OCT4, SOX2, NANOG, KLF4, ESRRB, PRDM14, NR5A2, TCF3, SMAD3, TCFCP2L1, c-MYC, CTCF, TBX3, TFE3, KAP1, YY1, ZFX, RONIN and respective input/control samples at super-enhancer and typical enhancer constituents were downloaded from a previous study (Hnisz et al., 2013). For STAT3, the normalized read density at super-enhancer and typical enhancer constituents was calculated as described (Hnisz et al., 2013; Lin et al., 2012). Briefly, ChIP-Seq reads aligning to the region were extended by 200bp and the density of reads per basepair (bp) was calculated using bamToGFF (https://github.com/BradnerLab/pipeline). The density of reads in each region was normalized to the total number of million mapped reads producing read density in units of reads per million mapped reads per bp (rpm/bp). These data were used for the correlation heat maps on Figure 3A. For the analysis of H3K27Ac density in wild type mESC and in cells where the E3 of the *Prdm14* was deleted (Figure 2D), the input-subtracted ready density values of the regions corresponding to +/-500bp around the SE-constituents were used.

Correlating transcription factor occupancy at super-enhancer or typical enhancer constituents. To determine quantitative relationships between transcription factor binding

at either super-enhancer or typical enhancer constituents read densities of the 20 transcription factors in ESCs were used. The read densities calculated in the respective input/control samples were subtracted from the values calculated in the ChIP-Seq samples. Pairwise correlations between transcription factor binding patterns were determined using a Pearson correlation statistic across either super-enhancer constituents (n=646) or typical enhancer constituents (n=9,981). Transcription factor binding patterns were clustered using a correlation score-derived distance metric and pairwise correlations were shaded according to correlation statistic (Figure 3A).

ChIP-Seq data processing. Mouse ChIP-Seq datasets Mouse ChIP-Seq reads were mapped to the mouse reference genome (mm9/NCBI37) using Bowtie (Langmead et al., 2009) with parameters –k 1 –m 1 –best –S. Human ChIP-Seq reads were mapped to the human reference genome (hg19/GRCh37) using Bowtie with parameters –k 2 –m 2 –best –S. For display of ChIP signal (Figure 1A-E, 1G, 2A-D, 4A-B, 4E, Figure S1C, S2B-F, S3B-F, S3H, S4B-C), WIG files were created using MACS (Model based analysis of ChIP-Seq) peak finding algorithm (Zhang et al., 2008) with parameters –w –S –space=50 –nomodel –shiftsize=200 and uploaded to the UCSC Genome Browser (Kent et al., 2002). We used the MACS version 1.4.2 to identify regions of ChIP-Seq enrichment over background for TCF3, SMAD3 and STAT3 in mouse ESCs. A p-value threshold of enrichment of 1e-6 was used for these datasets.

Calculating constituent overlap at super-enhancers and typical enhancers. To determine the overlap of super-enhancers and typical enhancers with signaling transcription factors, each super-enhancer or typical enhancer was first decomposed into constituent enhancers (as defined by OCT4/SOX2/NANOG co-binding (Whyte et al., 2013)). Binary overlap (>1bp) for TCF3, SMAD3 and STAT3 transcription factors was determined across all constituents and aggregated to define a binary binding vector for each super-enhancer or typical enhancer (Figure 3B). On Figure S2A, normalized read densities for TCF3, SMAD3, STAT3 at super-enhancer constituents and typical enhancer constituents are displayed as a heatmap. For this analysis the ChIP-Seq read density values were normalized against the mean read density at the union of all constituent enhancers in the respective datasets. Rows (constituents) were ranked by their binary vector of peak presence/absence in combinations of the signaling TFs. Simulated domains comprised of typical enhancer constituents were created for Figure 3B (in orange) to compare domains made of the same number of constituents as super-

enhancers. A distribution of constituents per super-enhancer domain was created using information from a previous study (Whyte et al., 2013), and domains of typical enhancer constituents were created following the same distribution of constituents per domain. Using the per-constituent list of signaling TFs bound, a per-simulated domain count of bound TFs was created. This process was iterated 10,000 times. The values plotted are the counts of simulated domains (10,000 x 231) converted to a percent of simulated domains bound by counts of TFs.

Motif analysis. Motif occurrences in murine ESC constituent enhancers of both super-enhancers and typical enhancers for TCF3, SMAD3, and STAT3 were identified using FIMO (Grant et al., 2011) as described previously (Hnisz et al., 2013). Briefly we identified all motif occurrences within super-enhancer constituents and typical enhancer constituents with a P-value threshold of $< 10^{-4}$ for STAT3 (M01595), TCF3 (M01594), SMAD3 (M00701). Motif enrichments in human and mouse ESC super-enhancer constituents were calculated using TRAP multiple sequences (Thomas-Chollier et al., 2011) (Figure 3C). Murine ESC motif analysis was previously described, using mouse promoters as the control set of regions (Hnisz et al., 2013; Whyte et al., 2013). For human ESC (H1), super-enhancers and SE-constituents were downloaded from a previous study (Hnisz et al., 2013). Briefly, we identified peaks of H3K27ac significantly enriched over control using MACS with parameters –p 1e-9 –keep-dup=auto. We intersected constituent regions with super-enhancer regions and filtered for those super-enhancer constituents under 1000 bases in length. Sequences from these regions were used in TRAP multiple sequences using the Transfac 2010.1 vertebrate motif library and human_promoters control. For CTCF motif M01259 was used. P-values displayed on Figure 3C refer to the enrichment of motifs vs the control set of regions (promoters in the respective species).

Assigning genes to enhancers. Super-enhancer to gene assignments and typical enhancer to gene assignments for mESC and HCT-116 colorectal cancer cells were downloaded from a previous study (Hnisz et al., 2013). Briefly, both in mESC and HCT-116 cells super-enhancers and typical enhancers were assigned to the nearest active gene. Active gene were determined using H3K27Ac ChIP-Seq read densities around the transcription start site of the respective gene (Hnisz et al., 2013).

Identifying super-enhancers acquired in colorectal cancer. To identify super-enhancers acquired in colorectal cancer cells (HCT-116) relative to sigmoid colon (Figure 4C),

these two sets of SEs were pooled. Specifically, regions unique to both sets were retained and regions contacting another were collapsed. The density of H3K27Ac in both samples was calculated and normalized to mapped reads using bamToGFF (https://github.com/BradnerLab/pipeline/blob/master/bamToGFF.py). The top 100 collapsed regions with the highest ratio of HCT-116 H3K27Ac signal vs. sigmoid colon H3K27Ac were considered acquired (Figure 4C). Acquired SEs in breast cancer were identified the same way.

Metagene analysis. Metagenes in Figure 4C and Figure 4F at the acquired super-enhancers were constructed in three parts: a 2000bp upstream of the region, a size-normalized collapsed region, and 2000bp downstream of the region. Read densities normalized to million mapped reads were calculated in 200 windows per part.

Gene expression analyses

RNA-Seq RPKMs were calculated for two replicates each of murine ESCs treated with LIF for 1h (E-MTAB-1796 Arrayexpress dataset (Martello et al., 2013)) . Reads for each replicate were aligned to the mm9 reference genome using Tophat2 (Trapnell et al., 2009) version 2.0.11, using Bowtie version 2.2.1.0 and Samtools version 0.1.19.0. RPKMs per Refseq transcript were calculated from aligned reads using RPKM_count.py from RSeQC (Wang et al., 2012). Fold-changes for +/-LIF conditions were calculated by averaging RPKMs for each condition for all transcripts with the same gene name, dividing the –LIF by the +LIF average RPKM (adding one pseudocount each), and transforming by log2. Gene expression changes after blocking TGF-β signaling by the inhibitor SB431542 were downloaded from a previous study (Mullen et al., 2011). Gene expression changes after stimulation of the Wnt pathway by Wnt 3a conditioned medium were downloaded from a previous study (Cole et al., 2008).

RNA-Seq data for HCT-116 is from Array Express experiment E-MTAB-651. For Wnt stimulation and blockage, the APC and CTNNB1 (β-catenin) knockdown datasets were used, respectively. Prior to alignment, files downloaded from Array Express were split into csfasta and qual files of phred quality using SeqIO from BioPerl (http://www.bioperl.org/). Reads for each replicate were aligned to the a colorspace version of the hg19 reference genome using Tophat2 (Trapnell et al., 2009) version 2.0.11, using Bowtie version 1.0.1.0 and Samtools version 0.1.19.0 and parameters –C

–bowtie1 --quals. RPKMs per Refseq transcript were calculated from aligned reads using RPKM_count.py from RSeQC (Wang et al., 2012). Fold-changes were calculated by averaging RPKMs for each condition for all transcripts with the same gene name, dividing the average RPKM in the control sample (adding one pseudocount each), and transforming by log2.

Gene set enrichment analysis (GSEA)

GSEA (Mootha et al., 2003) analyses were performed using the tool available at http://www.broadinstitute.org/gsea/index.jsp. In brief, fold change (log2) in gene expression from two experimental conditions were calculated and the list was then used as a ranked list in the Pre-Ranked function of the GSEA software. GSEA was performed using 1000 iterations collapsing probes to the highest value. Gene sets were based on the super-enhancer and typical enhancer –associated genes in the respective cell types.

SUPPLEMENTAL REFERENCES

Brantjes, H., Roose, J., van De Wetering, M., and Clevers, H. (2001). All Tcf HMG box transcription factors interact with Groucho-related co-repressors. Nucleic Acids Res *29*, 1410-1419.

Cole, M.F., Johnstone, S.E., Newman, J.J., Kagey, M.H., and Young, R.A. (2008). Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes & development *22*, 746-755.

Daniels, D.L., and Weis, W.I. (2005). Beta-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. Nature structural & molecular biology *12*, 364-371.

Dowen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K.*, et al.* (2014). Control of cell identity genes occurs in insulated neighborhoods in Mammalian chromosomes. Cell *159*, 374-387.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017-1018.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-Andre, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell *155*, 934-947.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome research *12*, 996-1006.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nature protocols *1*, 729-748.

Lin, C.Y., Loven, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A. (2012). Transcriptional amplification in tumor cells with elevated c-Myc. Cell *151*, 56-67.

Loven, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell *153*, 320-334.

Martello, G., Bertone, P., and Smith, A. (2013). Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. The EMBO journal *32*, 2561-2574.

Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E.*, et al.* (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nature genetics *34*, 267-273.

Mullen, A.C., Orlando, D.A., Newman, J.J., Loven, J., Kumar, R.M., Bilodeau, S., Reddy, J., Guenther, M.G., DeKoter, R.P., and Young, R.A. (2011). Master transcription factors determine cell-type-specific responses to TGF-beta signaling. Cell *147*, 565-576.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nature genetics *24*, 372-376.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell *141*, 432-445.

Roose, J., Molenaar, M., Peterson, J., Hurenkamp, J., Brantjes, H., Moerer, P., van de Wetering, M., Destree, O., and Clevers, H. (1998). The Xenopus Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. Nature *395*, 608-612.
Thomas-Chollier, M., Hufton, A., Heinig, M., O'Keeffe, S., Masri, N.E., Roider, H.G., Manke, T., and Vingron, M. (2011). Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. Nature protocols *6*, 1860-1869.
Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.
Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell *153*, 910-918.
Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. Bioinformatics *28*, 2184-2185.
Whyte, W.A., Bilodeau, S., Orlando, D.A., Hoke, H.A., Frampton, G.M., Foster, C.T., Cowley, S.M., and Young, R.A. (2012). Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. Nature *482*, 221-225.
Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307-319.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome biology *9*, R137.