Stem Cell Reports, Volume *8*

Supplemental Information

An Embryonic Stem Cell-Specific NuRD Complex Functions through Interaction with WDR5

Ly-Sha Ee, Kurtis N. McCannell, Yang Tang, Nancy Fernandes, W. Rod Hardy, Michael R. Green, Feixia Chu, and Thomas G. Fazzio

SUPPLEMENTAL INFORMATION

SUPPLEMENTAL FIGURES

Figure S1. MBD3C contains a unique 50-amino acid N-terminus and forms an ESC-specific NuRD complex, related to Figure 1. (A) Left: Schematic of the three MBD3 isoforms in ESCs. MBD3A contains a methyl-binding domain (MBD) that is truncated in MBD3b and absent in MBD3C. The purple box signifies the unique MBD3 N-terminal 50 amino acids. Right: Western blot of MBD3 in ESCs and MEFs. Actin serves as a loading control. **(B)** Western blots of NuRD subunits from purified H3F-WDR5 complexes in WT, *Mbd3c* KO or *Mbd3abc* KO ESCs. **(C)** Schematic of the N-terminal *Mbd3c* DNA and MBD3C protein sequences determined by 5'RACE. Exons are indicated by numbered blue boxes, introns by connecting black lines. The sequence of *Mbd3* intron 2 is shown, with the *Mbd3c* N-terminus in red. The amino acid sequence of the MBD3C N-terminus derived from intron 2 is shown (represented by the gray box in the *Mbd3c* gene). **(D)** Table of peptide counts from mass spec analysis of individually FLAG-tagged MBD3 isoforms. **(E)** Glycerol gradient analysis of *Mbd3-H3F* nuclear extracts.

Figure S2. Characterization of MBD3 isoform KO ESC lines, related to Figure 3. (A) Western blot of MBD3 from lysates of the indicated mouse tissues. Actin blot and Coomassie stain are shown as loading controls. **(B)** Western blot of MBD3 expression in MBD3 isoform KO ESCs generated through CRISPR/Cas9. **(C-D)** Growth curves for *Mbd3c* KO **(C)** and *Mbd3ab* and *Mbd3abc* KO **(D)** ESC lines, relative to wildtype (WT) ESCs. Error bars represent +/- standard deviation of three replicate experiments performed on each clonal replicate. **(E)** Western blots of indicated proteins during differentiation of WT, *Mbd3abc* KO ESCs overexpressing *Mbd3c-H3F*, or *Mbd3abc* KO ESCs overexpressing *Mbd3cΔN-H3F.* **(F)** Western blots of indicated proteins during differentiation of WT or *Mbd3cΔN–ab* KO ESCs.

Figure S3. Images of dox-induced reprogrammed colonies, related to Figure 3. (A) Representative images of EOS-EGFP positive/OSK-mCherry negative iPSCs at reprogramming Day 30, 10 days after dox removal (left) and an iPSC line derived from a single colony (right). Scale bars = 400μm. **(B)** Representative AP staining of iPSCs at reprogramming Day 31.

Figure S4. Chromatin extraction assay, related to Figure 4. (A) Western blot of MBD3 from salt fractionation of chromatin from WT ESC nuclei. Protein released at indicated concentrations of NaCl is shown. RNA Pol II is included as a control.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cell Culture and Generation of ESC lines

Murine ESCs are derived from E14 and cultured on gelatin-coated plates as previously described (Chen et al., 2013). *Mbd3a-H3F*, *Mbd3c-H3F*, and *Mbd3cΔN-H3F* ESCs were generated by infection of E14 with pLJM1 lentiviral vectors carrying the respective constructs.

The H3F-WDR5 targeting construct was made by inserting PCR-produced homology arms (959 and 561 bp) and an H3F tag into pBluescript II SK+ (Stratagene). Oligos were inserted into pX330-puro^R to target to the N-terminus of WDR5. The plasmids were transfected with FuGENE HD (Promega) into E14, *Mbd3c* KO, and *Mbd3abc* KO cells. Clones were selected with puromycin and screened by PCR, sequencing, and western blot.

Expression of MBD3C and WDR5 domain mapping mutants

MBD3C-H3F, V5-WDR5 constructs and the indicated WDR5 mutants (Yang et al., 2014) were cloned into pCAGGS-IRES-Hygro^R using the 5' XhoI site and the 3' EcoRI site. pCAGGS-V5-WDR5 was created from a synthesized full-length mouse *Wdr5*. MBD3A-H3F and truncations of MBD3C-H3F with XhoI and MfeI restriction sites were derived by PCR on *Mbd3*-H3F sequences and cloned into the vector. The MBD3C Δ41-50 and R43A mutant constructs were made by PCR on pCAGGS-MBD3C-H3F with primers incorporating the mutations and flanking primers and then digesting the new sequences and ligating them into the vector. pCAGGS–cN41-50–MBD3A-H3F was made by synthesizing the N-terminus and inserting it into the MBD3A plasmid using the XhoI site and a gene-internal BamHI site. Plasmids were transiently transfected into 293T cells using FuGENE HD (Promega), and the cells were harvested for IP after 2 days.

Generation of MBD3 isoform KO ESCs

Mbd3ab, *Mbd3c*, *Mbd3abc*, and *Mbd3cΔN-ab* KO ESC lines were generated using the CRISPR/Cas9 system (Cong et al., 2013) to introduce mutations into *Mbd3* exon 2 (*Mbd3ab* KO and *Mbd3cΔN−ab* KO), intron 2 (*Mbd3c* KO), or exon 5 (*Mbd3abc* KO). Guide RNAs targeting *Mbd3ab*, *Mbd3c*, or *Mbd3abc* (see Table S3 for sequences) were cloned into the pX330-puro^R vector and transfected into E14 ESCs as described (Hainer et al., 2015). Individual clones were screened by TOPO cloning (Invitrogen) and sequencing to verify the presence of homozygous frameshift mutations. ESC lines were further screened by western blot to verify loss of the appropriate MBD3 isoform(s). To create the *Mbd3cΔN−ab* KO line, we first transfected the pX330 plasmid targeting *Mbd3c* into E14 ESCs along with a donor plasmid containing the *Mbd3c* coding sequence with a 50-amino acid N-terminal deletion. The donor plasmid was generated by annealing oligos to make the *Mbd3cΔN* cDNA construct and cloning along with ~2kb homology arms flanking the *Mbd3c* start site into pBluescript SK II+ (Stratagene). Verified *Mbd3cΔN* ESC lines were then retargeted using the *Mbd3ab* KO guide RNA plasmid and screened as described above.

Mbd3abc KO/Mbd3c o/e ESC lines were made by replacing the endogenous *Mbd3* locus with a construct containing PCR-produced homology arms (2070 and 1649 bp), the full CAGGS promoter, and a modified *Mbd3c*-H3F-polyA sequence without CpGs in the gene or tag. The *Mbd3abc* KO – cΔN o/e construct was made using PCR and restriction digestion to delete the 150 bp corresponding to the unique N-terminus of MBD3C. Oligos were inserted into two CRISPR plasmids (pX330-puro^R) and transfected into E14 cells as described above. Clones were selected with puromycin and screened by PCR, sequencing, and western blot.

5'RACE
5'RACE was performed on 4µg total RNA using the 5'RACE System Version 2.0 kit (Invitrogen). See Table S3 for primer sequences.

MBD3 Purification, WDR5 Purification, and LC-MS/MS
MBD3/NuRD complex was purified from MBD3-H3F, MBD3A-H3F, MBD3C-H3F, and MBD3CΔN-H3F ESCs as described (Yildirim et al., 2011). Purified samples were separated by SDS-PAGE, stained with SimplyBlue SafeStain (Invitrogen), and LC-MS/MS was performed as described in (Chen et al., 2013). For WDR5 complex purification, nuclear fractions were isolated from H3F-WDR5 ESCs using the NE-PER kit (Thermo), diluted 1:3 in MVL buffer (50mM Tris, pH 7.5; 250mM NaCl; 0.1% Triton X-100), and purified similarly, omitting the His purification step.

Western Blotting and Immunoprecipitation
Western blots were performed with the following antibodies: anti-MBD3 (Bethyl A302-528A and Abgent AM2203B), anti-FLAG (M2, Sigma F1804), anti-WDR5 (Bethyl A302-429A and A302-430A), anti-OCT4 (Santa Cruz sc-8628), anti-NANOG (Bethyl A300-398A), anti-MTA1 (Bethyl A300-911A), anti-MTA2 (Santa Cruz sc-28731), anti-CHD4 (Bethyl A301-082A) anti–β-ACTIN (Sigma A1978), anti–RNA POLYMERASE II (Santa Cruz sc-899), anti-RBBP4 (Bethyl A301-206A), anti-RBBP7 (Bethyl A300-958A), anti-LSD1 (Bethyl A300-215A), anti-p66α (Bethyl A302-358A), anti-p66β (Bethyl A301-281A), anti-HDAC1 (Bethyl A300-713A), anti-HDAC2 (Bethyl A300-705A), anti-ASH2L (Bethyl A300-112A), anti-V5 (Invitrogen 46-0705). Mouse tissue lysates were prepared by homogenizing indicated tissues in lysis buffer (50mM Tris-HCl pH 7.5; 150mM NaCl; 0.5% Triton X-100; 5% glycerol; 1mM PMSF). For IP, nuclear lysates were prepared using the NE-PER kit (Thermo). FLAG IP was performed as described in (Chen et al., 2013). MBD3 IPs in ESCs and 293T cells were performed similarly, except washes were performed in MVL buffer + 1mM EDTA.

Glycerol Gradient Analysis
Nuclear extracts were prepared from *Mbd3-H3F* ESCs using the NE-PER kit (Thermo). Nuclear extract (1470µg) was diluted in MVL buffer and spun in a 10-40% glycerol gradient at 37krpm for 17 hours in a Beckmann L-90K ultracentrifuge. 29 fractions were collected and odd fractions western blotted with the indicated antibodies.

Embryoid Body Differentiation
ESCs were differentiated to embryoid bodies (EBs) in suspension culture. 2.5 x 10^6 cells were plated in ESC medium without LIF in bacteriological plates. Cells were replated to non-gelatinized cell culture plates after 3 days and harvested for western blots at the indicated timepoints.

Chromatin Extraction Assay
Chromatin was extracted from WT ESCs by salt fractionation as described (Henikoff et al., 2009). Briefly, 4 x 10^7 cells were pelleted, washed in PBS, and resuspended in TM2 buffer (10mM Tris-HCl pH 7.5, 2mM $MgCl_2$, 1.5% NP-40, 1x HALT protease inhibitor

cocktail) on ice for 5min. Cells were pelleted and the nuclei incubated in TM2 + 70mM NaCl for 2h at 4°C. Nuclei were pelleted and incubated as before in TM2 + 140mM NaCl, and further pelleted and incubated in TM2 + 600mM NaCl overnight at 4°C. Supernatants from each incubation were saved as 0, 70,140 and 600mM NaCl fractions, clarified by centrifugation at full speed, and western blotted with the indicated antibodies.

Reprogramming to pluripotent stem cells

Lentiviral plasmids for dox-inducible Oct4, Sox2, and Klf4 expression (pLenti-Tet-OSK-mCherry) and for L-Myc expression (pLenti-Tet-L-Myc) were generated by digestion of CMV-OSK and CMV-L-Myc cDNA from FUW-OSKM and pMXs-Ms-L-Myc (Addgene 20328 and 26023 respectively) and cloning into pcDNA3.1 with HIV1-based 5' and 3' LTRs from pGIPZ. To package lentivirus, 293T cells were transfected with 5µg pLenti-Tet-OSK-mCherry, pLenti-Tet-L-Myc, FUW-rTtA (Addgene 20342), or EOS-EGFP (Addgene 21313) lentiviral reporter plasmids along with packaging plasmids (5µg psPAX2 and 2.5µg pCMV-VSV-G (Addgene)). Primary MEFs were infected with day 2 viral supernatant using 8µg/mL hexadimethrine bromide (Sigma), and re-infected after 24 hours. MEFs were replated in ESC media after 48 hours and induced with 2µg/ml dox 4 days after the first infection. Dox was removed at day 20 and cells were cultured for an additional 10 days. On Day 30 cells were imaged to assess loss of OSK transgenes and EOS-EGFP reporter activation, and on day 31 were stained for alkaline phosphatase according to kit instructions (Millipore). Single colonies were picked on Day 30, expanded, and western blotted with the indicated antibodies. Media were changed every other day.

Bisulfite pyrosequencing

WT ESCs were subjected to embryoid body differentiation and harvested at the indicated timepoints. Briefly, genomic DNA was phenol-chloroform extracted from cells incubated in ES cell lysis buffer (10 mM Tris, pH 7.5; 10 mM EDTA; 10 mM NaCl; 0.5% sarkosyl) with 1 µg/µL proteinase K at 55°C overnight. The DNA was bisulfite converted using the EpiTect Bisulfite Kit (QIAGEN). Primers were designed for the *Mbd3c* CpG island using PyroMark Q24 software (QIAGEN, see Table S3 for sequences), with one PCR primer in a pair biotinylated, and PCR was performed on the converted DNA with KAPA HiFi HotStart Uracil+ ReadyMix (Kapa Biosystems). PCR products were bound to streptavidin sepharose beads (GE Healthcare) and sequenced using a PyroMark Q24 (QIAGEN). Data were analyzed using PyroMark Q24 software.

RNA-seq

Total RNA was isolated from two biological replicate WT, *Mbd3ab* KO, *Mbd3c* KO, *Mbd3abc* KO, and *Mbd3cΔN−ab* KO ESC lines using TRIzol (Life Technologies), and purified with the Zymo RNA Clean and Concentrator kit. 2µg RNA was used for library preparation. RNA was rRNA-depleted (NEB and Clontech) and strand-specific libraries were prepared using the TruSeq Stranded mRNA LT kit (Illumina).

RNA-seq data analysis

Reads were mapped to the mouse mm10 genome with TopHat2 (Kim et al., 2013), using parameters --library-type fr-firststrand --segment-length 38. Mapped reads were processed in HOMER (Heinz et al., 2010) using the “analyzeRepeats” command to calculate raw counts and normalized reads per kilobase per million mapped reads (rpkm) for each gene. Differential gene expression was calculated with DESeq2 (Love et al., 2014) using the “getDiffExpression” command in HOMER. Genes with an adjusted p value < 0.05 and log2 (fold change) were considered significantly changed.

To map WDR5 and Mof binding at TSSs of significantly changed genes, WDR5 (Ang et al., 2011) and Mof (Li et al., 2012) ChIP-seq data were downloaded from GEO (GSE22934 and GSE37268) and aligned to the mouse mm10 genome using Bowtie (Langmead et al., 2009). Mapped reads were processed in HOMER using the “annotatePeaks” command. For Mof ChIP-seq, 2 replicate libraries were averaged in the aggregation plot.

To map WDR5 at MBD3-bound DNase I hypersensitive sites (DHSs), peaks were called from MBD3 ChIP-seq data (Yildirim et al., 2011; GSE31690) using the HOMER “findPeaks” command, and MBD3-bound DHSs were identified using the “mergePeaks” command with peaks called from mouse ENCODE DHSs (GSM1014154) without TSSs. The WDR5 ChIP-seq library was aligned to the MBD3-bound DHS peak data using the “annotatePeaks” command.

SUPPLEMENTAL TABLES

Table S1, related to Figure 1. Proteins identified in LC-MS/MS of endogenous MBD3-H3F ESCs.

Table S2, related to Figure 1. Proteins identified in LC-MS/MS of pLJM1- MBD3A-H3F, MBD3C-H3F, and MBD3CΔN-H3F ESCs.

Table S3, related to Supplemental Experimental Procedures. Oligonucleotides used in this study.

SUPPLEMENTAL REFERENCES

Ang, Y.-S., Tsai, S.-Y., Lee, D.-F., Monk, J., Su, J., Ratnakumar, K., Ding, J., Ge, Y., Darr, H., Chang, B., et al. (2011). Wdr5 Mediates Self-Renewaland Reprogramming via the Embryonic Stem Cell Core Transcriptional Network. Cell *145*, 183–197.

Chen, P.B., Hung, J.-H., Hickman, T.L., Coles, A.H., Carey, J.F., Weng, Z., Chu, F., and Fazzio, T.G. (2013). Hdac6 regulates Tip60-p400 function in stem cells. eLife *2*, e01557.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Hainer, S.J., Gu, W., Carone, B.R., Landry, B.D., Rando, O.J., Mello, C.C., and Fazzio,

T.G. (2015). Suppression of pervasive noncoding transcription in embryonic stem cells by esBAF. Genes Dev. *29*, 362–378.

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

Henikoff, S., Henikoff, J.G., Sakai, A., Loeb, G.B., and Ahmad, K. (2009). Genome-wide profiling of salt fractions maps physical properties of chromatin. Genome Res. *19*, 460–469.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

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.

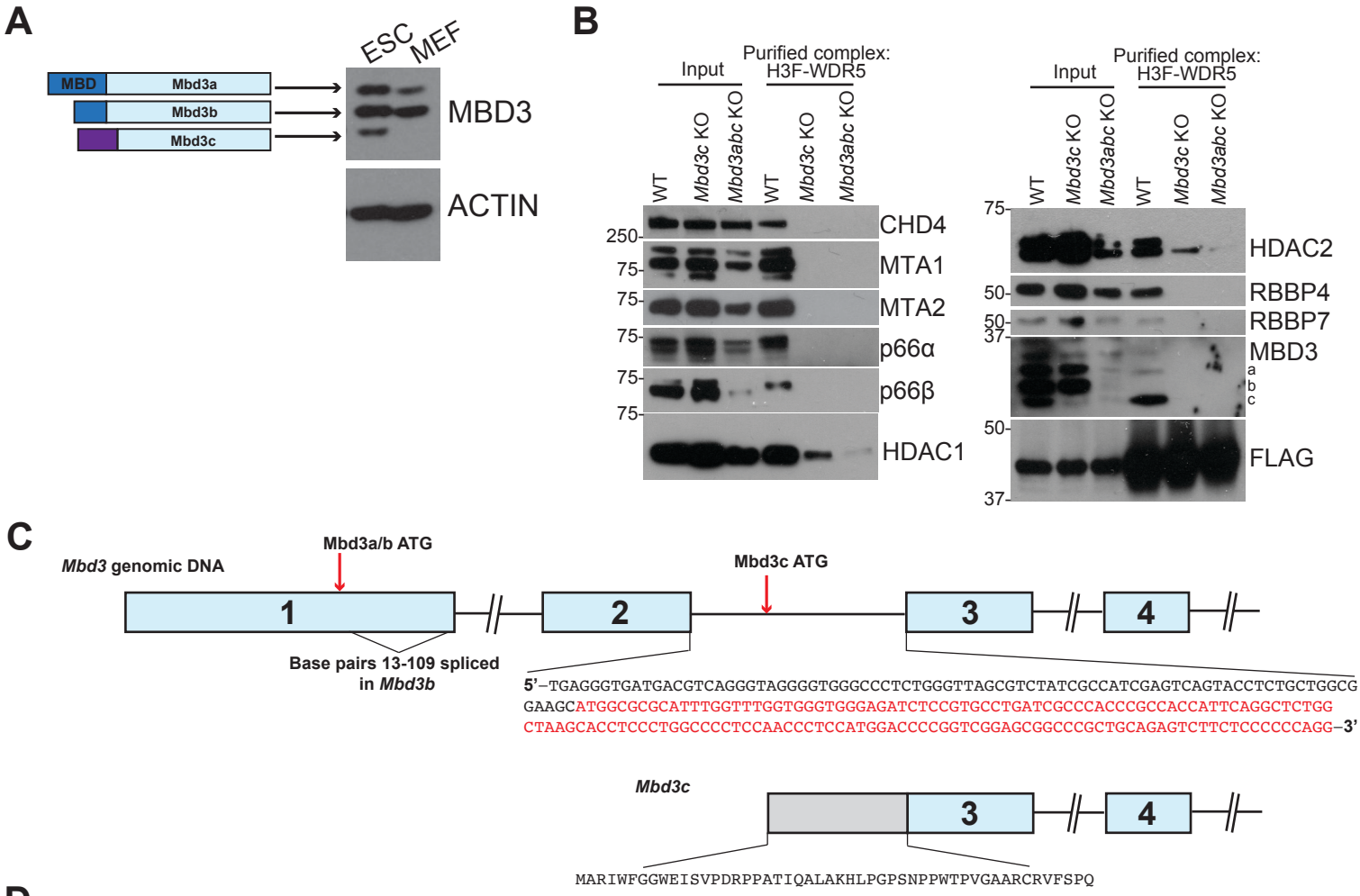
Li, X., Li, L., Pandey, R., Byun, J.S., Gardner, K., Qin, Z., and Dou, Y. (2012). The histone acetyltransferase MOF is a key regulator of the embryonic stem cell core transcriptional network. Cell Stem Cell *11*, 163–178.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Yang, Y.W., Flynn, R.A., Chen, Y., Qu, K., Wan, B., Wang, K.C., Lei, M., and Chang, H.Y. (2014). Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. eLife *3*, e02046.

Yildirim, O., Li, R., Hung, J.-H., Chen, P.B., Dong, X., Ee, L.-S., Weng, Z., Rando, O.J., and Fazzio, T.G. (2011). Mbd3/NURD Complex Regulates Expression of 5-Hydroxymethylcytosine Marked Genes in Embryonic Stem Cells. Cell *147*, 1498–1510.

Figure S1



D

Peptide count			NuRD complex protein
MBD3A-H3F	MBD3C-H3F	MBD3CΔN-H3F	
322	291	365	Chromodomain-helicase-DNA-binding protein 4
396	503	493	Metastasis-associated protein MTA2
382	203	425	Metastasis-associated protein MTA1
381	375	445	Metastasis-associated protein MTA3
428	527	674	Transcriptional repressor p66-alpha
343	414	543	Transcriptional repressor p66-beta
268	271	308	Histone deacetylase 1
193	183	213	Histone deacetylase 2
213	208	320	Methyl-CpG-binding domain protein 3
200	250	280	Histone-binding protein RBBP4
163	157	180	Histone-binding protein RBBP7
2	157	3	WD repeat-containing protein 5

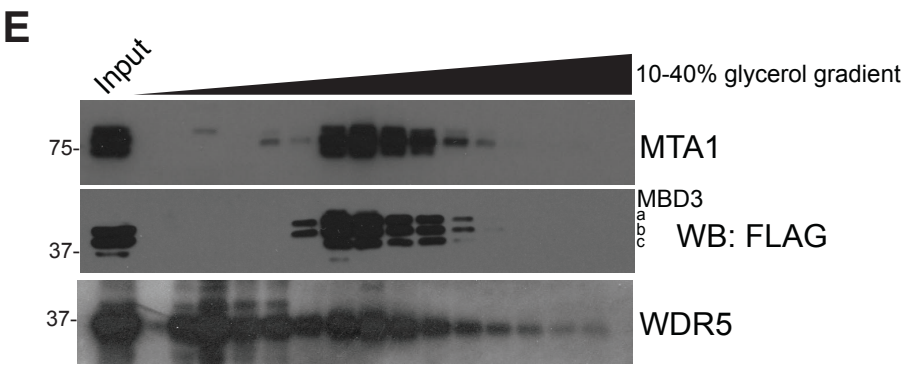


Figure S2

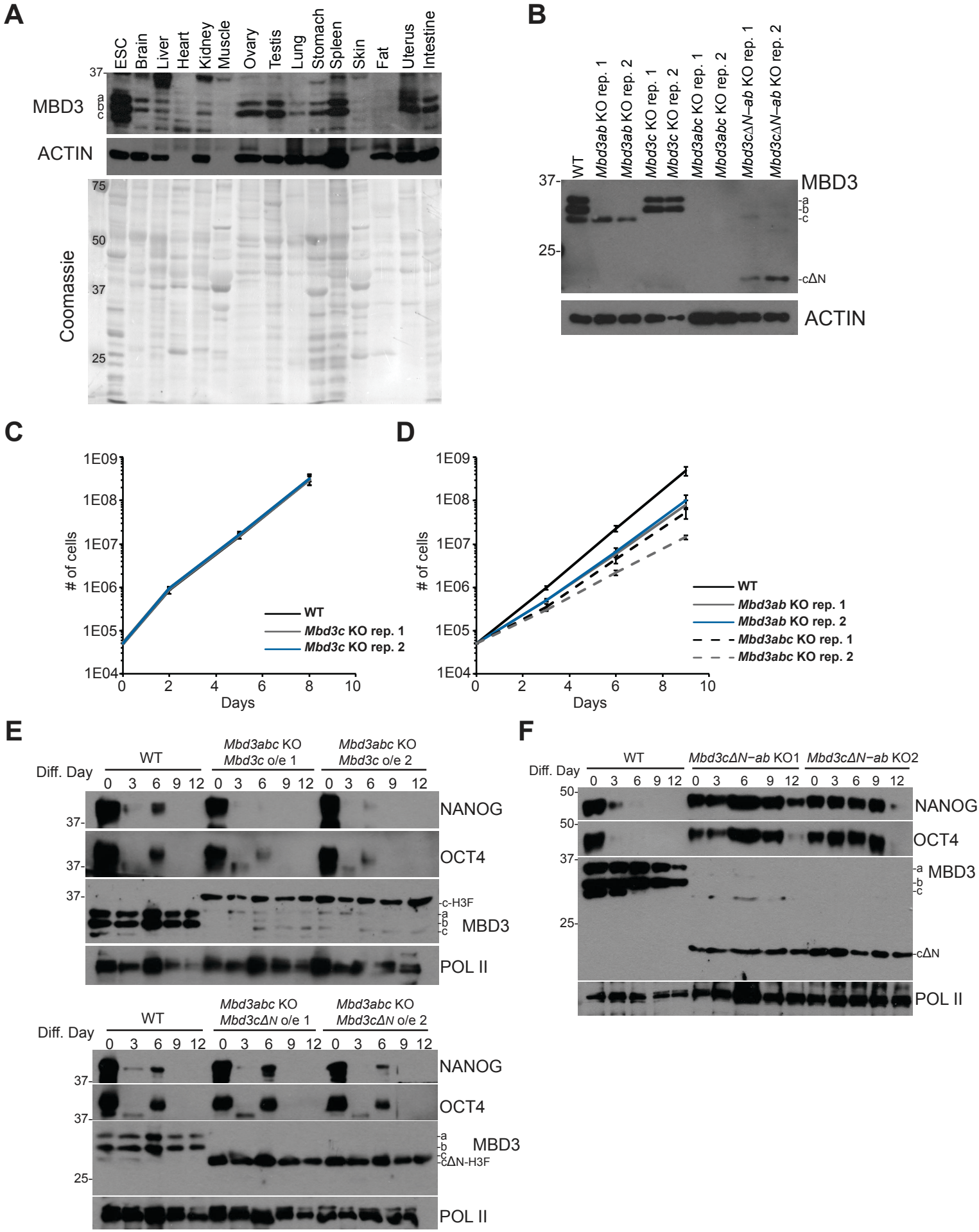

A
ESC
Brain
Liver
Heart
Kidney
Muscle
Ovary
Testis
Lung
Stomach
Spleen
Skin
Fat
Uterus
Intestine
37
MBD3
a
b
c
ACTIN
75
50
37
25
Coomassie
B
WT
Mbd3ab KO rep. 1
Mbd3ab KO rep. 2
Mbd3c KO rep. 1
Mbd3c KO rep. 2
Mbd3abc KO rep. 1
Mbd3abc KO rep. 2
Mbd3cΔN–ab KO rep. 1
Mbd3cΔN–ab KO rep. 2
37
25
MBD3
a
b
c
cΔN
ACTIN
C
1E09
1E08
1E07
1E06
1E05
1E04
# of cells
0
2
4
6
8
10
Days
WT
Mbd3c KO rep. 1
Mbd3c KO rep. 2
D
1E09
1E08
1E07
1E06
1E05
1E04
# of cells
0
2
4
6
8
10
Days
WT
Mbd3ab KO rep. 1
Mbd3ab KO rep. 2
Mbd3abc KO rep. 1
Mbd3abc KO rep. 2
E
Diff. Day
WT
Mbd3abc KO Mbd3c o/e 1
Mbd3abc KO Mbd3c o/e 2
0 3 6 9 12 0 3 6 9 12 0 3 6 9 12
37
NANOG
37
OCT4
37
c-H3F
a
b
c
MBD3
POL II
Diff. Day
WT
Mbd3abc KO Mbd3cΔN o/e 1
Mbd3abc KO Mbd3cΔN o/e 2
0 3 6 9 12 0 3 6 9 12 0 3 6 9 12
37
NANOG
37
OCT4
37
25
a
b
c
MBD3
cΔN-H3F
POL II
F
Diff. Day
WT
Mbd3cΔN–ab KO1
Mbd3cΔN–ab KO2
0 3 6 9 12 0 3 6 9 12 0 3 6 9 12
50
NANOG
50
OCT4
37
25
a
b
c
MBD3
cΔN
POL II

Figure S3

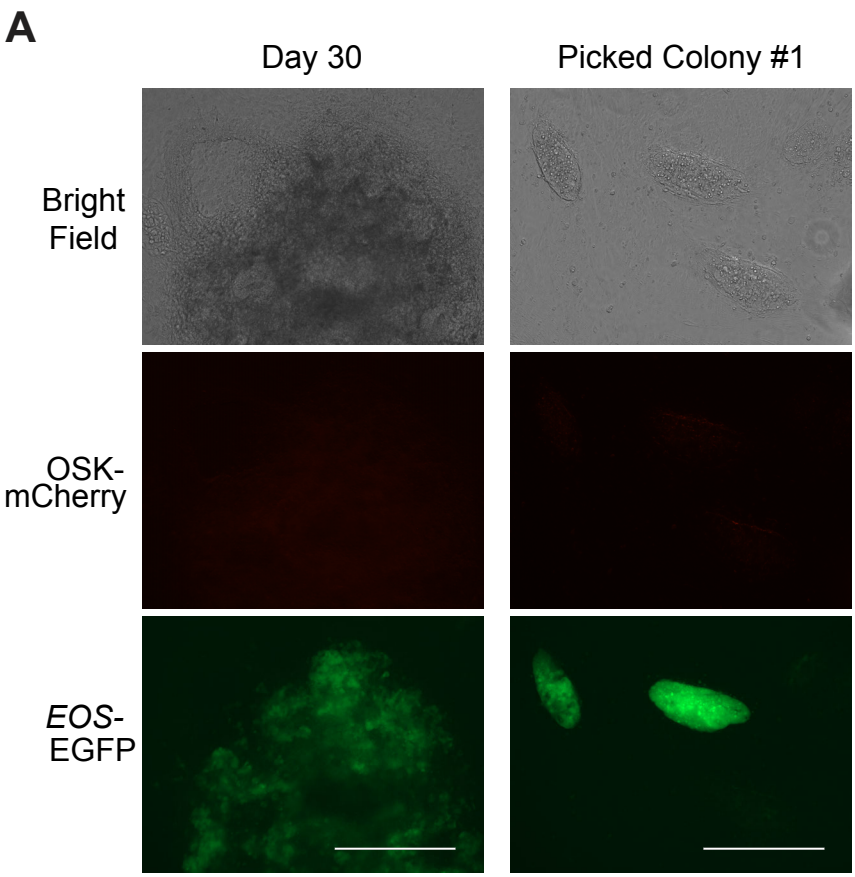

A
Day 30
Picked Colony #1
Bright Field
OSK-mCherry
EOS-EGFP
B
Day 31
Alkaline phosphatase

Figure S4