Antagonistic Chromatin Remodeling Complexes, Mbd3 and Brg1, Regulate a Set of Common Target Genes in Embryonic Stem Cells

A Major Qualifying Project
submitted to the Faculty of
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfillment of the requirements for the
degree of Bachelor of Science

by
Ruowang Li

Date:
Apr 28, 2011

Report Submitted to:

Professor Elizabeth Ryder
Worcester Polytechnic Institute

Contents

Acknowledgment 3
Abstract 3
Introduction 5
Embryonic Stem Cells 5
Chromatin regulation 6
Nucleosome Remodeling Factors 8
Study Goals 10
Results 11
Brg1 and Mbd3 oppositely regulate a set of common genes 11
Genome-wide mapping of Mbd3 14
Interaction of Mbd3 and Brg1 in ES cells 16
Characterization of Mbd3 binding 18
Mbd3 and Brg1 regulate the recruitment of RNA Pol II 21
Methods 24
Generating aggregation plot from ChIP-Seq data 24
Microarray data analysis 26
Enrichment of factor binding at transcription factor binding sites 27
Discussion 28
Reference 29

Acknowledgment

I would like to acknowledge and extend my gratitude to the following persons who have made the completion of this project possible:

Dr.Zhiping Weng

Dr.Elizebeth Ryder

Dr.Oliver Rando

Dr.Thomas Fazzio

Abstract

Two chromatin remodeling factors Mbd3 and Brg1 affect the pluripotency in mouse embryonic stem (ES) cells. Knock down of the two factors showed that they oppositely affect a common set of genes. Genome wide mapping of Mbd3 and Brg1 using ChIP-Seq showed that the two factors preferentially bind to promoter regions and physically interact in ES cells. Mbd3 binding is found to be enriched at Polycomb and bivalent target genes marked with H3K4me3 and H3K27me3. Higher level of Mbd3 binding also correlated with lower gene expression in mouse ES cells.

Introduction

Embryonic stem cell research is one of the most important research areas in biology because of the unique characteristics of these cells. In human, traditional disease treatments such as organ transplant or chemotherapy have significant draw backs, such as the lack of organ donors, and sometimes harm the patient's body. Embryonic stem cell therapy has the potential to avoid these problems because the cells have the ability to develop into any cell types. Due to ethical reasons surrounding the use of human embryonic stem cells for research, mouse embryonic stem cells are the next best model for such research.

A long-standing goal in biology is to be able to understand and control the ability of cells to retain a stem cell character vs. the ability to differentiate. These states are known to be affected by a number of protein complexes, such as transcription factors, chromatin remodelers, and histone modifiers. The goal of our study was to better understand the relationships between these various complexes as well as gaining new insight into two chromatin modelers, Mbd3 and Brg1.

Embryonic Stem Cells

Mouse embryonic stem (ES) cells are a widely used model for biological research. The cells are obtained from the inner cell mass of the mouse blastocyst, which has the ability to develop into any cell type. (Figure 1) (Openbio)

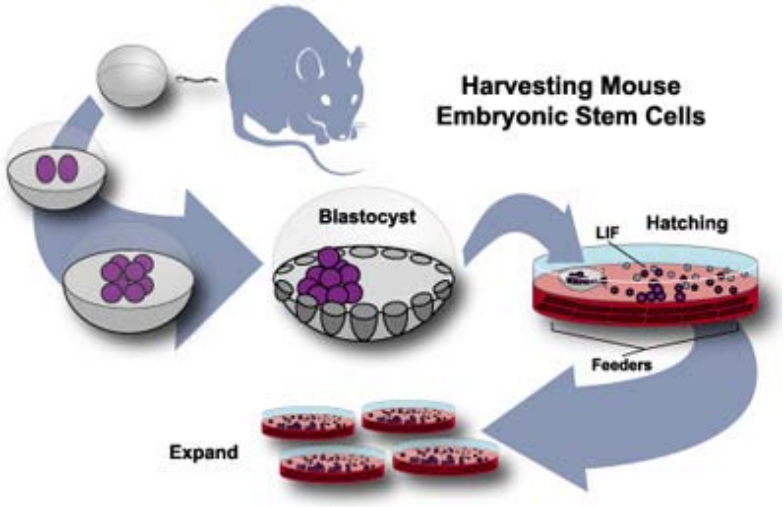

Figure 1. **Schematic of the ES cell harvesting**

Because of the pluripotent property of the ES cells, these cells have unique transcription regulation and chromatin modification compared to somatic cells. In human and mouse ES cells, pluripotency is maintained by a network of transcription factors. In general, transcription factors are DNA binding proteins that bind to regulatory sequences of genes that regulate a cell's ability to self-renew (maintain its population) or differentiate. (Boyer, 2005;Loh, 2006 ;Chen, 2008 ;Kim, 2008). In mouse, several transcription factors such as Oct4, Sox2, and Nanog, are the "master regulators" of the ES cell that allow it to retain pluripotency. (Nichols, 1998;Niwa, 2000; Avilion, 2003;Chambers, 2003;Mitsui, 2003).

Chromatin regulation

Many chromatin regulators also play essential roles in ES cell gene regulation, self-renewal, and differentiation. As shown in Figure 2, DNA is organized in the cell by histones, small proteins composed mostly of arginine and lysine. A nucleosome is formed by DNA wrapped around eight histone proteins that is 146 nucleotide pairs long. Further folding of nucleosomes by other proteins results in the compact chromatin structure (Lewis, 2002)

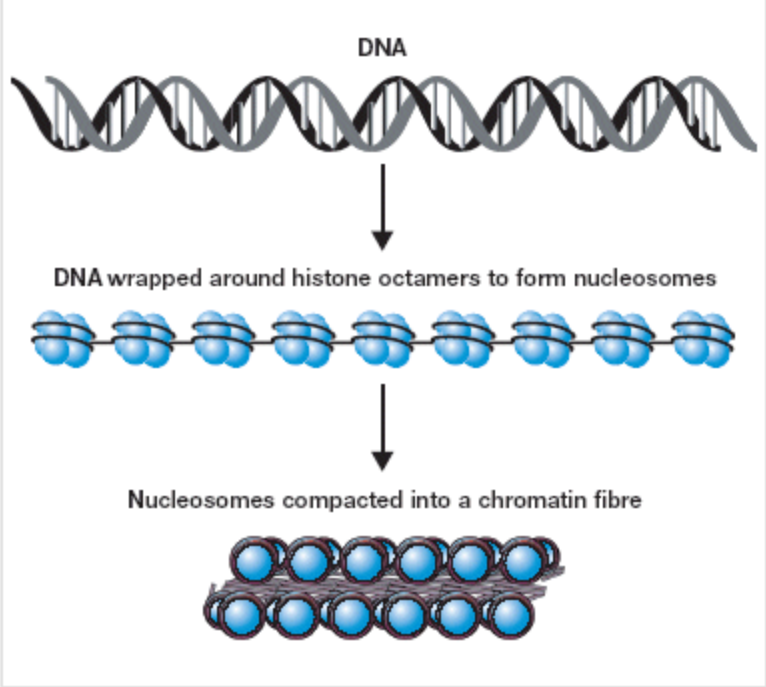

Figure 2. **The structure of chromation** (Abcam)

Histone methylation describes the process in which methyl groups are added to the histones, whereas acetylation is the addition of an acetyl group (Figure 3) (Rice, 2001). These modifications typically have specific effects on transcription of nearby genes; for example, H3K4me3 marks genes that are up regulated and H3K27me3 marks genes that are down regulated. A number of protein complexes that catalyze histone methylation, acetylation, and ubiquitylation, as well as proteins catalyzing removal of those marks, also have essential roles in ES cells (Niwa, 2007; Surface, 2010).

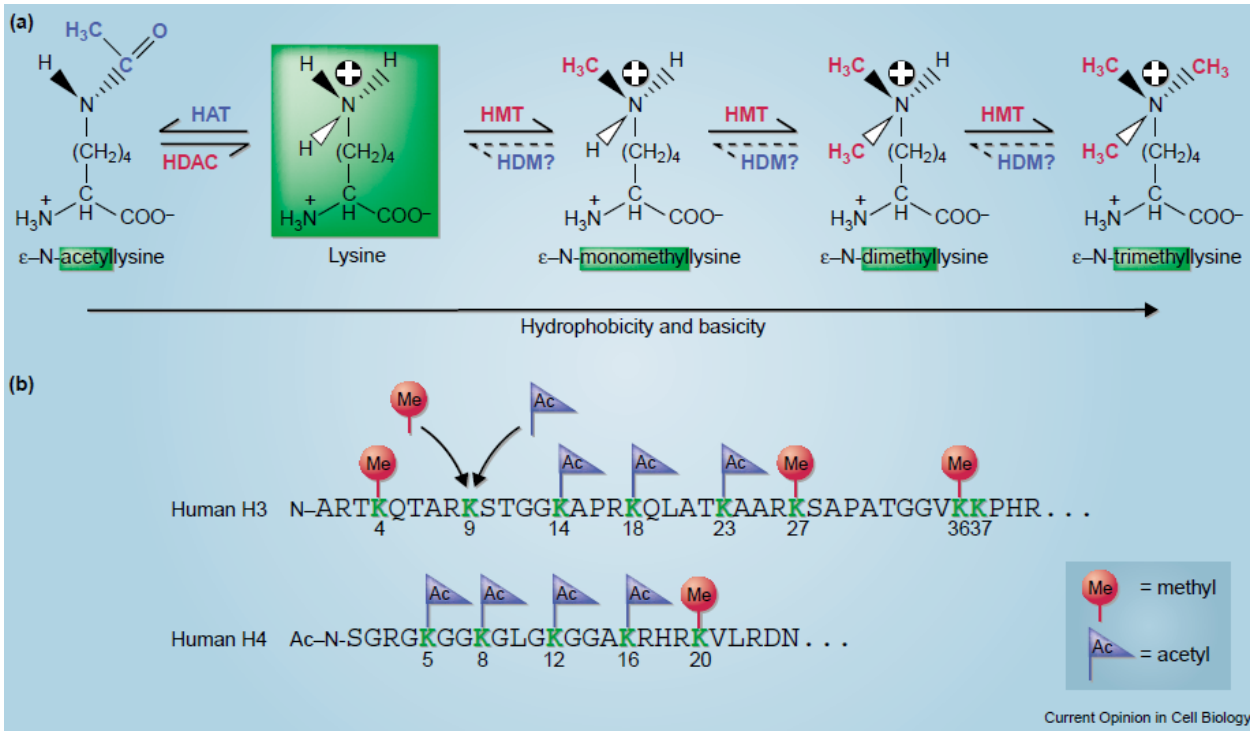

Figure 3. **Mechanism of histone methylation and acetylation** (a). Addition of methyl group or acetyl group to Lysine residue (b) Example of histone methylation and acetylation on Human histone H3 and H4

Mouse ES cells have an unusual bivalent histone modification pattern compared with somatic cells. Near the promoters of developmentally regulated genes, they contain both an activating mark histone H3 trimethylated at lysine 4 (H3K4me3) and a repressive mark histone H3 trimethylated at lysine 27 (H3K27me3) (Bernstein, 2006). The chromatin modifying complexes, MLL/SET1 Complex and Polycomb Repressive Complex 2 (PRC2), which catalyze the addition of these histone marks, respectively, have important roles in development (van Lohuizen, 1998).

Nucleosome Remodeling Factors

Several ATP dependent nucleosome remodeling factors, which use energy harvested from ATP hydrolysis to modify the spatial and compositional features of the nucleosomes, also

have essential roles in ES cells (Fazzio, 2010; Keenen, 2009). BAFs (Brahma/Brg1 Associated Factor) are a family of related ATP-dependent nucleosome remodeling factors which function to both activate and repress transcription by remodeling nucleosomes near promoters. They are homologous to yeast SWI/SNF factors (de la Serna, 2006; Trotter, 2008; Racki, 2008). The predominant form of the BAF complexes is namded esBAF (Ho, 2009) and is shown in Figure 4.

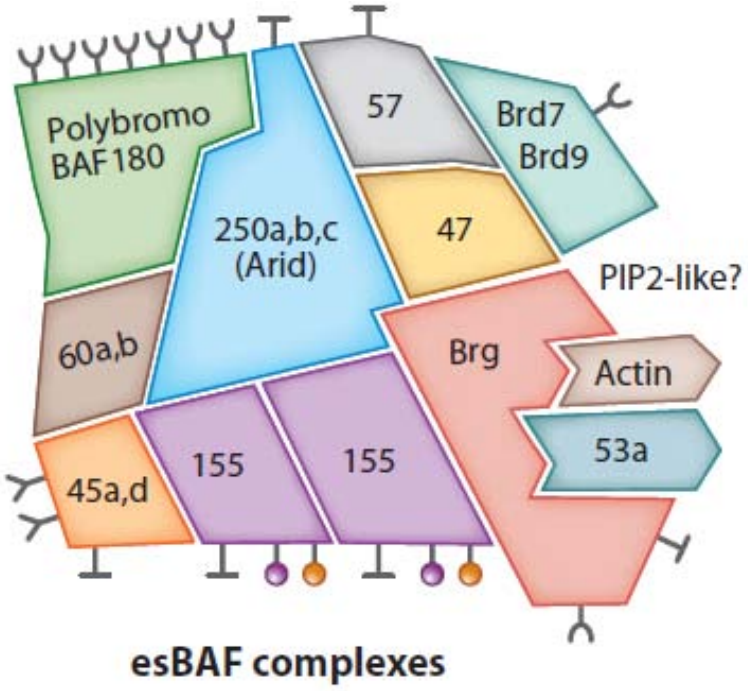

Figure 4. **Composition of esBAF nucleosome remodeling factor** (Lessard, 2010)

Homozygous knockout (KO) or knockdown (KD) of several BAF subunits induces deficiency in ES cell self-renewal and pluripotency, indicating their crucial roles in ES cell gene expression pattern (Gao, 2008; Ho, 2009; Kidder, 2009; Yan, 2008).

In addition, NURD (Nucleosome Remodeling and Deacetylase) is also an ATP dependent chromatin remodeling factor which creates repressive chromatin structure in vitro and in vivo (Denslow, 2007).The NURD complex is also found in humans, and the structure is shown in

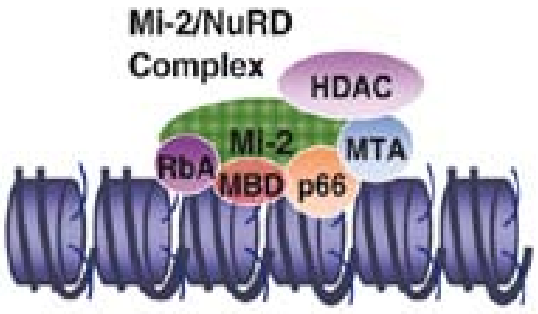

Figure 5. **Human NuRD complex structure** (Denslow, 2007)

KD or homozygous KO of the NURD subunit Mbd3 in ES cells results in defects in differentiation and changed developmental potency (Kaji, 2006; Kaji, 2007; Zhu, 2009). Despite the fact that chromatin modelers have been identified in previous studies, the genes that are regulated by these factors are unknown. The relationship between the transcription factors and chromatin modelers, such as Brg1 and Mbd3, has not been studied.

Study Goals

In a previous RNA interference experiment, it was found that chromatin regulators Tip60 and p400 are required for ES self-renewal and pluripotency. Two other factors, Ash2l and Suz12, are required for H3K4me3 and H3K27me3 catalysis, respectively. In this study, we would like to understand the interactions of these four chromatin regulators in addition to Mbd3 and Brg1. A genomic-scale genetic interaction assay of these six factors identified a set of common targets of the NURD and BAF complexes. Many of the identified common target genes are important for ES cell self renewal or differentiation. In this study, we identified the localization of Mbd3 and

Brg1 using ChIP-Seq technology. We determined that these two factors physically interact in ES cells, and that Brg1 facilitates the binding of Mbd3. In addition, we found that Mbd3 is enriched at Polycomb and bivalent genes, and that high Mbd3 binding correlates with a higher gene expression level.

Results

The following work is the result of a highly collaborative effort. In order to have a complete picture for the study, some sections of the results include work solely done by Dr.Oliver Rando, as indicated in the figure legends. The rest of the work was done collaboratively.

Brg1 and Mbd3 oppositely regulate a set of common genes

Previous experiments using genome wide RNA interference to knock down chromatin regulatory factors in mouse ES cells identified 68 genes that have altered self-renewal, pluripotency or viability (Fazzio, 2010). Knockdown of subunits of Tip60, p400 or BAF complexes have shown deficiencies in ES cell self-renewal and pluripotency. The Tip60-p400 complex requires the activating mark H3K4me3 for binding to its targets. The complex colocalizes with genes marked by this modification and genes with bivalent histone marks that includeH3K4me3 and the repressing mark H3K27me3. Interestingly, Ash2l and Suz12 genes are required for catalysis of H3K4me3 and H3K27me3, respectively, but their KD did not show similar deficiencies as the Tip60-p400 complex KD. Other KD of genes, such as Mbd3 KD, has shown alterations in cell morphology, but had no similar effect on self-renewal as the KD of theTip60-p400 complex.

In order to understand the relationship and identify the targets of the six chromatin regulators, Tip60, p400, Mbd3, Brg1, Ash2l, and Suz12, changes in mRNA level for each KD or double KD of these factors were profiled. KD involving Tip60 or p400 had similar gene expression changes (Figure 6).

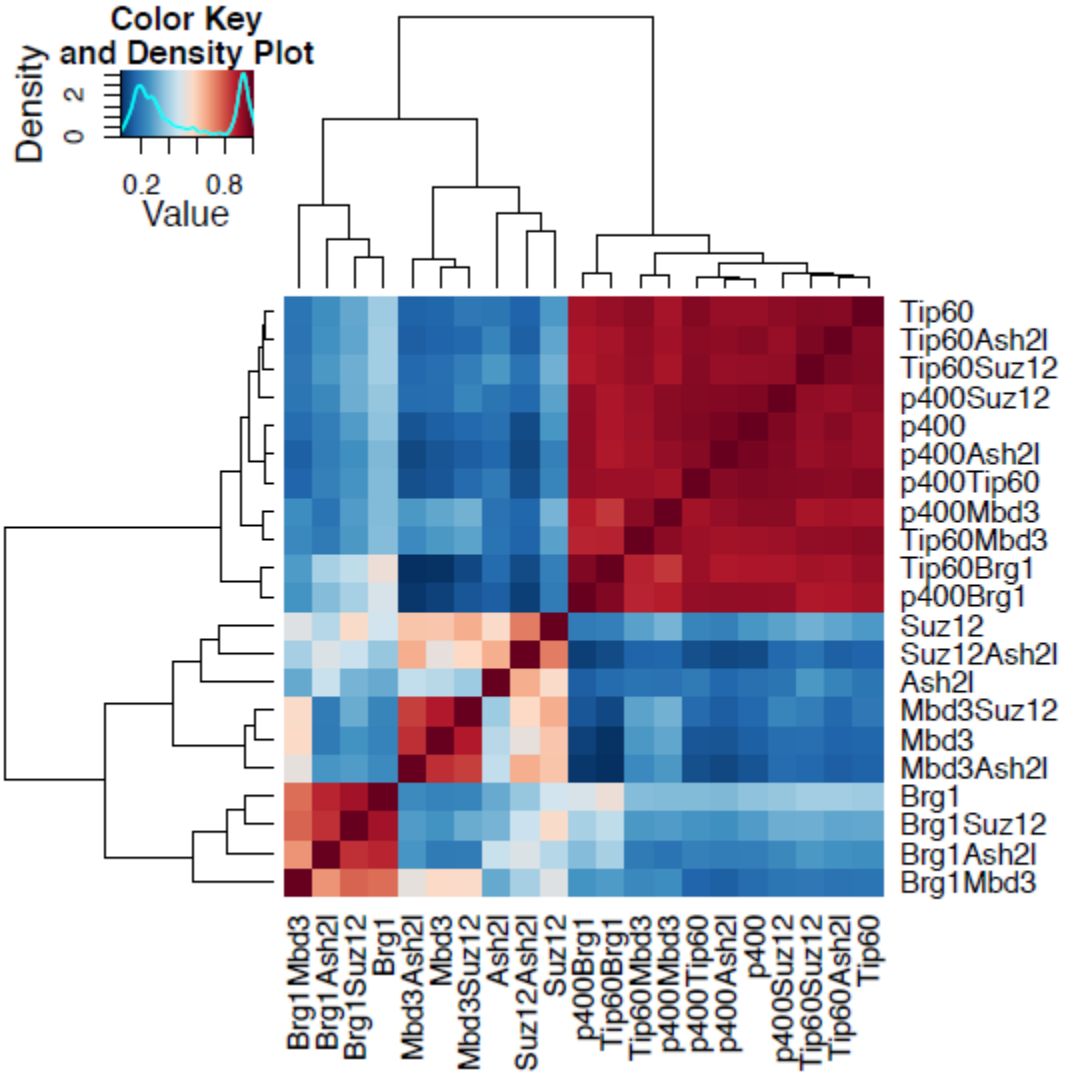

Figure 6. **Correlation of expression between 21 cell conditions.** The expression levels of single and double KD of Brg1, Mbd3, Ash2l, Suz12, Tip60, and P400 were correlated. The Pearson correlation coefficients for each correlation pair are shown in the above heat map. Red: more correlation Blue: less correlation. Rando and Fazzio, unpublished results.

The Mbd3 KD has similar gene expression profile with Ash2l KD and Suz12KD, suggesting that these three factors might share similar regulatory targets. Using principal component analysis (PCA) method, it was determined that Tip60 KD or p400 KD caused the biggest changes in the overall expression profile. Mbd3 KD and Brg1(BAF subunit) KD affected the same principle component but in opposite directions, implying that these two factors antagonistically regulate the gene expression of some common targets, and the overlap of genes affected by Brg1 KD or Mbd3 KD was statistically significant ($p<2.2 \times 10^{-16}$). In addition, a group of genes were upregulated in Mbd3 KD and down regulated in Brg1 KD, and a smaller set of genes exhibit the converse pattern (Figure 7).

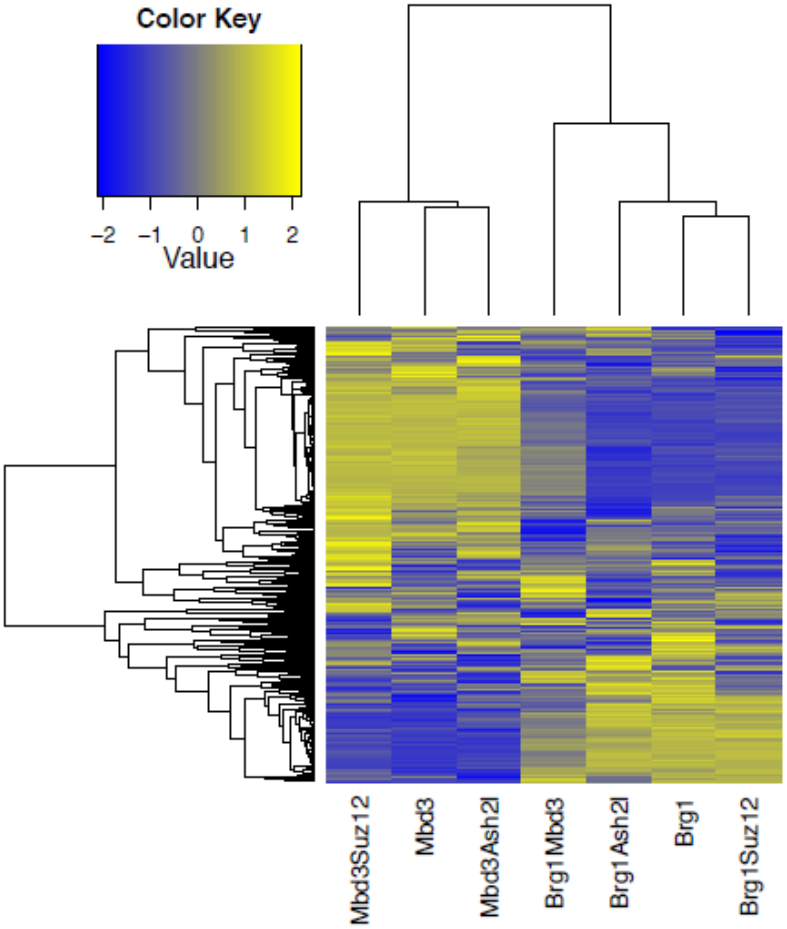

Figure 7. **Mbd3 and Brg1 oppositely regulate a set of genes.** The expression levels after KD of combinations of Mbd3, Brg1, Ash2l and Suz12 in ES cells were compared to wild type. Yellow: genes upregulated. Blue: genes downregulated. Transparent: unaffected genes. Rando and Fazzio, unpublished results.

Next, double KD of Brg1 and Mbd3 was performed in ES cells. The gene expression profile that resulted from the double KD showed a more wild type profile compared to either single KD, suggesting that these two factors may antagonistically regulate their common target genes.

Genome-wide mapping of Mbd3

Genome wide mapping of Brg1 factor has been done in ES cells; however, the locations of Mbd3 binding are unknown. In order to investigate the localization and Brg1 and Mbd3 factors, ChIP-Seq was carried out for Mbd3 in murine ES cells. Three point six million uniquely mapped reads for Mbd3 were generated. As shown in Figure 8, Mbd3 showed a strong preference for binding around the promoter regions in the ES cell; the peak of the binding was slightly downstream of the transcription start site (TSS). Higher levels of Mbd3 binding were seen in genes that were upregulated in Mbd3 KD compared with unaffected genes (Figure 9). This result is consistent with Mbd3 normally acting to repress the bound genes. Also shown in Figure 8 is the previous mapping of Brg1 around the TSS. The binding sites of the two factors have extensive overlap. It should also be noted that the binding of Mbd3 is downstream of the binding of Brg1, suggesting that the two complexes have an orientation preference with respect to the direction of transcription.

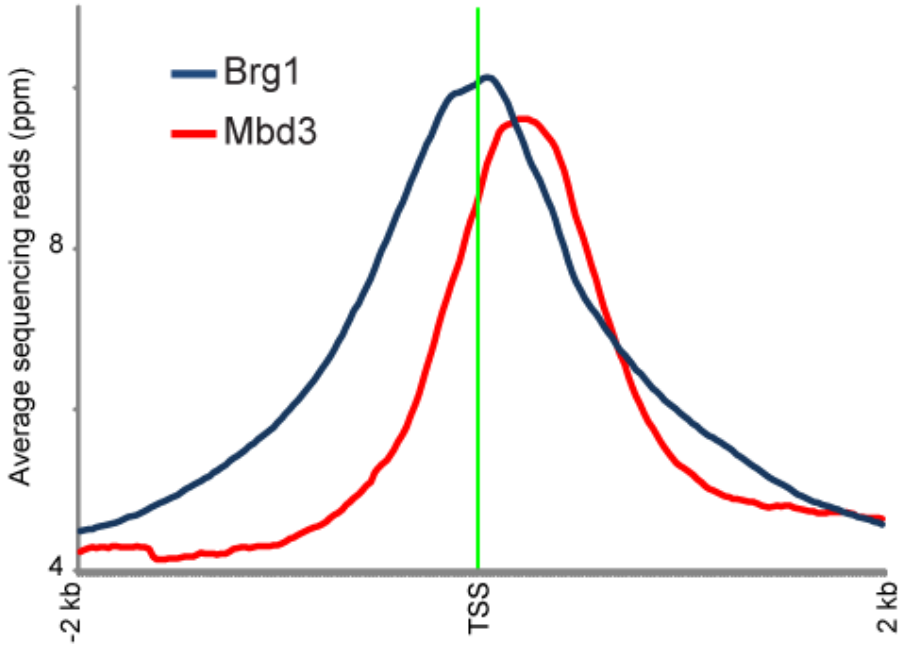

Figure 8. **Localization of Mbd3 at the promoter region**. Mbd3 was mapped to the genome in murine ES cells by ChIP using Mbd3 antibody followed by Solexa deep sequencing. The average binding profile of Mbd3 and Brg1 is shown around the all Transcription Start Sites (TSS) provided by the RefSeq.

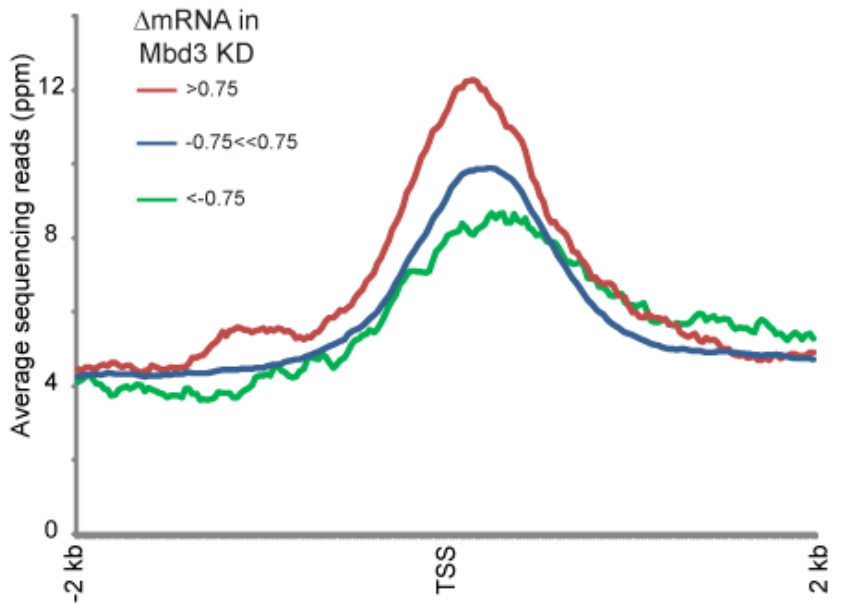

Figure 9. **Mbd3 binding level at upregulated, unaffected, and down regulated gene targets in Mbd3 KD ES cell.** All Genes were separated into three groups according to the change of mRNA level in Mbd3 KD ES cell. The average binding profiles of Mbd3 for all the genes in these three groups were shown. Red: upregulated genes. Blue: unaffected genes. Green: down regulated genes.

Interaction of Mbd3 and Brg1 in ES cells

In order to see if Brg1 and Mbd3 physically interact in vivo, IP-Westerns were used to pull down one factor using an antibody to the other factor. Immunoprecipitation using anti-Mbd3 antibody immunoprecipitated Brg1 (Figure 10), indicating the physical interactions between the two factors. In addition, Brg1 involvement in Mbd3 binding was tested. Genome wide mapping of Mbd3 in Brg1 KD ES cells was performed (Figure 11) and showed a loss of binding of Mbd3 at the promoter region. This result could suggest that Mbd3's binding, which constitutes its functionality, depends on the presence of Brg1. However, this inference seems to be opposite to the observations that down regulation of a set of genes in Brg1KD cells was returned to normal after further KD of Mbd3 (Figure 7), which implies that Mbd3 is capable of functioning without Brg1.This result could be explained by a model where Mbd3 can only achieve steady state binding with the presence of Brg1, but in the absence of Brg1, Mbd3 can still loosely bind to the genes to perform its function. Taken together, the results have shown that Brg1 and Mbd3 have a direct association and regulate a set of common genes in ES cells.

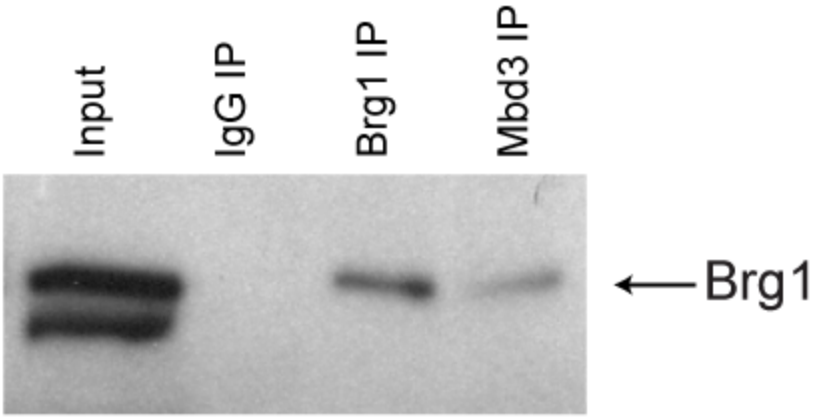

Figure 10. **Physical binding of Mbd3 and Brg1 in ES cells.** Western blot for Brg1 after immunoprecipitation with shown antibodies Lane input: ES cell lysates. Lane IgG: IgG antibody. Lane Brg1: Brg1 antibody. Lane Mbd3: Mbd3 antibody Rando, unpublished results.

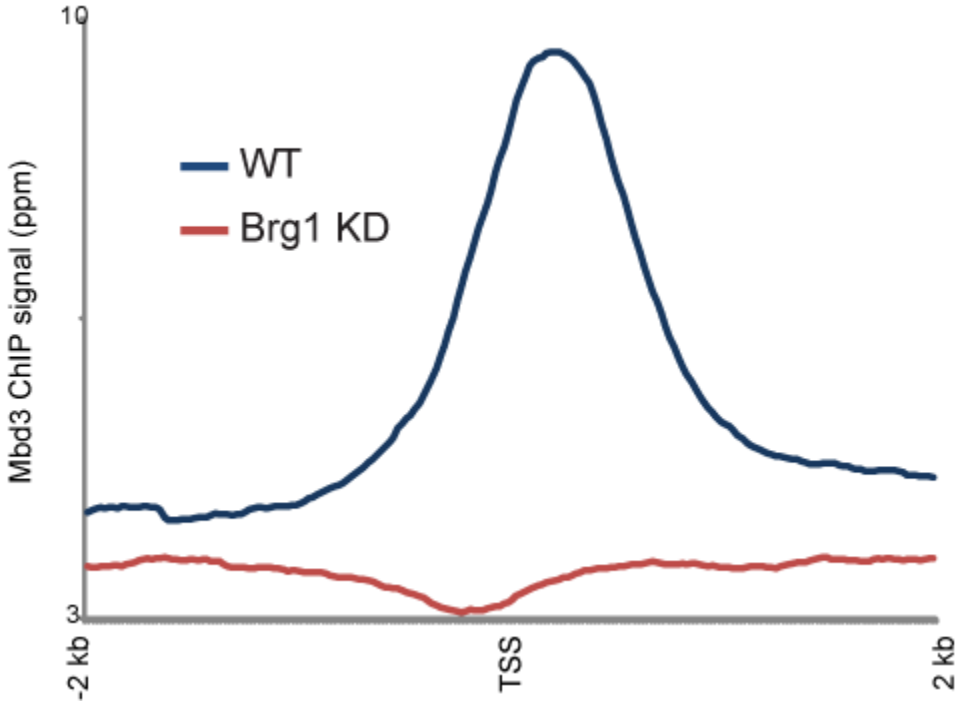

Figure 11. **Loss of Mbd3 binding in Brg1 KD ES cell** Mbd3 binding profile in Brg1 KD was generated as in Figure 6. WT line is the Mbd3 binding profile shown in Figure 8. Brg1 KD line is the Mbd3 binding profile in Brg1 KD ES cell. Mbd3 binding was lost in Brg1 KD ES cells.

Characterization of Mbd3 binding

To see the relationship of Mbd3 binding profile with the known transcriptional network, the binding profile was compared with previously published ChIP-Seq datasets for transcription factors and chromatin regulators. The difference in Mbd3 binding level between genes bound and unbound by a factor indicated on the x-axis was shown in Figure 12. The highest enrichment of Mbd3 binding was seen at the targets of the Polycomb proteins Ezh1, Suz12, Phc1, and Eed, as well as the bivalent histone modifications H3K27me3 catalyzed by PRC2 complex and H3K4me3 catalyzed by MLL/SET1 complexes.

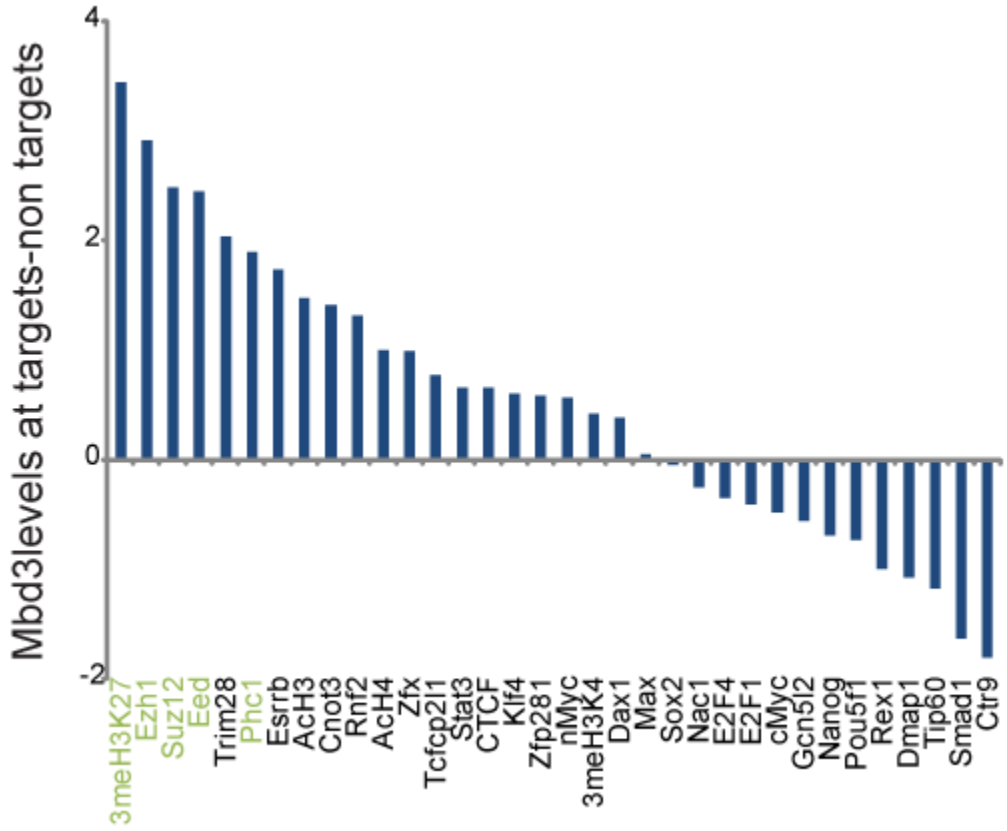

Figure 12. **Mbd3 binding is enriched at Polycomb and bivalent targets.**

Mbd3 binding level was calculated at the target sites of 34 factors shown on the x-axis. The binding level was determined by the sum of the ChIP signal of a 1 kb window centered on the

+200 position relative to the binding sites. The differences of Mbd3 binding level was obtained by subtracting the Mbd3 binding level at genes bound and genes not bound by these factors. The histogram was sorted with the highest enrichment of Mbd3 binding to the left.

The enrichment of Mbd3 at these targets could be due to the NURD complex function together with Polycomb or MLL/SET1 complexes or both to regulate bivalent gene expression. As shown in Figure 6, the gene expression of Mbd3 KD ES cells had a similar profile with Suz12 KD ES cells and Ash2l KD ES cells. This could suggest the similarity between H3K4me3 and H3k27me3 in regulating bivalent genes as Ash2l and Suz12 are required for the catalysis of those two histone marks, respectively and Mbd3 could also regulate these genes. Most genes up or down regulated in Mbd3 KD were also changed in Suz12 KD or Ash2l KD (Figure 13).

It is known that Polycomb complexes have functions such as repressing genes that are expressed during development. In order to see if Mbd3 factor has a similar mechanism, the Mbd3 binding data was compared with gene expression profiles during ES cell development into embryoid bodies. As shown in Figure 14, genes bound significantly by Mbd3 were upregulated during development, suggesting the repressing role on the “poised” developmental genes.

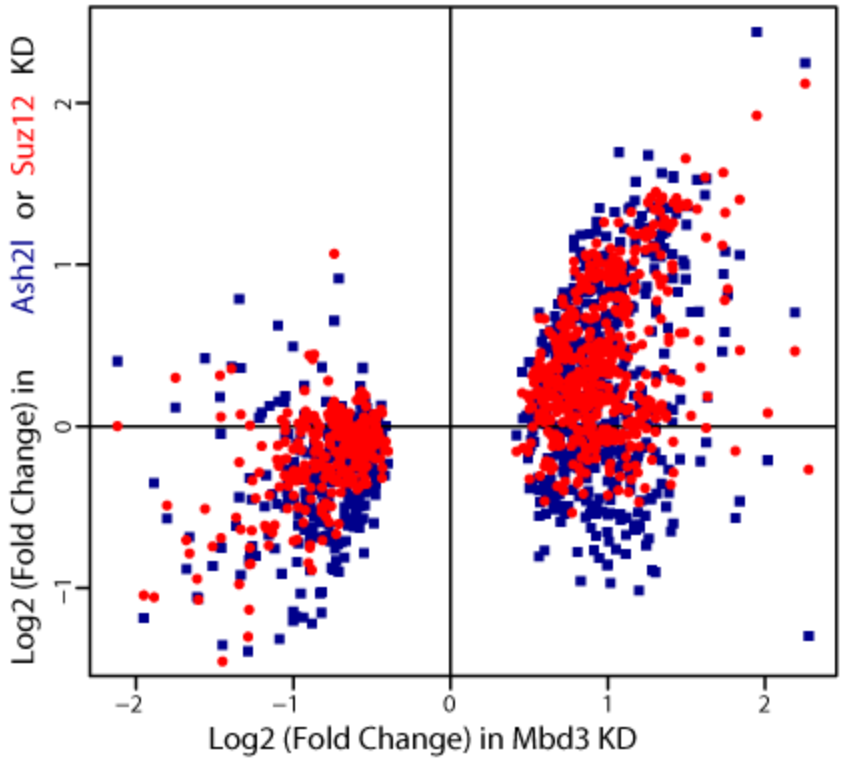

Figure 13. **Genes affected by Mbd3 KD have similar change in Suz12 KD or Ash2l KD** Dr. Oliver Rando, unpublished result

Scatterplot of Mbd3 KD gene expression vs Ash KD and Suz12 KD in log2(fold change). Blue: gene expression change in Ash2l KD. Red: gene expression change in Suz12 KD. A group of genes clustered in the first quadrant was upregulated in Mbd3 KD as well as in Ash2l KD and Suz12 KD. A second group of genes clustered in the third quadrant was down regulated in Mbd3 KD as well as in Ash2l KD and Suz12 KD.

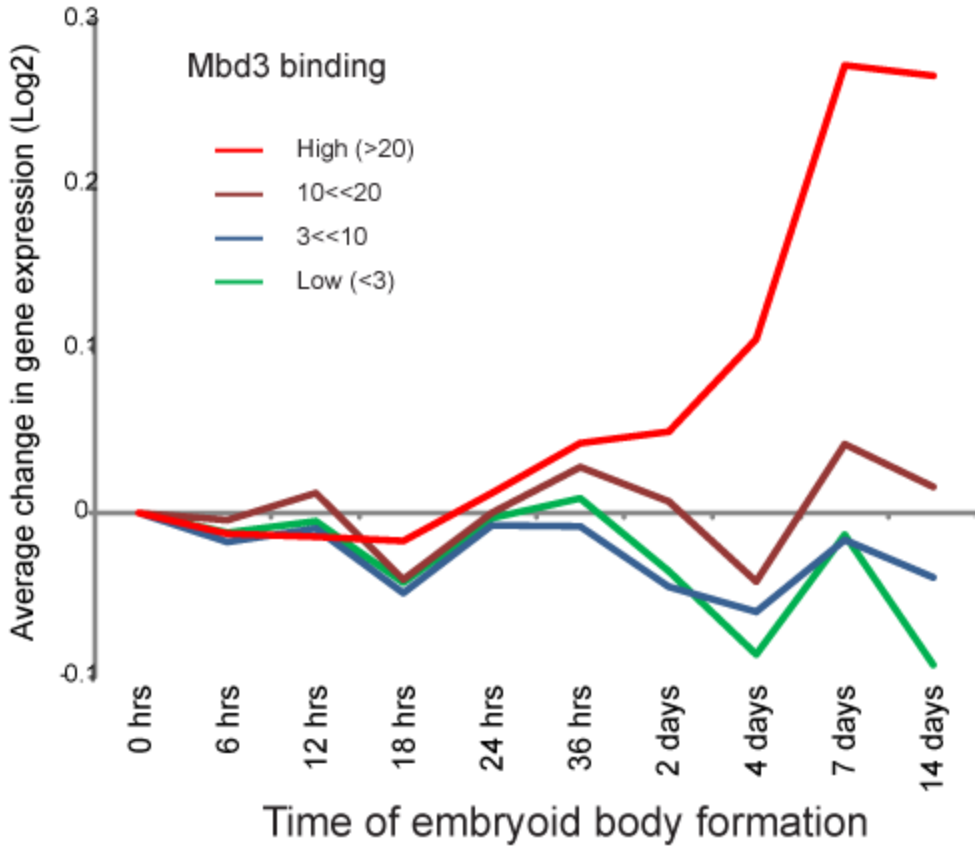

Figure 14. **Highly Mbd3 bound genes were upregulated during embryoid development occupancy** Dr. Oliver Rando, unpublished result

Mbd3 binding level was separated into 4 groups shown on the figure. The average mRNA level for each group was plotted against time. Genes with highest Mbd3 binding showed the most increased mRNA level over time.

Mbd3 and Brg1 regulate the recruitment of RNA Pol II

The NuRD complex contains ATPase Swi2/Snf2 and histone deacetylase subunits Hdac1 and Hdac2. It is therefore interesting to see how Mbd3 (NuRD subunit) and Brg1 has controlled gene expression by observing the histone profile at the common target genes. The nucleosome occupancy of H3 and H4 acetylation at Mbd3/Brg1 common target sites were analyzed using ChIP-qPCR. In Mbd3 KD ES cells, genes normally repressed by Mbd3 had decreased H3 nucleosome occupancy and increased H4 acetylation. This result suggests that in wild type cells,

Mbd3 functions to stabilize and deacetylate nucleosomes at promoters. In contrast, genes in Brg1 KD has increased H3 occupancy and varied H4 acetylation (Figure 15). It can then be concluded that Brg1 and Mbd3 antagonistically regulate the nucleosome occupancy at the promoter sites of the target genes. Thus, In Brg1 binding genes, the gene activation is associated with the loss of nucleosome and in Mbd3 binding genes, the gene repression is associated with nucleosome stabilization and deacetylation.

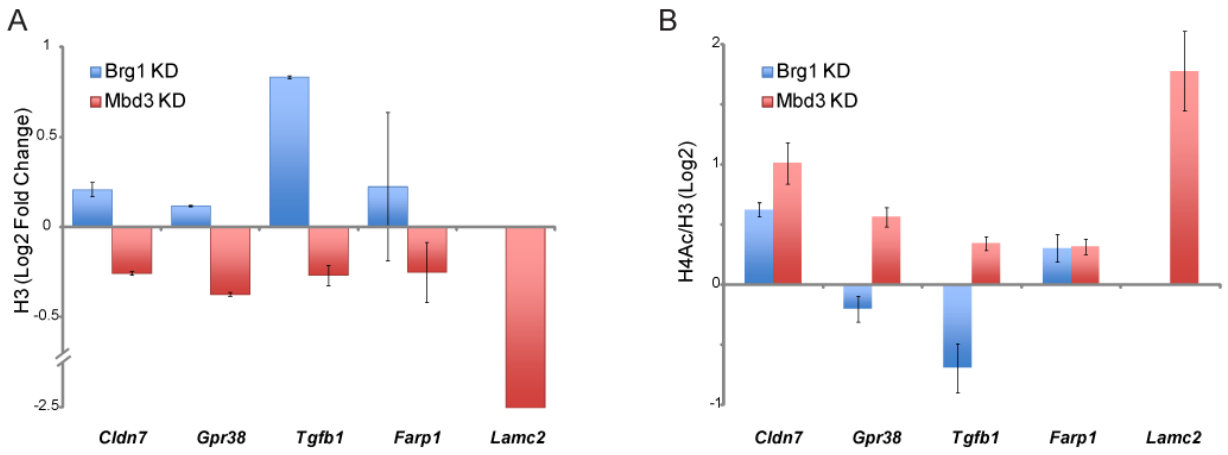

Figure 15. **Mbd3 and Brg1 antagonistically regulate nuclosome occupancy (Dr. Oliver Rando, unpublished result)**

(A) Nucleosome occupancy measured by H3 histone level at Mbd3/Brg1 targets. The plot shows the fold change between Mbd3 KD and Brg1 KD vs control KD. H3 level increased in Brg1 KD ES cells and decreased n Mbd3 KD ES cells

(B). H4 acetylation at Mbd3/Brg1 targets. Similar to A, acetylation level was compared with control KD. Acetylation level is varied at targets in Brg1 KD ES cells and increased in Mbd3 KD ES cells

Existing studies have shown that genes in mammals and eukaryotes are regulated by transcriptional pausing, the temporary suspension of RNA Polymerase II (Pol II) unit during transcription [Core, 2008]. Because of the downstream binding of Mbd3 proximal to the promoter, it is interesting to see whether the NURD complex is involved in this mechanism.

Whole genome mappings of Pol II in GFP KD, Mbd3 KD, and Brg1 KD were carried out in ES cells. The binding profiles of Pol II around the promoter region were similar among the three conditions. (Figure 16) The Pol II profile for a group of genes that are repressed by Mbd3 and activated by Brg1 was further analyzed. (Figure 17) As shown in the figure, Pol II binding increased at the 5' end of these genes in Mbd3 KD and decreased in Brg1 KD. Because the Pol II profile was changed only at the promoter region instead of over the gene body, this result suggests that the role of the factors is in the recruitment of Pol II rather than in the transcriptional pausing.

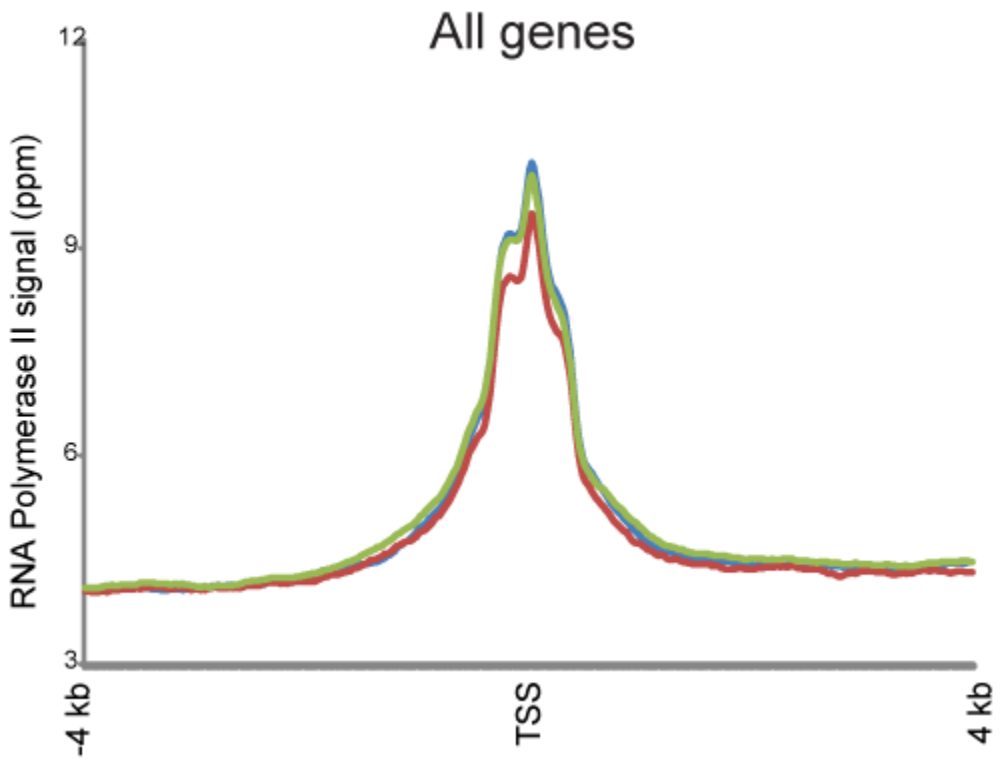

Figure 16. **Similar RNA Polymerase binding profile at promoters in GFP KD, Mbd3 KD and Brg1 KD.**

RNA Polymerase was mapped to the genome in GFP KD, Mbd3 KD and Brg1 KD ES cells. The average binding profile of RNA Pol II was generated around the all the TSSs in mouse ES cells. Shown in figure is the RNA Pol II profile from -4kb to +4kb relative to the TSSs.

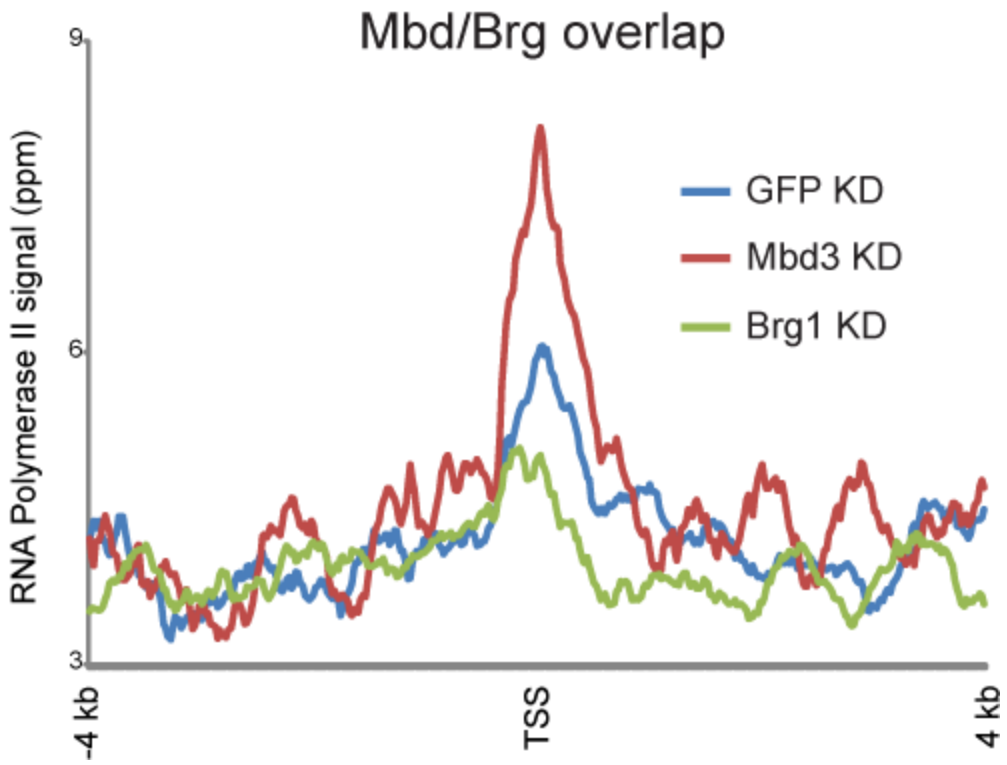

Figure 17. **RNA Pol II profile at the promoter of Mbd3/Brg1 overlapping targets in GFP KD, Mbd3 KD and Brg1 KD ES cells.**

RNA Pol II binding level generated in the same way as Figure 11, but only at a subset of genes which are repressed by Mbd3 and activated by Brg1. 5' RNA Pol II increased in Mbd3 KD cells and decreased in Brg1 KD cells.

Methods

Generating aggregation plot from ChIP-Seq data

ChIP-Sequencing (ChIP-Seq) of Mbd3 was done in wild type and Brg1 knockdown ES cells. ChIP-Seq of RNA Pol II was performed under Mbd3 KD, Brg1KD and GFP KD ES cells. In addition, ChIP-Seq data of Brg1, c-Myc, CTCF, E2fl, Esrrb, Klf4, Nanog, n-Myc, Oct4, Smad1, Sox2, STAT3, Suz12, Zfx, Tcfcp2Il, H3K4me3 and H3K27me3 were obtained from published

papers. Each ChIP-Seq dataset contained from 5 million to 31 million 25 or 36 base pairs sequencing tags. Each short sequencing tag was mapped to the most current mouse genome (mm9) using Bowtie (Langmead, 2010), allowing 1 mismatch. The tags that mapped to only one location or unique mappers to the mouse genome were retained for further analysis. After the tags were mapped to the genome, they were extended by 400 base pairs (the average size of cloned DNA fragment) in the 3' end direction. Each extended tag now represents a genomic location of the factor binding in the mouse genome. Using the extended tags, the aggregation plots of factor binding around the transcription start sites (TSS), as annotated by the RefSeq track of the UCSC genome browser, were generated. As shown here, each row represents a unique TSS in the mouse genome. For each TSS, the chromosome, strand, TSS location, gene information and the tags profile between a -4kb to +4kb window around the TSS was generated. The 8kb window was represented using a 20 base pairs bin size, so each number represents the number of tags in a 20 base pair region. The tags profile represents the distribution of tags around the TSS, e.g if at location -2000 relative to the TSS, the reads profile has a value of 3, it means that there are three tags found at 2000 bp to 1980bp upstream of the transcription direction. Normalization by the sequencing depth was done by multiplying the tag profile to (1,000,000/number of unique mappers for that ChIP-Seq dataset) x 20, so that the tables can be compared from one experiment to another.

chrom	strand	location	Info_1	Info_2	Info_3	Info_4	Info_5	Info_6	Info_7	Info_8	Info_9	Info_10	Info_11	Info_12	-4000
chrM	+	5326	n/a,,,,												20.00647
chrM	+	7009	n/a,,,,												66.52152
chrM	+	7765	n/a,,,,												20.00647
chrM	+	9875	n/a,,,,												25.50825
chrM	+	12405	n/a,,,,												0
chrY	-	31737	NM_011667,MGI:98891,Ube1y1,ubiquitin-activating enzyme E1, Chr Y 1,Gene												0
chrY	-	133851	NM_009570,MGI:99212,Zfy1,zinc finger protein 1, Y linked,Gene												0
chrY	+	155091	NM_011667,MGI:98891,Ube1y1,ubiquitin-activating enzyme E1, Chr Y 1,Gene												0
chrY	+	155155	NM_011667,MGI:98891,Ube1y1,ubiquitin-activating enzyme E1, Chr Y 1,Gene												0
chrY	+	234230	NM_0114	n/a,,,,											0
chrY	+	347054	NM_012011,MGI:1349430,Eif2s3y,eukaryotic translation initiation factor 2, subunit 3, structural gene Y-linked,Gene												0
chrY	-	395310	n/a,,,,												0
chrY	-	582180	NM_009484,MGI:894810,Uty,ubiquitously transcribed tetratricopeptide repeat gene, Y chromosome,Gene												0
chrY	-	582198	NM_0094	n/a,...											0

-3980	-3960	-3940	-3920	-3900	-3880	-3860	-3840	-3820	-3800	-3780	-3760	-3740	-3720
20.00647	20.00647	20.50663	28.50922	23.5076	30.00971	30.00971	30.00971	27.5089	20.00647	33.01068	40.01294	40.01294	40.01294
80.02589	89.0288	90.02912	90.02912	80.02589	90.02912	90.02912	74.02395	70.02265	74.02395	88.02848	90.02912	90.02912	90.02912
20.00647	20.00647	20.00647	20.00647	20.00647	15.00485	11.00356	20.00647	23.5076	32.01036	44.01424	50.01618	52.01683	50.01618
30.00971	31.01003	40.01294	52.51699	60.01942	68.52217	69.52249	60.01942	52.51699	50.01618	50.01618	50.01618	59.51925	75.02427
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 18. **Example of a tags profile generated for a dataset**. The figure shows a representation of a tags profile generated for a dataset which contains chromosome location, gene information, and an 8kb window around each TSS. Genes in the top 4 rows have significant number of tags in the -4000 bp to -3720bp region, suggesting binding of factor in this region

After generating the tags profile for each TSS, the aggregation plot can be generated by adding all the rows together as shown in Figure 8 in the result section. The height of the curve represents the relative binding level of the factor at that position.

Microarray data analysis

Gene expressions of 41,175 genes were measured by microarray in Mbd3 KD and Brg1 KD ES cells. Dr.Fazzio has identified 55 genes that were down regulated in Mbd3 KD and up regulated in Brg1KD and 98 genes that were up regulated in Mbd3 KD and down regulated in Brg1KD. The p value cut off was less than 0.05. The aggregation plots for the two sets of genes were obtained by extracting from the full table shown above.

Enrichment of factor binding at transcription factor binding sites

The binding sites for thirteen mouse ES transcription factors were obtained from the Ng group. The binding site locations from the paper were recorded according to the mm8 mouse genome. First, the binding site coordinates were converted to mouse mm9 genome. Then, the tags profiles around the binding sites were generated in a similar fashion as with the TSS binding profile. The difference was that we were looking at tags profile around binding sites instead of transcription start sites. The tags profiles from all ChIP-Seq datasets were generated around the binding site locations.

		-4000	-3980	-3960	-3940	-3920	-3900	-3880	-3860	-3840	-3
2	Brg1KD_cMyc_Ozlem	0.203589	0.204215	0.206645	0.207308	0.20814	0.207038	0.206106	0.205543	0.206459	0.2
3	Brg1Kd_Mbd3ab	0.192584	0.193975	0.193945	0.193975	0.193984	0.192846	0.192309	0.193542	0.192241	0.192
4	Brg1KD_Pol2_Fazzio	0.288945	0.28815	0.287286	0.287004	0.286498	0.285582	0.284457	0.283877	0.283544	0.28
5	Brg1KD_Pol2_Ozlem	0.208625	0.208524	0.208888	0.209133	0.20971	0.209891	0.210018	0.210266	0.210129	0.209
6	BRG	0.263609	0.264432	0.26498	0.26519	0.265133	0.264511	0.264542	0.264161	0.263262	0.261
7	EGFPKD_Pol2_Fazzio	0.279177	0.279228	0.279309	0.28002	0.279569	0.279405	0.279428	0.279435	0.280441	0.280
8	GFPKD_cMyc_Ozlem_75bp	0.22811	0.230489	0.227857	0.224847	0.225542	0.226973	0.231625	0.232657	0.233625	0.237
9	GFPKd_Mbd3ab7A	0.201476	0.201436	0.20293	0.203284	0.207281	0.208264	0.207714	0.203638	0.204608	0.205
10	GFPKD_Pol2_Ozlem_75bp	0.287189	0.288342	0.288577	0.288923	0.288655	0.286683	0.285332	0.283892	0.282217	0.281
11	Mbd3_Fazzio	0.253859	0.253543	0.252401	0.252467	0.25264	0.252079	0.250519	0.249303	0.247906	0.245
12	Mbd3Kd_cMycab2	0.230793	0.231257	0.231097	0.233023	0.233048	0.231748	0.230855	0.229244	0.229491	0.230
13	Mbd3Kd_cMycab3	0.213195	0.215159	0.214918	0.213278	0.213864	0.214076	0.214499	0.215277	0.213958	0.210
14	Mbd3KD_Pol2_Fazzio	0.268925	0.268088	0.267593	0.268029	0.268193	0.268074	0.267131	0.266771	0.266644	0.267
15	Mbd3KD_Pol2_Ozlem	0.210731	0.210543	0.210769	0.211079	0.211644	0.211447	0.211221	0.210833	0.210611	0.20
16	Mbd3_Ozlem	0.247145	0.248955	0.250124	0.249848	0.250086	0.250035	0.248602	0.248372	0.249299	0.249
17	Tet1KD_Mbd3	0.195016	0.19491	0.195282	0.195605	0.196918	0.197125	0.196165	0.197247	0.197356	0.197
18	Young_Pol2_run1	0.71694	0.71587	0.716354	0.715518	0.713892	0.714688	0.712711	0.709917	0.707866	0.705
19											
20											
21											
22											
23											
24											
25											

TSS | Readme | Sheet1 | Oct4 | c-Myc | CTCF | E2fl | Essrb | Klf4 | Nanog | n-Myc | Smad1 | Sox2 | STAT3 | Tcfcp2l1 | Zfx

Figure 19. **Example of read profiles around thirteen mouse ES transcription factor binding sites.** Shown on the bottom of the figure are the names of the transcription factors. The read profiles of the

datasets shown on the left were generated around each transcription factor binding site obtained from the published paper (Chen, 2008)

Discussion

The pluripotency and self-renewal of mouse ES cells have been extensively studied because mouse ES cells share many similarities to human cells and are easier to manipulate. The core transcription factors, such as Oct4, Sox2, and Nanog have been mapped and their functional relationships have been studied in prior work. However, in mouse ES cells, many additional transcription factors and chromatin remodelers are also important for pluripotency and self-renewal and their properties have not been determined. In this study, chromatin remodelers Brg1 and Mbd3 were studied and they have shed new light on the ES cell regulation network.

In Mbd3 KD and Brg1 KD ES cells, we have identified a group of genes that have opposite expression changes under these two conditions. Several hundred genes experienced up regulation in Mbd3 KD ES cells, and the same set of genes were down regulated in Brg1 KD ES cells. Another smaller set of genes exhibited the opposite expression changes. In Mbd3 and Brg1 double KD cells, the expression change is less severe and more similar to the WT condition. This observation suggests that Mbd3 and Brg1 have antagonistic functions in mouse ES cells. A similar scenario in ES cells that exhibit the antagonistic actions is the bivalent histone marks, H3K4me3 and H3K27me3. Many genes have both the activating H3K4me3 mark and repressing H3K27me3 mark in ES cells. This may suggest that antagonistic regulation may be a common mechanism in mouse ES cells.

Mbd3 and Brg1 were both enriched at the promoter region and regulate many common genes in the ES cells. It was also found that the two factors physically interact using immunoprecipitation. This result may suggest that the two factors function as a unit in ES cells. In addition, Mbd3 binding levels were reduced in Brg1 KD ES cells, which further supports the model that Brg1 interacts with Mbd3 in vivo. Interestingly, double KD of Mbd3 and Brg1 reverted the gene expression of common targets to close to WT expression compared to Brg1 single KD. Thus, even though Mbd3 binding was lost in Brg1 KD condition, some functionality of Mbd3 was still present.

Genome-wide mapping of Mbd3 showed that it localized near the promoter regions. Specifically, the peak of binding was about 200 base pairs downstream of the promoters. The biological reason for the downstream preference requires further examination. Previous ChIP-Seq of Brg showed that it also preferentially binds to the promoter region. The mechanism for such preference is also unclear. However, it is known that RNA Pol II also binds to the promoter region. One hypothesis is that Mbd3 and Brg1 have a functional relationship with RNA Pol II. As seen from the result, the Pol II binding over a set of genes misregulated in Mbd3 and Brg1 KD ES cells deviated from the control group, indicating the Mbd3 and Brg1 have roles in the recruitment of RNA Pol II.

Reference

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev *17*, 126-140.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks

key developmental genes in embryonic stem cells. Cell 125, 315-326.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122, 947-956.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell 113, 643-655.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells. Cell 133, 1106-1117.

Cuthbert, Graeme. "Guide to Chromatin Immunoprecipitation (ChIP) Methods." Abcam. Web. 28 Apr. 2011. <http://www.abcam.com/index.html?pageconfig=resource&rid=9>.

de la Serna, I.L., Ohkawa, Y., and Imbalzano, A.N. (2006). Chromatin remodelling in mammalian differentiation: lessons from ATP-dependent remodellers. Nat Rev Genet 7, 461-473.

Denslow, S.A., and Wade, P.A. (2007). The human Mi-2/NuRD complex and gene regulation. Oncogene 26, 5433-5438.

Fazzio, T.G., and Panning, B. (2010). Control of embryonic stem cell identity by nucleosome remodeling enzymes. Curr Opin Genet Dev 20, 500-504.

Gao, X., Tate, P., Hu, P., Tjian, R., Skarnes, W.C., and Wang, Z. (2008). ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling component BAF250a. Proc Natl Acad Sci U S A 105, 6656-6661.

Ho, L., Jothi, R., Ronan, J.L., Cui, K., Zhao, K., and Crabtree, G.R. (2009a). An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. Proc Natl Acad Sci U S A 106, 5187-5191.

Julie.L., Gerald, C (2010) Chromatin Regulatory Mechanisms in Pluripotency. Annual Review of Cell and Developmental Biology Vol 26: 503-532

Kaji, K., Caballero, I.M., MacLeod, R., Nichols, J., Wilson, V.A., and Hendrich, B. (2006). The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. Nat Cell Biol 8, 285-292.

Kaji, K., Nichols, J., and Hendrich, B. (2007). Mbd3, a component of the NuRD corepressor

complex, is required for development of pluripotent cells. Development 134, 1123-1132.

Keenen, B., and de la Serna, I.L. (2009). Chromatin remodeling in embryonic stem cells: regulating the balance between pluripotency and differentiation. J Cell Physiol 219, 1-7.

Kidder, B.L., Palmer, S., and Knott, J.G. (2009). SWI/SNF-Brg1 regulates self-renewal and occupies core pluripotency-related genes in embryonic stem cells. Stem Cells 27, 317-328.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008). An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells. Cell 132, 1049-1061.
Lewis, Molecular Biology of the Cell. 4th edition. Alberts B, Johnson A, Lewis J, et al. New York: Garland Science; 2002.

Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25.

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., et al. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet 38, 431-440.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 113, 631-642.

"Mouse Embryonic Stem Cells." Open Biosystem. Web. 28 Apr. 2011. <http://www.openbiosystems.com/GeneTargeting/MEScells/>.
Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell 95, 379-391.

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

Niwa, H. (2007). Open conformation chromatin and pluripotency. Genes Dev 21, 2671-2676.
Racki, L.R., and Narlikar, G.J. (2008). ATP-dependent chromatin remodeling enzymes: two heads are not better, just different. Curr Opin Genet Dev 18, 137-144.

Rice, JC., Allis CD (2001). Histone methylation versus histone acetylation: new insights into epigenetic regulation. Curr Opin Cell Biol 2001: 263-73

Surface, L.E., Thornton, S.R., and Boyer, L.A. (2010). Polycomb group proteins set the stage for early lineage commitment. Cell Stem Cell 7, 288-298.

Trotter, K.W., and Archer, T.K. (2008). The BRG1 transcriptional coregulator. Nucl Recept Signal 6, e004.

van Lohuizen, M. (1998). Functional analysis of mouse Polycomb group genes. Cell Mol Life Sci 54, 71-79.

Yan, Z., Wang, Z., Sharova, L., Sharov, A.A., Ling, C., Piao, Y., Aiba, K., Matoba, R., Wang, W., and Ko, M.S. (2008). BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells. Stem Cells 26, 1155-1165.

Zhu, D., Fang, J., Li, Y., and Zhang, J. (2009). Mbd3, a component of NuRD/Mi-2 complex, helps maintain pluripotency of mouse embryonic stem cells by repressing trophectoderm differentiation. PLoS One 4, e7684.