creative commons
COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

의학박사 학위논문

Ctbp2 modulates chromatin states during embryonic stem cell differentiation

배아 줄기 세포의 분화 과정에서 Ctbp2에 의한 후생유전학적 조절 기전 연구

2016년 2월

서울대학교 대학원

의과학과 의과학전공

강병희

A thesis of the Degree of Doctor of Philosophy

배아 줄기 세포의 분화 과정에서 Ctbp2에 의한 후생유전학적 조절 기전 연구

Ctbp2 modulates chromatin states during embryonic stem cell differentiation

February 2016

The Department of Biomedical Sciences
Seoul National University Graduate School
Byung-Hee Kang

ABSTRACT

The chromatin status is essential for cell fate determination. Accordingly, chromatin regulators are important for cell state transition process, such as deve1elopment and differentiation. Embryonic stem cells (ESCs) are maintained by core transcription factors (CTFs), which are closely connected with chromatin regulators to modulate chromatin environment and gene expression. The transcriptional corepressor C-terminal binding protein 2 (Ctbp2), which makes chromatin structures less accessible for transcription, is known as gating factor for pluripotency exit, but the underlying mechanism of Ctbp2 in ESC differentiation is poorly understood. Here, we show that Ctbp2 and ESC CTFs are co-occupied in active ESC genes and Ctbp2 mediates H3K27ac deacetylation in association with nucleosome remodeling and deacetylation (NuRD) complex during mouse embryonic stem cell (mESC) differentiation. In addition, it was confirmed that Ctbp2 occupies most of super-enhancers, regulatory region of CTFs. In undifferentiated conditions, knockdown of *Ctbp2* causes aberrantly high H3K27ac level but the gene expression levels are not significantly changed. Even if differentiation signal is given, *Ctbp2*-knockdown cells fail to down-regulate ESC CTF expression, thereby sustaining ESC maintenance. The fact that Ctbp2 and NuRD complex preoccupy in super-enhancers accounts for the recent finding that super-enhancers are sensitive to perturbation. We validated above findings in human breast cancer cell line MCF7 by analyzing public genomewide ChIP-

sequencing data, which shows that CTBP1/2 is associated with master transcription factors ERα, FOXA1, and GATA3 in ER positive breast cancer. Furthermore, CTBP1/2 is enriched in active chromatin regions, especially in most of super-enhancers. These suggest that CTBP1/2 is critical for not only mESC differentiation but also cancer differentiation and progression processes in human.

* This work is published in Stem Cells; published online May 26, 2015 doi:10.1002/stem.2046

Keywords: embryonic stem cells (ESCs), differentiation, core transcription factors (CTFs), C-terminal binding protein 2 (Ctbp2), nucleosome remodeling and deacetylation (NuRD), histone deacetylation, super-enhancer

Student number: 2009-22240

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

mESC (Mouse embryonic stem cell)

CTF (Core transcription factor)

Ctbp2 (C-terminal binding protein 2)

Lsd1 (Lysine-specific demethylase 1)

NuRD (Nucleosome remodeling and deacetylation)

ChIP-seq (Chromatin immunoprecipitation sequencing)

RNA-seq (RNA sequencing)

IGV (Integrated genomics viewer)

TSS (Transcription start site)

scr (Scrambled)

I. INTRODUCTION

Embryonic stem cells (ESCs) are pluripotent cells that have the ability to self-renew and differentiate into any cell type. Pluripotency and self-renewal are maintained by a small number of core transcription factor (CTF) networks. Meanwhile, CTFs determine the progression of ESCs by activating stage-specific genes and repressing lineage-specific genes (1-3). Thus, maintaining the appropriate level of CTFs is important for decisions with regard to cell identity and commitment toward natural lineages (4). Accordingly, chromatin repressive complexes are required for coordinated differentiation.

Nucleosome remodeling and deacetylation (NuRD) complex, has been implicated in a general transcriptional corepressor crucial for embryogenesis and ESC differentiation by which one of the main functions is deacetylation of H3K27ac in active ESC genes (5-9). Mbd3 (Methyl-CpG Binding Domain Protein 3), one of the core NuRD components, knockout ESCs are maintained in undifferentiated condition, but fail to lineage commitment upon induction of differentiation due to improper down regulation of active ESC genes (6,9). In addition, lysine-specific demethylase 1 (Lsd1), which associates with the NuRD complex in ESCs, is required for normal ESC differentiation by enhancer decommissioning of active ESC genes during differentiation (10). Four independent groups reported their purification of the Oct4 complex from mESCs (11-14), in which the NuRD components were core Oct4-associating proteins, indicating that gene regulatory network between Oct4 and the NuRD is essential for ESC pluripotency and/or lineage commitment.

Another important repressive epigenetic factor for embryogenesis and ESC differentiation is Polycomb repressive complex2 (PRC2). PRC2 catalyzes the trimethylation of histone H3K27 to form repressive chromatin that silences genes stably (15). PRC2 occupies bivalent genes in ESCs, which are defined by the presence of H3K4me3 and H3K27me3 (16). These regions are found primarily in development-related genes in ESCs and are poised for expression on ESC differentiation (17,18).

CtBP isoforms are transcriptional corepressors that recruit chromatin regulators (19-21). In vertebrates, Ctbp2 mutant mice show early embryonic lethality, while Ctbp1-null mice are viable (22). In addition, knockdown of *Ctbp2* fails to initiate differentiation of mESCs (23,24). Like the NuRD components, Ctbp2 was one of the 10 most common Oct4-interacting proteins in mESC (11-14), indicating that Ctbp2 contributes to Oct4 function with regard to establishing ESC identity. However, the epigenetic function of Ctbp2 in ESCs is unknown.

II. MATERIALS AND METHODS

1. ChIP-Sequencing

Sequencing of ChIPed DNAs was performed by BML (Korea) and KRIBB (Korean Research Institute of Bioscience & Biotechnology). Briefly, DNA fragments were ligated to a pair of adaptors for sequencing on the Illumina Hiseq-2000. The ligation products were size fractioned to obtain 200–300 bp fragments on a 2% agarose gel and subjected to 18 cycles of PCR amplification. Each library was diluted to 8 pM for 76 cycles of single-read sequencing on the Illumina Hiseq-2000 following the manufacturer's recommended protocol.

2. ChIP-Sequencing Data Analysis

2 - 1. Mapping and Identification of Enriched Regions

Reads after sequencing were mapped against to the mouse genome (NCBI build 37/mm9) using Bowtie v.0.12.7 (25) with the default parameters. The BAM format outputs were sorted by genomic coordinates (samtools sort) and reliable reads based on mapping score were used in subsequent processes. We used MACS v.1.4.0 tool to select binding sites for transcription factors and histone modification regions. We applied default settings and found significant regions (P-value <= 10^{-5}) compared to matched control samples.

2 - 2. Reference Data Information

We collected public ChIP-seq data that were related to core transcriptional factors and histone modifications in mESCs and MCF7 from previous publications. The list of data sets and GEO accession numbers are described in Table 1.

2 - 3. Annotation of Enriched Regions

Using HOMER (26), Ctbp2-enriched regions were annotated for promoters, exons, introns, intergenic regions and other features according to RefSeq transcripts. To annotate chromatin state for Ctbp2-enriched regions, we applied a genomewide segmentation program ChromHMM v. 1.0.6 (27). Firstly, we segmented the E14 chromatin into ten different states (putative active promoter, putative weak promoter, putative proximal weak enhancer, putative active enhancer, putative distal, weak enhancer, transcription transition, transcription elongation, insulator, weak signal1, weak signal2) determined by genomic coordinates and the distribution of the histone modification profile (H2AZ, H3K4me1, H3K4me2, H3K4me3, H3K27ac, H3K27me3, H3K36me3, H3K9ac) (28,29) and CTCF (30) binding profiles. Afterwards, we intersected segmented E14 chromatin with Ctbp2-enriched regions using BEDOPS v2.2.0 tools (31), in order to determine chromatin states for Ctbp2-enriched regions. GO annotation was performed using GREATv2.0.2 (32).

2 - 4. Motif and Co-bound Factor Analysis

HOMER was used for motif and co-bound factor analysis. De novo motif and known motif discovery algorithms were applied with default settings. We considered two factors to be co-bound when their peak centers were located within 200bp each other. Visualization of histone modification profiles for Ctbp2-enriched regions and refGene TSS spanning regions was accomplished according to the following step. First, ±50kb peak spanning regions were binned into 200bp windows, and for each window, normalized read densities were generated using HOMER. Second, applying Cluster v3.0 (33) (distance measure: euclidean, linkage: pairwise single linkage, k-means: k =3) read densities were clustered into three groups. Third, their heatmaps were created by Java TreeView (http://jtreeview.sourceforge.net).

3. RNA-Sequencing

Sequencing library was prepared using Truseq RNA Sample Preparation kit v2 (IIIumina, CA, USA) according the manufacture's protocols. Briefly, mRNA was purified from total RNA using poly-oligoT-attached magnetic beads, fragmented, and converted into cDNAs. Then, adapters were ligated and the fragments were amplified on a PCR. Sequencing was performed in paired end reads (2x100bp) using Hiseq-2000 (IIIumina).

4. RNA-Sequencing Data Analysis

Reads for each sample were aligned to the mouse genome (NCBI build 37/mm9) using TopHat v2.0.7 (34) with default settings. Cufflinks v2.0.2 (35) was used to quantify FPKM values of genes defined from RefSeq transcripts.

Normalized gene expression table was collected by DESeq v1.16.0 (36). Unsupervised hierarchical clustering analysis of gene expression was done by Cluster 3.0 and visualized the results with Java TreeView. Differential expressed genes (DEGs) were defined applying edgeR v2.12 (37) with FDR < 0.05. For defined DEGs, gene ontology analysis was conducted by DAVID (38). Other statistical tests were performed in R statistical packages.

Source	Cell line	Data type
GSE38596	E14	MeDIP-seq, RNA-seq, H2AZ, H3K27ac, H3K27me3 H3K36me3, H3K4me1, H3K4me2, H3K4me3
GSE24211	E14	H3K9Ac
GSE44288	V65	OCT4, SOX2, NANOG
GSE28247	E14	P300, CTCF
GSE11431	E14	NANOG, OCT4, SOX2, SMAD1, E2F1, TCFCP2I1, CTCF, ZFX, STAT3, KLF4, ESRRB, c-MYC, n-MYC, P300, SUZ12
GSE27841	V65	HDAC1, HDAC2, LSD1, Mi2B, REST, COREST
GSE49431	E14	EZH2
GSE19165	E14	RNA-seq
GSE36546	MCF7	CTBP1/2
GSE60270	MCF7	ER-a, FOXA1, P300, GATA3
ENCODE	MCF7	RNA-seq, CTCF, EGR1, ELF1, FOSL2, FOXM1, GABP, GATA3, HDAC2, JUND, MAX, NR2F2, NRSF, P300, PML, RAD21, RXL, SIN3A, SRF, TAF1, TCF12, TEAD4, ZNF217, c-MYC, POL2, HAE2
ENCODE	E14, MEF, NIH3T3, CH12	RNA-seq

Table 1. The list of re-analyzed data sets

III. RESULTS

1. Ctbp2 is associated with ESC CTFs

To access the role of Ctbp2 in ESC differentiation, we explored the expression level of *Ctbp2* using publically available RNA-seq data (29,39). When we compared the expression level of *Ctbp2* and *Ctbp1* in various sorts of mouse cell line, *Ctbp2* was highly expressed in mESC, unlike *Ctbp1* (Figure 1a). Meanwhile, *Ctbp2* level was decreased during mESC differentiation (Figure 1a). Additionally, the expression level of *Ctbp2* was associated with ESC CTFs, such as *Pou5f1*, *Sox2*, and *Nanog* (Figure 1b). At the same time, ESC CTFs occupied Ctbp2 regulatory elements, which indicate that Ctbp2 are regulated by ESC CTFs (Figure 1a).

Next, we mapped the genomewide occupancy of Ctbp2 in mESC by ChIP-seq with Ctbp2-specific antibody. We identified 7,406 high-confidence peaks of Ctbp2 (with $P<10^{-5}$) that were enriched in promoters, introns and intergenic regions (Figure 2a). By gene ontology (GO) analysis using GREAT (32), Ctbp2-occupied regions were linked to stem cell maintenance and development (Figure 2b). Oct4, Sox2, Klf4, and Esrrb motifs were significantly enriched in regions Ctbp2 peaks, based on motif discovery algorithms (26) (Figure 3a). In addition, Ctbp2-enriched regions were significantly overlapped with ESC CTFs (Figures 3b, 3c, and 3d), consistent with previous reports which Ctbp2 associates with the Oct4 complex (11-14). Taken together, these results suggest that Ctbp2 plays a role in concert with ESC CTFs at early differentiation phase in regulating differentiation.

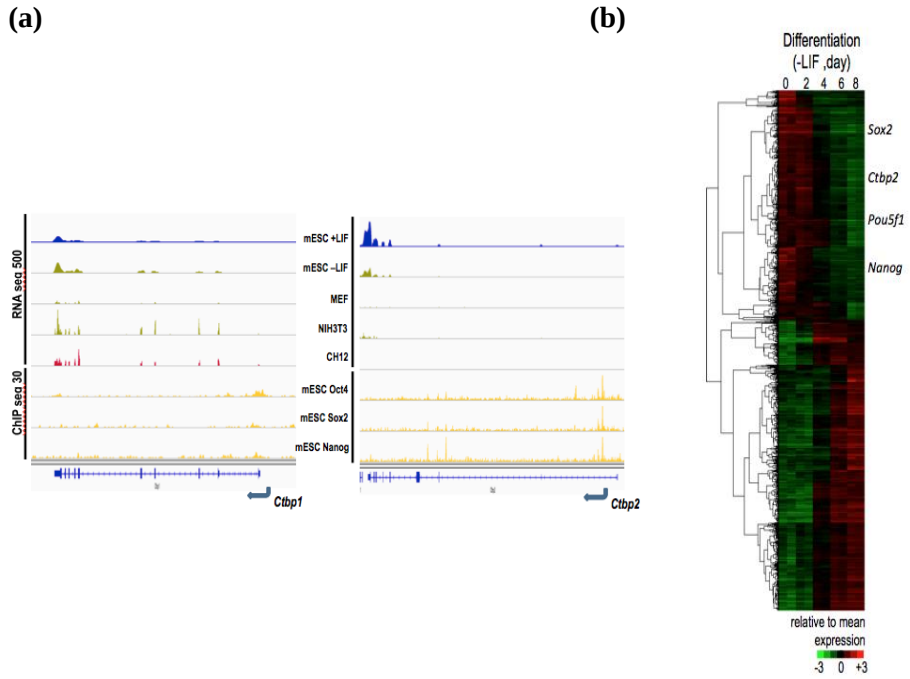
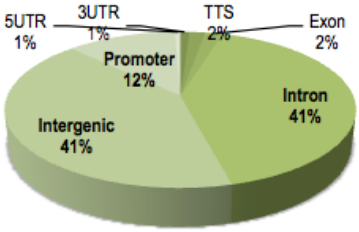


Figure 1. Ctbp2 is CTF downstream target and up-regulated in undifferentiated mESC.

(a) Integrated genomics viewer (IGV) representation of *Ctbp1*, and *Ctbp2* loci. The expression levels of *Ctbp1* was higher in differentiated cell line, such as E14 (mESC) LIF withdrawal day 6, MEF (mouse embryonic fibroblast), NIH-3T3 (fibroblast), CH12 (B cell lymphoma) than undifferentiated ESC. *Ctbp2* is highly expressed in undifferentiated mESC and occupied by Oct4, Sox2, and Nanog in regulatory elements. **(b)** ESC CTFs and Ctbp2 were presented with clustered image.

(a)



(b)

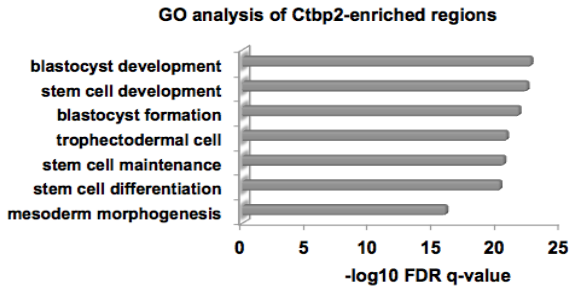


Figure 2. Annotation for Ctbp2-enriched regions.

(a) Pie chart showing the proportion of genomic regions for Ctbp2-enriched regions. **(b)** Functional annotation of Ctbp2 peaks using GREAT (32), showing enrichment of stem cell maintenance and development genes.

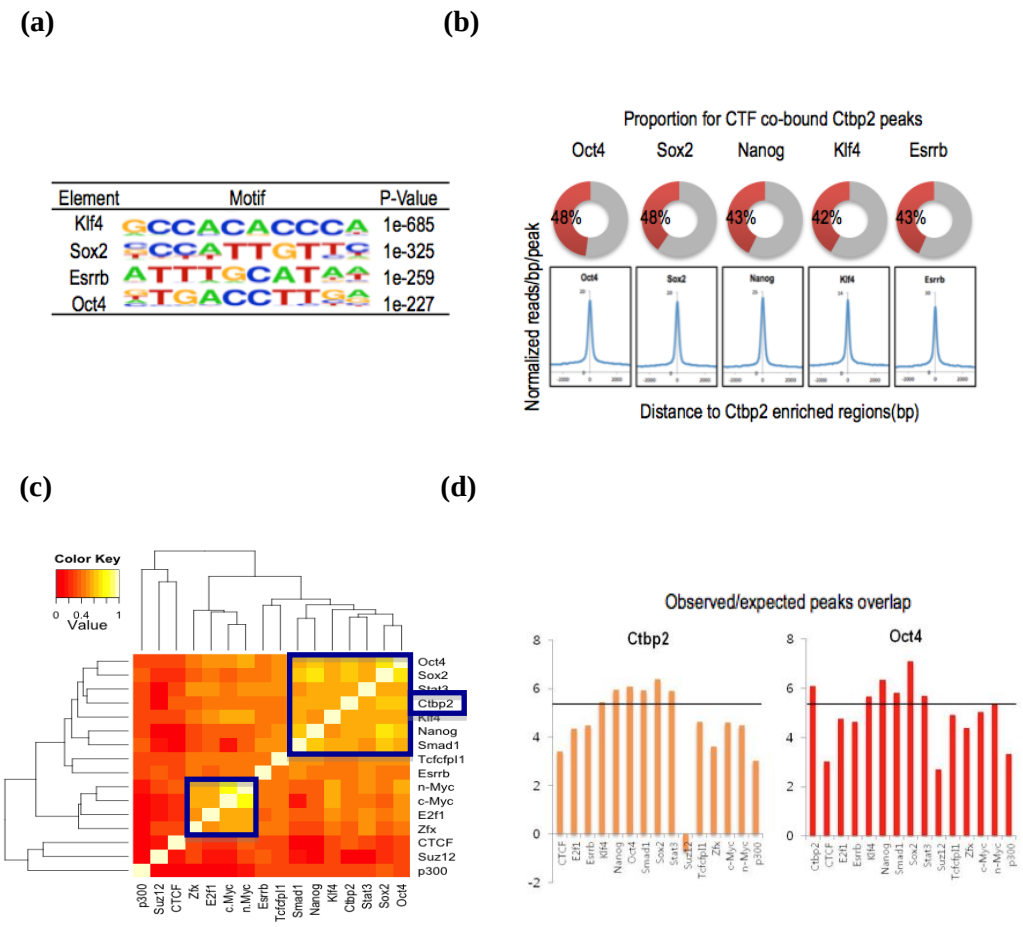


Figure 3. Ctbp2 is co-occupied with ESC CTFs.

(a) Using motif discovery algorithms (26), the Oct4, Sox2, Klf4, and Esrrb motifs were predicted with high scores in Ctbp2-enriched regions **(b)** Upper panel depicts the percentages of regions co-occupied by Ctbp2 and CTFs (Oct4, Sox2, Nanog, Klf4, Esrrb) (40,41) in Ctbp2-enriched regions in ESCs. Lower panel shows the binding profiles of CTFs in Ctbp2-enriched regions**. (c, d)** Analysis of Ctbp2-enriched regions for previous ChIP-seq data in ESCs (40) reveals that CTFs are co-bound with Ctbp2. CTFs and Ctbp2 were clustered along both axes based on the similarity for scaled peak overlap ratio.

2. Ctbp2 is enriched in active chromatin regions

To analyze the chromatin states in Ctbp2-enriched regions, we annotated a combination of histone modifications to the mouse genome using ChromHMM (27) (Figure 4). According to our dissected chromatin states, Oct4-enriched regions were mostly located in active enhancers and promoters, Suz12 in weak promoters and enhancers, P300 in active enhancers, and H3K4me3 in active promoters (Figure 5a), consistent with previous studies (10,29,30). Ctbp2-enriched regions are generally comprised of active enhancers and promoters (58%, $P<10^{-16}$). Also, Ctbp2 occupied weak enhancer and promoter regions (32%; Figure 5b).

To profile histone modification for H3K4me1, H3K4me3, H3K27ac, and H3K27me3, 30,392 refGene m99 TSS surrounding 100kb regions were clustered according to Ctbp2 and chromatin enrichment status. Consistent with the ChromHMM chromatin annotation results, Ctbp2-enriched regions were significantly overlapped with active enhancer marker H3K27ac and H3K4me1 (Figures 6a, and 6b). When searching for the individual gene tracks at *Pou5f1* and *Nr5a2* loci, representatives of active ESC genes, we found that Ctbp2 occupies enhancer and promoter regions together with the active chromatin states and ESC CTFs, but not housekeeping gene, *ActB* (Figure 7)

. Considering the recent findings that super-enhancers in ESCs are linked to cell type-specific genes and regulate ESC identity (41), we hypothesized that Ctbp2 occupied super-enhancer regions and controlled their chromatin status

during differentiation. Indeed, most ESC super-enhancers were overlapped with Ctbp2-occupied regions (203/231, 87.8%; Figure 8a). The levels of H3K27ac at regions bound by both Ctbp2 and p300 were higher than those of p300 alone regions (Figure 8b). These results suggest that Ctbp2 was associated with ESC super-enhancers.

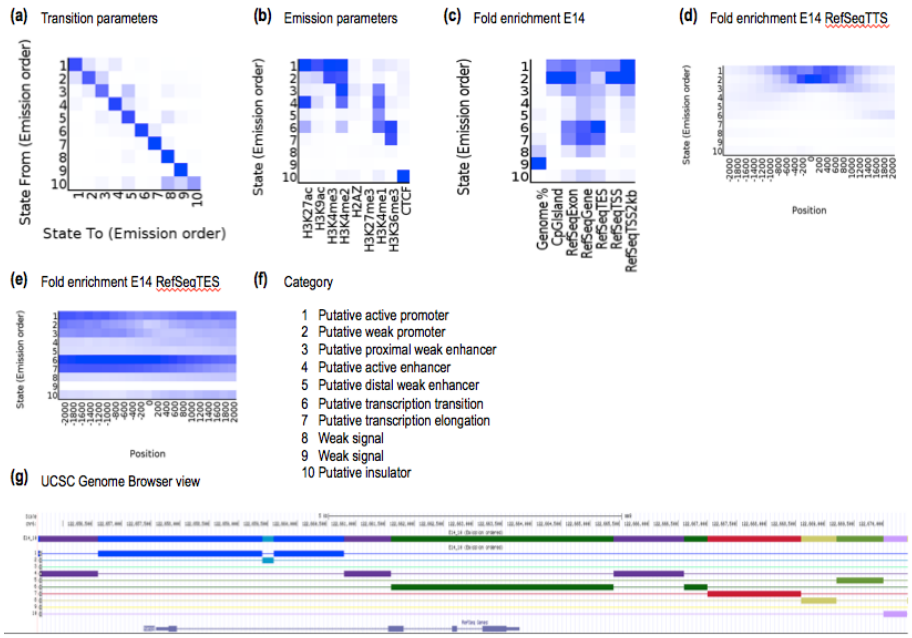
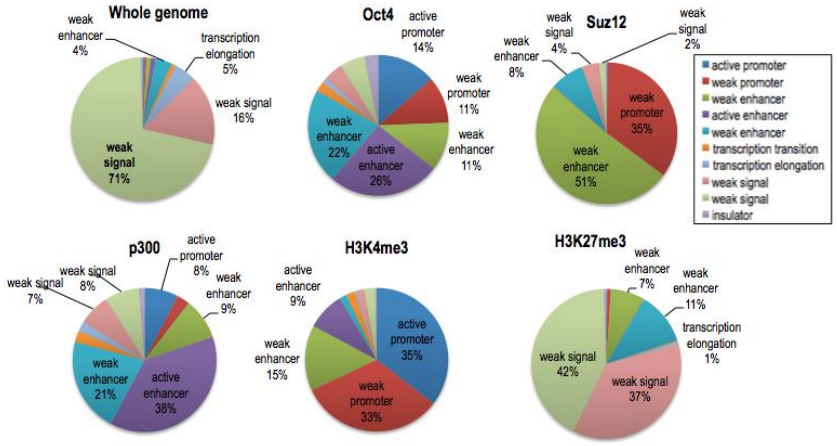


Figure 4. Annotation of chromatin states using combinations of histone modification patterns.

(a) The heatmap indicates transition probabilities of the HMM model. **(b)** Emission parameters of chromatin states using enrichment of nine histone marks. **(c-e)** Fold enrichment of chromatin states on genomic positions. **(f)** Candidate annotation for each chromatin state. **(g)** UCSC genome browser track (Kent et al, 2002) example for *Nanog* locus.

(a)



(b)

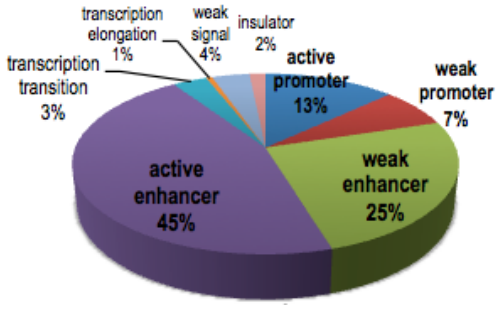


Figure 5. Proportion of chromatin states for genomic regions.

(a) In consistent with previous researches, Oct4-enriched regions were linked to active enhancers and promoters, Suz12 to weak promoters and proximal enhancers, P300 to active enhancers, and H3K4me3 to active promoters. **(b)** Pie chart showing the proportion of chromatin states in Ctbp2-enriched regions. Ctbp2-occupied regions (58%) are predominantly active enhancers and promoters.

(a) **(b)**

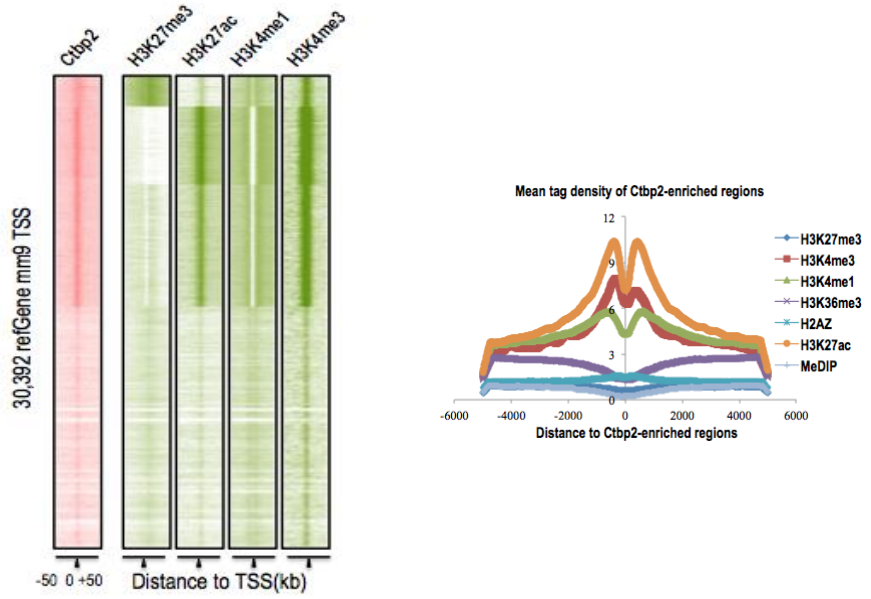


Figure 6. Ctbp2-enriched regions are significantly overlapped with active enhancer and promoter marker.

(a) Heatmap of the RefSeq genes (mm9) shows that Ctbp2-enriched regions correlate highly with active enhancer markers, (H3K3me1, H3K27ac, and H3K4me3) (29,30) but weakly with H3K27me3 (42). **(b)** Distribution of histone marks (H3K27me3 ,H3K27ac, H3K4me3, H3K4me1, H3K36me3, H2AZ) and DNA methylation (29,30,41), which were normalized ChIP-seq signal across the 7,406 Ctbp2-enriched regions and 10 kb surrounding regions, indicates that Ctbp2 occupies actively transcribed genes in ESCs.

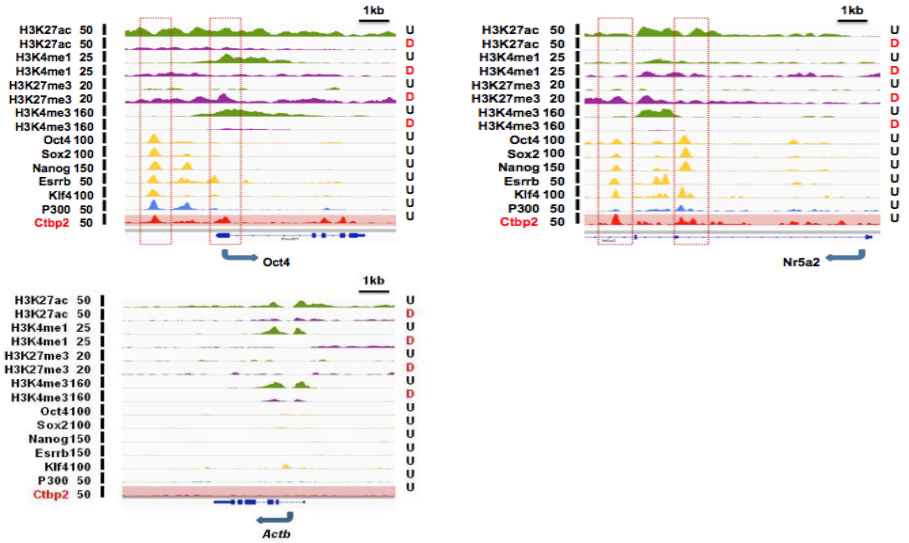


Figure 7. Ctbp2 is enriched in active ESC genes.

IGV representation of Ctbp2, Oct4, Sox2, Nanog, Esrrb, Klf4, H3K4me3, H3H4me1, H3K27ac, and H3K27me3 peaks in *Pou5f1*, *Nr5a2*, *ActB* loci in ESCs, indicating that Ctbp2-enriched regions are co-occupied by ESC core transcription factors that are highly enriched with active histone marks (H3K4me3, H3K27ac, and H3Kme1) but not with H3K27me3 in undifferentiated ESCs. This histone pattern was reversed in differentiated ESCs, as evidenced by the enrichment in H3K27me3 but not active histone marks. Unlike ESC active genes, in *ActB* loci the occupancies of Ctbp2, Oct4, Sox2, Nanog, Esrrb, and Klf4 were not shown in this locus. The enrichment of H3K27me3 did not appear after differentiation. [U; undifferentiated ESCs, D; differentiated ESCs (day 6)].

(a) **(b)**

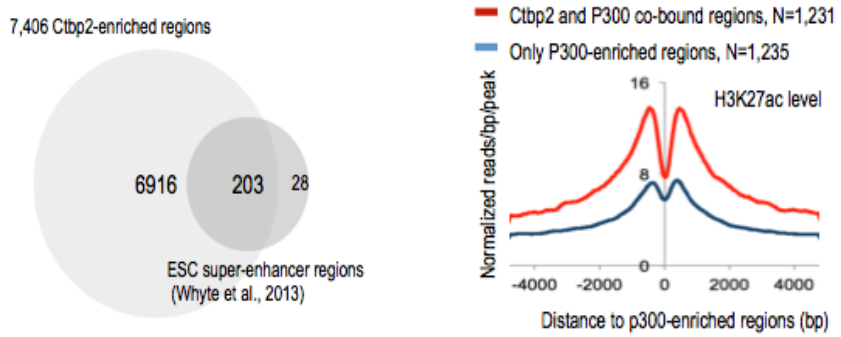


Figure 8. Ctbp2 is enriched in super-enhancers

(a) Most (87.8%) ESC super-enhancers (41) are 7,406 Ctbp2-enriched regions. **(b)** H3K27ac levels are higher at Ctbp2 and p300 co-occupied regions versus the 1,235 regions occupied by p300 alone (30).

3. Knockdown of *Ctbp2* attenuates gene expression changes during differentiation

Considering that Ctbp2 is generally known for a repressor (19) and enriched in active ESC genes, we aimed to examine the function of Ctbp2 in transcriptome. We compared transcriptomes between wild-type and *Ctbp2*-knockdown ESCs under undifferentiated and differentiating conditions by RNA sequencing. Consistent with Ctbp2 playing an important role in the exit from pluripotency, not in the maintenance (24), we identified that *Ctbp2* knockdown made more drastic changes on the transcriptome in differentiated ESCs than in undifferentiated ESCs. Furthermore, we found that LIF-withdrawal had profound impacts only in wild-type ESCs (Figures 9a and 9b).

To discover direct targets for Ctbp2 among differentially expressed genes (DEGs) during differentiation, we focused on the 1,430 genes that were occupied by Ctbp2 in ChIP-seq data and differentially expressed upon wild-type ESC differentiation in RNA-seq data. We identified 811 genes, which include ESC CTFs, was significantly down-regulated in wild-type ESCs after differentiation, and these expressional changes were not apparent in *Ctbp2*-knockdown ESCs (Figures 10a and 10b). These genes were shown to be related to stem cell maintenance and differentiation in gene ontology analysis (Figure 10b). By contrast, we identified 619 genes that were up-regulated in wild-type ESCs after differentiation, but not in *Ctbp2*-knockdown ESCs. These genes were associated with specific lineage commitment programs,

such as skeletal development (Figure 10b). Even though there were also 619 up-regulated genes found in Ctbp2 targets during differentiation, we focused on down-regulated genes, which represent ESC active genes, since ESC CTFs are master regulator for cell fate determination.

Next, we explored the expression level of the genes Ctbp2-occupied super-enhancers. Notably, these genes were significantly down-regulated in wild-type ESCs after LIF withdrawal, but not in *Ctbp2*-knockdown ESCs (Figures 11a and 11b).

(a) **(b)**

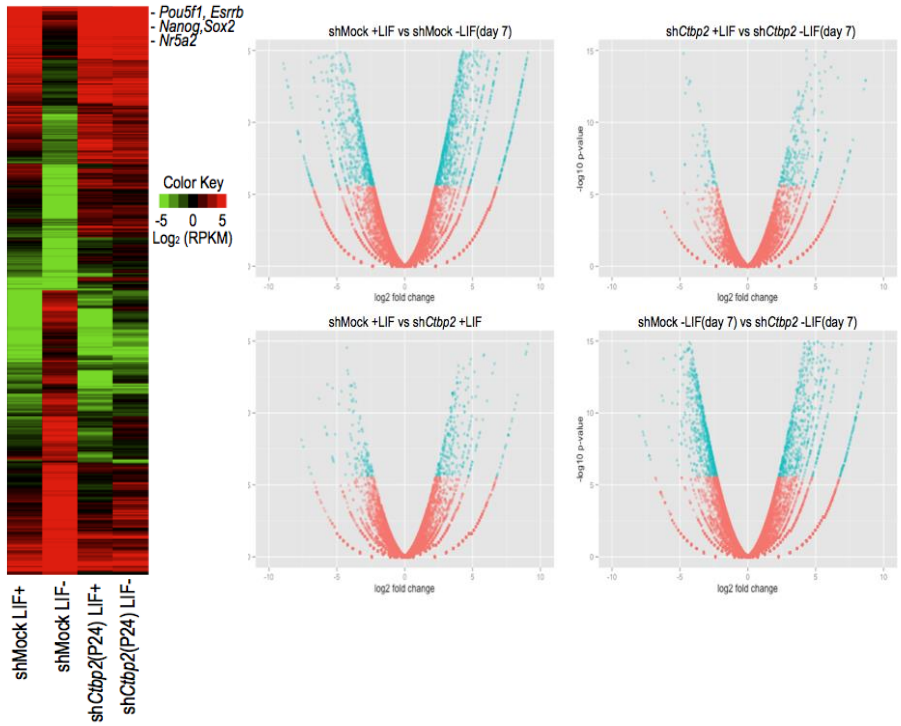


Figure 9. *Ctbp2* knockdown causes aberrant expression during ESC differentiation.

(**a**) Gene expression profiles were compared between wild-type and *Ctbp2*-knockdown ESCs in undifferentiated (+LIF) / differentiated (-LIF day 7) states. DEGs (FDR <0.01) were sorted according to their fold differences. Red/green colors represent expression changes relative to the mean of DEGs. Plot represents hierarchical clustering of the 500 most up-regulated or down-regulated genes. *Ctbp2* knockdown was shown to diminish the overall expression changes occurring during ESC differentiation. (**b**) RNA-sequencing of wild-type and *Ctbp2*-knockdown E14 in both undifferentiated (+LIF) and differentiated (-LIF day 7) states. Differential gene expressions for

19,305 genes were presented as volcano plot. The horizontal axis stands for Log2 fold changes and vertical axis stands for P values. Statistical tests were done by edgeR. Blue dots represent significantly down or up-regulated genes. The cut-off was bonferroni-corrected P value 0.05. Drastic changes on transcriptome occurred only two following cases, shMock +LIF versus shMock –LIF (day 7) and shMock –LIF (day 7) versus shCtbp2 (P24) –LIF (day 7).

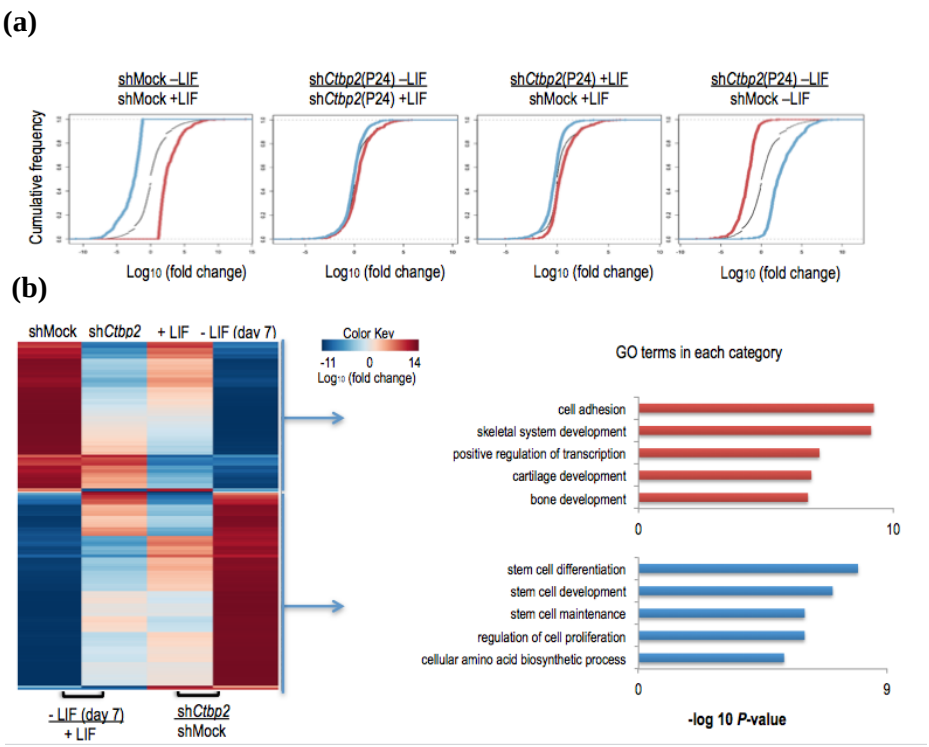


Figure 10. Effect of *Ctbp2* knockdown on Ctbp2 target genes.

(a) Gene expression profiles were analyzed for wild-type and *Ctbp2*-knockdown ESCs in differentiated (+LIF) and undifferentiated (-LIF) states. Gene expression profiles were similar in both ESC types in undifferentiated state. After differentiation, up-regulated and down-regulated genes were observed in wild-type ESCs but significantly reduced in *Ctbp2*-knockdown ESCs. [black line; all genes (19,305 genes), blue line; Ctbp2-occupied and down-regulated genes during differentiation (811 genes), red line; Ctbp2-occupied and up-regulated genes during differentiation (619 genes)]. **(b)** By gene ontology analysis using DAVID (Huang da et al., 2009) among differentially expressed genes (DEGs) which were bound by Ctbp2 during wild-type ESC differentiation, the first set, composed of 811 genes is related

with stem cell maintenance and development. It was strongly decreased after differentiation in wild-type ESCs but not in *Ctbp2*-knockdown ESCs. The other set, 619 genes, is related with cell adhesion and specific lineage commitments such as skeletal system development. It was strongly increased after differentiation only in wild-type ESCs.

(a) **(b)**

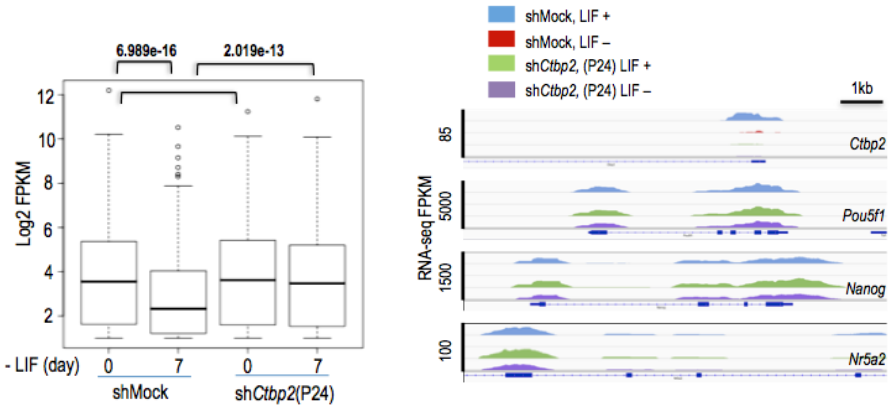


Figure 11. Effect of *Ctbp2* knockdown on genes regulated by super-enhancer.

(a) Expression of genes Ctbp2 occupied super-enhancer regions was significantly decreased in wild-type ESC after differentiation but not in *Ctbp2*-knockdown ESCs. **(b)** mRNA levels of *Ctbp2, Pou5f1*, *Nanog*, and *Nr5a2* represented by IGV.

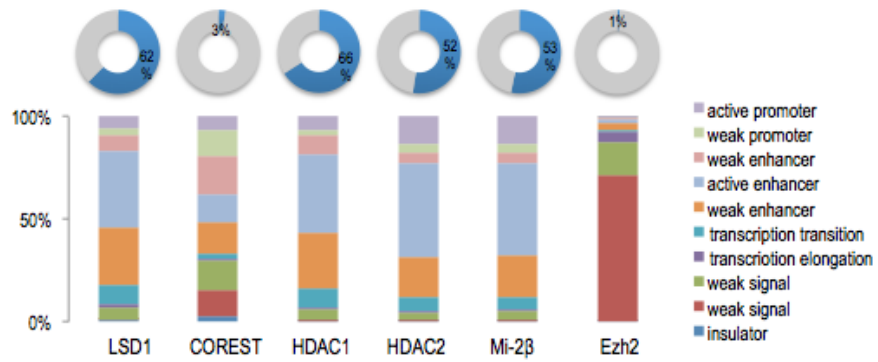
4. Ctbp2 controls the H3K27ac level during differentiation

Recently, by CtBP2 proteomic analysis, several components of the NuRD complex were identified as CtBP2-interacting proteins in HeLa cell line (43). Drosophila CtBP is a binding partner of Polycomb group protein (PcG), and CtBP mutant fails to recruit PcG to its targets (44). So far, the association between Ctbp2 and its repressive partners in ESCs is unknown. To gain further insight into the role of Ctbp2 as transcriptional corepressor, we compared Ctbp2-enriched regions with previously reported genomewide occupancy of repressor complexes (10,45), such as NuRD, CoREST, and PRC2. Ctbp2-enriched regions are well associated with the NuRD components (Lsd1, Hdac1, Hdac2, and Mi-2β), but not with CoREST or PRC2 complex (Figure 12a, upper). The NuRD components, Mbd3 and Hdac1, are dispensable for ESC maintenance but are essential for exit from pluripotency (6,9,46). In addition, the function of Lsd1 in ESCs is associated with the NuRD complex and is required for proper ESC differentiation (10). According to our dissected chromatin state analysis, Ctbp2 and NuRD co-occupied regions are highly associated with active enhancers (Figure 12a, lower), suggesting that Ctbp2 at active ESC genes has a function on exit from pluripotency like a NuRD components.

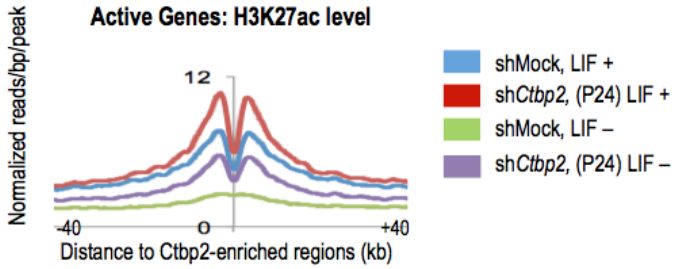
The NuRD complex has an ability to deacetylate H3K27, which is known for gene repression. It mediates active ESC gene silencing which is necessary for proper ESC differentiation (9). To examine whether Ctbp2 modulates the

NuRD-associated depression, we measured H3K27ac levels of Ctbp2-enriched regions in active ESC genes both in wild-type and *Ctbp2*-knockdown ESCs during differentiation. The H3K27ac level was decreased in wild-type ESCs during differentiation, whereas such decrease was marginal and disappeared in *Ctbp2*-knockdown ESCs (Figure 12b). The mean peak density of H3K27ac at active ESC genes was obviously higher in *Ctbp2*-knockdown ESCs than wild-type ESCs during ESC differentiation (Figure 12c), indicating that Ctbp2 has a function in repression of active ESC genes by deacetylation of H3K27ac.

(a)



(b)



(c)

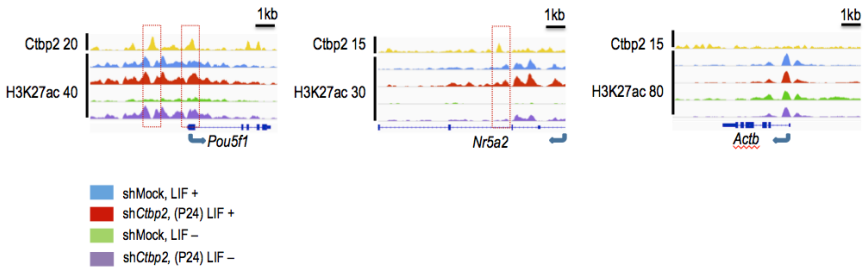


Figure 12. Ctbp2 controls the H3K27ac level in active ESC genes during ESC differentiation.

(**a**) Ctbp2 is co-occupied with the NuRD complex in active enhancers in ESCs. Upper panel shows the percentages of co-occupied regions by Ctbp2

and chromatin repressive complexes (Lsd1, CoREST, Hdac1, Hdac2, Mi-2β, Ezh2, Suz12). Lower panel shows the proportion of chromatin states in co-occupied regions by Ctbp2 and chromatin repressive complex.

(**b**) The H3K27ac level around active ESC genes of Ctbp2-occupied regions between wild-type and *Ctbp2*-knockdown ESCs during differentiation. The mean ChIP-seq density of H3K27ac increased after *Ctbp2* knockdown in wild-type ESCs. Meanwhile, H3K27ac level decreased after differentiation in wild-type ESCs, but *Ctbp2*-knockdown ESCs were shown to lessen the decrease of H3K27ac. **(c)** IGV representation of *Pou5f1, Nr5a2, and Actb* loci. The decrease of H3K27ac was shown around Ctbp2-enriched regions (*Pou5f1, Nr5a2*) in wild-type ESCs after differentiation, but such increase was diminished in *Ctbp2*-knockdown ESCs. In *Actb* regulatory locus (unoccupied by Ctbp2), no such difference of H3K27ac was observed in both ESCs.

5. CTBP1/2 occupies active chromatin regions in association with master transcription factors in MCF7.

Since super-enhancer and cell-type-specific master transcription factors are also found at not only ESCs but also differentiated cells, we questioned whether similar molecular mechanisms may exist in cancer cells. We analyzed the genomewide occupancy of CTBP1/2 in MCF7 human breast cancer cells (47), which are known for a well differentiated estrogen receptor positive luminal A subtype. CTBP1/2 occupied regions were significantly enriched for active promoters (19%) and active enhancers (19%; Figure 13a).

To investigate the functional role of CTBP1/2 in breast cancer cell, we searched for CTBP1/2 co-binding transcription factors. CTBP1/2 occupied regions were enriched for Forkhead-box protein A1 (FOXA1) and GATA-binding protein 3 (GATA3) motifs, which are known for estrogen receptor (ER) cooperating transcription factors (Figure 13c). Cluster analysis reveals a high overlap between CTBP1/2, ER, FOXA1, GATA3, ZNF217 and TCF7l2 (Figure 13d). By gene set enrichment analysis using GREAT (32), CTBP1/2-occupied regions were linked to luminal-like cancer cell specific gene sets and gene sets which are bound by ER (Figure 13e). ER is an essential transcription factor for the growth of ER positive breast cancer cell. GATA3, FOXA1, and ER are appeared to cooperate in mediating the estrogen response (48,49). Furthermore, ER, ZNF217 and GATA3 are proposed as master transcription factor, which are associated with super-enhancers (50).

In addition, most MCF7 super-enhancer regions were overlapped with CTBP1/2-occupeid regions like mESCs super-enhancer regions (106/117, 90.5%; Figure 14a). Moreover, CTBP1/2 enrichment was significantly higher in super-enhancers than in typical enhancers (Figure 14b). Taken together, these results indicate that CTBP1/2 plays important role in concert with master transcription factor networks in breast cancer actively transcribed regions.

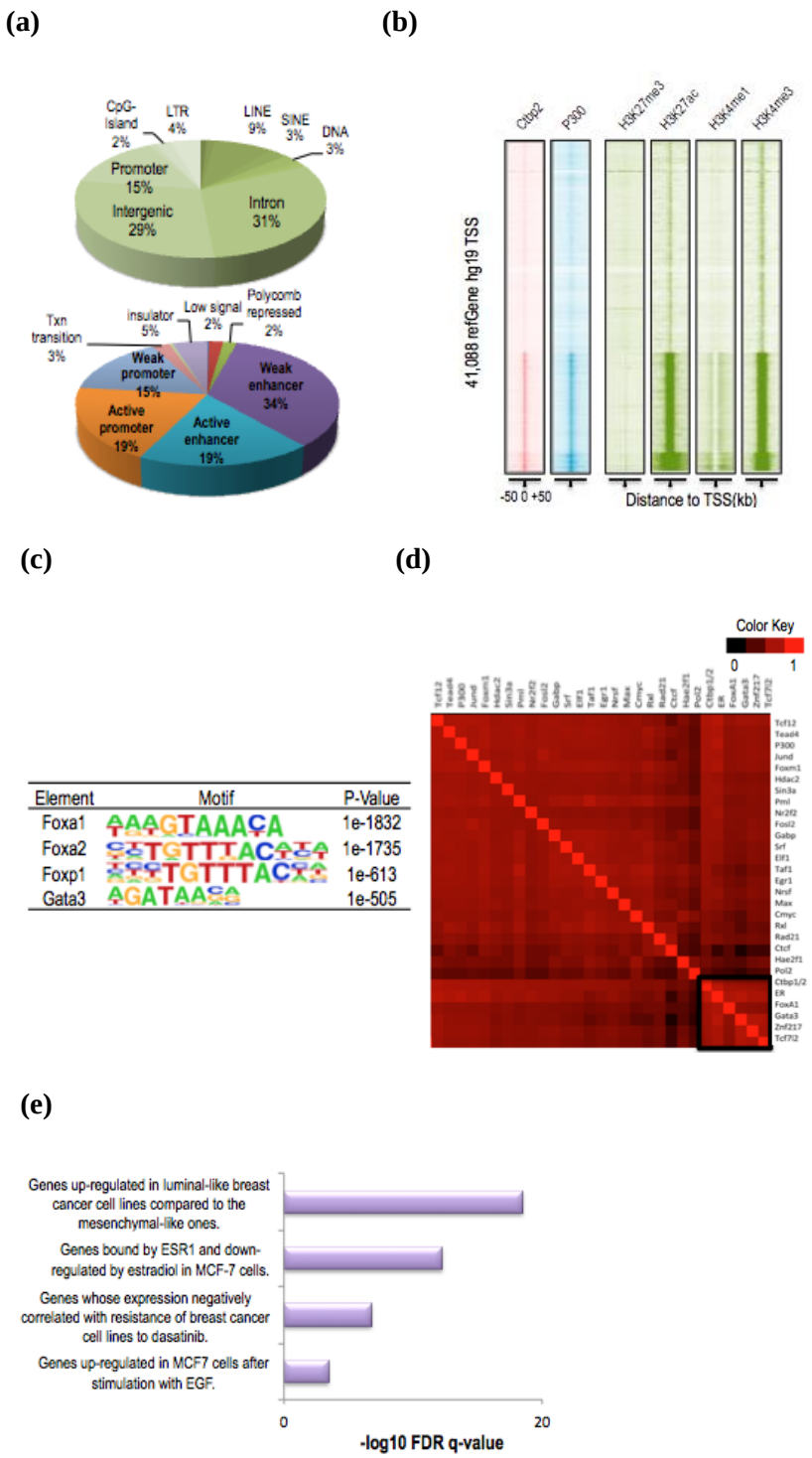


Figure 13. CTBP1/2 is enriched at actively transcribed genes in MCF7.

(**a**) Upper chart showing the proportion of genomic regionsfor CTBP1/2-enriched regions. Lower chart showing the proportion of chromatin states in CTBP1/2-enriched regions. **(b)** Distribution of histone marks (H3K27me3, H3K27ac, H3K4me3 and H3K4me1) and P300, which were normalized ChIP-seq signal across the 41,088 refGene hg19 TSS and 100 kb surrounding regions, indicates that CTBP1/2 occupies actively transcribed genes in ESCs. **(c)** Using motif discovery algorithms (26), the FOXA1, FOXA2, FOXP1, and GATA3 motifs were predicted with high scores in CTBP1/2-enriched regions. **(d)** Analysis of CTBP1/2-enriched regions for previous ChIP-seq data in MCF7 cells (40) reveals that breast cancer cell master transcription factor such as ER, FOXA1, and GATA3 are co-bound with CTBP1/2. **(e)** Functional annotation of CTBP1/2 peaks using GREAT (32), showing enrichment with luminal-like breast cancer specific genes and gene sets which are bound by ESR1**.**

(a) **(b)**

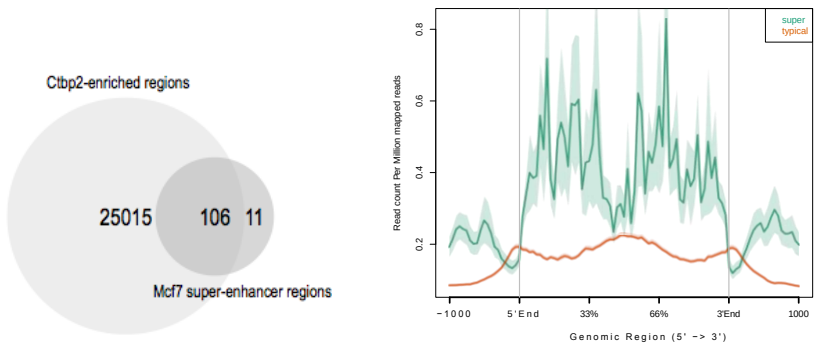


Figure 14. CTBP1/2 is enriched in MCF7 super-enhancers.

(a) Most (90.5%) MCF7 super-enhancers (41,50) are overlapped with 25,394 CTBP1/2-enriched regions. **(b)** Mean ChIP-seq density of CTBP1/2 across the 3,938 typical enhancers and the 117 super-enhancers.

IV. DISCUSSION

In previous studies Ctbp2 was identified as gating factor for exit from pluripotency (23,24). This study we aimed to define underlying mechanisms for Ctbp2 gated pluripotency exit. We showed that significant portion of Ctbp2 was occupied at active promoters and enhancers, especially 90% of the super-enhancers. Since ESC master transcription factors, ESC CTFs, which determines cell identity, are controlled by super-enhancers, epigenetic repression of super-enhancer will play a key role in differentiation and pluripotency exit (41). We verified that NuRD complex combined Ctbp2 mediate H3K27ac deacetylation, and LIF withdrawal in *Ctbp2*-knockdown ESCs fail to proper shutdown of the super-enhancer associated ESC CTFs. Additionally, we show that Ctbp2 was associated with super-enhancers and master transcriptional factors in human breast cancer cell line as well as in mESCs.

ESCs have unique transcriptional programs—active ESC genes and bivalent lineage specific genes— for maintaining self-renewal and pluripotency. Since ESCs need to activate any lineage specific genes during differentiation, bivalent chromatin status may provide the flexibility needed to prime for activation or repression (42). It was reported that the function of Lsd1 in ESCs is associated with NuRD complex and is required for proper ESC differentiation (10). NuRD function in ESCs have been shown to repress a subset of active ESC genes, such as Rex1, Tbx3, Klf4 and Klf5, in undifferentiated condition which is possibly necessary process for appropriate

ESC differentiation (9). Our study and above papers suggest that ESC active chromatin are primed for activation and repression, likewise bivalent chromatin are primed for activation and repression. The ESC super-enhancers are known for sensitivity to perturbation, Oct4 depletion preferentially reduces the expression of super-enhancer associated (41). This can be explained by the co-occupancy of the activator and repressor complex. When activation signals from LIF-Stat3 are functioning, NuRD-Ctbp2-Lsd1 mediate repressor signals remain silent, but when activation signals are disappeared repressor signal take the place of activators, and led to rapid declination of expression for ESC CTFs (Figure 15).

Although in this paper we focused only on Ctbp2 in active ESC genes, 32% of Ctbp2 are occupied weak enhancer and promoter regions. In human ESCs, there are meaningful overlaps with Ctbp2-enriched regions and H3K27me3 enriched regions (data not shown). Also, Polycomb core component SUZ12 was colocalized with CTBP2 in human ESCs (51). H3K27ac and H3K27me3 are mutually exclusive and high levels of HAT-mediated H3K27ac prevent PRC2-mediated H3K27me3 at the same genomic locus (52,53). As assumed from above, Ctbp2 may recruit PRC2 after Ctbp2-mediated deacetylation of H3K27ac; however, more studies are needed for further validation.

Cancer cells appear to acquire super-enhancers at oncogenic drivers during tumorigenesis (50). The CTBP1/2 occupancy on breast cancer super-enhancers suggests that CTBP1/2 could be a gate for cell state transitions,

which is directed to differentiated phenotype from undifferentiated stem cell-like phenotype. The recent report that low levels of *CTBP2* expression was associated with bad prognosis in ovarian cancer patient, which can be attributed to increased stemness of cancer cells, supports proceeding hypothesis (54). While CTBP is previously known to be associated with bad prognosis of cancer by mediating EMT (epithelial-mesenchymal transition) (55,56). This discrepancy is attributed to the effect of chromatin repressive complex, which is determined by the function of genes they regulate. Because of the context dependent roles of CTBP in cancer cells (oncogenic or tumor suppressive), CTBP is necessary to be investigated in each cancer or condition. However, it would be very important for further studies knowing that CTBP mediates silencing of genes in active chromatin regions when activating signals disappeared.

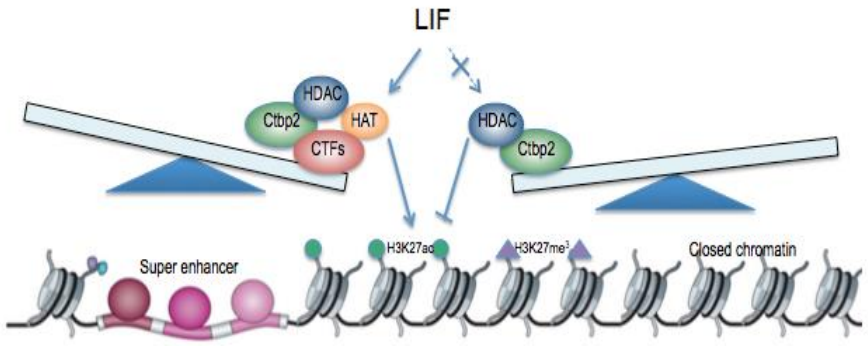


Figure 15. Schematic model for Ctbp2-mediated differentiation initiation.

V. REFERENCES

1. Boiani, M. and Schöler, H.R. (2005) Regulatory networks in embryo-derived pluripotent stem cells. *Nature Reviews Molecular Cell Biology*, **6**, 872-881.

2. Jaenisch, R. and Young, R. (2008) Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell*, **132**, 567-582.

3. Young, R.A. (2011) Control of the embryonic stem cell state. *Cell*, **144**, 940-954.

4. Ng, H.-H. and Surani, M.A. (2011) The transcriptional and signalling networks of pluripotency. *Nature cell biology*, **13**, 490-496.

5. Hu, G. and Wade, P.A. (2012) NuRD and pluripotency: a complex balancing act. *Cell stem cell*, **10**, 497-503.

6. 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. *Nature cell biology*, **8**, 285-292.

7. Kaji, K., Nichols, J. and Hendrich, B. (2007) Mbd3, a component of the NuRD co-repressor complex, is required for development of pluripotent cells. *Development*, **134**, 1123-1132.

8. McDonel, P., Costello, I. and Hendrich, B. (2009) Keeping things quiet: roles of NuRD and Sin3 co-repressor complexes during mammalian development. *The international journal of biochemistry & cell biology*, **41**, 108-116.

9. Reynolds, N., Latos, P., Hynes-Allen, A., Loos, R., Leaford, D., O'Shaughnessy, A., Mosaku, O., Signolet, J., Brennecke, P. and Kalkan, T. (2012) NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment. *Cell stem cell*, **10**, 583-594.

10. 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.

11. Ding, J., Xu, H., Faiola, F., Ma'ayan, A. and Wang, J. (2012) Oct4 links multiple epigenetic pathways to the pluripotency network. *Cell research*, **22**, 155-167.

12. Esch, D., Vahokoski, J., Groves, M.R., Pogenberg, V., Cojocaru, V., Vom Bruch, H., Han, D., Drexler, H.C., Araúzo-Bravo, M.J. and Ng, C.K. (2013) A unique Oct4 interface is crucial for reprogramming to pluripotency. *Nature cell biology*, **15**, 295-301.

13. Pardo, M., Lang, B., Yu, L., Prosser, H., Bradley, A., Babu, M.M. and Choudhary, J. (2010) An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell stem cell*, **6**, 382-395.

14. van den Berg, D.L., Snoek, T., Mullin, N.P., Yates, A., Bezstarosti, K., Demmers, J., Chambers, I. and Poot, R.A. (2010) An Oct4-centered protein interaction network in embryonic stem cells. *Cell stem cell*, **6**, 369-381.

15. Simon, J.A. and Kingston, R.E. (2009) Mechanisms of polycomb gene silencing: knowns and unknowns. *Nature reviews Molecular cell biology*, **10**, 697-708.

16. Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M. and Plath, K. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, **125**, 315-326.

17. Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A. and Ray, M.K. (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *nature*, **441**, 349-353.

18. Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F. and Isono, K.-i. (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell*, **125**, 301-313.

19. Chinnadurai, G. (2002) CtBP, an unconventional transcriptional corepressor in development and oncogenesis. *Molecular cell*, **9**, 213-224.

20. Kim, J.-H., Cho, E.-J., Kim, S.-T. and Youn, H.-D. (2005) CtBP represses p300-mediated transcriptional activation by direct association with its bromodomain. *Nature structural & molecular biology*, **12**, 423-428.

21. Shi, Y., Sawada, J.-i., Sui, G., Affar, E.B., Whetstine, J.R., Lan, F., Ogawa, H., Luke, M.P.-S., Nakatani, Y. and Shi, Y. (2003) Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature*, **422**, 735-738.

22. Hildebrand, J.D. and Soriano, P. (2002) Overlapping and unique roles for C-terminal binding protein 1 (CtBP1) and CtBP2 during mouse development. *Molecular and cellular biology*, **22**, 5296-5307.

23. Betschinger, J., Nichols, J., Dietmann, S., Corrin, P.D., Paddison, P.J. and Smith, A. (2013) Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell*, **153**, 335-347.

24. Tarleton, H.P. and Lemischka, I.R. (2010) Delayed differentiation in embryonic stem cells and mesodermal progenitors in the absence of CtBP2. *Mechanisms of development*, **127**, 107-119.

25. 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.

26. 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. *Molecular cell*, **38**, 576-589.

27. Ernst, J. and Kellis, M. (2012) ChromHMM: automating chromatin-state discovery and characterization. *Nature methods*, **9**, 215-216.

28. Hezroni, H., Tzchori, I., Davidi, A., Mattout, A., Biran, A., Nissim-Rafinia, M., Westphal, H. and Meshorer, E. (2011) H3K9 histone acetylation predicts pluripotency and reprogramming capacity of ES cells. *Nucleus*, **2**, 300-309.

29. Yu, P., Xiao, S., Xin, X., Song, C.-X., Huang, W., McDee, D., Tanaka, T., Wang, T., He, C. and Zhong, S. (2013) Spatiotemporal clustering

of the epigenome reveals rules of dynamic gene regulation. *Genome research*, **23**, 352-364.

30. Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W.H., Ye, C., Ping, J.L.H. and Mulawadi, F. (2011) CTCF-mediated functional chromatin interactome in pluripotent cells. *Nature genetics*, **43**, 630-638.

31. Neph, S., Kuehn, M.S., Reynolds, A.P., Haugen, E., Thurman, R.E., Johnson, A.K., Rynes, E., Maurano, M.T., Vierstra, J. and Thomas, S. (2012) BEDOPS: high-performance genomic feature operations. *Bioinformatics*, **28**, 1919-1920.

32. McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M. and Bejerano, G. (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology*, **28**, 495-501.

33. de Hoon, M.J., Imoto, S., Nolan, J. and Miyano, S. (2004) Open source clustering software. *Bioinformatics*, **20**, 1453-1454.

34. Trapnell, C., Pachter, L. and Salzberg, S.L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, **25**, 1105-1111.

35. Roberts, A., Trapnell, C., Donaghey, J., Rinn, J.L. and Pachter, L. (2011) Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol*, **12**, R22.

36. Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome biol*, **11**, R106.

37. Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139-140.

38. Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research*, **37**, 1-13.

39. Consortium, E.P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57-74.

40. Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W. and Jiang, J. (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*, **133**, 1106-1117.

41. 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.

42. Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K. and Koche, R.P. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, **448**, 553-560.

43. Zhao, L.-J., Subramanian, T., Vijayalingam, S. and Chinnadurai, G. (2014) CtBP2 proteome: Role of CtBP in E2F7-mediated repression and cell proliferation. *Genes & cancer*, **5**, 31.

44. Srinivasan, L. and Atchison, M.L. (2004) YY1 DNA binding and PcG recruitment requires CtBP. *Genes & development*, **18**, 2596-2601.

45. Kaneko, S., Son, J., Shen, S.S., Reinberg, D. and Bonasio, R. (2013) PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. *Nature structural & molecular biology*, **20**, 1258-1264.

46. Dovey, O.M., Foster, C.T. and Cowley, S.M. (2010) Histone deacetylase 1 (HDAC1), but not HDAC2, controls embryonic stem cell

differentiation. *Proceedings of the National Academy of Sciences*, **107**, 8242-8247.

47. Di, L.-J., Byun, J.S., Wong, M.M., Wakano, C., Taylor, T., Bilke, S., Baek, S., Hunter, K., Yang, H. and Lee, M. (2013) Genome-wide profiles of CtBP link metabolism with genome stability and epithelial reprogramming in breast cancer. *Nature communications*, **4**, 1449.

48. Zaret, K.S. and Carroll, J.S. (2011) Pioneer transcription factors: establishing competence for gene expression. *Genes & development*, **25**, 2227-2241.

49. Theodorou, V., Stark, R., Menon, S. and Carroll, J.S. (2013) GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome research*, **23**, 12-22.

50. Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, 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.

51. Nielsen, C.B., Younesy, H., O'Geen, H., Xu, X., Jackson, A.R., Milosavljevic, A., Wang, T., Costello, J.F., Hirst, M. and Farnham, P.J. (2012) Spark: a navigational paradigm for genomic data exploration. *Genome research*, **22**, 2262-2269.

52. Klymenko, T. and Müller, J. (2004) The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO reports*, **5**, 373-377.

53. Tie, F., Banerjee, R., Stratton, C.A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M.O., Scacheri, P.C. and Harte, P.J. (2009) CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. *Development*, **136**, 3131-3141.

54. Cui, T.X., Kryczek, I., Zhao, L., Zhao, E., Kuick, R., Roh, M.H., Vatan, L., Szeliga, W., Mao, Y. and Thomas, D.G. (2013) Myeloid-derived suppressor cells enhance stemness of cancer cells by inducing microRNA101 and suppressing the corepressor CtBP2. *Immunity*, **39**, 611-621.

55. Chinnadurai, G. (2009) The transcriptional corepressor CtBP: a foe of multiple tumor suppressors. *Cancer research*, **69**, 731-734.

56. Sanchez-Tillo, E., Lazaro, A., Torrent, R., Cuatrecasas, M., Vaquero, E., Castells, A., Engel, P. and Postigo, A. (2010) ZEB1 represses E-cadherin and induces an EMT by recruiting the SWI/SNF chromatin-remodeling protein BRG1. *Oncogene*, **29**, 3490-3500.

ABSTRACT IN KOREAN

국문 초록

세포는 같은 유전 서열을 공유하지만, 후생유전학적 조절을 통해 서로 다른 특성을 띠게 된다. 후생 유전학적 조절은 염색질의 구조 변화를 통하여 이루어지는데 이는 유전자의 발현을 조절하기 때문에 세포의 기능 및 운명 결정에 필수적이다. 이 때문에 염색질의 조절 인자들은 개체의 발생, 줄기 세포의 유지 및 분화 등 세포의 특성을 유지 또는 변화시켜가는 과정에서 주요한 역할을 담당한다. 염색질 조절 인자이며 전사 억제 기능을 담당하는 C-terminal binding protein 2(Ctbp2)는 배아 줄기 세포가 전 분화능을 잃고 분화해 가는 과정에서 필수적인 것으로 알려졌지만, 그 세부 기전이 잘 알려지지 않았다.

본 연구에서는 Ctbp2 가 전사가 활성화된 유전자의 조절 영역에서 배아 줄기 세포 핵심 전사 인자와 함께 있으며, 분화 과정에서 Nucleosome remodeling and deacetylation(NuRD) 복합체와 함께 히스톤 H3-K27 의 탈 아세틸화를 매개함을 밝힐 수 있었다. 또한, 줄기 세포 핵심 전사 인자의 조절 영역인 super-enhancer 의 대부분에 Ctbp2 가 위치하였음을 확인하였다. 분화 신호가 주어지기 전, Ctbp2 기능을 억제하는 경우 H3-K27 의 아세틸화가 증가 되지만 핵심 전사 인자들의 유전자 발현 정도는 변화하지 않는다. 반면 Ctbp2 의 기능이 억제된 상태에서는 분화

신호가 주어져도 H3-K27 의 탈 아세틸화를 통한 핵심 전사 인자들의 유전자 발현 억제가 이루어지지 않아 세포는 분화하지 못하고 전 분화 능을 유지한 상태에 머물게 된다. 이를 통해 super-enhancer 영역이 다른 일반적 enhancer 와 달리 분화 신호에 민감하게 유전자 발현이 빠르게 감소하는 데에는 분화 전부터 super-enhancer 에 있는 Ctbp2 와 NuRD 복합체의 기능이 중요하다는 것을 알 수 있었다. 사람의 유방암세포에서 genomewide ChIP-sequencing 분석을 통하여 super-enhancer 영역에 CTBP1/2 가 위치 함을 확인하여 앞선 기전이 마우스 배아 줄기 세포뿐만이 아니라 사람 암세포의 분화 및 암의 진행 등의 과정에서도 중요한 역할을 담당할 수 있음을 알 수 있었다.

주요어 : 배아 줄기 세포, 분화, 핵심 전사 인자, C-terminal binding protein 2 (Ctbp2), nucleosome remodeling and deacetylation (NuRD), 히스톤 탈 아세틸화, super-enhancer

학 번 : 2009-22240