이학박사 학위논문

The epigenetic modification in embryonic stem cell pluripotency

후성유전학에 의한 줄기세포 전분화능 조절

2014 년 02 월

서울대학교 대학원

의과학과 의과학 전공

김 태 완

A thesis of the Degree of Doctor of Philosophy

후성유전학에 의한 줄기세포 전분화능 조절

The epigenetic modification in embryonic stem cell pluripotency

February 2014

The Department of Biomedical Science
Seoul National University
College of Medicine
Tae Wan Kim

ABSTRACT

Pluripotent stem cells, including embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs) are promising sources for regenerative medicine because of their ability to differentiate into all of the cell types in the body. In particular, iPSCs, which can be derived from patient tissue, have great potential to be used in patient- specific cell therapy. However, the practical application of such cells remains a distant goal due to our limited understanding of molecular mechanisms underlying pluripotency. Epigenetic modifications on chromatin are essential for appropriate cell state transition, such as developmental process, and uncontrolled these modifications give rise to disease. So, it is necessary to study epigenetic modifications in ESC pluripotency for regenerative medicine.

First, we investigated that the role of *O*-linked-*N*-acetylglucosamine (*O*-GlcNAc) in ESC maintenance and differentiation, and in somatic cell reprogramming. Because, *O*-GlcNAc has emerged as a critical regulator of diverse cellular processes, but its role in ESCs and pluripotency has not been investigated yet. Here we show that *O*-GlcNAcylation directly

regulates core components of the pluripotency network. Blocking *O*-GlcNAcylation disrupts ESC self-renewal and reprogramming of somatic cells to iPSCs. The core reprogramming factors Oct4 and Sox2 are *O*-GlcNAcylated in ESCs, but the *O*-GlcNAc modification is rapidly removed upon differentiation. *O*-GlcNAc modification of Threonine 228 in Oct4 regulates Oct4 transcriptional activity and is important for inducing many pluripotency related genes, including *Klf2, Klf5, Nr5a2, Tbx3* and *Tcl1*. A T228A point mutation that eliminates this *O*-GlcNAc modification reduces the capacity of Oct4 to maintain ESC self-renewal and reprogram somatic cells. Overall, our study makes a direct connection between *O*-GlcNAcylation of key regulatory transcription factors and the activity of the pluripotency network.

Second, we investigated the stable silencing of active ESC genes during differentiation. Pluripotency is governed by a core transcription factor (CTF) network to establish autoregulatory circuitry in ESCs. For proper exit from pluripotency to lineage commitments, CTF circuitry should be extinguished in an orderly manner by epigenetic regulation. Yet, how the stable silencing marker H3K27 trimethylation (H3K27me3) is

recruited to active ESC genes during differentiation is poorly understood. Here we show that Ctbp2 is involved in Polycomb repressive complex 2 (PRC2)–mediated silencing of active ESC genes during differentiation. Ablation of Ctbp2 leads to inappropriate gene silencing in ESCs, thereby sustaining ESC maintenance during differentiation. Ctbp2 in ESCs associates with core components of PRC2 and Oct4. Ctbp2 regulates the recruitment of Ezh2 to active ESC genes and stimulates H3K27me3 during differentiation, which is pivotal for lineage commitment.

Keywords: ESC, *O*–GlcNAcylation, Reprogramming, Ctbp2, Pluripotency, Core transcription factors, PRC2, Gene silencing, iPSC

Student number: 2009–30600

CONTENTS

LIST OF FIGURES

Results (Part1)

Results (Part2)

LIST OF ABBREVIATIONS

ESC (Embryonic stem cell)

MEF (Mouse embryonic fibroblast)

iPSC (Induced pluripotent stem cell)

GlcNAc (*N*–acetylglucosamine)

OGT (*O*–GlcNAc transferase)

OGA (*O* –GlcNAcase)

EB (Embryoid body)

AP (Alkaline phosphatase)

STZ (Streptozotocin)

Dox (Doxycycline)

WGA (Wheat germ agglutinin)

CHX (Cycloheximide)

CTF (Core transcription factor)

Ctbp2 (C–terminal binding protein 2)

Ctbp1 (C–terminal binding protein 1)

PRC2 (Polycomb repressive complex 2)

I . INTRODUCTION

Embryonic stem cells (ESCs) are self-renew indefinitely and pluripotent that they can differentiate into all cell types representative of all three embryonic germ layers. Differentiated somatic cells can be reprogrammed into induced pluripotent stem cells (iPSCs), which resemble ESCs. Because of their ability to differentiate into all cell types, stem cells are promising sources for regenerative medicine. Especially iPSCs have great potential to be used in patient- specific cell therapy. However, the practical application of such cells remains a distant goal due to our limited understanding of molecular mechanisms underlying pluripotency (Hanna et al., 2010; Ng and Surani, 2011; Okita and Yamanaka, 2010; Patel and Yang, 2010).

Epigenetic modifications on chromatin are indispensable for appropriate embryogenesis and/or cellular differentiation and uncontrolled these modifications fail to specialized lineage commitments, indicating that it is essential to study epigenetic modifications in ESC pluripotency for successful patient-specific cell therapies (Portela, A. and Esteller, M, 2010).

Metabolites caused by metabolism are used for epigenetic

modifacations as substrates. For exemple, kinases use ATP for phosphorylation , methyltrasnferases use SAM for methylation, acethyltransferases use acethyl-CoA for acethylation, and *O*-glunactransferase uses UDP-GlcNAc for *O*-GlcNAcylation, suggesting that epigenetic modifications and metabolism are deeply interconnected.

Oct4 was known as a master regulator for ESC pluripotency and reprogramming. Because of its functional importance, four independent groups have published purification of Oct4 complex from mouse ESCs (Van den Berg et al., 2010; Pardo et al., 2010; Ding et al., 2012; Esch et al., 2013). Among them, Ogt and Ctbp2 are selected as 10 common Oct4-associating proteins from these published reports, indicating that Ogt and Ctbp2 probably contribute to Oct4 function in ESC identity. Also, the function of these genes is associated with metabolism, such as Ogt uses UDP-GlcNAc as a substrate for *O*-GlcNAcylation and Ctbp2 act as a redox sensor (Zhang et al., 2002), suggesting that these genes contribute to ESC pluripotency by epigenetic modification(s).

O-GlcNAcylation is an *O*-β-glycosidic attachment of single *N*-acetylglucosamine (GlcNAc) to serine and threonine

residues of nucleocytoplasmic proteins. Similar to phosphorylation, this posttranslational modification rapidly cycles on and off hundreds of protein substrates in response to diverse biological cues. However, unlike phosphorylation , only two enzymes catalyze the addition and removal of *O*−GlcNAc in mammals; *O*−GlcNAc transferase (OGT) adds the modification and *O*−GlcNAcase (OGA) removes it (Hanover et al., 2010; Hu et al., 2010; Love et al., 2010; Zeidan and Hart, 2010). Several studies have shown that *O*−GlcNAcylation controls diverse aspects of cellular physiology, such as nutrient and growth factor sensing, cell cycle progression, and stress response by regulating protein function, localization, and stability (Hanover et al., 2010; Hart and Copeland, 2010; Hart et al., 2007; Love et al., 2010). In particular, *O*−GlcNAcylation adjusts protein function according to the nutritional status of the cell (Zeidan and Hart, 2010).

Glucose/Nutrient conditions in culture affect ESC maintenance and differentiation, but the underlying mechanisms are poorly understood. *In vitro* development of early human and mouse embryos is enhanced in glucose−deficient media (Conaghan et al., 1993; Quinn, 1995) while addition of glucosamine during *in*

vitro maturation inhibits embryo development (Sutton-McDowall et al., 2006). In addition, glucose levels during formation of embryoid bodies (EBs) influence their differentiation propensity, with high glucose generating beating cardiac muscle more efficiently, and low glucose promoting neuronal lineages (Mochizuki et al., 2011). Complete loss of Ogt function is lethal to mouse ES cells (Shafi et al., 2000). EBs with elevated *O*-GlcNAc levels, through treatment with glucosamine or OGA inhibitors, have impaired differentiation into cardiac cells (Kim et al., 2009). Together these studies suggest that, *O*-GlcNAcylation promotes ES cell maintenance, and decreases in *O*-GlcNAcylation may be required for ES cell differentiation.

The transcription factors, Oct4, Sox2, Klf4 (collectively termed OSK), and Nanog are part of a core pluripotency network that controls ESC self-renewal and pluripotency, and are also core factors used in reprogramming somatic cells to iPSCs (Ng and Surani, 2011; Stadtfeld and Hochedlinger, 2010). Previous studies have shown that Sox2 undergoes *O*-GlcNAcylation in rat forebrain (Khidekel et al., 2004), and Oct4 interacts with Ogt (Pardo et al., 2010; van den Berg et al., 2010)

and may undergo *O*-GlcNAcylation (Webster et al., 2009) in ESCs. These data suggest that *O*-GlcNAc could directly affect core components of the pluripotency network to functionally regulate pluripotency and somatic cell reprogramming.

In the present study, we analyzed the role of *O*-GlcNAc in ESC maintenance and differentiation, and in somatic cell reprogramming. We found that blocking *O*-GlcNAcylation inhibited ESC self-renewal and the efficiency of iPS cell generation, while increasing *O*-GlcNAcylation during EB formation inhibited normal ESC differentiation. Moreover, we determined that Oct4 and Sox2 are modified by *O*-GlcNAc in mouse ESCs and *O*-GlcNAc regulates their transcriptional activity. Our study reveals that *O*-GlcNAc controls pluripotency by directly regulating transcriptional activities of core components of the pluripotency network. These findings will expand our understanding of the reprogramming process and pluripotency at a molecular level.

Furthermore, pluripotency is governed by a core transcription factor (CTF) network to establish autoregulatory circuitry in ESCs (Jaenisch and Young, 2008). For proper exit from pluripotency to lineage commitments, CTF circuitry should be

extinguished in an orderly manner by epigenetic regulation (Mikkelsen et al., 2007). Epigenetic factors, including the LSD1-NuRD, silence active ESC genes (Reynolds et al., 2012; Whyte et al., 2012). Yet, how the stable silencing marker H3K27 trimethylation (H3K27me3) is recruited to active ESC genes during differentiation is poorly understood. C-terminal binding protein 2 (Ctbp2), a transcriptional co-repressor, is a core Oct4-associating protein (Esch et al., 2013), but its epigenetic function in ESCs is unknown. Here we show that Ctbp2 is involved in Polycomb repressive complex 2 (PRC2)-mediated silencing of active ESC genes during differentiation. Ablation of Ctbp2 leads to inappropriate gene silencing in ESCs, thereby sustaining ESC maintenance during differentiation. Ctbp2 in ESCs associates with core components of PRC2 and Oct4. Ctbp2 regulates the recruitment of Ezh2 to active ESC genes and stimulates H3K27me3 during differentiation. We demonstrate that Ctbp2 guides PRC2 toward active ESC genes to mark H3K27me3 for stable silencing, which is pivotal for lineage commitment.

Ⅱ. MATERIALS AND METHODS

1. Mice

C57BL/6 and pOct4-GFP mice (Jackson Laboratory) used for mouse embryonic fibroblasts (MEFs) preparation were maintained under specific pathogen-free conditions at the Center for Animal Resource Development of Seoul National University College of Medicine as described (Cho et al., 2010). Animal experimental procedures were approved by Institutional Animal Care and Use Committee (IACUC) of Seoul National University, Korea (SNU-070419-4).

2. Cell Culture

E14 mouse ESCs were cultured on γ-irradiated MEFs or STO feeder layers in a 0.1% gelatin (Sigma-Aldrich) coated tissue culture dish as described (Jang et al., 2009). The mouse ESC medium was composed of DMEM (Welgene) and 15% (v/v) fetal bovine serum (FBS; PAA) or KNOCKOUT^{TM} DMEM (Gibco) and 15% (v/v) KNOCKOUT^{TM} SR (Gibco), supplemented with 2 mM L-glutamine, 55 μM β-mercaptoethanol, 1% (v/v) non-essential amino acid, 100 U/ml penicillin and 100 μg/ml

streptomycin (all from Gibco), and 1000 U/ml ESGRO (Millipore). To prepare RNAs and cellular proteins without feeder cell contamination, ESCs and iPSCs were at least 2 passaged in a 0.1% gelatin coated tissue culture dish without feeder layers.

3. Reprogramming

Retroviral expression vectors for murine Oct4, Sox2, Klf4, and their point mutants were cloned into IRES-GFP-deleted pMIG vector (Addgene). Retroviruses were produced using Plat-E system (Morita et al., 2000). Plat-E was kindly provided by T. Kitamura (University of Tokyo, Japan). Reprogramming was principally done per the Lim group (Han et al., 2010). Briefly, equal amounts of virus encoding Oct4, Sox2, and Klf4 were applied to 5×10^4 plated MEFs in 10-cm^2 dishes in 10% FBS DMEM media containing 8 $ngml^{-1}$ polybrene. After 24 hr, inactivated feeder cells and fresh mESC culture media were added, and the culture was then maintained for up to 21 days. Alkaline phosphatase (AP) staining was done using AP detection kit (Millipore). The expression of endogenous Oct4 was monitored by the expression of Oct4 promoter-driven GFP

using a fluorescent microscope (OLYMPUS, IX71). For comparing the effect of glucose concentration on reprogramming efficiency, we made mESC culture media using both high and low glucose DMEM (Gibco 11995 and 11885).

4. Retroviral shRNA-Mediated Knockdown of Ogt

For construction of plasmids for shRNA synthesis, 19 base-pair gene-specific regions were cloned into pMKO.1 puro (Addgene). Retroviruses were produced using Plat-E system (Morita et al., 2000). Oligonucleotide sequences are

#1 : CTGATTTCATTGATGGTTA,

#2 : CTGGATACTCCTTTGTGTA,

#3: CGACTACACAGATTAATAA. pMKO.1 puro vector was used as a control.

5. Lentiviral shRNA-Mediated Knockdown of *Ctbp2*

PLKO-puro constructs expressing shRNAs were purchased from Sigma. The shRNA constructs targeting Ctbp2 are TRCN0000109335 and TRCN0000307495. Lentiviruses were produced with 293FT cell by co-transfection of 1g each of pMD2.G, pMDLg/pRRE, pRSV-rev, pLKO-shRNA using

Lipofectamin (Invitrogen). 48 h after transfection, virus-containing medium was collected and filtered by 0.45m filters. Polybrene (10ug ml^{-1}) was added just before infection to target cells and infection was performed for 5 h. Puromysin selection (2ug ml^{-1}) was performed 48 h after infection for 2 days. As a control, we used pLKO.1 puro vector.

6. Self-Renewal Assay

Self-renewal assay (colony-forming assay) was done as described (Chambers et al., 2007). Briefly, ESCs were trypsinized to a single cell suspension and re-plated at 500 cells per 1 well in 6-well plates. Dox was added 1 day after re-plating. After incubation for 7 days with/without Dox, plates were stained for alkaline phosphatase and scored for differentiation status.

7. DNA Constructs and Chemicals

cDNAs of murine Oct4, Sox2, Klf4, and Nanog were generated by PCR amplification from E14 cells and were cloned into pcDNA3.1/myc-His(-) B (Invitrogen). Expression vector for HA-tagged Oct4 and HA-tagged Oct4 deletion mutants were

generated by subcloning PCR product into pDEST-HA and Flag-tagged Oct4 deletion mutants (⊿POU and ⊿POU-h) were generated by subcloning PCR product into pcDNA3.0-Flag (Invitrogen). To produce retroviruses co-expressing Flag-tagged Oct4, Sox2, hOGT and a puromycin resistant gene, genes were subcloned into pMSCV-Flag puro, which was generated by inserting a Flag tag into the pMSCVpuro (Clonetech). For long-term transgene expression in ESCs, Flag-tagged Oct4 and Oct4 T228A mutant were subcloned into a pCAG-IP vector which were kindly provided by Hitoshi Niwa (RIKEN, Japan). Various point mutants of Oct4 and Sox2 were generated by site-directed mutagenesis (iNtRON). Streptozocin (streptozotocin, STZ), glucosamine, and N-acetylglucosamine (GlcNAc) were purchased from Sigma.

8. *In Vitro* Differentiation of ESCs

ESC differentiation was induced by LIF withdrawal of ESC medium in monolayer cultures. For embryoid body (EB) formation, ESCs were cultured in low attachment dishes containing ESC medium without LIF. Oct4 depletion mediated differentiation of ZHBTc4ESCs was done as described (Niwa et

al., 2000).

9. Generation of Stable Expression of Either Flag-tagged CTBP2 or EED ESCs

For long-term transgene expression in E14 ESCs, a Flag-tagged human CTBP2 was cloned into a pCAG-Flag-IP vector which was generated by inserting a Flag tag into a pCAG-IP vector which was kindly provided by Hitoshi Niwa (RIKEN, Japan). For production of EED overexpression ESC, human EED (Addgene plasmid 24231) was subcloned into a pCAG-Flag-IP vector. To generate ESC overexpressing either Flag-tagged CTBP2 or EED, either pCAG-Flag-CTBP2-IP or pCAG-Flag-EED-IP vector was transfected into ESCs using lipofectamine (Invitrogen). After 48h trasnfection, 2ug ml^{-1} puromycin selection was performed for stable integration. Resistant cells were expanded and tested for CTBP2 and EED expression by western blotting.

10. Generation of Rescued-CTBP2/*Ctbp2*-knockdown ESCs

To rescue Ctbp2 in *Ctbp2*-knockdown ESCs, human CTBP2

gene was cloned into a pCAG-Flag-IB vector which was generated by substitution of blasticidine for puromycin into a pCAG-Flag-IP vector. To generate rescued-CTBP2 in *Ctbp2*-knockdown ESC, a pCAG-Flag-CTBP2-IB vector was transfected into *Ctbp2*-knockdown ESC using lipofectamine (Invitrogen) and 8ug ml^{-1} blasticidine selection was performed for stable integration after 48 h. Resistant cells were expanded and tested for CTBP2 expression by western blotting.

11. sWGA Purification, Immunoprecipitation, Western blot and Immunofluorescence.

Agarose bound succinylated wheat germ agglutinin (sWGA) was purchased from Vector laboratories; *O*-N-Acetyl-hexosaminidase (*O* -hex) and PNGase F were from New England BioLabs. sWGA pull-down assay and treatment *O* -hex or PNGase F were done per the manufacture's instruction. Immunoprecipitation, Immunofluorescence and Western blot were done as described (Jang et al., 2009). Anti-Ogt (O0624) antibody was purchased from Sigma; anti-Klf4 (sc-20691) and anti-Oct4 (sc-5279) antibodies were from Santa Cruz Biotechnology; anti-Sox2 (MAB2018) antibody was from R&D

systems; anti-Nanog (ab14959) and anti-*O*-GlcNAc (RL Ⅱ) antibodies were from Abcam; anti-*O*-GlcNAc (CTD110.6) antibody was from Covance. Anti-Ctbp1 (612042), anti-Ctbp2 (612044) and anti-Ezh2 (612667) antibodies were purchased from BD Transduction Laboratories; anti-Mi-2b (ab72418), anti-Hdac1(ab7028) and anti-Lsd1 (ab17721) antibodies were from Abcam; anti-Esrrb (H6707) and anti-Nr5a2 (H2325) antibodies were from R&D System; anti-Hdac2 (51-5100) antibody was from Invitrogen.

12. Real-Time qPCR

Total RNA was extracted by Trizol (Invitrogen) and cDNA synthesis was done according to the manufacturer's protocol (AMV Reverse Transcriptase, Takara). Quantitative Real time-PCR was performed using the SYBR Green qPCR Kit (Finnzymes, F-410L) on the iQ5 Real-time PCR Detection System (Bio-Rad). By the $2\Delta\Delta^{CT}$ calculation method, we analyze relative mRNA expression levels. Primers used for real-time PCR are listed below.

13. Chromatin Immunoprecipitation (ChIP) Assay

Chromatin immunoprecipitation (ChIP) assay were done as described[1]. Briefly, cells were cross-linked with 1% (v/v) formaldehyde for 10 min at 37° C. After washing twice with cold phosphate buffer saline, cross-linked cells were lysed with buffer containing 85mM KCl, 5mM PIPES (pH 8.0) and 0.5% (v/v) NP-40. After centrifugation, nuclear pellets were resuspended in buffer containing 10mM EDTA, 100mM Tris-Cl (pH 8.1) and 1% SDS. To solubilize and shear cross-linked DNA, lysates were sonicated. 10-fold diluted lysates in IP buffer containing 16.7mM Tris-Cl (pH 8.1), 167mM NaCl, 0.01% (w/v) SDS, 1.1% (v/v) Triton X-100 and 1.2mM EDTA were incubated overnight at 4° C with protein A/G PLUS agarose and 1~2ug appropriate antibodies. Beads were washed with 1X 20mM Tris-Cl (pH 8.1), 150mM NaCl, 0.1% (w/v) SDS, 1% Triton X-100, 2mM EDTA, 1X 20mM Tris-Cl (pH 8.1), 150mM NaCl, 0.1% (w/v) SDS, 1% Triton X-100, 2mM EDTA, 1X 0.25M LiCl, 1% (v/v) NP40, 1% (v/v) deoxycholate, 1mM EDTA, 10mM Tris-Cl (pH 8.1) and 2X with TE. Immune complexes were then eluted by adding 250㎕ of elution buffer containing 1% (w/v) SDS and 0.1M $NaHCO_3$ two times. Cross-

linking was reversed by adding of 20㎕ of 5M NaCl and incubated overnight at 65° C. DNA was precipitated with ethanol and kept in −20° C for the further use. Primers used for ChIP−qPCR are listed below.

	Gene	5' primer	3' primer
Real-time PCR	*Oct4* (total)	GGCGTTCTCTTTGGAAAGGTGTTC	CTCGAACCACATCCTTCTCT
	Oct4 (endo)	AGGCAAGGGAGGTAGACAAGAGAA	AAATGATGAGTGACAGACAGGCCAGG
	Oct4 (transgene)	GGGCTGTATCCTTTCCTCTGCC	CGATAAGCTTGGCTGCAGGAATT
	Sox2 (total)	TCCGGGAAGCGTGTACTTATCCTT	CCAAGATGCACAACTCGGAGATCA
	Sox2 (endo)	GAGAGCAAGTACTGGCAAGACCG	ATATCAACCTGCATGGACATTTTT
	Nanog	ATGAAGTGCAAGCGGTGGCAGAAA	CCTGGTGGAGTCACAGAGTAGTTC
	Ogt	TGCTCTCAAAGAGAAGGGCAGTGT	TGTTCCCGTTTGATGTTGGCAAGG
	18s rRNA (control)	TTAGAGTGTTCAAAGCAGGCCCGA	TCTTGGCAAATGCTTTCGCTCTGG
	Actb(control)	ATCACTATTGGCAACGAGCG	TCAGCAATGCCTGGGTACAT
	Gata4	CCGAGCAGGAATTTGAAGAGG	GCCTGTATGTAATGCCTGCG
	Afp	TCGTATTCCAACAGGAGG	AGGCTTTTGCTTCACCAG
	SMA	ATTGTGCTGGACTCTGGAGATGGT	TGATGTCACGGACAATCTCACGCT
	Pitx2	CGGCAGAGGACTCATTTCAC	CTTCCGTAAGGTTGGTCCAC
	Klf2	CGCACCTAAAGGCGCATCTG	TTCGGTAGTGGCGGGTAAGC
	Klf4	ACAGGCGAGAAACCTTACCACTGT	GCCTCTTCATGTGTAAGGCAAGGT
	Esrrb	AGTACAAGCGACGGCTGGAT	CCTAGTAGATTCGAGACGATCTTAGTCA
	Zfp42(Rex1)	CAGCTCCTGCACACAGAAGA	ACTGATCCGCAAACACCTG
	Utf1	GGACCCTTCGATAACCAGATCC	CAGGTTCGTCATTTTCCGCA
ChIP (Oct4 RE)	*Nanog*	TTAGATCAGAGGATGCCCCCTAAG	CTCCTACCCTACCCACCCCCTATT
	Klf2	AAGAGCCTCCAGGTGGGAAG	CTGGAGACCTGGGCTCATTG
	Zfp42	CGATCGGGATTCTGAAGAGGC	GCCCAATGAGAAGCGCTCAA

	Oct4	GGCGTTCTCTTTGGAAAGGTGTTC	CTCGAACCACATCCTTCTCT
	Sox2	TCCGGGAAGCGTGTACTTATCCTT	CCAAGATGCACAACTCGGAGATCA
	Nanog	ATGAAGTGCAAGCGGTGGCAGAAA	CCTGGTGGAGTCACAGAGTAGTTC

Real – time PCR	*Nr0b1 (Dax1)*	TCCAGGCCATCAAGAGTTTC	ATCTGCTGGGTTCTCCACTG
	Esrrb	AGTACAAGCGACGGCTGGAT	CCTAGTAGATTCGAGACGATCTTAGTCA
	Zfp42(Rex1)	CAGCTCCTGCACACAGAAGA	ACTGATCCGCAAACACCTG
	Klf2	CGCACCTAAAGGCGCATCTG	TTCGGTAGTGGCGGGTAAGC
	Klf4	ACAGGCGAGAAACCTTACCACTGT	GCCTCTTCATGTGTAAGGCAAGGT
	Fgf4	CTCTTCGGTGTGCCTTTCTTTACCGA	AGGAAGGAAGTGGGTTACCTTCATGG
	Nr5a2	TCTCACACACAGAAGTCGCGTTCA	TGCAGGTTCTCCAGGTTCTTCACA
	Ctbp1	ACGTCCTCTTCTATGATCCA	TTGTCTCATCTGCTTGACAG
	Ctbp2	ACCACACACAGCCTGGTACA	GTGAATTGCTTGCTGGTCAA
	18s rRNA (control)	TTAGAGTGTTCAAAGCAGGCCCGA	TCTTGGCAAATGCTTTCGCTCTGG
	Actb(control)	ATCACTATTGGCAACGAGCG	TCAGCAATGCCTGGGTACAT
ChIP	*Pecam1*	CTTGAACTGTGGTGGAGAGTGCTG	TAAGGAAGGGCTAGGACGAGAGG
	Nanog	TTAGATCAGAGGATGCCCCCTAAG	CTCCTACCCTACCCACCCCCTATT
	Klf2	AAGAGCCTCCAGGTGGGAAG	CTGGAGACCTGGGCTCATTG
	Zfp42	CGATCGGGATTCTGAAGAGGC	GCCCAATGAGAAGCGCTCAA
	Nr5a2	CGGTTGCAATTACAAAGGATGA	CGCCGAGGCTGGATAAA
	T	ACAGATCCGCATTGAGCTTC	CCCATGGACTGCCACTAACT
	Fgf5	AGGGACGGTCAAGATTCCTT	AGAACCAGCAGAGTCCCAGA
	Gata6	CCTTCCCATACACCACAACC	CCCCTCCTTCCAAATTAAGC
	Hoxd9	GGATAATCGCCTAGGTGTGACTTAG	CATCTCTTCTTGCCTCTCTGGG

Primers used in a thesis

14. Reporter Gene Assay

The reporter plasmid pOct4 (10×) TATA luc and 6×O/S *Fgf4* promoter luc were kindly provided by Jungho Kim (Sogang University, Korea) and Lisa Dailey (New York University School of Medicine). 10×Oct4 RE-luc and 6×O/S RE-luc were generated by subcloning the provided plasmids into pMSCVpuro (Addgene). To stably incorporate reporter genes into genomic DNA, cells infected with retrovirus containing reporter genes were selected with puromycin for at least 2 weeks. Reporter gene activity was measured as described (Jang et al., 2009).

15. Nano-LC-ESI-MS/MS Analysis of *O*-GlcNAc Sites in Oct4

For the identification of *O*-GlcNAc sites in Oct4, gels were treated with endoproteinase Glu-C (Roche, Mannheim, Germany) at 37℃ for overnight. Following digestion, proteolytic peptides were extracted and dried with SpeedVac. Resuspended samples in 0.1% formic acid were purified and concentrated using C18 ZipTips (Millipore, MA) before MS analysis. The proteolytic peptides were loaded onto a fused

silica microcapillary column (12 cm x 75 μm) packed with C18 reversed phase resin (5 μm, 200 Å). Peptide separation was conducted a series of step gradients composed of initial isobaric flow for 5 min with 3% solvent B (0.1% formic acid in acetonitrile), then linear gradient from 3% to 35% for 30 min, followed by linear gradient to 40% for 5 min. At the end of each running, 90% of solvent B was eluted for 10 min with the rate of 250 nL/min. The % gradient of solvent B was against solvent A (0.1% formic acid in H_2O). The column was directly connected to LTQ linear ion-trap mass spectrometer (Finnigan, CA) equipped with a nano-electrospray ion source. Each full MS scan was followed by five MS/MS scan corresponding from the most intense to the fifth intense peaks of full MS scan. The acquired LC-ESI-MS/MS fragment spectra were searched in the BioWorksBrowser™ (version Rev. 3.3.1 SP1, Thermo Fisher Scientific Inc., CA) with the SEQUEST search engines against in the data in FASTA format generated from Pou5f1, (NCBI accession number NM_013633) in National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov). The conditions for the search were as follws; Glu-C as enzyme specificity, a permissible level for two missed cleavages,

peptide tolerance; ±2 amu, a mass error of ±1 amu on fragment ions and variable modification of cysteine (+57 Da), oxidation of methionine (+16 Da) and *O*-GlcNAc of serine or threonine (+203.2 Da) residues.

16. ChIP-sequencing

DNA fragments were ligated to a pair of adaptors for sequencing on the Illumina Genome Analyzer. 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 Genome Analyzer II following the manufacturer's recommended protocol.

17. Mapping and Identification of Marked Region

Reads after sequencing were mapped against the mouse reference genome (UCSC genome browser, ver. mm9) using bwa version 0.6.0 (Li and Durbin, 2009) 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 (ver. 1.4.0) (Zhang et al., 2008) tool to select binding sites for transcription factors. We applied default settings and found significant regions compared to matched control samples. Candidates having concordant regions among the three TFs were selected and the closest distance from summit peaks to TSS or TES were calculated for finding nearest genes. Finally, summit peaks having significant level (adjusted *P*-value <= 0.01) within a range of 10 kb from TSS and TES were analyzed and visualized by using IGV (Robinson et al., 2011).

18. Gel-Filtration Analysis

Undifferentiated and differentiated E14 cells were lysed in IP150 buffer (150mM NaCl, 25mM Tris 8.0, 0.1% NP-40, 10% glycerol, 1mM EDTA). Whole cell lysates were fractionized by Superdex 200 gel filtration column using AKTA FPLC system (Amersham Pharmacia Biotech). Ctbp2 and PRC2 components in fractions of lysates were analyzed by SDS-PAGE and western blotting.

19. Mass Spectrometry Analysis

To identify proteins that are specifically interacting with Ctbp2

in ESCs, cell lysates of E14 ESCs overexpressing Flag-tagged CTBP2 were immunoprecipitated using anti-Flag M2 Affinity Gel. The FLAG-tagged proteins were then eluted from the resin by competition with 3XFlag-peptide (Sigma). To remove non-specific proteins that were co-eluted from the elution step, a negative control reaction was carried out, where anti-Flag M2 Affinity Gel was incubated with wild-type E14 ESC cell lysates. Eluents from the FLAG-tag and the negative control reactions were separated by SDS-PAGE. Each protein gel band was cut and subjected to in-gel tryptic digestion following the general protocol. Briefly, protein bands were excised, destained, and washed. Proteins were reduced with 20 mM dithiothreitol and alkylated with 55 mM iodoacetamide. After dehydration with acetonitrile, the proteins were digested with 12.5 ng μl^{-1} sequencing-grade modified porcine trypsin (Promega, Madison, WI) in 50 mM ammonium bicarbonate overnight at 37° C. Peptides were extracted from the gel slices with 50% (v/v) acetonitrile in 5% (v/v) formic acid.

Peptides isolated from the in-gel digestion were resuspended in 20 μl of Solvent A (0.1% formic acid in water), and 3 μL of sample was injected into nano-ultra-high performance liquid

chromatography (LC) system (Easy nLC, Thermo Fisher Scientific, San Jose, CA) interfaced with Q-Exactive mass spectrometer (Thermo Fisher Scientific). Peptides were separated on an in-house packed 75 μm (inner diameter) x 10cm C_{18} column and eluted with a linear gradient of 2~38% Solvent B (0.1% formic acid in acetonitrile) for 70 min at a flow rate of 300 nl min^{-1}. The standard mass spectrometric condition of the spray voltage was set to 1.9 kV and the temperature of the heated capillary was set to 250° C. The full scans were acquired in the mass analyzer at 400-1600 *m/z* with a resolution of 70,000 and the tandem mass spectrometry (MS/MS) scans were obtained with a resolution of 17,500 by normalized collision energy (NCE) of 27 eV for high energy collisional dissociation fragmentation. The advanced gain control target was $5e^9$, maximum injection time was 120 ms, and the isolation window was set to 3 m/z. The Q-Exactive was operated in data-dependent mode with one survey MS scan followed by ten MS/MS scans, and the duration time of dynamic exclusion was 30 seconds.

20. Database Search

Collected MS/MS data were converted into mzXML files through the Trans Proteomic pipeline (version 4.5) software and searched against the decoy IPI Mouse database (version 3.87, 119072 entries) for the estimation of the false discovery rate (FDR) with the SEQUEST® (Thermo Fisher Scientific, San Jose, CA; version v.27, rev. 11) program in the SORCERER™ (Sage-N Research, Milpitas CA, version 3.5) search platform. Precursor and fragment ion tolerance were set to 10 ppm and 0.5 Da, respectively. Trypsin was chosen as the enzyme with a maximum allowance of up to two missed cleavages. Carbamidomethyl of cysteine was considered as the fixed modification, while the variable modification was set for methionine oxidation. The Scaffold software package (version 3.4.9, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide and protein identifications were accepted if they could be established at greater than 95% and 99% probability, respectively, as specified by the Peptide and Protein Prophet algorithm (Nesvizhskii et al., 2003; Keller et al., 2002). The protein identification contained at least two identified peptides. To select Ctbp2-specific interacting proteins, any proteins

identified in the negative control experiment were discarded as a false positive.

Ⅲ. RESULTS (PART1)

O-GlcNAc Regulates Pluripotency and Reprogramming by Directly Acting on Core Components of the Pluripotency Network

This research was published in *Cell Stem Cell* on June 2012.

1. *O*-GlcNAc Regulates ESC Self-Renewal and Differentiation, and Somatic Cell Reprogramming

To investigate the role of *O*-GlcNAc in ESC maintenance, we perturbed *O*-GlcNAcylation using either short hairpin RNAs (shRNAs) against *Ogt* (shOgt) or chemical inhibition of Oga. We selected shOgts that achieve incomplete knockdown and are doxycline-inducible (Figure 1A) because complete elimination of Ogt expression is lethal to ESCs (Shafi et al., 2000). Dox-treated E14 tTS Tet-shOgt cells (Tet-shOgt) exhibited reduced numbers of colonies compared to control shRNA-expressing cells (Tet-shMock) in colony forming assays with the majority of colonies staining negative for alkaline phosphatase (AP), suggesting that Ogt affects both ESC proliferation (via cell cycle regulation or apoptosis) and self-renewal (Figures 1B, 1C). *O*-GlcNAc levels can also be manipulated using chemicals. E14 ESCs that were treated with either streptozocin (streptozotocin; STZ), which is known to inhibit OGA via the production of a transition state analog (Toleman et al., 2006), exhibited elevated global levels of *O*-GlcNAc. During EB formation, the normal increase in lineage-

specific marker expression and concordant decrease in pluripotency marker expression was hampered by STZ treatment (Figures 1D). These results suggest that *O*-GlcNAc also affects ESC differentiation.

We also investigated the role of *O*-GlcNAcylation in reprogramming fibroblasts to iPSCs. Transduction of Oct4-promoter-driven GFP MEFs (pOct4-GFP MEFs) with OSK transgenes induced colonies with ESC colony-like morphology. Some of them expressed GFP, a marker of endogenous Oct4 expression in our system (Figure 1E). All GFP-positive colonies were also positive for AP staining. In GFP-positive colonies, the expression of ESC-specific genes (Oct4, Sox2, and Nanog) was comparable to that of mouse ESCs, and the transgene was nearly silenced.

In MEFs, treatment of shOgts reduced global *O*-GlcNAc levels (Figure 1F). Addition of shOgts during OSK-driven iPSC generation reduced reprogramming efficiency (Figure 1G), whereas, overexpression of human OGT in MEFs elevated global *O*-GlcNAc levels (Figure 1H) and increased reprogramming efficiency (Figure 1I). Low glucose culture medium reduces cellular *O*-GlcNAc levels compared to high

glucose medium (Figure 1J), and MEFs reprogrammed under low glucose medium developed less iPSC colonies (Figure 1K). Collectively, these results demonstrate that *O*-GlcNAcylation is important for somatic cell reprogramming.

Figure 1.

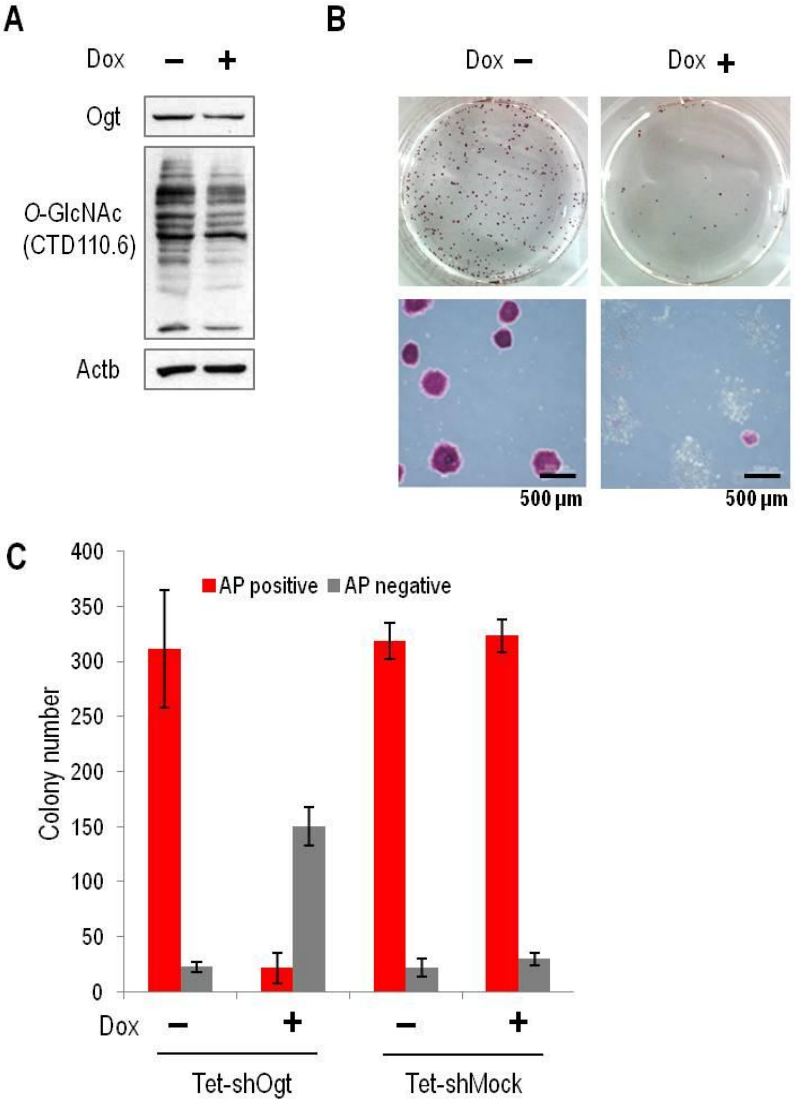

A
Dox – +
Ogt
O-GlcNAc
(CTD110.6)
Actb
B
Dox –
Dox +
500 μm
500 μm
C
AP positive
AP negative
Colony number
400
350
300
250
200
150
100
50
0
Dox – + – +
Tet-shOgt
Tet-shMock

Figure 1.

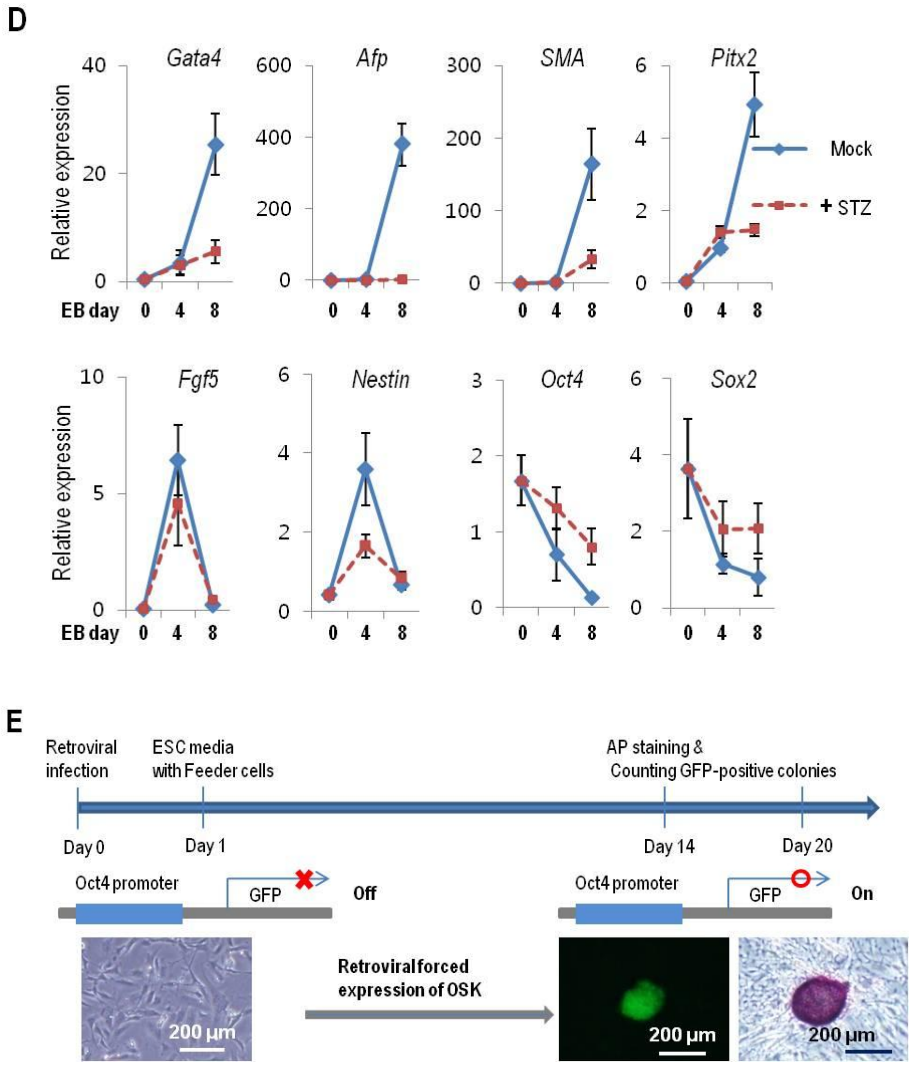

D
Gata4
Afp
SMA
Pitx2
Fgf5
Nestin
Oct4
Sox2
Relative expression
EB day
0 4 8
Mock
+ STZ
E
Retroviral infection
ESC media with Feeder cells
AP staining & Counting GFP-positive colonies
Day 0
Day 1
Day 14
Day 20
Oct4 promoter
GFP
Off
On
Retroviral forced expression of OSK
200 μm

Figure 1.

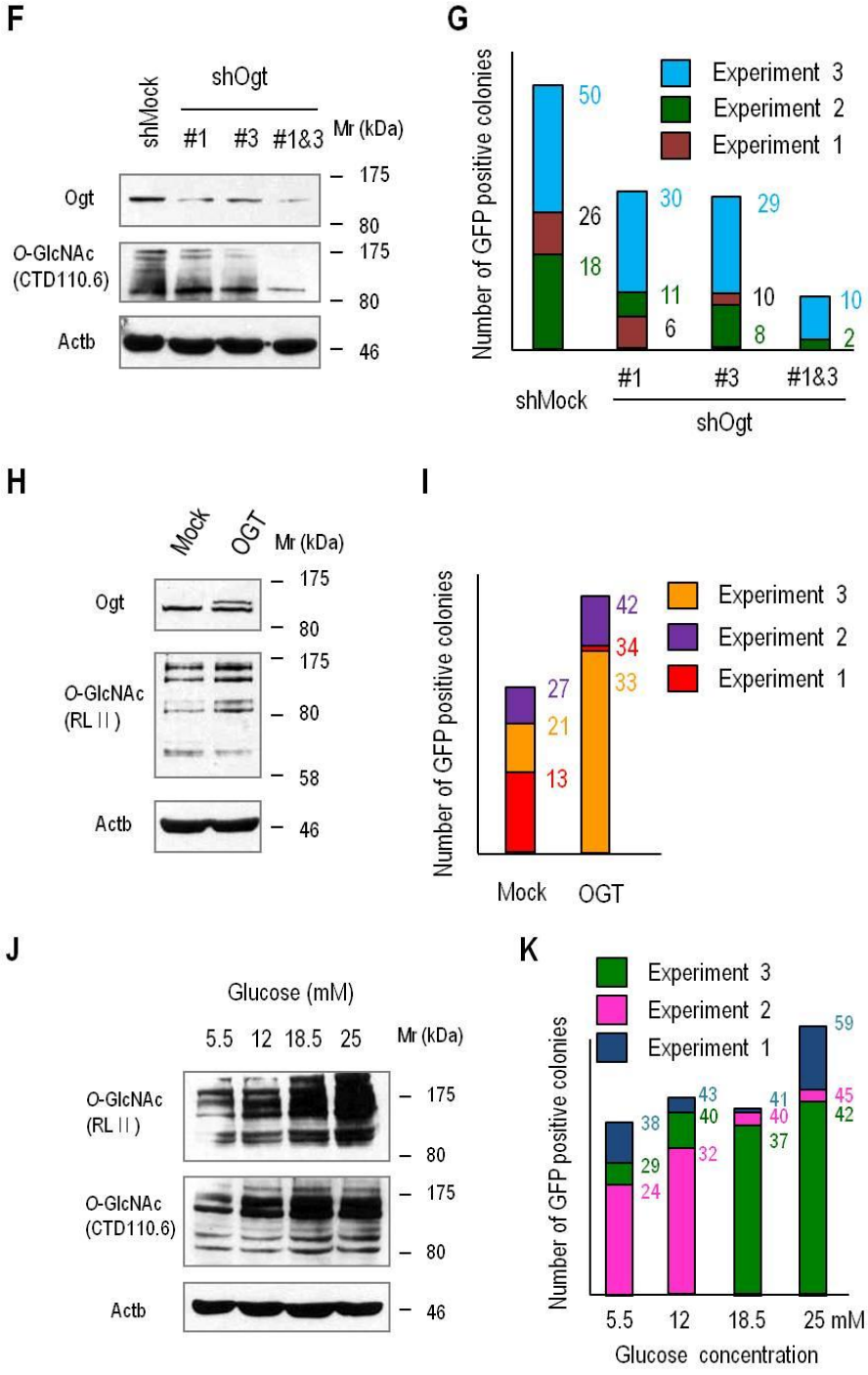

F
shMock
shOgt
#1
#3
#1&3
Mr (kDa)
Ogt
O-GlcNAc
(CTD110.6)
Actb
175
80
175
80
46
G
Number of GFP positive colonies
Experiment 3
Experiment 2
Experiment 1
50
26
18
30
11
6
29
10
8
10
2
shMock
#1
#3
#1&3
shOgt
H
Mock
OGT
Mr (kDa)
Ogt
O-GlcNAc
(RL II)
Actb
175
80
175
80
58
46
I
Number of GFP positive colonies
Experiment 3
Experiment 2
Experiment 1
27
21
13
42
34
33
Mock
OGT
J
Glucose (mM)
5.5
12
18.5
25
Mr (kDa)
O-GlcNAc
(RL II)
O-GlcNAc
(CTD110.6)
Actb
175
80
175
80
46
K
Number of GFP positive colonies
Experiment 3
Experiment 2
Experiment 1
38
29
24
43
40
32
41
40
37
59
45
42
5.5
12
18.5
25 mM
Glucose concentration

Figure 1. *O*-GlcNAc Regulates ESC Self-Renewal and Differentiation, and Somatic Cell Reprogramming

(A) Western Blot showing reduced overall *O*-GlcNAc levels in E14 tTS Tet-shOgt cells upon addition of Doxycycline (+/- Dox). *O*-GlcNAcylation of proteins was detected by Western blot with *O*-GlcNAc-specific antibody (CTD110.6).

(B, C) Mild knockdown of Ogt reduced ESC self-renewal capacity. E14 tTS Tet-shOgt cells were plated for self-renewal assay with /without Dox. The undifferentiated state was assessed by alkaline phosphatase (AP) staining. Colony numbers are expressed as mean ± standard deviation ($n \geq 3$).

(D) Enhancement of *O*-GlcNAcylation perturbs ESC differentiation. E14 cells were differentiated into embryoid bodies (EBs) with/ without streptozocin (STZ). The mRNA expression levels of lineage specific markers and pluripotency markers were analyzed by real-time qPCR. The relative mRNA expression levels are expressed as mean ± standard deviation ($n \geq 3$).

(E) Schematic of reprogramming of pOct4-GFP MEFs. Reprogrammed cells were detected by alkaline phosphatase (AP) staining and GFP expression.

(F) Reduction in overall *O*-GlcNAc level by shRNA mediated knockdown of Ogt in MEFs was determined by Western blot.

(G) Decrease in *O*-GlcNAc level reduces the efficiency of reprogramming. pOct4-GFP MEFs were reprogrammed by forced expression of retroviral Oct4, Sox2, and Klf4 (OSK) in the presence of shOgts. Reprogrammed cells were detected by AP staining and GFP expression. The number of GFP-positive colonies was presented (n=3).

(H) Elevation in overall *O*-GlcNAc level by overexpression of human OGT in MEFs was determined by Western blot.

(I) Increase in *O*-GlcNAc level elevates the efficiency of reprogramming. MEFs were reprogrammed in the presence of

OGT. The number of GFP-positive colonies was presented (n=3).

(J) Gradual increase in overall *O*-GlcNAc level by increasing glucose concentration in MEFs was determined by Western blot.

(K) Reprogramming efficiency is increased in high-glucose media. pOct4-GFP MEFs were reprogrammed in various glucose concentrations. The number of GFP-positive colonies was presented (n=3).

2. Core Components of the Pluripotency Network Are Modified by *O*-GlcNAc in Undifferentiated ESCs

To investigate the molecular mechanism linking *O*-GlcNAcylation and pluripotency, we postulated that *O*-GlcNAcylation directly controls core components of the pluripotency network. To test this hypothesis, we initially determined whether OSK and Nanog were modified by *O*-GlcNAcylation. *O*-GlcNAcylated proteins have high affinity for succinylated wheat germ agglutinin (sWGA) and are easily eluted by competition with GlcNAc. Thus, we purified sWGA-bound proteins from ESC lysates and probed for OSK and Nanog by Western blot. All endogenous OSK but not Nanog bound specifically to sWGA in ESCs, and competition with GlcNAc completely eliminated their binding to sWGA (Figure 2A).

Because sWGA binds terminal GlcNAc residues, we determined if the OSK binding to sWGA was mediated by *O*-GlcNAc. Removal of terminal *O*-linked glycosidic modifications by *O*-*N*-Acetyl-hexosaminidase (*O*-hex) (Gandy et al., 2006) during sWGA pull-down reduced Oct4 and Sox2 binding

to sWGA (Figure 2B). However, removal of *N*-linked GlcNAc by using excess PNGase F (Gandy et al., 2006) did not affect Oct4 and Sox2 binding to sWGA (Figure 2C). Thus sWGA-Oct4 and -Sox2 interactions are due primarily to *O*-GlcNAc on Oct4 and Sox2.

We reconfirmed the *O*-GlcNAcylation of Oct4 and Sox2 by Western blot with anti-*O*-GlcNAc (clone CTD110.6), which recognizes *O*-GlcNAcylated serine and threonine residues (Figure 2D). Endogenous Oct4 and Sox2 interacted with Ogt (Figure 2E) in undifferentiated ESCs.

During ESC differentiation into EBs, *O*-GlcNAcylation levels of Oct4 and Sox2 decreased rapidly. Oct4 *O*-GlcNAcylation levels decreased rapidly by day one and were nearly undetectable by day two, but total Oct4 protein levels remained unchanged over the same period (Figures 2F - 2H). As controls, Ogt and Gapdh *O*-GlcNAcylation levels were not altered.

Figure 2.

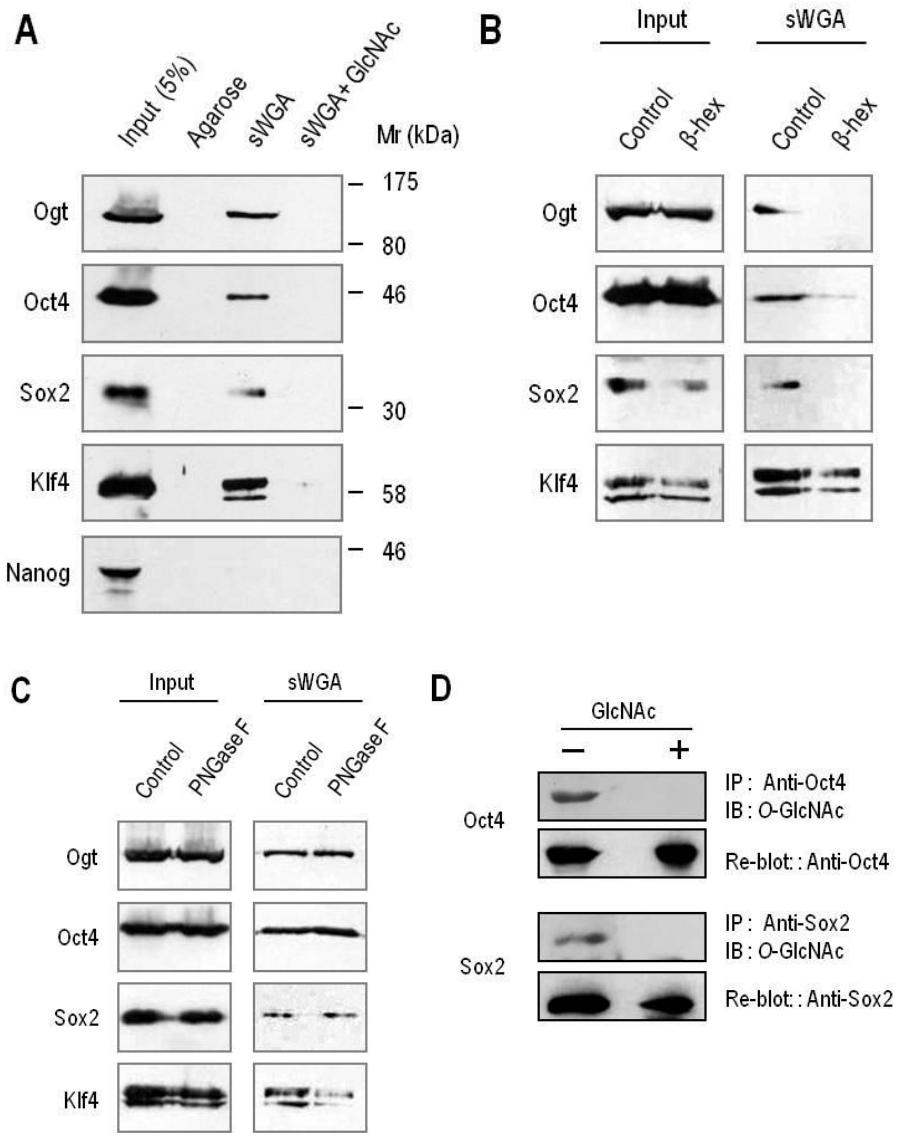

A
Input (5%)
Agarose
sWGA
sWGA+GlcNAc
Mr (kDa)
Ogt
Oct4
Sox2
Klf4
Nanog
175
80
46
30
58
46
B
Input
sWGA
Control
β-hex
Control
β-hex
Ogt
Oct4
Sox2
Klf4
C
Input
sWGA
Control
PNGase F
Control
PNGase F
Ogt
Oct4
Sox2
Klf4
D
GlcNAc
–
+
Oct4
IP : Anti-Oct4
IB : O-GlcNAc
Re-blot: : Anti-Oct4
Sox2
IP : Anti-Sox2
IB : O-GlcNAc
Re-blot: : Anti-Sox2

Figure 2.

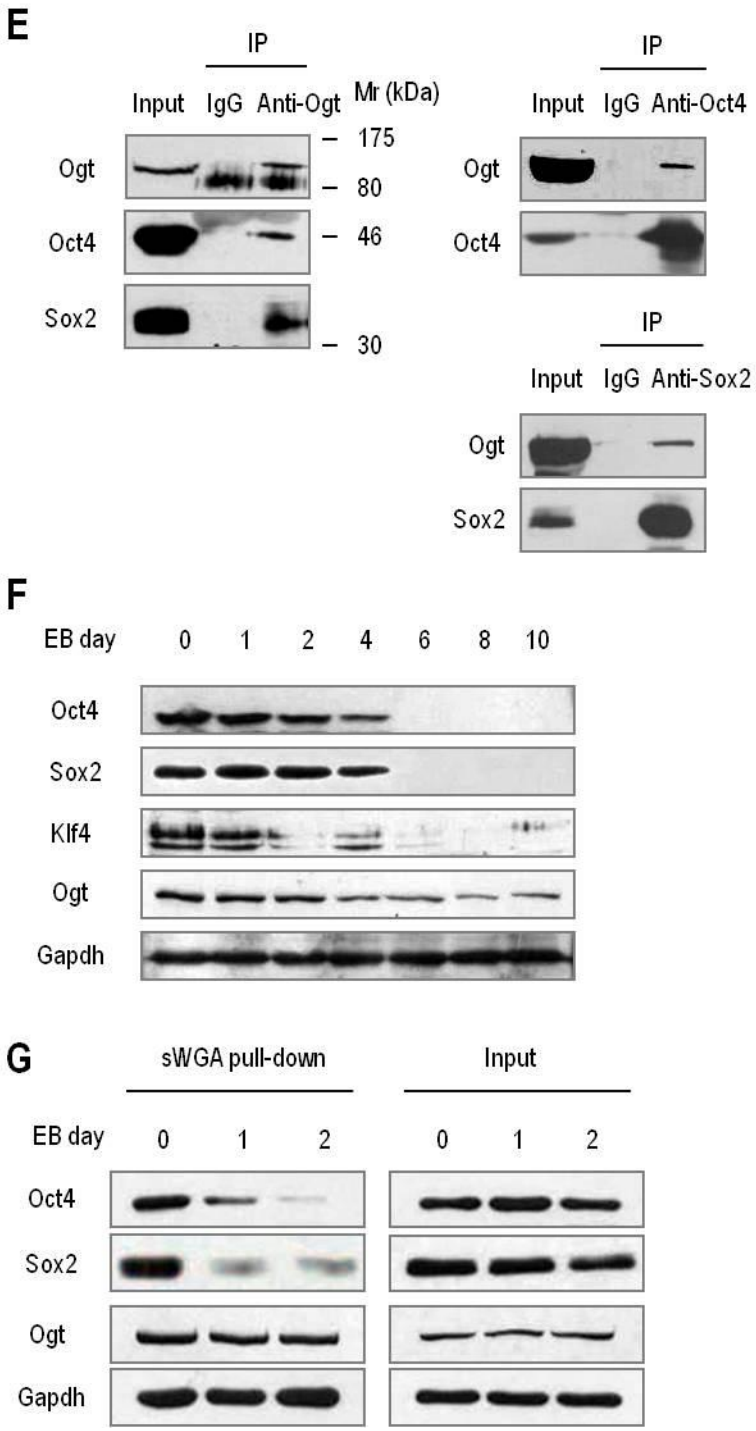

E
IP
Input IgG Anti-Ogt
Mr (kDa)
175
80
46
30
Ogt
Oct4
Sox2
IP
Input IgG Anti-Oct4
Ogt
Oct4
IP
Input IgG Anti-Sox2
Ogt
Sox2
F
EB day 0 1 2 4 6 8 10
Oct4
Sox2
Klf4
Ogt
Gapdh
G
sWGA pull-down
Input
EB day 0 1 2 0 1 2
Oct4
Sox2
Ogt
Gapdh

Figure 2.

H

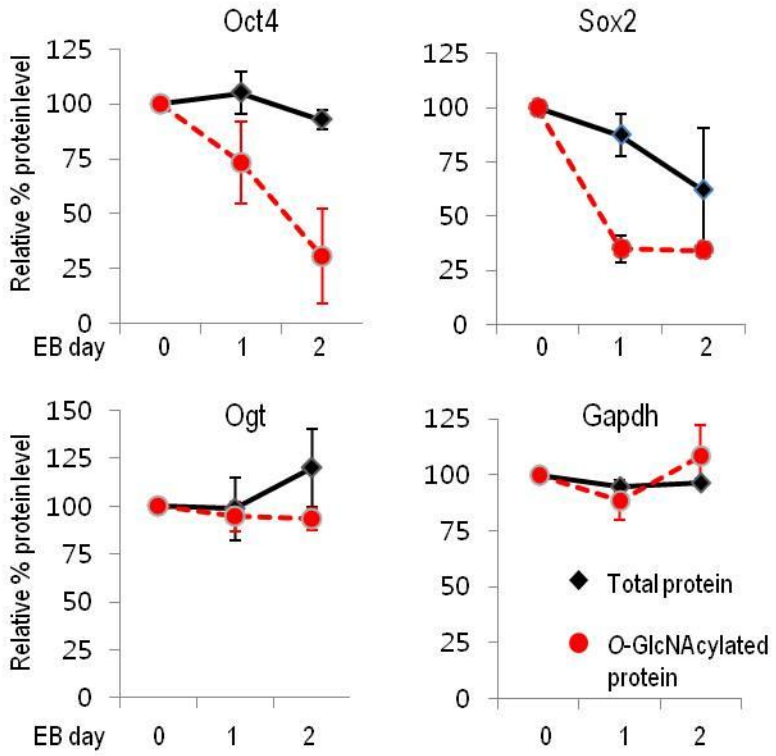

Oct4
Sox2
Ogt
Gapdh
Relative % protein level
EB day
0
1
2
125
100
75
50
25
0
150
Total protein
O-GlcNAcylated protein

Figure 2. Core Components of the Pluripotency Network Are Modified by *O*-GlcNAc in Undifferentiated ES Cells

(A) Whole cell lysates from E14 cells were incubated with succinylated wheat germ agglutinin (sWGA)-conjugated agarose beads with/without *N*-acetyl-*D*-glucosamine (GlcNAc, 0.5 M). Bound proteins were analyzed by Western blot.

(B, C) During sWGA pull down assay, sWGA-bound proteins were treated with N-Acetyl-hexosaminidase (hex) or PNGase F.

(D) Whole cell lysates from E14 cells were immunoprecipitated with antibodies for Oct4 or Sox2. *O*-GlcNAcylation of proteins was detected by Western blot with *O*-GlcNAc-specific antibody (CTD110.6). Specificity of the antibody was checked by competition with GlcNAc.

(E) Endogenous interactions between Ogt and Oct4, or Ogt and Sox2 in E14 cells were detected by co-immunoprecipitation assay.

(F) Changes in protein levels of OSK and Ogt during ESC differentiation into EBs were assessed by Western blot. Gapdh level was measured as a control.

(G, H) *O*-GlcNAcylation levels of Oct4 and Sox2 decrease during ESC differentiation into EBs. Whole cell lysates from E14 EBs were pulled down with sWGA-agarose. Bound proteins were analyzed by Western blot. Relative % of total Oct4 and Sox2, and sWGA-bound Oct4 and Sox2 levels are expressed as mean ± standard deviation ($n \geq 3$). Ogt and Gapdh *O*-GlcNAc levels were measured as controls.

3. Mapping *O*-GlcNAcylation Sites on Oct4 and Sox2

To identify the *O*-GlcNAcylation site(s) in Oct4, we examined the interaction between several truncated mutants of Oct4 and sWGA. Oct4-truncated mutants, lacking transactivation domain 1 (TAD1) and TAD2, still bound to sWGA, but Oct4 deletion mutants that lacked the POU homeodomain (POU-h) failed to bind to sWGA (Figures 3A). The *O*-GlcNAcylation of the POU domain of Oct4 was reconfirmed by Western blot with anti-*O*-GlcNAc (Figure 3B) in HEK293 cells. Therefore, we chose to mutate additional putative *O*-GlcNAcylation sites within the POU-h. Threonine 228 (T228) and serine 229 (S229) were chosen because S229 has been identified as a phosphorylation site in the ESC phosphoproteome (Swaney et al., 2009) and because *O*-GlcNAcylation often competes with phosphorylation at the same or adjacent residues. To determine whether these mutations also reduce Oct4 *O*-GlcNAcylation in undifferentiated ESCs, we stably expressed wild-type (WT) and point mutants of Flag-tagged Oct4 into E14 cells. Although the undifferentiated state of ESCs is sensitive to Oct4 levels (Niwa et al., 2000), slight

over-expression of Oct4 in our experiments maintained ESCs in an undifferentiated state. Through Western blot analyses, we determined that both T228A and S229A reduced Oct4 *O*-GlcNAc levels in ESCs (Figure 3C).

To directly detect *O*-GlcNAc on Oct4 in ESC, we purified *O*-GlcNAcylated Oct4 from ZHBTc4 F-Oct4 cells, which express Flag-tagged Oct4 instead of endogenous Oct4 (van den Berg et al., 2010), by serial pull-down with anti-Flag and sWGA. Purified protein was analyzed by nano-LC-ESI-MS/MS. *O*-GlcNAc on T228 was detected from three independent experiments in succession (Figure 3D). Of note, Oct4 T228 residues are highly conserved between species and among POU family members (Figures 3E and 3F).

The *O*-GlcNAcylation sequence in Sox2 has been previously determined in rat brain (Khidekel et al., 2004), although the exact sites have not been identified (Figure 3G). We mutated all nine serine and threonine sites in this sequence to alanine. By using the sWGA binding assay and Western blots with anti-*O*-GlcNAc (CTD110.6), we identified S248, T258, and S259 as *O*-GlcNAcylation sites (Figures 3H and 3I). To determine whether these three sites are responsible for *O*-GlcNAcylation

in undifferentiated ESCs, we stably expressed Flag-tagged Sox2 WT and mutant (S248A, T258A, and S259A) into E14 cells. Slight overexpression of Sox2 in our experiments maintained ESCs in an undifferentiated state (Figure 3J). The triple mutant exhibited reduced binding to sWGA in ESCs (Figure 3K). Recently, both S248 and T258 were also identified as *O*-GlcNAcylated residues in Sox2 by ETD MS/MS in ESCs (Myers et al., 2011), which is consistent with our results.

Figure 3.

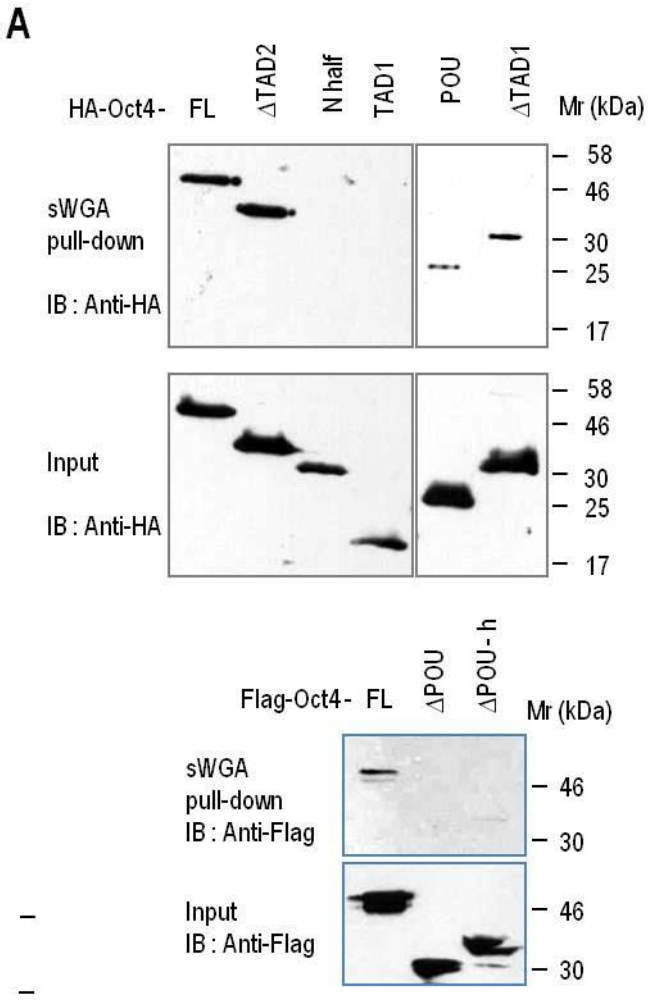

A
HA-Oct4-
FL
ΔTAD2
N half
TAD1
POU
ΔTAD1
Mr (kDa)
sWGA
pull-down
IB : Anti-HA
58
46
30
25
17
Input
IB : Anti-HA
58
46
30
25
17
Flag-Oct4-
FL
ΔPOU
ΔPOU-h
Mr (kDa)
sWGA
pull-down
IB : Anti-Flag
46
30
Input
IB : Anti-Flag
46
30

Figure 3.

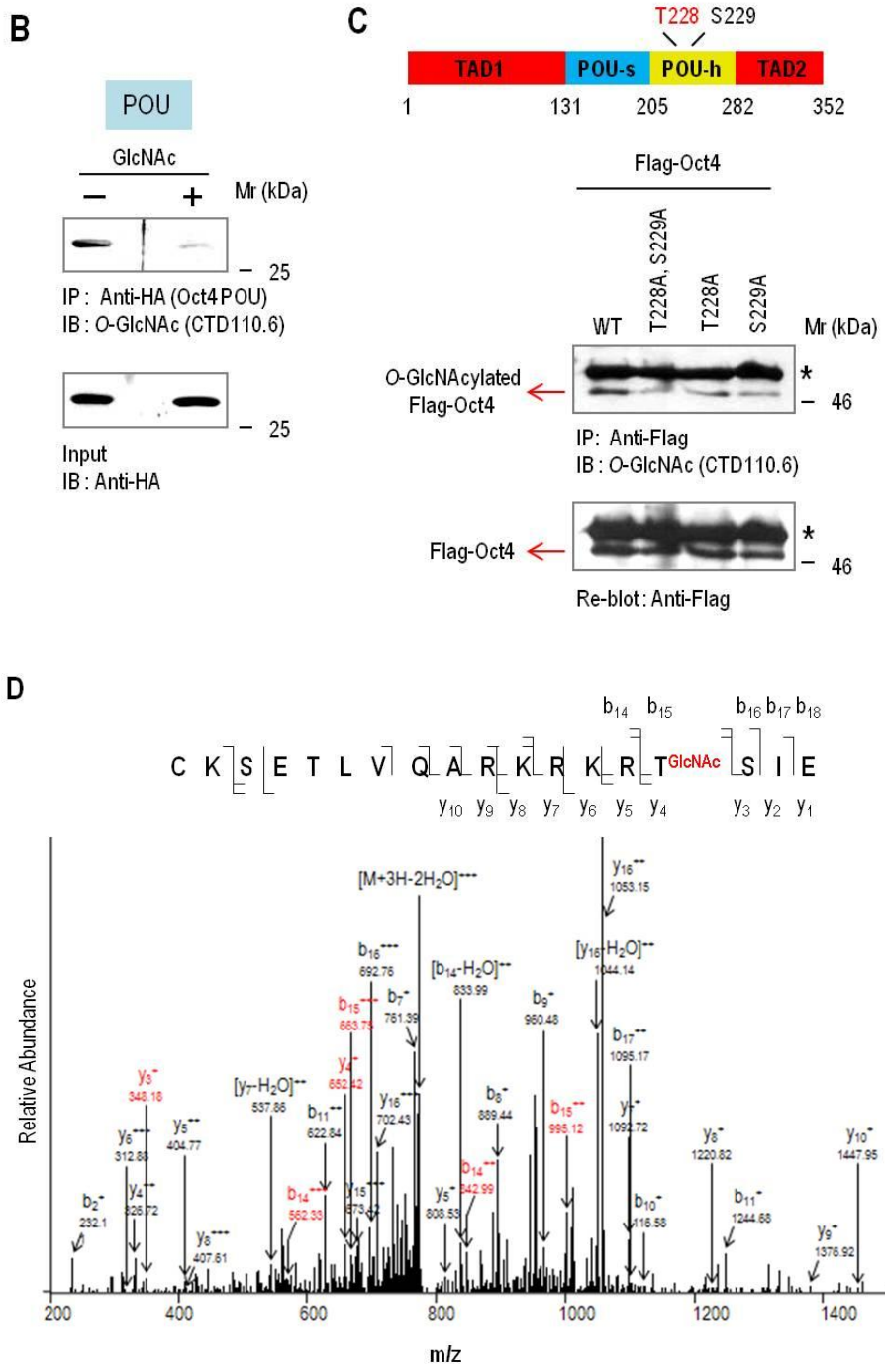

B
POU
GlcNAc
–
+
Mr (kDa)
25
IP : Anti-HA (Oct4 POU)
IB : O-GlcNAc (CTD110.6)
25
Input
IB : Anti-HA
C
T228 S229
TAD1
POU-s
POU-h
TAD2
1
131
205
282
352
Flag-Oct4
WT
T228A, S229A
T228A
S229A
Mr (kDa)
O-GlcNAcylated Flag-Oct4
46
IP: Anti-Flag
IB : O-GlcNAc (CTD110.6)
Flag-Oct4
46
Re-blot : Anti-Flag
D
C K S E T L V Q A R K R K R T GlcNAc S I E
Relative Abundance
m/z
200
400
600
800
1000
1200
1400

Figure 3.

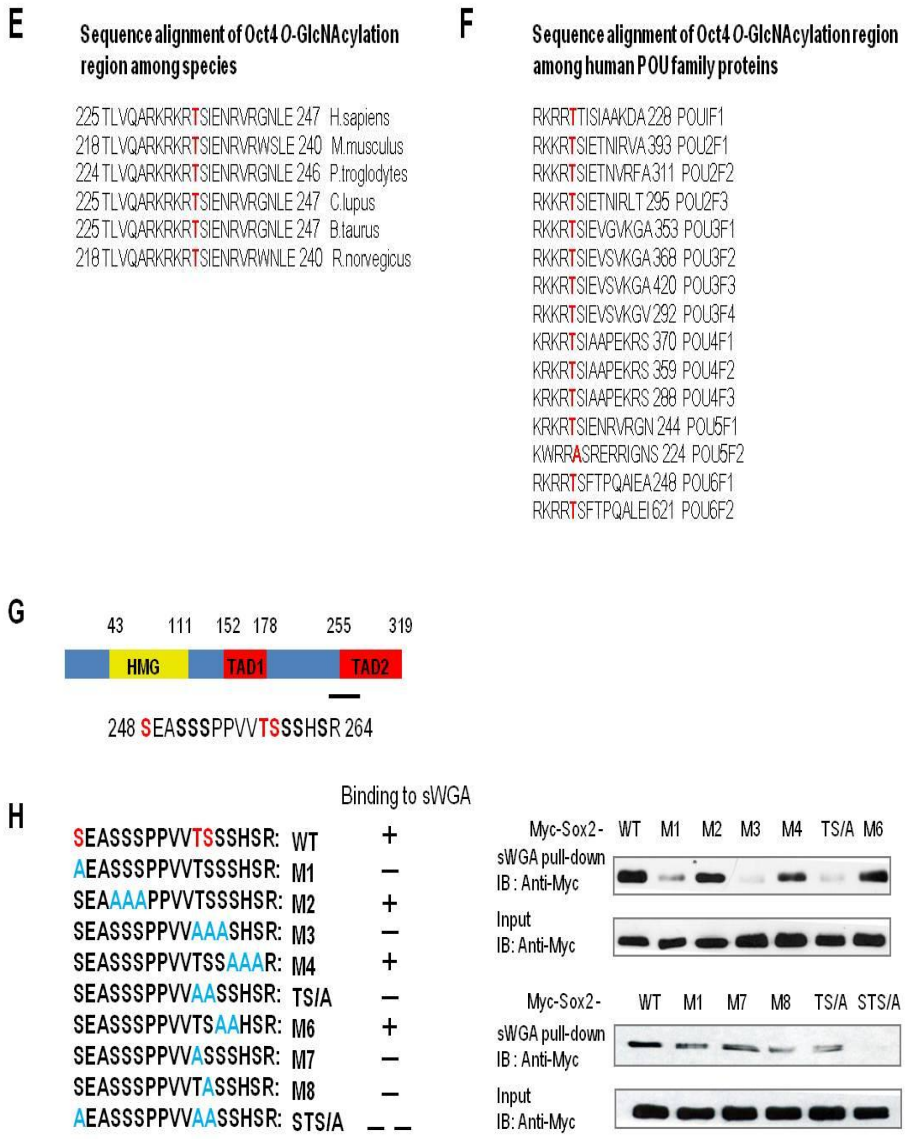

E
Sequence alignment of Oct4 O-GlcNAcylation region among species
225 TLVQARKRKRTSIENRVRGNLE 247 H.sapiens
218 TLVQARKRKRTSIENRVRWSLE 240 M.musculus
224 TLVQARKRKRTSIENRVRGNLE 246 P.troglodytes
225 TLVQARKRKRTSIENRVRGNLE 247 C.lupus
225 TLVQARKRKRTSIENRVRGNLE 247 B.taurus
218 TLVQARKRKRTSIENRVRWNLE 240 R.norvegicus
F
Sequence alignment of Oct4 O-GlcNAcylation region among human POU family proteins
RKRRTTISIAAKDA 228 POUIF1
RKKRTSIETNIRVA 393 POU2F1
RKKRTSIETNVRFA 311 POU2F2
RKKRTSIETNIRLT 295 POU2F3
RKKRTSIEVGVKGA 353 POU3F1
RKKRTSIEVSVKGA 368 POU3F2
RKKRTSIEVSVKGA 420 POU3F3
RKKRTSIEVSVKGV 292 POU3F4
KRKRTSIAAPEKRS 370 POU4F1
KRKRTSIAAPEKRS 359 POU4F2
KRKRTSIAAPEKRS 288 POU4F3
KRKRTSIENRVRGN 244 POU5F1
KWRRASRERRIGNS 224 POU5F2
RKRRTSFTPQAIEA 248 POU6F1
RKRRTSFTPQALEI 621 POU6F2
G
43 111 152 178 255 319
HMG
TAD1
TAD2
248 SEASSSPPVVTSSHSR 264
H
Binding to sWGA
SEASSSPPVVTSSHSR: WT +
AEASSSPPVVTSSHSR: M1 –
SEAAAAPPVVTSSSHSR: M2 +
SEASSSPPVVAAASHSR: M3 –
SEASSSPPVVTSSAAAR: M4 +
SEASSSPPVVAASSHSR: TS/A –
SEASSSPPVVTSAAHSR: M6 +
SEASSSPPVVASSSHSR: M7 –
SEASSSPPVVTASSHSR: M8 –
AEASSSPPVVAASSHSR: STS/A – –
Myc-Sox2- WT M1 M2 M3 M4 TS/A M6
sWGA pull-down
IB : Anti-Myc
Input
IB: Anti-Myc
Myc-Sox2- WT M1 M7 M8 TS/A STS/A
sWGA pull-down
IB : Anti-Myc
Input
IB: Anti-Myc

Figure 3.

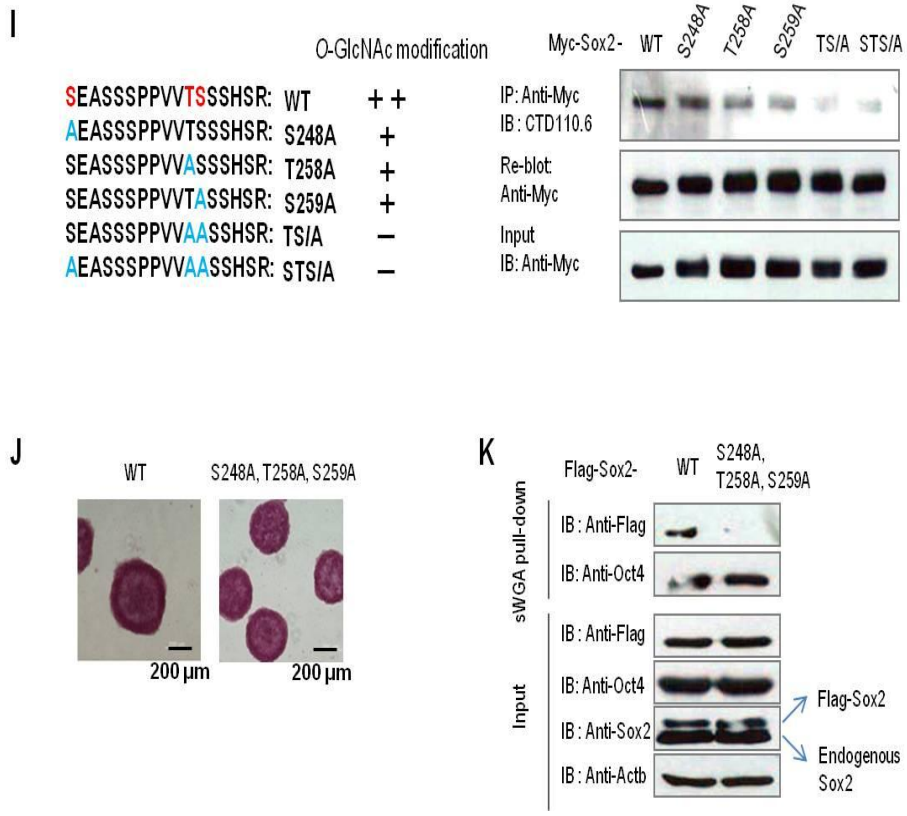

I
O-GlcNAc modification
SEASSSPPVVTSSSHSR: WT ++
AEASSSPPVVTSSSHSR: S248A +
SEASSSPPVVASSSHSR: T258A +
SEASSSPPVVTASSHSR: S259A +
SEASSSPPVVAASSHSR: TS/A –
AEASSSPPVVAASSHSR: STS/A –
Myc-Sox2- WT S248A T258A S259A TS/A STS/A
IP: Anti-Myc
IB: CTD110.6
Re-blot:
Anti-Myc
Input
IB: Anti-Myc
J
WT
S248A, T258A, S259A
200 μm
200 μm
K
Flag-Sox2- WT S248A, T258A, S259A
sWGA pull-down
IB: Anti-Flag
IB: Anti-Oct4
Input
IB: Anti-Flag
IB: Anti-Oct4
IB: Anti-Sox2
IB: Anti-Actb
Flag-Sox2
Endogenous Sox2

Figure 3. Mapping *O*-GlcNAcylation Sites on Oct4 and Sox2

(A) POU domain of Oct4 undergoes *O*-GlcNAcylation. Whole cell lysates from HEK 293 cells overexpressing various truncation mutants of Oct4 were pulled down with sWGA-agarose. Bound proteins were eluted by competition with 0.5 M GlcNAc and analyzed by Western blot.

(B) Whole cell lysates from HEK 293 cells overexpressing HA-tagged POU domain of Oct4 were immunoprecipitated with anti-HA. *O*-GlcNAcylation of the protein was detected by Western blot with *O*-GlcNAc antibody (CTD110.6). Specificity of the antibody was checked by competition with GlcNAc.

(C) Whole cell lysates from E14 cells expressing Flag-Oct4 WT and point mutants were immunoprecipitated with anti-Flag. *O*-GlcNAcylation of Oct4 proteins was detected by Western blot with *O*-GlcNAc antibody (CTD110.6). *: IgG heavy chain.

(D) T228 on Oct4 is *O*-GlcNAcylated in ES cells. *O*-GlcNAcylated Oct4 was purified from ZHBTc4 F-Oct4 cells by immunoprecipitation with anti-Flag, and then sWGA pull down. Purified Oct4 was analyzed by nano-LC-ESI-MS/MS. *O*-GlcNAcylation on Oct4 T228 was detected from three independent experiments.

(E, F) Sequence alignment of Oct4 *O*-GlcNAcylation region between species and among POU family members.

(G) The *O*-GlcNAcylation sequence in Sox2 (Khidekel et al., 2004).

(H) Serine 248, threonine 258, and serine 259 on Sox2 undergo *O*-GlcNAcylation. Whole cell lysates from HEK 293 cells overexpressing various Sox2 point mutants were pulled down with sWGA-agarose beads. Bound proteins were eluted by competition with 0.5 M GlcNAc and analyzed by Western blot.

(I) Various Sox2 point mutants were immunoprecipitated and the presence of *O*-GlcNAc modification was determined by Western blot with anti-*O*-GlcNAc antibody (CTD110.6).

(J) Flag-Sox2 WT and mutant (S248A, T258A, and S259A) were stably expressed in E14 cells. The undifferentiated state of stable cells was assessed by AP staining.

(K) Whole cell lysates from E14 cells stably expressing Flag-Sox2 WT and mutant were pulled down with sWGA-agarose. Bound proteins were eluted by competition with 0.5 M GlcNAc and analyzed by Western blot.

4. An *O*-GlcNAcylation-Defective Mutant of Oct4 and Sox2 Reduces Reprogramming Efficiency and ESC Self-Renewal

Having determined the *O*-GlcNAcylation sites in Oct4 and Sox2, we next determined the effects of *O*-GlcNAc on Oct4 and Sox2 function. During OSK-driven reprogramming of Oct4-promoter-driven GFP MEFs, we replaced wild-type Oct4 and Sox2 with their *O*-GlcNAcylation-defective point mutants. Substitution of Oct4 WT with T228A inhibited the development of GFP-positive colonies, with less than 5% GFP-positive colonies compared to Oct4 WT (Figures 4A). We also examined the effect of S229 phosphorylation on reprogramming, because S229 has been previously identified as a phosphorylation site (Swaney et al., 2009). Substitution of Oct4 WT with S229D (replacement of serine 229 to aspartate), which mimics phosphorylation, completely blocked the induction of GFP-positive colonies (Figures 4A). Oct4 S229A, which cannot undergo phosphorylation, had a milder effect inducing ~30 % GFP-positive colonies compared to Oct4 WT. These differences in reprogramming efficiency were not caused by disparities in viral titer (data not shown). These results

demonstrate that *O*-GlcNAcylation of Oct4 at T228 leads to an increase in Oct4 activity, whereas phosphorylation at S229 inhibits its function.

Substitution of Sox2 WT with the triple mutant (S248A, T258A, and S259A) did not alter reprogramming efficiency. However, expression of the Sox2 double mutant (T258A and S259A) induced about half as many GFP-positive colonies compared with Sox2 WT (Figures 4B). There were no disparities in viral titer of Sox2 WT and mutants (data not shown). These seemingly confusing results could be due to crosstalk between phosphorylation and *O*-GlcNAcylation. Because Sox2 S248 can be mutually modified by *O*-GlcNAcylation (Myers et al., 2011) and phosphorylation (Rigbolt et al., 2011; Swaney et al., 2009), *O*-GlcNAc and phosphate may compete with each other at this site. S248A blocks both phosphorylation and *O*-GlcNAcylation, so if one of these modifications severely reduces Sox2 activity, S248A may increase Sox2 function. Because the Sox2 double mutant (T258A and S259A) represses Sox2 activity, all three mutations may cancel out their effects on Sox2 function.

Since we determined that Oct4 *O*-GlcNAcylation at T228

reduces Oct4 activity during reprogramming, we next investigated whether it affects ESC maintenance. We replaced endogenous Oct4 with *O*-GlcNAc-defective Oct4 using ZHBTc4 cells (Niwa et al., 2000), which completely lose Oct4 protein within 24 h after Dox addition (Lee et al., 2010). ZHBTc4 cells were infected with Flag-Oct4 WT or mutants of equivalent viral titer and Dox was added 24 h after infection. 8 days after Dox treatment, approximately one hundred colonies with undifferentiated ESC morphology appeared in Flag-Oct4 WT infected plates whereas no colonies appeared in Flag-Mock infected plates and only a few colonies appeared in Flag-Oct4 T228A infected plates (Figures 4C and 4D), suggesting that Oct4 T228A is reduced in its capacity to maintain ESCs compared to Oct4 WT.

Although rare, Flag-Oct4 T228A was capable of forming undifferentiated colonies in infected ESCs. We attempted to examine the undifferentiated colonies from Oct4 T228A cells but failed because retroviral promoters are commonly silenced during long term ESC culture (Pannell and Ellis, 2001). Thus, we stably expressed Flag-Oct4 WT and T228A under the control of the constitutive CAG expression unit (Niwa et al.,

2002; Niwa et al., 1991) in Dox treated ZHBTc4 cells. When Flag-Oct4 WT cells were cultured normally with undifferentiated state, however, Flag-Oct4 T228 cells were prone to differentiate during culture. Colony-forming assays showed that ~50% of Flag-Oct4 T228 colonies were differentiated (Figures 4E), thus showing that *O*-GlcNAc defective Oct4 cannot promote self-renew efficiently.

Figure 4.

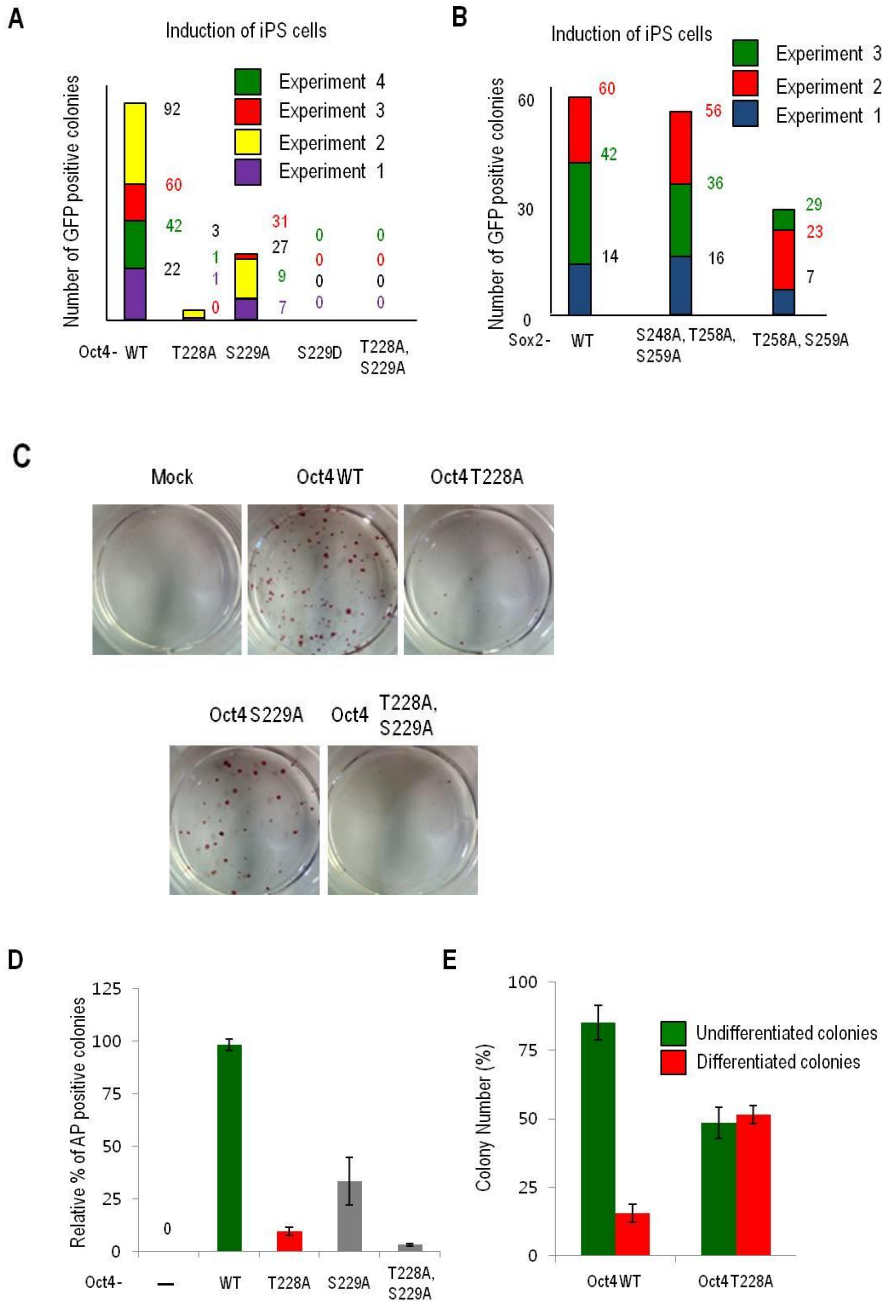

A
Induction of iPS cells
Number of GFP positive colonies
Experiment 4
Experiment 3
Experiment 2
Experiment 1
92
60
42
22
3
1
1
0
31
27
9
7
0
0
0
0
0
0
Oct4- WT
T228A
S229A
S229D
T228A, S229A
B
Induction of iPS cells
Number of GFP positive colonies
Experiment 3
Experiment 2
Experiment 1
60
30
0
60
42
14
56
36
16
29
23
7
Sox2- WT
S248A, T258A, S259A
T258A, S259A
C
Mock
Oct4 WT
Oct4 T228A
Oct4 S229A
Oct4 T228A, S229A
D
Relative % of AP positive colonies
125
100
75
50
25
0
0
Oct4- —
WT
T228A
S229A
T228A, S229A
E
Colony Number (%)
100
75
50
25
0
Undifferentiated colonies
Differentiated colonies
Oct4 WT
Oct4 T228A

Figure 4. An *O*-GlcNAcylation-Defective Mutant of Oct4 and Sox2 Reduces Reprogramming Efficiency and ESC Self-Renewal

(A) During OSK-driven reprogramming of pOct4-GFP MEFs, wild-type Oct4 was replaced with various Oct4 point mutants. Reprogrammed cells were identified by morphology, AP staining, and GFP expression. Results from 4 independent experiments are presented.

(B) During OSK-driven reprogramming of pOct4-GFP MEFs, wild-type Sox2 was replaced with Sox2 point mutants. Reprogrammed cells were identified by morphology, AP staining, and GFP expression. Results from 3 independent experiments were presented.

(C, D) Equal numbers of ZHBTc4 cells were infected with same titer of retroviral Oct4 point mutants. After selecting infected cells in the presence of doxycycline (Dox) to eliminate endogenous Oct4, ES cell colonies were stained with AP. AP positive colony numbers are expressed as relative % mean ± standard deviation ($n \geq 3$).

(E) CAG promoter driven Flag-Oct4 WT and T228A were introduced into ZHBTc4 cells in the presence of Dox. After selecting Flag-Oct4 stably expressing cells, cells were re-plated for self-renewal assay. The undifferentiated state was assessed by both morphology and AP staining. Colony numbers are expressed as relative % mean ± standard deviation ($n \geq 3$).

5. *O*-GlcNAcylation Regulates Oct4 Transcriptional Activity

Next, we investigated the mechanism by which *O*-GlcNAcylation of Oct4 influences reprogramming and ESC self-renewal. Oct4 T228A showed no apparent difference in protein stability compared to WT (Figure 5A). In addition, various Oct4 and Sox2 mutants used in the reprogramming assay exhibited no difference in cellular localization (Figures 5B and 5C).

To determine the effect of *O*-GlcNAcylation on Oct4 transcriptional activity, we used 10×Oct4 RE- luc, which contains 10 copies of the Oct4 response element (Oct4 RE) upstream of the TATA box (Lee et al., 2005). To rule out the activity of endogenous Oct4, NIH3T3 cells, which do not express Oct4 (Figures 5D), were stably incorporated with 10× Oct4 RE- luc. These NIH3T3 stable cells were infected with retroviral Oct4 WT and mutants and luciferase activity was measured 4 days after infection. In this reporter assay, Oct4 T228A (TA), and S229D (SD) barely activated luciferase expression, and S229A (SA) had ~60 % of WT activity (Figure 5D). Reporter activity of Oct4 mutants closely resembled their abilities to promote somatic cell reprogramming (Figures 4A

and 5D).

Oct4 and Sox2 cooperate in regulating many of their targets. Therefore, we investigated the effect of *O*-GlcNAcylation on Oct4 activity at promoters that are synergistically regulated by Oct4 and Sox2. The reporter plasmid 6×O/S RE-luc, which contains 6 copies of the Oct4 and Sox2 response elements (O/S RE) upstream of the minimal *Fgf4* promoter (Ambrosetti et al., 2000) was used for this purpose. NIH3T3 cells, which have no endogenous Oct4 and Sox2 expression, were stably incorporated with 6×O/S RE-luc. Expression of Oct4 or Sox2 alone slightly increased reporter activity, and simultaneous expression of both synergistically increased reporter activity. Oct4 TA and SD did not synergize with Sox2 to activate the reporter gene (data not shown).

Next, we investigated the dependence of Oct4 activity on *O*-GlcNAcylation by perturbing *O*-GlcNAcylation in our reporter assay. Over-expression of human OGT in NIH3T3 cells elevated overall *O*-GlcNAc levels (Figure 5E). OGT overexpression increased reporter activity in Oct4 WT-expressing cells but not in Oct4 TA-expressing cells (Figure 5F). Inhibition of *O*-GlcNAcylation with shOgt reduced the

Oct4-dependent reporter activity (Figure 5G). Simultaneous infection of retroviral human OGT with shOgts restored *O*-GlcNAc level and reporter activity demonstrating that the changes in reporter activity were not an off-target effect of shRNAs (Figures 5G and 5H). These results suggest that *O*-GlcNAcylation of Oct4 enhances its transcriptional activity, and the T228 is a major *O*-GlcNAcylation site that controls Oct4 transcriptional activity.

Next, we wanted to see whether *O*-GlcNAcylation regulates endogenous Oct4 transcriptional activity in ESCs. We stably incorporated 10×Oct4 RE-luc reporter gene into the genome of E14 cells or ZHBTc4 cells. During differentiation to EBs, reporter activities of E14 10×Oct4 RE-luc stable cells sharply decreased to below 10 % after two days of differentiation (Figure 5I). At that time point, total Oct4 protein levels remained high, but were almost depleted in *O*-GlcNAc (Figure 5I). STZ partially recovered both the *O*-GlcNAc levels and reporter activity (Figures 5I). Luciferase activity in ZHBTc4 10×Oct4 RE-luc stable cells was similarly decreased during differentiation and restored by STZ (Figure 5J). However, STZ did not increase reporter activity in the presence of Dox

(Figure 5K), suggesting that STZ induced reporter activity via *O*-GlcNAc-Oct4. E14 10×Oct4 RE-luc stable cells cultured in low glucose (5.5 mM) media had reduced *O*-GlcNAc-Oct4 level and reporter activity compared to cells cultured in high glucose (25 mM) media, despite comparable levels of total Oct4 protein (Figure 5L). These results demonstrate that the Oct4 activity is dependent on the *O*-GlcNAcylation level of Oct4.

Figure 5.

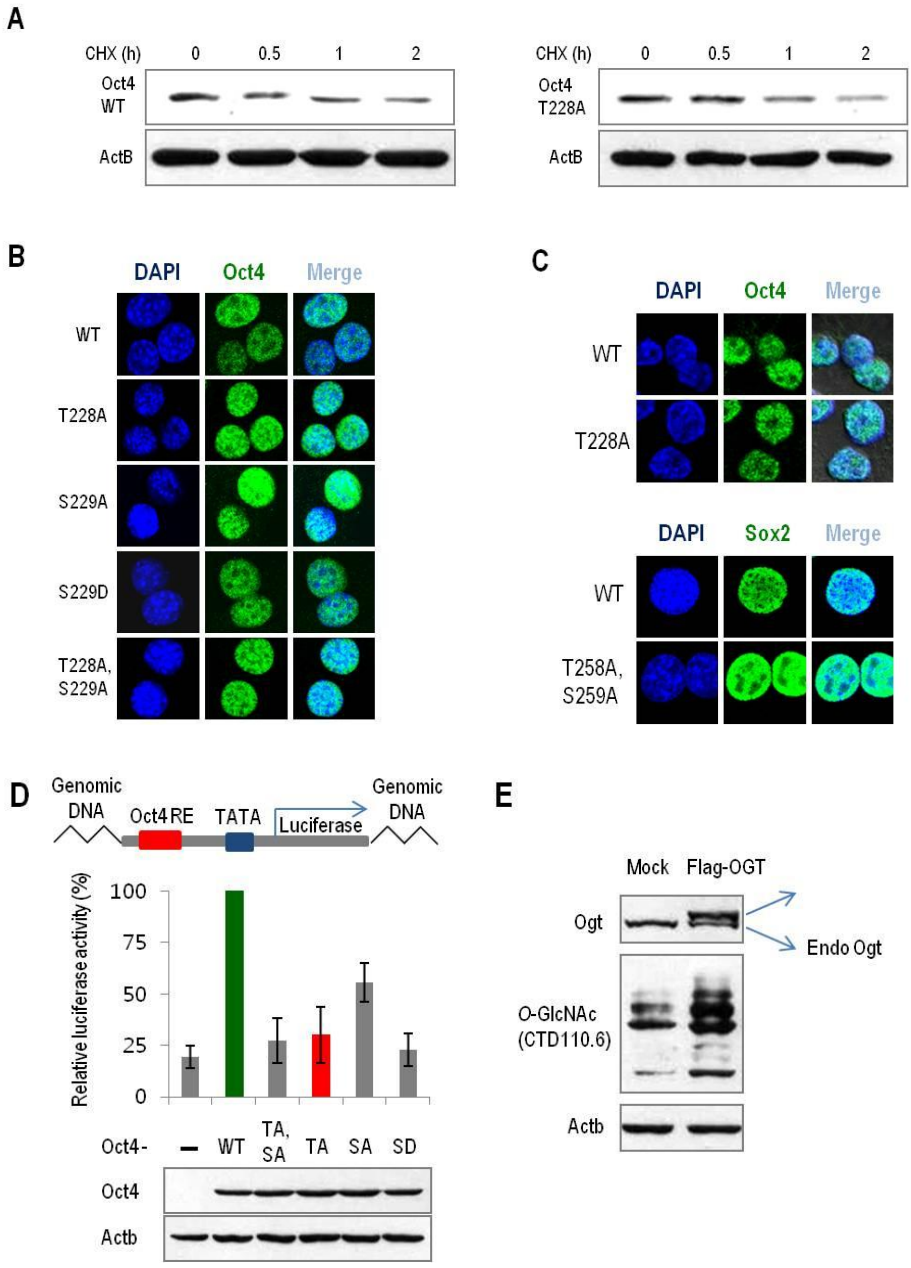

A
CHX (h) 0 0.5 1 2
Oct4 WT
ActB
CHX (h) 0 0.5 1 2
Oct4 T228A
ActB
B
DAPI Oct4 Merge
WT
T228A
S229A
S229D
T228A, S229A
C
DAPI Oct4 Merge
WT
T228A
DAPI Sox2 Merge
WT
T258A, S259A
D
Genomic DNA
Oct4 RE
TATA
Luciferase
Genomic DNA
Relative luciferase activity (%)
100
75
50
25
0
Oct4- — WT TA, SA TA SA SD
Oct4
Actb
E
Mock Flag-OGT
Ogt
Endo Ogt
O-GlcNAc (CTD110.6)
Actb

Figure 5.

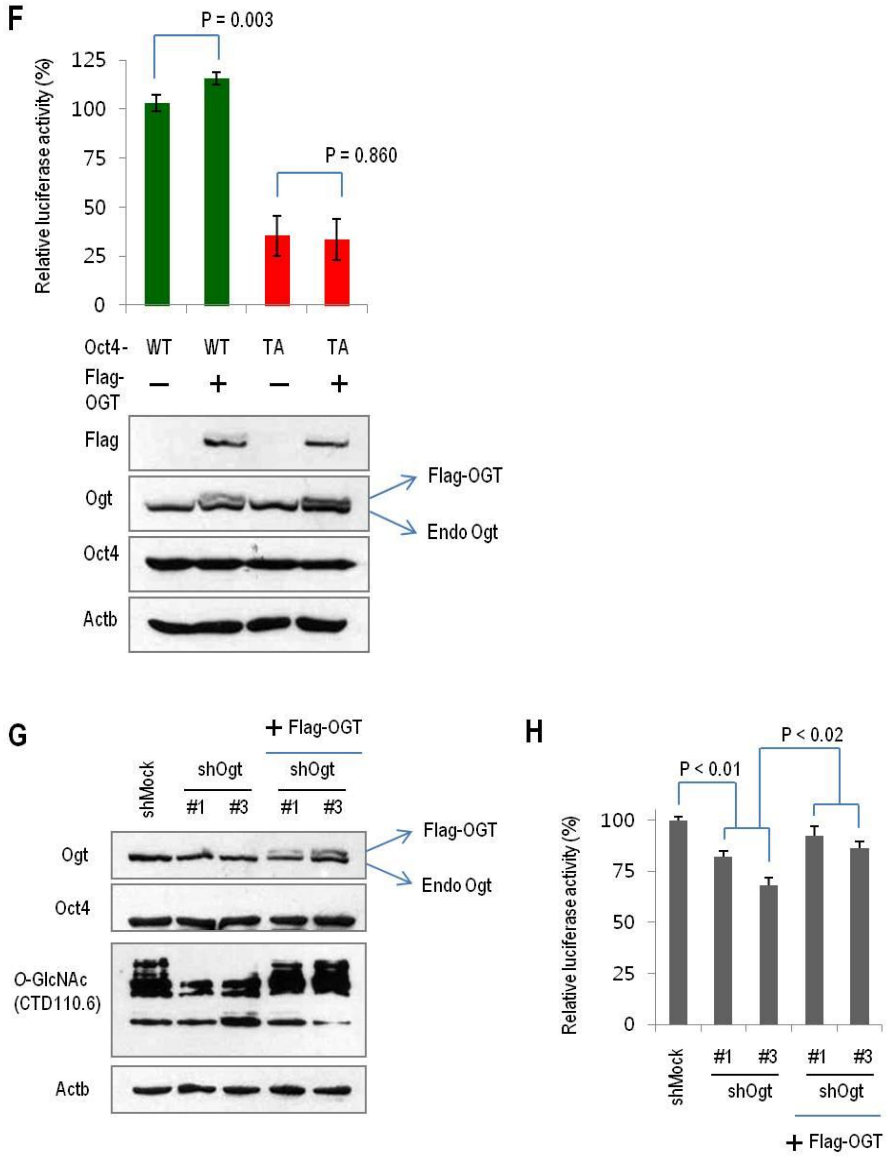

F
P = 0.003
P = 0.860
Relative luciferase activity (%)
125
100
75
50
25
0
Oct4-
WT
WT
TA
TA
Flag-OGT
–
+
–
+
Flag
Ogt
Oct4
Actb
Flag-OGT
Endo Ogt
G
+ Flag-OGT
shMock
shOgt
shOgt
#1
#3
#1
#3
Ogt
Oct4
O-GlcNAc
(CTD110.6)
Actb
Flag-OGT
Endo Ogt
H
P < 0.02
P < 0.01
Relative luciferase activity (%)
100
75
50
25
0
shMock
#1
#3
#1
#3
shOgt
shOgt
+ Flag-OGT

Figure 5.

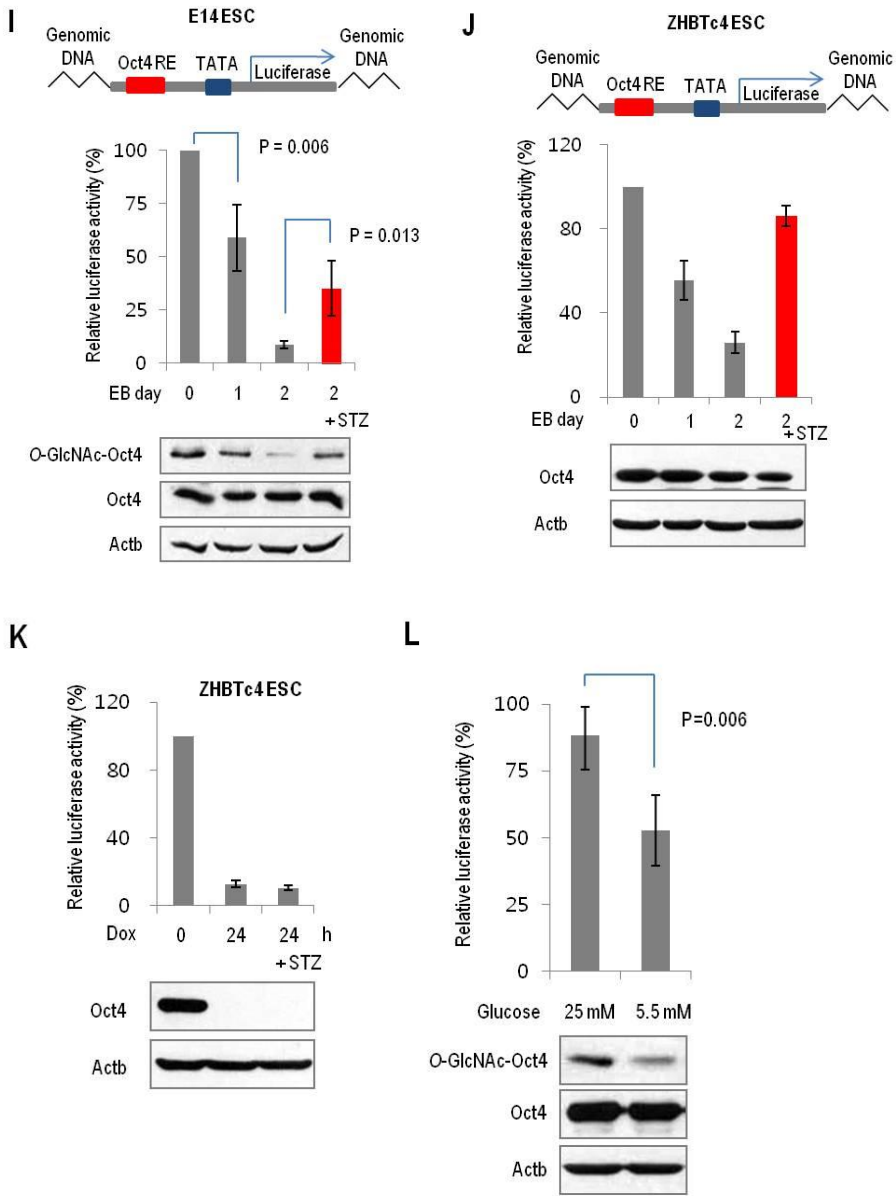

I
E14ESC
Genomic DNA
Oct4 RE
TATA
Luciferase
Genomic DNA
Relative luciferase activity (%)
P = 0.006
P = 0.013
EB day
+STZ
O-GlcNAc-Oct4
Oct4
Actb
J
ZHBTc4ESC
K
ZHBTc4ESC
Dox
h
L
P=0.006
Glucose
25 mM
5.5 mM

Figure 5. *O*-GlcNAcylation Regulates Oct4 Transcriptional Activity

(A) No apparent difference in protein stability between Oct4 WT and T228A. Endogenous Oct4 was replaced with Flag-Oct4 WT and T228A in ZHBTc4 cells. These cells were treated with cycloheximide (CHX; 20ug/ml) for indicated time periods and the Oct4 protein levels were determined by Western blot.

(B-C) No difference in the cellular localization of Oct4 and Sox2 mutants. (B) NIH3T3 cells were overexpressed with Oct4 WT and mutants. The localization of Oct4 was determined by immunofluorescent staining with anti-Oct4 antibody. (C) Upper panel shows that endogenous Oct4 was replaced with Flag-Oct4 WT and T228A in ZHBTc4 cells. The localization of Oct4 was determined by immunofluorescent staining with anti-Flag antibody. Lower panel shows that NIH3T3 cells were overexpressed with Sox2 WT and mutant. The localization of Sox2 was determined by immunofluorescent staining with anti-Sox2 antibody.

(D) 10 copies of Oct4 responsive element-driven luciferase (10×Oct4 RE-luc) reporter gene was stably incorporated into the genome of NIH3T3 cells. These stable cells, which have no endogenous Oct4, were infected with indicated retroviral Oct4 wild type (WT) and mutants. Luciferase activity was measured 4 days after infection with a luminometer. The values are expressed as relative % mean luminescent units ± standard deviation (n≥3). TA, SA : T228A and S229A, TA : T228A, SA : S229A, SD : S229D.

(E, F) Increasing *O*-GlcNAc by OGT overexpression elevates Oct4 WT, but not Oct4-TA, driven reporter activity. NIH3T3 cells stably incorporated with 10×Oct4 RE-luc reporter gene were infected with retroviral Oct4 WT or TA. Then, overall *O*-GlcNAc level was increased by additional infection of retroviral Flag-tagged OGT. Luciferase activity was measured 4 days after infecting OGT. P : P-value.

(G, H) Decreasing *O*-GlcNAc by shOgts reduces Oct4 driven reporter activity. NIH3T3 10×Oct4 RE-luc stable cells were infected with retroviral Oct4 WT. Then, overall *O*-GlcNAc level

was decreased by additional infection of retroviral shRNAs against Ogt. Luciferase activity was measured 4 days after infecting shOgts. To check specific effect of shOgts, Flag-tagged human OGT was simultaneously infected with shOgts.

(I) E14 cells stably incorporated with 10×Oct4 RE-luc reporter gene were differentiated into EBs with/ without STZ (3 mM). Luciferase activity was measured at the indicated time points. At the same time, total Oct4 protein levels and *O*-GlcNAcylated Oct4 levels were determined by Western blot and sWGA pull down assay.

(J) ZHBTc4 stable cells were differentiated into EBs with/ without STZ (3 mM). Luciferase activity was measured at the indicated time points. At the same time, Oct4 levels were determined by Western blot.

(K) STZ does not elevate reporter activity in the absence of Oct4. ZHBTc4 stable cells were treated with Dox with/ without STZ (3 mM) and luciferase activity was measured.

(L) E14 cells stably incorporated with 10×Oct4 RE-luc reporter gene were cultured in different glucose concentration. Luciferase activity was measured ~ 1 week after culture. At the same time, total Oct4 protein levels and *O*-GlcNAcylated Oct4 levels were determined by Western blot and sWGA pull down assay.

6. Identification of Genes Regulated by Oct4 *O*-GlcNAcylation

We next identified Oct4 target genes upregulated by *O*-GlcNAcylation of Oct4. Because *O*-GlcNAc levels of Oct4 are significantly reduced after two days of differentiation to EBs without an accompanying change in Oct4 protein levels (Figure 5I), we postulated that expression of *O*-GlcNAcylated Oct4 target genes should be changed by this time point. In addition, because STZ sustained *O*-GlcNAc levels on Oct4 during EB formation (Figure 5I), we postulated that the reduced expression of target genes should be recovered by STZ. However, STZ-upregulated genes may also include genes regulated by *O*-GlcNAcylation of proteins other than Oct4. To rule out those genes, we used ZHBTc4 cells, which lose Oct4 upon Dox treatment. Genes whose expression levels were recovered from STZ addition in Dox-treated ZHBTc4 cells may not be regulated by Oct4 *O*-GlcNAcylation, and so these genes were eliminated from our subsequent analyses.

In mRNA microarray experiments, 203 probes (174 genes) among 12,222 validated probes satisfied our criteria for potential *O*-GlcNAcylated Oct4 target genes (Figure 6A, left).

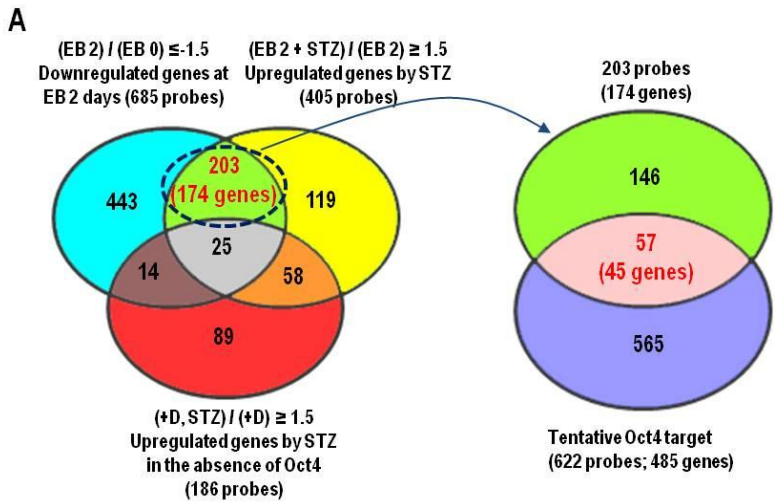
These 174 genes could represent both direct and indirect targets of *O*-GlcNAcylated Oct4. To identify the direct targets of *O*-GlcNAcylated Oct4, we chose a subset of known Oct4 target genes among the list of validated probes. Among 485 tentative Oct4 target genes (622 probes) from published data (Sharov et al., 2008), 45 genes (57 probes) were included in our set of potential *O*-GlcNAcylated Oct4 target genes (Figure 6A, right). We classified these 45 genes as potential genes directly upregulated by *O*-GlcNAcylated Oct4 (Figure 6B). Using a similar logic, we found 3 potential genes directly downregulated by *O*-GlcNAcylated Oct4.

We confirmed a subset of our microarray data using real-time qPCR. *O*-GlcNAcylated Oct4 targets selected from our dataset (*Nanog, Klf4, Klf2, Esrrb*, and *Zfp42*) were indeed reduced in expression during ESC differentiation into EBs, and their expression was also recovered upon addition of STZ. In contrast, expression of *Utf1,* which was not identified as a potential *O*-GlcNAcylated Oct4 target gene, was not reduced in day 2 EBs nor significantly affected by STZ treatment (Figure 6C). In ZHBTc4 cells, addition of Dox nearly eliminated Oct4 protein, and STZ did not affect expression of the putative

targets (Figure 6D).

If the selected genes are direct targets of *O*-GlcNAcylated Oct4, their promoters should be occupied by *O*-GlcNAcylated Oct4. Using serial chromatin immunoprecipitation (ChIP) and sequencing, we tested whether *O*-GlcNAcylated Oct4 occupied these potential target gene promoters. ZHBTc4 F-Oct4 cells were ChIPed with anti-Flag antibody. After eluting the bound chromatin by competition with Flag peptide, eluates were re-ChIPed by sWGA agarose. Bound chromatin was eluted by competition with GlcNAc and sequenced using genome Analyzer Ⅱ. 485 tentative Oct4 target genes (Sharov et al., 2008) were analyzed (Figure 6E). Among 485 genes, 271 genes showed ChIP-Sequencing (ChIP-Seq) signals in their gene region (defined as -10kb from transcription stat site to +10kb from transcription ending site). Among 45 potential *O*-GlcNAcylated Oct4 target genes, 29 genes, including *Klf2, Klf5, Nr5a2, Tbx3, Tcl1* and *Zfp42*, showed ChIP-Seq signals in their gene region (Figure 6F). We present these 29 genes as refined tentative *O*-GlcNAcylated Oct4 target genes. Key regulators of ES cells, such as *Klf2, Klf5, Nr5a2, Tbx3* and *Tcl1* are included in the list (Figure 6G).

Figure 6.


A
(EB 2) / (EB 0) ≤-1.5
Downregulated genes at
EB 2 days (685 probes)
(EB 2 + STZ) / (EB 2) ≥ 1.5
Upregulated genes by STZ
(405 probes)
203 probes
(174 genes)
443
203
(174 genes)
119
25
14
58
89
146
57
(45 genes)
565
(+D, STZ) / (+D) ≥ 1.5
Upregulated genes by STZ
in the absence of Oct4
(186 probes)
Tentative Oct4 target
(622 probes; 485 genes)

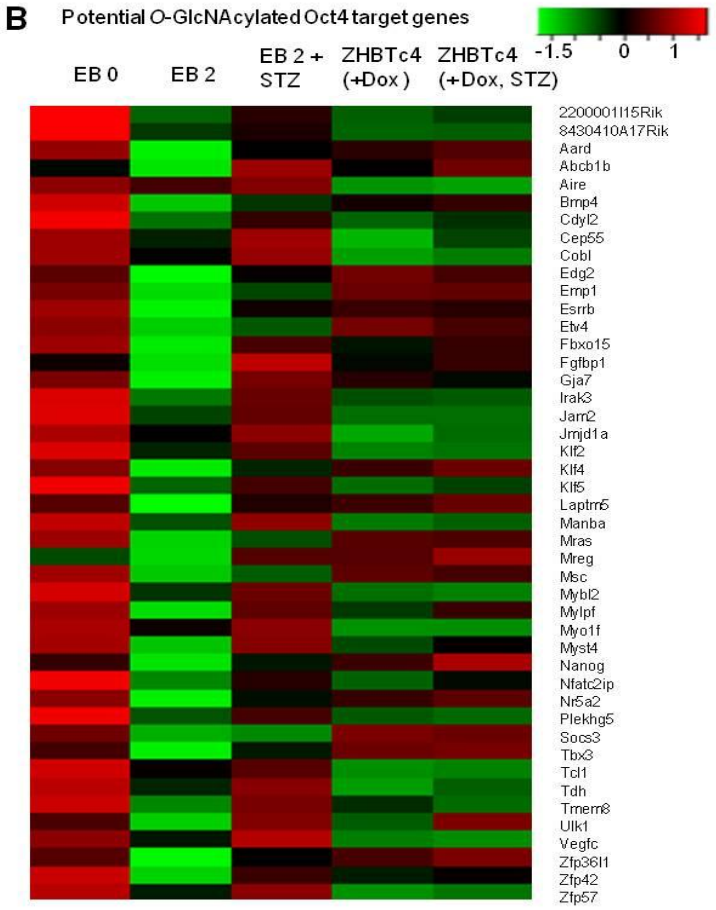

B
Potential O-GlcNAcylated Oct4 target genes
-1.5
0
1
EB 0
EB 2
EB 2 + STZ
ZHBTc4 (+Dox)
ZHBTc4 (+Dox, STZ)
2200001I15Rik
8430410A17Rik
Aard
Abcb1b
Aire
Bmp4
Cdyl2
Cep55
Cobl
Edg2
Emp1
Esrrb
Etv4
Fbxo15
Fgfbp1
Gja7
Irak3
Jam2
Jmjd1a
Klf2
Klf4
Klf5
Laptm5
Manba
Mras
Mreg
Msc
Mybl2
Mylpf
Myo1f
Myst4
Nanog
Nfatc2ip
Nr5a2
Plekhg5
Socs3
Tbx3
Tcl1
Tdh
Tmem8
Ulk1
Vegfc
Zfp36l1
Zfp42
Zfp57

Figure 6.

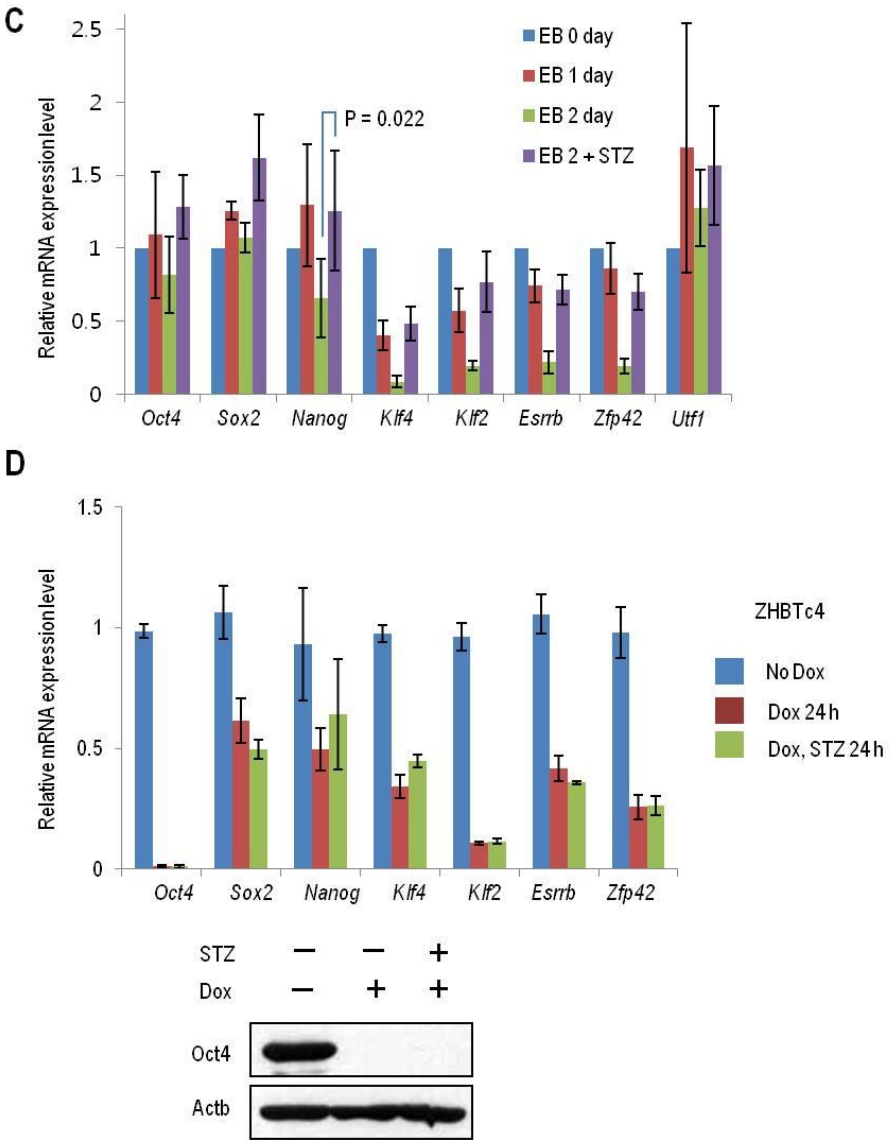
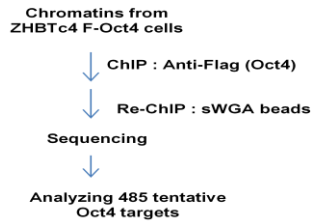
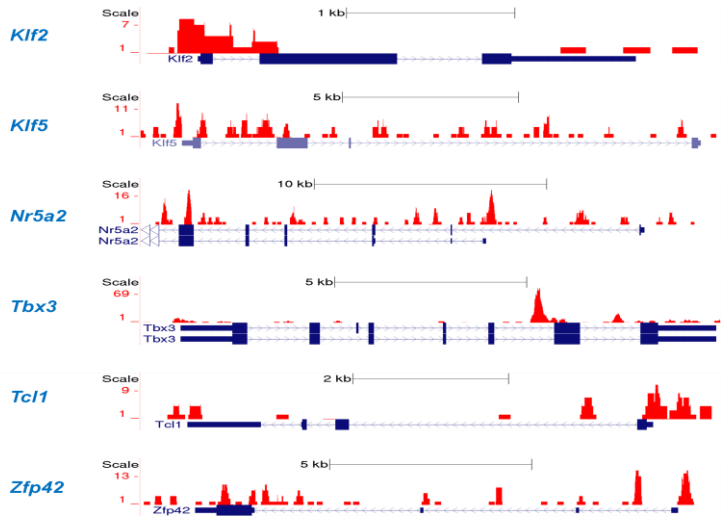

C
Relative mRNA expression level
2.5
2
1.5
1
0.5
0
P = 0.022
EB 0 day
EB 1 day
EB 2 day
EB 2 + STZ
Oct4
Sox2
Nanog
Klf4
Klf2
Esrrb
Zfp42
Utf1
D
Relative mRNA expression level
1.5
1
0.5
0
Oct4
Sox2
Nanog
Klf4
Klf2
Esrrb
Zfp42
ZHBTc4
No Dox
Dox 24 h
Dox, STZ 24 h
STZ
Dox
Oct4
Actb

Figure 6.

E



F



G

Refined tentative *O*-GlcNAcylated Oct4 target genes

8430410A17Rik	Irak3	Nanog	Ulk1
Aard	Jam2	Nfatc2ip	Vegfc
Abcb1b	Klf2	Nr5a2	Zfp36l1
Aire	Klf5	Plekhg5	Zfp42
Bmp4	Mreg	Socs3	Zfp57
Cep55	Msc	Tbx3	
Cobl	Mybl2	Tcl1	
Etv4	Myo1f	Tdh	

Figure 6. Identification of Genes Regulated by Oct4 *O*–GlcNAcylation

(A) mRNA expression levels in E14 cells before differentiation (EB 0), 2 days after differentiation to EB (EB 2), and 2 days after differentiation to EB in the presence of STZ (EB 2+ STZ) were analyzed by microarray. In addition, mRNA expression levels in ZHBTc4 cells treated with Dox for 24 h in the presence (+D,STZ) or absence(+D) of STZ were also analyzed by microarray. The probes (genes) were grouped according to the categories indicated and the number of probes (genes) corresponding to the categories were represented using Venn diagrams. The genes in the selected category were compared with tentative Oct4 target genes from published data (Sharov et al., 2008) .

(B) Potential *O*–GlcNAcylated Oct4 target genes were presented with clustered images.

(C) E14 cells were differentiated into EBs in the presence or absence of STZ. mRNA levels of selected putative *O*–

GlcNAcylated Oct4 target genes were analyzed by real-time qPCR.

(D) Oct4 protein levels in ZHBTc4 cells treated with Dox (1 μg/ml) for 1 day with/ without STZ (3 mM) were assessed by Western blot. At the same time, mRNA expression levels of selected genes were analyzed by real-time qPCR.

(E) *O*-GlcNAcylated Oct4 bound chromatin were purified by chromatin immunoprecipitation (ChIP) assays with anti-Flag, and then re-ChIP with sWGA from ZHBTc4 F-Oct4 cell chromatin. DNA fragments present in the purified chromatin were sequenced using genome Analyzer II. 485 tentative Oct4 target genes were analyzed.

(F) ChIP-seq binding profiles for some of potential *O*-GlcNAcylated Oct4 target genes.

(G) Refined tentative direct *O*-GlcNAcylated Oct4 target genes were presented.

Figure 7.

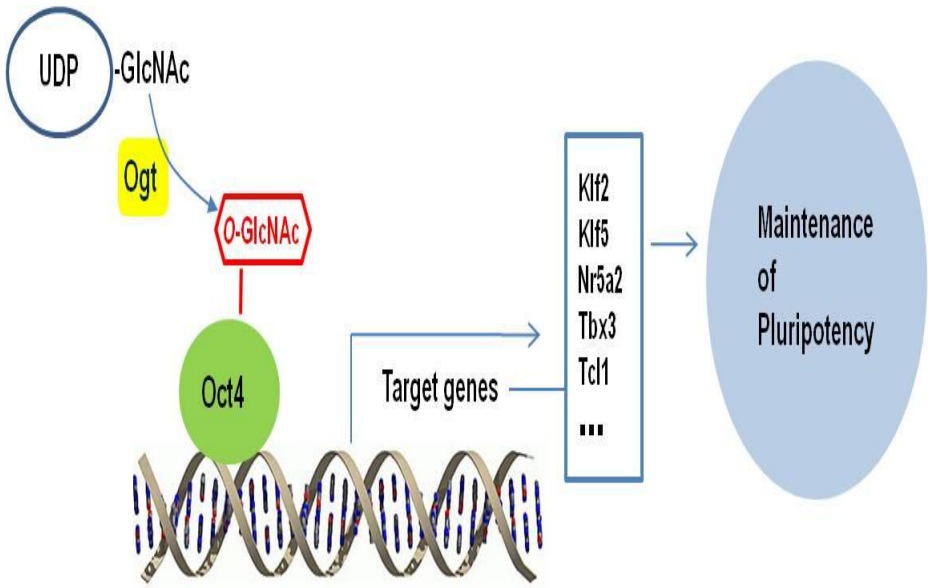

UDP
-GlcNAc
Ogt
O-GlcNAc
Oct4
Target genes
Klf2
Klf5
Nr5a2
Tbx3
Tcl1
...
Maintenance
of
Pluripotency

Figure 7. A Model Directly Linking *O*-GlcNAc with Pluripotency

Ogt adds *O*-GlcNAc on Oct4 and Sox2, the core components of pluripotency network in ESCs. *O*-GlcNAc on Oct4 is important for the expression of *O*-GlcNAcylated Oct4 target genes such as *Klf2, Klf5, Nr5a2, Tbx3* and *Tcl1*, which have all been implicated in the regulation of pluripotency.

Ⅲ. RESULTS (PART2)

Exit from pluripotency by Ctbp2 via silencing of active ESC genes during differentiation

1. Ctbp2, but not Ctbp1, is associated with undifferentiated state and is a downstream target of core transcription factors (CTFs) in ESCs

CtBP isoforms are transcriptional co-repressors via recruiting chromatin regulators (Chinnadurai, 2002; Kim et al., 2005; Shi et al., 2003). *In vertebrate*, *Ctbp2*-null mice, but not *Ctbp1*-null mice, experience early embryonic lethality (Hildebrand and Soriano, 2002). In addition, knockdown of *Ctbp2* leads to incomplete ESC differentiation and exit from pluripotency (Betschinger et al., 2013; Tarleton and Lemischka, 2010). Yet, the molecular mechanisms of these phenomena in ESCs are unknown.

To this end, we first examined the expression level of Ctbp1 and Ctbp2 in mouse ESCs during differentiation and in mouse embryonic fibroblast (MEF) cells. Unlike Ctbp1, the expression level of Ctbp2 was higher in undifferentiated versus differentiated ESCs and MEFs (Figures 8A-E). Further, on depletion of Oct4, treatment of ZHBTc4 ESCs with doxycycline for 1 day downregulated Ctbp2 (Figure 8F). Using published

ChIP-seq data in ESCs (Chen et al., 2008), we observed that ESC CTFs, such as Oct4, Sox2, Klf4, and Nanog (OSKN), enriched regions around *Ctbp2* but not *Ctbp1* (Figure 8G), indicating that Ctbp2 is an ESC factor that is highly expressed in the undifferentiated state.

Figure 8.

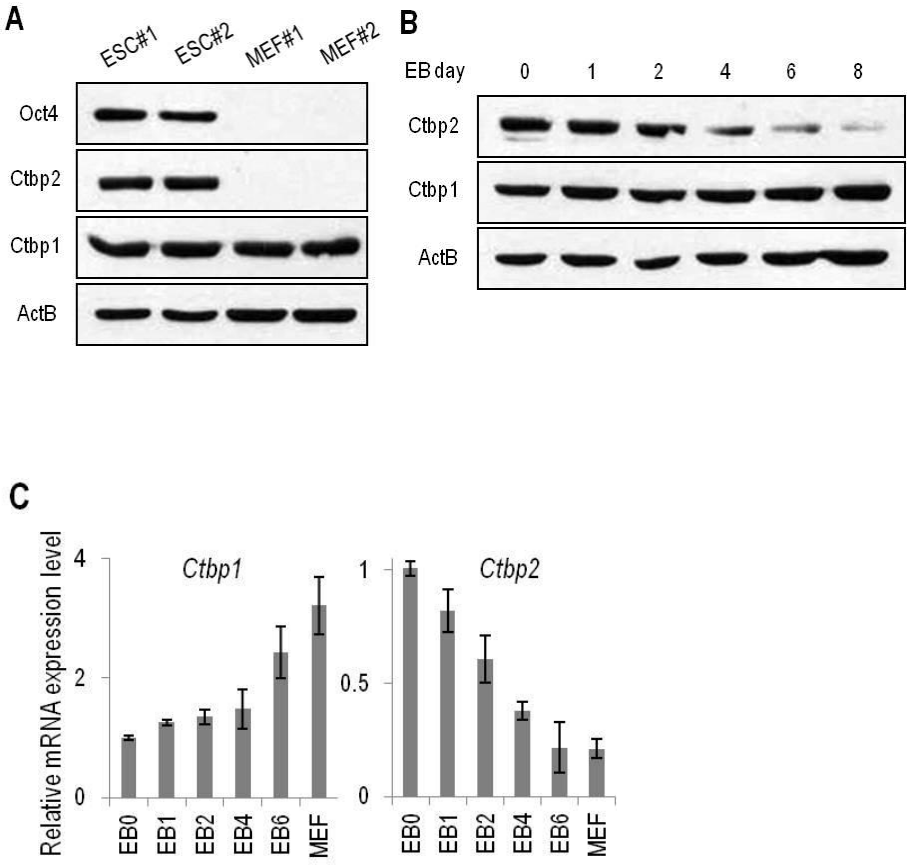

A
ESC#1
ESC#2
MEF#1
MEF#2
Oct4
Ctbp2
Ctbp1
ActB
B
EB day
0
1
2
4
6
8
Ctbp2
Ctbp1
ActB
C
Relative mRNA expression level
Ctbp1
Ctbp2
EB0
EB1
EB2
EB4
EB6
MEF

Figure 8.

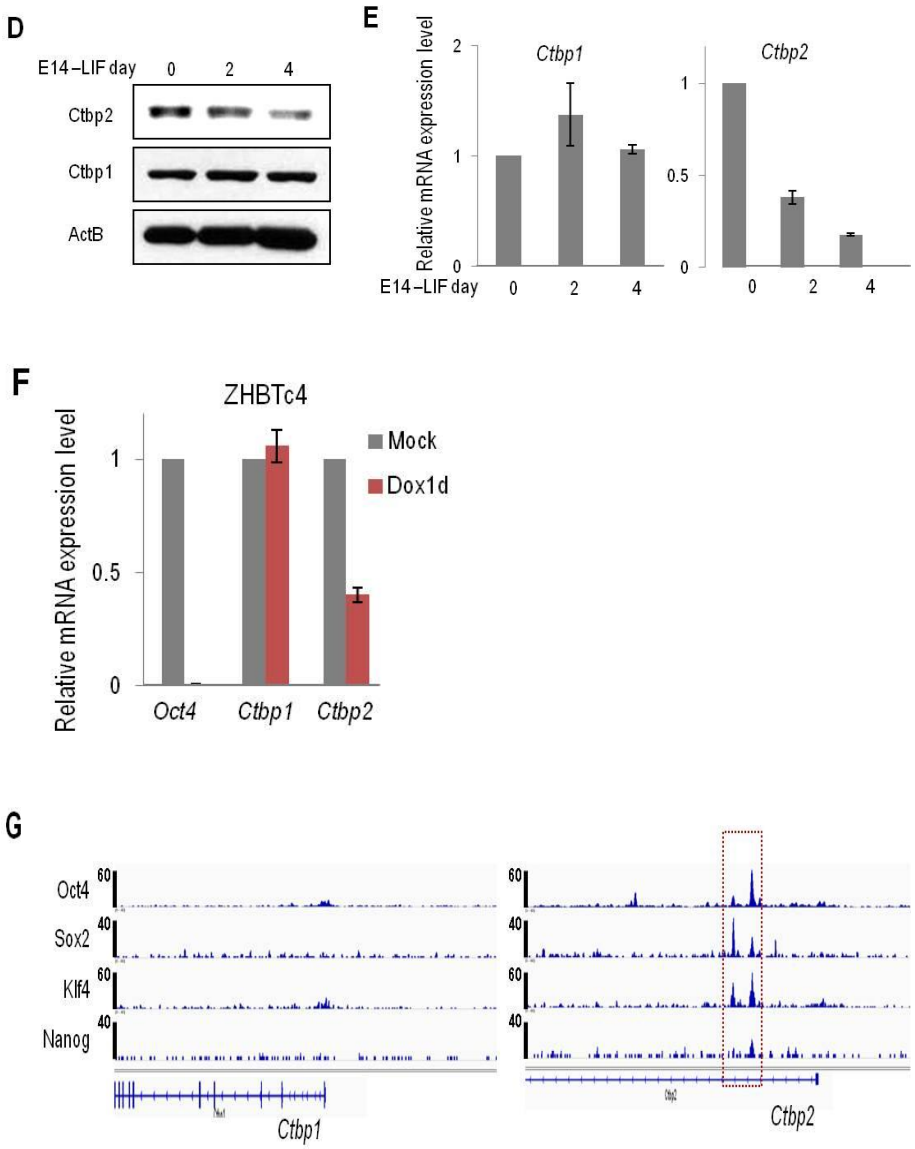

D
E14–LIF day 0 2 4
Ctbp2
Ctbp1
ActB
E
Relative mRNA expression level
Ctbp1
Ctbp2
E14–LIF day 0 2 4
F
ZHBTc4
Mock
Dox1d
Relative mRNA expression level
Oct4
Ctbp1
Ctbp2
G
Oct4
Sox2
Klf4
Nanog
Ctbp1
Ctbp2

Figure 8. Ctbp2, but not Ctbp1, is associated with undifferentiated state and is a downstream target of core transcription factors (CTFs) in ESCs

(A) In western blot analysis, expression level of Ctbp2, but not Ctbp1, is higher in undifferentiated E14 ESCs than in lineage-committed cells, MEFs. Expression level was detected with indicated antibodies and ActB was used as a control.

(B-E) During E14 ESC differentiation for indicated period by embryoid body (EB) formation and LIF withdrawal, expression levels of Ctbp2 and Ctbp1 were measured by western blotting and real-time qPCR. Ctbp2 expression level was decreased during ESC differentiation, while Ctbp1 expression level was in constant. Values represent mean ± standard deviation ($n \geq 3$).

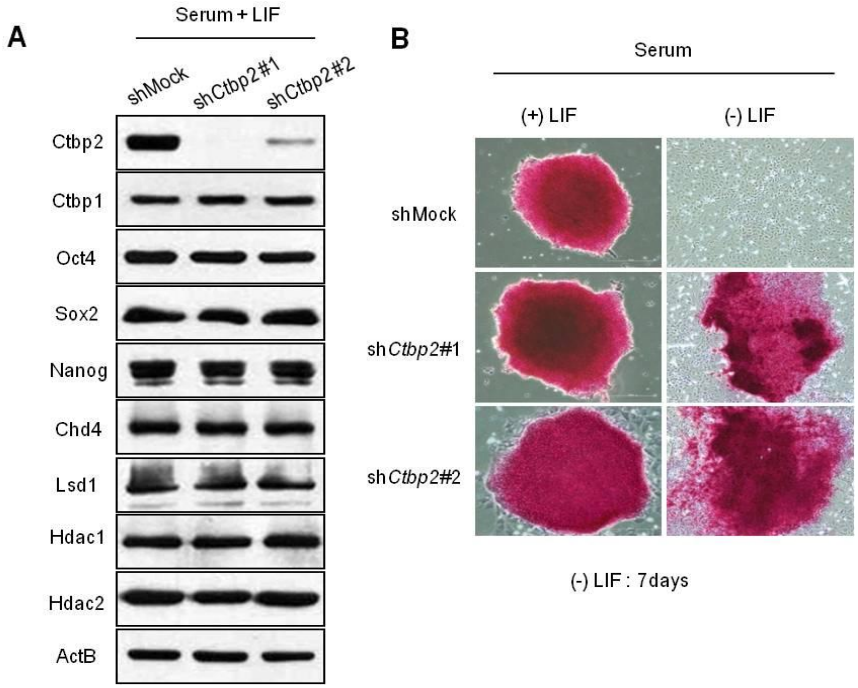
(F) In real-time qPCR analysis, *Ctbp2* expression, but not *Ctbp1*, was decreased after treatment of doxycycline (2 ug ml^{-1}) for 1 day in ZHBTc4 ESCs that lead to Oct4 depletion. Values represent mean ± standard deviation ($n \geq 3$). **g,** In published

ChIP-seq data (Whyte *et al.*, 2013; Chen et al., 2008), ESC CTFs [Oct4, Sox2, Klf4, and Nanog (OSKN)] enriched the regions around *Ctbp2*, but not *Ctbp1*.

2. Knockdown of *Ctbp2* in E14 ESCs did not affect ESC maintenance, but impaired ESC exit from pluripotency during differentiation

To determine the function of Ctbp2 in ESC maintenance or exit from pluripotency, 2 shRNA-mediated *Ctbp2*-knockdown ESCs were produced that did not affect the expression of Ctbp1 or pluripotency-related ESC genes (Figure 9A). Depletion of *Ctbp2* delayed the exit from pluripotency during differentiation, as evidenced by sustained alkaline phosphatase (AP) activity, ESC morphology, and expression of ESC pluripotency-related genes (Figures 9B-F).

Figure 9.


A
Serum + LIF
shMock
shCtbp2#1
shCtbp2#2
Ctbp2
Ctbp1
Oct4
Sox2
Nanog
Chd4
Lsd1
Hdac1
Hdac2
ActB
B
Serum
(+) LIF
(-) LIF
shMock
shCtbp2#1
shCtbp2#2
(-) LIF : 7days

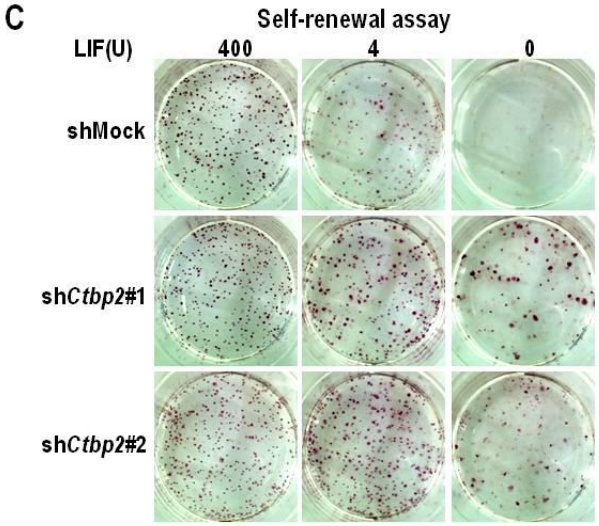

C
Self-renewal assay
LIF(U)
400
4
0
shMock
shCtbp2#1
shCtbp2#2

Figure 9.

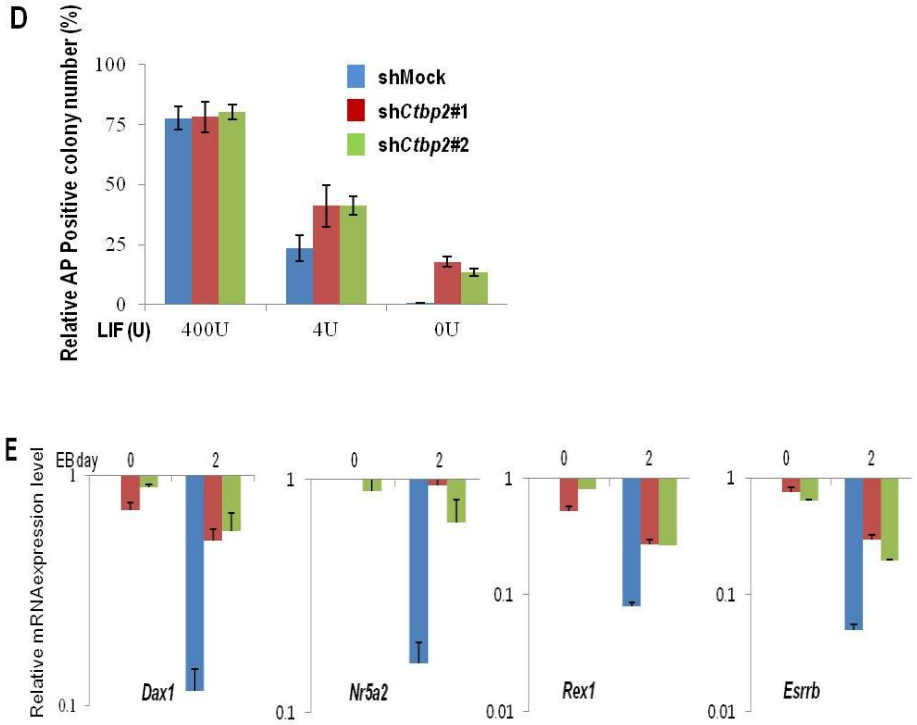

D
Relative AP Positive colony number (%)
100
75
50
25
0
shMock
shCtbp2#1
shCtbp2#2
LIF (U)
400U
4U
0U
E
Relative mRNA expression level
EB day
0
2
1
0.1
0.01
Dax1
Nr5a2
Rex1
Esrrb

Figure 9.

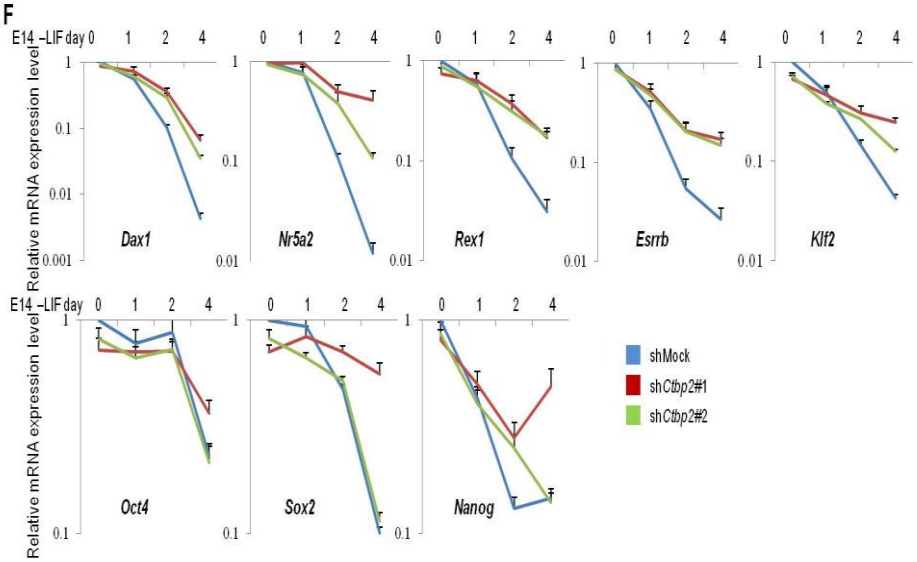

F
E14 –LIF day 0 1 2 4
Relative mRNA expression level
Dax1
Nr5a2
Rex1
Esrrb
Klf2
Oct4
Sox2
Nanog
shMock
shCtbp2#1
shCtbp2#2

Figure 9. Knockdown of *Ctbp2* in E14 ESCs did not affect ESC maintenance, but impaired ESC exit from pluripotency during differentiation

(A) There were no differences in Ctbp1 or pluripotent-associated gene expression between shMock-ESCs and two shRNA-mediated *Ctbp2*-knockdown ESCs by western blotting. ActB was used as a control.

(B-D) *Ctbp2*-knockdown ESCs were resistant to exit from pluripotency during ESC differentiation. The undifferentiated state was assessed by AP staining and ESC morphology. Values represent mean ± standard deviation ($n \geq 3$).

(E, F) Real-time qPCR analysis depicts that the expression levels of a subset of ESC active genes decreased during ESC differentiation in wild-type ESCs but partly sustained in *Ctbp2*-knockdown ESCs. Values represent mean ± standard deviation ($n \geq 3$).

3. Rescue of *Ctbp2* depletion phenotype during ESC differentiation

The phenotype was rescued with exogenous human CTBP2 (Figures 10A–D). Notably, when *Ctbp2*–knockdown ESCs were cultured continuously for more than 24 passages in the absence of LIF, restrained exit–from–pluripotency populations were enriched. These enriched populations had comparable AP activity and mRNA expression of pluripotency–related genes with undifferentiated wild–type ESCs (Figures 10E and 10F). This phenotype was also rescued with a transgene, indicating that the exit from pluripotency depends on Ctbp2 (Figures 10G and 10H). Also, by qRT–PCR and western blot analysis, the pluripotent related genes were down–regulated after ESC differentiation in wild–type ESCs but not in *Ctbp2*–knockdown ESCs (Figures 10I and 10J). Collectively, these results suggest that Ctbp2 is required to shut down active ESC genes to effect the proper transition of cell states during differentiation but that it is dispensable for ESC maintenance.

Figure 10.

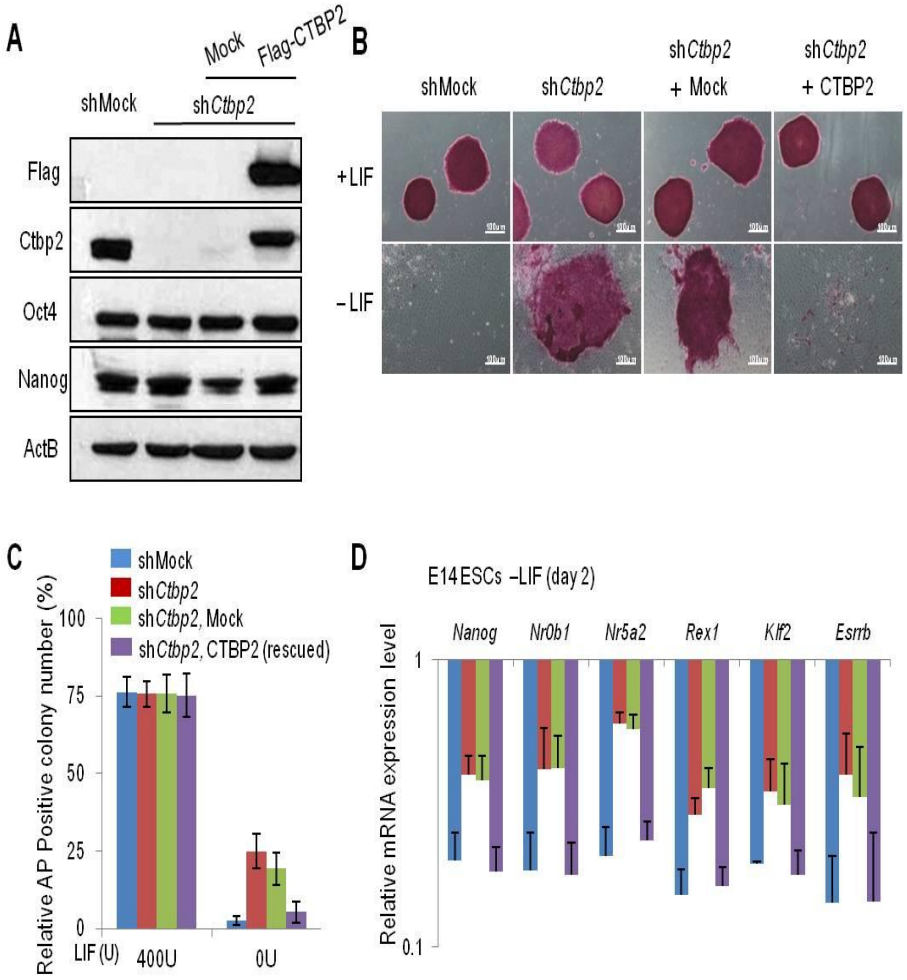

A
Mock
Flag-CTBP2
shMock
shCtbp2
Flag
Ctbp2
Oct4
Nanog
ActB
B
shMock
shCtbp2
shCtbp2 + Mock
shCtbp2 + CTBP2
+LIF
-LIF
C
shMock
shCtbp2
shCtbp2, Mock
shCtbp2, CTBP2 (rescued)
Relative AP Positive colony number (%)
100
75
50
25
0
LIF (U)
400U
0U
D
E14 ESCs -LIF (day 2)
Nanog
Nr0b1
Nr5a2
Rex1
Klf2
Esrrb
Relative mRNA expression level
1
0.1

Figure 10.

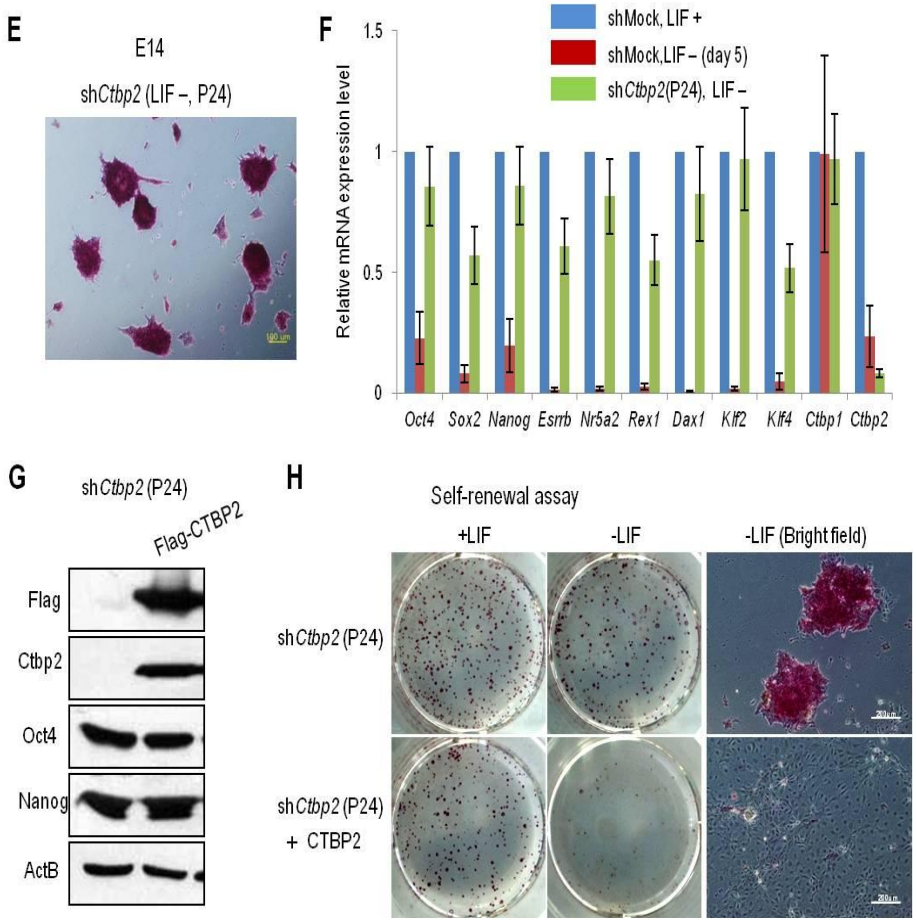
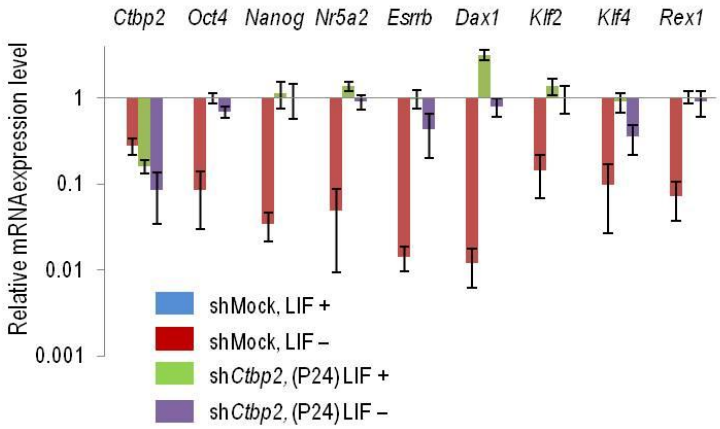

E
E14
shCtbp2 (LIF –, P24)
F
shMock, LIF +
shMock,LIF – (day 5)
shCtbp2(P24), LIF –
Relative mRNA expression level
1.5
1
0.5
0
Oct4
Sox2
Nanog
Esrrb
Nr5a2
Rex1
Dax1
Klf2
Klf4
Ctbp1
Ctbp2
G
shCtbp2 (P24)
Flag-CTBP2
Flag
Ctbp2
Oct4
Nanog
ActB
H
Self-renewal assay
+LIF
-LIF
-LIF (Bright field)
shCtbp2 (P24)
shCtbp2 (P24)
+ CTBP2

Figure 10.

I


Ctbp2
Oct4
Nanog
Nr5a2
Esrrb
Dax1
Klf2
Klf4
Rex1
Relative mRNAexpression level
1
0.1
0.01
0.001
shMock, LIF +
shMock, LIF –
shCtbp2, (P24) LIF +
shCtbp2, (P24) LIF –

J

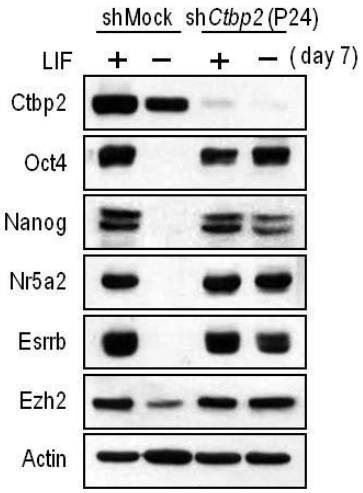

shMock
shCtbp2 (P24)
LIF + – + – (day 7)
Ctbp2
Oct4
Nanog
Nr5a2
Esrrb
Ezh2
Actin

Figure 10. Rescue of *Ctbp2* depletion phenotype in ESCs

(A) Western blot analysis with Flag and Ctbp2 antibodies of *Ctbp2* knocked down by mouse-specific sh*Ctbp2* in ESCs. *Ctbp2* expression was rescued by Flag-tagged CTBP2 overexpression in *Ctbp2*-knockdown ESCs (CTBP2-rescued ESCs). Oct4 and Nanog protein levels were not affected by Ctbp2. ActB was used as an internal control.

(B, C) Self-renewal assay and alkaline phosphatase (AP) staining of knockdown of *Ctbp2,* leading to incomplete exit from pluripotency during ESC differentiation, which was rescued by CTBP2 overexpression.

(D) In real-time qPCR analysis, knockdown of *Ctbp2* led to partly incomplete repression of a subset of ESC genes after LIF-withdrawal for 2 days compared with wild-type ESCs. However, those genes were properly repressed by ectopic expression of human CTBP2 in *Ctbp2*-knockdown ESCs (CTBP2-rescued *Ctbp2*-knockdown ESCs). Values represent mean ± standard deviation ($n \geq 3$).

(E) Continuous culture of more than 24 passages of *Ctbp2*-knockdown ESCs in the absence of LIF (sh*Ctbp2*, P24), with nearly positive AP staining and ESC morphology.

(F) Continuous culture of *Ctbp2*-knockdown ESCs (sh*Ctbp2*#1) in the absent of LIF for more than 24 passages (sh*Ctbp2*, P24), they had comparable mRNA expression levels of a subset of ESC active genes with wild-type undifferentiated ESCs. mRNA expression was analyzed by real-time qPCR. Values represent mean ± standard deviation ($n \geq 3$).

(G) *Ctbp2*-knockdown ESCs (sh*Ctbp2*, P24) were cultured in the presence of LIF and rescued by Flag-tagged CTBP2 overexpression. Expression of Oct4 and Nanog was similar between them.

(H) AP staining and ESC morphology showing that rescued ESCs exit normally from pluripotency during differentiation.

(I, J) Downregulation of a subset of active ESC genes that are

normally observed in wild-type ESCs after differentiation but not in *Ctbp2*-knockdown ESCs (I, real-time PCR; J, western blot).

4. Ctbp2 is a component of PRC2 complex

To examine the molecular function of Ctbp2 during ESC differentiation, we searched for Ctbp2 interacting proteins with Flag-tagged CTBP2-expressing ESCs by mass spectrometry (Figures 11A-E). We identified proteins that represented transcription factors and chromatin regulators that are associated with stem cell pluripotency and differentiation, such as the repressor complexes NuRD, PRC2, and Oct4 (Boyer et al., 2006; Esch et al., 2013; Reynolds et al., 2012; Whyte et al., 2012)(Figure 11F).

We confirmed that Ctbp2 interacts endogenously with Chd4, Hdac1, Lsd1, PRC2, and Oct4 by immunoprecipitation (Figures 11G-J). The NuRD complex represses a subset of active ESC genes under undifferentiated conditions, and its ablation impairs ESC differentiation due to the upregulation of active ESC genes (Reynolds et al., 2012). This finding suggests that the function of Ctbp2 in ESCs is distinct from that of NuRD, because *Ctbp2*-knockdown ESCs maintain comparable expression levels of active ESC genes in undifferentiated conditions.

Diminished PRC2 component impairs the exit from

pluripotency but is unessential for ESC maintenance (Simon and Kingston, 2009), similar to the phenotype of *Ctbp2*–knockdown ESCs. Jarid2, a newly identified component of PRC2, targets a subset of bivalent developmental genes (Landeira et al., 2010; Li et al., 2010; Pasini et al., 2010). Jarid2 was not found in the Ctbp2 complex in ESCs (Figure 11F), indicating that Ctbp2–PRC2 has a disparate function from Jarid2–PRC2 in ESCs. Ctbp has been identified as a partner of Polycomb group protein (PcG) in *Drosophila*, and PcG recruitment to its targets is abolished with mutant CtBP (Srinivasan and Atchison, 2004). Yet, the significance of the connection between Ctbp2 and PcG in mammals remains unknown.

By gel filtration assay, Ctbp2 co–eluted with core components of PRC2 in undifferentiated and differentiated states (Figure 11K). Thus, we speculate that Ctbp2 cooperates with PRC2 to silence active ESC genes and consequently affect proper exit from pluripotency during differentiation.

Figure 11.

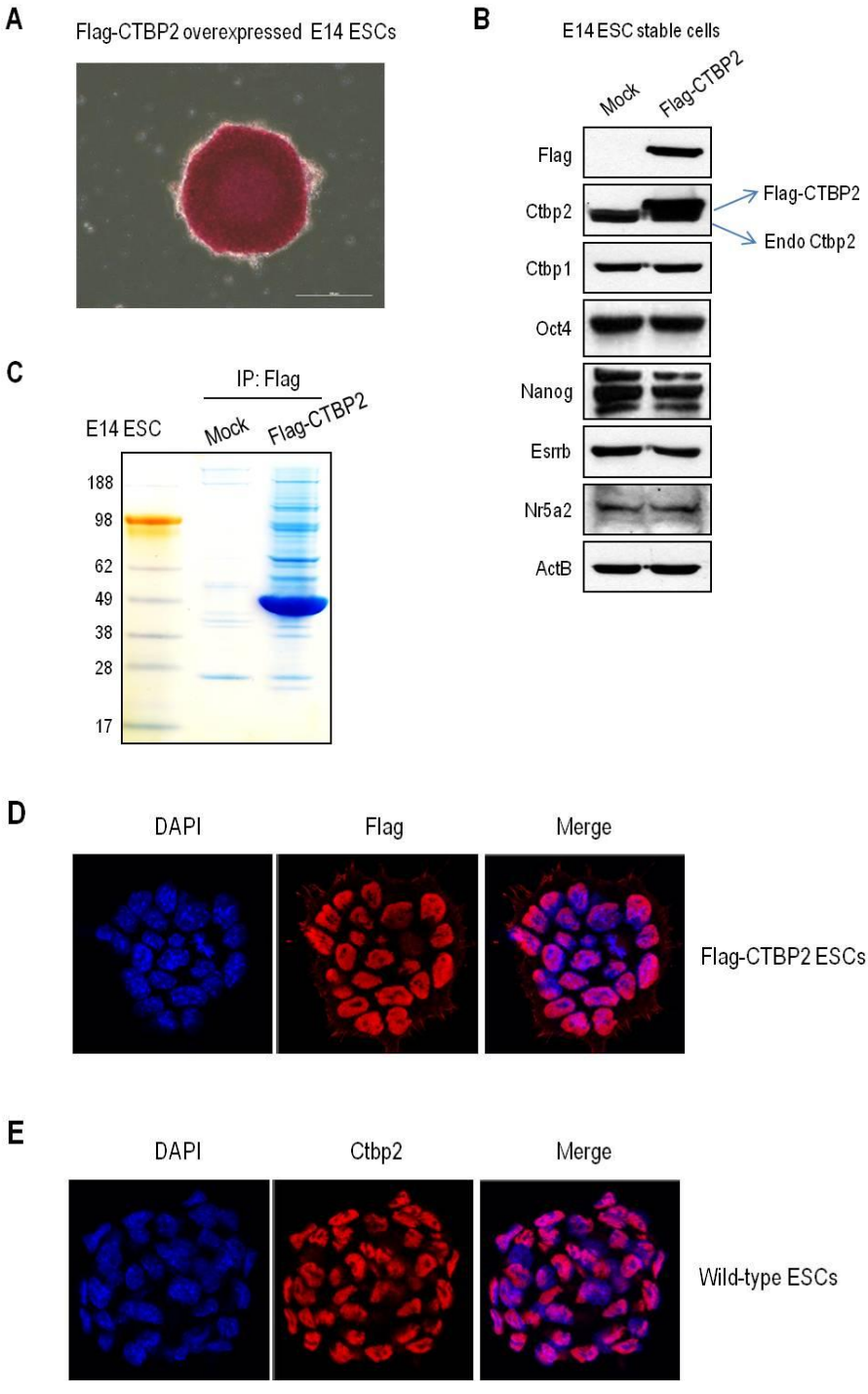

A
Flag-CTBP2 overexpressed E14 ESCs
B
E14 ESC stable cells
Mock
Flag-CTBP2
Flag
Ctbp2
Flag-CTBP2
Endo Ctbp2
Ctbp1
Oct4
Nanog
Esrrb
Nr5a2
ActB
C
IP: Flag
E14 ESC
Mock
Flag-CTBP2
188
98
62
49
38
28
17
D
DAPI
Flag
Merge
Flag-CTBP2 ESCs
E
DAPI
Ctbp2
Merge
Wild-type ESCs

Figure 11.

F

CtBP2 interacting proteins	ID	NurD related components		Transcription Factors		Oct4-interating components
Chd4	IPI00857771	O			O	
Gatad2a	IPI00229784	O			O	
Gatad2b	IPI00128615	O			O	
Mta2	IPI00128230	O			O	
Mta3	IPI00221805	O			O	
Hdac1	IPI00114232	O		O	O	O
Hdac2	IPI00137668	O		O		O
Mbd3	IPI00830957	O			O	
Rbbp4	IPI00122696	O		O		
Rbbp7	IPI00122698	O		O		
Apc	IPI00896736					
Axin1	IPI00875849					
Ctnnb1(Beta-catenin)	IPI00125899					
Gsk3beta	IPI00125319					
Nanog	IPI01019143		O		O	
Esrrb	IPI00874774		O		O	
Pou5f1	IPI00117218		O		O	
Klf8	IPI00119232		O			
E2f7	IPI00420139		O			
Suz12	IPI00396676			O		
Kdm1a (Lsd1)	IPI00453837			O	O	O
Ogt	IPI00420870			O	O	
Ehmt2 (G9a)	IPI00170261			O		O
Ezh2	IPI00468525			O		
Rcor1(CoRest1)	IPI00265217			O		O
Rcor2(CoRest2)	IPI00226581			O		O
Smarcb1	IPI00129145			O	O	
Smarca5	IPI00396739			O		
Smarcd1	IPI00121270			O		
EED				O		
Ctbp2 (BAIT)	IPI00114237				O	O
Ctbp1	IPI00128155					O
Zeb1	IPI00133262					O
Zfp217(Znf217)	IPI00758403					O
Zfp516	IPI00380736					O

Figure 11.

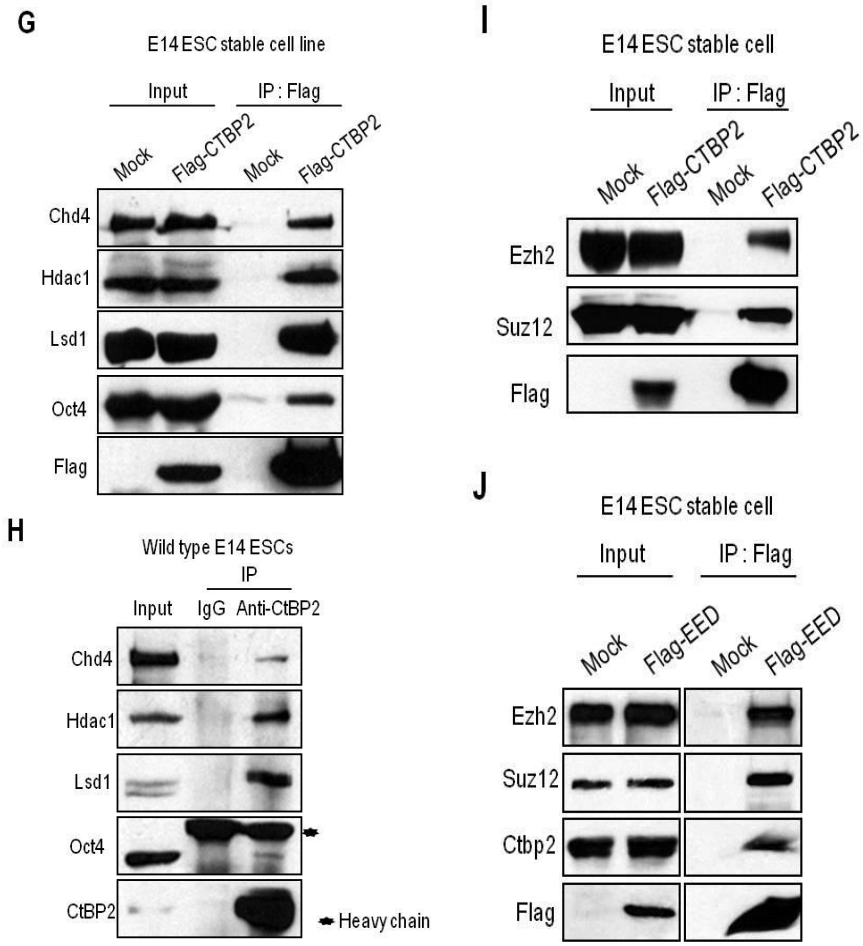

G
E14 ESC stable cell line
Input
IP : Flag
Mock
Flag-CTBP2
Mock
Flag-CTBP2
Chd4
Hdac1
Lsd1
Oct4
Flag
I
E14 ESC stable cell
Input
IP : Flag
Mock
Flag-CTBP2
Mock
Flag-CTBP2
Ezh2
Suz12
Flag
H
Wild type E14 ESCs
IP
Input
IgG
Anti-CtBP2
Chd4
Hdac1
Lsd1
Oct4
CtBP2
Heavy chain
J
E14 ESC stable cell
Input
IP : Flag
Mock
Flag-EED
Mock
Flag-EED
Ezh2
Suz12
Ctbp2
Flag

Figure 11.

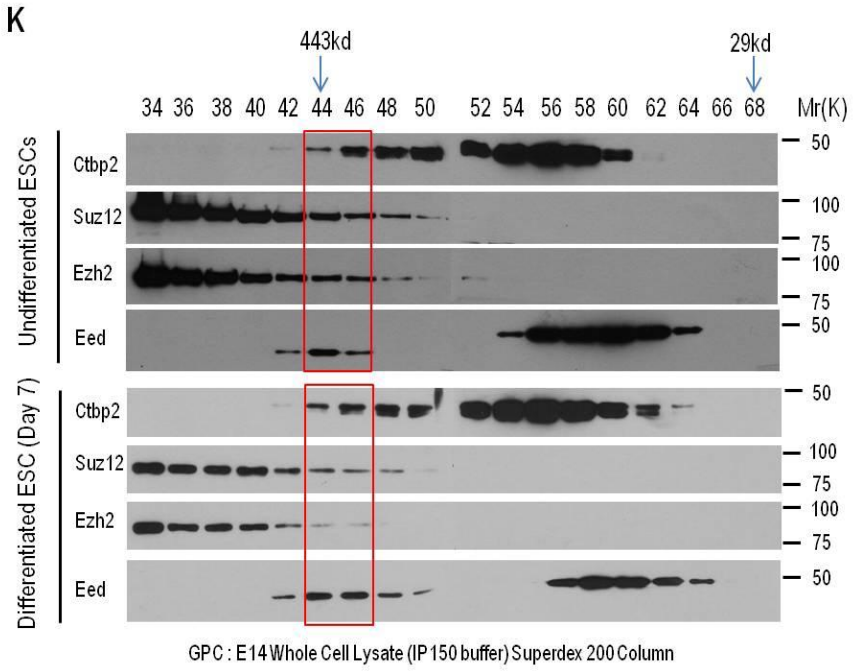

K
443kd
29kd
34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 64 66 68 Mr(K)
Undifferentiated ESCs
Ctbp2
Suz12
Ezh2
Eed
Differentiated ESC (Day 7)
Ctbp2
Suz12
Ezh2
Eed
50
100
75
100
75
50
50
100
75
100
75
50
GPC : E14 Whole Cell Lysate (IP 150 buffer) Superdex 200 Column

Figure 11. Ctbp2 is a component of PRC2 complex

(A, B) Flag-tagged human CTBP2 was stably transfected to ESCs. These cells have similar AP activity, ESC morphology and expression levels of ESC genes such as Oct4, Nanog, Esrrb, and Nr5a2 to those of wild-type ESCs. The expression of Flag, Ctbp1, Ctbp2, Oct4, Nanog, Esrrb, and Nr5a2 was detected by western blotting. ActB used as a control.

(C) Commassie blue-stained gel of affinity-purified Ctbp2 complex in ESCs stably overexpressed with Flag-CTBP2. b, By mass spectrometry, core PRC2 components were identified with reasonable peptide counting. c, Proteins were immunoprecipitated with FLAG antibodies from ESCs stably overexpressed with Flag-CTBP2; endogenous Ezh2 and Suz12 were detected with the indicated antibodies.

(D, E) The location of Ctbp2 in ESCs was observed by confocal microscopy. Indicated antibodies were used for immunostaining in Flag-tagged overexpressed-CTBP2 ESCs and wild-type ESCs.

(F) Ctbp2 interacting proteins were shown.

(G) Whole cell lysates from ESCs that stably overexpressed with Flag-tagged CTBP2 were immunoprecipitated by anti-Flag M2 affinity gel and eluted with 3XFlag-peptides. Interactions between Ctbp2 and Chd4, Hdac1, Lsd1 or Oct4 from elutes were detected by western blotting with indicated antibodies.

(H) E14 ESCs were immunoprecipitated by Ctbp2 and interactions between endogenous Ctbp2 and Chd4, Hdac1, Lsd1 or Oct4 were evaluated by western blotting with indicated antibodies.

(I) Proteins were immunoprecipitated with FLAG antibodies from ESCs stably overexpressed with Flag-CTBP2; endogenous Ezh2 and Suz12 were detected with the indicated antibodies.

(J) Proteins were immunoprecipitated with FLAG antibodies from ESCs stably overexpressed with Flag-EED; endogenous Ezh2, Suz12, and Ctbp2 were detected with the indicated antibodies.

(K) Fractionation of protein extracts from undifferentiated and differentiated E14 ESCs (day 7) by gel-filtration analysis, showing that Ctbp2 co-migrates with the core PRC2 components. ESC lysates were separated by molecular weight, indicated above the panels, and the antibodies for western blot are shown on the left.

5. Ctbp2 regulates PRC2 recruitment, but not Oct4 recruitment

Ezh2 is a critical component of PRC2 that catalyzes H3K27me3 to its target genes to firm repressive chromatin, regulating development and lineage commitment in ESCs (Simon and Kingston, 2009). To determine whether Ctbp2 regulates PRC2-associated target genes, a subset of loci that were co-bound by Ctbp2 and PRC2 was subjected to ChIP with Ezh2 and Ctbp2. Notably, under undifferentiated conditions, Ezh2 recruitment to a subset of PRC2 target genes that were co-bound by Ctbp2 and Ezh2 decreased in *Ctbp2*-knockdown ESCs versus wild-type ESCs (Figure 12A). Instead, Oct4 recruitment to ESC gene regions that were co-occupied by Ctbp2 and Oct4 was similar between *Ctbp2*-knockdown and wild-type ESCs (Figure 12B).

During differentiation, Ezh2 enrichment in active ESC genes was observed in wild-type ESCs, whereas such occupancy decreased significantly in *Ctbp2*-knockdown ESCs (Figure 12C). As expected, wild-type ESCs harbored increased H3K27me3 peaks in a subset of active ESC genes, whereas *Ctbp2*-knockdown ESCs had significantly lower peaks during

differentiation (Figure 12D). These results indicate that Ctbp2 controls Ezh2 recruitment to a subset of PRC2 target genes in undifferentiated ESCs and to ESC active genes during differentiation, all of which are regions that require Ezh2 for silencing.

In active ESC loci that were co-bound by Ctbp2 and Oct4, Ctbp2 did not affect the expression of their target genes in undifferentiated ESCs, and active ESC genes did not have Ezh2 peaks or H3K27me3 marks, even high expression of Ctbp2. These results might be attributed to findings that histone acetyltransferase-mediated epigenetically active states antagonize PcG-mediated gene silencing (Klymenko and Muller, 2004; Tie et al., 2009) and that the repressive activity of Ctbp2 decreases in the presence of p300 (Zhao et al., 2006).

Figure 12.

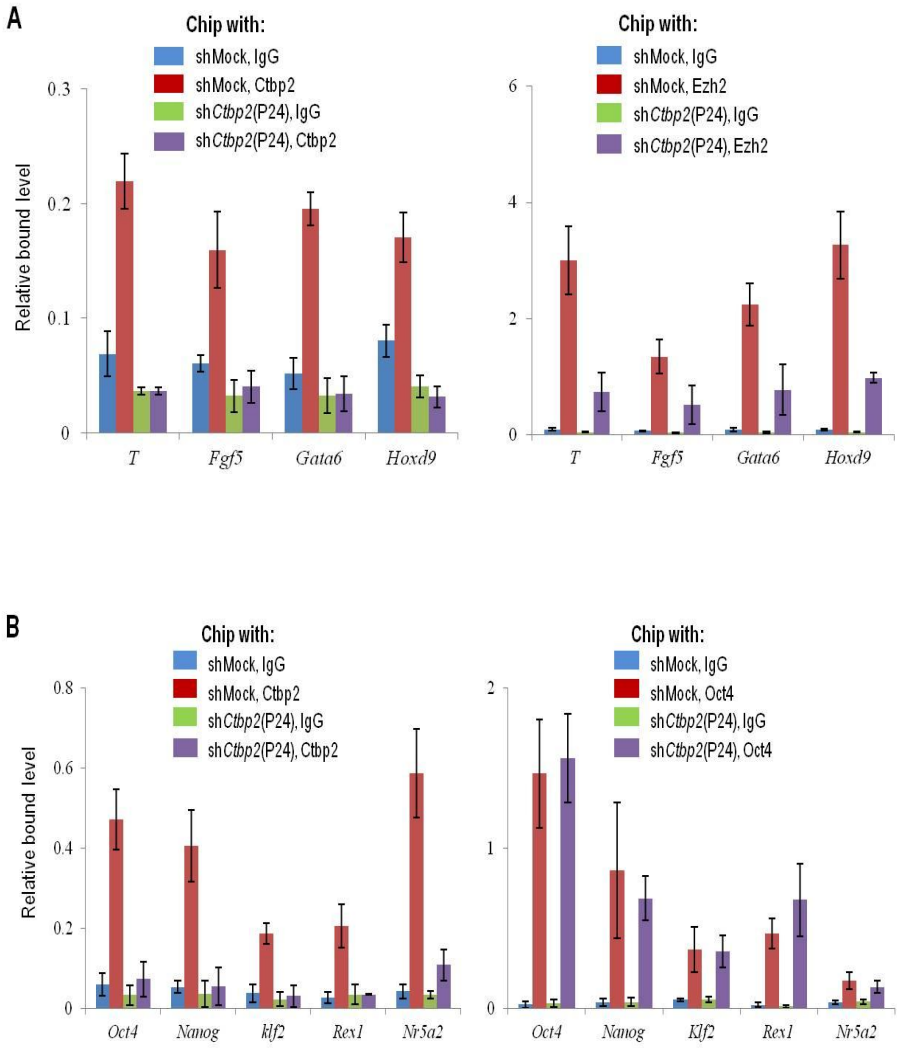

A
Chip with:
shMock, IgG
shMock, Ctbp2
shCtbp2(P24), IgG
shCtbp2(P24), Ctbp2
Relative bound level
0.3
0.2
0.1
0
T
Fgf5
Gata6
Hoxd9
Chip with:
shMock, IgG
shMock, Ezh2
shCtbp2(P24), IgG
shCtbp2(P24), Ezh2
6
4
2
0
T
Fgf5
Gata6
Hoxd9
B
Chip with:
shMock, IgG
shMock, Ctbp2
shCtbp2(P24), IgG
shCtbp2(P24), Ctbp2
Relative bound level
0.8
0.6
0.4
0.2
0
Oct4
Nanog
klf2
Rex1
Nr5a2
Chip with:
shMock, IgG
shMock, Oct4
shCtbp2(P24), IgG
shCtbp2(P24), Oct4
2
1
0
Oct4
Nanog
Klf2
Rex1
Nr5a2

Figure 12.

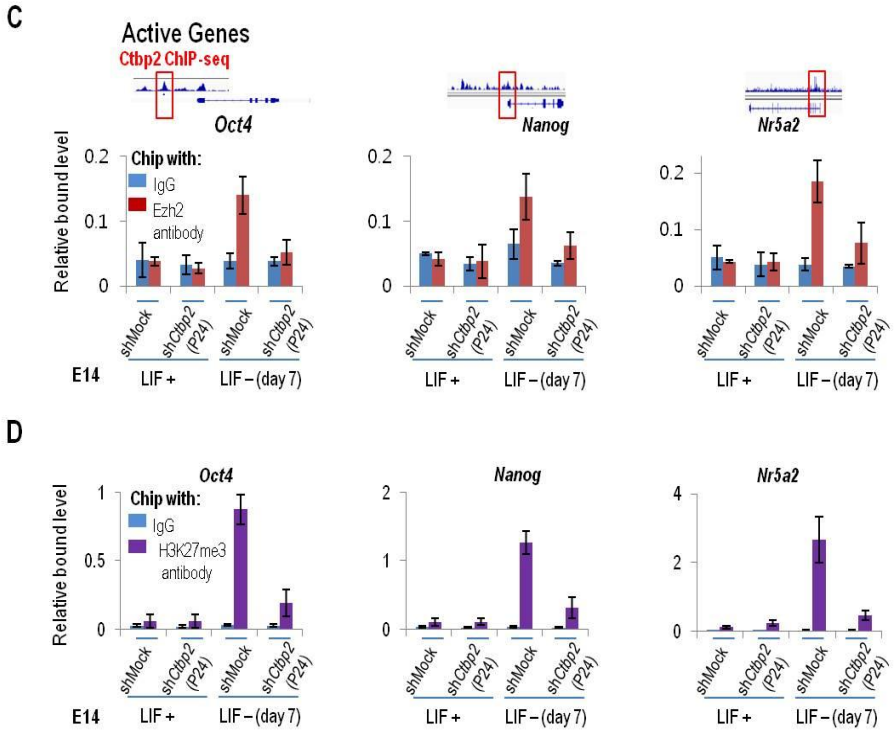

C
Active Genes
Ctbp2 ChIP-seq
Oct4
Nanog
Nr5a2
Relative bound level
Chip with:
IgG
Ezh2 antibody
0.2
0.1
0
shMock
shCtbp2 (P24)
E14
LIF +
LIF – (day 7)
D
Chip with:
IgG
H3K27me3 antibody
1
0.5
2
4

Figure 12. Ctbp2 regulates PRC2 recruitment, but not Oct4 recruitment

(A) ChIP-qPCR analysis of Ezh2 and Ctbp2 in undifferentiated wild-type and *Ctbp2*-knockdown E14 ESCs have been shown. Left graph shows that Ctbp2 binds to a subset of loci occupied by Ezh2 in wild-type ESCs. Ezh2 binding was significantly decreased in *Ctbp2*-knockdown ESCs compared with wild-type ESCs (right graph). Values represent mean ± standard deviation (n≥3).

(B) ChIP-qPCR analysis of Ctbp2 and Oct4 in undifferentiated wild-type and *Ctbp2*-knockdown E14 ESCs reveal that, regardless of Ctbp2-knockdown, there is no significant change in Oct4 binding to chromatins of several ESC active genes where co-bound loci by Oct4 and Ctbp2 in our ChIP-seq analysis described above. Values represent mean ± standard deviation (n≥3).

(C, D) ChIP-qPCR analysis of Ezh2 and H3K27me3 in regions of a subset of active ESC genes where Ctbp2 and Oct4 co-

occupy during ESC differentiation. Ezh2 and H3K27me3 occupancy in active ESC genes increased after differentiation in wild-type ESCs (LIF- for 7days) but declined significantly in *Ctbp2*-knockdown ESCs. Values represent mean ± standard deviation (n≥3).

6. Oct4 depletion-mediated differentiation is impeded in the absence of Ctbp2

Next, we examined whether Ctbp2 influences the recruitment of H3K27me3 to active ESC genes using ZHBTc4 ESCs, based on previous results that treatment of these cells with doxycycline decreases the occupancy of active Oct4-occupied genes by p300, H3K27ac, and H3K4me3 and results in differentiation (Ang et al., 2011; Niwa et al., 2000). After a 3-day depletion of Oct4, *Ctbp2*-knockdown ESCs had partially sustained AP positivity, ESC morphology, and ESC genes (Figures 13A-C). Further, the levels of H3K27me3 in active ESC genes declined significantly in *Ctbp2*-knockdown ESCs compared with wild-type ESCs on depletion of Oct4 (Figure 13D).

The results are consistent with our model, in which Ctbp2 is essential for the epigenetically inactive state of active ESC genes when they must be stably silenced.

Figure 13.

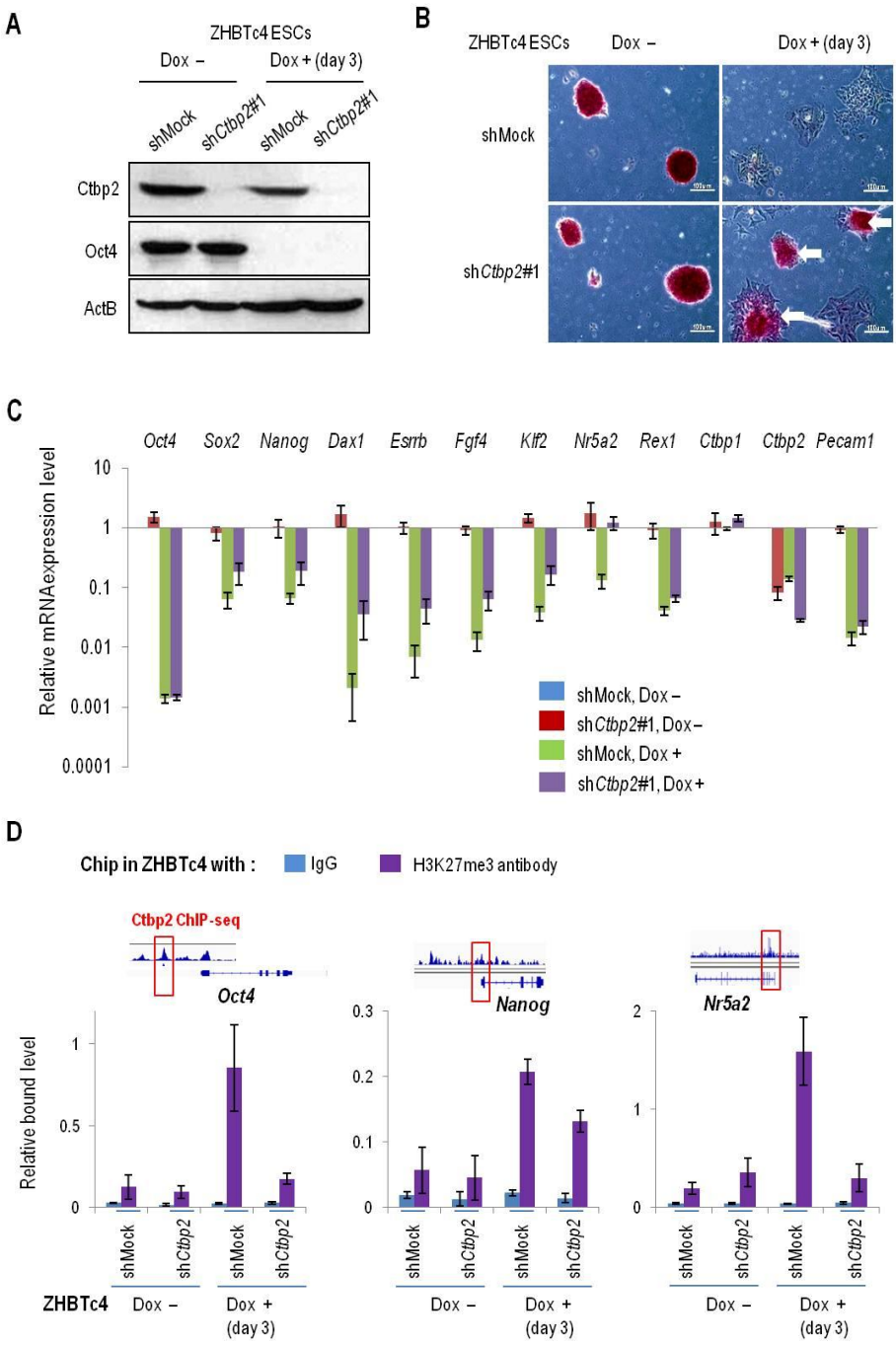

A
ZHBTc4 ESCs
Dox –
Dox + (day 3)
shMock
shCtbp2#1
shMock
shCtbp2#1
Ctbp2
Oct4
ActB
B
ZHBTc4 ESCs
Dox –
Dox + (day 3)
shMock
shCtbp2#1
C
Oct4
Sox2
Nanog
Dax1
Esrrb
Fgf4
Klf2
Nr5a2
Rex1
Ctbp1
Ctbp2
Pecam1
Relative mRNAexpression level
10
1
0.1
0.01
0.001
0.0001
shMock, Dox –
shCtbp2#1, Dox –
shMock, Dox +
shCtbp2#1, Dox +
D
Chip in ZHBTc4 with :
IgG
H3K27me3 antibody
Ctbp2 ChIP-seq
Oct4
Nanog
Nr5a2
Relative bound level
shMock
shCtbp2
ZHBTc4
Dox –
Dox +
(day 3)

Figure 13. Oct4 depletion-mediated differentiation is impeded in the absence of Ctbp2

(A) Depletion of *Ctbp2* in ZHBTc4 ESCs stably infected with shCtbp2, confirmed by western blot with Ctbp2 antibody. Oct4 depletion by doxycycline treatment (2ug ml^{-1}) for 3 days in wild-type and *Ctbp2*-knockdown ZHBTc4 ESCs was detected with Oct4 antibody. ActB was used as an endogenous control.

(B, C) *Ctbp2*-knockdown ZHBTc4 ESCs sustained AP activity, ESC morphology, and a subset of active ESC genes after treatment with doxycycline for 3 days compared with wild-type ZHBTc4 ESCs. Values represent mean ± standard deviation ($n \geq 3$).

(D) ChIP-qPCR analysis of H3K27me3 in regions of a subset of active ESC genes where Ctbp2 and Oct4 co-occupy in ZHBTc4 cells after treatment with or without doxycycline for 3 days. Increased binding of H3K27me3 to active genes after doxycycline treatment in wild-type ESCs was significantly

attenuated in *Ctbp2*-knockdown ZHBTc4 ESCs. Values represent mean ± standard deviation (n≥3).

Figure 14.

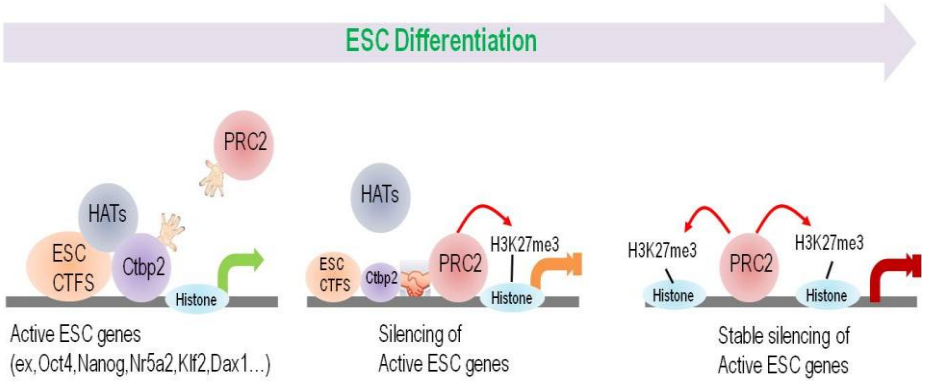

ESC Differentiation
PRC2
HATs
ESC
CTFS
Ctbp2
Histone
Active ESC genes
(ex,Oct4,Nanog,Nr5a2,Klf2,Dax1...)
HATs
ESC
CTFS
Ctbp2
PRC2
H3K27me3
Histone
Silencing of
Active ESC genes
H3K27me3
PRC2
H3K27me3
Histone
Histone
Stable silencing of
Active ESC genes

Figure 14. A schematic model. In undifferentiated ESCs, Ctbp2 primes their position in actively transcribed ESC genes and their repressive activity was compromised by HATs. During ESC differentiation, Ctbp2 recruits PRC2 toward the active ESC genes to mark H3K27me3 for stable silencing of active ESC genes, which is pivotal for lineage commitment.

Ⅳ. DISCUSSION

First part of results demonstrates that *O*-GlcNAc is an important regulator of pluripotency. *O*-GlcNAc is critical for ESC self-renewal and somatic cell reprogramming to pluripotency. Mild knockdown of Ogt reduces ESC potential to self-renew. Increasing global *O*-GlcNAc levels by STZ inhibits normal ESC differentiation. Decreasing *O*-GlcNAc levels by shOgts reduced the reprogramming efficiency of generating iPSCs, and increasing *O*-GlcNAc levels by OGT overexpression increased reprogramming efficiency. Together, these results suggest that *O*-GlcNAc is important for both maintaining and inducing pluripotency. We also demonstrate here that *O*-GlcNAcylation is critical for Oct4 transcriptional activity and that inhibition of Oct4 *O*-GlcNAcylation disrupts pluripotency. *O*-GlcNAcylation of Oct4 is important for Oct4 to activate many pluripotency-related genes such as *Klf2, Klf5, Nr5a2, Tbx3* and *Tcl1* (Figure 7).

ESCs are commonly cultured in high-glucose (25 mM) media, and the restriction of nutrients such as glucose and amino acids is a basic strategy employed in the *in vitro* differentiation of ES cells (Mochizuki et al., 2011). Because *O*-GlcNAc signaling is

sensitive to nutrient status, we hypothesized that nutrient status may affect pluripotency and differentiation via *O*-GlcNAcylation. In accordance with this idea, we found that reprogramming is inefficient under low-glucose (5.5 mM) conditions compared to high-glucose conditions.

The threonine 228 residue on Oct4 undergoes *O*-GlcNAcylation increased Oct4 activity. We could not detect *O*-GlcNAc at Oct4 Serine 229 using mass spectroscopy, although Oct4 S229A was reduced in *O*-GlcNAc level. This does not completely eliminate the possibility that S229 is also *O*-GlcNAcylated, however, the effect of S229A is much milder than that of T228A. The double alanine mutant of Oct4 at T228 and S229 was not completely depleted in *O*-GlcNAcylation, leaving the possibility that additional *O*-GlcNAcylation site(s) may exist. However, we demonstrate that T228 is the major *O*-GlcNAcylation site that regulates Oct4 transcriptional activity. In contrast to Oct4 WT, Oct4 T228A transcriptional activity was not affected by perturbation of *O*-GlcNAc levels.

Serine 229 on Oct4 has been identified as a phosphorylation site (Swaney et al., 2009), and has been predicted to be modified/catalyzed by PKA (protein kinase A) through

bioinformatic analyses (Saxe et al., 2009). Oct4 S229D, which mimics phosphorylation, decreased Oct4 reporter activity. However, contradictory results have been obtained at the same paper with treatment of a PKA activator, 8-Br-cAMP, which elevated both Oct4 protein levels and reporter activity (Saxe et al., 2009). ESCs treated with 8-Br-cAMP showed conflicting results with ESC differentiation occurring in the presence of LIF, and ES cell self renewal maintained in the absence of LIF (Faherty et al., 2007). Our results show that Oct4 S229D completely loses transcriptional activity, and cannot induce reprogramming to iPSC. Considering that Oct4 S229A retains ~50 % of WT transcriptional activity, and ~30 % of WT ability to induce iPSCs, Oct4 S229D may faithfully mimic phosphorylation. While it is clear that phosphorylation at S229 negatively regulates Oct4 activity, the kinase(s) responsible for the modification is (are) not apparent yet.

Using microarray we identified potential target genes that are dominantly regulated by *O*-GlcNAcylated Oct4. To select direct target genes, we used serial ChIP assay to purify Oct4 and then the *O*-GlcNAcylated protein bound DNA. Several pluripotency related factors, such as *Klf2, Klf5, Nr5a2, Tbx3* and *Tcl1*, are

identified as direct targets. Among them, Klf2 and Klf5 can reprogram MEFs in conjunction with Oct4 and Sox2 (Nakagawa et al., 2008) and sustain ESC self-renewal (Hall et al., 2009). Nr5a2 can replace Oct4 (Heng et al., 2010) in the reprogramming process and Tbx3 improves the germ-line competency of iPSCs (Han et al., 2010). By regulating the expression levels of these genes, *O*-GlcNAc of Oct4 may adjust the pluripotent state according to the external signals.

Polycomb group (PcG) complexes have been suggested to link *O*-GlcNAc cycling to epigenetic regulation of chromatin and ESC differentiation (Love et al., 2010). PcG complexes set the stage in ESCs for the acquisition and maintenance of specific developmental gene expression programs during development. PcG complexes are essential for establishment of bivalent chromatin, which carry both active and repressive histone marks, and are a characteristic of ESCs (Sauvageau and Sauvageau, 2010; Surface et al., 2010). OGT is a component of PcG complexes, and its activity is essential for full repression by PcG (Gambetta et al., 2009). Thus, OGT may regulate ESC differentiation in part through PcG complexes. However, PcG complexes are not required for the maintenance of self-

renewal in ESCs (Sauvageau and Sauvageau, 2010; Surface et al., 2010) and cannot sufficiently account for the critical role of *O*-GlcNAc in ESC self-renewal.

In this first part of study, we reveal that *O*-GlcNAc regulates the function of Oct4 and Sox2, which are core components of the pluripotency network. In undifferentiated ES cells, Oct4 is *O*-GlcNAcylated and upregulates several genes critical for maintenance of ESC identity. Upon differentiation, *O*-GlcNAc is rapidly detached from Oct4 and the expression of pluripotency-related Oct4 target genes is decreased. These findings may explain why *O*-GlcNAc is critical for pluripotency.

Based on second part of results, Ctbp2 resides in active ESC genes and co-occupies regions with ESC CTFs in undifferentiated ESCs by genome-wide ChIP-sequencing. Further, ablation of Ctbp2 leads to inappropriate gene silencing in ESCs, thereby sustaining ESC maintenance during differentiation. Ctbp2 in ESCs associates with core components of PRC2 and Oct4. Ctbp2 regulates the recruitment of Ezh2 to active ESC genes and stimulates H3K27me3 during differentiation. We demonstrate that Ctbp2 pre-occupies regions of active ESC genes in undifferentiated ESCs and

guides PRC2 toward active ESC genes to mark H3K27me3 for stable silencing. We propose that Ctbp2 is a fine-tuning chromatin regulator in ESCs, ensuring adequate transition by priming its position in active ESC genes to shut them down epigenetically on induction of lineage commitment programs. This finding might explain why ablation of PRC2 components prevents the differentiation of ESCs into lineage-committed cells.

V. REFERENCES

Ambrosetti, D.C., Scholer, H.R., Dailey, L., and Basilico, C. (2000). Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer. J Biol Chem *275*, 23387-23397.

Ang, Y.S., Tsai, S.Y., Lee, D.F., Monk, J., Su, J., Ratnakumar, K., Ding, J., Ge, Y., Darr, H., Chang, B., *et al.* (2011). Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. Cell *145*, 183-197.

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.

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

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline

development. Nature *450*, 1230-1234.

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.

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

Cho, H.J., Lee, C.S., Kwon, Y.W., Paek, J.S., Lee, S.H., Hur, J., Lee, E.J., Roh, T.Y., Chu, I.S., Leem, S.H., *et al.* (2010). Induction of pluripotent stem cells from adult somatic cells by protein-based reprogramming without genetic manipulation. Blood *116*, 386-395.

Conaghan, J., Handyside, A.H., Winston, R.M., and Leese, H.J. (1993). Effects of pyruvate and glucose on the development of human preimplantation embryos in vitro. J Reprod Fertil *99*, 87-95.

Ding J., Xu H., Faiola F., Ma'ayan A., Wang J. (2012). Oct4 links multiple epigenetic pathways to the pluripotency network. Cell Res 22(1), 155-67

Esch, D., Vahokoski, J., Groves, M.R., Pogenberg, V., Cojocaru,

V., Vom Bruch, H., Han, D., Drexler, H.C., Arauzo-Bravo, M.J., Ng, C.K., *et al.* (2013). A unique Oct4 interface is crucial for reprogramming to pluripotency. Nat Cell Biol *15*, 295-301.

Faherty, S., Fitzgerald, A., Keohan, M., and Quinlan, L.R. (2007). Self-renewal and differentiation of mouse embryonic stem cells as measured by Oct4 expression: the role of the cAMP/PKA pathway. In Vitro Cell Dev Biol Anim *43*, 37-47.

Gambetta, M.C., Oktaba, K., and Muller, J. (2009). Essential role of the glycosyltransferase sxc/Ogt in polycomb repression. Science *325*, 93-96.

Gandy, J.C., Rountree, A.E., and Bijur, G.N. (2006). Akt1 is dynamically modified with O-GlcNAc following treatments with PUGNAc and insulin-like growth factor-1. FEBS Lett *580*, 3051-3058.

Hall, J., Guo, G., Wray, J., Eyres, I., Nichols, J., Grotewold, L., Morfopoulou, S., Humphreys, P., Mansfield, W., Walker, R., *et al.* (2009). Oct4 and LIF/Stat3 additively induce Kruppel factors to sustain embryonic stem cell self-renewal. Cell Stem Cell *5*, 597-609.

Han, J., Yuan, P., Yang, H., Zhang, J., Soh, B.S., Li, P., Lim, S.L., Cao, S., Tay, J., Orlov, Y.L., *et al.* (2010). Tbx3 improves the

germ-line competency of induced pluripotent stem cells. Nature *463*, 1096-1100.

Hanna, J.H., Saha, K., and Jaenisch, R. (2010). Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. Cell *143*, 508-525.

Hanover, J.A., Krause, M.W., and Love, D.C. (2010). The hexosamine signaling pathway: O-GlcNAc cycling in feast or famine. Biochim Biophys Acta *1800*, 80-95.

Hart, G.W., and Copeland, R.J. (2010). Glycomics hits the big time. Cell *143*, 672-676.

Hart, G.W., Housley, M.P., and Slawson, C. (2007). Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins. Nature *446*, 1017-1022.

Heng, J.C., Feng, B., Han, J., Jiang, J., Kraus, P., Ng, J.H., Orlov, Y.L., Huss, M., Yang, L., Lufkin, T., *et al.* (2010). The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. Cell Stem Cell *6*, 167-174.

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

Hu, P., Shimoji, S., and Hart, G.W. (2010). Site-specific interplay between O-GlcNAcylation and phosphorylation in cellular regulation. FEBS Lett *584*, 2526-2538.

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

Jang, H., Choi, S.Y., Cho, E.J., and Youn, H.D. (2009). Cabin1 restrains p53 activity on chromatin. Nat Struct Mol Biol *16*, 910-915.

Jang, H., Kim, T.W., Yoon, S., Choi, S.Y., Kang, T.W., Kim, S.Y., Kwon, Y.W., Cho, E.J., and Youn, H.D. (2012). O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network. Cell Stem Cell *11*, 62-74.

Keller, A., Nesvizhskii, A.I., Kolker, E. & Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383-5392.

Kim, H.S., Park, S.Y., Choi, Y.R., Kang, J.G., Joo, H.J., Moon, W.K., and Cho, J.W. (2009). Excessive O-GlcNAcylation of proteins suppresses spontaneous cardiogenesis in ES cells.

FEBS Lett *583*, 2474-2478.

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. Nat Struct Mol Biol *12*, 423-428.

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

Khidekel, N., Ficarro, S.B., Peters, E.C., and Hsieh-Wilson, L.C. (2004). Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain. Proc Natl Acad Sci U S A *101*, 13132-13137.

Lee, J., Go, Y., Kang, I., Han, Y.M., and Kim, J. (2010). Oct-4 controls cell-cycle progression of embryonic stem cells. Biochem J *426*, 171-181.

Landeira, D., Sauer, S., Poot, R., Dvorkina, M., Mazzarella, L., Jorgensen, H.F., Pereira, C.F., Leleu, M., Piccolo, F.M., Spivakov, M., *et al.* (2010). Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators. Nat Cell Biol *12*, 618-624.

Lee, J., Rhee, B.K., Bae, G.Y., Han, Y.M., and Kim, J. (2005). Stimulation of Oct-4 activity by Ewing's sarcoma protein. Stem Cells *23*, 738-751.

Li, G., Margueron, R., Ku, M., Chambon, P., Bernstein, B.E., and Reinberg, D. (2010). Jarid2 and PRC2, partners in regulating gene expression. Genes & development *24*, 368-380.

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

Love, D.C., Krause, M.W., and Hanover, J.A. (2010). O-GlcNAc cycling: emerging roles in development and epigenetics. Semin Cell Dev Biol *21*, 646-654.

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

Mochizuki, H., Ohnuki, Y., and Kurosawa, H. (2011). Effect of glucose concentration during embryoid body (EB) formation from mouse embryonic stem cells on EB growth and cell differentiation. J Biosci Bioeng *111*, 92-97.

Myers, S.A., Panning, B., and Burlingame, A.L. (2011).

Polycomb repressive complex 2 is necessary for the normal site-specific O-GlcNAc distribution in mouse embryonic stem cells. Proc Natl Acad Sci U S A *108*, 9490-9495.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol *26*, 101-106.

Nesvizhskii, A.I., Keller, A., Kolker, E. & Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* 75, 4646-4658.

Ng, H.H., and Surani, M.A. (2011). The transcriptional and signalling networks of pluripotency. Nat Cell Biol *13*, 490-496.

Niwa, H., Masui, S., Chambers, I., Smith, A.G., and Miyazaki, J. (2002). Phenotypic complementation establishes requirements for specific POU domain and generic transactivation function of Oct-3/4 in embryonic stem cells. Mol Cell Biol *22*, 1526-1536.

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., Yamamura, K., and Miyazaki, J. (1991). Efficient

selection for high-expression transfectants with a novel eukaryotic vector. Gene *108*, 193-199.

Okita, K., and Yamanaka, S. (2010). Induction of pluripotency by defined factors. Experimental Cell Research *316*, 2565-2570.

Pannell, D., and Ellis, J. (2001). Silencing of gene expression: implications for design of retrovirus vectors. Rev Med Virol *11*, 205-217.

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.

Pasini, D., Cloos, P.A., Walfridsson, J., Olsson, L., Bukowski, J.P., Johansen, J.V., Bak, M., Tommerup, N., Rappsilber, J., and Helin, K. (2010). JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. Nature *464*, 306-310.

Patel, M., and Yang, S. (2010). Advances in reprogramming somatic cells to induced pluripotent stem cells. Stem Cell Rev *6*, 367-380.

Portela, A. & Esteller, M. (2010).Epigenetic modifications and human disease. *Nature biotechnology* **28**, 1057-1068, doi:

10. 1038/nbt.1685

Quinn, P. (1995). Enhanced results in mouse and human embryo culture using a modified human tubal fluid medium lacking glucose and phosphate. J Assist Reprod Genet *12*, 97-105.

Rigbolt, K.T., Prokhorova, T.A., Akimov, V., Henningsen, J., Johansen, P.T., Kratchmarova, I., Kassem, M., Mann, M., Olsen, J.V., and Blagoev, B. (2011). System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation. Sci Signal *4*, rs3.

Sauvageau, M., and Sauvageau, G. (2010). Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. Cell Stem Cell *7*, 299-313.

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

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.

Saxe, J.P., Tomilin, A., Scholer, H.R., Plath, K., and Huang, J. (2009). Post-translational regulation of Oct4 transcriptional activity. PLoS One *4*, e4467.

Shafi, R., Iyer, S.P., Ellies, L.G., O'Donnell, N., Marek, K.W., Chui, D., Hart, G.W., and Marth, J.D. (2000). The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. Proc Natl Acad Sci U S A *97*, 5735-5739.

Sharov, A.A., Masui, S., Sharova, L.V., Piao, Y., Aiba, K., Matoba, R., Xin, L., Niwa, H., and Ko, M.S. (2008). Identification of Pou5f1, Sox2, and Nanog downstream target genes with statistical confidence by applying a novel algorithm to time course microarray and genome-wide chromatin immunoprecipitation data. BMC Genomics *9*, 269.

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

Simon, J.A., and Kingston, R.E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol *10*, 697-708.

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

Stadtfeld, M., and Hochedlinger, K. (2010). Induced pluripotency: history, mechanisms, and applications. Genes Dev *24*, 2239-2263.

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.

Sutton-McDowall, M.L., Mitchell, M., Cetica, P., Dalvit, G., Pantaleon, M., Lane, M., Gilchrist, R.B., and Thompson, J.G. (2006). Glucosamine supplementation during in vitro maturation inhibits subsequent embryo development: possible role of the hexosamine pathway as a regulator of developmental competence. Biol Reprod *74*, 881-888.

Swaney, D.L., Wenger, C.D., Thomson, J.A., and Coon, J.J. (2009). Human embryonic stem cell phosphoproteome revealed by electron transfer dissociation tandem mass spectrometry. Proc Natl Acad Sci U S A *106*, 995-1000.

Tarleton, H.P., and Lemischka, I.R. (2010). Delayed differentiation in embryonic stem cells and mesodermal

progenitors in the absence of CtBP2. Mech Dev *127*, 107–119.

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.

Toleman, C., Paterson, A.J., Shin, R., and Kudlow, J.E. (2006). Streptozotocin inhibits O-GlcNAcase via the production of a transition state analog. Biochem Biophys Res Commun *340*, 526–534.

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.

Webster, D.M., Teo, C.F., Sun, Y., Wloga, D., Gay, S., Klonowski, K.D., Wells, L., and Dougan, S.T. (2009). O-GlcNAc modifications regulate cell survival and epiboly during zebrafish development. BMC Dev Biol *9*, 28.

Whyte, W.A., Bilodeau, S., Orlando, D.A., Hoke, H.A., Frampton, G.M., Foster, C.T., Cowley, S.M., and Young, R.A. (2012). Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. Nature *482*, 221–225.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307-319.

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 Res *23*, 352-364.

Zeidan, Q., and Hart, G.W. (2010). The intersections between O-GlcNAcylation and phosphorylation: implications for multiple signaling pathways. J Cell Sci *123*, 13-22.

Zhao, L.J., Subramanian, T., Zhou, Y., and Chinnadurai, G. (2006). Acetylation by p300 regulates nuclear localization and function of the transcriptional corepressor CtBP2. The Journal of biological chemistry *281*, 4183-4189.

Zhang, Q., Piston, D.W and Goodman, R.H. (2002). Regulation of corepressor function by nuclear NADH.Science295,1895-1897

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Ⅵ. Abstract in Korea

줄기세포 및 유도만능줄기세포(iPSC)는 소비자 맞춤형 줄기세포로써의 임상적 적용 가능성, 특이질병 및 환자맞춤형 약물 스크리닝을 할 수 있는 점등으로 인하여 각광을 받고 있다. 이러한 줄기세포는 같은 유전체 배경을 가짐에도 불구하고 여러 다른 종류의 조직들로 분화하는 능력 (전분화능)을 가지고 있기 때문이다. 하지만 유도만능줄기세포로부터 임상적 치료 시 그 목적과는 다르게 다른 조직 등으로 분화가 진행되어 암과 같은 치명적 손실을 가져올 수 있는 단점 또한 가지고 있다. 이런 다른 종류의 조직들로 바뀔 수 있는 능력은 유전정보 상위의 epigenetic modification (후성유전학적인 변화)에 의해 조절된다고 알려져 있고, 정상적이지 않은 후성유전학적인 변화요인들 또한 각종 질병이나 암 등을 유발한다고 알려져 있다. 따라서 실질적인 치료 형 줄기세포 개발을 위해서는 줄기세포의 후성유전학적인 변화 메커니즘 이해는 반드시 선행되어야 하는 문제이다.

첫째로, 본 연구자는 *O*-GlcNAc modification (오글루넥화)가 줄기세포의 유지 및 분화 그리고 역분화에 어떤 영향을 주는지 관찰해 보았다. 왜냐하면 오글루넥화는 여러 다른 일반세포들에 중요한 역할을 한다고 알려져 있었지만 줄기세포 및 전분화능에 대한 역할은 밝혀져

있지 않았다. 줄기세포에서 오글루넥화를 막으면 자가분열 및 역분화가 잘 되지 않는 것을 관찰하였다. 또한 줄기세포 유지에 중요한 전사인자인 Oct4 와 Sox2 가 오글루넥화가 되고 줄기세포가 분화할 시 빠르게 오글루넥화가 없어지는 것을 관찰하였다. 결국 Oct4 의 228 Threonine 의 오글루넥화가 Oct4 의 전사인자 활성 및 줄기세포 전분화능에 관련된 유전자군 Klf2, Klf5, Nr5a2, Tbx3, Tcl1 의 활성을 조절하는데 중요한 역할을 하는 것을 밝혔다. Oct4 의 228 Threonine 에 오글루넥화가 되지 않는 mutant Oct4(Oct4 T228A)는 줄기세포의 자가 분열 및 역분화를 시키지 못하는 것 역시 관찰되었고, 결국 줄기세포 핵심인자의 오글루넥화가 줄기세포 전분화능에 중요한 역할을 하는 것을 제시하였다.

두 번째로, 본 연구자는 줄기세포 분화 시 줄기세포에서 특이적으로 많이 발현되는 주요 유전자 군의 안정적인 억제에 관련된 메커니즘 연구를 하였다. 전분화능은 주요 전사인자의 포지티브 피드백 (positive feedback) 네트워크에 의해 조절되고 있으며, 다른 조직이나 세포 등으로 분화가 되기 위해서는, 후성유전학에 의해 줄기세포 주요 전사인자들의 발현 억제가 필요하지만 줄기세포 분화 시 안정적인 억제 후성유전학 마커인 H3K27me3 이 어떡해 주요 전사인자의 크로마틴에 접근하여 발현을 억제하는지에 대해서는 모르는 상황이었다. 따라서 본 연구자는 줄기세포 분화 시 Ctbp2 유전자가 Polycomb repressive complex 2(PRC2)를 매개로 하여 주요 줄기

세포 인자의 억제에 관련된 것을 밝혀 냈다. shRNA 를 통하여 Ctbp2 유전자 발현을 낮춘 줄기세포에서는 줄기세포 분화 시 주요 줄기세포 유전자 군의 발현이 억제되지 못하였고, 결국 줄기세포의 전분화능이 유지되는 것을 관찰하였다. 줄기세포에서 Ctbp2 는 PRC2 와 Oct4 와 결합하는 것을 관찰하였고, 줄기세포 분화 시 Ctbp2 가 주요 줄기세포 유전자 군으로의 Ezh2 의 크로마틴 접근성을 조절하고 결국 억제 후성유전학 마커인 H3K27me3 의 크로마틴 접근성을 조절하는 것을 밝혀냈다. 이는 결국 Ctbp2 가 올바른 줄기세포 분화에 중요한 역할을 하는 것을 제시하였다.

주요어: 줄기세포, 오글루넥화(O-GlcNAcylation), 역분화, Ctbp2, 전분화능, 주요 전사인자, PRC2, 유전자 억제, 유도만능줄기세포