이학박사학위논문

배반포 분화 과정에서 PHF6의 기능 연구

Functional studies of PHF6 in trophectoderm differentiation

2020년 2월

서울대학교 대학원

생명과학부

오 성 룡

Functional studies of PHF6 in trophectoderm differentiation

by

Sungryong Oh

Advisor

Professor Sung Hee Baek, Ph.D.

A Thesis for the Degree of Doctor of Philosophy

February, 2020

School of Biological Sciences

Seoul National University

ABSTRACT

Sungryong Oh

School of Biological Sciences

The Graduate School

Seoul National University

Epigenetic regulation is important for lineage determination during early development in mammal. In early development, expression of specific transcription factors plays an important role in differentiation of cells into different lineages, such as endoderm, mesoderm, ectoderm, and trophectoderm. For lineage specification, histone modification patterns are dramatically reiwired on the lineage-specific genes to regulate their expressions. To date, transcription factor network guiding to trophectoderm determination has been studied. However, the mechanism by which epigenetic state is reprogrammed on the trophectodermal genes was poorly

understood. Here, I identified *Plant Homeodomain Finger 6* (*Phf6*) is crucial for transcriptional activation of several genes including *caudal homeodomain transcription factor 2* (*Cdx2*), during trophectoderm differentiation. Importantly, PHF6 possesses intrinsic E3 ubiquitin ligase activity on histone H2BK120 for transcriptional activation. PHF6 recognizes H2BK12 acetylation via the extended PHD2 (ePHD2) domain of PHF6, and H2BK12 acetylation concomitantly stimulates H2BK120 mono-ubiquitination by the ePHD1 domain of PHF6 *in vitro*. Furthermore, the acetylation-dependent E3 ubiquitin ligase activity of PHF6 plays a key epigenetic regulatory function for expressions of trophectodermal genes. Together, my data provide evidence that H2BK12ac and H2K120ub linkage via PHF6 is crucial for trophectoderm differentiation through activating gene expression.

Key words

Plant Homeodomain finger 6 (*Phf6*), trophectoderm, lineage, Histone H2B K12 Acetylation (H2BK12Ac), Histone H2B K120 mono-ubiquitination (H2BK120ub), E3 Ubiquitin ligase, *caudal homeodomain transcription factor 2* (*Cdx2*)

Student Number: 2011-20343

CONTENTS

LIST OF FIGURES AND TABLES

of H2BK120ub for *Cdx2* expression during trophectoderm differentiation

Table I. Several transcription factors known to have a function to trophectoderm differentiation

CHAPTER I

Introduction

I-1. Mammalian embryogenesis

1.1 Blastocyst formation in pre- and post-implantation stage

The first step in the early development of mammals is when the sperm and the ovulated egg merge in the fallopian tube to form a fertilized egg, which then successfully attaches on the maternal uterus along the fallopian tube. At this time, the fertilized egg undergoes a very rapid division, and these divided cells form a non-uniform distribution of the fertilized egg before implantation, called a blastocyst (Beddington and Robertson, 1999; Hamilton and Laing, 1946; Lu et al., 2001; Rossant and Tam, 2009) (Fig I-1). The blastocyst consists of three major lineages: epiblast (EPI) in Inner Cell Mass (ICM), primitive endoderm (PE), and trophectoderm (TE). Each isolated cells are referred to as Embryonic Stem cells (ESCs) derived from ICM populations (Evans and Kaufman, 1981; Martin, 1981), Xen cells derived from PE (Kunath et al., 2005), and Trophoblast Stem cells (TSCs) derived from TE (Tanaka et al., 1998).

In the post-implantation stage, they differentiate into cells that perform different functions. ESCs in ICM population and Xen cells in PE are differentiated into the three germ layer of endoderm, mesoderm, and ectoderm at the inner area of blastocyst together, and are destined to further develop into functional organisms and individuals.

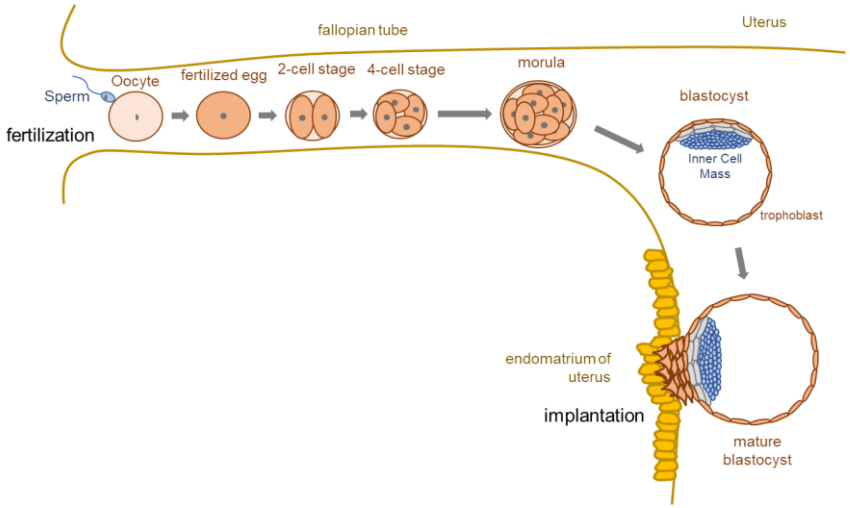


Figure I- 1. Illustration of the stages from fertilization to implantation

Fertilized eggs is divided rapidly to form blastocyst through morula stages. At the blastocyst stage, divided cells are distinguished at three types: Inner Cell Mass (ICM), Xen cells, and trophoblast cells. For proper implantation on the endometrium of mother's uterus, differentiation of trophoblast cells is an essential excess.

1.2 Gene expression regulations in early development

In mammalian embryogenesis, the differentiation of cells into their respective lineages requires very sophisticated and dynamic gene expression control process. Some genes play an important role throughout embryogenesis, but there are also genes that play a specific lineage dependent role at a particular time, thus defining them as lineage-marker genes or lineage-specification factors (Braude et al., 1988; Prioleau et al., 1994; Schultz et al., 1999). At pre-implantation stage, it has been identified that isolated cells from ICM populations, called embryonic stem cells (ESCs), have an ability to generate almost every organisms, a potential called pluripotency. In these cells, several transcription factors like *POU class 5 Homeobox 1* (*Pou5f1*, also known as *Oct3/4*), *SRY-Box Transcription factor 2* (*Sox2*), *Kruppel-like factor 4* (*Klf4*) and *Myc,* have been identified to maintain pluripotency and repress expressions of differentiation marker genes (Takahashi and Yamanaka, 2006).

When blastocysts are successfully implanted on the uterine wall,

different cells of blastocyst begin to differentiate into their own lineages. Cell fates of each lineage have been known to be determined by cell-cell polar axis and signaling pathways from local environment (Beddington and Robertson, 1999; Lu et al., 2001; Rossant and Tam, 2009). Once cell fates regarding which lineage of 3 germ layers the cell will differentiate has been determined, lineage-specific transcription factors begin to express to form organs of interest through differentiation.

Epigenetic controls, especially histone modifications, are one of the best known regulatory system in embryonic stem cell maintenance and differentiation. In pre-implantation stages of blastocyst, promoters of certain key lineage-specification genes are regulated by both transcription activation histone markers and transcription repression histone markers, an epigenetic state called the bivalent state (Bernstein et al., 2006; Hu et al., 2013; Jiang et al., 2011). After successful implantation of blastocysts, repressive histone markers are removed, and transcription of these lineage-specific genes are fully activated (Cantone and Fisher, 2013; Karlic et al., 2010; Meissner, 2010).

1.3 Trophectoderm differentiation and placenta formation

Among cell types of blastocysts, trophoblast, located outside the blastocyst, not only plays an important role in proper implantation of blastocyst on the uterine wall, but also continue to differentiation after the implantation, eventually forming the placenta (Cross et al., 1994; Lu et al., 2001) (Fig I-2).

Trophectoderm lineage is the first determined cell fate in embryogenesis. To date, one of the most important factor for this lineage determination is the cellular polarity generated by uneven distribution of cells in the transition from morula to blastocyst. This asymmetric division results in selective activation of Hippo pathway only at the inner cell populations (Cockburn et al., 2013; Hirate et al., 2013; Nishioka et al., 2009). In addition, activation of Notch signaling pathway only in the trophoblast stem cells have been known to be important for trophectoderm specification and maintenance (Rayon et al., 2014).

Majority of studies were performed on animal models to identify key factors of trophectoderm differentiation. Among them, *Caudal Type*

Homeobox 2 (*Cdx2*) is the most important transcription factor for trophectoderm differentiation. It was first reported that *Cdx2* is expressed only in the trophecotderm lineages at E3.5 days pre-implantation blastocysts (Beck et al., 1995). In addition, it was also reported that the lack of *Cdx2* resulted in the formation of immature blastocyst (Strumpf et al., 2005). Not only in mice but also in other mammalian blastocysts, *Cdx2* expression and importance toward trophectoderm specification and maintenance was reported (Bou et al., 2017; Sritanaudomchai et al., 2009). Based on these studies, *Cdx2* is regarded as a master factor for trophectoderm differentiation in mammalian blastocysts to date.

The upstream factors for *Cdx2* expression are also important for trophectoderm differentiation. Moreover, it was reported that *Cdx2* is not sufficient for initiation of trophectoderm lineage specification (Wu et al., 2010), suggesting that and other factors for trophectoderm differentiation are needed. *TEA domain transcription factor 4* (*Tead4*) has been reported as an upstream factor of *Cdx2* and regulator of initiation of trophectoderm lineage specification (Nishioka et al., 2009; Yagi et al., 2007). *Tead4* is

expressed in pre-implantation blastocyst, and functions as a transcription factor of *Cdx2* in enhancer region along with Yap protein.

Gata3 is another critical factor for trophectoderm differentiation. (Ralston et al., 2010). *Gata3* is also identified as a downstream factor of *Tead4,* and is important for expression of trophoblast maturation factors in TSCs. However, *Gata3* working mechanism is independent of *Cdx2*, suggesting that *GATA3* works in parallel with *Cdx2* in trophectoderm differentiation. Not only *GATA3*, but also *T-box gene Eomesodermin* (*Eomes*), *Transcription factor AP-2 gamma* (*Tcfap2c*) were known as important transcription factors for early stage of trophectodem differentiation (Kuckenberg et al., 2010; Russ et al., 2000). Moreover, several other transcription factors are known to regulate gene expressions for placenta formation, (Table 1).

It has been also reported that suppression of pluripotency marker genes in trophoblasts are also important for maintaining trophectoderm properties. In blastocyst, it was reported that trophectoderm marker *Cdx2* and pluripotency marker *Oct4* suppressed the expression of each other

during developmental lineages (Niwa et al., 2005; Yeap et al., 2009). In addition, *Sox2* was also known as a critical factor to form trophectoderm or cavitate in morula stage of embryos (Keramari et al., 2010; Li et al., 2007). These reports suggest that trophoblast cells and ICM cells in the blastocyst are not completely independent, and continue to interact with each other as the organism develop.

Unlike histone modifications during the differentiation of 3 germ layers, histone modifications of trophectoderm genes have not been studied well. It has been only reported that the repressive epigenetic state for suppressing pluripotency genes exist in trophectoderm. The repression markers of histones like Histone H3 tri-methylation of lysine 9 (H3K9me3), are regulated in order to suppress pluripotency genes like *Oct4* in trophectoderm differentiation (Lohmann et al., 2010; Wang et al., 2018; Yeap et al., 2009). However, histone modifications for transcription activation in trophectoderm genes have not been established yet.

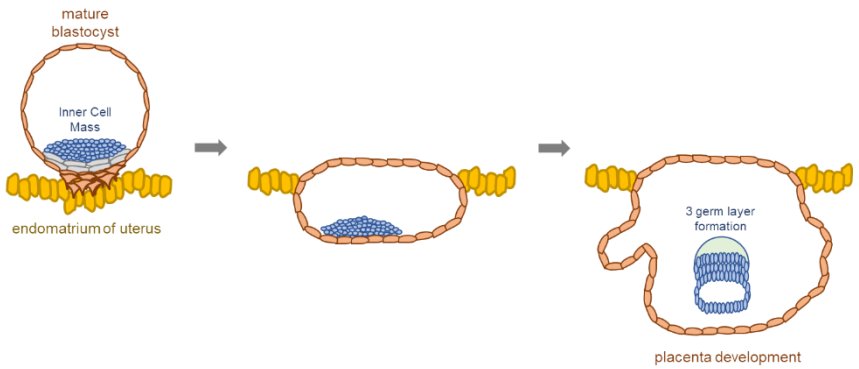


Figure I- 2. Illustration of the differentiation of blastocysts after implantation.

Just after proper implantation, trophoblast cells of blastocyst start integration in the maternal endometrium of uterus, and differentiate into several types of cells for forming placenta. Inner Cell Mass populations

Gene	Full name	Function
ASCL2(MASH2)	Achaete-scute family BHLH Transcription factor 2	Spongiotrophoblast development
HAND1	Heart And Neural Crest Derivatives Expressed 1	trophoblast cells differnetiation
DLX3	Distal-Less Homeobox 3	placental vascularization
GCM1	Glial Cells Missing Transcription Factor 1	syncytiotrophoblast and labyrinth formation
CEBPα	CCAAT Enhancer Binding Protein Alpha	labyrinth formation
PRDM1	PR/SET Domain 1	endovascular trophoblast giant cells differentiation
OVOL2	Ovo Like Zinc Finger 2	Labrynth formation and placental vascularization
CITED1	Cbp/P300 Interacting Trans-activator With Glu/Asp Rich Carboxy-Terminal Domain 1	spongiotrophoblast differnetiation
CITED2	Cbp/P300 Interacting Trans-activator With Glu/Asp Rich Carboxy-Terminal Domain 2	spongiotrophoblast, invasive trophoblast differnetiation
PPARγ	Peroxisome Proliferator Activated Receptor Gamma	labyrinth and giant cell layer differentiation
FOSL1	FOS Like 1, AP-1 Transcription Factor Subunit	Labrynth formation and placental vascularization

Table 1. Several transcription factors known to have a function to trophectoderm differentiation

I-2. Histone modification recognizing factors

2.1 Histone modifications and histone modifying enzymes

In eukaryotic cells, genomic DNAs is wound around histone octamers composed of four types of histone proteins – H2A, H2B, H3, and H4 -, called nucleosomes. Because of this structure, transcription process is closely related to nucleosome states. The N-terminal region of the histone protein exists in a linear form, which is called the histone tails. Several residues in the tails undergo post-translational modifications (PTMs) (Fig I-3). Transcription regulation by these epigenetic modifications is reported abundantly in various biological systems (Heintzman et al., 2009; Karlic et al., 2010; Wang et al., 2008) including in developmental stages (Bernstein et al., 2006; Gan et al., 2007; Grewal and Moazed, 2003; Meissner, 2010; Mikkelsen et al., 2007).

Several histone modifications like methylation, acetylation, phosphorylation, and ubiquitination, are well established (Fig I-3). For

regulation of histone modifications, it is important for the enzymatic complexes, called writer and eraser, to modulate the histone states. For example, H3K4me3, the best known form of histone methylation involved in transcription activation, is specifically distributed on the Transcription Start Site (TSS) of genes with transcriptional activation (Wang et al., 2008). The mixed-lineage leukemia protein (MLL) complex is the writer enzyme that attaches tri-methylation to H3K4 residues. During ESC differentiation, MLL complex is involved in the regulation of H3K4me3 of lineage-specific genes (Denissov et al., 2014; Gregory et al., 2007; Jiang et al., 2011; Wu et al., 2013).

Histone acetylation on lysine residues of histone H3 and H4 are another largely studied regarding histone modification and transcription. It is well known that acetylation on histone residues contributes to the formation of open chromatin and induces transcription activation (Grunstein, 1997). Many Histone Acetyl Transferases (HATs) including CBP/p300 are identified, and considered as a sufficient factors for transcriptional regulation (Martinez-Balbas et al., 1998; Struhl, 1998).

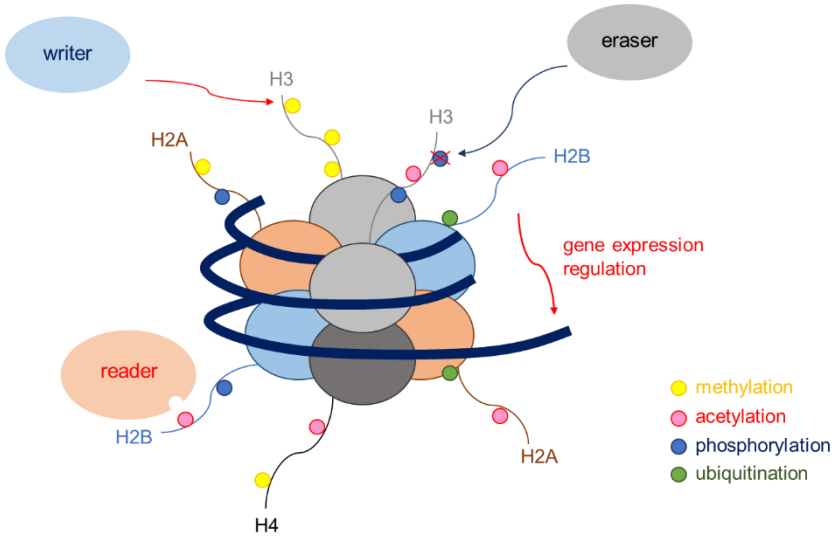


Figure I- 3. Histone modifications and writer, eraser, reader

Illustration of histone modifications and modifying enzymes. Nucleosomes are composed of histone octamer. Each histones have N-terminal tails and several modifications are exsited in these regions for regulating transcription. Several enzymes work for regulation of histone modifications, called writers and eraser. Several proteins, called readers, recognize these modifications and recruit several other factors.

2.2 Readers of Histone Modifications

For regulation of histone modifications, it is important for the enzymatic complexes, called writer and eraser, to modulate the histone states. However, other factors are needed to explain how these regulated histone modifications actually affect gene expression. Many studies have identified epigenetic modifiers, also known as histone modification reader, that regulate the transcription process by recognizing histone modifications. By functioning through recognition of histone modifications, reader proteins not only recruit transcription activation factors or transcription repression factors (Candau et al., 1997; Musselman et al., 2012; Yang et al., 2008), but also other histone modifying enzymes and cause sequential changes on modification of other residues, which is called histone modifications crosstalk (Taverna et al., 2006) (Kleine-Kohlbrecher et al., 2010).

The domains for recognizing specific histone modifications have been well established. For example, bromo-domain is the most well-known reader domain for histone acetylation (LeRoy et al., 2008). Not only

several HATs or HDACs have bromo-domains (Delvecchio et al., 2013; Dhalluin et al., 1999), but also several chromatin remodeling factor or transcription mediators have bromo-domains (Yang et al., 2008), suggesting that many transcription machinery operate through acetylation recognition. For recognizing histone methylations, there are several reader domains reported including chromo-, WD40, Tudor, and PHD domain (Musselman et al., 2012). Like bromo-domains of acetylation reading, these methylation reader-containing proteins have ability to recruit other factors or have an intrinsic enzymatic activity through its separate protein domains.

2.3. PHD domain

Among methylation reading domains, Plant Homeodomain(PHD)-fingers domain are known as an H3K4me2/3 reading domain (Sanchez and Zhou, 2011), thereby worked as a transcription coactivator. PHD domain belongs to the Zinc-binding superfamily. Tandem PHD domain catches two Zinc ions by conserved cysteine and histidine residues. Some different types of PHD domain recognizes non-methylation of H3 lysine 4(Lan et al., 2007; Org et al., 2008) or acetylation of H3 (Zeng et al., 2010).

PHD domain is very similar in structure to another Zinc finger family, the Really New Gene (RING) domain (Capili et al., 2001; Matthews et al., 2009). However, RING domain has an E3 ligase activity, whereas PHD domain has a methylation reading activity. Traditionally, the distinction between PHD domain and RING domain was found in the order of cysteine-histidine conserved residues. Some recent studies, however, have reported cases where the PHD domain has an E3 ligase activity (Ivanov et al., 2007; Lu et al., 2002).

I-3. Plant Homeodomain-finger family 6 (PHF6)

3.1. Structure of PHF6

Among PHD domain containing genes, PHF6 is a protein-coding gene that contains two PHD domains (Fig 3A). The PHF6 gene is located on X chromosome, and the two PHD domains of PHF6 are extended form (ePHD). Compared to typical PHD domain, the extended PHD domains have additional half of the PHD domains, so that they capture three zinc ions (Fig I-4).

3.2. PHF6 in Leukemia

The role of PHF6 is well studied in the leukemia system. In leukemia, several *phf6* somatic mutations have been reported, and it has been considered that functional defects by PHF6 mutation as a leukemia driving factor (Van Vlierberghe et al., 2010; Van Vlierberghe et al., 2011). Other study reported that PHF6 interacts with Nucleosome Remodeling

Deacetylase (NuRD) complex (Todd and Picketts, 2012). It has also been reported the interaction between PHF6 and PAF1 transcription elongation complex (Zhang et al., 2013). These studies suggest that the function of PHF6 is similar to the function of other typical reader activities. However, it was reported that the ePHD2 domain of PHF6 binds with double-stranded DNA in sequence-independent manner, not with Histones, *in vitro* (Liu et al., 2014). However, later studies of the function of PHF6 in leukemia show results that suggest that PHF6 will still function as a histone modification reader (Meacham et al., 2015; Soto-Feliciano et al., 2017). Nevertheless, still the precise molecular function of PHF6 has not been known yet.

3.3. PHF6 in Börjeson-Forssman-Lehmann (BFL) syndrome

In addition to leukemia, it has been reported that somatic mutations of PHF6 is a critical factor Börjeson-Forssman-Lehmann (BFL) syndrome (Berland et al., 2011; Crawford et al., 2006; Zweier et al., 2013). Several somatic mutations of PHF6 in BFL syndrome have been identified, and

some studies based on genetics proved these mutations as a critical driving factor of BFL syndrome.

Typical symptoms of BFL syndrome include severe intellectual disability, epilepsy, hypogonadism, obesity, several characteristic craniofacial features, microcephaly, dental abnormalities, abnormal shapes of fingers, toes, and ears. These symptoms appear to be typical problems that takes place at perinatal developmental stages, strongly suggesting that BFL syndrome patients may have some early stages of fetal development.

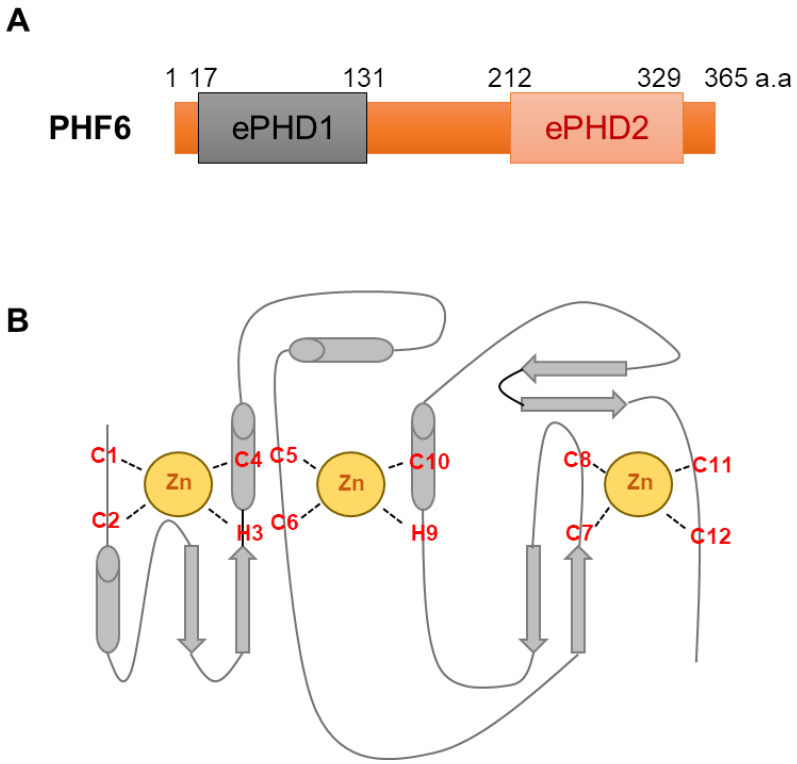


Figure I- 4 Illustrations of PHF6 structure

(A) Schematics of PHF6 protein. PHF6 has two extended PHD domains. (B) Modeling of extended PHD domain structure. Extended PHD domain has additional half-PHD domain, results in capturing 3 Zinc ions.

CHAPTER II

PHF6 regulates trophectoderm differentiation as a transcription regulator

II-1. Summary

In mammalian early development, precise blastocyst formation is an important step in affecting the development of post-implantation. With the symptoms of BFL syndrome, to which *Phf6* mutation is a critical factor, PHF6 can be inferred as a factor that regulates early stages of development. In the mouse blastocyst at embryonic day 3.5~4.0 pre-implantation stage, PHF6 was well expressed in both the Inner Cell Mass (ICM) population and the trophoblast, which are leading to differentiate into fetal organisms and placenta each.

In order to confirm the effect of PHF6 in detail, I constructed *Phf6*-knockout (KO) cells from ZHBTc4 mouse embryonic stem cells (mESCs) line. As a result, *Phf6* had no effect on maintaining the properties of embryonic stem cells. However, when differentiating wild-type (WT) and *Phf6*-KO mESCs by embryoid bodies formation (EB formation), I found that insufficient differentiation of overall lineages, especially trophectoderm, is occurred in *Phf6*-KO EBs. In addition, when driving

mESCs into trophectoderm by removing *Oct4*, I also confirmed that the expressions of the overall trophectoderm genes were significantly decreased when *Phf6* is absent. Taken together, I concluded that PHF6 is important for trophectoderm lineage by regulating gene expressions in early development stages.

II-2. Introduction

Trophectoderm differentiation is a differentiation process that forms the placenta after implantation and plays an important role in the embryonic development of mammals (Cross et al., 1994). To preserve the cell fate, epigenetic regulators that govern histone modification function as critical players in the maintenance of transcriptional memory. However, histone modifications for transcription activation of trophectoderm lineage have not been established yet, whereas histone modifications for the 3 germ layer lineage specification from embryonic stem cell is well established.

At the pre-implantation stage of blastocysts, transcription repression mechanism between ICM and trophoblasts are also important for maintaining their specific cell fates. It is well known that *Oct4* and *Cdx2*, master transcription factors in ICM and trophoblast respectively, inhibit each other's expression in their respective lineages (Nishioka et al., 2009; Niwa et al., 2005). According to this regulatory mechanism, it was also

well known that ESCs were differentiated into early trophectoderm lineage by *Oct4* depletion (Guo et al., 2008; Hay et al., 2004; Whyte et al., 2012).

Mutations in *Phf6* have been found to induce Börjeson-Forssman-Lehmann (BFL) syndrome by several genetics studies (Berland et al., 2011; Crawford et al., 2006; Zweier et al., 2013). Patients with this syndrome show several symptoms like irregular teeth, intellectual disability, and hypogonadism, which are typical features involving defects during early development. Moreover, according to recent report about *Phf6* C99F-knockin (KI) mice, which is the model mice for BFL syndrome by *Phf6* mutations known in BFL syndrome, KI mice have been shown to live well but have only brain defects. However they also reported homozygote *Phf6*-knockout (KO) mice, and unlike to *Phf6* KI mice, the mics have been shown to exhibit perinatal lethality (Cheng et al., 2018). This difference between *Phf6* KO and KI mice also strongly indicate that *Phf6* plays an important role during early development, not only limited on neural development.

II-3. Results

PHF6 is highly expressed in ICM and trophoblasts in pre-implantation stage blastocysts

From the symptoms of patients with BFL syndrome, I first examined the possibility that *Phf6* plays an important role during the early developmental stages. To confirm the function of *Phf6*, I first confirmed whether *Phf6* was expressed during early development. To confirm this, I extracted blastocysts from mating female mice at the pre-implantation stages. Mice are known to take 21 days from fertilization to childbirth, and it is well known that blastocysts form 3.5 days after mating in pre-implantation state. As a result of whole-mount immunofluorescence staining with PHF6 antibody in the E3.5 blastocyst, I confirmed that PHF6 is expressed in inner cell mass and trophoblast in blastocyst of pre-implantation stage (Fig II-1). Through this, I could confirm the possibility that PHF6 could function in the early differentiation stage where cell fate is determined.

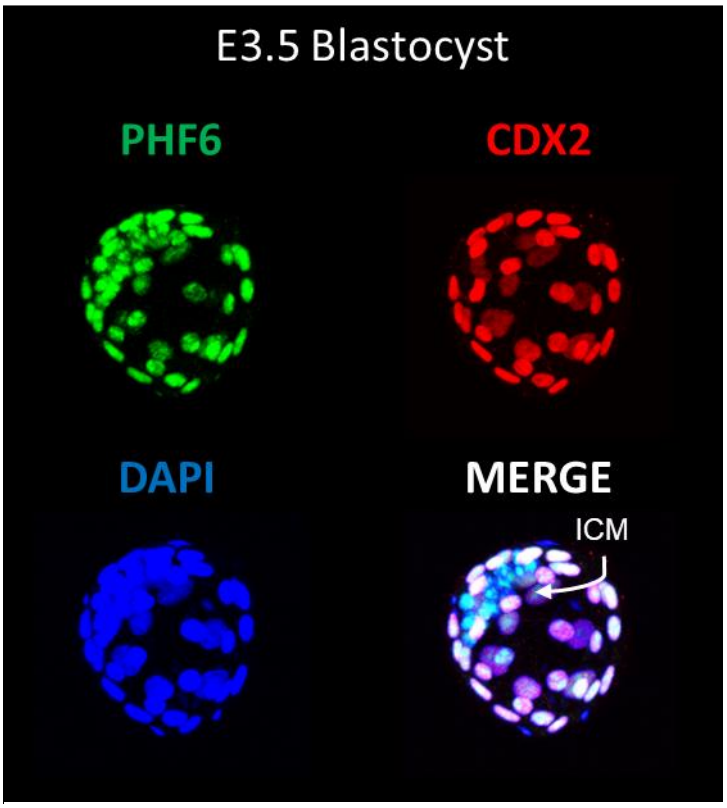


Figure II- 1. PHF6 levels in pre-implantation stage blastocyst

Whole-mount immunofluorescence staining of PHF6 in pre-implantation stage blastocyst. 3D-imaging was obtained by confocal microscopy. 3D-rendering process was performed by Imaris x64 software.

Generation of Phf6 Knockout (KO) mouse embryonic stem cells (mESCs)

From the expression pattern of *Phf6*, I generated *Phf6* KO in ZHBTc4 cell lines to study the function of PHF6 at blastocyst. The ZHBTc4 line is engineered to allow *Oct4* inducible depletion by tetracycline or doxycycline treatment, and then differentiate into trophectoderm lineage (Guo et al., 2008; Hay et al., 2004). This gave me the advantage of simultaneously identifying the role of *Phf6* both in ESCs and in trophectoderm.

Generation of *Phf6* KO mESCs is performed using CRISPR-Cas9 system. Using GPP sgRNA Designer (CRISPRko) (Doench et al., 2016), I designed sgRNA which targets genome sequence at Exon4 of *Phf6* (Fig II-2A). After Cas9 generation, I confirmed the generation of *Phf6* KO ZHBTc4 cell lines by western blotting analysis (Fig II-2B). Sanger sequencing of targeted Exon4 showed that *Phf6* KO ZHBTc4s, which I made, are added 5 nucleotides at the 100th amino acid histidine-encoding genome sequence and resulted in triplet code broken (Fig II-2C).

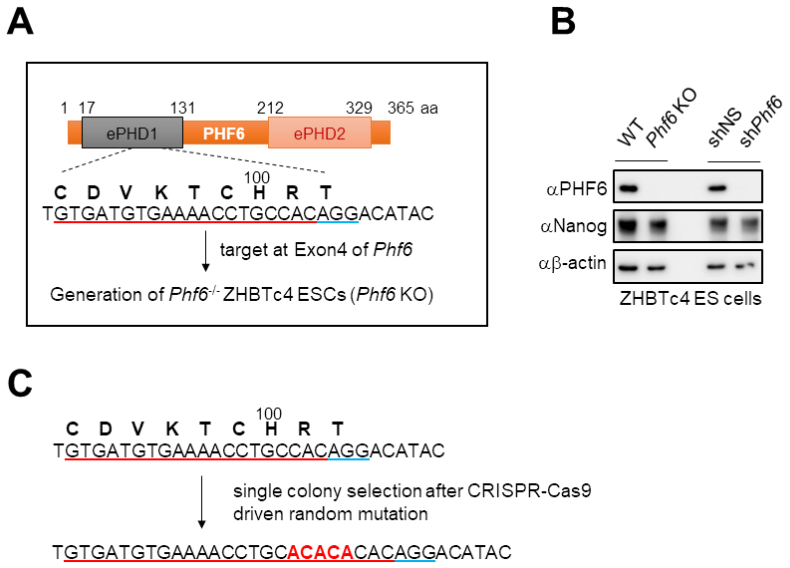


Figure II- 2. Generation of *Phf6* KO mouse embryonic stem cell lines

(**A**) Schematics of generation of *Phf6* KO ZHBTc4 cell lines. sgRNA was designed using GPP sgRNA designer. (**B**) Western blot analysis for validation of *Phf6* KO ZHBTc4. Knockdown (KD) of PHF6 is control for Phf6 KO validation. (**C**) Sanger sequencing for genome sequence change of *Phf6* KO ZHBTc4.

Phf6 Knockout dose not affect the maintenance of mouse embryonic stem cells

Using this generated *Phf6* KO ZHBTc4, I utilized Embryoid Bodies (EBs) as a model of *in vitro* differentiation from ESCs. PHF6 mRNA and protein levels were decreased during *in vitro* differentiation after EB formation (Fig II-3A, B), suggesting that PHF6 has a potential role in embryonic stem cell maintenance. Since the features of embryonic stem cells (ESCs) are self-renewal and pluripotency, I first measured the growth rates of WT and *Phf6* KO ESCs for checking the effects of PHF6 at self-renewal of ESCs. As a result, the cellular growth of wild type (WT) and *Phf6* KO ESCs did not differ significantly from each other (Fig II-3C, upper), suggesting that PHF6 dose not affect self-renewal ability of ESCs. Further, I detected alkaline phosphatase (AP) activity, which is highly activated in undifferentiated ESCs but decreases during ESC differentiation, one of the golden-standard method for measuring pluripotency (Jiang et al., 2008; ten Berge et al., 2011; Wobus et al., 1984). Both WT and *Phf6* KO ESCs showed comparable AP staining without

loss of ESC morphological features, indicating that PHF6 deficiency did not affect ESC pluripotency (Fig II-3D, upper). Consistent with the results of *Phf6* KO ESCs, the ESC pluripotency and growth properties of PHF6 shRNA knockdown ESCs did not significantly differ from those of control shRNA knockdown ESCs (Fig II-3C, D, lower). The expression levels of pluripotency genes including *Oct4*, *Tbx3*, and *Klf5* were comparable between WT and *Phf6* KO ESCs (Fig II-3E), indicating that PHF6 is not responsible for ESC maintenance.

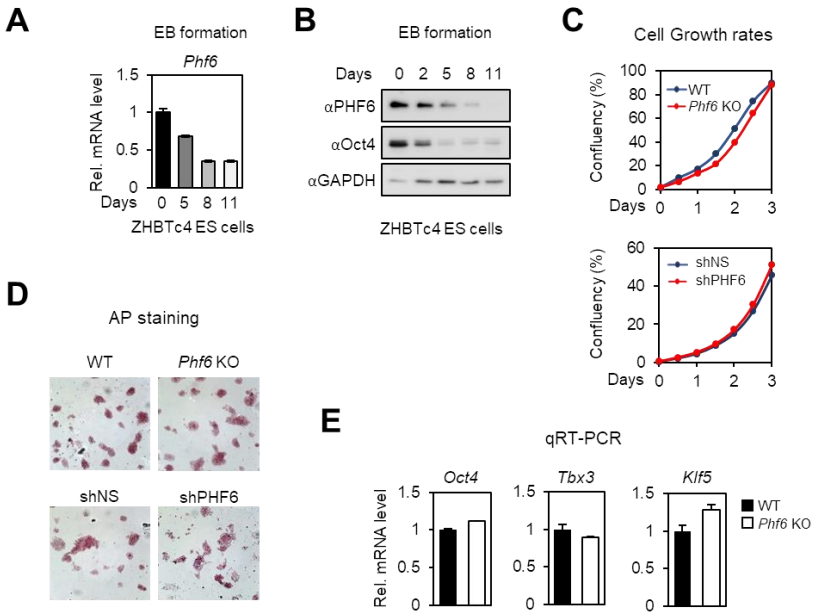


Figure II- 3. PHF6 does not affect the maintenance of embryonic stem cell

(**A**) Quantatitive Real-time PCR (qRT-PCR) analysis for checking mRNA level of Phf6 according to differentiate by Embryoid Bodies (EB) formation. (**B**) Western blot analysis of Phf6 in EB formation. (**C**) Cell proliferation assay for measuring cell growth rates between WT and *Phf6* deficient ZHBTc4. Cell confluencies were measured by photographing every 12 hours using JuLI Stage Live Cell Imaging. (**D**) AP staining for measuring pluripotency between WT and *Phf6* deficient ZHBTc4. (**E**) qRT-PCR analysis of pluripotency genes between WT and *Phf6* KO ZHBTc4.

Phf6 deficiency impairs trophectoderm differentiation in Embryoid Bodies formation

Since *Phf6* did not affect the maintenance of embryonic stem cells, I looked at how the disruption of PHF6 affects the differentiation stage. For checking the role of *Phf6* at differentiation states, I followed the formation of EBs with WT and *Phf6* KO ZHBTc4 by adherent culture condition, which is mimic of blastocyst outgrowth assay. Comparison of EBs derived from *Phf6* KO and WT ESCs showed that *Phf6* KO EBs failed to form outer cell layer (FigII-4A). In the EB formation process, these outer cell layers are known to be cell populations with trophectoderm properties in other species, especially in human induced pluripotent stem cells (iPSCs) (Gerami-Naini et al., 2004; Xu et al., 2017). Moreover, immunostaining of WT EBs with an Cdx2 antibody, a trophectoderm marker, showed that *Cdx2* is highly expressed in the outer cell layer of the EBs comprised trophoblast cells (FigII-4B). Intriguingly, PHF6 expression overlapped with *Cdx2* expression in the outer cell layers, including migratory zone (FigII-4C). Further, I compared the expression

level of each lineage marker genes between WT and *Phf6* KO EBs. As expected, the mRNA levels of trophectodermal genes including *Cdx2, Plac1, Ascl2, and Gata2* were not sufficiently expressed in *Phf6* KO EBs those in WT EBs (FigII-4D). Thus, I could confirm that *Phf6* loss impaired trophectoderm lineage differentiation from ESCs. Moreover, it was also observed that expressions of 3 germ layer marker genes are decreased in *Phf6* KO EBs, except mesodermal genes (FigII-4E, F, G). These data suggest that PHF6 has an important function to regulated mammalian early development.

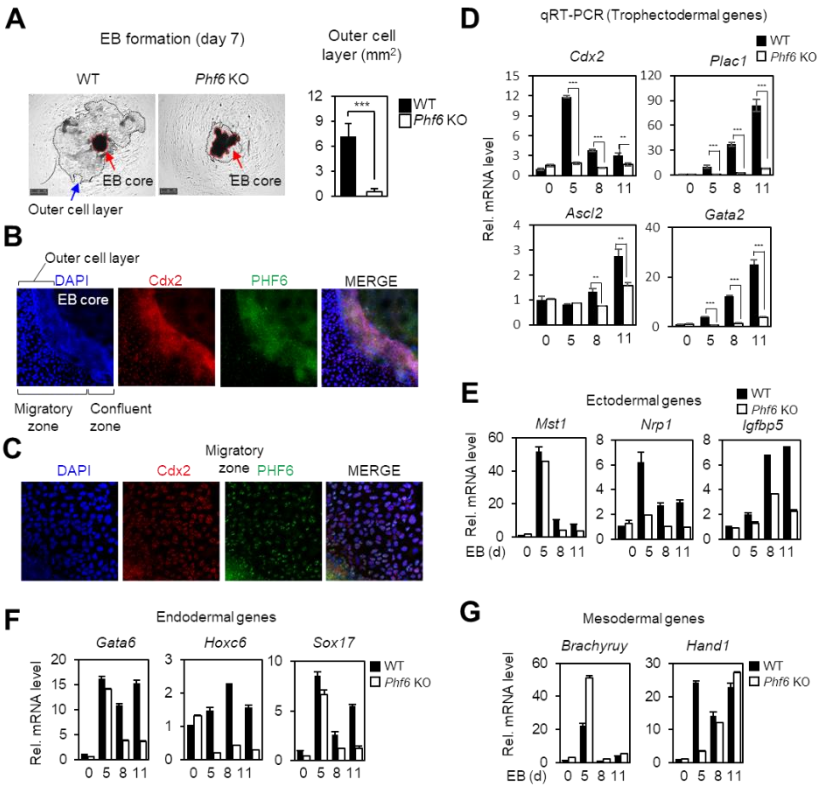


Figure II- 4. *Phf6* KO shows defects on Embryoid Bodies differentiation

(**A**) Growth of EB out-layer cells between WT and *Phf6* KO EBs. Area were calculated by ImageJ software. (**B-C**). Immunoflourescence staining of EB core and out-layer cells between WT and *Phf6* KO EBs. Images are taken using Zeiss microscopy. (**D-G**). qRT-PCR analysis of mRNA levels of each lineage-specific marker genes between WT and *Phf6* KO EBs.

PHF6 directly regulates trophectoderm differentiation in transcriptional level

Since *Phf6* KO EBs show insufficient expressions of differentiation genes of overall lineages, I checked if *Phf6* is actually involved in the overall differentiation situation. For ESC differentiation into specific lineage, I treated Doxycycline (Dox) and Retinoic acid (RA) in WT and *Phf6* KO ZHBTc4 cell lines, which result in mESCs differentiated into trophectoderm and neural ectoderm, respectively. ZHBTc4 cell line is an ESC line that has been modified to be tetracycline or doxycycline-inducible *Oct4* depletion, resulted in early trophectoderm differentiation (Hay et al., 2004). In the case of retinoic acid, it is well known to induce ESCs differentiated into neural ectoderm lineage (Okada et al., 2004; Rochette-Egly, 2015). After treating these two chemicals with WT and *Phf6* KO ZHBTc4, I performed mRNA sequencing and compared differentiated expressed genes (DEGs) of samples by fold change. As a result, I found that the number of genes which are not sufficiently expressed in *Phf6* KO ESCs were significantly higher in Dox treatment

than RA treatment (FigII-5A). In addition, when analyzing the biological processes of signal-dependent up-regulated genes, I found that two biological processes - development and transcription – showed significant *Phf6* dependency just when trophectoderm differentiation is induced by Dox treatment (FigII-5B). qRT-PCR analysis also confirmed that the trophoectodermal genes are regulated in *Phf6*-dependent manner (FigII-5C). These comparisons suggests that PHF6 only affects directly in trophectoderm differentiation by transcriptional regulatory function.

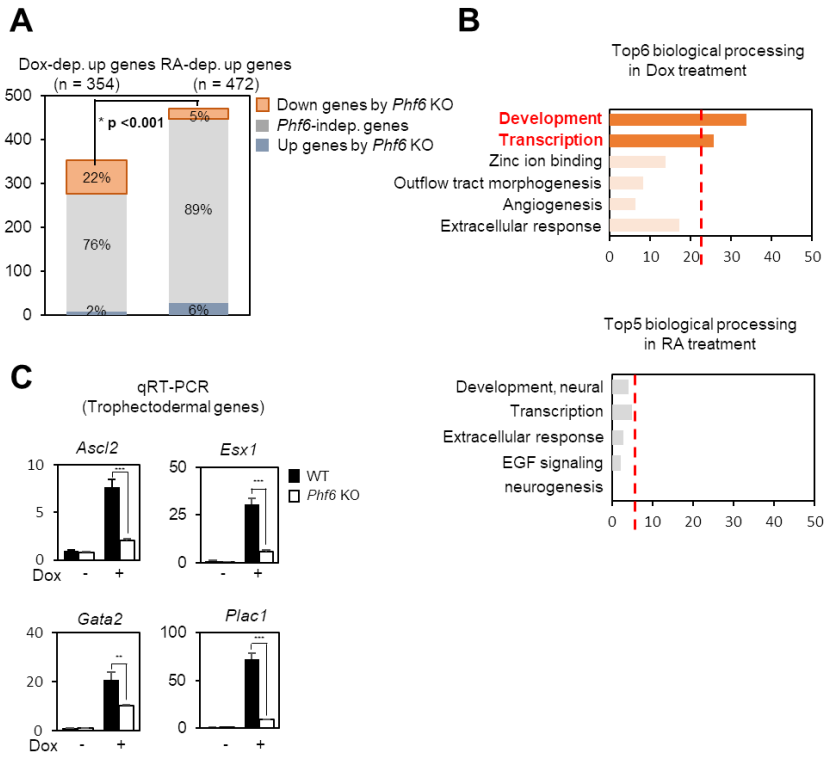


Figure II- 5. PHF6 directly regulates trophectoderm differentiation.

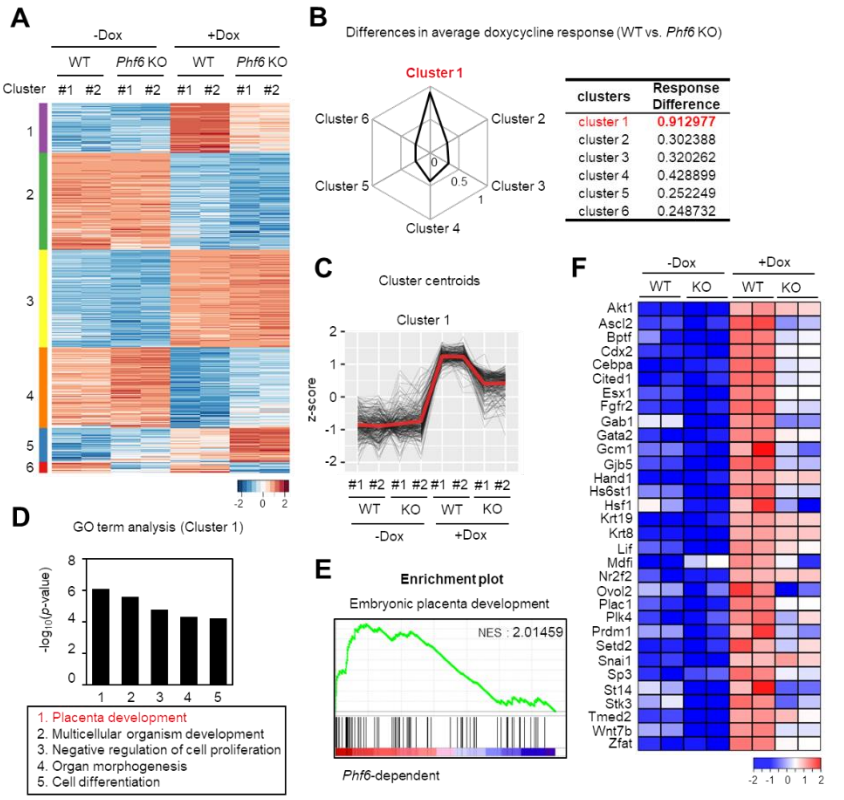
(**A**) Comparison of PHF6 effect on Differentiated Expressed Genes (DEGs) between Doxycycline (Dox) and Retinoic acid (RA). DEGs were sorted by Fold Changes (cutoff : |log2(FC)| > 1.5). (**B**) Overall biological processes of PHF6-dependently expressed genes. The red dotted line represents the overall ratio and cutoff. (**C**) qRT-PCR of trophectoderm genes between WT and *Phf6* KO ESCs upon Dox treatment. *p*-values were calculated by one-way ANOVA test. (* : $p < 0.05$, ** : $p < 0.01$, ***: $p < 0.001$).

RNA-sequencing analysis reveals PHF6 as a transcriptional activator for trophectoderm differentiation of ESCs

To identify *Phf6*-dependent gene clusters during trophectoderm differentiation, I performed unsupervised *k* -means clustering (k = 6) and compared gene expression changes in each cluster between WT and *Phf6* KO ESCs upon Dox treatment (FigII-6A). In each cluster, I first measured the differences in Dox response between WT and *Phf6* KO ESCs. As a result, cluster 1 showed the strongest differential Dox response between WT and *Phf6* KO ESCs (FigII-6B), suggesting that regulation of gene expression within cluster 1 is highly dependent on PHF6. Gene expression within cluster 1 exhibited little or no difference between WT and *Phf6* KO in the ESC states (-Dox) but considerable decrease in *Phf6* KO ESCs after induction during differentiation states (+Dox) (FigII-6C). The most significant term of Gene Ontology (GO) for cluster 1 was placenta development that is biologically associated with trophectoderm differentiation (FigII-6D), indicating that PHF6 positively regulates the transcriptional activation of trophectodermal genes during

differentiation.

To identify the subset of genes whose expression depends on *Phf6* during differentiation, Gene Set Enrichment Analysis (GSEA) was performed by correlating with PHF6 dependency in cluster 1. GSEA revealed that the gene sets for embryonic placenta development was significantly enriched in cluster 1 (FigII-6E). From the gene sets, PHF6-dependent genes included *Ascl2*, *Esx1*, *Gata2*, and *Plac1* that are crucial for placenta development (FigII-6F). Together, the transcriptome profiles indicate that PHF6 directly regulates the expression of genes involved in trophectoderm lineage differentiation.



(legend on the next page)

Figure II- 6. Trophectoderm marker genes are transcriptionally regulated by PHF6

(**A**) Heatmap of *k* -means clustering of variably expressed genes (n = 1775, *k* = 6). Genes were grouped into six clusters on the basis of expression similarity. (**B**) Differences in Dox response for WT and *Phf6* KO ESCs. Dox response is calculated using average of z-score differences of genes for Dox treatment in ESCs. (**C**) Z-score centroids of Cluster 1 genes. The center of the cluster is highlighted as red. (**D**) Functional Gene Ontology (GO) analysis of Cluster 1. Placenta development was the most enriched terms in Cluster 1. (**E**) Representative gene set enrichment analysis (GSEA) of DEGs. The pattern of media gene expression from centroids in Figure C was used for standards of significance. (**F**) Genes that were up- and down-regulated in *Phf6* KO ESCs in comparison with WT ESCs. The colour bar represents the gradient of log2-fold-changes in each comparison. Positively correlated genes from GSEA were selected.

II-4. Discussions

In this study, I confirmed the existence of PHF6 expression in early embryonic development, and generated *Phf6*-Knockout ZHBTc4 mouse embryonic stem cell lines. There was no information on the expression of *Phf6* at the stages before the formation of fetal organs and the lineage specification had just begun. With the symptoms of BFL syndrome, my results strongly suggest that *Phf6* may have an important role in this early development stage.

Using Embryoid Bodies (EB) formation *in vitro* differentiation methods from embryonic stem cells, I found that *Phf6* plays an important role in blastocyst differentiation, especially in trophectoderm differentiation. EB formation is the best method of mimic *in vivo* differentiation among the differentiation methods in cultured systems, with differentiation of all 3 germ layers. However, it has not been reported that EBs from mouse embryonic stem cells have a function to form trophoblast cells. When I tried EB differentiation by mimicking the

in vitro blastocyst outgrowth assay, I could find that EBs from mouse embryonic stem cells also have the ability to form the cell out-layers around the EB core, which is proved as trophectoderm lineage cells. However, EB formation from *Phf6* KO ESCs shows defects in the formation of this trophectoderm outer cell layers, suggesting that *Phf6* is important for trophectoderm differentiation.

In 2018, Cheng et, al. reported *Phf6* C99F knock-in (KI) mice as a model of BFL syndrome with *Phf6* KO mice (Cheng et al., 2018). In this report, *Phf6* C99F mice show brain abnormalities and problems with neural development, and KO mice were analyzed in the same direction. However, as they also mentioned, the KI mice did not show all the various defects of BFL syndrome, suggesting that the role of *Phf6* is not limited to neuronal differentiation. In addition, from the phenotype difference that KI mice lived well while KO mice showed perinatal lethality, this report also suggested that the role of *Phf6* in early development is not limited to neural development. When I compared the transcriptome differences between WT and *Phf6* KO ESCs by chemical-

induced lineage specification, I could not find any significant effects of *Phf6* on neural lineage specification by retinoic acid treatment. Moreover, when comparing the expressions of neural-ectodermal lineage marker genes between WT and *Phf6* KO ESCs by RA treatment, most neural marker genes were expressed well in both (data not shown). These findings suggest that *Phf6* has no direct function to neural differentiation, at least in early lineage specification stages. However, with my observation that the expression of neural marker genes was not sufficiently increased in *Phf6* KO EBs, it can be speculated that the problem of trophectoderm differentiation and placenta formation may have an omni-directional effect on the overall growth process of the fetus. In short, at least when cell lineage is determined, *Phf6* has a direct function as a transcription activator in trophectoderm differentiation, and defects in trophectoderm differentiation affect precise differentiation of other lineages.

II-5. Materials and Methods

Mice

For isolation of blastocysts at embryonic day 3.5, C57BL/6 wild-type mice at 10-15 weeks-old ages were used for mating and sacrifice. Mice were housed in a specific pathogen-free Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC)-accredited facility. Mice mating was checked by varginal plug check. Mice sacrifice was proceeded by cervical dislocation. All animal studies and procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of National Cancer Center Research Institute.

Cell culture and transfection

A conditional Oct4-depleted (ZHBTc4) ES cell line was described previously39. Briefly, in the first 1~2 passages, mESCs were maintained on mouse primary embryonic fibroblast feeders. After stabilization, mESCs were cultured under feeder-free conditions, using 0.1% gelatin-

coated culture dishes. mESCs were maintained in Dulbecco's modified Eagle medium (DMEM; Hyclone), supplemented with 15 % fetal bovine serum (FBS; Hyclone), 0.055 mM β-mercaptoethanol, 2 mM L-glutamine, 0.1 mM nonessential amino acid, 5,000 units/ml of penicillin/ streptomycin (GIBCO), and 1,000 units/ml of leukemia inhibitory factor (LIF) (Chemicon) in a humidified incubator at 37□ with 5 % CO2. HEK293T cell lines were cultured at 37□ in Dulbecco's modified Eagle's medium (DMEM) containing 10% FBS and antibiotics in a humidified incubator with 5 % CO2. All cell lines used in the study were regularly tested for mycoplasma contamination. All transfection in mESCs was performed Lipofectamine 3000 (Invitrogen) according to the manufacturer's protocol. All transfection in H2K293T was performed with PEI.

Virus for shRNA-based knockdown

To generate shRNA knockdown cells, lentiviral shRNA constructs were transfected along with viral packaging plasmids (psPAX2 and pMD2.G)

into HEK293T cells. 1 day after transfection at 60~70% confluency of cells, transfected cells were changed new fresh media and maintained 24h for virus collecting. After vicus collecting, the media were filtered and mixed with Lenti-X collector (#631232, TAKARA), and incubated O/N at 4□. These concentrated viruses were infected into cells with polybrene.

Generation of *Phf6* KO mESCs

CRISPR-Cas9 system was used for generation of *Phf6* KO mESCs. sgRNA designs for *Phf6* targeting were performed from GPP sgRNA Designer (CRISPRko) (Doench et al., 2016). Selected sgRNAs were cloned into the pRGEN-U6 vector, and transfected into ZHBTc4. After single colony selection with puromycin treatment, obtained *Phf6* KO colonies were confirmed first by western blotting and then by Sanger-sequencing for identifying frame-shift mutations.

Antibodies and reagents

The following commercially available antibodies were used: anti-PHF6 (68262) antibody was purchased from Novus; anti-Nanog (ab21624) and anti-Cdx2 (ab157524) antibodies were purchased from Abcam; anti-Oct4 (sc-5279) and anti-GAPDH (sc-25778) antibodies were purchased from Santa Cruz Biotechnology; anti-β-actin (A1978) antibody was purchased from Sigma.

Western blotting

All concentrated cells were lysed in the lysis buffer (50mM Tris-HCl pH 8.0, 200mM NaCl, 0.5% NP-40, protease inhibitors (Aprotinin and Leupeptin)) at 4□, 10 minutes. For sampling of western blot of histone modifications and enzymes at the same conditions, we performed sonication shortly of all lysed samples. After sonication and discarding debris, sample density was calculated by Bradford assay for normalization.

Embryoid Bodies (EB) formation

From ZHBTc4 mESCs, EBs were created by the hanging-drop method, starting with 1000cells/drop. After 2 days in the drop state, each drop begins to grow on a non-coated sterile cover glass for Immuno-fluorescence staining, and on a round-shaped 96-well plate for live-cell imaging. In all culture situations, media were fed through addition per day and plate movement was also fixed to a minimum for minimizing the physical impact on the EBs. All EBs were cultured using ESC media without LIF.

Alkaline Phosphatase (AP) Staining

For AP staining, WT and *Phf6*-depletion ZHBTc4 were cultured on the 0.1% gelatin-coated cover glass in 12-well plates. 1 104 cells were seeded per well. After 2 days, AP staining was performed using the Alkaline Phosphatase Detection Kit (#SCR004, Millipore). Staining was performed according to the provided protocol. Briefly, each cell was fixed by 4% paraformaldehyde for 1-2 min, and react with the solution (Naphthol/Fast Red Violet mix) 15-20 minutes at RT. After the reaction,

washout solutions and make slides for imaging.

Live Cell imaging

For tracking growth of ZHBTc4 mESCs and EBs, JuLI Stage real-time cell history recorder (NanoEnTeK) was used. For tracking mESCs growth, $1*10^4$ cells were seeded into 0.1% gelatin-coated 12-well culture dishes, and cell confluency was recorded every 12 hours. After the recording is complete, cell growth curves were analyzed by software provided by JuLI. In the case of EB formation, each drop was moved into each well of 96-well round-shaped dishes and cell morphology was recorded every 12 hours. Then, movies were produced by the provided software. To calculate the out-layer cell area, images at 7 days were calculated using the Image-J program.

Immuno-fluorescence Staining

EB spheroids for immunofluorescence microscopy were grown on coverslips prior to the experiment (1 EB spheroid per coverslip). EBs were fixed with 2% paraformaldehyde in PBS for 10 minutes and then were washed two times with DPBS at room temperature. Fixed cells were permeabilized with 0.5% Triton X-100 in PBS (PBS-T) for 5 min at room temperature. Blocking was performed with 10% FBS in 0.1% PBS-T for 30 min. For staining, cells were incubated with primary antibodies (PHF6 1:250, CDX2 1:50) for 4hr at room temperature, washed four times with 0.1% PBS-T, and incubated for 1hr with fluorescent labelled secondary antibodies (Invitrogen). Cells were washed and mounted by VECTASHIELD (H-1200, Vector Laboratories) with DAPI (Sigma). Fluorescence was visualized on a Zeiss LSM700 confocal microscope (Carl Zeiss).

Quantitative RT-PCR

Total RNAs were extracted using Trizol (Invitrogen) and reverse transcription was performed from 1~2μg total RNAs using the M-MLV cDNA Synthesis kit (Enzynomics). The abundance of mRNA was detected by an ABI prism 7500 system or BioRad CFX384 with SYBR TOPreal qPCR 2x PreMix (Enzynomics). The quantity of mRNA was calculated using ddCt method and *Gapdh* and *β-actin* were used as controls. All reactions were performed as triplicates. The following mouse primers were used:

Phf6 Forward : 5'-TGGCGGGAGAAGACGAAGA -3' ,

Phf6 Reverse : 5'- TCTGTAGGGAGCACCGACAAT -3' ,

β-actin Forward : 5'- TAGCCATCCAGGCTGTGCTG -3',

β-actin Reverse : 5'- CAGGATCTTCATGAGGTAGTC -3',

Oct4 Forward : 5'- AGAGGATCACCTTGGGGTACA -3',

Oct4 Reverse : 5'- CGAAGCGACAGATGGTGGTC -3',

Tbx3 Forward : 5'- AGAGGATCACCTTGGGGTACA -3',

Tbx3 Reverse : 5'- CCATTGCCAGTGTCTCGAAAAC -3'.

Klf5 Forward : 5'- CCGGAGACGATCTGAAACAC -3',

Klf5 Reverse : 5’- CAGATACTTCTCCATTTCACATCTTG -3’,

Gapdh Forward : 5’- CATGGCCTTCCGTGTTCCTA -3’,

Gapdh Reverse : 5’- CCTGCTTCACCACCTTCTTG -3’,

Cdx2 Forward : 5’- CAAGGACGTGAGCATGTATCC -3’,

Cdx2 Reverse : 5’- GTAACCACCGTAGTCCGGGTA -3’,

Plac1 Forward : 5’- CTTCAGCTACTCGGAGCAAAA -3’,

Plac1 Reverse : 5’- GTGAACATGATTGGGAGGGC -3’,

Ascl2 Forward : 5’- TCCTGGTGGACCTACCTGCTT -3’,

Ascl2 Reverse : 5’- AGGTCAGTCAGCACTTGGCATT-3’,

Gata2 Forward : 5’- CCTCCAGCTTCACCCCTAA -3’,

Gata2 Reverse : 5’- CAGAGAGGGGTGGCTGTG -3’,

Msx1 Forward : 5’- CCCCTTCTGGGAAGAAAAAG -3’,

Msx1 Reverse : 5’- GGGCTCATCTCTTGAAGCAC -3’,

Nrp1 Forward : 5’- GACAAATGTGGCGGGACCATA -3’,

Nrp1 Reverse : 5’- TGGATTAGCCATTCACACTTCTC -3’,

Igfbp5 Forward : 5’- TTCCCTCCAGGAGTTCAAAGC -3’,

Igfbp5 Reverse : 5’- GGTCACAGTTGGGCAGGTACA -3’,

Gata6 Forward : 5’- TTGCTCCGGTAACAGCAGTG-3’,

Gata6 Reverse : 5’- GTGGTCGCTTGTGTAGAAGGA-3’,

Sox17 Forward : 5’- GATGCGGGATACGCCAGTG -3’,

Sox17 Reverse : 5’- CCACCACCTCGCCTTTCAC -3’,

Hoxc6 Forward : 5’- CCGAGTTAGGTAGCGGTTGAA -3’,

Hoxc6 Reverse : 5’- CCTGGATGCAGCGAATGAAT -3’

Brachyruy Forward : 5’- GGGTATTCCCAATGGGGGTG-3’,

Brachuruy Reverse : 5’- CGGTGGTTCCTTAGAGCTGG -3’.

Hand1 Forward : 5’- CACCACCTACCACCGCAGTA-3’,

Hand1 Reverse : 5’- CCTTCTTGGGTCCTGAGCCTTT-3’

Esx1 Forward : 5’- TTGGAGGGAGCAGACTACCAG-3’,

Esx1 Reverse : 5’- CCAAAGTCGGAGTAGAAAGTTGT-3’.

Statistical analysis

All experiments were performed independently at least three times. Values are expressed as mean ± SEM. Significance was analyzed using ANOVA test. *p*-value of less than 0.05 was considered statistically

significant.

RNA-sequencing analysis

Total RNA was extracted from WT and KO cells, respectively, in the presence or absence of doxycycline. Then, the stranded mRNA-seq library was prepared following Illumina's TruSeq protocol. After the production of raw data by HiSeq platform, the reads were pre-processed to remove adaptors and bases with low quality by Trimmomatic (v0.36) (Bolger et al., 2014). Next, STAR (v2.5.3) was used to align the reads (Dobin et al., 2013), and Transcripts Per Million (TPM) was calculated with RSEM (v1.3.0) (Li and Dewey, 2011). Using read counts per gene, differentially expressed genes (DEG) were identified with DESeq2 (v1.18.1) for all 4 possible pair-wise comparisons between WT and KO under normal condition and doxycycline treatment (Love et al., 2014). k-means clustering was performed for the union of 4 sets of DEGs in R (v 3.4.3). For Gene Set Enrichment Analysis (GSEA), a phenotype label

was defined as 2 : 2 : 50 : 10 = WT - doxycycline : KO – doxycycline : WT + doxycycline : KO + doxycycline, and Pearson correlation coefficient per gene was used for ranking. Enrichment score was then computed for the gene sets in molecular signature database (MSigDB) v6.2. (Mootha et al., 2003; Subramanian et al., 2005).

To analyze the difference in doxycycline reactivity between WT and *Phf6* KO cells for each cluster, we first obtained the z-score mean of the genes in each cluster for each sample. Thereafter, the difference in the mean values according to the presence or absence of doxycycline in each of the WT and KO cells was determined, which was regarded as the doxycycline reactivity of WT and *Phf6* KO. Finally, the differences in doxycycline reactivity between WT and KO cells by clusters were calculated and considered to be doxycycline reactivity according to the presence or absence of *Phf6* per cluster.

Response Difference = |(∑(+Dox z-scores) – ∑(–Dox z-scores) in WT) /n - (∑(+Dox z-scores) – ∑(–Dox z-scores) in KO) /n |

CHAPTER III

PHF6 has a function as E3 ubiquitin ligase on H2BK120 in H2BK12Ac-dependent manner

III-1. Summary

PHF6 has two extended PHD domains. These domains of PHF6 have one additional pre-PHD domain compared to a typical PHD domain, and have a structure that captures three zinc ions. There are several reports about PHF6 as a transcription coactivator, but it has not been established what histone modification PHF6 recognizes.

As a result of screening for histone modifications recognized by PHF6, I confirmed that PHF6 recognizes H2B K12 acetylation (H2BK12Ac) *in vitro* through ePHD2 domain. Moreover, I also confirmed that the specific acidic-rich regions, especially the E223 glutamic acid residue, in ePHD2 domain of PHF6 is critical for H2BK12Ac recognition.

Between several histone modifications related to transcription regulation, I confirmed that global level of H2B K120 mono-ubiquitination (H2BK120ub) was significantly decreased in *Phf6 KO* mESCs during trophectoderm differentiation. By checking the H2BK120ub levels at the promoters of trophectodermal genes, I

confirmed that H2BK120ub level at the promoters was significantly increased in *Phf6*-dependent manner when mESCs were differentiated into trophectoderm. H3K4me3 level was also increased *Phf6*-dependently according to the level of H2BK120ub.

Combined with the molecular function of PHF6 as an H2BK12Ac reader, I confirmed that PHF6 has a H2BK12Ac-dependent E3 ligase activity for H2BK120 residue *in vitro*. In addition, I also confirmed that both functionally dead mutant of ePHD1 domain and H2BK12Ac unrecognizable mutant of ePHD2 domain were not able to ubiquitinate H2BK120 residue. These *in vitro* data suggest that PHF6 has an acetyl-dependent E3 ligase activity for H2BK120ub and both ePHD domains are important for the activity. Together, I concluded that PHF6 regulates expressions of trophectodermal genes through acetylation-dependent E3 ubiquitin ligase activity.

II-2. Introduction

Among PHD domain-containing family, PHF6 is a gene with two extended PHD domains, which has an additional half of PHD domain to capture three Zinc ions. The importance of PHF6 somatic mutations has been known not only in BFL syndrome but also in leukemia. Several studies independently reported the PHF6 interaction partners such as Nucleosome Remodeling Deacetylase (NuRD) complex in the nucleus (Todd and Picketts, 2012), RNA polymerase II associated Factor 1 (PAF1) complex (Zhang et al., 2013), or a rDNA transcriptional activation factor UBF (Wang et al., 2013) , but the precise function of PHF6 itself has not been reported yet. The crystal structure of ePHD2 domain of PHF6 was reported and by focusing on this domain, ePHD2 domain of PHF6 recognizes sequence-free dsDNAs *in vitro* while unable to recognize histone modifications (Liu et al., 2014). However, following researches of *Phf6* in leukemia reported that PHF6 has a promoter-specific binding affinity at transcriptionally activated genes and physically interact with

histones *in vivo* (Meacham et al., 2015; Soto-Feliciano et al., 2017). For this reason, the precise molecular function of *Phf6* is still controversial, but in the context, PHF6 is more likely to have a function as a reader of histone modifications, rather than nonspecific dsDNAs.

Among histone modifications, tri methylation on histone H3 lysine 4 residue (H3K4me3) is the best known histone marker for transcription activation in various systems, including stem cell differentiation (Chen et al., 2015; Shen et al., 2015). One of the major regulatory modulation preceding H3K4me3 is mono-ubiquitination of lysine K120 residue of H2B (H2BK120ub). In detail, this mono-ubiquitination is recognized by *Ash2l*, one of the core subunits of Mixed Lineage Leukemia (MLL) complex (Nakanishi et al., 2009; Wu et al., 2013), then recruits MLL complex on the existing promoter region.

Acetylation is one of the well-known modifications of transcription regulation. Many subsequent studies have shown that the acetylation readers play a crucial role in gathering several factors involved in transcription. Bromo-domain is one of the most famous reader of histone

acetylation (Musselman et al., 2012). Recently, not only bromo-domain but also Plant Homeodomain (PHD)-fingers domain was reported to have acetylation-recognition activity (Zeng et al., 2010). Many PHD domains have been known as a reader for histone methylation, but in certain case, PHD domain could also recognize acetylation through the certain acidic residues. In the case of H2B, there are several acetylation sites including H2BK12 residue. H2BK12Ac is known to be written by CBP/p300, and locate on the promoter regions of genes that are transcriptionally active (Wang et al., 2008). However, the precise mechanism describing how H2BK12Ac is involved in transcription activation is not studied yet.

III-3. Results

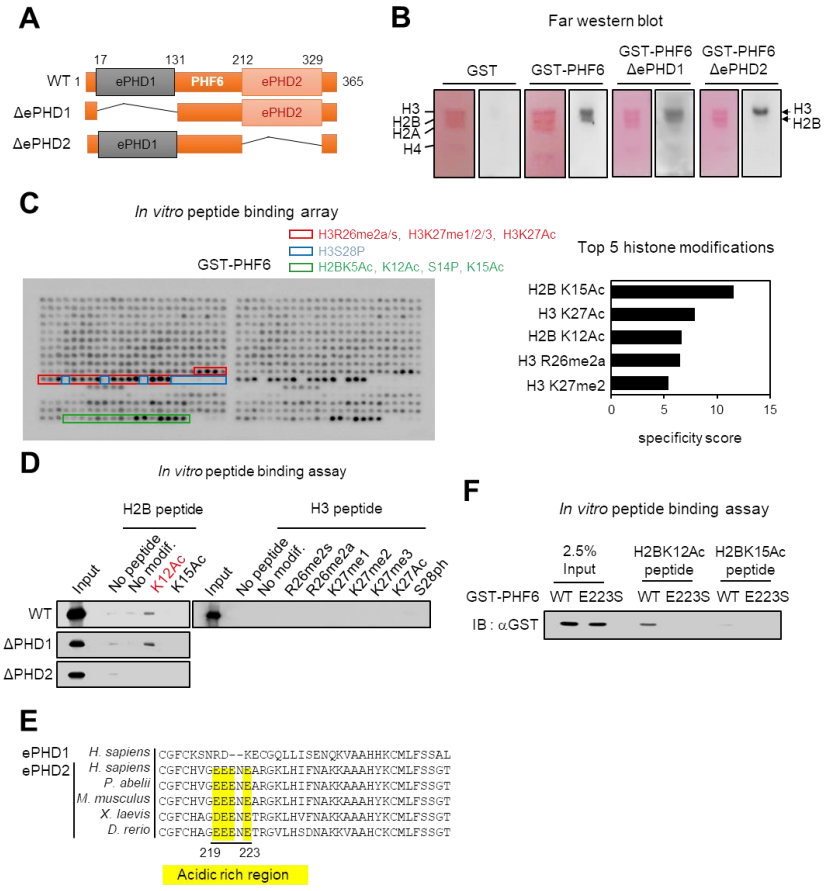
PHF6 recognizes H2BK12Ac through glutamic acid residues in ePHD2 domain

PHF6 has two highly conserved ePHDs, containing a C2H2 zinc finger and an imperfect PHD finger (FigIII-1A). To examine transcription activation by PHF6 via recognition of histone modification, I performed far-western blot analysis on histone extracts obtained from ZHBTc4 ESCs along with recombinant proteins including GST-tagged PHF6 WT (GST-PHF6 WT), to check the binding between PHF6 and histones. As a result, I could find that recombinant GST-PHF6 proteins selectively interact with histones H2B and H3. In addition, when performing same assay with deletion mutants for the ePHD1 or ePHD2 domain of PHF6, I found that ePHD2 deletion mutant of PHF6 selectively did not interact with Histone H2B (FigIII-1B). These data suggest that PHF6 has a function as a reader of any modification of histone H2B through the ePHD2 domain.

Next, I used an *in vitro* peptide binding array to identify specific modifications of H2B or H3 recognized by PHF6. Consistent with previous far-western immunoblot data, this assay confirmed the specific binding of PHF6 to acetylations of histones H2B and methylations or acetylation in R26-K27 residue of H3 (FigIII-1C, left). Comparing the significance, the top five histone modifications recognized by PHF6 were H2BK15Ac, H3K27Ac, H2BK12Ac, H3R26me2a, and H3K27me2 (FigIII-1C, right). Based on these results, I performed an *in vitro* peptide pulldown assay for each H2B and H3 modification to identify specific histone modification recognized by PHF6, and confirmed that PHF6 specifically binds to the acetylation of Histone H2BK12 residues (FigIII-1D). In addition, this specific interaction between PHF6 and H2BK12Ac peptide was dramatically disappeared only when ePHD2 domain is deleted (FigIII-1D), indicating that PHF6 recognizes H2BK12Ac through ePHD2 domain.

To clarify how the ability of PHF6 to read H2BK12Ac occurs only by its ePHD2 domain, I compared the amino acid sequences of the two PHD

domains of PHF6. Interestingly, the ePHD2 domain, but not ePHD1 domain, specifically contains a negatively charged amino acids-rich region (E219-E223) (FigIII-1E). According to previous case of the acetylation-recognizing PHD domain, it is important that the carbonyl oxygen of the negative charged amino acid makes a hydrogen bond with the acetyl amide of acetylation (Zeng et al., 2010). The crystal structure of the ePHD2 domain suggests that four glutamic acids (E219, E220, E221, and E223), especially E223, are possible to be key residues forming an acetylated substrate recognition motif (Liu et al., 2014). For this reason, I constructed a mutant without the negative charge of the E223 residue of PHF6 (PHF6 E223S) and measured the H2BK12Ac peptide binding affinity compaing with PHF6 WT. As a result, the E223S mutant resulted in significant loss of H2BK12Ac recognition *in vitro* (FigIII-1F). These data indicated that glutamic acid-rich motifs in PHF6 ePHD2 domain are crucial for H2BK12Ac reading activities.



(legend on the next page)

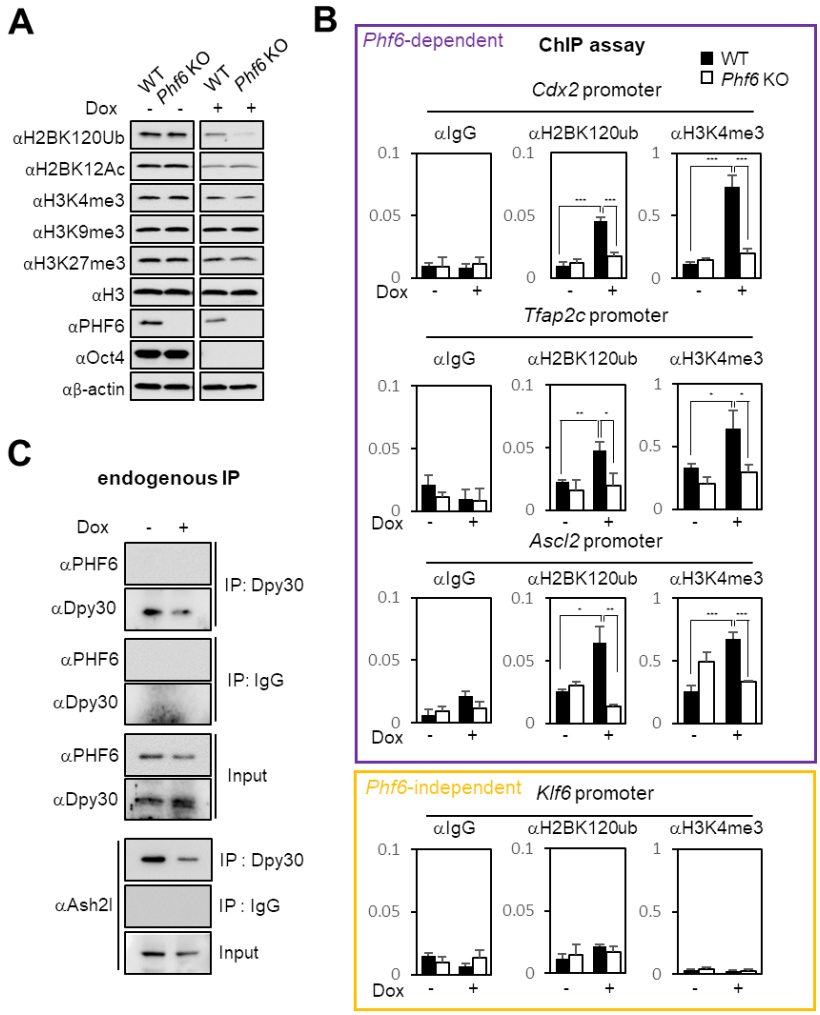
Figure III- 1. PHF6 recognizes H2BK12Ac via acidic amino acids in ePHD2 domain.

(A) Schematics of functional units in PHF6 and PHF6 deletion mutants. (B) Far western blot analysis of PHF6 WT and deletion mutants on histone extracts from ZHBTc4. 4 types of core histones were separated by molecular size in SDS-PAGE gel. (C) *in vitro* peptide binding array with GST-PHF6. Duplicated reactions were perfomed in kit. Array kit were purchased from ActiveMotif. Specificity of histone modifications were calculated by provided analysis tool. (D) *in vitro* histone peptide pulldown analysis with PHF6 WT and deletion mutants. Top five modified histone peptide candidates and nearby modifications were selected. (E) Amino acid sequences of the ePHD1 and ePHD2 domains of PHF6 orthologues in diverse species are aligned. The amino acids with a yellow background indicate the negatively charged region in the ePHD2 domain. The amino acid alignment was performed the ClustalX (F) *in vitro* peptide binding assay of PHF6 WT or E223S mutant with H2BK12Ac and K15Ac peptides.

PHF6 regulates H2BK120ub levels during trophectoderm differentiation

Like other PHD-containing proteins, acetylation reading activity of PHF6 through ePHD2 domain could subsequently regulate different epigenetic changes by recruiting other enzymes, or having enzymatic activity through other domain within itself. Therefore, I examined various histone modifications like H2BK120ub, H3K4me3, H3K27me3, and H3K9me3, known to regulate transcription activation or repression during ESC differentiation, along with H2BK12Ac in WT and *Phf6* KO ESCs. Surprisingly, only the global level of H2BK120ub, transcriptional activation markers in stem cell differentiation and other systems, was significantly lower in *Phf6* KO ESCs than WT upon Dox treatment (FigIII-2A). H2BK120ub is known as a preceding modification of H3K4me3 by recruiting MLL H3K4me3 methyl-transferase complex at the promoter of transcribed activated genes (Wu et al., 2013). When I checked H2BK120ub and H3K4me3 levels at the promoters of several doxycycline-dependent transcribed genes by chromatin immuno-

precipitation, I found that H2BK120ub level and H3K4me3 level of PHF6 target genes are significantly decreased at their promoter regions (FigIII-2B). Moreover, I found that PHF6 dose not interact with DPY30, one of MLL complex subunit, during trophectoderm differentiation. (FigIII-2C). These data suggest that PHF6 regulates transcriptional activation of trophectoderm marker genes by increasing H2BK120ub level during trophectoderm differentiation.



(legend on the next page)

Figure III- 2. PHF6 regulates H2BK120ub levels during trophectoderm differentiation

(A) Global levels of several histone modifications in undifferentiated and Dox-induced differentiated WT and *Phf6* KO ESCs **(B)** ChIP assay at the promoters of PHF6 target genes (*Cdx2, Ascl2, Tfap2c*) and negative control gene (*Klf6*) with H2BK120ub and H3K4me3 antibodies. p-values were calculated by one-way ANOVA test (n.s : not-significant, * : $p < 0.05$, ** : $p < 0.01$, ***: $p < 0.001$). **(C)** Endogenous Immmuno-precipitation assay between PHF6 and DPY30, a subunit of MLL complex. ASH2L as a positive control

H2BK12Ac is a preceding modification of H2BK120ub, and PHF6 links H2BK12Ac and H2BK120ub

Since I confirmed that PHF6 is a H2BK12Ac reader *in vitro* and transcriptional activator by H2BK120ub regulator *in vivo*, I checked for potential crosstalk between H2BK120ub and H2BK12Ac. Given that CBP/p300 functions as an acetyltransferase of several lysine residues of histone H2B, including K12, I knocked down p300 with siRNA and checked the H2BK120ub level in differentiated WT ESCs. As a result, interestingly, p300 knockdown decreased H2BK120ub levels upon Dox treatment (FigIII-3A), suggesting that there is a portion of the H2BK120ub regulated by p300-dependent acetylation. For directly confirming the relationship between H2BK12Ac and H2BK120ub, I over-expressed H2BK12R mutant and H2BK120R mutant, which are depletion of H2BK12Ac and H12BK120ub, respectively, in differentiated ESCs, and checked the levels of H2BK12Ac and H2BK120ub in these exogenous H2B-containing mono-nucleosomes. For purifying only exogenous-H2B containing nucleosomes,

Micrococcal Nucleases (MNase) were treated into the cell lysates, and mono-nucleosomes with exogenous H2B-Flag were purified by Flag-M2 pulldown assay. As a result, nucleosomes with H2BK12R acetylation deficient mutation led to marked reduction in H2BK120ub compared to nucleosomes with H2B WT, whereas nucleosomes with H2BK120R ubiquitination deficient mutation did not affect H2BK12Ac levels (FigIII-3B). These data suggest that there are some H2BK120 ubiquitination portions exist only when H2BK12Ac are preceded. Next, I checked that H2BK12Ac reading activity of PHF6 influences this H2BK12Ac-H2BK120ub cascade. I over-expressed PHF6 WT and E223S, an H2BK12Ac recognition mutant, in differentiated ESCs, and found that H2BK120ub levels decreased when PHF6 K12Ac reading activity was broken (FigIII-3C). These data suggest that K12Ac functions as a preceding modification of H2BK120ub and PHF6 regulates the H2BK120ub level via H2BK12Ac recognizing activity.

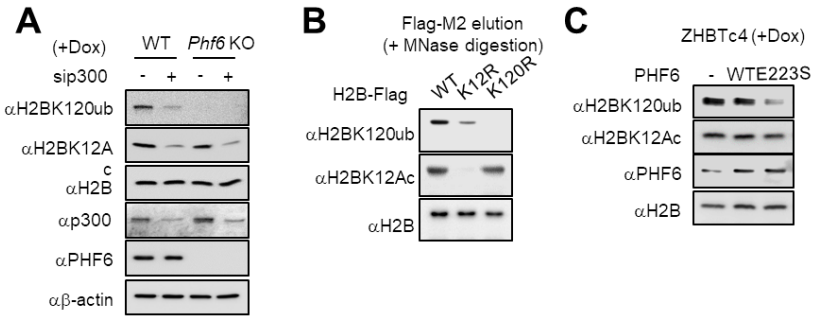


Figure III- 3. H2BK12Ac act as a preceding modification to H2BK120ub, and acetyl-reading activity of PHF6 is important for H2BK120ub

(A) Western blot analysis of p300 knockdown in WT and *Phf6* KO ESCs with Dox treatment. **(B)** Comparison of correlation between K12Ac and K120ub using H2B K12R / K120R mutants. MNase digestion was performed to elute mono-nucleosomes containing these H2B WT and MTs, and the modification states of these ectopic-H2B containing mono-nucleosomes were confirmed by immunoblot. Mono-nucleosome states by MNase digestion were confirmed by DNA agarose gel electrophoresis. (**C**) Global levels of histone modifications by immunoblot analysis after introducing PHF6 E223S mutant compared to PHF6 WT.

PHF6 does not interact with RNF20/40

Many genes with PHD domains work as a linker to recruit other modifying enzymes and alter modification status. Therefore, I checked whether PHF6 recruits other E3 ubiquitin ligases to exert its E3 ubiquitin ligase activity. In mammals, RNF20/40 is known to date as an E3 ligase for H2BK120ub and USP44 as a deubiquitinase in several systems including ESC differentiation (Fuchs et al., 2012). However, the co-immunoprecipitation assay revealed that PHF6 bound neither RNF20/40 nor USP44 when ESCs were differentiated into trophectoderm (FigIII-4A). In addition, RNF20/40 and USP44 expression levels did not differ between *Phf6* KO and WT ESCs during trophectoderm differentiation (FigIII-4B). Moreover, on the promoter of *Cdx2* during trophectoderm differentiation, there was no significant difference of RNF20 recruitment between WT and *Phf6* KO ESCs, whereas H2BK120ub level is significantly decreased in *Phf6* KO ESCs (FigIII-4C). These data indicate that RNF20/40 and USP44 are not responsible for the portion of H2BK120ub regulated by PHF6 during trophectoderm differentiation.

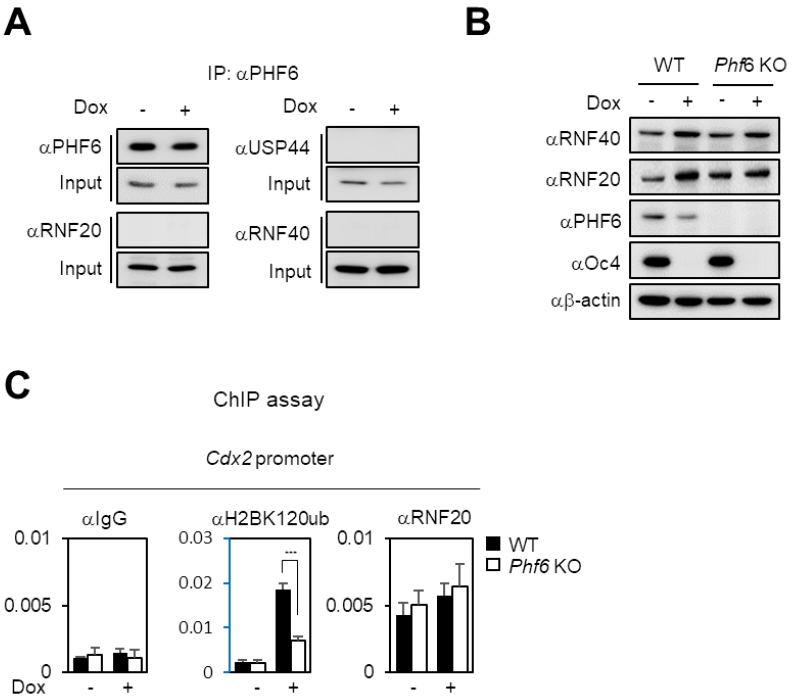


Figure III- 4. PHF6 is independent of RNF20/40 E3 ligases.

(**A**) Endogenous co-IP between PHF6 and RNF20, RNF40 and USP44 during trophectoderm differentiation. (**B**) Western blot analysis of RNF20/40 between WT and *Phf6* KO ESCs during trophectoderm differentiation. (**C**) ChIP assay on the promoter of *Cdx2* during trophectoderm differentiation. *p*-values were calculated by one-way ANOVA test. (*** : $p < 0.001$).

PHF6 has an intrinsic E3 Ub ligase activity for H2BK120ub in H2BK12Ac-dependnet manner

Since PHF6 did not have any interactions with already known enzymes in trophectoderm differentiation, I further hypothesized that decreased level of H2BK120ub in *Phf6* KO ESCs upon Dox treatment could be attributed to intrinsic E3 ligase activity of PHF6 specifically on H2BK120. Recent studies have reported that PHD domain itself has an E3 ligase activity, like RING domain (Ivanov et al., 2007; Lu et al., 2002). Since I have proven that PHF6 recognizes H2BK12Ac through ePHD2 domain that there exists a correlation between K12Ac and K120ub in H2B, I examined whether PHF6 has an E3 ligase activity for H2BK120ub, and the activity is depending on H2BK12Ac. First, I tested whether PHF6 has intrinsic E3 ligase activity by checking the ability to form poly-ubiquitin chains. As a result of *in vitro* ubiquitination assay, I could confirm that PHF6 has the ability to form poly-ubiquitin chains by working with UBCH3 or UBCH6 among several E2 enzymes (Fig III-5A). Next, for checking that PHF6 could work as an E3 ligase to

H2BK120ub, I eluted mono-nucleosomes that contain WT or K12R mutant of H2B with Flag tags from *Phf6* KO ESCs using MNase digestion. Then, I performed *in vitro* ubiquitination assay by mixing eluted exogenous H2B-Flag containing mono-nucleosomes as substrate and purified GST-PHF6 as E3 ligase. As a result, I found that PHF6 functions properly as E3 ligase to H2BK120ub only when UBCH3, not UBCH6, works together as its E2 partner (FigIII-5C). Moreover, it was also confirmed that H2BK120ub by UBCH3-PHF6 is not conjugated in H2BK12R nucleosomes (FigIII-5C). These data indicate that PHF6 has an E3 ligase activity and this activity works in acetylation-dependent manner to H2BK120ub.

Next, I determined ePHD1 domain is responsible for exerting the E3 ubiquitin ligase activity of PHF6. In the cases of Zinc fingers as an E3 ligases, it was previously reported that Zinc ion captured through cysteine residues is important for the function of PHD domains, and nearby hydrophobic residues are also important for E3 ligase activity (Plechanovova et al., 2012; Zheng et al., 2000). From this reasoning, I

generated point mutants of PHF6 ePHD1 domain in several conserved residues and checked the E3 ligase activity of each PHF6 point mutants. *In vitro* ubiquitination assay revealed that PHF6 WT is able to ubiquitinate H2BK120; however, both E223S mutant (H2BK12Ac unrecognizable mutant) or several ePHD1 domain mutants were not able to ubiquitinate H2BK120 (FigIII-5D). Interestingly, PHF6 C20A mutant, which is located in extended region of ePHD1, did not lose the enzymatic activity, suggesting that extended region of ePHD1 domain is not important for E3 ligase activity. Previously, I validated that the ePHD1 domain of PHF6 does not affect the binding of PHF6 with H2B (FigIII-1B). Taken together, these results illustrate that PHF6 recognizes H2BK12Ac through ePHD2 domain and subsequently deposits ubiquitin at H2BK120 residue, thus suggesting that PHF6 has E3 ligase enzymatic activity through ePHD1 domain *in vitro*.

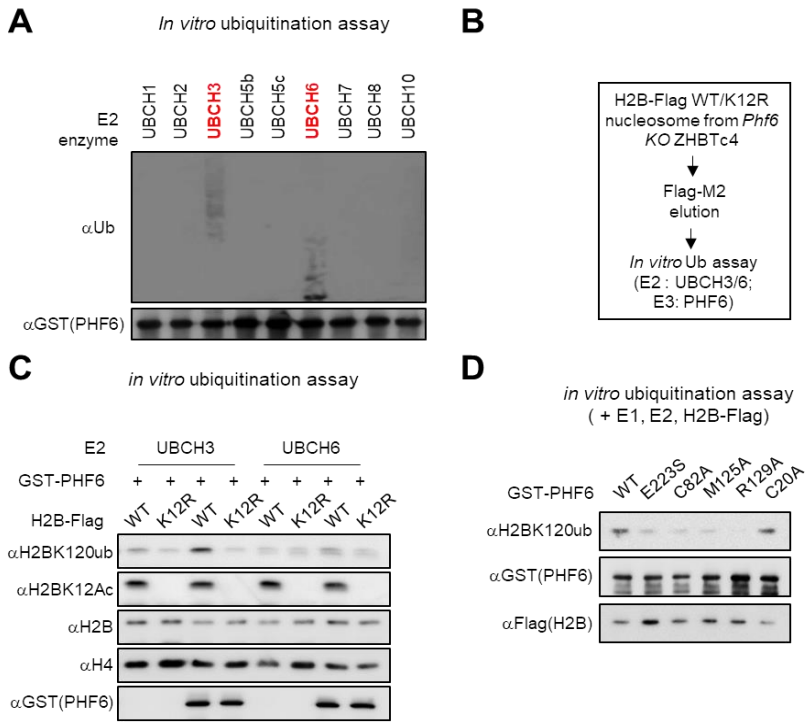


Figure III- 5. PHF6 is an H2BK12Ac-dependent E3 ligase of H2BK120ub

(**A**) *in vitro* ubiquitination assay of PHF6 and several E2 enzymes for poly-ubiquitin chains formation. (**B**) Schematics of *in vitro* nucleosome ubiquitination assay. (**C**) *in vitro* ubiquitination assay of H2B WT or K12R containing nucleosomes by PHF6 and UBCH3/6. Histone H4 blots for nucleosome maintenance. (**D**) *in vitro* ubiquitination assay of H2B WT containing nucleosomes with PHF6 WT or functional mutants of each domain.

II-4. Discussions

My results revealed that PHF6 recognized H2BK12 acetylation via the ePHD2 domain, triggered H2BK120 ubiquitination through the E3 ligase enzymatic activity in preceding acetylation-dependent manner, and facilitated trophectoderm lineage determination through activation of *Cdx2* expression. I have established the molecular basis of acetylation-dependent H2B ubiquitination by two ePHD domains in PHF6 and uncovered previously unidentified histone modification axis for gene expression activation during trophectoderm differentiation.

H2BK12Ac was reported to be generally required for transcriptional activation. To be precise, it has been reported that H2BK12Ac is detected dominantly within the promoter region of transcriptionally activated genes, but how H2BK12Ac affects transcription activation has not been known yet. Upon the transcription regulation machinery of other acetylations on Histone H3 and H4 with their specific readers, it has been predicted that the transcription activation mechanism through

H2BK12Ac would have a reader for recruiting other factors like transcription factors, chromatin remodeling factors, or histone modifying enzymes. Here I reported that ePHD2 domain of PHF6 has a function to recognize H2BK12Ac *in vitro*. Interestingly, the ePHD2 domain-specific acidic amino acid-rich regions are important for H2BK12Ac recognition. With the previous only report for acetylation recognition activity of PHD domain in DPF3b, these results suggest that the acidic amino acids are important for the acetylation reading activity by PHD domain.

With the activity of H2BK12Ac reading activity, I proved that PHF6 has an intrinsic E3 ubiquitin ligase activity for H2BK120ub *in vitro*. I also proved that this enzymatic activity is dependent on the acetylation reading activity. In several cases, enzymes that have PHD domains show their enzymatic activity in a PHD domain- dependent manner. Moreover, there are some reports that PHD domain has E3 ligase activity, based on the structural similarity with RING domain. Because PHF6 has only two conserved ePHD domains, I hypothesized the ePHD1 domain as an E3 ligase domain, and proved that PHF6 has an H2B binding affinity with

or without ePHD1 domain, but lose the E3 ligase activity when ePHD1 structure was broken. With these results, I first confirmed the precise transcription activation machinery of H2BK12Ac histone marker with the E3 ligase enzymatic activity of PHF6, and proposed the transcription activation cascade by H2BK12Ac-H2BK120ub-H3K4me3 histone modifications crosstalk.

With these *in vitro* results, I initially expected that H2BK12Ac levels would increase at the target gene promoters upon Dox treatment, with or without PHF6. Contrary to expectations, however, H2BK12Ac levels on target genes declined during Dox treatment, just like the case of global H2BK12Ac levels. In addition, unlike the global H2BK12Ac level, the H2BK12Ac level at the target promoter was further decreased in *Phf6* KO cells. This result suggests that the deacetylase of H2BK12Ac will be increased of its activity during trophectoderm differentiation, and binding of PHF6 at the promoters inhibits the effects of deacetylase at the target genes.

RNF20/40 is a known E3 ligase of H2BK120ub in several systems. In

embryonic stem cells differentiation, it was known that RNF20/40 regulates H2BK120ub level at the promoter of neural ectodermal genes upon Retinoic Acids (RA) treatment. However, I found that there were no effects of RNF20/40 to *Cdx2* and *Gata2* expression during trophecotderm differentiation. Interestingly, of total H2BK120ub levels in trophectoderm differentiation, not all H2BK120ub levels depend on H2BK12Ac. This suggests that not all H2BK120ub levels are regulated by PHF6 during trophectoderm differentiation, and PHF6-dependent H2BK120ub levels are strictly concentrated in the trophectoderm differentiation genes. Taken together, my study is the first to show that PHF6 is a key epigenetic regulator for trophectoderm differentiation through activating gene expression by H2BK12Ac-H2BK120ub-H3K4me3 histone modification-cascade axis.

II-5. Materials and Methods

Antibodies and reagents

The following commercially available antibodies were used: anti-PHF6 (68262) antibody was purchased from Novus; anti-H3 (ab1791), anti-H2B (ab1790), anti-H3K4me3 (ab8580), anti-H3K27me3 (ab6002), anti-H3K9me3 (ab8898), anti-H2BK12Ac (ab195494), anti-RNF40 (ab191309) and anti-Cdx2 (ab157524) antibodies were purchased from Abcam; anti-OCT4 (sc-5279), anti-USP44 (sc-377203), anti-p300 (sc-8981), anti-HDAC1 (sc-6298), anti-GST (sc-459) and anti-GAPDH (sc-25778) antibodies were purchased from Santa Cruz Biotechnology; anti-β-actin (A1978) and anti-Flag (F3165) antibodies was purchased from Sigma; anti-H2BK120ub (39623) antibody was purchased from ActiveMotif; anti-RNF20 (A300-714A), anti-DPY30 (A304-296A), and anti-ASH2L (A300-489A) antibodies were purchased from Bethyl Laboratories; anti-HA (MMs-101R-500) antibody was purchased from Covance; anti-His (G020) antibody was purchased from Abm.

Knockdown of genes

For knockdown of target genes, I used both siRNA and lenti-viral shRNA systems. For knockdown by siRNA, lipofectamine 3000 Transfection Reagents were used. For knockdown by lenti-viral shRNA system, shRNA constructs were transfected in HEK293T cells with packaging system. After collecting, viruses were concentrated and infected into cells with polybrene. Sequences of siRNA oligomers and shRNAs are as follows:

shPHF6 Forward : 5'- CCGGGTTCAGCTCACAACAACATCACTCG AGTGAT GTTGTTGTGAGCTGAACTTTTTG -3',

shPHF6 Reverse : 5'- AATTCAAAAAGTTCAGCTCACAACAACATC ACTCGAGTGATGTTGTTGTGAGCTGAAC 3',

siRNF20 : 5'- AGTCACAGTTCTCTGTCCTGTATA -3',

sip300 : 5'- CCCTGGATTAAGTTTGATAAA -3'.

Western blotting

All concentrated cells were lysed in the lysis buffer (50mM Tris-HCl pH

8.0, 200mM NaCl, 0.5% NP-40, protease inhibitors (Aprotinin and Leupeptin) at 4□, 10 minutes. For sampling of western blot of histone modifications and enzymes at the same conditions, we performed sonication shortly of all lysed samples. After sonication and discarding debris, sample density was calculated by Bradford assay for normalization.

Co-Immunoprecipitation (co-IP) analysis

For co-Immunoprecipitation assay, cells were lysed in lysis buffer as same condition of western blotting analysis, except sonication. In the lysates, antibodies for pull-down assay were mixed with lysates, and reaction was done at 4□, O/N. After overnight antibody reaction, 50% slurry of nProtein A Sepharose (17-5280-04, GE Healthcare) and nProtein G Sepharose (17-0618-02, GE Healthcare) were mixed with reacting tube at 4□, 2 hours for concentrating proteins recognized by antibodies and partner proteins. After washing beads 5 times, beads were

boiled with sampling buffers and loaded to SDS-PAGE gel for western blotting analysis.

Far-Western blot analysis

GST-PHF6 proteins were purified using GST beads in *E. coli*. Histone extracts from ZHBTc4 ES cells were separated by SDS-PAGE. After separating histone octamers in order of H3, H2B, H2A and H4 according to their sizes, 0.1 μg/ml of purified GST-PHF6 proteins were added to binding buffer (100 mM NaCl, 20 mM Tris [pH 7.6], 10 % glycerol, 0.1 % Tween-20, 50 μM $ZnCl_2$, 2 % skim milk powder and 1 mM DTT), and allowed to conjugate with histone extracts on membrane. After binding of GST-PHF6 on histones, western blot analysis was performed using anti-GST antibody.

***In vitro* Histone peptide pulldown assay**

0.5 μg/μl of biotinylated histone peptides mixed with 0.3μg of GST-

eluted proteins in assay buffer (250 mM NaCl, 50 mM Tris-HCl [pH 7.5], 0.05 % NP-40, and 50 μM $ZnCl_2$) were incubated overnight. Then, 50 % slurry of streptavidin beads were added and further incubated for 1 hr. After removing non-specific interactions by washing, beads were boiled in sampling buffer and peptide-protein interaction was detected by immunoblotting. Biotinylated histone peptides of modified histone H3 were purchased from Boston BioChems, and those of modified histone H2B were from JPT.

***In vitro* Histone peptide binding array**

Histone peptide binding array kit was purchased from ActiveMotif (#13005). The following assays were progressed according to the provided protocol. Briefly, blocking a kit by 5% milk in TTBS (10mM Tris-HCl pH 7.5, 150mM NaCl, 0.05% tween-20) at 4□ Overnight. After blocking, the kit was incubated with 0.3μg of eluted GST-PHF6 in binding buffer (100mM KCl, 20mM HEPES pH 7.9, 1mM EDTA, 10%

glycerol, 0.1mM DTT) RT 1h. Then, GST primary antibody and secondary antibody were incubated on the kit sequentially at 1h RT each. Detection was by ECL develop. Washed three times with TTBS between each step.

Micrococcal Nuclease (MNase) digestion

For making nucleosomes without H2BK12Ac or H2BK120ub, H2B-Flag WT/K12R/K120R were transfected into cells. Then, cells were lysed by MNase lysis buffer (50mM Tris-HCl pH 7.5, 1mM CaCl2, 0.2% NP-40, 100×Inhibitor Complex) and treated 50U MNase (#M0247S, NEB) for making mono-nucleosomes. After MNase digestion, H2B-Flag containing mono-nucleosomes were pull down by Flag-M2 bead overnight, then these mono-nucleosomes were eluted by 3×Flag peptides (#F4799, Sigma).

***In vitro* Ubiquitination Assay**

For substrates, H2B-Flag WT/K12R were transfected into *Phf6* KO ZHBTc4 and H2B-Flag containing mono-nucleosomes were purified by MNase digestion. After purification, these nucleosomes were mixed with E1, E2, E3 enzyme, ubiquitin, 50μM ZnCl2, and 10x buffer (500mM Tris-HCl pH 7.5, 20mM ATP, 10mM MgCl2, 2mM DTT) in 50μl volume, then incubated at 37℃, 1h. To stop the reaction, 10×sampling buffer was added and boiled at 100℃ 10 minutes. E1, E2 (UBCH3, UBCH6), ubiquitin were purchased from Boston Biochems. Purified GST-PHF6 WT/MTs were considered as E3 ligases in reaction.

Quantitative RT-PCR

Total RNAs were extracted using Trizol (Invitrogen) and reverse transcription was performed from 1~2μg total RNAs using the M-MLV cDNA Synthesis kit (Enzynomics). The abundance of mRNA was detected by an ABI prism 7500 system or BioRad CFX384 with SYBR TOPreal qPCR 2x PreMix (Enzynomics). The quantity of mRNA was

calculated using ddCt method and *Gapdh* and *β-actin* were used as controls. All reactions were performed as triplicates. Primer sequences for qRT-PCR analysis are as follows:

Phf6 Forward : 5'-TGGCGGGAGAAGACGAAGA -3' ,

Phf6 Reverse : 5'- TCTGTAGGGAGCACCGACAAT -3' ,

Gapdh Forward : 5'- CATGGCCTTCCGTGTTCCTA -3',

Gapdh Reverse : 5'- CCTGCTTCACCACCTTCTTG -3'.

Tead4 Forward : 5'- GCACCATTACCTCCAACGAG-3' ,

Tead4 Reverse : 5'- GATCAGCTCATTCCGACCAT-3' ,

Oct4 Forward : 5'- AGAGGATCACCTTGGGGTACA -3',

Oct4 Reverse : 5'- CGAAGCGACAGATGGTGGTC -3',

Cdx2 Forward : 5'- CAAGGACGTGAGCATGTATCC -3',

Cdx2 Reverse : 5'- GTAACCACCGTAGTCCGGGTA -3'.

ChIP assays

ChIP assay was basically performed according to Iyer lab protocol. Briefly, cells were cross-linked by 1% formaldehyde 10 min, 37□. After

cross-linking, cells were harvested and lysed for nuclear fraction. Nuclear pellets were sonicated in RIPA buffer, then incubated with A/G agarose beads about 30mins for pre-clearing. After pre-clearing, primary antibodies were added and incubates overnight. After beads reaction, salt-gradient washout beads was performed, and protein-DNA complex bound to beads were eluted with elution buffer (1% SDS, 100mM NaHCO3). Then, eluted complex were de-crosslinked between DNAs and proteins at 55□, overnight. After de-crosslinking, DNAs were isolated from the mixed elutes by purification columns (#28105, QIAGEN). Eluted DNAs were detected by qRT-PCR. All reactions were performed as triplicates. Primer sequences used for ChIP-qPCR are as follows:

Cdx2 promoter : Forward : 5'– TCAACGTTTGTCCCCAGACA –3',
Reverse : 5'- ACGTCCTTGTCCAGAAGGTAGCT -3';

Ascl2 promoter : Forward : 5'- TTCATGGCCTACTGACGTGC -3',
Reverse : 5'- CATTCGGACCCTCTCTCGGA -3'.

Tfap2c promoter : Forward : 5'- GTCGGACAGGTTGAGCTAGAG-3',

Reverse : 5'- CAGGGGTCAGAAGGGTAGCA-3'.

Klf6 promoter : Forward : 5'- TGGCATAGTCTCCATGCTTCC-3',

Reverse : 5'- TTTTCGGGTACTGGGCTCTG-3'.

Statistical analysis

All experiments were performed independently at least three times. Values are expressed as mean ± SEM. Significance was analyzed using ANOVA test. A *p*-value of less than 0.05 was considered statistically significant.

CHAPTER IV

Conclusions

I found the sequential event of histone crosstalk for lineage-specific gene expressions by histone reader-containing modifying enzyme, PHF6. PHF6 has two extended PHD domains, and ePHD2 domain of PHF6 specifically recognize acetylation of H2B lysine 12 residue. Through ePHD2 domain mediated physical interaction with H2BK12Ac, PHF6 works as an E3 ligase for the mono-ubiquitination of H2B lysine 120 residue, which is one of the transcription activation marker and preceding modification for H3K4me3. By this enzymatic activity, PHF6 transcriptionally regulates the lineage-specific gene expression during trophectoderm differentiation (FigIV-1). Moreover, 3 germ layer differentiation was also insufficient when Phf6 is depleted due to malformation of trophectoderm. Therefore, it could be explained that defective traits observed in the BFL syndrome patients is derived from immature placental formation.

My work first established the precise mechanism of H2B acetylation on transcription. Compared to the acetylations on H3/H4, acetylations on H2A/H2B have been less studied. Among many forms of acetylations of

H2B, it has been reported that the genome distributions H2BK12Ac is enriched at the promoters of actively transcribed genes, suggesting the role as a transcription activation marker. However, the precise machinery of how transcription activation is regulated by H2BK12Ac still remains unknown. In order for histone modifications to work in transcription regulation, the role of histone readers is important for recruitment of several factors. I found that PHF6 is a reader of H2BK12Ac that functions through ePHD2 domain. More specifically, among two ePHD domains of PHF6, only ePHD2 domain has an acidic amino acids-enriched motif, and these residues are localized at the surface in crystal structure. With this prediction, I experimentally confirmed that E223 residue in ePHD2 domain is critical for H2BK12Ac reading activity. Together, these results suggest that the acidic amino acids are important for PHD domain as an acetyl-reader, but not as an methyl-reader.

My work was also meaningful in discovery of another example that PHD domain has an E3 ligase activity. PHD reader domain and RING E3 ligase domain have a similar structure and sequences. I proved that PHF6

has an E3 ligase activity, and mutations of PHF6 on ePHD1 domain or ePHD2 domain both leads loss of intrinsic E3 ligase activity. Because PHF6 ePHD2 domain works as an acetylation reader, these results suggest that ePHD1 domain of PHF6 has an E3 ligase activity *in vitro*. In addition, the decreased global level of H2BK120ub by *Phf6* deficiency also suggest the enzymatic activity of PHF6 as a ubiquitin writer.

Taken together, I first described the transcription activation mechanism by histone modifications crosstalk during trophectoderm differentiation. In detail, I found the precise function of H2BK12Ac through PHF6, working as a preceding modification of H2BK120ub. PHF6 has an E3 ligase activity and H2BK12Ac reading activity conferred by each domains respectively. With the previous reports stating the H2BK120ub-H3K4me3 links, my data suggest that the H2BK12Ac-H2BK120ub-H3K4me3 histone crosstalk at the promoter of transcribed activated genes is crucial during trophectoderm differentiation.

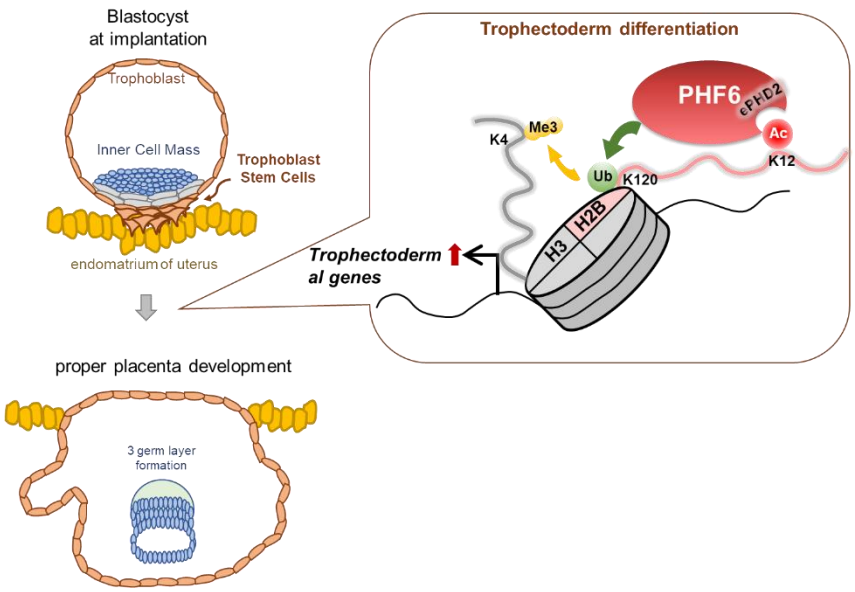


Figure IV- 1. Schematic model to show PHF6 function as an acetylation-dependent E3 ligase of H2BK120ub for *Cdx2* expression during trophectoderm differentiation

Schematic model of the role of PHF6 in trophectoderm differentiation. During trophectoderm differentiation, PHF6 recognizes H2BK12Ac on the *Cdx2* promoter, and ubiquitinates H2BK120 residue. H2BK120ub by PHF6 results in the recruitment of MLL complex and thereby increase of H3K4me3 at the promoter regions of trophectodermal genes. These H2BK12Ac-H2BK120ub-H3K4me3 histone modification crosstalks by PHF6 result in fully activation of gene expression during trophectoderm differentiation.

References

Beck, F., Erler, T., Russell, A., and James, R. (1995). Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. Dev Dyn *204*, 219-227.

Beddington, R.S., and Robertson, E.J. (1999). Axis development and early asymmetry in mammals. Cell *96*, 195-209.

Berland, S., Alme, K., Brendehaug, A., Houge, G., and Hovland, R. (2011). PHF6 Deletions May Cause Borjeson-Forssman-Lehmann Syndrome in Females. Mol Syndromol *1*, 294-300.

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

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

Bou, G., Liu, S., Sun, M., Zhu, J., Xue, B., Guo, J., Zhao, Y., Qu, B., Weng, X., Wei, Y.*, et al.* (2017). CDX2 is essential for cell proliferation and polarity in porcine blastocysts. Development *144*, 1296-1306.

Braude, P., Bolton, V., and Moore, S. (1988). Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. Nature *332*, 459-461.

Candau, R., Zhou, J.X., Allis, C.D., and Berger, S.L. (1997). Histone acetyltransferase activity and interaction with ADA2 are critical for GCN5 function in vivo. Embo j *16*, 555-565.

Cantone, I., and Fisher, A.G. (2013). Epigenetic programming and reprogramming during development. Nat Struct Mol Biol *20*, 282-289.

Capili, A.D., Schultz, D.C., Rauscher, I.F., and Borden, K.L. (2001). Solution structure of the PHD domain from the KAP-1 corepressor: structural determinants for PHD, RING and LIM zinc-binding domains. Embo j *20*, 165-177.

Chen, K., Chen, Z., Wu, D., Zhang, L., Lin, X., Su, J., Rodriguez, B., Xi, Y., Xia, Z., Chen, X.*, et al.* (2015). Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. Nat Genet *47*, 1149-1157.

Cheng, C., Deng, P.Y., Ikeuchi, Y., Yuede, C., Li, D., Rensing, N., Huang, J., Baldridge, D., Maloney, S.E., Dougherty, J.D.*, et al.* (2018). Characterization of a Mouse Model of Borjeson-Forssman-Lehmann Syndrome. Cell Rep *25*, 1404-1414.e1406.

Cockburn, K., Biechele, S., Garner, J., and Rossant, J. (2013). The Hippo pathway member Nf2 is required for inner cell mass specification. Curr Biol *23*, 1195-1201.

Crawford, J., Lower, K.M., Hennekam, R.C., Van Esch, H., Megarbane, A., Lynch, S.A., Turner, G., and Gecz, J. (2006). Mutation screening in Borjeson-Forssman-Lehmann syndrome: identification of a novel de

novo PHF6 mutation in a female patient. J Med Genet *43*, 238-243.

Cross, J.C., Werb, Z., and Fisher, S.J. (1994). Implantation and the placenta: key pieces of the development puzzle. Science *266*, 1508-1518.

Delvecchio, M., Gaucher, J., Aguilar-Gurrieri, C., Ortega, E., and Panne, D. (2013). Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation. Nat Struct Mol Biol *20*, 1040-1046.

Denissov, S., Hofemeister, H., Marks, H., Kranz, A., Ciotta, G., Singh, S., Anastassiadis, K., Stunnenberg, H.G., and Stewart, A.F. (2014). Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. Development *141*, 526-537.

Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K., and Zhou, M.M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. Nature *399*, 491-496.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R.*, et al.* (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol *34*, 184-191.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-156.

Fuchs, G., Shema, E., Vesterman, R., Kotler, E., Wolchinsky, Z., Wilder, S., Golomb, L., Pribluda, A., Zhang, F., Haj-Yahya, M.*, et al.* (2012). RNF20 and USP44 regulate stem cell differentiation by modulating H2B monoubiquitylation. Mol Cell *46*, 662-673.

Gan, Q., Yoshida, T., McDonald, O.G., and Owens, G.K. (2007). Concise review: epigenetic mechanisms contribute to pluripotency and cell lineage determination of embryonic stem cells. Stem Cells *25*, 2-9.

Gerami-Naini, B., Dovzhenko, O.V., Durning, M., Wegner, F.H., Thomson, J.A., and Golos, T.G. (2004). Trophoblast differentiation in embryoid bodies derived from human embryonic stem cells. Endocrinology *145*, 1517-1524.

Gregory, G.D., Vakoc, C.R., Rozovskaia, T., Zheng, X., Patel, S., Nakamura, T., Canaani, E., and Blobel, G.A. (2007). Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. Mol Cell Biol *27*, 8466-8479.

Grewal, S.I., and Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. Science *301*, 798-802.

Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. Nature *389*, 349-352.

Guo, Y., Mantel, C., Hromas, R.A., and Broxmeyer, H.E. (2008). Oct-4 is critical for survival/antiapoptosis of murine embryonic stem cells subjected to stress: effects associated with Stat3/survivin. Stem Cells *26*, 30-34.

Hamilton, W.J., and Laing, J.A. (1946). Development of the egg of the cow up to the stage of blastocyst formation. J Anat *80*, 194-204.

Hay, D.C., Sutherland, L., Clark, J., and Burdon, T. (2004). Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. Stem Cells *22*, 225-235.

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W.*, et al.* (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature *459*, 108-112.

Hirate, Y., Hirahara, S., Inoue, K., Suzuki, A., Alarcon, V.B., Akimoto, K., Hirai, T., Hara, T., Adachi, M., Chida, K.*, et al.* (2013). Polarity-dependent distribution of angiomotin localizes Hippo signaling in preimplantation embryos. Curr Biol *23*, 1181-1194.

Hu, D., Garruss, A.S., Gao, X., Morgan, M.A., Cook, M., Smith, E.R., and Shilatifard, A. (2013). The Mll2 branch of the COMPASS family regulates bivalent promoters in mouse embryonic stem cells. Nat Struct Mol Biol *20*, 1093-1097.

Ivanov, A.V., Peng, H., Yurchenko, V., Yap, K.L., Negorev, D.G., Schultz, D.C., Psulkowski, E., Fredericks, W.J., White, D.E., Maul, G.G.*, et al.* (2007). PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. Mol Cell *28*, 823-837.

Jiang, H., Shukla, A., Wang, X., Chen, W.Y., Bernstein, B.E., and Roeder, R.G. (2011). Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. Cell *144*, 513-525.

Jiang, J., Chan, Y.S., Loh, Y.H., Cai, J., Tong, G.Q., Lim, C.A., Robson, P., Zhong, S., and Ng, H.H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat Cell Biol *10*, 353-360.

Karlic, R., Chung, H.R., Lasserre, J., Vlahovicek, K., and Vingron, M. (2010). Histone modification levels are predictive for gene expression. Proc Natl Acad Sci U S A *107*, 2926-2931.

Keramari, M., Razavi, J., Ingman, K.A., Patsch, C., Edenhofer, F., Ward, C.M., and Kimber, S.J. (2010). Sox2 is essential for formation of trophectoderm in the preimplantation embryo. PLoS One *5*, e13952.

Kleine-Kohlbrecher, D., Christensen, J., Vandamme, J., Abarrategui, I., Bak, M., Tommerup, N., Shi, X., Gozani, O., Rappsilber, J., Salcini, A.E., *et al.* (2010). A functional link between the histone demethylase PHF8 and the transcription factor ZNF711 in X-linked mental retardation. Mol Cell *38*, 165-178.

Kuckenberg, P., Buhl, S., Woynecki, T., van Furden, B., Tolkunova, E., Seiffe, F., Moser, M., Tomilin, A., Winterhager, E., and Schorle, H. (2010). The transcription factor TCFAP2C/AP-2gamma cooperates with CDX2 to maintain trophectoderm formation. Mol Cell Biol *30*, 3310-3320.

Kunath, T., Arnaud, D., Uy, G.D., Okamoto, I., Chureau, C., Yamanaka,

Y., Heard, E., Gardner, R.L., Avner, P., and Rossant, J. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. Development *132*, 1649-1661.

Lan, F., Collins, R.E., De Cegli, R., Alpatov, R., Horton, J.R., Shi, X., Gozani, O., Cheng, X., and Shi, Y. (2007). Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. Nature *448*, 718-722.

LeRoy, G., Rickards, B., and Flint, S.J. (2008). The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. Mol Cell *30*, 51-60.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics *12*, 323.

Li, J., Pan, G., Cui, K., Liu, Y., Xu, S., and Pei, D. (2007). A dominant-negative form of mouse SOX2 induces trophectoderm differentiation and progressive polyploidy in mouse embryonic stem cells. J Biol Chem *282*, 19481-19492.

Liu, Z., Li, F., Ruan, K., Zhang, J., Mei, Y., Wu, J., and Shi, Y. (2014). Structural and functional insights into the human Borjeson-Forssman-Lehmann syndrome-associated protein PHF6. J Biol Chem *289*, 10069-10083.

Lohmann, F., Loureiro, J., Su, H., Fang, Q., Lei, H., Lewis, T., Yang, Y., Labow, M., Li, E., Chen, T.*, et al.* (2010). KMT1E mediated H3K9

methylation is required for the maintenance of embryonic stem cells by repressing trophectoderm differentiation. Stem Cells *28*, 201-212.

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

Lu, C.C., Brennan, J., and Robertson, E.J. (2001). From fertilization to gastrulation: axis formation in the mouse embryo. Curr Opin Genet Dev *11*, 384-392.

Lu, Z., Xu, S., Joazeiro, C., Cobb, M.H., and Hunter, T. (2002). The PHD domain of MEKK1 acts as an E3 ubiquitin ligase and mediates ubiquitination and degradation of ERK1/2. Mol Cell *9*, 945-956.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A *78*, 7634-7638.

Martinez-Balbas, M.A., Bannister, A.J., Martin, K., Haus-Seuffert, P., Meisterernst, M., and Kouzarides, T. (1998). The acetyltransferase activity of CBP stimulates transcription. Embo j *17*, 2886-2893.

Matthews, J.M., Bhati, M., Lehtomaki, E., Mansfield, R.E., Cubeddu, L., and Mackay, J.P. (2009). It takes two to tango: the structure and function of LIM, RING, PHD and MYND domains. Curr Pharm Des *15*, 3681-3696.

Meacham, C.E., Lawton, L.N., Soto-Feliciano, Y.M., Pritchard, J.R., Joughin, B.A., Ehrenberger, T., Fenouille, N., Zuber, J., Williams, R.T.,

Young, R.A., *et al.* (2015). A genome-scale in vivo loss-of-function screen identifies Phf6 as a lineage-specific regulator of leukemia cell growth. Genes Dev *29*, 483-488.

Meissner, A. (2010). Epigenetic modifications in pluripotent and differentiated cells. Nat Biotechnol *28*, 1079-1088.

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.

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

Musselman, C.A., Lalonde, M.E., Cote, J., and Kutateladze, T.G. (2012). Perceiving the epigenetic landscape through histone readers. Nat Struct Mol Biol *19*, 1218-1227.

Nakanishi, S., Lee, J.S., Gardner, K.E., Gardner, J.M., Takahashi, Y.H., Chandrasekharan, M.B., Sun, Z.W., Osley, M.A., Strahl, B.D., Jaspersen, S.L., *et al.* (2009). Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. J Cell Biol *186*, 371-377.

Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A.,

Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N., *et al.* (2009). The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. Dev Cell *16*, 398-410.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell *123*, 917-929.

Okada, Y., Shimazaki, T., Sobue, G., and Okano, H. (2004). Retinoic-acid-concentration-dependent acquisition of neural cell identity during in vitro differentiation of mouse embryonic stem cells. Dev Biol *275*, 124-142.

Org, T., Chignola, F., Hetenyi, C., Gaetani, M., Rebane, A., Liiv, I., Maran, U., Mollica, L., Bottomley, M.J., Musco, G., *et al.* (2008). The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. EMBO Rep *9*, 370-376.

Plechanovova, A., Jaffray, E.G., Tatham, M.H., Naismith, J.H., and Hay, R.T. (2012). Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. Nature *489*, 115-120.

Prioleau, M.N., Huet, J., Sentenac, A., and Mechali, M. (1994). Competition between chromatin and transcription complex assembly regulates gene expression during early development. Cell *77*, 439-449.

Ralston, A., Cox, B.J., Nishioka, N., Sasaki, H., Chea, E., Rugg-Gunn, P., Guo, G., Robson, P., Draper, J.S., and Rossant, J. (2010). Gata3 regulates

trophoblast development downstream of Tead4 and in parallel to Cdx2. Development *137*, 395-403.

Rayon, T., Menchero, S., Nieto, A., Xenopoulos, P., Crespo, M., Cockburn, K., Canon, S., Sasaki, H., Hadjantonakis, A.K., de la Pompa, J.L.*, et al.* (2014). Notch and hippo converge on Cdx2 to specify the trophectoderm lineage in the mouse blastocyst. Dev Cell *30*, 410-422.

Rochette-Egly, C. (2015). Retinoic acid signaling and mouse embryonic stem cell differentiation: Cross talk between genomic and non-genomic effects of RA. Biochim Biophys Acta *1851*, 66-75.

Rossant, J., and Tam, P.P. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. Development *136*, 701-713.

Russ, A.P., Wattler, S., Colledge, W.H., Aparicio, S.A., Carlton, M.B., Pearce, J.J., Barton, S.C., Surani, M.A., Ryan, K., Nehls, M.C.*, et al.* (2000). Eomesodermin is required for mouse trophoblast development and mesoderm formation. Nature *404*, 95-99.

Sanchez, R., and Zhou, M.M. (2011). The PHD finger: a versatile epigenome reader. Trends Biochem Sci *36*, 364-372.

Schultz, R.M., Davis, W., Jr., Stein, P., and Svoboda, P. (1999). Reprogramming of gene expression during preimplantation development. J Exp Zool *285*, 276-282.

Shen, J., Jiang, D., Fu, Y., Wu, X., Guo, H., Feng, B., Pang, Y., Streets, A.M., Tang, F., and Huang, Y. (2015). H3K4me3 epigenomic landscape

derived from ChIP-Seq of 1,000 mouse early embryonic cells. Cell Res *25*, 143-147.

Soto-Feliciano, Y.M., Bartlebaugh, J.M.E., Liu, Y., Sanchez-Rivera, F.J., Bhutkar, A., Weintraub, A.S., Buenrostro, J.D., Cheng, C.S., Regev, A., Jacks, T.E.*, et al.* (2017). PHF6 regulates phenotypic plasticity through chromatin organization within lineage-specific genes. Genes Dev *31*, 973-989.

Sritanaudomchai, H., Sparman, M., Tachibana, M., Clepper, L., Woodward, J., Gokhale, S., Wolf, D., Hennebold, J., Hurlbut, W., Grompe, M.*, et al.* (2009). CDX2 in the formation of the trophectoderm lineage in primate embryos. Dev Biol *335*, 179-187.

Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. Genes Dev *12*, 599-606.

Strumpf, D., Mao, C.A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development *132*, 2093-2102.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem

cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. Science *282*, 2072-2075.

Taverna, S.D., Ilin, S., Rogers, R.S., Tanny, J.C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L.P., Chait, B.T.*, et al.* (2006). Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. Mol Cell *24*, 785-796.

ten Berge, D., Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R.K., and Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. Nat Cell Biol *13*, 1070-1075.

Todd, M.A., and Picketts, D.J. (2012). PHF6 interacts with the nucleosome remodeling and deacetylation (NuRD) complex. J Proteome Res *11*, 4326-4337.

Van Vlierberghe, P., Palomero, T., Khiabanian, H., Van der Meulen, J., Castillo, M., Van Roy, N., De Moerloose, B., Philippe, J., Gonzalez-Garcia, S., Toribio, M.L.*, et al.* (2010). PHF6 mutations in T-cell acute lymphoblastic leukemia. Nat Genet *42*, 338-342.

Van Vlierberghe, P., Patel, J., Abdel-Wahab, O., Lobry, C., Hedvat, C.V., Balbin, M., Nicolas, C., Payer, A.R., Fernandez, H.F., Tallman, M.S.*, et*

al. (2011). PHF6 mutations in adult acute myeloid leukemia. Leukemia *25*, 130-134.

Wang, C., Liu, X., Gao, Y., Yang, L., Li, C., Liu, W., Chen, C., Kou, X., Zhao, Y., Chen, J.*, et al.* (2018). Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. Nat Cell Biol *20*, 620-631.

Wang, J., Leung, J.W., Gong, Z., Feng, L., Shi, X., and Chen, J. (2013). PHF6 regulates cell cycle progression by suppressing ribosomal RNA synthesis. J Biol Chem *288*, 3174-3183.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q.*, et al.* (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet *40*, 897-903.

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.

Wobus, A.M., Holzhausen, H., Jakel, P., and Schoneich, J. (1984). Characterization of a pluripotent stem cell line derived from a mouse embryo. Exp Cell Res *152*, 212-219.

Wu, G., Gentile, L., Fuchikami, T., Sutter, J., Psathaki, K., Esteves, T.C., Arauzo-Bravo, M.J., Ortmeier, C., Verberk, G., Abe, K.*, et al.* (2010). Initiation of trophectoderm lineage specification in mouse embryos is

independent of Cdx2. Development *137*, 4159-4169.

Wu, L., Lee, S.Y., Zhou, B., Nguyen, U.T., Muir, T.W., Tan, S., and Dou, Y. (2013). ASH2L regulates ubiquitylation signaling to MLL: trans-regulation of H3 K4 methylation in higher eukaryotes. Mol Cell *49*, 1108-1120.

Xu, X., Shi, D., Liu, Y., Yao, Y., Dai, J., Xu, Z., Chen, D., Teng, H., and Jiang, Q. (2017). In vivo repair of full-thickness cartilage defect with human iPSC-derived mesenchymal progenitor cells in a rabbit model. Exp Ther Med *14*, 239-245.

Yagi, R., Kohn, M.J., Karavanova, I., Kaneko, K.J., Vullhorst, D., DePamphilis, M.L., and Buonanno, A. (2007). Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. Development *134*, 3827-3836.

Yang, Z., He, N., and Zhou, Q. (2008). Brd4 recruits P-TEFb to chromosomes at late mitosis to promote G1 gene expression and cell cycle progression. Mol Cell Biol *28*, 967-976.

Yeap, L.S., Hayashi, K., and Surani, M.A. (2009). ERG-associated protein with SET domain (ESET)-Oct4 interaction regulates pluripotency and represses the trophectoderm lineage. Epigenetics Chromatin *2*, 12.

Zeng, L., Zhang, Q., Li, S., Plotnikov, A.N., Walsh, M.J., and Zhou, M.M. (2010). Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. Nature *466*, 258-262.

Zhang, C., Mejia, L.A., Huang, J., Valnegri, P., Bennett, E.J., Anckar, J.,

Jahani-Asl, A., Gallardo, G., Ikeuchi, Y., Yamada, T.*, et al.* (2013). The X-linked intellectual disability protein PHF6 associates with the PAF1 complex and regulates neuronal migration in the mammalian brain. Neuron *78*, 986-993.

Zheng, N., Wang, P., Jeffrey, P.D., and Pavletich, N.P. (2000). Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. Cell *102*, 533-539.

Zweier, C., Kraus, C., Brueton, L., Cole, T., Degenhardt, F., Engels, H., Gillessen-Kaesbach, G., Graul-Neumann, L., Horn, D., Hoyer, J.*, et al.* (2013). A new face of Borjeson-Forssman-Lehmann syndrome? De novo mutations in PHF6 in seven females with a distinct phenotype. J Med Genet *50*, 838-847.

국문 초록 / ABSTRACT IN KOREAN

포유류의 초기 배아 발생 단계에서 세포 운명 결정을 위한 특정 유전자들의 발현 조절이 매우 중요하다. 이들 분화 유전자의 특이적 발현을 조절하는 데에 있어서 후성유전학적 조절이 중요한 역할을 한다. 그 중 배반포 분화는 태반 형성의 운명을 가지게 되는 세포들을 말하는데, *Caudal Type Homeobox 2* (*Cdx2*)를 중심으로 배반포 분화로의 운명 결정에 관여하는 여러 중요한 전사 인자들은 밝혀졌지만, 이들의 발현을 조절하는 히스톤 변형에 대해서는 거의 밝혀진 바가 없다.

Plant-Homeodomain Finger 6 (PHF6)는 PHD 도메인을 두 개 가지고 있는 유전자로, 이들 도메인을 통해 히스톤 변형을 인지하는 기능을 가질 것으로 추측하고 있다. PHF6는 백혈병과 Börjeson-Forssman-Lehmann (BFL) 증후군에서 가장 연구가 활발히 되었는데, 여러 유전학적 방법을 통한 연구를 통해 이 질병들에서 PHF6에 다양한 종류의 돌연변이

종류가 형성되었다는 것이 밝혀져 있다. 이 중 BFL 증후군 환자들은 초기 배아 단계의 발생 과정에서 여러 가지 문제가 생겼을 때 나타나는 전형적인 증상들을 보이고 있는데, 이를 통해서 PHF6가 초기 발생 단계에서 중요한 역할을 할 것이라고 추측할 수 있다. 그러나, PHF6의 정확한 분자적 기작에 대해서는 아직 밝혀진 바가 없다.

본 연구에서 PHF6가 결핍된 마우스 배아 줄기 세포를 만들었고, 이를 통해 초기 발생 및 분화 과정에서 PHF6가 배반포 초기 발생에 매우 중요한 역할을 한다는 것을 확인하였다. *Phf6*가 결핍된 배아 줄기 세포는 배반체로 분화하였을 때 배반포 세포들로 분화하는 세포들을 거의 형성하지 못하는 것을 확인하였다. 또한, 배아 줄기 세포를 배반포 세포로만 분화시켰을 때, 주요 전사 인자들이 *Phf6* 결핍 시에 발현하지 못하는 것을 전사체 분석을 통해 확인하였다. 이들의 전사 조절에 대한 PHF6의 분자적 역할을 분석해 본 결과, PHF6는 히스톤 H2BK12의 아세틸화를 인지하는 기능을 통해 전사 활성 마커인 히스톤 H2BK120 잔기를 유비퀴틴화 하는 리가아제

효소 활성을 지니고 있음을 확인할 수 있었다. 또한 배반포 분화 마커 유전자들의 프로모터 위에서의 H2BK120의 유비퀴틴화가 PHF6 의존적으로 증가하는 것을 확인함으로써, PHF6의 유비퀴틴화 효소 활성이 배반포 분화 마커 유전자들의 전사 조절에 중요하다는 것을 밝혔다.

본 연구에서는 아직 밝혀지지 않은 PHF6의 분자적 기작을 유비퀴틴화 리가아제 효소 활성을 가지고 있는 것으로 밝혀내었다. 히스톤 H2BK12 아세틸화의 경우 전사 활성에 관여하는 메커니즘이 밝혀진 바가 없었는데, 본 연구에서 PHF6를 통한 H2BK120 유비퀴틴화를 통해 H2BK12 아세틸화가 전사 활성에 기여한다는 사실을 밝혀내었다. H2BK120 유비퀴틴화는 히스톤 H3K4 삼중메틸화의 선행으로 작용한다는 사실이 잘 밝혀져 있고, 이러한 기존의 보고와 더불어 본 연구는 H2BK12 아세틸화 – H2BK120 유비퀴틴화 – H3K4 삼중메틸화로 이어지는 히스톤 변형 간의 상호 작용을 밝혀내었다. 또한, 초기 발생 중 배반포 분화에서는 히스톤 변형 조절을 통한 배반포 분화 마커 유전자들의 발현 조절 기작이 밝혀진 바

가 거의 없었는데, 본 연구에서 새롭게 밝힌 히스톤 상호 작용이 배반포 발생에 중요한 역할을 한다는 것을 처음으로 밝혀내었다.

주요어:

PHF6, 배반포 분화, Cdx2, 히스톤 H2BK12 아세틸화, 히스톤 H2BK120 유비퀴탄화, 리가아제

학번: 2011-20343