Functional analysis of Zfp819 in pluripotency and embryonic development

DISSERTATION

to obtain the Ph. D. degree

in the International Ph. D. Program for Agricultural Sciences in

Göttingen (IPAG)

Georg-August-University Göttingen, Germany

Presented by

Xiaoying Tan

Born in Sichuan, P.R. China

Göttingen, 2012

D7

Name of referee: Prof. Dr. Dr. Bertram Brenig

Name of co-referee: Prof. Dr. med. Dr. h. c. Wolfgang Engel

Date of dissertation: 26.09.2012

Dedicated to my family,

whose love, encouragement and faith in me were ever present and always appreciated.

"The significant problems we face cannot be solved by
the same level of thinking that created them"

-Albert Einstein

Table of Contents

I List of publications

1. Zheng Y*, **Tan X*** (co-first author), Pyczek J, Nolte J, Pantakani DV, Engel W. (2012). Generation and characterization of Yeast two-hybrid cDNA libraries derived from two distinct mouse pluripotent cell types. Mol Biotechnol. DOI 10.1007/s12033-012-9561-4

2. **Tan X**, Xu X, Zechner U, Nolte J, Engel W, Pantakanai DV. Zfp819, a novel KRAB-zinc finger protein, interacts with KAP1 and functions in genomic integrity maintenance of mouse embryonic stem cells. Stem Cells and Development. (Manuscript under review).

3. Xu X, Pantakani DV, Luhrig S, **Tan X**, Khromov T, Nolte J, Dressel R, Zechner U, Engel W. (2011). Stage-specific germ-cell marker genes are expressed in all mouse pluripotent cell types and emerge early during induced pluripotency. PLoS One 6(7): e22413.

4. Xu X, **Tan X**, Lin Q, Schmidt B, Engel W, Pantakani DV. (2012) Dazl and its novel splice variant functions in translational repression of target mRNAs in mouse embryonic stem cells. BBA - Gene Regulatory Mechanisms journal. (Manuscript under review).

II Abstract

Pluripotency is established and maintained by many pluripotent stem cell-specific transcription factors such as *Oct3/4, Nanog*, and *Sox2* and by the function of their coordinated regulatory network. Several studies have explored how these pluripotency-related factors regulate their thousands of targets by interaction with their known and unknown interacting partners. In the present study, we identified *Zfp819* as a novel pluripotency-related factor and elucidated its function in pluripotent stem cells.

In the first part of this thesis, we generated two distinct pluripotent cell type-specific yeast two-hybrid (Y2H) cDNA libraries which aid in identification of potential interaction partners and to get deeper insights into the function of a protein of interest, namely Zfp819. Further, ESCs cDNA library was used to screen for potential interaction partners of Zfp819. This screen led us to identify 17 putative interaction partners thereby to generate protein interactome of Zfp819. The gene ontology (GO) categorization of putative Zfp819 interaction proteins suggested that *Zfp819* might function in regulation of transcription, in genome integrity maintenance, and in cell cycle/apoptosis.

In the second part of this study, preferential expression of *Zfp819* in undifferentiated pluripotent cell lines and epigenetic marks associated with transcriptional activation at the *Zfp819* promoter region were identified. Additionally, we identified the region(s) responsible for nuclear localization of *Zfp819*. Further, we identified that *Zfp189* functions in the transcriptional repression of specific endogenous retroviral elements (*ERV*s) in pluripotent cells. Through cellular and biochemical studies we show that Zfp819 interacts with several proteins especially Kap1 and Chd4, which are involved in genomic integrity maintenance. And indeed the loss of *Zfp819* in ESCs results in susceptibility for DNA damage and impairment in DNA damage repair. Collectively, the identification of Zfp819 interaction partners together with the molecular and functional studies revealed that *Zfp819* functions in the regulation of genomic stability of pluripotent cells by suppressing a subset of *ERVs*.

Zusammenfassung

Pluripotenz wird durch viele Stammzell-spezifische Transkriptionsfaktoren wie *Oct3/4*, *Nanog* und *Sox2* sowie deren Funktion in ihrem regulatorischen Netzwerk etabliert und aufrechterhalten. Viele Studien haben gezeigt, wie diese Pluripotenz-assoziierten Faktoren ihre Zielgene regulieren. Dies geschieht durch die Interaktion mit bekannten und unbekannten Interaktionspartnern. In der vorliegenden Arbeit haben wir *Zfp819* als einen neuen Pluripotenz-assoziierten Faktor beschrieben und dessen Funktion in pluripotenten Stammzellen untersucht.

Im ersten Teil der vorliegenden Arbeit haben wir zwei cDNA-Banken für Yeast two Hybdrid (Y2H)-Assays aus unterschiedlichen pluripotenten Stammzelltypen generiert. Dies hatte zum Ziel, potentielle Interaktionspartner eines Kandidatenproteins zu identifizieren um dadurch Eindrücke über die Funktion des Proteins zu gewinnen. Für die Identifizierung von potentiellen Interaktionspartnern von Zfp819 haben wir die cDNA-Bank aus embryonalen Stammzellen benutzt. Wir konnten 17 putative Interaktionspartner identifizieren und daraus ein hypothetisches „Interaktom" von Zfp819 generieren. Die Einordnung der putativen Interaktionspartner nach ihrer Gen-Ontologie (GO) ließ vermuten, dass *Zfp819* eine Rolle in der Regulation der Transkription, der Aufrechterhaltung der genetischen Integrität und im Zellzyklus bzw. bei der Apoptose spielt.

Im zweiten Teil der vorliegenden Arbeit wurde die sehr intensive Expression von *Zfp819* in undifferenzierten pluripotenten Zelllinien gezeigt. Desweiteren konnte die Promotorregion von *Zfp819* identifiziert werden, und es wurde gezeigt, dass diese mit epigenetischen Mustern ausgestattet ist. Zusätzlich konnten wir Regionen im *Zfp819*-Gen identifizieren, die für die nukleäre Lokalisation von *Zfp819* verantwortlich sind. Desweiteren konnten wir zeigen, dass *Zfp819* in der transkriptionellen Repression von spezifischen endogenen, retroviralen Elementen (*ERV*s) in pluripotenten Zellen eine Rolle spielt. Durch zelluläre und biochemische Studien konnten wir zeigen, dass Zfp819 mit vielen Proteinen interagiert (z.B. Kap1 und Chd4), welche für die Aufrechterhaltung der genomischen Integrität von Bedeutung sind. Tatsächlich resultierte der Verlust von *Zfp819* in embryonalen Stammzellen in einer erhöhten Anfälligkeit für DNA-Schäden und in einer verminderten DNA-Reparatur.

Zusammenfassend lassen die Identifizierung der Interaktionspartner sowie die Ergebnisse der molekularen und der funktionellen Studien vermuten, dass *Zfp819* durch die Unterdrückung von ausgewählten *ERV*s eine Rolle in der Regulation der genomischen Stabilität von pluripotenten Zellen spielt.

III Abbreviations

°C	Degree celsius
Ac	Acetylation
AP	Alkaline phosphatase
BLAST	Basic local alignment search tool
bp	Base pair
BSA	Bovine serum albumin
c-	Centi
cDNA	Complementary DNA
Chd4	Chromodomain helicase DNA binding protein 4
ChIP	Chromatin immunoprecipitation
d	Day(s)
DAPI	4',6-diamidino-2-phenylindole
dH2O	Destilled water
DMEM	Dulbecco's Modified Eagle's Medium
DN	Down regulation
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease
DSBs	DNA double strand breaks
ECCs	Embryonic carcinoma cells
EDTA	Ethylenediaminetetraacetic acid
EGCs	Embryonic germ cells
EGFP	Enhanced green fluorescent protein
EpiSCs	Epiblast stem cells
ESCs	Embryonic stem cells
et al.	Et alii (and others)
EtBr	Ethidium bromide
FACS	Fluorescence activated cell sorting
FCS	Fetal calf serum
FITC	Fluorescein isothiocyanate
g	Gram
GDNF	Glial cell-derived neurotrophic factor

5

GO	Gene ontology
GSCs	Germline stem cells
h	Hour
H3K27	Histone 3 lysine 27
H3K27me3	Trimethylation of histone 3 lysine 27
H3K4	Histone 3 lysine 4
H3K4me3	Trimethylation of histone 3 lysine 4
H3K9ac	Acetylation of histone 3 lysine 9
H3K9me3	Trimethylation of histone 3 lysine 9
HATs	Histone acetyl transferases
HDACs	Histone deacetylases
HMT	Histone methyltransferase
HPRT	Hypoxanthin-phosphoribosyl-transferase
HRP	Horseradish peroxidase
ICC	Immunocytochemistry
ICM	Inner cell mass
ICRs	Imprinting control regions
iPSCs	Induced pluripotent stem cells
k-	Kilo
Kap1	Krüppel-associated protein1
kb	Kilobase
kDa	Kilodalton
KO	Knock-out
l	Liter
LIF	Leukemia inhibitory factor
m	Meter
M	Molar
m-	Milli
maGSCs	Multipotent adult germline stem cells
Me	Methylation
MEFs	Mouse embryonic fibroblasts
mg	Milligram
min	Minute

miRNA	Micro-RNA
mRNA	Messenger RNA
MS	Mass spectrometry
NCBI	National Center for Biotechnology
NCS	Neocarzinostatin
Neo	Neomycin
ng	Nanogram
NLS	Nuclear localization signal
nt	Nucleotide
OD	Optical density
OE	Overexpression
ORF	Open reading frame
p-	Pico
PBS	Phosphate buffered saline
PFA	Paraformaldehyde
PGCs	Primordial germ cells
pH	Preponderance of hydrogen ions
PMSF	Phenylmethanesulphonylfluoride
qRT-PCR	Quantitative reverse transcription-PCR
RA	Retinoic acid
RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase-PCR
Sdha	Succinate dehydrogenase
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
shRNA	Short hairpin RNA
SSCs	Spermatogonial stem cells
ssDNA	Single strand DNA
Tris	Trishydroxymethylaminomethane
UTR	Untranslated region
V	Volt
Vol.	Volume
WT	Wild type

x g	Gravity
X-α-Gal	5-bromo-4-chloro-3-indolyl-α-D-galactopyranoside
X-Gal	5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside
µ-	Micro

x g   Gravity

IV List of figures and tables

9

1 Introduction

Pluripotent stem cells have the potential to differentiate into all cell types of the body and therefore hold a great promise for future regenerative medicine applications. Understanding of regulatory factors and mechanisms responsible for maintenance of pluripotency and differentiation potential of pluripotent cells is of prime focus. In search of identifying novel factors responsible for maintenance of pluripotency, we performed transcriptome analysis (Meyer et al., 2010) and identified *Zinc finger protein 819* (*Zfp819*) as being specifically expressed in pluripotent cells but not in differentiated cells. The goal of the thesis was to elucidate the function of *Zfp819* in pluripotent cells.

1.1 Stem cells

Stem cells were first discovered in early 1980's and are known to exist in various organisms of a variety of species. Based on their origins, stem cells were categorized into embryonic stem cells (ESCs) and adult stem cells. Embryonic stem cells are generated from *in vitro* culture of inner cell mass (ICM) of blastocysts (Evans and Kaufmann, 1981; Martin, 1981; Thomson et al., 1998). Adult stem cells are also known as somatic stem cells, which are found to deposit in the niche of mature tissues throughout the body to maintain the tissue homeostasis (Prindull et al., 1978; Reynolds and Weiss, 1992; Shi et al., 2005).

Stem cells are defined as a cell type possessing two main properties, namely the ability to divide unlimitedly while maintaining the undifferentiated status and the potential to differentiate and give rise to specialized cell types (reviewed by Sylvester and Longaker, 2004). According to the differentiation potential, stem cells can be divided into totipotent, pluripotent, multipotent and unipotent stem cells. Totipotent stem cells are either zygotes or the cells from early stage embryos (until ~8-16 cell stage), which have the capacity to form an entire organism including all three germ layers, namely endoderm, mesoderm, and ectoderm as well as extra-embryonic tissues (Suwinska et al., 2008). Unlike totipotent stem cells, pluripotent stem cells such as ESCs can also give rise to all three germ layers, but not to extra-embryonic tissues (Evans and Kaufmann, 1981). Multipotent stem cells hold the differentiation potential into several distinct cell types (Orkin and Zon, 2008; Blanpain and Fuchs, 2009). Unipotent stem cells, such as

hepatocytes, which constitute most of the liver, are more specialized and are restricted to produce only a single cell type (Cantz et al., 2008).

1.2 Types of pluripotent stem cells

Based on their origin of establishment, pluripotent stem cells are named as: 1) Embryonic Carcinoma Cells (ECCs), which are derived from malignant germ-cell teratocarcinomas (Finch and Ephrussi, 1967; Kahan and Ephrussi, 1970). 2) Embryonic Germ Cells (EGCs), which are derived from Primordial Germ Cells (PGCs) within gonadal ridges of E8.5 to E11.5 mouse embryos (Matsui et al., 1992). 3) Epiblast embryonic Stem Cells (EpiSC) generated from epiblast of post implantation embryos E5.5 to E6.5 (Brons et al., 2007; Tesar et al., 2007). 4) Multipotent Germline Stem Cells (mGSCs) from neonatal mouse testis (Kanatsu-Shinohara et al., 2004) and multipotent adult Germline Stem Cells (maGSCs) derived from adult mouse testis (Guan et al., 2006). In recent years, the milestone event in stem cell biology was the discovery of induced pluripotency from somatic cells by four transcription factors (*Oct3/4*, *Sox2*, *Klf4*, and *c-Myc*). The factors are introduced into mouse and human somatic cells and reprogram them into pluripotent stem cells, widely known as induced Pluripotent Stem Cells (iPSCs) (Takahashi and Yamanaka, 2006; Takahashi et al., 2007; Yu et al., 2007). The origin of distinct pluripotent stem cells, mentioned above, is depicted in Fig.1.1.

Because of the wide differentiation potential of pluripotent stem cells, they offer a great promise for clinical treatment of individual patients suffering from cellular degeneration because of disease or injury (Ankrum and Karp, 2010). However, before any clinical attempts on human being can be tried, basic research is undoubtedly needed to uncover the underlying mechanisms of stem cell therapy and at the same time patients' safety is also a key concern. Especially, what has to be mentioned is that the majority of achievements about stem cell research come from intensively investigated ESCs, "the gold standard of pluripotency" and mouse models. Thus, elucidation of mouse ESCs at basic molecular level is still of importance and prerequisite.
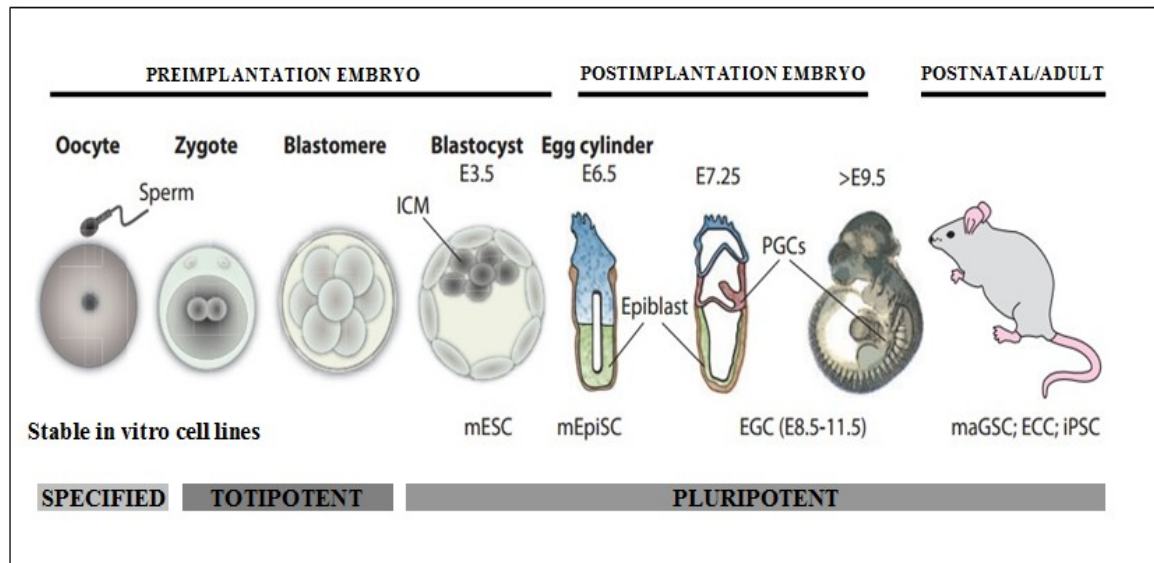
Figure 1.1 Various origins of different pluripotent stem cell types. Different types of pluripotent stem cell can be derived at various embryonic developmental stages. ESCs originate from ICM of E3.5 embryos; egg cylinder stage embryos can give rise to EpiSCs; between E8.5 and E11.5, PGCs can give rise to EGCs. Testes from both neonatal and adult mice can give rise to mGSC and maGSC, respectively. Moreover, germ cell tumors can give rise to ECCs. iPSCs can be obtained after forced expression of the four Yamanaka factors into somatic cells (Figure modified from Dejosez and Zwaka, 2012).

1.3 Embryonic Stem Cells (ESCs)

Since their first derivation in 1981, ESCs have become a widely used tool in biological research. ESCs proliferate rapidly, while maintaining their differentiation potential into all germ layers. Once exposed to extrinsic differentiation induction factors, they start to undergo transitions to relatively differentiated states from a self-renewing and pluripotent state.

In recent years, much of the research has been engaged in fractionation of subtypes in ESC culture. This gave a solid evidence for ESC heterogeneity even though the global presence of core transcriptional factors such as *Oct3/4* and *Nanog*. Moreover, different fractions of cells exhibit varying differentiation potential (Toyooka et al., 2008). Due to the culture condition, there is also a portion of cells which are getting relatively differentiated. Apparently only a subpopulation of pluripotent portion of stem cells is qualified to be induced for a broad range of customized specialized cell types. Therefore, to understand the molecular mechanisms of how pluripotency is maintained in ESCs, further studies are necessary for the applications of regenerative medicine.

1.4 Maintenance of pluripotency and differentiation potential of ESCs

1.4.1 Transcriptional regulation of pluripotency

Previous genome wide approaches have revealed transcriptional regulatory networks as primarily responsible for maintaining ESCs pluripotency (Zuo et al., 2009). However, only some core transcriptional factors occupy the most important positions in this network, such as *Nanog* and *Oct 3/4*, whose knockdown result in loss of pluripotency and induction of differentiation (Loh et al., 2006). Furthermore, those core factors are not only important for the early development of embryos and ESCs (Niwa et al., 2000; Mitsui et al., 2003), but also transcriptionally regulate a number of targets, including components of the *TGF-β* (e.g., *TDGF1, LEFTY2/EBAF*) and Wnt (e.g., *DKK1, FRAT2*) signaling pathways (Loh et al., 2006). In 2006, Sato et al. found out that Wnt signaling pathway is active in undifferentiated ESCs and is required for pluripotency specific gene expression to maintain pluripotency and self-renewal. The balance between two cytokines, leukemia inhibitor factor (LIF) and BMP (bone morphogenetic protein) facilitates self-renewal in mouse ESCs (Ying et al., 2003). However, in human ES cells, in apparent contrast, repression of *BMP* by adding inhibitor *Noggin* and *bFGF* is reported to be essential for maintaining the undifferentiated state (Xu et al., 2005). Question: Is there a common pathway for hESCs and mESCs to maintain pluripotency and how is the interaction relationship between different pathways? Are there other signature markers for pluripotency? Many questions still need to be addressed.

mESCs provide an objective to characterize the molecular pathways that regulate the maintenance of the undifferentiated state as well as the events in commitment towards the specialized cell lineages happening in the earliest stage of embryo development. Loss of pluripotency results in differentiation into various cell types of the three primary germ layers, which is generally accompanied by downregulation of core transcription factors and signaling molecules that maintain the pluripotent phenotype, and upregulation of transcription factors involved in differentiation (Niwa et al., 2000; Velkey and O'Shea, 2003). One of the important signaling pathways responsible for this process is the bone morphogenetic protein (BMP) pathway. Analysis of null mice for both *BMP2* and *BMP4* gene, revealed a crucial role of BMP proteins in primordial germ cell induction (Lawson et al., 1999), mesoderm formation (Fujiwara et al., 2001), and extra-embryonic mesoderm development (Zhang and Bradley, 1996). Similarly, Serine/threonine kinase40 (*Stk40*), a

negatively controlled target of *Oct3/4*, can activate the Erk/MAPK pathway and induce endoderm differentiation (Li et al., 2010).

Little is known about how markers of pluripotent cells are regulated during the differentiation of ESCs. Recent studies have shown that the Germ Cell Nuclear Factor (*GCNF*), an orphan nuclear receptor, is required for the inactivation of pluripotency marker genes during ESCs differentiation (Gu et al., 2005). *GCNF* can directly repress *Oct3/4* and *Nanog* expression by binding to their promoters, when ESCs are induced to differentiate by retinoic acid (RA) (Gu et al., 2005). Another possible mechanism for *GCNF* mediated repression of *Oct3/4* is that *GCNF* recruits DNA methyltransferases (*Dnmt3a and Dnmt3b*) to the *Oct3/4* promoter and facilitates methylation (Sato et al., 2006). Similarly, nucleosome remodeling and histone deacetylation (NuRD) complex is also known to play an important role during ESCs differentiation. ESCs lacking *Mbd3* (a component of NuRD) display severe defects in differentiation (Kaji et al., 2007). It is possible that NuRD is important for silencing pluripotent marker genes, such as *Nanog* during ES cells differentiation.

1.4.2 Epigenetic regulation of pluripotency

Epigenetic modifications including DNA methylation and histone post-translational modifications (e.g. acetylation, methylation, phosphorylation, ubiquitination and ADP-ribosylation) can regulate gene transcription and are also important in maintenance of ESCs pluripotency (Jenuwein and Allis, 2001). Modifications of histone proteins and changes in DNA methylation dictate the final outcome in transcriptional hierarchy mediated by transcriptional factors and are designated to alter the gene expression without changes of the DNA sequence. The nuclear architecture and its associated epigenetic modifications are depicted in Fig.1.2.

The functional evidence for DNA methylation in early embryogenesis and pluripotency came from knockout mouse models. The deficiency of *Dnmt1*, a key DNA methylation maintenance enzyme, leads to a loss of ~90% of DNA methylation and embryos die early in embryogenesis (Li et al., 1992). *Dnmt3a* knockout mice can develop to term but die several weeks after birth, while *Dnmt3b* deficient mice even could not survive until birth because of multiple developmental defects (Okano et al., 1999). Another DNA methylation related gene, CpG binding protein (*CGBP*), possesses a DNA-binding specificity for unmethylated CpG dinucleotides and its deletion in mice gives rise

to reduced levels of DNA methyltransferase activity and thereby to global methylation (Carlone et al., 2005). The *CGBP* knockout ESCs exhibit increased apoptosis, and failed to differentiate even when exposed to differentiation conditions (Carlone et al., 2005).

In ESCs the loss of pluripotency in response to either extrinsic cues or dedicated inductions results in progressive gene silencing of expression. Genes which are active in ESCs become gradually silenced upon differentiation induction, whereas lineage specification marker genes start to express with gradual loss of methylation level at their CpG rich regions (Deb-Rinker et al., 2005). For example, *Oct3/4* and *Nanog*, whose CpG rich regions are hypomethylated in undifferentiated ESCs, become silenced with hypermethylation at CpG rich regions of their promoters (Lagarkova et al., 2010).
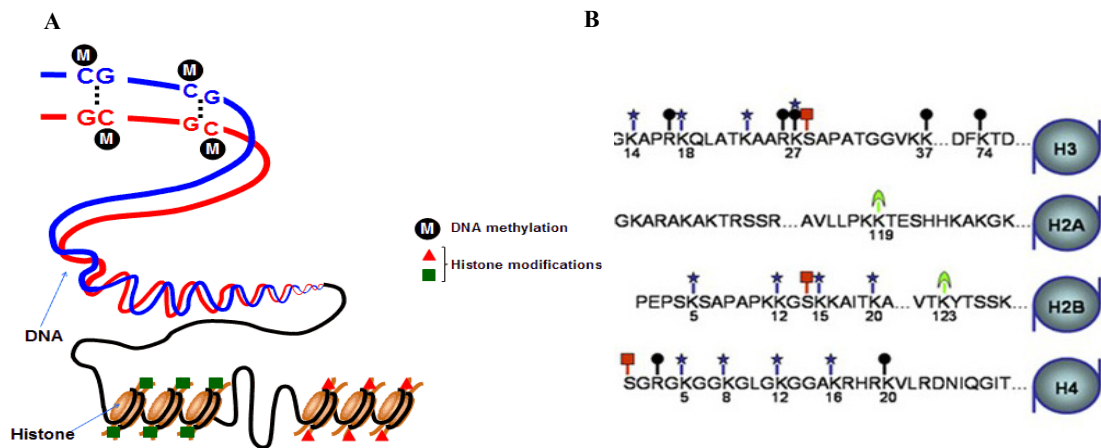
Figure 1.2 Two main components of the epigenetic modifications. A. CpG methylation: methyl groups are added to 5 position of the cytosine in context of CpG, and repress gene activity. **B.** Histone modifications. Various post translational modifications including acetylation (blue star), methylation (black circle), phosphorylation (red box), ubiquitination (green halfmoon) occur at the N terminus amino acid lysine (K) and serine of histones H2A, H2B, H3 and H4. (Figure modified from http://sgugenetics.pbworks.com/w/page/47848618/Introduction%20to%20Epigentic, 2012 and http://www.epigenetics.ch/mod.html, 2012)

The second way of epigenetic regulation is mediated through post transcriptional modifications of histone N-terminal tails (Jenuwein and Allis, 2001). These histone modifications are known to lead to chromatin compaction and gene transcription by altering the interaction between histone proteins and DNA (Narlikar et al., 2002).

Globally, the chromatin feature of undifferentiated ESCs is characterized by dispersed heterochromatin structure and more widely spread euchromatin, exhibiting an enrichment of acetylation in histone 3 (H3) and histone 4 (H4) proteins. During differentiation, the chromatin structure changes into more concentrated heterochromatin and loses the acetylation of histone proteins (Meshorer and Misteli, 2006). Expression of

ESCs hall mark genes (e.g. *Oct3/4* and *Nanog*) is also related with the other activation histone modification marks like trimethylation at lysine 4 of H3 (3meH3K4) and methylation at lysine 36 of H3 (meH3K36) (Efroni et al., 2008). Interestingly, Bernstein et al. (2006) stated that many lineage-specific transcriptional factors such as *Hox* and *Pax* clusters in ESCs have bivalent modification patterns (enrichment for both activation histone mark (3meH3K4) and repression mark (3meH3K27)), reflecting a possible molecular mechanisms for regulation of ESC differentiation and maintenance of pluripotency.

1.4.3 Maintenance of genome stability in ESCs

The most important hallmark of ESCs is to give rise to all cell types of organism. The accumulation of mutations in the ESCs genome would be fatal to the developing organism and the viability of the individual. Therefore ESCs have more extensive mechanisms to maintain the genomic integrity and to avoid failure of organic formation especially for clinical application.

Because of the nature of extraordinary length and chemical characters, DNA is prone to various lesions (Geacintov and Broyde, 2010). DNA damage can happen in response to many exogenous genotoxic stresses, including ionization such as UV, IR, and chemotherapeutic agents as well as endogenous DNA processing events (DNA replication) (Kuo et al., 1984; Torres-Ramos et al., 2002; Short et al., 2005). As shown in Fig.1.3, DNA damage signals immediately activate related PI3K-related kinase (*PIKK*) family proteins, Ataxia telangiectasia mutated (*ATM*) and Ataxia telangiectasia and Rad3 related protein (*ATR*) in mammalian cells (Abraham, 2001). *ATM* is mainly activated in response to double stranded DNA breaks resulting from ionizing radiation for example, first modifies early response proteins such as *MDC1*, which later amplify the DNA damage signal and transduce the signals to downstream effectors (eg, *Chk2*). At the end, tumor suppressor protein *Brca1* for DNA repair can be activated (Elledge, 1996). *ATM* also phosphorylates *MDM2* and *p53* (Harris, 2005) and results in subsequent transcription of gene *p21* which leads to apoptosis. *ATR–ATRIP* complex is recruited to the other stresses such as replication errors, UV exposure, it phosphorylates *Chk1* which mediates the degradation of *Cdc25A* and induces S phase checkpoints. Activation of *P53* by *ATR* is involved in apoptosis.
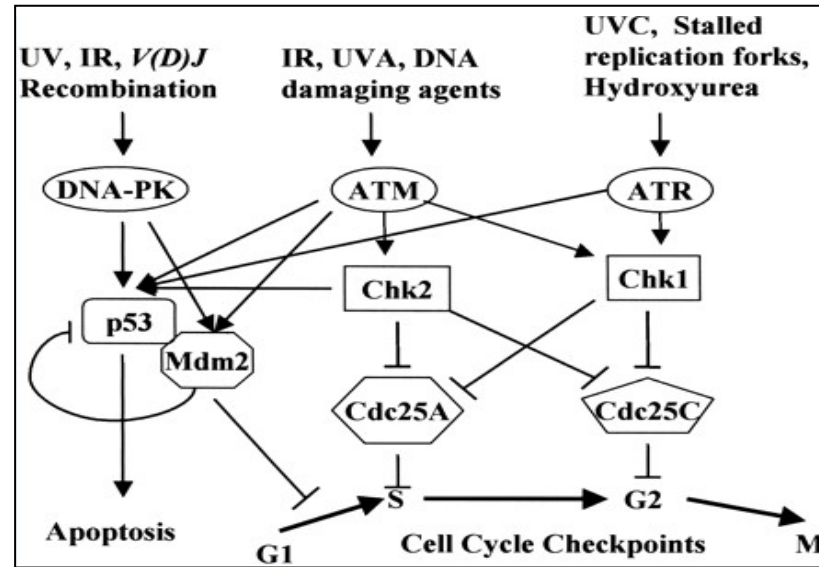
Figure 1.3 Distinct DNA damage response pathways activated by various stresses. DNA double strand breaks (DSB) resulting from ionizing radiation and DNA damage agents can be mainly recognized by *ATM* and *ATR*. In response to IR, *ATM* becomes activated and works as a DNA damage sensor, transducing the signals to down stream effectors (eg. *Chk2*). On the other hand, UV light and replication errors lead to activation of *ATR* and its complex. The down stream protein *P53* can be subsequently phosphorylated in *ATM/Chk2* or *ATR/Chk1* dependent manner, respectively, in two pathways and later perform the function in cell cycle arrest and apoptosis. (Figure modified from Yang et al., 2003)

In order to ensure the integrity of the genome, ESCs employ distinct mechanisms to guard. Once a break has occurred, cells respond by recruiting DNA repair proteins to the DSB sites and initiate an efficient repairing pathway to eliminate mutations (Essers et al., 2006; Mortusewicz et al., 2008). Actually, in line with that ESCs are more stringent in preserving their genomic integrity than somatic cells. Spontaneous mutation frequency in ESCs was shown to be 100-fold lower than in somatic cells (Hong et al., 2006; Stambrook, 2007). By global gene expression assay, it was found that the mRNA levels of DNA repair related genes involved in several types of repair pathways was increased more significantly in ESCs than in differentiated cells (Stambrook, 2007). Alternatively, ESCs escape a G1 checkpoint due to moderate activation of *p53* or even not activated by checkpoint kinase2 (*Chk2*), rather undergo higher incidence of apoptosis in a *p53* independent manner (Hong and Stambrook, 2004; Fluckiger et al., 2006). However, Lin et al. (2005) found that *Nanog* promoter is bound by *p53* and proposed a novel function of *p53* to promote the differentiation of mESCs by repressing *Nanog*, and subsequent apoptosis activation.

Chromodomain helicase DNA-binding protein 4 (*Chd4*) was newly identified as a component of genome guardian machinery (Polo et al., 2010; Urquhart et al., 2011). *Chd4*

is a member of the class II family of CHD ATPases and a major subunit of Nucleosome Remodeling and histone Deacetylation (NuRD) complex. Knockdown of *Chd4* leads to spontaneous DNA damage and increased irradiation (IR) sensitivity (Larsen et al., 2010). At the same time, *Chd4* depletion disrupts DNA repair protein *Brca1* assembly and impairs DNA damage repairing (Smeenk and Lohrum, 2010). *Chd4* deficiency in bone marrow causes loss of lymphoid cell due to the differentiation defects in hematopoietic stem cell (Yoshida et al., 2008).

Endogenous Retroviruses (*ERV*s), whose reactivation is often observed in cancer cells (Howard et al., 2008; Lamprecht et al., 2010) are involved in genome stability of ESCs. *ERV*s are variant of retroviruses, which become permanently integrated in the host genome and are inherited to the next generation after infection of germline cells (Urnovitz and Murphy, 1996; Lower, 1996). *ERV* sequences comprise a large proportion of the genome of human (10%) and mouse (8%), giving rise to extensive genomic alteration (International Human Genome Sequencing Consortium, 2001; International Mouse Genome Sequencing Consortium, 2002). Even though they are largely silenced, some *ERV*s are expressed in placenta of human and animals (Kalter et al., 1973; Levinson et al., 1978), lymphoid cells (Medstrand et al., 1992) and mammary carcinoma cell lines (Patience et al., 1996). The expression of *ERV* was also reported in mouse germ cells and preimplantation embryos (Brulet et al., 1985; Poznanski and Calarco, 1991; Kigami et al., 2003; Evsikov et al., 2004). Comparative expression patterns of *ERV*s in oocytes and preimplantation embryos revealed differential activation of several types of *ERV*s, due to their multiple stage specific promoters (Peaston et al., 2004). In this year, it was again emphasized that their existence is helpful for derivation of extra-embryonic tissues in placental mammals (Macfarlan et al., 2012). Although a large number of *ERV*s is expressed in zygotic and 2-cell stage embryos, these *ERV*s, especially murine endogenous retrovirus *MuERV-L/MERVL*, are highly repressed in ESCs (Maksakova et al., 2011), but the mechanism behind this silencing is still largely unknown.

*ERV*s are able to insert into the host genome usually by copying themselves to a second place or migrating from the original site into a new site (known as retrotransposition), thereby affecting other genes and resulting in genome instability (Doolittle et al., 1982; Maksakova et al., 2006). As retrotransposition can be harmful for the genome, therefore a large number of pathways have evolved to repress the expression of these elements so that the harmful impact of host fitness can be minimized (Walsh and

Bestor, 1999; Pannell et al., 2000; Matsui et al., 2010). DNA methylation is required for this repression in postimplantation embryos and in several cell types derived from postimplantation embryos (Walsh and Bestor, 1999). However, ESCs lacking DNA methytransferases are normal regarding methylation level of *ERV,* which may reflect that repression of *ERV* occurs independent of DNA methylation and probably through histone modification (Leung and Lorincz, 2011).

Kap1 (KRAB-associated protein 1), also known as *TRIM28* (tripartite motif-containing protein 28) is reported to play a very important role in embryogenesis. The deficient mice die before gastrulation stage at E5.5. In 2010, it was reported that Kap1 can repress *ERVs* in ESCs by binding at 5'untranslated region (5'UTR) in particular *IAP* elements. Rowe et al. (2010) showed that *Kap1* deletion leads to a marked upregulation of a range of *ERVs*, in particular *IAP* elements in ESCs as well as in early embryos. In this context, binding between H3K9 methyltransferase *SETDB1 (ESET)* and *Kap1*, which facilitates H3K9me3 of *ERVs*, is indispensable for the repression of *ERVs* (Matsui et al., 2010). *Kap1* is thought as a universal co-repressor protein which enables the repressing target genes of Zinc finger proteins, especially KRAB (Krüppel Related Associated Box)-zinc finger proteins (Friedman et al., 1996; Ryan et al., 1999; Sripathy et al., 2006).

1.5 Zinc finger proteins and their roles in ESCs

Zinc finger proteins are believed to belong to the most abundant of eukaryotic proteins, 2-3% of human proteins are zinc finger proteins (Lander et al., 2001). Many of them harbor a KRAB domain, which is responsible for either gene activation or repression and protein-protein interactions (Margolin et al., 1994). In addition, they bear several classic Cys2His2 zinc finger motifs which enable the recognition of a specific DNA sequence, thus functioning in transcriptional regulation of target genes (Klug, 1999; Wolfe et al., 2000).

A growing body of literature suggests that zinc finger proteins play an important role in maintenance of ESCs pluripotency, differentiation potential, proliferation, and cell cycle control. Zinc finger protein 206 (*Zfp206*) is known to regulate ESCs gene expression and differentiation and its expression is also considered as a hallmark of pluripotent cells (Zhang et al., 2006b; Wang et al., 2007). Likewise, expressed *Zfp42*, also known as *Rex1*, is restricted to undifferentiated stem cells and is known to maintain the undifferentiation state, as the depletion of *Rex1* was found to promote the expression of

three germ layer markers (Scotland et al., 2009). Recently, CtBP-interacting BTB Zinc Finger Protein (*CIBZ*) was shown to be a key transcriptional regulator of ESCs proliferation and played a role in G1/S transition partly depending on Nanog protein expression (Nishii et al., 2012).

About one third of zinc finger proteins bear a KRAB domain with a total number of about 200 genes in human and in mouse (Lander et al., 2001; Waterston et al., 2002). They have transcriptional repression function facilitated by the co-repressor *Kap1*. *Kap1* recruits the heterochromatin proteins (*HP1*), histone deacetylases (*HDACS*), and *SETDB1* and form into a protein complex, which promotes the heterochromatin formation and leads to target genes silencing (Urrutia, 2003). KRAB zinc finger proteins and *Kap1* mediated gene repressions are connected to a variety of processes, including stem cell self-renewal and pluripotency, early embryogenesis (development and differentiation), DNA damage, and genomic imprinting (Cammas et al., 2002; Cammas et al., 2004; Li et al., 2008; Hu et al., 2009; Tian et al., 2009). Interestingly, the KRAB and *Kap1* mediated transcriptional regulation even facilitates repressing of promoters located several tens of kilobases away from the primary repressor site (Groner et al., 2010). *Zfp568* is indispensible to control morphogenesis of embryonic and extra-embryonic tissues (Shibata and Garcia-Garcia, 2011), but *Kap1* is differentially required for this function in a tissue specific manner (Shibata et al., 2011). Moreover, *Zfp57* maintains the genomic imprints in ESCs in a *Kap1* dependent manner, its loss leads to depletion of methylation at several imprinted regions (Zuo et al., 2011). In ESCs, KRAB zinc finger protein *Zfp809* was found as a retrovirus restriction factor, by recruiting *Kap1* to regulate the silencing of *ERVs* (Wolf and Goff, 2009). Even though numerous investigations were documented about functional analysis of KRAB zinc finger proteins, the function of a large number of KRAB zinc finger proteins is still largely unknown.

1.6 Aims

This study was aimed to elucidate the function of *Zfp819*, a novel putative mouse ESC specific marker, which we have identified as being highly expressed in mESCs but not in their differentiated counterparts. Towards this end, the present study was divided into four aspects as listed below:

1. Spatial and temporal expression analysis of *Zfp819*.
2. Functional analysis of *Zfp819* in pluripotent stem cells, especially in ESCs.
3. Identification and characterization of Zfp819 interaction proteins.
4. Generation of loss-of-function and gain-of-function mouse models to decipher the function of *Zfp819* during early embryogenesis.

2 Materials and methods

2.1 Materials

2.1.1 Animals

The used mouse strains C57BL/6J, 129/Sv, CD-1 were initially purchased from Charles River Laboratories, Wilmington, USA, and maintained as colonies at the Animal Facility of the Institute of Human Genetics, Goettingen. *ROSA26CreERT* (Vooijs et al., 2001) and *EIIaCre* (Lakso et al., 1996) mice were kindly provided by Prof. Hahn, Institute of Human Genetics, Goettingen. All experiments involving mice were performed according to protocols authorized by the Institutional Animal Care and Use Committee of the University of Goettingen.

2.1.2 Equipment

The equipment used in this study is listed in Table 2.1.

Table 2.1 Equipment used in present study

Equipment	Company	Location of Company
7900HT Real Time PCR System	Applied Biosystems	USA
Autoclave	Webeco	Germany
Bioruptor Sonicator	Diagenode	Belgium
Bisulfite Pyrosequencing System	Biotage	Sweden
CO_2 Incubator (MCO-20AIC)	SANYO	Japan
Eppendorf Centrifuge (5415D)	Eppendorf	Germany
GS Junior Sequencing System	Roche	USA
Heidolph Reax Vortex	Heidolph	Germany
Heraeus Megafuge 16R	Thermo Scientific	Germany
Heraeus Multifuge X3R	Thermo Scientific	Germany
Invert Microscope (Primo Vert)	Carl Zeiss	Germany
Laminar Airflow	Thermo Scientific	Germany
Microcentrifuge (Heraeus Fresco 21)	Thermo Scientific	Germany
Microscope (BX60)	Olympus	Germany
NanoDrop Spectrophotometer (ND-1000)	Thermo Scientific	Germany
PCR Machine	PeQlab	Germany
Pipettes	Gilson	France
Refrigerator (+4°C)	Privileg	Germany
Refrigerator (-152°C)	SANYO	Japan

Table 2.1 Continued

Equipment	Company	Location of Company
Refrigerator (-20°C)	LIEBHERR	Germany
Refrigerator (-80°C)	SANYO	Japan
Semi-Dry-Blot	Biometra	Germany
Sonicator (Branson sonifier 250)	Branson	USA
Thermomixer 5436	Eppendorf	Germany
Water Bath	Labortechnik	Germany
X-Ray Automatic Processor (Curix 60)	Agfa	Germany

2.1.3 Chemicals and kits

The chemicals and kits used for this thesis are included in Table 2.2.

Table 2.2 Chemicals and kits used in present study

Chemicals and kits	Company	Location of Company
2-Mercaptoethanol	Gibco/ BRL	Eggenstein, Germany
1-bromo-3-chloropropane (BCIP)	Boehringer	Mannheim, Germany
4′,6-Diamidino-2-Phenylindole Dihydrochloride (DAPI)	Vector Labs	Burlingame, Germany
Anti-Actin antibody	Sigma-Aldrich	Deisenhofen,Germany
Anti-Chd4 antibody	Abcam	Cambridge, UK
Anti-E2-tag antibody (Mouse monoclonal)	Abcam	Cambridge, UK
Anti-E2-tag antibody (Rabbit polyclonal)	Lifespan	Eching, Germany
Anti-H3f3b antibody	Abcam	Cambridge, UK
Anti-H3K27me3 antibody	Millipore	Darmstadt, Germany
Anti-H3K4me3 antibody	Active Motif	La Hulpe, Belgium
Anti-H3K9ac antibody	Abcam	Cambridge, UK
Anti-H3K9me3 antibody	Millipore	Darmstadt, Germany
Anti-Kap1 antibody	Cell Signaling	Frankfurt, Germany
Anti-Myc-tag antibody	Millipore	Darmstadt, Germany
Anti-Nanog antibody	Abcam	Cambridge, UK
Anti-Oct4 antibody	Abcam	Cambridge, UK
Anti-p-H2A.X (Ser139) antibody	Cell Signaling	Frankfurt, Germany
Anti-p-p53 (Ser15) antibody	Cell Signaling	Frankfurt, Germany
Anti-Sox2 antibody	Abcam	Cambridge, UK
Anti-α-tubulin (Mouse monoclonal) antibody	Sigma-Aldrich	Deisenhofen,Germany
Anti-α-tubulin (Rabbit polyclonal) antibody	Sigma-Aldrich	Deisenhofen,Germany
Bacto-Tryptone	Difco	Detroit, USA
Bacto-Yeast-Extract	Difco	Detroit, USA

Table 2.2 Continued

Chemicals and kits	Company	Location of Company
Bisulfite Pyrosequencing Kit	Biotage	Uppsala, Sweden
BL21 (DE3) Competent Cells	Novagen	Darmstadt, Germany
Chemiluminescent Substrate	Pierce	Rockford, USA
Chloroform	Baker	Deventer, Netherlands
Collagenase	Sigma-Aldrich	Deisenhofen,Germany
Coomassie Blue G-250	Sigma-Aldrich	Deisenhofen, Germany
DH5α-Bacteial-Competent Cells	Invitrogen	Karlsruhe, Germany
Diethylpyrocarbonate (DEPC)	Sigma-Aldrich	Deisenhofen, Germany
Dropout (DO) Supplement	Clontech	Heidelberg, Germany
Dulbecco's-Modified-Eagles-Medium (DMEM)	PAN	Aidenbach, Germany
Endo Free Plasmid Maxi Kit	Qiagen	Hilden, Germany
Fetal Bovine Serum (FBS)	PAN	Aidenbach, Germany
Gelatin	Sigma-Aldrich	Deisenhofen, Germany
Goat anti-mouse IgG-alkaline phosphatase conjugate	Sigma-Aldrich	Deisenhofen,Germany
Goat anti-mouse IgG-Cy3 conjugate	Sigma-Aldrich	Deisenhofen, Germany
Goat anti-mouse IgG-FITC conjugate	Sigma-Aldrich	Deisenhofen, Germany
Goat anti-rabbit IgG-alkaline phosphatase conjugate	Sigma-Aldrich	Deisenhofen,Germany
Goat anti-rabbit IgG-Cy3 conjugate	Sigma-Aldrich	Deisenhofen, Germany
Goat anti-rabbit IgG-FITC conjugate	Sigma-Aldrich	Deisenhofen, Germany
GST•BindTM Kits	Novagen	Darmstadt, germany
Hybridisation Solution	Amersham	Freiburg, Germany
Hydrochloric Acid (HCl)	Merk	Darmstadt,Germany
Immunoprecipitation Kit (Protein A)	Roche	Penzberg, Germany
Leukemia Inhibitory Factor (LIF)	Chemicon	Temecula, USA
L-Gultamine	PAN	Aidenbach, Germany
Lipofectamine 2000	Invitrogen	Karlsruhe, Germany
Maxi Plasmid Kit	Qiagen	Hilden, Germany
Mega Plasmid Kit	Qiagen	Hilden, Germany
Midi Plasmid Kit	Invitrogen	Karlsruhe, Germany
Mitomycin C	Sigma-Aldrich	Deisenhofen, Germany
Nitro-blue Tetrazolium (NBT)	Roche	Penzberg, Germany
Nitrocellulose Membrane	Amersham	Freiburg, Germany
Non Essential Amino Acids (NEAA)	Invitrogen	Karlsruhe, Germany
NuPAGE MOPS SDS Running Buffer	Invitrogen	Karlsruhe, Germany
NuPAGE Novex Bis-Tris 4-12% Gels	Invitrogen	Karlsruhe, Germany
Nylon Membrane	Amersham	Freiburg, Germany
OneDay ChIP Kit	Diagendode	Liège, Belgium

Table 2.2 Continued

Chemicals and kits	Company	Location of Company
OptiMEM I	Invitrogen	Karlsruhe, Germany
Paraformaldehyde (PFA)	Merck,	Darmstadt,Germany
Penicillin/Streptomycin	PAN	Aidenbach, Germany
Peptone	Carl Roth	Karlsruhe, Germany
pGEM-T Vector System	Promega	Mannheim, Germany
Phosphate Buffered Saline (PBS)	PAN	Aidenbach, Germany
Platinum Taq DNA Polymerase	Invitrogen	Karlsruhe, Germany
Protein A/G PLUS Agarose	SantaCruz	Heidelberg, Germany
Proteinase K	Roche	Penzberg, Germany
Proximity Ligation Assay (PLA) kit	Bethyl	Hamburg, Germany
QIAprep Spin Miniprep Kit.	Qiagen	Hilden, Germany
QIAquick PCR Purification Kit	Qiagen	Hilden, Germany
Rapid Immunoprecipitation (RIPA) Buffer	Millipore	Darmstadt, Germany
Restriction Enzymes	Invitrogen	Karlsruhe, Germany
RNase A	Qiagen	Hilden, Germany
SeeBlue Plus2 Pre-Stained Standard	Invitrogen	Karlsruhe, Germany
Shearing ChIP kit - 100	Diagendode	Liège, Belgium
Sodium Chloride (NaCl)	Merck	Darmstadt,Germany
Sodium Pyruvate	Gibco/ BRL	Eggenstein, Germany
SuperScript II Reverse Transcriptase	Invitrogen	Karlsruhe, Germany
SYBR Green Master Mix	Qiagen	Hilden, Germany
T4 DNA Ligase	Promega	Mannheim, Germany
Triton X-100	Serva	Heidelberg, Germany
Trizol Reagent	Invitrogen	Karlsruhe, Germany
Trypsin	PAN	Aidenbach, Germany
Yeast extract	Carl Roth	Karlsruhe, Germany

2.1.4 Primers

The primers used in this study are listed in Table 2.3.

Table 2.3 Primers used for ChIP PCR, cloning, and genotyping

Name	Primer sequence (5′-3′)	Application
Actin chip F	CCCCAACACACCTAGCAAAT	ChIP-PCR
Actin chip R	ACTGCCCCATTCAATGTCTC	ChIP-PCR
Chd4 F	GAATTCCAGTGGAAAAGGCAGAGGTGA	Cloning
Chd4 R	GCGGCCGCTCACTGCTGCTGGGCTACCTGCTG	Cloning
ChipZfp819_NBRF1	CCGGGGTACTGGTTAGTTCA	ChIP-PCR
ChipZfp819_NBRF2	ATCTCCATGGGGAATATCCA	ChIP-PCR
ChipZfp819_NBRR1	CTGGCAAACACAGAAGTGGA	ChIP-PCR
ChipZfp819_NBRR2	AAGCAGAAATCTGCGTGTGA	ChIP-PCR
ChipZfp819_TBRF1	TTTCTAGGATCCAGCCCGTA	ChIP-PCR
ChipZfp819_TBRF2	CTCCATGACTGACTGGGGTTA	ChIP-PCR
ChipZfp819_TBRR1	AACCCCAGTCAGTCATGGAG	ChIP-PCR
ChipZfp819_TBRR2	TGAGACCAGCTTGGGCTACT	ChIP-PCR
GL3_IAP U3_F	GGTACCTTCAGTGTCCTAGTTCCCTTCC	Cloning
GL3_IAP U3_R	GCTAGCGCCTGGCTTACAGGTTCAGA	Cloning
H3f3b F	GAATTCAAATGGCCCGAACCAAGCAGA	Cloning
H3f3b R	GCGGCCGCTTAAGCTCTCTCCCCCCGTA	Cloning
Hnf4 qRT _F	CCACATGTACTCCTGCAGGTTTAG	qPCR
Hnf4 qRT _R	CGCTCATTTTGGACAGCTTC	qPCR
Hoxa11 chip F	AGAAGTGCCTCTGGCTCTGA	ChIP-PCR
Hoxa11 chip R	GATTTGCACGGTGACTTGATT	ChIP-PCR
Hprt F	AGCCCCAAAATGGTTAAGGTTGC	qPCR
Hprt R	TTGCAGATTCAACTTGCGCTCAT	qPCR
IAP qRT _F	AGCAGGTGAAGCCACTG	qPCR
IAP qRT _R	CTTGCCACACTTAGAGC	qPCR
IAP_5′UTR F	CGGGTCGCGGTAATAAAGGT	ChIP-PCR
IAP_5′UTR R	ACTCTCGTTCCCCAGCTGAA	ChIP-PCR
IAP_gag F	TTTCTTAAAATGCCCAGGCTTT	ChIP-PCR
IAP_gag R	CTTGCCCTTAAAGGTCTAAAGCA	ChIP-PCR
IAP_Poly F	CTTGCCCTTAAAGGTCTAAAGCA	ChIP-PCR
IAP_Poly R	GCGGTATAAGGTACAATTAAAAGATATGG	ChIP-PCR
IAP_U3 F	CGAGGGTGGTTCTCTACTCCA	ChIP-PCR
IAP_U3 R	GACGTGTCACTCCCTGATTGG	ChIP-PCR
KAP1_RBCC_F	GAATTCCTTCTGGAGCACTGCGGCGTGTGT	Cloning
KAP1_RBCC_R	CCCGGGCTAATCCACAATCATTTTGAGGGC	Cloning

Table 2.3 Continued

Name	Primer sequence (5′-3′)	Application
Klf4 qRT _F	TCAGGTACCCCTCTCTCTTCTTC	qPCR
Klf4 qRT _R	CGCTTCATGTGAGAGAGTTCCT	qPCR
Line1 qRT _F	TTTGGGACACAATGAAAGCA	qPCR
Line1 qRT _R	CTGCCGTCTACTCCTCTTGG	qPCR
Mest F	CAGCAGCTTCTGGCATGTGG	ChIP-PCR
Mest R	AACCCCAGATTCTAGTGAAG	ChIP-PCR
Nanog qRT _F	TTACAAGGGTCTGCTACTGAGTG	qPCR
Nanog qRT _R	CAGGACTTGAGAGCTTTTGTTTG	qPCR
Nestin qRT _F	AGCAGGAGAAGCAGGGTCTAC	qPCR
Nestin qRT _R	GCTGTCACAGGAGTCTCAAGG	qPCR
Oct4 chip F	TGGGCTGAAATACTGGGTTC	ChIP-PCR
Oct4 chip R	TTGAATGTTCGTGTGCCAAT	ChIP-PCR
Oct4 qRT _F	CGGAAGAGAAAGCGAACTAGC	qPCR
Oct4 qRT _R	GCCTCATACTCTTCTCGTTGG	qPCR
Overexpression E2 F	GTCGACGATGGCTGCTGACATGAATTTCTG	Cloning
Overexpression E2 R	GCGGCCGCTTAGCGATCTCTAAAATCAGAAGAA GTACTTGAACCACCAGGGTTCTCACCAGTGTGAG	Cloning
Zfp819_GST_N_F	CCATGGAGATGGCTGCTGACATGAATTTC	Cloning
Zfp819_GST_N_R	GCGGCCGCCAGGCTGGATGTACTGGGAAG	Cloning
Zfp819_GST_FL_F	CATGGAGATGGCTGCTGACATGAATTTC	Cloning
Zfp819_GST_FL_R	GCGGCCGCAGGGTTCTCACCAGTGTGAG	Cloning
Sine B1 qRT _F	GTGGCGCACGCCTTTAATC	qPCR
Sine B1 qRT _R	GACAGGGTTTCTCTGTGTAG	qPCR
Vimentin qRT _F	TGCAGTCATTCAGACAGGATGT	qPCR
Vimentin qRT _R	ATCTCTTCATCGTGCAGTTTCTTC	qPCR
Zfp206 qRT _F	GAGAGGAGGTGGTACAGCTATTG	qPCR
Zfp206 qRT _R	AGGTGGAGGTAACTCATTCAGTG	qPCR
Zfp819 qPCR F	GTATTCCAGGCCAGAGGTCA	qPCR
Zfp819 qPCR R	CAGCCTTCACACACATTTGG	qPCR
Zfp819_3MA_F	AAGCTTACTGTGGGGTGACTGAGGAG	Cloning
Zfp819_3MA_R	GGATCCGGCTGAAGAAAGCACCAGAA	Cloning
Zfp819_3ProbeF	CCCTCAGCAACCATTTATGC	Cloning
Zfp819_3ProbeR	GACAAACAATTTCCCGCAGT	Cloning
Zfp819_5MA_F	GCGGCCGCGCATGCACAAAATAGGCAGA	Cloning
Zfp819_5MA_R	AAGCTTGGGGTAGGGGGATAGCATTA	Cloning
Zfp819_5ProbeF	TCGGGAAAGAGGATGATGTC	Cloning
Zfp819_5ProbeR	CCTGAGCTACGGAACTGGAG	Cloning

Table 2.3 Continued

Name	Primer sequence (5′-3′)	Application
Zfp819_AA473_F	AGCTTCGTGTGGGAAGTCGTTCAGGAGAAAGTCCCA CCTTAAAGTGCATAGCCGAACACACACTGGTGAGAA GCCTTATGAATGCCCTGACTGCG	Cloning
Zfp819_AA501_R	GATCCGCAGTCAGGGCATTCATAAGGCTTCTCACCA GTGTGTGTTCGGCTATGCACTTTAAGGTGGGACTTTC TCCTG AACGACTTCCCACACGA	Cloning
Zfp819_Chip_F	TGCTTGCATAACTGGGAGTG	ChIP-PCR
Zfp819_Chip_R	CACCTGAACAAAGGCTCCTC	ChIP-PCR
Zfp819_CLSNeoF	GAGGCAGGGTTTCTCCGTGTAGCCCTGGCTGTCCTGGA ACTCACTCTGTAGCCCAATTCCGATCATATTC	Cloning
Zfp819_CLSNeoR	GGGATGAGTAAGTCAGTAGCTGATTACACCTCTCTGCTA	Cloning
Zfp819_E2geno_F	TCTACCTCCACGAATTCGGCCGCTCTAGAACTAGTGG CACCGGTGAGAAACCTTACAG	Genotyping
Zfp819_E2geno_R	TCTAAAATCAGAAGAAGTACT	Genotyping
Zfp819_GST_F	CCATGGAGATGGCTGCTGACATGAATTTC	Cloning
Zfp819_GST_R	GCGGCCGCCAGGCTGGATGTACTGGGAAG	Cloning
Zfp819_KRAB_F	GAATTCAGATTGGTGTCCTTTGAGGATG	Cloning
Zfp819_KRAB_R	CCCGGGCTAACACCTCTGATGTGGGAAATC	Cloning
Zfp819_NCL_NF	GTCGACCATGGCTGCTGACATGAATTTC	Cloning
Zfp819_NCL_NR	GCGGCCGCTTACAGGCTGGATGTACTGGGAAG	Cloning
Zfp819_NCL_CF	GTCGACCCCTCAGCAACCATTTATGC	Cloning
Zfp819_NCL_ CR	GCGGCCGCCTAAGGGTTCTCACCAGTGTG	Cloning
Zfp819_YH_ CF	CCATGGACCCTCAGCAACCATTTATGC	Cloning
Zfp819_YH_CR	CCCGGGCTAAGGGTTCTCACCAGTGTG	Cloning
Zfp819_YH_NF	CCATGGCAATGGCTGCTGACATGAATTTC	Cloning
Zfp819_YH_NR	CCCGGGTTACAGGCTGGATGTACTGGGAAG	Cloning
Zfp819NCL_C328F	AAGCTTCGCATGGATGTCCTGGCTGTCGG	Cloning
Zfp819NCL_C394R	GGATCCGCTAAAGGCCTTCCCACAGTC	Cloning
Zfp819NCL_C469F	AAGCTTCGAGATGCTGCGACTGTGGGAAG	Cloning
Zfp819NCL_C534R	GGATCCAGACATGGCCTTCCCACAGTC	Cloning
Zfp819TS3U0102F	TGTGTGCAGTTGTTTGCTCA	RT-PCR
Zfp819TS3U0102R	AAAGCCTTTCACAGGTCCAA	RT-PCR
Zfp819TS3U01R	CAAATTCTTTATTCTTTTCCT CAG GT	RT-PCR
Zfp819TS5U0103F	CCAAGACATTCATGCTGCTG	RT-PCR
Zfp819TS5U02F	GGATTCTGAACACTGGGGAAT	RT-PCR
Zfp819TS5U03F	GATGACTGAGGCAGCTGTAGG	RT-PCR

Table 2.3 Continued

Name	Primer sequence (5′-3′)	Application
Zfp819TS5UR	CACTCCTCCTGGCTGAAGTT	RT-PCR
Znf175qRT _F	GAGTGGAGATGCCTGCTGAT	qPCR
Znf175qRT_R	CACTCCTCCCTGCTGAAGTC	qPCR

2.1.5 Solutions, buffers, and culture media

Solution, buffers, and culture media are prepared as shown in Table 2.4.

Table 2.4 Solution, buffers, and culture media used in this study

Name	Compositions
Alkaline Phosphatase (AP)-Buffer	100 mM NaCl
	50 mM MgCl$_2$
	100 mM Tris/HCl (pH 9.5)
Blocking Buffer for Immunocytochemistry	0.2%Triton in PBS
Blocking Buffer for Western Blot	10 mM Tris-HCl (pH 7.3)
	137 mM NaCl
	0.1 % Tween 20
Denaturation Solution	1.5 M NaCl
	0.5 M NaOH
Depurination Solution	250 mM HCl
DNA Extraction Buffer	100 mM Tris-HCl (pH8.0)
	5 mM EDTA
	0.2% SDS
	200 mM NaCl
	100 μg/ml Proteinase K
Embryonic Stem (ES) Cell Culture Medium	DMEM
	20% FBS
	1 mM NEAA
	1 mM Sodium pyruvate
	10 μM 2-Mercaptoethanol
	2 mM L-Glutamine
	1000 U/ml LIF
	1 mM Penicillin/streptomycin
Fixation Buffer	4% PFA in PBS
Freezing Medium	25% FBS
	10% DMSO
	65% Culture medium

Table 2.4 Continued

Name	Compositions
Luria Bertani (LB) Medium for Bacteria	1 % (w/v) Tryptone
	0,5 % (w/v) Yeast Extract
	1 % (w/v) NaCl
Mouse Embryonic Fibroblasts (MEFs) Culture Medium	DMEM
	2 mM L-Glutamine
	10% FBS
	1% Penicillin/streptomycin
Neutralisation Solution	1.5 M NaCl
	1 M Tris/HCl (pH 7.0)
Protein Extraction Buffer	10 mM Tris–HCl (pH 8.0)
	1 mM EDTA
	2.5% SDS
	1 Protease inhibitor cocktail
	tablet/10 ml buffer
	100 mM PMSF
Protein Transfer Buffer (10x)	25 mM Tris-HCl (pH 8.3)
	150 mM Glycine
	20% Methanol
Saline Sodium Citrate (SSC) (20x) Solution	3 M NaCl
	0.3 M sodium citrate (pH 7.0)
Tris/Borate/EDTA (TBE) Buffer (5x)	450 mM Tris base
	450 mM Boric acid
	20 mM EDTA (pH 8.0)
Tris/EDTA (TE) Buffer	10 mM Tris/HCl (pH 8.0)
	1 mM EDTA
	20 g/L Difco peptone
Yeast Peptone Dextrose Adenine	10 g/L Yeast extract
(YPDA) Medium for Yeast	20 g/L Agar (for plates only)
	935 ml H_2O
	50 ml 40% glucose (to be added freshly)
	15 ml 0.2% Adenine sulfate (to be added freshly)
Yeast Standard (SD) Medium	1.7 g Yeast nitrogen base without amino acids
	5 g Ammonium sulfate
	600 mg DO Supplement
	20 g Agar (for plates only)
	950 ml H_2O (pH 5.8)
	50 ml 40% Glucose (to be added freshly)

2.2 Methods

2.2.1 Cell culture

The mouse ESC line (ES R1) with 129/Sv genetic background was maintained as described previously (Zechner et al., 2009). Briefly, the undifferentiated ESCs were cultured on gelatin-coated flask, plated with Mitomycin C-inactivated MEFs as feeder layer using ESC culture medium (see section 2.1.5). MEFs were cultured in MEFs culture medium (see section 2.1.5) until cells became confluent, then treated with 10 µg/ml Mitomycin C in MEFs culture medium. After 2-3 h of incubation, cells were washed twice with PBS and collected by trypsinization and centrifugation. Next, they were resuspended in MEFs culture medium and plated onto culture flasks, which were pretreated with 0.1% gelatin for 30 min. The feeder cells were left to attach by incubation overnight (O/N) at 37°C, 5% CO_2. Before plating ESCs on feeder layer, the medium was changed to ESC culture medium. NIH-3T3 cells obtained from Dr. S. A. Aaronson (Bethesda, U.S.A.), were cultured using MEFs culture medium.

2.2.2 Genomic DNA extraction, genotyping, and promoter methylation analysis

Mouse tail tips or cell pellets were incubated in 700 µl of DNA extraction buffer at 55°C O/N in thermomixer. The resulting lysates were centrifuged at 13,000 xg for 10 min and genomic DNA (gDNA) was purified from the supernatant by phenol chloroform extraction, followed by precipitation with isopropanol and washing the pellet with 70% ethanol. The gDNA pellet was then dissolved in 50 µl of dH_2O and used for either genotyping or promoter methylation analysis. For the latter experiment, EpiTect Bisulfite Kit was used for bisulfite treatment of gDNA, and then the methylation status of the promoter region of the gene of interest was analyzed by bisulfite pyrosequencing using PSQTM 96MA Pyrosequencing System with the PyroGold SQA reagent kit. Pyro Q-CpG software was applied for pyrosequencing result analysis (Zechner et al., 2009).

2.2.3 Total RNA isolation, cDNA synthesis, RT-PCR, and quantitative real time PCR

Total RNA was extracted from cells or tissues using Trizol Reagent following the manufacturer's protocol. In order to avoid any RNase activity, special RNase-free

Eppendorf cups and DEPC-treated dH_2O were used during the procedure. About 100 mg of tissue or $\sim10^6$-10^7 cells were homogenized in 1 ml of Trizol Reagent using a glass-teflon homogenizer. The homogenate was vortexed and incubated on ice for 5 min to permit the complete dissociation of nucleoprotein complexes. Then 0.2 ml of cold chloroform was added, mixed vigorously, kept on ice for 10 min. After centrifugation at 13,000 xg for 15 min at 4°C, the upper aqueous phase was transferred into a new Eppendorf cup. By adding 0.5 ml isopropanol, RNA was precipitated and subsequently washed with 75% ethanol. Finally, the RNA pellet was dissolved in ~50 μl of DEPC-dH_2O and stored at -80°C.

For cDNA synthesis, 5 μg of total RNA was digested with DNaseI to avoid any gDNA contamination and used in SuperScriptII Reverse Transcriptase system. 0.3 μl of cDNA was used for checking the cDNA synthesis quality by RT-PCR with housekeeping gene HPRT. For quantitative Real Time PCR (qPCR) analysis, the reaction composed of 1 μl diluted cDNA (1/20), 5 μl 2x SYBR Green Master Mix, 1 μl forward primer, 1 μl reverse primer, 2 μl dH_2O was run in ABI7900 Real-Time PCR System. Primers used in RT-PCR and qPCR are listed in section 2.1.4 Primers. The qPCR data were analyzed by the delta-delta-Ct method.

2.2.4 Protein extraction and Western blotting

For protein extraction, ~100 mg of mouse tissues or 1 x 107 cells were homogenized in 300-500 μl of protein extraction buffer, incubated on ice for 40 min and sonicated and followed by centrifugation at 13,000 xg for 20 min at 4°C. The protein extracts (40 μg) were resolved on 4-12% SDS-PAGE and transferred onto nitrocellulose membrane using semi dry blotting method at 150 mA for 15 min and 300 mA for 45 min. Next, the nitrocellulose membrane was blocked in 5% blocking buffer at room temperature (RT) for 1 h and then incubated with primary antibody for O/N at 4°C. For primary antibody, mouse monoclonal anti-myc tag (1:1500), mouse or rabbit monoclonal anti-E2-tag (1:1000), rabbit monoclonal anti-Kap1 (1:1000), mouse monoclonal anti-Chd4 (1:1000), and mouse monoclonal anti-α-tublin (1:10000) diluted in 2 % non-fat milk in blocking buffer were used. Next day, the membrane was washed three times for 10 min with 2 % non-fat milk in blocking buffer at RT and incubated for 1-2 h at RT with either alkaline phosphatase- or horseradish peroxidase-conjugated anti mouse/rabbit IgG secondary antibody (1:5000). Then, the membrane was washed three times for 10 min with 2 %

non-fat milk in blocking buffer at RT and one time for 5 min in PBS at RT. Finally, signals were visualized using either BCIP-NBT system or Chemiluminescence detection method.

2.2.5 Plasmid DNA transfection and immunocytochemistry

For transfection, approximately 0.5×10^5 NIH-3T3 cells were plated one day before transfection on cover slips in 24-well plates. For transfection into ESCs, 0.5×10^5 ESCs were plated 5 h before transfection. 1 µg of plasmid DNA and 2.5 µl of Lipofectamine 2000 reagent were diluted in 50 µl of OptiMEM, separately, and incubated at RT for 5 min. Subsequently, the diluted Lipofectamine and DNA were mixed together and incubated for additional 20 min at RT for allowing complex formation. Meanwhile, cells were washed twice with PBS and 400 µl of DMEM medium was applied in each well. After completion of the incubation, the DNA-Lipofectamine complex was added drop by drop to the cells and incubated for 3 h at 37°C .Then medium was changed back to normal MEF culture medium or ESC culture medium and the cells were further incubated for 24-48 h at 37°C and 5% CO_2.

For immunostaining, the cells were fixed with 4% PFA for 30 min at RT and then incubated with 50 mM NH_4Cl for 10 min. Later the cells were permeabilized by three times incubation with 0.2 % Triton in PBS for 4 min each at RT. 1:100 diluted primary antibodies were added followed by incubation for 1 h at RT. After washing, the secondary antibody conjugated with either FITC or Cy3 in PBS containing 0.2% Triton was added and incubated for additional 1 h. After incubation, cells were washed three times with 0.2 % Triton in PBS and were mounted with DAPI and proceeded to fluorescence microscopic analysis. Cells expressing EGFP fluorescence proteins were directly mounted with DAPI after fixation with 4% PFA.

For transfection in 6-well plate, 2×10^5 cells were used and transfected with 2 µg of plasmid DNA and 5 µl of Lipofectamine complex prepared in a total of 200 µl OptiMEM, after 24-48 h of transfection, cells were subjected to protein extraction and used either directly for Western blot analysis or for co-immunoprecipitation experiments.

2.2.6 Co-immunoprecipitation and Proximity Ligation Assay

For immunoprecipitation experiments, ESCs were transiently transfected with an appropriate construct using Lipofectamine 2000. After 48 h, proteins were isolated by

using RIPA buffer, followed by centrifugation at 13,000 xg, 4°C, 20min. Co-immunoprecipitation was done using the Immunoprecipitation Kit (Protein A) in accordance with the manufacturer's protocol. Briefly, to reduce the background caused by non-specific adsorption of irrelevant cellular proteins to protein A/G-agarose, a pre-clearing step was done by incubating 300 μg -500 μg of protein extract with 50 μl protein A/G-agarose suspension for 3 h. Then, the samples were centrifuged at 5,000 × g for 30 sec and the resulting supernatants were transferred into fresh tubes in which 6μg of appropriate antibody was added. Protein-antibody mixtures were incubated at 4°C for 2-3 h. After that, mixtures were incubated O/N at 4°C with 50 μl freshly prepared protein A/G-agarose beads. The next day, samples were centrifuged at 5,000 × g for 30 sec, and the beads were washed with washing buffer I, buffer II and buffer III for 20 min respectively. Finally, the beads were resuspended in 50 μl SDS-PAGE-loading buffer, and then eluted proteins were denatured and loaded onto 4-12% SDS-PAGE. Western blot analysis was performed with indicated antibodies.

In Proximity Ligation Assay (PLA), NIH-3T3 cells which were plated on cover slip in a 4-well plate were transiently transfected with indicated constructs. After 24 h of transfection, the cells were processed using the protocol provided with the PLA kit. Briefly, one drop of blocking solution was added onto the cover slip and the plate was incubated in a pre-heated humidity chamber for 30 min at 37°C. After that, the 1:100 diluted indicated primary antibodies in antibody diluent were added onto each well after tapping off the blocking solution and incubated at RT. After 1 h of incubation, cover slips were washed two times for 5 min each with wash buffer A, then the 1:5 diluted PLA probes PLUS and MINUS in antibody diluent were added into each well and incubated in a pre-heated humidity chamber for additional 1 h at 37°C. Next, the probe solution was removed followed by two times washing with wash buffer A for 5 min each. Subsequently, 40 μl Ligation-Ligase solution diluted in pure water was added onto cover slips and incubated in humidity chamber for 30 min at 37°C. After washing two times for 2 min each with wash buffer A, cover slips were exposed to 40 μl Amplification-Polymerase solution diluted in pure water and left in a pre-heated humidity chamber for 100 min at 37°C. Finally, the cover slips were washed two times for 10 min each with wash buffer B followed by an additional washing with 0.01 x wash buffer B. The coverslips were mounted by DAPI and analyzed by fluorescence microscope.

2.2.7 Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) assay was done according to the protocol provided with the Shearing Kit and OneDay ChIP Kit. Briefly, the wild type ESCs or ESCs transfected with hEF1α-Zfp819-E2 construct were trypsinized to get the cell pellet. For each preparation, 3×10^6 cells were used for shearing and subsequent chromatin preparation. After washing with PBS, cells were resuspended with 600 µl ESCs medium supplemented with 45 µl buffer A and 20 µl formaldehyde, incubated at RT for fixation and cross-linking. After 10 min, cross-linking was immediately quenched with the addition of 66 µl of 1.25 M glycine. Cells were lysed by 10 min incubation in lysis buffer B and C on ice, followed by centrifugation at 1,600 xg for 5 min at 4°C. Afterwards, the cell pellets were resuspended in buffer D containing protease inhibitors and sonicated (30 seconds "ON" / 30 seconds "OFF" and 30 cycles) using Bioruptor to obtain an average DNA fragment chromatin length of ~200-500 bp.

For immunoprecipitation, soluble chromatin was pre-cleaned by incubation with protein A/G sepharose beads for 3 h at 4°C. At the same time, freshly washed protein A/G sepharose beads were incubated with 5 µg antibody against E2, Oct4, Sox2, and Nanog, separately, for 1 h at 4°C to form the beads-antibody complex. As a negative control either rabbit or mouse IgG was used. After pre-cleaning, the samples were added into washed bead-antibody complex and kept for O/N at 4°C. Next day, beads complex was washed twice with 1 ml washing buffer and one more washing with 12 ml washing buffer in 15 ml falcon. Further, 100 µl of the provided DNA purifying slurry was used for purification coupled with boiling for de-crosslinking. Finally, 100 µl ChIP grade water was applied to elute target DNA and 2 µl of this DNA was used in a 10 µl final q-PCR reaction.

2.2.8 Yeast-two hybrid screening and direct-yeast-two hybrid analysis

To identify the interaction partners of Zfp819, we constructed a bait containing the N-terminal part (a.a. 1-343 corresponding to nucleotides (nt) 228-1256 of NM_028913.3) of Zfp819 (Zfp819_N) and screened the previously generated ESC cDNA library (Zheng et al., 2012) in a Yeast Two-Hybrid (Y2H) screening method. The GAL4-BD fusion bait construct of *Zfp819* was prepared by cloning a PCR fragment of Zfp819_N into pGBKT7 vector resulting in pGBKT7-Zfp819_N. After excluding the autoactivation by co-

transformation of pGBKT7-Zfp819_N with empty pGADT7 vector into AH109 yeast strain, the bait construct was transformed into Y187 yeast strain. The Y2H screen was performed on ESC cDNA library using Matchmaker pre-transformed library protocol (Clontech). Briefly, the pre-transformed ESC library in yeast strain AH109 was mixed and mated together with strain Y187 containing the pGBKT7-Zfp819_N. After 24 h of mating, the culture was spread on SD/-Leu/-Trp/-His/-Ade plates and the surviving colonies were verified on SD/-Leu/-Trp/-His/-Ade/+ X-α-Gal plates. The positive clones that were blue on X-α-Gal were inoculated for O/N culture and the plasmid DNA was isolated using QIAprep Spin Miniprep Kit. By PCR amplification and sequencing using vector specific primers, we obtained the sequences of cDNA inserts. Identities of prey cDNA clones were determined by BLAST analysis.

To test the auto-activation or interaction, the prey cDNA clones were co-transformed with either empty pGBKT7 vector or pGBKT7-Zfp819_N into AH109 strain, respectively. The co-transformants were first selected on SD/-Leu/-Trp plates and later tested for the reporter gene expression on SD/-Leu/-Trp/-His/-Ade and X-α-Gal plates. The surviving clones without autoactivation were identified as putative interaction partners of Zfp819 and characterized further.

2.2.9 Recombinant GST-Zfp819_N protein generation and GST-pull down assay

2.2.9.1 Generation of pET-41a-Zfp819_N construct and protein purification

The N-terminal region (a.a. 1-343 corresponding to nucleotides (nt) 228-1256 of NM_028913.3) of Zfp819 (Zfp819_N) was PCR amplified using Zfp819_GST_F Zfp819_GST_R primers and cloned in-frame to the downstream part of GST-tag in pET-41a (+) vector to generate pET-41a-Zfp819_N.

The recombinant construct as well as pET-41a alone were transformed separately into BL21competent cells to induce and purify GST-Zfp819_N and GST-only, respectively. Firstly, a single bacterial colony was inoculated into 20 ml of LB medium containing 50 µg/ml Kanamycin and incubated at 37°C with shaking. The next morning, 5 ml O/N culture was diluted into 500 ml LB medium containing 50 µg/ml Kanamycin and cultured further at 37°C until the OD600 reached to ~0.6-0.8. Then 1 mM IPTG was added to induce the protein expression and incubated at 30°C for additional 5 h.

Cells were harvested by centrifugation at 5,000 x g for 15 min at 4°C. Subsequently, recombinant GST-Zfp819_N and GST-only were purified from bacterial cell extracts using GST•Bind Kit. Briefly, cells were harvested and resuspended in 20 ml of cold 1x GST Bind/Wash buffer and sonicated three times, 4 min each with an interval of 2 min to avoid the heat. The cell debris was removed by centrifugation at 20,000 x g for 20 min at 4°C and the supernatant was filtered through a 0.45 μm membrane to prevent clogging of the resin. The resulting supernatant was applied onto a column made up of 2 ml washed glutathione-agarose resin and incubated for 1 h at 4°C. The bound resin was washed by washing with 10 ml 1x GST Bind/Wash buffer to remove the non-bound protein. Thereafter, the bound protein was eluted with 5 ml of 1x GST Elution buffer and collected as five 1 ml fractions. The elution sample (20 μl) from each fraction as well as input and flow through were resolved on 4-12% SDS-PAGE and stained with Coomassie blue staining solution. Purified protein was dialyzed O/N at 4°C against 1x GST Bind/Wash buffer prior to performing GST-pull down assay.

2.2.9.2 GST-pull down assay

Glutathione-agarose beads were mixed gently by inverting several times and 50 μl of beads for each reaction was transferred to a clean microcentrifuge tube followed by three times washing with 500 μl of 1x GST Bind/Wash buffer and centrifugation at 2,000 rpm for 2 min. The supernatant was removed and beads were resuspended in 100 μl 1x GST Bind/Wash buffer. Then, 30 μg of either GST-Zfp819_N fusion protein or GST-only were added to Glutathione-agarose beads and incubated at RT for 1 h followed by centrifugation at 2,000 rpm for 2 min and two times washing with 1x GST Bind/Wash buffer. At the meantime, total protein extracts were prepared from ESCs using RIPA buffer and 50μl of Glutathione-agarose beads were used for pre-cleaning of protein extracts by incubation at 4°C. After 3 h of incubation, the beads were centrifuged at 2,000 rpm for 2 min and the resulting pre-cleaned protein extract was added to the centrifuge tube containing the bead-protein complex and the mixture was incubated O/N at 4°C on the roller. The next morning, beads were washed three times with 1 ml of 1x GST Bind/Wash buffer. Finally, the beads were resuspended in 2x GST Elution buffer, centrifuged and subjected to Western blot analysis for detecting either Chd4 or Kap1 endogenous protein.

2.2.10 Southern blotting

After electrophoresis of restriction enzyme digested gDNA, the gel was treated with 0.25 M HCl for depurination for 7 min, with denaturation solution for 25 min and neutralisation for 20 min. The transfer of the DNA to the nylone membrane was done in a Turbo-Blot apparatus by capillary flow. Twenty six Whatman filter papers were layered on a Stack Tray, followed by 1 Whatman filter paper soaked in 2 x SSC. Then the equilibrated nitrocellulose membrane that was soaked in 2 x SSC was laid on the top. The treated agarose gel was placed on the membrane and covered by 2 x SSC soaked Whatman filter paper and 1 dry Whatman paper. Then the buffer tray was placed and filled with 2 x SSC. Finally a wick cover was put on top of the blot. The transfer was carried out for O/N. The next morning, DNA on the membrane was fixed by UV crosslinking in UV stratalinker 1800.

Prehybridization was done in 12 ml of Rapid-hyb buffer with sheared denaturated salmon sperm DNA and incubated for 1 h in the hybridisation oven at 65°C. Then, the labelled probe was denaturated at 95°C for 10 min, chilled on ice for 5 min and added to the hybridisation solution. The hybridisation was carried out overnight in the oven. Next day, the membrane was washed for 10 min with 2 x SSC, then with 2 x SSC containing 0.2% SDS at 65°C for 10-20 min. Finally, the membrane was washed with 0.2 x SSC containing 0.1 % SDS at the hybridisation temperature. After drying the membrane, it was sealed in a cassette and exposed to autoradiography overnight at -80°C. The film was developed in X-Ray Automatic Processor Curix 60.

2.2.11 Computer analyses

For the analysis of nucleotide sequences, programs like BLAST, BLAST2 and other programs from National Centre for Biotechnology Information (NCBI) were used (www.ncbi.nlm.nih.gov). For restriction analysis of DNA, NEB cutter V2.0 or Webcutter program were used (http://tools.neb.com/NEBcutter2/index.php; http://rna.lundberg.gu.se/cutter2). For primer design, Primer 3 software was used (http://Frodo.wi.mit.edu). ExPASy tools were used for proteins and their domains information (www.Expasy.org). Mouse genome sequence and other information about mouse genes were referenced from Ensembl (www.ensembl.org/index.html).

3 Results

Ever since their first derivation in 1981, pluripotent stem cells have attracted biologists as well as physicians because of their potential in regenerative medicine applications. At the molecular level, pluripotency is established and maintained by both known as well as unknown regulatory mechanisms. Our previous efforts to identify unknown factors which might play a crucial role in establishment and maintenance of pluripotency uncovered *Zfp819* as a putative pluripotency-related gene. In the present thesis, we endeavored to identify the function of *Zfp819*, a KRAB-zinc finger protein, in pluripotent stem cells as well as in mouse embryonic development. Most results of this thesis are included in the two following manuscripts, whereas unpublished results are included in the discussion part.

The following result sections contain a brief description of the aim and results of the study in context of the complete thesis, the status of each manuscript as well as authors contributions.

3.1 Generation and characterization of yeast two-hybrid cDNA libraries derived from two distinct mouse pluripotent cell types

Identification of protein-protein interaction networks will help us to dissect the pluripotency and differentiation potential of pluripotent cells, hence their efficient use in clinical applications. In the first part of this thesis, we developed two yeast two-hybrid (Y2H) cDNA libraries to identify potential interaction partners of a protein of interest. We used these libraries to identify interaction partners of Stra8 and Zfp819, which are highly expressed in pluripotent stem cells, yet the molecular function is unknown. These screens led us to identify several putative interaction proteins of both Stra8 and Zfp819, where some of these interactions were confirmed using both *in vitro* and *in vivo* approaches. Gene ontology (GO) categorization of putative Stra8 interaction proteins led us to suggest that *Stra8* might function in chromatin organization and transcription regulation. Similarly, Zfp819 interaction proteins categorization revealed that *Zfp819* might play a role in transcription regulation, cell cycle, and apoptosis regulation as well as in telomere maintenance. Collectively, our Y2H libraries are a useful tool to identify potential protein interaction networks and their possible function in pluripotent stem cells.

Authors: Ying Zheng*, **Xiaoying Tan***, Joanna Pyczek, Jessica Nolte, D.V. Krishna Pantakani and Wolfgang Engel.

* These authors contributed equally to this work.

Status: Published in *Molecular Biotechnology* journal, DOI10.1007/s12033-012-9561-4.

Impact Factor: 2.17

Authors contributions:

1. Ying Zheng: Conception and design, generation of cDNA libraries, data collection, data analysis, data interpretation.

2. Xiaoying Tan: Data collection, data analysis, data interpretation, helped in drafting the manuscript.

3. Joanna Pyczek: Data collection.

4. Jessica Nolte: Conception and design, data interpretation.

5. D.V. Krishna Pantakani: Conception and design, discussion, data interpretation, drafted the manuscript.

6. Wolfgang Engel: Conception and design, discussion, financial support, helped in drafting the manuscript.

RESEARCH

Generation and Characterization of Yeast Two-Hybrid cDNA Libraries Derived From Two Distinct Mouse Pluripotent Cell Types

Ying Zheng · Xiaoying Tan · Joanna Pyczek · Jessica Nolte · D. V. Krishna Pantakani · Wolfgang Engel

Abstract Pluripotent stem cells have the therapeutic potential in future regenerative medicine applications. Therefore, it is highly important to understand the molecular mechanisms governing the pluripotency and differentiation potential of these cells. Our current knowledge of pluripotent cells is largely limited owing to the candidate gene/protein approach rather than studying the complex interactions of the proteins. Experimentally, yeast two-hybrid system (Y2H) is by far the most useful and widely used method to detect the protein–protein interactions in high-throughput screenings. Unfortunately, currently there is no GAL4-based pluripotent stem cell-specific cDNA library available for screening the interaction proteins impeding the large-scale studies. In this study, we report the construction of Y2H cDNA libraries derived from mouse pluripotent embryonic stem cells (ESCs) and multipotent adult germ-line stem cells (maGSCs) in GAL4-based Y2H vector system with very high transformation efficiency. Furthermore, we have constructed two different baits and screened for interaction partners in an effort to characterize the libraries and also as a part of our ongoing studies. Consequently, many putative interaction proteins were identified in both cases and their interaction was further validated by direct-Y2H. The observed interactions between bait proteins and their respective analyzed putative interaction proteins were further confirmed using two independent approaches in mammalian cells, thus highlighting the biological significance of the identified interactor (s). Finally, we would like to make these cDNA libraries as a resource that can be distributed to the research community.

Keywords Pluripotent stem cells · Protein–protein interactions · Yeast two-hybrid cDNA library · ESC · maGSC

Ying Zheng and Xiaoying Tan contributed equally to this work.

Y. Zheng · X. Tan · J. Pyczek · J. Nolte ·
D. V. K. Pantakani (✉) · W. Engel
Institute of Human Genetics, University of Goettingen,
Heinrich-Dueker-Weg 12, 37073 Goettingen, Germany
e-mail: krishna.if1@gmail.com; pkrishn@gwdg.de

Y. Zheng
e-mail: yzzkl@163.com

X. Tan
e-mail: mytanxy@yahoo.com.cn

J. Pyczek
e-mail: joanna.pyczek@uj.edu.pl

J. Nolte
e-mail: jnolte1@gwdg.de

W. Engel
e-mail: wengel@gwdg.de

Y. Zheng
Department of Histology and Embryology, Medical College
of Yangzhou University, 11 Huaihai Road, Yangzhou 225001,
Jiangsu, China

Background

Embryonic stem cells (ESCs) derived from inner cell mass of the pre-implantation stage embryos are pluripotent and have the competence to differentiate into all the germ layers including germ cells [1, 2]. Human ESCs hold the promise for future regenerative medicine therapies as they can self-renew without loosing the pluripotency and have the potential to differentiate into all the cell types of the body [3] albeit the ethical issues associated with ESCs derivation from embryo. On the other hand, multipotent

Humana Press

adult germ-line stem cells (maGSCs) which are generated from adult mouse testis are as pluripotent as ESCs [4–8], and if successful, the human maGSCs hold great potential in regenerative medicine applications and can bypass the ethical issues associated with ESCs. However, it is important to fully understand how the pluripotency is established and maintained and how the differentiation is initiated and maturated to a desired cell-type(s) before we proceed with any clinical applications.

Previous efforts to understand the pluripotency of ESCs at molecular level have uncovered genes such as Oct3/4, Nanog, Sox2, Rex1, and Sall4 as important regulators of pluripotency [9–12]. However, our knowledge on pluripotency of ESCs is largely limited owing to the candidate gene/protein approach rather than studying the complex interactions of the protein, as the function of a specific protein may depend on its interacting protein. Therefore, studies aimed at understanding the protein–protein interaction networks (protein interactomes) of pluripotent cells have identified some important regulatory networks implicated in the pluripotency [13, 14]. Experimentally, protein–protein interactions can be detected using a variety of techniques such as yeast two-hybrid system (Y2H), immunoprecipitation-coupled mass spectrometry (IP-MS), protein microarrays, synthetic lethality, targeted releasable affinity probe (TRAP), stable isotope labeling by amino acids in cells (SILAC). Among the above mentioned, Y2H system allows the cost-effective and genomic-scale screening for protein–protein interactions in a relatively short period of time. This approach, which relies on the activation of downstream reporter genes by the GAL4-based system [15, 16], in which the transcriptional activator GAL4 is split into DNA-binding domain (BD) and activation domain (AD) and fused to bait and prey, respectively. The interaction between bait and prey proteins bring the GAL4 domains into close proximity to each other and lead to the transcriptional activation of reporter genes. Based on these properties, even weak and transient interactions, which are difficult to detect in immunoprecipitation based experiments can be detected easily.

Unfortunately, currently there is no pluripotent stem cell-specific cDNA library available for screening the interaction proteins impeding the large-scale studies. Here, we report the construction of high quality cDNA libraries of mouse ESCs and maGSCs in GAL4-based Y2H vector system. Furthermore, we have constructed two different baits [Zinc finger protein 819 (Zfp819) and Stimulated by retinoic acid (Stra8)] and screened for interaction partners in an effort to characterize the libraries and also as a part of our ongoing studies. Consequently, many interaction partners were identified in both cases and the putative interaction partners were validated by direct-Y2H. Further, one interaction partner from each screen was validated in

mammalian cell culture system using co-immunoprecipitation (Co-IP) or glutathione S-transferase (GST) pull-down assay, and co-localization methods, highlighting the biological significance of our Y2H screen.

Results and Discussion

Generation of Mouse ESCs and maGSCs Y2H cDNA Libraries

Although many different techniques are available to analyze protein–protein interactions, only two techniques, IP-MS and Y2H screen are widely used in high-throughput screening applications. The use of IP-MS allows the identification of most of the components of a large protein complex, whereas Y2H identifies interactions between two individual proteins and their minimal interacting domains. For high-throughput protein interaction analyses, the Y2H screen is by far the most useful and widely used as it is based on the in vivo genetic screening approach and has the potential to identify low affinity as well as transient protein interactions [17]. In this study, we have successfully created Y2H cDNA libraries derived from mouse ESCs and maGSCs. ESCs, the "gold standard of pluripotency" are being studied extensively at the molecular level to understand the pluripotent cell characteristics and their potential in clinical applications. While maGSCs are a recently derived cell type and possess all the analyzed pluripotent cell characteristics [4–8], hence might be an alternative to ESCs to avoid ethical and immunological concerns.

Currently, there are no available cDNA libraries of mouse pluripotent cells for analyzing the protein–protein interactions in GAL4-based Y2H. To overcome this, we have created Y2H cDNA libraries from high quality polyA+ mRNA derived from mouse ESCs as well as from maGSCs using Oligo-(dT) priming method (Fig. 1). After first-strand cDNA synthesis, we obtained good smear from both ESCs and maGSCs which was comparable to the human placenta polyA+ mRNA, a positive control provided in the Matchmaker Library Construction and Screening Kits (Fig. 1a). Further, ds-cDNA (double strand-cDNA) was prepared using SMART technology and was cloned into pGADT7-Rec vector with the help of recombination in yeast AH109. The transformation efficiency was higher than the expected ($\geq 1.0 \times 10^6$) with $\sim 2.3 \times 10^6$ and $\sim 2.4 \times 10^6$ transformants for ESCs and maGSCs, respectively (Table 1). We then performed colony PCR on ~ 100 randomly picked transformants using vector specific primers to check the insert size and to analyze the recombination efficiency (Fig. 1b). The PCR analysis of the ESC cDNA library revealed the insert ranging from ~ 0.2 to ~ 2.0 kb with an average insert size

Mol Biotechnol

Fig. 1 Construction and verification of Y2H cDNA libraries of mouse ESCs and maGSCs. **a** Agarose gel electrophoresis showing the smear of first strand cDNA synthesized from polyA+ mRNA isolated from mouse ESCs and maGSCs using oligo-(dT) priming method. The polyA+ mRNA from human placenta was used as a positive control. **b** Colony PCR amplification on randomly picked yeast recombinant clones from mouse ESCs and maGSCs Y2H cDNA libraries. The molecular weights of 1 kb DNA ladder are indicated in both **a** and **b**
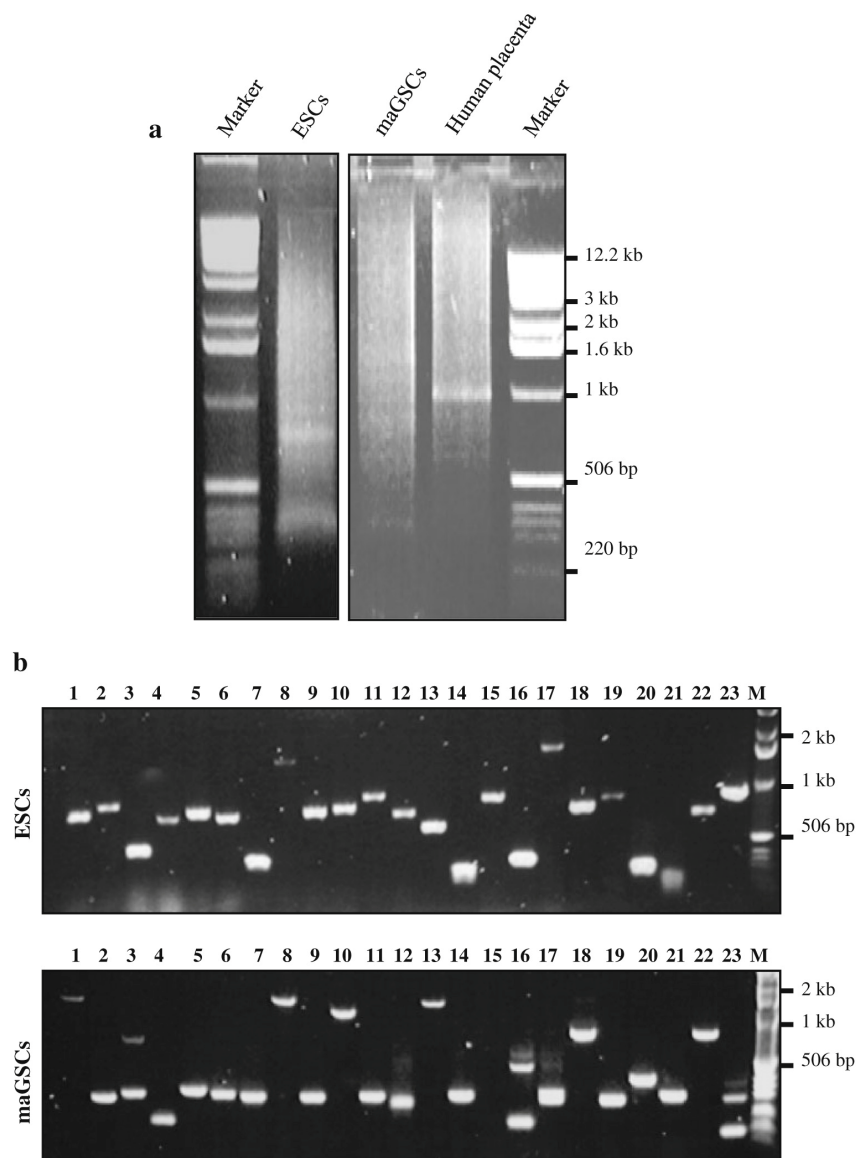
of ∼0.8 kb (Table 1). Similarly, maGSC library was found to contain inserts ranging from ∼0.1 to ∼2.0 kb with an average insert size of ∼0.6 kb (Table 1). The observed average insert size of ∼0.6 to ∼0.8 kb of the prey clone will minimize the protein interaction domain. The percentage of positive recombinant clones was ∼90 and ∼94 % in ESC and maGSC libraries, respectively (Table 1). The remaining recombinant clones might have a longer insert size, hence could not be amplified. We then prepared frozen stocks of ESCs and maGSCs with a cell

density of $\sim 8 \times 10^7$ and $\sim 7 \times 10^7$ cells/ml, respectively (Table 1). Furthermore, we estimated the titer of the libraries as $\sim 4 \times 10^7$ and $\sim 3.6 \times 10^7$ cfu/ml for ESCs and maGSCs libraries, respectively (Table 1).

Characterization of Y2H cDNA Libraries

The quality of both ESC and maGSC libraries was evaluated by screening interaction proteins for Zfp819 and Stra8, respectively. *Zfp819* was identified in a comparative

Table 1 The mouse ESCs and maGSCs Y2H cDNA libraries transformation efficiency, insert size and, the quality of libraries

	ESCs cDNA library	maGSCs cDNA library
Transformation efficiency (expected $\geq 1 \times 10^6$ transformants/3 μg pGADT7-Rec)	2.3×10^6	2.4×10^6
Insert size (kb)		
Minimum	0.2 kb	0.1 kb
Maximum	2.0 kb	2.0 kb
Average	0.8 kb	0.6 kb
% of positive recombinant clones	90	94
Cell density of frozen library (cells/ml)	8×10^7	7×10^7
cDNA library titer (cfu/ml)*	4×10^7	3.6×10^7

* cfu/ml – colony forming units/ml

transcriptome analysis of undifferentiated and differentiated ESCs as a novel gene which is expressed highly in pluripotent cells but not in their differentiated counterparts ([6] and unpublished data). The protein encoded by *Zfp819* belongs to C2H2-zinc finger (C2H2-Znf) family of proteins and bears a functional KRAB (Krueppel-associated box) domain on its N-terminal region, yet the function is not known. Recently, *Zfp819* was shown to be highly expressed in partially and fully reprogrammed induced pluripotent cells (iPSCs), but not in parental somatic cells [18]. Collectively, these results suggest a possible crucial role for Zfp819 in establishment and maintenance of pluripotency. To elucidate the function of Zfp819 in pluripotent cells, we screened ESC cDNA library with N-terminal region of Zfp819 (Zfp819_N) as a bait. This screen yielded a total of ∼800 colonies on high stringency nutritional selection plates (SD/-Leu/-Trp/-His/-Ade), of which 180 candidates were analyzed by sequencing, resulting in BLAST hits for 150 prey clones (Table 2). Further analysis revealed that 46 % of clones are in-frame of the target genes, of which 64 % of clones were identified as putative interaction proteins, whereas the remaining showed auto-activation in direct-Y2H assay. The putative interaction partners of Zfp819 were further categorized based on the GO (gene ontology) term biological process (Table 3)

indicating that Zfp819 might function as a transcriptional and cell cycle/apoptosis regulator.

In order to validate the authenticity of our Y2H screen, we performed protein interaction studies between Zfp819 and its putative interaction partner, Chromodomain helicase DNA-binding protein 4 (Chd4), using direct-Y2H and mammalian cell culture system (Fig. 2). Chd4 is a component of NuRD chromatin remodeling complex, which functions in translational repression by histone deacetylation [19]. We have identified the C-terminal region (aa 1,658–1,915) of Chd4 as a prey in our Y2H screen, narrowing down the Chd4 interaction region with Zfp819 (Fig. 2a). Co-transformation of purified prey Chd4 clone and bait, Zfp819, confirmed the interaction between Chd4 and Zfp819 in a direct-Y2H assay and showed no auto-activation when Chd4 was co-transformed together with empty bait vector (Fig. 2b and data not shown). To perform GST pull-down assay, we purified the GST-Zfp819_N fusion protein to the near homogeneity (Fig. 2c). We also purified GST protein alone to use as a negative control (Fig. 2c). The GST pull-down assay with cell extracts from ESCs and the subsequent western blot analysis with Chd4 specific antibodies confirmed the interaction between Zfp819 and Chd4 (Fig. 2d). Co-localization studies with E2-Zfp819 and endogenous Chd4 revealed that both proteins partially co-localize in the nucleus as discrete spots (Fig. 2e).

On the other hand, we used Stra8 as bait to screen maGSCs library. Stra8 is a retinoic acid responsive gene essential for meiosis, but is also known to be expressed in pluripotent cells, yet the molecular function is unknown [20–24]. Screening of maGSCs cDNA library with glutamic acid (GA)-rich region of Stra8 (Stra8^{GA}) resulted in a total of ∼300 colonies on high stringency nutritional selection plates. Out of 300 positive clones, 81 were analyzed by sequencing, resulting in a BLAST hits for 76 prey clones (Table 2). In contrast to ESCs screen, maGSCs screen resulted in 26 % of in-frame clones; while 37 % clones contained 3′ UTR sequences (Table 2). The direct-Y2H analysis of in-frame clones identified 40 % as putative interaction proteins, while the remaining 60 % showed auto-activation. The GO analysis of Stra8 putative interaction partners revealed that Stra8 might function in chromatin assembly/modification and transcription

Table 2 Y2H screening of mouse ESCs and maGSCs cDNA libraries with bait proteins as a proof-of-concept

	No. of positive clones	Analyzed sequences	Sequences with BLAST hit	In-frame (%)	Out-of-frame (%)	Minus (%)	3′ UTR (%)
ESCs library (Bait: Zfp819_N)	800	180	150	46.0	35.4	1.3	17.3
maGSCs library (Bait: Stra8^{GA})	300	81	76	26.0	37.0	0.0	37.0

Table 3 List of Zfp819 putative interaction proteins identified in a Y2H screen

Clone no.	NCBI accession No.	Gene/protein description	Gene ontology
174, 362	NM_145979	Chromodomain helicase DNA-binding protein 4 (Chd4)	Chromatin modification, regulation of transcription
270, 298, 313, 326, 372, 546, 548, 691	NM_008211	H3 histone, family 3B (H3f3b)	Nucleosome assembly
201, 532	NM_025828	Lectin, mannose-binding 2 (Lman2)	Protein transport
334, 354, 711, 720, 749	NM_008143	Guanine nucleotide binding protein (G protein), beta polypeptide 2 like 1 (Gnb2l 1)	Positive regulation of protein phosphorylation, negative regulation of translation
524	NM_010329	Podoplanin (Pdpn)	Cell morphogenesis
559	NM_012342	BMP and activin membrane-bound inhibitor, homolog (Bambi)	Positive regulation of cell proliferation, regulation of cell shape
314	NM_028388	NADH dehydrogenase (ubiquinone) flavoprotein 2 (Ndufv2)	Mitochondrial electron transport, NADH to ubiquinone
315	NM_023202	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 7 (Ndufa7)	Mitochondrial electron transport, NADH to ubiquinone
318	NM_027204	Mitochondrial ribosomal protein L12 (Mrpl12)	Transcription from mitochondrail promoter, translation
330	NM_011157	Serglycin (Srgn)	Apoptosis, platelet degradation
194	NM_026274	Ring finger and SPRY domain containing 1 (Rspry1)	NA
196, 285, 265, 264, 236, 252, 271, 276, 700, 729	NM_010860	Myosin, light polypeptide 6, alkali, smooth muscle and non-muscle (Myl6)	Skeletal muscle tissue development, muscle contraction
356	NM_013535	Gene rich cluster, C10 gene (Grcc10)	NA
267, 302, 360	NM_001143790	RIKEN cDNA 1500010J02 gene	Positive regulation of DNA replication, telomere maintenance
561	NM_025849	RIKEN cDNA 3110001D03 gene	NA
368, 550, 568	NM_026566	RIKEN cDNA9430023L20 gene	Autophagy
286, 248	NR_015585	RIKEN cDNA 4933439C10 gene	NA

regulation processes (Table 4). To validate the Stra8 Y2H screen, we characterized the interaction between Stra8 and its putative interaction partner, AT-rich interactive domain 4B (Arid4B), in both ex vivo and in vitro studies (Fig. 3). The C-terminal region (aa 1,119–1,314) consisting of coiled–coiled (CC) domain of Arid4B has been identified as a prey, indicating that this region is sufficient to mediate the interaction with Stra8 (Fig. 3a). The direct-Y2H assay further confirmed the interaction of Arid4B with Stra8 and showed no auto-activation (Fig. 3b and data not shown). Transient overexpression of c-Myc-tagged Stra8 (c-Myc-Stra8) and HA-tagged Arid4B (HA-Arid4B), and subsequent Co-IP studies revealed the interaction between these two proteins (Fig. 3c). Further, overexpression and immunostaining studies revealed that both Stra8 and Arid4B co-localize in the nucleus with diffused pattern (Fig. 3d).

Generally, Y2H screen is prone to result in detection of false positives, albeit at low rate, hence the relevance of the identified putative interaction proteins in a physiological context has to be validated using one or more independent

interaction methods [25]. The confirmation of Chd4 and Arid4b interaction with Zfp819 and Stra8, respectively, strongly suggests that the other identified interaction partners of our screens are potentially to be biologically significant, but have to be verified using independent methods. Previously, several studies have successfully used Y2H screening method to detect physiologically relevant interaction partners for DNA-binding proteins, including zinc finger proteins [26–29]. The study by Kalenik et al. [27] has successfully used zinc finger protein YY1, which is implicated in the negative regulation of myogenic differentiation, in Y2H screen and identified interaction proteins such as YY1-associated factor 2 (YAF2). The interaction between YY1 and YAF2 was confirmed through several independent methods employing mammalian cell culture system and further studies revealed that YAF2 binds to YY1 and enhances the proteolytic cleavage of this factor during myogenic differentiation [27]. Likewise, Rodel et al. [28] have used Gfi-1, a zinc finger protein with a dominant oncogene function, to screen for interaction proteins using Y2H assay and could identify PIAS3, an
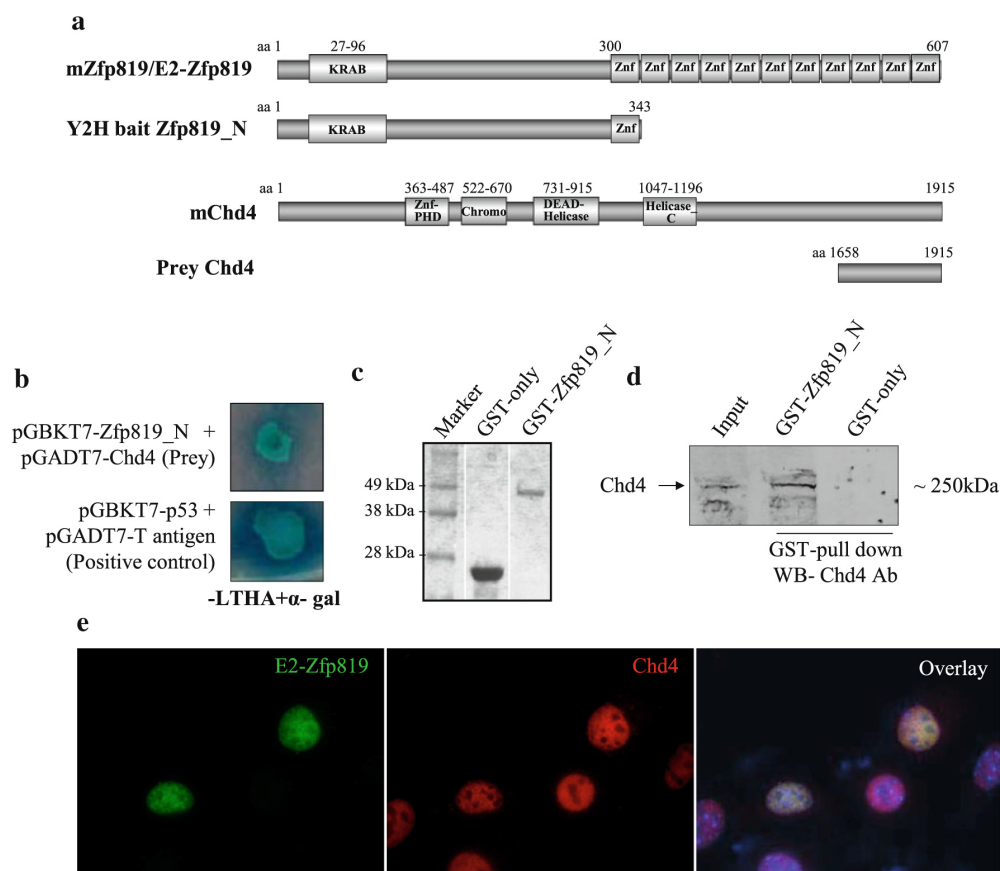
Fig. 2 Zfp819 interacts with Chd4 in both ex vivo and in vivo. **a** Illustration of mouseZfp819 and Chd4 proteins together with Y2H/mammalian expression constructs used for the analysis. Zfp819 contains a kruppel-associated box (KRAB) domain at the N-terminus and 11 tandemly arranged zinc finger (Znf) motifs at the C-terminus of the protein. The Chd4 contains a Znf-plant homeodomain (Znf-PHD), a chromatin organization modifier (Chromo) domain and two centrally located helicase domain types, DEAD-helicase (DEAD) and helicase carboxyl-terminal domain (helicase_C). **b** Confirmation of Chd4 interaction with Zfp819 using direct-Y2H. The AH109 yeast cells transformed with bait and prey plasmids, as indicated, were selected on nutritional selection medium, SD-LTHA and for α-galactosidase (α-gal) activity. As a positive control, the interaction between p53 and T-antigen was assayed in direct-Y2H experiments. **c** Coomassie blue stained SDS-PAGE gel showing the homogeneity of purified GST-only or GST-Zfp819_N fusion proteins. **d** GST pull-down assay and subsequent Western blot analysis with Chd4 antibodies confirmed the interaction between Zfp819 and endogenous Chd4. **e** IFC analysis of NIH-3T3 cells transiently transfected with E2-Zfp819 (*green*) showed co-localization of Zfp819 with endogenous Chd4 (*red*) in the nucleus. The cells were counterstained with DAPI to visualize the nucleus (Color figure online)

inhibitor of STAT signaling, as an interaction partner. The authors could confirm the physical interaction between Gfi-1 and PIAS3 through various independent experimental methods and show that Gfi-1 is a novel component of STAT signaling pathway with a function in relieving PIAS3 block and activation of STAT signaling [28]. Moreover, two other zinc finger proteins, INSM1 and A20 were successfully used in Y2H screens to identify physiologically relevant interaction proteins such as CAP and TXBP151, respectively [26, 29]. Collectively, all of these findings together with results from our current Y2H screen point to the identification of biologically significant interaction proteins in Y2H screening approach.

Conclusions

In conclusion, we have generated high quality Y2H cDNA libraries from two distinct mouse pluripotent cell types, ESCs and maGSCs. Further, as a proof of concept, we have performed Y2H screen on these two libraries with two different bait proteins and could identify several putative

Table 4 List of Stra8 putative interaction proteins identified in a Y2H screen

Clone no.	NCBI accession no.	Gene/protein description	Gene ontology
62	NM_194262	AT-rich interactive domain 4B (Arid4B)	Chromatin assembly, regulation of transcription
50	NM_030241	SET domain containing (lysine methytransferase) 8 (Setd8)	Chromatin modification, negative regulation of transcription
81, 169	NM_001167922	General transcription factor II E, polypeptide 2 (Gtf2e2)	Regulation of transcription, transcription elongation
52, 166	NM_010480	Heat shock protein 90, alpha (cytosolic), class A member1 (Hsp90AA1)	Mitotic cell cycle, protein folding, ATP catabolic process
108	NM_009288	Serine/threonine kinase 10 (STK10)	Protein phosphorylation
231	NM_026396	Biogenesis of ribosome, homolog (S. cerevisiae) (Brix1)	Translation, ribosome biogenesis
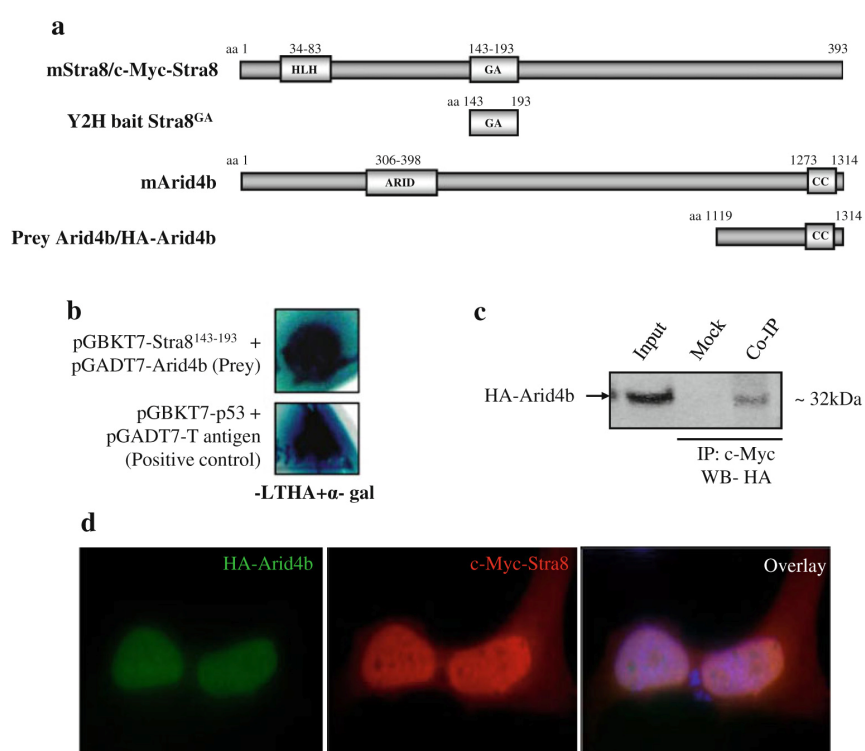
Fig. 3 Stra8 interacts with Arid4b in both ex vivo and in vivo. **a** Illustration of mouseStra8 and Arid4b protein domain organization together with Y2H/mammalian expression constructs used for the analysis. Stra8 contains an N-terminal helix–loop–helix (HLH) domain and a centrally located GA-rich region. Arid4b bears an AT-rich interaction domain (ARID) and C-terminally located coiled-coil domain. **b** Confirmation of Arid4b interaction with Stra8 using direct-Y2H. The yeast cells transformed with bait and prey plasmids, as indicated, were selected on SD–LTHA and for α-galactosidase (α-gal) activity. The interaction between p53 and T-antigen was assayed as a positive control in direct-Y2H experiments. **c** Co-IP analysis using c-Myc antibodies on cell lysates prepared form NIH-3T3 cells transiently transfected with c-Myc-Stra8 and HA-Arid4b. Western blot analysis with HA-tag antibody confirmed the interaction between Stra8 and Arid4b. **d** IFC analysis of NIH-3T3 cells transiently transfected with c-Myc-Stra8 (green) and HA-Arid4b (red) showed co-localization of both proteins in the nucleus. The cells were counterstained with DAPI to visualize the nucleus (Color figure online)

interaction partners, thus validating the quality of these libraries. We also validated the interaction between bait proteins and their respective putative interaction protein using in vivo, ex vivo, and in vitro protein–protein interaction methods, thus highlighting the biological significance of the identified interactor. Finally, our Y2H cDNA

libraries are useful tools to analyze the protein–protein interactions in pluripotent stem cells and will be distributed freely for the research community.

Methods

Cell Culture

The derivation and maintenance of mouse ESCs and maGSCs from 129/Sv genetic background was described previously [7]. Briefly, the undifferentiated ESCs and maGSCs were maintained on mitomycin C-inactivated mouse embryonic fibroblasts (MEFs) and cultured in DMEM (PAN, Germany) supplemented with 20 % defined FBS (fetal bovine serum) (PAN, Germany), 1 % penicillin/streptomycin, 0.1 mM non-essential amino acids, 2 mM L-glutamine, 1 mM sodium pyruvate, 0.1 mM β-mercaptoethanol (all the above ingredients are from Life Technologies, Germany) and 1,000 U/ml LIF (Chemicon, USA). For feeder depletion, the ESC and maGSC cultures were trypsinized and replated on gelatine coated culture dishes for 20 min. The resulting non-adherent ESCs and maGSCs were collected and used for further analysis. NIH-3T3 cells were maintained as previously described [30].

Construction of GAL4-AD Fusion cDNA Libraries of ESCs and maGSCs

The GAL4-AD fusion cDNA libraries of ESCs and maGSCs were generated using Matchmaker Library Construction and Screening kits (Clontech, Germany) and the protocols provided therein. Briefly, total RNA was isolated from feeder depleted ESCs and maGSCs using peqGOLD TriFast-Reagent (PeqLab, Germany). Then, 1 mg of total RNA from each cell type was used to isolate polyA+ mRNA using Oligotex Direct mRNA kit (Clontech, Germany). Further, the polyA+ mRNA was used to synthesize first-strand cDNA using SMART cDNA synthesis technology (Clontech, Germany). To prepare sufficient double strand-cDNA (ds-cDNA) for transformation into yeast, first-strand cDNA was PCR amplified using Advantage 2 PCR kit (Clontech, Germany). The purified ds-cDNA was co-transformed with linear pGADT7-Rec vector into Yeast AH109 using lithium acetate transformation method [31] and selected on SD/-Leu agar plates. The resultant transformants were pooled and stored at −80°C after estimating the transformation efficiency. Also, cDNA inserts were PCR amplified from randomly picked colonies using Advantage 2 PCR kit (Clontech, Germany) to analyze the average length of cDNA inserts and the recombination efficiency.

Y2H Screening using cDNA Libraries of ESCs and maGSCs, Respectively

To test the GAL4-AD fusion cDNA libraries of ESCs and maGSCs in Y2H screen, two bait proteins, Zfp819 and Stra8 were used, respectively. The GAL4-BD fusion bait construct of *Zfp819* was prepared by cloning PCR fragments of N-terminus Zfp819 (aa 1–343 corresponding to nucleotides (nt) 228–1,256 of NM_028913.3) into pGBKT7 vector resulting in pGBKT7-Zfp819_N. To prepare Stra8 bait construct (pGBKT7-Stra8^{GA}), the cDNA fragment spanning the GA-rich region of the mouse *Stra8* gene (aa 143–193 corresponding to nucleotides (nt) 528–680 of NM_009292.1) was PCR amplified and cloned into the pGBKT7 vector containing the GAL4 DNA-binding domain. The bait constructs were co-transformed with empty pGADT7 vector into AH109 yeast strain to test and exclude auto-activation of GAL4 activated reporter genes, HIS3, ADE, and LacZ. After verification, the bait constructs were transformed separately into Y187 yeast strain. The Y2H screen was performed on ESCs and maGSCs cDNA libraries using pGBKT7-Zfp819_N and pGBKT7-Stra8^{GA}, respectively, using Matchmaker pre-transformed library protocol (Clontech). Briefly, the pre-transformed ESCs and maGSCs libraries in yeast strain AH109 were mixed and mated together with strain Y187 containing the pGBKT7-Zfp819_N and pGBKT7-Stra8^{GA}, respectively. After 24 h of mating, the culture was spread on SD/-Leu/-Trp/-His/-Ade plates and the surviving colonies were further verified on SD/-Leu/-Trp/-His/-Ade/+ X-α-Gal. The positive clones that were blue on X-α-Gal were cultured and the plasmid DNA was isolated using QIAprep Spin Miniprep Kit (Qiagen, Germany). The cDNA inserts of the isolated prey clones were PCR amplified and sequenced using vector specific primers. Identities of prey cDNA clones were determined by BLAST analysis (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Direct-Y2H Assay

The prey cDNA clones were co-transformed with either empty pGBKT7 vector or bait construct into AH109 strain by lithium acetate method [31], to test the auto-activation or interaction, respectively. The co-transformants were first selected on SD/-Leu/-Trp plates and later tested for the reporter gene expression on SD/-Leu/-Trp/-His/-Ade and X-α-Gal plates.

Construction of Mammalian Expression Vectors

To generate Zfp819-E2 expression construct, firstly, the CMV promoter of pEGFP-N1 (Clontech) was replaced with the human EF1α promoter to obtain phEF1α-EGFP-N1.

Next, the ORF of mouse *Zfp819* was PCR amplified using primers containing the E2-tag at the C-terminus and cloned downstream of the human EF1α promoter in phEF1α-EGFP-N1 by replacing EGFP cassette to generate phEF1α-Zfp819-E2 construct. To generate c-Myc-tagged Stra8 expression vector, the ORF of the full-length mouse *Stra8* was PCR amplified and cloned into pCMV-Myc vector (Clontech). The cDNA insert of prey clone, *Arid4b* (aa 81–286 corresponding to nt 244–861 of NM_030241) was cloned into the pCMV-HA expression vector (Clontech).

Purification of GST-Fusion Protein and GST Pull-Down Assay

For generation of GST-Zfp819_N fusion protein, the N-terminal region of Zfp819 (aa 1–343 corresponding to nucleotides (nt) 228–1,256 of NM_028913.3) was PCR amplified using pGBKT7-Zpf819_N construct as a template, and cloned into *Nco* I and *Not* I restriction sites of pET-41a vector (Novagen). The primers used for the PCR amplification is as follows: forward primer: 5′ CCATG-GAGATGGCTGCTGACATGAATTTC and reverse primer: 5′ GCGGCCGCCAGGCTGGATGTACTGGGAAG. The expression and purification of either GST-only or GST-Zfp819_N fusion protein was performed as previously described [32]. For GST pull-down experiments, the ESCs from a confluent 10-cm cell culture plate were lysed with 1 ml of RIPA buffer and are processed as described [32].

Co-IP, Western Blot, and Co-localization Studies

For Co-IP experiments, NIH-3T3 cells were transiently transfected with indicated constructs, using Lipofectamine 2000 (Life technologies, Germany) and processed using the Immunoprecipitation kit (Protein G) (Roche, Germany). The protein complexes were immunoprecipitated using c-Myc tag (05-724, Millipore) antibodies and western blotting was performed with HA-tag (ab9110, Abcam) antibodies. The eluted protein complexes from GST pull-down experiments were subjected to Western blotting using Chd4 antibodies (ab70469). For IFC experiments, NIH-3T3 cells grown on round coverslips were transiently transfected with indicated constructs. After 24 h, the cells were washed with PBS and fixed in 4 % paraformaldehyde (PFA) before processing for IFC using a standard protocol using the antibodies mentioned above. Finally, the coverslips were mounted with DAPI mounting medium (Vector Laboratories) and visualized by Olympus BX60 fluorescence microscope. Images were acquired and processed using Cell^F software.

References

1. Evans, M. J., & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature, 292*(5819), 154–156.
2. Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America, 78*(12), 7634–7638.
3. Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., et al. (1998). Embryonic stem cell lines derived from human blastocysts. *Science, 282*(5391), 1145–1147.
4. Guan, K., Nayernia, K., Maier, L. S., Wagner, S., Dressel, R., Lee, J. H., et al. (2006). Pluripotency of spermatogonial stem cells from adult mouse testis. *Nature, 440*(7088), 1199–1203.
5. Khromov, T., Pantakani, D. V., Nolte, J., Wolf, M., Dressel, R., Engel, W., et al. (2011). Global and gene-specific histone modification profiles of mouse multipotent adult germline stem cells. *Molecular Human Reproduction, 17*(3), 166–174.
6. Meyer, S., Nolte, J., Opitz, L., Salinas-Riester, G., & Engel, W. (2010). Pluripotent embryonic stem cells and multipotent adult germline stem cells reveal similar transcriptomes including pluripotency-related genes. *Molecular Human Reproduction, 16*(11), 846–855.
7. Zechner, U., Nolte, J., Wolf, M., Shirneshan, K., Hajj, N. E., Weise, D., et al. (2009). Comparative methylation profiles and telomerase biology of mouse multipotent adult germline stem cells and embryonic stem cells. *Molecular Human Reproduction, 15*(6), 345–353.
8. Zovoilis, A., Nolte, J., Drusenheimer, N., Zechner, U., Hada, H., Guan, K., et al. (2008). Multipotent adult germline stem cells and embryonic stem cells have similar microRNA profiles. *Molecular Human Reproduction, 14*(9), 521–529.
9. Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., et al. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell, 113*(5), 643–655.
10. Ng, H. H., & Surani, M. A. (2011). The transcriptional and signalling networks of pluripotency. *Nature Cell Biology, 13*(5), 490–496.
11. Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., et al. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell, 95*(3), 379–391.
12. Rogers, M. B., Hosler, B. A., & Gudas, L. J. (1991). Specific expression of a retinoic acid-regulated, zinc-finger gene, Rex-1, in preimplantation embryos, trophoblast and spermatocytes. *Development, 113*(3), 815–824.
13. Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D. N., Theunissen, T. W., et al. (2006). A protein interaction network for

pluripotency of embryonic stem cells. *Nature, 444*(7117), 364–368.

14. Zuo, C., Liang, S., Wang, Z., Li, H., Zheng, W., & Ma, W. (2009). Enriching protein–protein and functional interaction networks in human embryonic stem cells. *International Journal of Molecular Medicine, 23*(6), 811–819.

15. Fields, S., & Song, O. (1989). A novel genetic system to detect protein–protein interactions. *Nature, 340*(6230), 245–246.

16. Young, K. H. (1998). Yeast two-hybrid: So many interactions, (in) so little time. *Biology of Reproduction, 58*(2), 302–311.

17. Yang, M., Wu, Z., & Fields, S. (1995). Protein-peptide interactions analyzed with the yeast two-hybrid system. *Nucleic Acids Research, 23*(7), 1152–1156.

18. Sridharan, R., Tchieu, J., Mason, M. J., Yachechko, R., Kuoy, E., Horvath, S., et al. (2009). Role of the murine reprogramming factors in the induction of pluripotency. *Cell, 136*(2), 364–377.

19. Tong, J. K., Hassig, C. A., Schnitzler, G. R., Kingston, R. E., & Schreiber, S. L. (1998). Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. *Nature, 395*(6705), 917–921.

20. Anderson, E. L., Baltus, A. E., Roepers-Gajadien, H. L., Hassold, T. J., de Rooij, D. G., van Pelt, A. M., et al. (2008). Stra8 and its inducer, retinoic acid, regulate meiotic initiation in both spermatogenesis and oogenesis in mice. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 14976–14980.

21. Baltus, A. E., Menke, D. B., Hu, Y. C., Goodheart, M. L., Carpenter, A. E., de Rooij, D. G., et al. (2006). In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication. *Nature Genetics, 38*, 1430–1434.

22. Bowles, J., & Koopman, P. (2007). Retinoic acid, meiosis and germ cell fate in mammals. *Development, 134*, 3401–3411.

23. Guan, K., Nayernia, K., Maier, L. S., Wagner, S., Dressel, R., Lee, J. H., et al. (2006). Pluripotency of spermatogonial stem cells from adult mouse testis. *Nature, 440*, 1199–1203.

24. Xu, X., Pantakani, D. V., Lührig, S., Tan, X., Khromov, T., Nolte, J., et al. (2011). Stage-specific germ-cell marker genes are expressed in all mouse pluripotent cell types and emerge early during induced pluripotency. *PLoS ONE, 6*(7), e22413.

25. Bruckner, A., Polge, C., Lentze, N., Auerbach, D., & Schlattner, U. (2009). Yeast two-hybrid, a powerful tool for systems biology. *International Journal of Molecular Sciences, 10*(6), 2763–2788.

26. De Valck, D., Jin, D. Y., Heyninck, K., Van de Craen, M., Contreras, R., Fiers, W., et al. (1999). The zinc finger protein A20 interacts with a novel anti-apoptotic protein which is cleaved by specific caspases. *Oncogene, 18*(29), 4182–4190.

27. Kalenik, J. L., Chen, D., Bradley, M. E., Chen, S. J., & Lee, T. C. (1997). Yeast two-hybrid cloning of a novel zinc finger protein that interacts with the multifunctional transcription factor YY1. *Nucleic Acids Research, 25*(4), 843–849.

28. Rodel, B., Tavassoli, K., Karsunky, H., Schmidt, T., Bachmann, M., Schaper, F., et al. (2000). The zinc finger protein Gfi-1 can enhance STAT3 signaling by interacting with the STAT3 inhibitor PIAS3. *EMBO Journal, 19*(21), 5845–5855.

29. Xie, J., Cai, T., Zhang, H., Lan, M. S., & Notkins, A. L. (2002). The zinc-finger transcription factor INSM1 is expressed during embryo development and interacts with the Cbl-associated protein. *Genomics, 80*(1), 54–61.

30. Pantakani, D. V., Czyzewska, M. M., Sikorska, A., Bodda, C., & Mannan, A. U. (2011). Oligomerization of ZFYVE27 (Protrudin) is necessary to promote neurite extension. *PLoS ONE, 6*(12), e29584.

31. Gietz, R. D., Schiestl, R. H., Willems, A. R., & Woods, R. A. (1995). Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast, 11*(4), 355–360.

32. Burnicka-Turek, O., Kata, A., Buyandelger, B., Ebermann, L., Kramann, N., Burfeind, P., et al. (2010). Pelota interacts with HAX1, EIF3G and SRPX and the resulting protein complexes are associated with the actin cytoskeleton. *BMC Cell Biology, 11*, 28.

3.2 Zfp819, a novel KRAB-zinc finger protein, interacts with KAP1 and functions in genomic integrity maintenance of mouse embryonic stem cells

In the second part of this thesis, we set out to reveal the expression pattern and function of *Zinc finger protein 819* (*Zfp819*), a novel pluripotency-related gene which we identified through the transcriptome comparison of undifferentiated and differentiated pluripotent stem cells. Firstly, *Zfp819* was found to be expressed preferentially in undifferentiated pluripotent stem cells and in testis compared to differentiated cell types and adult mouse tissues. In order to reveal the function of *Zfp819* in embryonic stem cells (ESCs), we performed both overexpression and downregulation studies in ESCs. However, we failed to detect any overt effect of *Zfp819* downregulation and upregulation on pluripotency and differentiation networks of ESCs. Interestingly, we observed a significant overexpression of endogenous retroviral elements (*ERV*s) such as *IAP* and *LINE 1* in *Zfp819* downregulated cells. In support of these results, we identified the physical association of *Zfp819* with the genomic regions of *IAP* elements as well as with *KAP1*, a universal co-repressor for KRAB-zinc finger proteins and known to repress *ERV*s. Moreover, we observed a spontaneous DNA damage and inefficiency of *Zfp819* downregulated cells to repair the induced DNA damage. Collectively, our findings reveal that *Zfp819* functions in repression of *ERV*s thereby maintaining the genome stability of ESCs.

Authors: Xiaoying Tan, Xingbo Xu, Ulrich Zechner, Jessica Nolte, Wolfgang Engel, D.V. Krishna Pantakani

Authors contributions:
1. Xiaoying Tan: She has done most of the experiments, conception and design, data collection, data analysis, data interpretation, helped in drafting the manuscript

2. Xingbo Xu: Conception and design, data collection, data analysis, helped in drafting the manuscript.

3. Ulrich Zechner: DNA methylation analysis and interpretation.

4. Jessica Nolte: Conception and design.

5. Wolfgang Engel: Conception and design, financial support, helped in drafting the manuscript.

6. D.V. Krishna Pantakani: Conception and design, data interpretation, drafted the manuscript.

Zfp819, a novel KRAB-zinc finger protein, interacts with KAP1 and functions in genomic integrity maintenance of mouse embryonic stem cells

Xiaoying Tan[1], Xingbo Xu[1], Ulrich Zechner[2], Jessica Nolte[1], Wolfgang Engel[1], D. V. Krishna Pantakani[1‡]

[1]Institute of Human Genetics, University of Goettingen, Goettingen, Germany. [2]Institute of Human Genetics, Johannes Gutenberg-University Mainz, Mainz, Germany

‡Correspondence: D. V. Krishna Pantakani, Tel: +49-551-399302; Fax: +49-551-399303; E-mail: krishna.if1@gmail.com

Abstract

Pluripotency is maintained by both known and unknown transcriptional regulatory networks. In the present study, we have identified Zfp819, a KRAB-zinc finger protein, as a novel pluripotency-related factor and characterized its role in pluripotent stem cells. We show that *Zfp819* is expressed highly in various types of pluripotent stem cells but not in their differentiated counterparts. We identified the presence of non-canonical nuclear localization signals in particular zinc finger motifs and identified them as responsible for the nuclear localization of Zfp819. Analysis of the *Zfp819* promoter region revealed the presence of a transcriptionally active chromatin signature. Moreover, we confirmed the binding of pluripotency-related factors, Oct4, Sox2, and Nanog to the distal promoter region of *Zfp819*, indicating that the expression of this gene is regulated by a pluripotency transcription factor network. We found that the expression of endogenous retroviral elements (ERVs) such as Intracisternal A Particle retrotransposons (IAPs) and Long Interspersed Nuclear Elements (LINE1s), is significantly upregulated in *Zfp819*-downregulated (Zfp819_DN) cells. In line with the activation of ERVs, we observed the occurrence of spontaneous DNA damage in Zfp819_DN cells. Furthermore, we tested whether Zfp819 can interact with KAP1, a KRAB-associated protein with a transcriptional repression function, and found the interaction between these two proteins in both *in vitro* and *in vivo* experiments. The challenging of *Zfp819*_DN cells with DNA damaging agent revealed that these cells are inefficient in repairing the damaged DNA, as cells showed presence of γH2A.X foci for a prolonged time. Collectively, our study identified Zfp819 as a novel pluripotency-related factor and unveiled its function in genomic integrity maintenance mechanisms of mouse embryonic stem cells.

Keywords: Zfp819, KAP1, Nuclear localization, Retroviral elements, DNA damage, ESCs

Introduction

Embryonic stem cells (ESCs) display a unique characteristic of pluripotency, namely, the maintenance of self-renewal as well as the potential to differentiate into diverse cell types of all three germ layers (Dejosez and Zwaka 2012). It is well known that ESCs employ a complex transcriptional regulatory network comprising of Oct3/4, Sox2 and Nanog, among several other factors, to regulate the pluripotency (Boyer et al. 2005; Ng and Surani 2011). Perturbations of expression of these core pluripotency factors lead to loss of pluripotency and subsequent nonspecific differentiation (Ivanova et al. 2006). Nanog may sustain the self renewal and undifferentiation state through the regulation of Oct3/4 and Sox2 expression, both of which in turn

control the expression of downstream genes important for maintenance of pluripotency or inhibition of differentiation (Loh et al. 2006). By employing genome-wide chromatin immunoprecipitation studies, thousands of direct target genes regulated by Oct3/4, Sox2, Klf4, and Nanog were identified (Loh et al. 2006; Chen et al. 2008; Kim et al. 2008; Marson et al. 2008). Interestingly, many of these targets encode for transcription factors such as Esrrb, Zfp281, and Sall4 which are also crucial for the maintenance of pluripotency (Mitsui et al. 2003; Zhang et al. 2006a; Wang et al. 2008; Zhang et al. 2008).

KRAB-associated protein 1 (KAP1, also known as TRIM28 and TIF1β) is a novel Oct3/4-dependent transcription factor that regulates pluripotency (Seki et al. 2010). KAP1 was initially identified as an interaction partner of KRAB-domain containing zinc finger transcription factors (Friedman et al. 1996). It is critical for the regulation of normal development as KAP1-deficient mice die before gastrulation (Cammas et al. 2000). Recently, KAP1 was shown to be implicated in repression of endogenous retroviruses (ERVs) in mouse ESCs, where the deletion of KAP1 resulted in a significant upregulation of a range of ERVs (Rowe et al. 2010). Additionally, ATM-mediated phosphorylation of KAP1 leads to its co-localization with γ-H2A.X, a DNA damage repair foci marker, in response to genotoxic stress. The phosphorylated KAP1 functions in derepression of its target genes to promote cell cycle arrest and apoptosis (White et al. 2006; Li et al. 2007; Li et al. 2010). In line with these observations, cells with loss of phosphorylated KAP1 were shown to be hypersensitive to induced DNA damage and show defects in chromatin decondensation (Ziv et al. 2006). Collectively, these studies highlight the important role of KAP1 in normal development and regulation of ESCs pluripotency and differentiation. Although, the precise molecular mechanism(s) by which KAP1 executes its diverse functions is unclear, it can be

hypothesized that the interaction between KAP1 and a wide variety of KRAB-zinc finger proteins may drive KAP1 divergent functions.

Zinc finger proteins belong to one of the most abundant protein families in eukaryotic genomes (Urrutia 2003). About 30% of these zinc finger proteins harbor a KRAB domain which is responsible for protein-protein interactions whereas several classical Cys_2His_2-zinc finger motifs present at the C-terminus of the protein enable the recognition of a specific DNA sequence thereby leading to either gene activation or repression (Witzgall et al. 1994; Looman et al. 2002). Several lines of evidence suggests that zinc finger proteins play an important role in maintenance of ESCs pluripotency and differentiation potential as well as proliferation and cell cycle control. Zinc finger protein 206 (Zfp206) is known to regulate ESCs gene expression and differentiation and its expression is also considered as a hall mark of pluripotent cells (Zhang et al. 2006c; Wang et al. 2007). Likewise, Zfp42 (also known as Rex1) expression is restricted to undifferentiated ESCs and is known to maintain their undifferentiation state (Eiges et al. 2001; Zhang et al. 2006b). Recently, CtBP-interacting BTB Zinc Finger Protein (CIBZ) was identified as a key transcriptional regulator of ESCs and was shown to play a role in G1/S transition partly depending on Nanog expression (Nishii et al. 2012). Moreover, ZFP57 was shown to maintain the genomic imprints in ESCs in a KAP1-dependent manner (Zuo et al. 2011). In ESCs, Zfp809 was found as a retrovirus restriction factor, which recruits KAP1 to regulate the silencing of ERVs (Wolf and Goff 2009). Although numerous studies have revealed the function of several KRAB-zinc finger proteins, the function of a large number of KRAB-zinc finger proteins is yet to be determined.

In the present study, we identified Zfp819, a KRAB-zinc finger protein with an unknown function, as a pluripotency-related gene and show that it is expressed in various pluripotent cell types. We

demonstrate that the *Zfp819* promoter region is highly enriched for transcriptionally active chromatin marks and is bound by pluripotency-related transcription factors, suggesting a possible crucial role for Zfp819 in pluripotent cells. Interestingly, the expression of selected retroviral elements was highly upregulated in Zfp819-downregulated cells, while the expression of pluripotency and differentiation marker genes was not significantly altered. In line with these observations, we show that Zfp819 interacts with KAP1, a known transcriptional repressor of retroviral elements, further strengthening the role of Zfp819 in regulation of these elements. Moreover, we observed the activation of DNA damage response markers such as γH2A.X in Zfp819-downregulated ESCs indicating its role in maintenance of genomic integrity.

Materials and methods

Cell Culture

The mouse ESC (ES R1) cell line was maintained as previously described (Zechner et al. 2009). Briefly, the undifferentiated ESCs were maintained on Mitomycin C-inactivated mouse embryonic fibroblasts (MEFs) and cultured using standard ESC culture medium (Zechner et al. 2009). For feeder depletion, the ESC cultures were trypsinized and replated on gelatin-coated culture dishes for 20 min and the non-attached ESCs were used for further analysis. NIH-3T3 cells were cultured in DMEM (PAN, Germany) supplemented with 10% fetal bovine serum (FBS) (PAN, Germany), 1% L-glutamine (Life Technologies, Germany) and 1% penicillin/streptomycin (Life Technologies, Germany).

For ESC differentiation experiments, feeder-depleted ESCs (1×10^6) were seeded onto 100 mm low attachment sterile cell plates (Greiner, Germany) in 10 ml ESC culture medium lacking leukemia inhibitory factor (LIF). The culture medium was exchanged every day until cells were harvested

at indicated time points. Alternatively, ESCs were seeded onto 0.1% gelatin-coated dishes, cultured without LIF and in the presence of 10^{-6}M retinoic acid (RA) for 20 days.

Transfection and Immunofluorescence staining

For transfection experiments, cells were plated onto 24-well plates containing glass cover slips, cultured overnight and transfected with Lipofectamine2000 (Life Technologies, Germany) according to the manufacturer's instructions. Briefly, the construct DNA and Lipofectamine 2000 were mixed in a ratio of 1:2 in a total volume of 100µl of Opti-MEM (Life Technologies, Germany) and allowed to form the complex by incubating 20 min at room temperature (RT). Then, the DNA-lipofectamine complex was added to the cells. After 3 h of incubation, the DMEM medium containing the complex was replaced with standard culture medium and the cells were allowed to grow for additional 24 h. The next day, cells were processed for immunofluorescence staining as described (Khromov et al., 2011). Microscopic images were acquired using an Olympus BX60 fluorescent microscope equipped with UplanFI 20X/0.50 objective lens and cellSens program. The acquired images were processed using the same cellSens software.

DNA methylation analysis and chromatin immunoprecipitation assays

Genomic DNA was isolated from undifferentiated and differentiated ESCs and maGSCs Bisulfite pyrosequencing of the *Zfp819* promoter region in the genomic DNA samples was performed as previously described (Zechner et al., 2009). Primer sequences are available on request.

Chromatin immunoprecipitation (ChIP) assays were performed on ESC chromatin using the Diagenode OneDay ChIP kit and protocols provided therein. Briefly, for each ChIP assay, 3×10^6 cells were cross-linked using 4% formaldehyde in PBS

and lysed with shearing kit (Diagenode) followed by bioruptor sonication (Diagenode, UCD-200 TM) to obtain a chromatin size of 200-500 bp. Then, the sheared chromatin was immunoprecipitated with 5μg of antibodies as indicated and processed with the Diagenode OneDay ChIP protocol to extract the chromatin bound to each specific antibody. Mouse/Rabbit IgG was used as a control in mock ChIP experiments. Quantitative analysis of the immunoprecipitated DNA was performed by real-time qPCR using the primers listed in supplementary table S1. The ChIP-qPCR data were analyzed using $\Delta\Delta Ct$ method in which the immunoprecipitated sample Ct value was normalized with the input DNA Ct value and the percentage of precipitation was calculated using the following formula (% Input= 2(Ct Input-Ct Iped) × DF × 100%, DF indicates the input dilution factor, i.e., 10).

Direct-Y2H Assay

For direct-Yeast Two-Hybrid (Y2H) experiments, the pGADT7-KAP1_RBCC construct and pGBKT7-Zfp819_N-ter were transformed together with empty pGBKT7 and pGADT7, respectively, into AH109 strain by the lithium acetate method (Gietz et al. 1995) to test for auto-activation. For testing interaction between KAP1 and Zfp819, the pGBKT7-Zfp819_N-ter and pGADT7-KAP1_RBCC constructs were transformed and first selected on SD/-Leu/-Trp plates and later tested for the reporter genes expression on SD/-Leu/-Trp/-His/-Ade and X-α-gal plates. The primers used for Y2H constructs generation are listed in supplementary table S3

GST-pull down assay and co-immunoprecipitation experiments

The generation of recombinant GST-Zfp819_N and subsequent GST-pull down assay were described previously (Burnicka-Turek et al. 2010; Zheng et al. 2012). Briefly, the protein extract of ESCs was incubated overnight on a roller with

Glutathione-agarose beads containing either GST-only or GST-Zfp819_N fusion proteins. Following incubation, the beads were washed to remove non-specifically bound proteins and resuspended in elution buffer. The eluted protein complex was collected and subjected to Western blot analysis with KAP1 antibodies. The source of KAP1 antibodies is listed in supplementary table S5.

For co-immunoprecipition, ESCs transiently transfected with indicated constructs for 48 h were used to prepare the protein extract using RIPA buffer (Millipore). The protein extracts were pre-cleared with protein A/G beads (Santa Cruz) for 3 h at 4°C. The supernatants were then incubated with the appropriate antibody for 1 h on ice with occasional mixing. After incubation, the beads were washed and the bound immune complexes were eluted using SDS-PAGE loading buffer, separated on 4-12% SDS-PAGE and subjected to Western blot analysis.

Luciferase reporter assay

Luciferase reporter assays were performed as previously described (Smorag et al. 2012). Briefly, one day before transfection, ESCs or NIH-3T3 cells were trypsinized and transferred (1-2 × 10^5 cells/ well) to a fresh well of a 6 well plate. Next day, the cells were transiently transfected in duplicates with indicated luciferase vectors; 2μg of the reporter plasmid was co-transfected with either Zfp819 overexpression construct (hEF1α_Zfp819_E2) or control vector (pEGFP-1, mock construct) using Lipofectamine2000 (Life technologies). At the same time, the cells were transfected with pRL Renilla Luciferase control reporter vector (Promega) for normalization of transfection efficiencies. After 48 h of transfection, the cells were lysed and assayed by using Dual-Glo Luciferase Assay System (Promega).

RNA extraction, RT-PCR and qPCR

Total RNA was extracted from cells using Trizol Reagent (Life Technologies, Germany) following the manufacturer's protocol. 5μg of total

RNA was digested with DNaseI (Sigma) and used for cDNA synthesis using the SuperScriptII system (Life Technologies, Germany). Diluted cDNA (1/10) was used as a template in QuantiFast SYBR Green (Qiagen) and run on an ABI 7900HT Real-Time PCR System (Applied Biosystems). Primers used in RT-PCR and qPCR are listed in supplementary data (Table S1).

Statistical Analysis

All qPCR data for RNA expression analysis (two or more biological replicates) were calculated using the delta $\Delta\Delta Ct$ method. The GraphPad Prism 4.0 software was used to calculate the statistical significance.

Results

Zfp819 is expressed preferentially in pluripotent stem cells

Pluripotency is established and maintained by both known and unknown transcription factors and their coordinated regulatory mechanisms (Boyer et al. 2005; Ng and Surani 2011). In an attempt to identify unknown transcription factors which might play an essential role in maintenance of pluripotency, we made use of our previous transcriptome analysis that compared undifferentiated and differentiated cell types of ESCs and multipotent adult germline stem cells (maGSCs) (Meyer et al. 2010). This comparative analysis led us to identify *Zinc finger protein 819* (*Zfp819*) as a novel pluripotency-related factor (Meyer et al. 2010). Zfp819 belongs to the zinc finger-containing KRAB family of proteins and bears a functional KRAB domain at its N-terminus, while the C-terminus harbors 11 zinc finger (Znf) motifs (Fig. 1A). Further analysis of the protein sequence revealed that Zfp819 contains moderately conserved KRAB-A and –B boxes, which are encoded by two different exons, hence classifying it as a KRAB-A+B type zinc finger protein (Fig. 1B). In order to validate the transcriptome data, we

performed the qPCR analysis on undifferentiated and differentiated ESCs and maGSCs and confirmed the dramatic downregulation of *Zfp819* upon differentiation (Fig. 1C). We performed qPCR analysis on various adult mouse tissues and found a very high expression of *Zfp819* in testis and moderate to weak expression in heart, kidney, liver, and ovary (Fig. 1D). Expression analysis of *Zfp819* in various pluripotent cell types revealed a strong expression in all analyzed pluripotent cell types as well as in blastocysts, but not in differentiated cell types such as MEFs and NIH-3T3 cells (Fig. 1E). Recently, *Zfp819* was shown to be expressed in both partially and fully reprogrammed iPSCs (Sridharan et al. 2009). To systematically analyze the transcriptional activation of *Zfp819* during reprogramming, we performed time course-based expression analysis and identified it to be started expressing from 10-days of reprogramming along with the endogenous *Oct4* activation (Fig. 1F).

Zfp819 is localized to the nucleus through non-canonical nuclear localization signals

Next, we performed sub-cellular localization studies and observed that the full-length Zfp819 protein localized only to the nucleus (Fig. 2B). To delineate the region responsible for this nuclear localization, we made several deletion constructs (Fig. 2A). The N-terminal region (Zfp819_N) containing KRAB domain was localized to the cytoplasm, whereas the C-terminal region (Zfp819_C) harboring all zinc finger motifs was found to be localized in the nucleus (Fig. 2 C). Our search for the presence of any canonical nuclear localization signal (NLS) in Zfp819 failed to identify such elements (data not shown). Recently, three non-canonical NLS consensus sequences named as PY-NLSs $-R/K/HX_{(2-5)}PY$, $\Phi G/A/S\Phi\Phi X_{(11-13)}PY$, and basic enriched$_{(5-8)}X_{(8-10)}PY$ were shown to be responsible for nuclear localization of several proteins in a Kapβ2-dependent manner (Lee et al., 2006). Our closer inspection revealed the presence of an
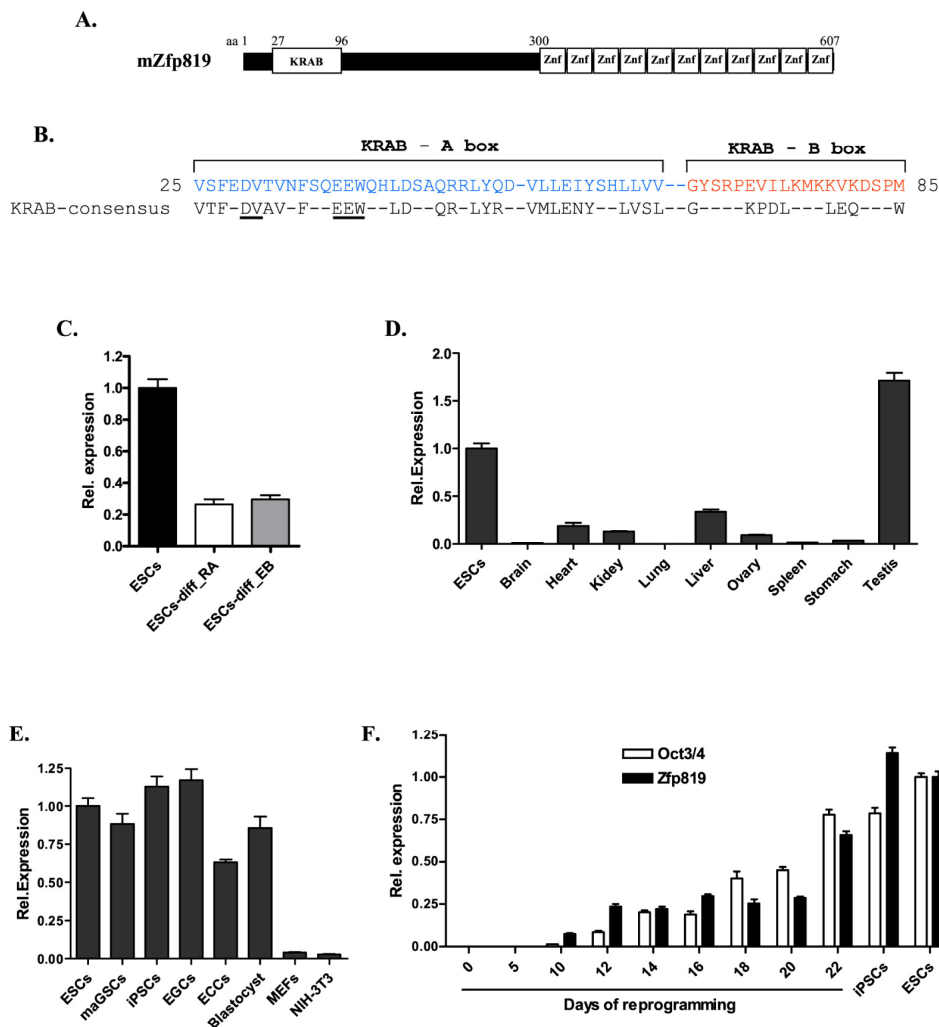
Figure 1. *Zfp819* **is highly expressed in pluripotent stem cells. (A)** Diagram showing the mouse Zfp819 protein architecture. KRAB - Krueppel-associated box, Znf - zinc finger motif. **(B)** The amino acid (aa) number is indicated on the top. The amino acid sequence of Zfp819 (25-96 aa) encoded by two different exons is indicated in blue and red. The consensus sequences of the KRAB-A box and the KRAB-B box are indicated below. The motifs responsible for interaction with KAP1 are underlined. **(C)** qRT-PCR data showing the expression of *Zfp819* in undifferentiated ESCs and either retinoic acid (RA)-differentiated (ESCs-diff_RA) or embryoid body (EB) differentiation (ESCs-diff_EB) counterparts. **(D)** Bar graph showing the qRT-PCR data of *Zfp819* expression in various adult mouse tissues. The expression level of *Zfp819* in ESCs was used to normalize the data. **(E)** qRT-PCR data showing the expression of *Zfp819* in various pluripotent cell types, blastocysts and differentiated cells. ESCs - embryonic stem cells, maGSCs - multipotent adult germline stem cells, iPSCs - induced pluripotent stem cells, EGCs - embryonic germ cells, ECCs - embryonic carcinoma cells, MEFs - mouse embryonic fibroblasts, NIH-3T3 - a widely used fibroblast cell line. **(F)** qRT-PCR data showing the expression of *Oct3/4* (white bars), a pluripotency-associated gene, and *Zfp819* (black bars) during the time course of somatic cell reprogramming using Yamanaka factors (Oct3/4, Sox2, Klf4 and c-Myc).
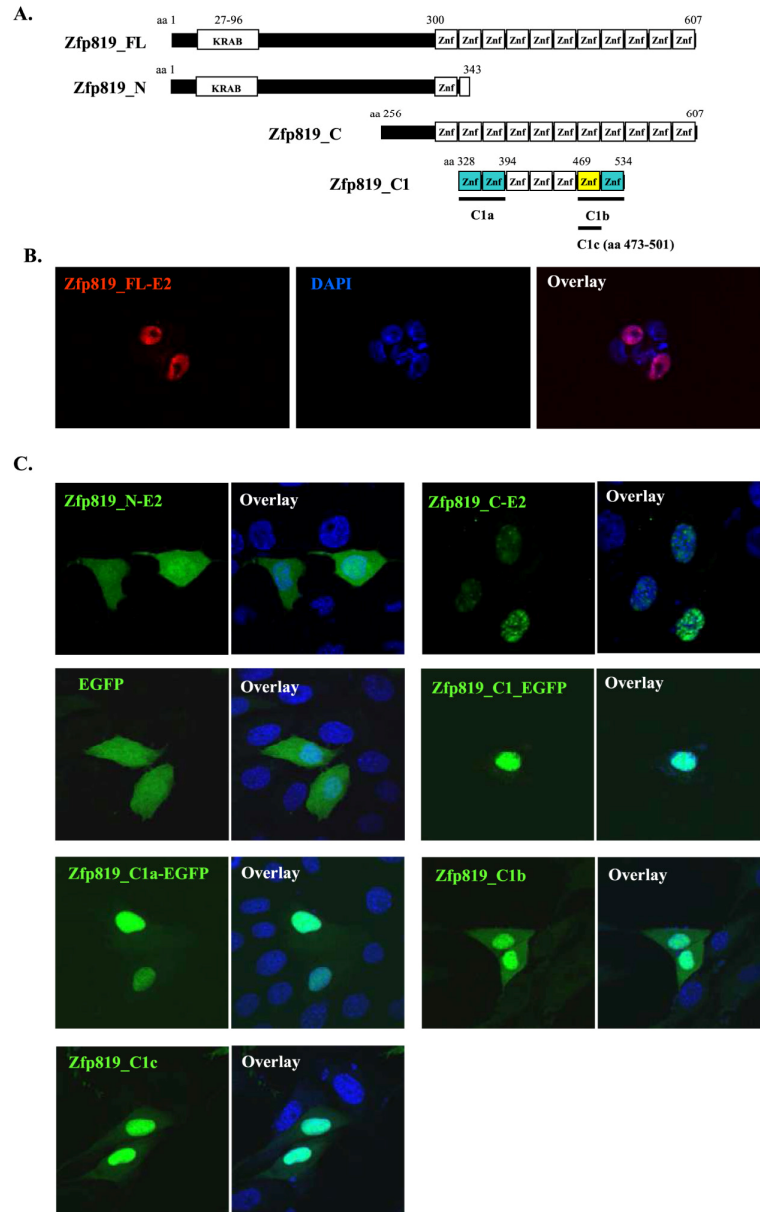
Figure 2. Nuclear localization of Zfp819 is mediated by non-canonical nuclear localization signals. (A) Schematic illustration of constructs used to determine the nuclear localization signal (NLS). Zfp819_FL, Zfp819_N, and Zfp819_C constructs were N-terminally fused to E2-epitope tag. The Zfp819_C1, C1a, C1b, and C1c sequences were N-terminally fused to an EGFP cassette. Aqua color indicates the presence of non-canonical PY-NLS (ΦG/A/S$\Phi\Phi$X$_{(11-13)}$PY), and yellow color indicates the presences of basic enriched$_{(5-8)}$X$_{(8-10)}$PY motifs. **(B)** Immunofluorescence images showing the localization of Zfp189_FL protein in ESCs. **(C)** Images showing the localization of Zfp819_N-E2, Zfp819_C-E2 and either EGFP-only or EGFP-fused Zfp819_C1-EGFP, Zfp819_C1a-EGFP, Zfp819_C1b-EGFP, and Zfp819_C1c-EGFP in NIH-3T3 cells. Nuclei were visualized by staining with DAPI (blue) in **B and C.**
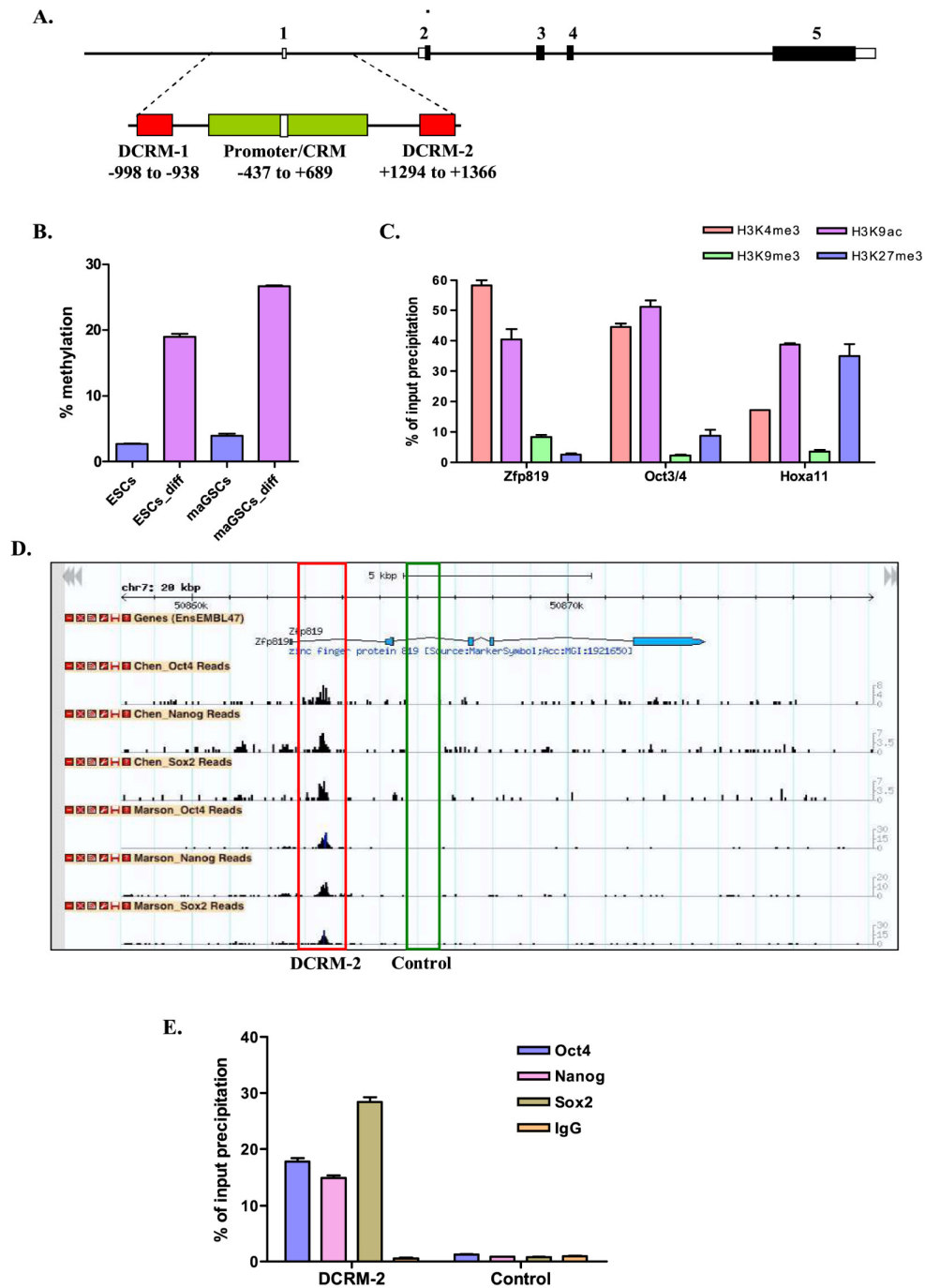
Figure 3. The promoter region of *Zfp819* is enriched for transcriptionally active epigenetic marks in ESCs. **(A)** Schematic diagram showing the genomic organization of *Zfp819* along with its non-coding (white boxes) and coding (black boxes) regions. The location of the putative promoter, a cis-regulatory module (CRM) (green) and two distal CRMs (DCRM-1 and DCRM-2, red) together with their position relative to the transcriptional start site (- or +) are shown in the enlarged diagram. **(B)** Bar graph showing the *Zfp819* promoter methylation levels in undifferentiated and RA-differentiated ESCs and maGSCs. **(C)** qPCR data showing the enrichment of

activating (H3K4me3 and H3K9ac) and depletion of repressive (H3K9me3 and H3K27me3) histone modification marks at the promoter region of *Zfp819*. The promoter regions of the transcriptionally active *Oct3/4* gene and the lineage-specific gene*Hoxa11*, with bivalent chromatin served as controls. **(D)** ChIP-Seq data from a publicly available database showing the enrichment of Oct4, Nanog, and Sox2 binding at DCRM-2 of *Zfp819* (red box). The region displaying no binding of transcription factors (green box) was used as a control region for ChIP and subsequent qPCR analysis. **(E)** ChIP followed by qPCR showing the binding of Oct4, Nanog and Sox2 to the DCRM-2 but not to the control region of *Zfp819*.

R/K/HX$_{(2-5)}$PY motif in zinc fingers 2-9 of Zfp819 (Supplementary Fig. S2). We also observed three fairly conserved ΦG/A/SΦΦX$_{(11-13)}$PY motifs in zinc fingers-2, -3, and -8, and one basic enriched$_{(5-8)}$X$_{(8-10)}$PY motif in zinc finger-7 of Zfp819 (Supplementary Fig. S2). The C-terminal fusion of zinc fingers-2-8 (Zfp819_C1) to EGFP resulted in the efficient localization of EGFP to the nucleus (Fig. 2C). Further, we split the Zfp819_C1 into three fragments, C1a, C1b, and C1c and fused these C-terminally with EGFP. All of them were found to drive the EGFP localization into the nucleus (Fig. 2C). Collectively, these results indicate that the non-canonical NLS present in selective zinc fingers of Zfp819 are responsible for its localization into the nucleus.

Transcript variants of *Zfp819* are expressed specifically in pluripotent stem cells

In the Ensembl database, four different transcripts were reported to be transcribed from the *Zfp819* gene, two out of these four are annotated to code for a functional protein (Supplementary Fig. S1A). To evaluate the expression of each transcript in pluripotent cells in comparison to adult mouse tissues, we designed PCR primers that specifically amplify each different transcript (Supplementary Fig. S1B, and data not shown). The RT-PCR analysis revealed the specific expression of two non-protein-coding transcripts (*Zfp819a* and *Zfp819b*) in pluripotent cells (Supplementary Fig. S1B). The protein-coding *Zfp819c* transcript is expressed in pluripotent cells as well as in testis, but not in any other analyzed tissues (Supplementary Fig. S1B). In

contrast, the other protein-coding *Zfp819d* transcript was found to be expressed in both pluripotent cells and in all analyzed tissues (Supplementary Fig. S1B).

The promoter region of *Zfp819* is epigenetically regulated and a target of core pluripotency transcription factors in pluripotent stem cells

The global chromatin state of pluripotent cells is highly dynamic and transcriptionally permissive. Hence, many genes are known to be expressed through leaky transcription (Meshorer and Misteli 2006). To evaluate whether the expression of *Zfp819* in pluripotent cells is due to inevitability or leaky expression, we studied the epigenetic regulation of the predicted promoter region of *Zfp819* (Fig. 3A). The DNA methylation studies revealed that the promoter region of *Zfp819* is largely hypomethylated (~5%) in undifferentiated ESCs and maGSCs, suggesting that the expression of *Zfp819* is inevitable in pluripotent cells and not due to leaky expression (Fig. 3B). In contrast, the promoter region of *Zfp819* became moderately methylated (~30%) in differentiated counterparts of ESCs and maGSCs (Fig. 3B). The promoter regions of pluripotency-related genes are known to be enriched for transcription-activating histone modification marks such as histone H3 lysine 4 trimethylation (H3K4me3) and histone H3 lysine 9 acetylation (H3K9ac) in pluripotent cells (Meshorer 2007; Mikkelsen et al. 2007; Pan et al. 2007). In contrast, the promoter regions of genes involved in lineage commitment are shown to contain a bivalent chromatin domain characterized by the presence of both transcriptionally active H3K4me3 and
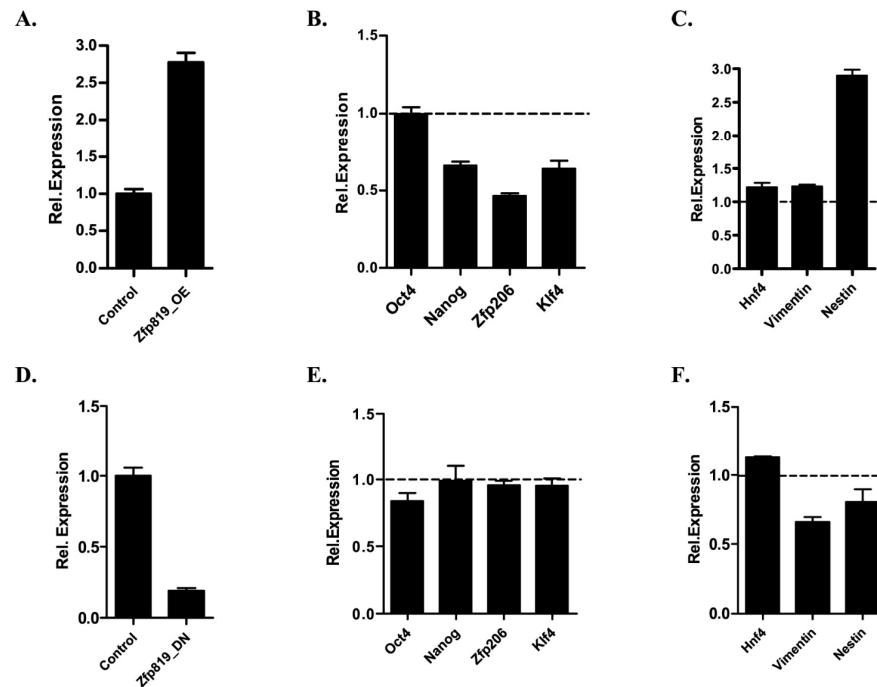
Figure 4. Zfp819 has no significant effect on the pluripotency network. (A) qRT-PCR data showing the expression of *Zfp819* in control and Zfp819-overexpressing (Zfp819_OE) cells. qRT-PCR data showing the expression of pluripotency-related genes **(B)** and differentiation marker genes **(C)** in Zfp819_OE cells. **(D)** qRT-PCR data showing the expression of *Zfp819* in control cells and Zfp819-stable downregulation cell line (Zfp819_DN). Bar graph showing the expression levels of pluripotency-related genes **(E)** and differentiation marker genes **(F)** in Zfp819_DN cells. The expression level of the corresponding genes in control ESCs is indicated with a dashed line in **B**, **C**, **E**, and **F**.

repressive histone H3 lysine 27 trimethylation (H3K27me3) histone modifications (Bernstein et al. 2006). To confirm the transcriptionally active chromatin state of *Zfp819* in pluripotent cells, we analyzed the histone modification patterns at the promoter region of *Zfp819* using chromatin immunoprecipitation (ChIP) followed by qPCR. These results revealed a very high enrichment for activating histone modification marks but not for repressive histone modification marks at the *Zfp819* promoter region, that was also observed for the pluripotency-related gene *Oct4* (Fig. 3C).

The pluripotency transcription factor network comprising Oct4, Sox2, Klf4, and Nanog regulates the expression of several genes which are involved in the establishment and maintenance of pluripotency (Mitsui et al. 2003; Kim et al. 2008; Yu et al. 2009). To address the question, whether *Zfp819* is a transcriptional target of the core pluripotency transcription factors, we searched the publicly available ChIP-sequencing data for these transcription factors and found *Zfp819* to be one of the targets of Oct4, Sox2, Klf4, and Nanog (Chen et al. 2008; Kim et al. 2008; Marson et al. 2008) (Fig. 3D). To further validate these results we performed ChIP on ESCs chromatin using specific antibodies against Oct4, Sox2, and Nanog. The qPCR analysis of immunoprecipitated chromatin revealed that the distal cis-regulatory module-2 (DCRM-2) region of *Zfp819* is bound by all three analyzed pluripotency
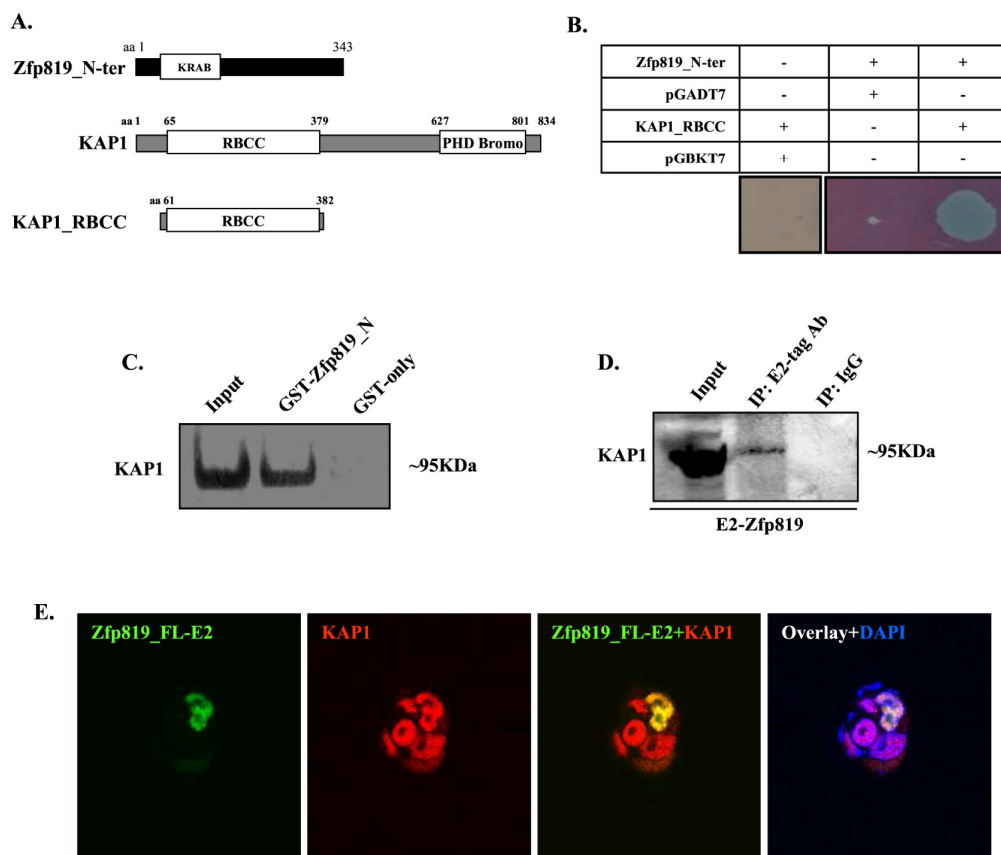
A.

Zfp819_N-ter

aa 1   343

| KRAB |

KAP1

aa 1 65   379   627   801 834

| RBCC | | PHD Bromo |

KAP1_RBCC

aa 61   382

| RBCC |

B.

Zfp819_N-ter	-	+	+
pGADT7	-	+	-
KAP1_RBCC	+	-	+
pGBKT7	+	-	-

C.

Input GST-Zfp819_N GST-only

KAP1   ~95KDa

D.

Input IP: E2-tag Ab IP: IgG

KAP1   ~95KDa

E2-Zfp819

E.

Zfp819_FL-E2   KAP1   Zfp819_FL-E2+KAP1   Overlay+DAPI

Figure 5. Zfp819 interacts with the universal co-repressor KAP1. (A) Schematic illustration of constructs used in direct yeast two-hybrid studies along with the representation of full length KAP1 protein architecture. KRAB – Krueppel-associated box; RBCC – Ring finger, B-box, Coiled-coil; PHD – Plant homeodomain; Bromo - Bromodomain. **(B)** The growth of reporter yeast strain AH109 transformed with Zfp819_KRAB and KAP1_RBCC on selective drop out (SD) medium. AH109 transformed with either KAP1_RBCC and pGBKT7 or Zfp819_KRAB and pGADT7 served as controls for the auto-activation test. **(C)** GST-pull down assay using either recombinant GST-Zfp819_N or GST-only and subsequent Western blot showing the binding of KAP1 to recombinant GST-Zfp819_N protein. The protein extract of ESCs (input) served as an input control. **(D)** Protein extracts from ESCs transiently transfected with E2-Zfp819 were immunoprecipitated with either E2-tag or non-specific IgG antibodies and analyzed by Western blot analysis with KAP1 antibodies. The protein extract from ESCs transfected with E2-Zfp819 served as an input control. **(E)** Immunofluorescence images showing the co-localization of E2-Zfp819 with endogenous KAP1 in ESCs.

transcription factors (Fig. 3E). Collectively, these results show that the expression of *Zfp819* is inevitable and regulated by the core pluripotency transcription factor network.

Zfp819 has no significant effect on the pluripotency and differentiation networks

The Zfp819 nuclear localization and its predicted role in transcriptional regulation prompted us to study whether perturbation of Zfp819 expression has any effect on the regulatory networks

of pluripotent stem cells. Towards this, we performed Zfp819 overexpression (Zfp819_OE) and downregulation (Zfp819_DN) studies in ESCs and analyzed the expression of genes involved in the maintenance of pluripotency as well in early differentiation of pluripotent cells. We transiently overexpressed Zfp819 in ESCs and found that the mRNA expression level of *Zfp819* is ~2.5-fold of that of the control cells (Fig. 4A). The qPCR analysis of pluripotency (*Oct4*, *Nanog*, *Zfp206*, and *Klf4*) and differentiation (*Hnf4*, *Vimentin*, and *Nestin*) marker genes revealed no significant differences in their expression upon Zfp819 overexpression (Fig. 4B, C). Next, we performed downregulation studies using stable overexpression of shRNA constructs against *Zfp819* and found a ~75% reduction in the mRNA expression of *Zfp819* (Fig. 4D). The Zfp819_DN cells showed no defects in morphology and proliferation (data not shown). Further, qPCR analysis for pluripotency and differentiation marker genes revealed no overt changes in their expression in Zfp819_DN cells (Fig. 4E, F). Taken together, these results reveal that Zfp819 has no significant role in the transcriptional regulation of pluripotency and differentiation genes.

Zfp819 interacts with the KRAB effector protein KAP1

We next investigated whether Zfp819 interacts with KAP1, a well known KRAB domain-binding protein with a transcriptional repression function. KAP1 is known to interact with many KRAB domain-containing proteins through its RBCC (Ring finger, B-box, Coiled-coil) domain (Friedman et al. 1996; Peng et al. 2000). To verify whether the KRAB domain of Zfp819 interacts with the RBCC domain of KAP1, we performed a direct-Yeast Two-Hybrid (direct-Y2H) assay using the constructs illustrated in figure 5A. The yeast AH109 strain transformed with empty vectors together with either Zfp819_N-ter or KAP1_RBCC showed no growth on SD/-Leu/-Trp/-His/-Ade/+α-gal, indicating no

autoactivation of Zfp819 as well as KAP1 constructs (Fig. 5B). However, co-transformation of Zfp819_N-ter and KAP1_RBCC showed strong growth with α-gal activity (blue color), indicating the interaction between Zfp819 and KAP1 (Fig. 5B). Next, we performed a Glutathione S-Transferase (GST)-pull down assay using the purified recombinant GST-Zfp819_N protein (Zheng et al. 2012). The GST-pull down assay using total protein extracts of ESCs and recombinant GST-Zfp819_N revealed the interaction between Zfp819 and KAP1, whereas the control assay with protein extracts incubated with GST-only failed to show interaction (Fig. 5C). To confirm these results at the physiological level, we performed co-immunoprecipitation (Co-IP) experiments with protein extracts isolated from cells transfected with E2-tagged Zfp819 (Zfp819_E2). Immunoprecipitation using E2-tag specific antibodies and the subsequent Western blot analysis for endogenous KAP1 revealed a specific interaction, whereas the control Co-IP with non-specific IgG antibodies failed to detect KAP1 (Fig. 5D). Further, we performed co-localization studies in ESCs and found both KAP1 and Zfp819 to co-localize in the nucleus (Fig. 5E).

Zfp819 binds and regulates the expression of endogenous retroviral elements in ESCs

KAP1 is known as a transcriptional repressor of endogenous retroviral elements (ERVs) such as IAPs, LINEs as well as short interspersed elements (SINEs) in ESCs (Rowe et al. 2010). The interaction of Zfp819 with KAP1 prompted us to study whether these ERV elements are transcriptionally activated upon Zfp819 downregulation in ESCs (Fig. 6A). qPCR analysis for the expression of IAPs, LINEs (LINE 1), and SINEs (SINE B1) in Zfp819_DN cells revealed that LINE1 and IAP elements are significantly upregulated, whereas SINE B1 showed no significant changes in expression (Fig. 6B-D). The KRAB domain-containing Znf proteins are
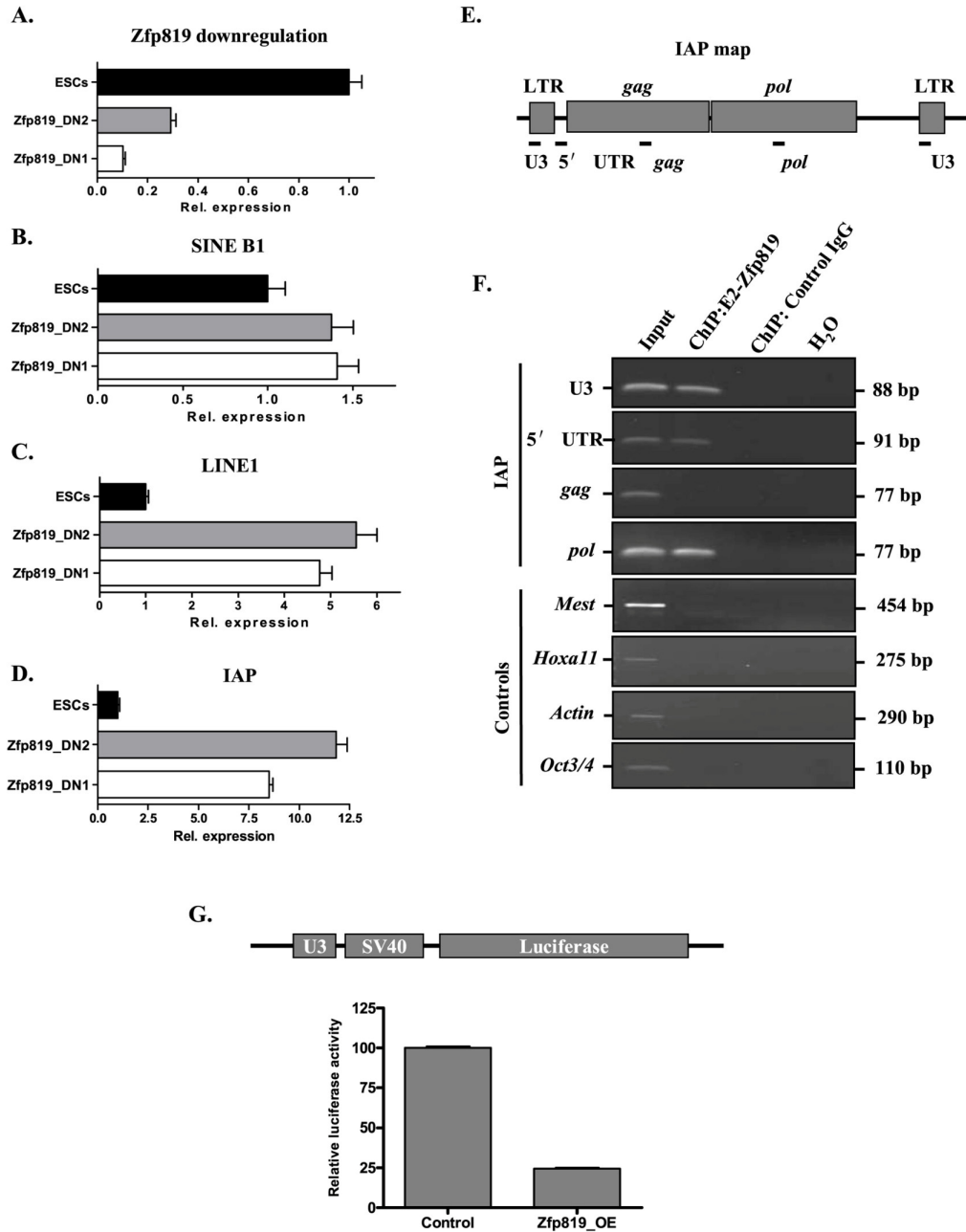
Figure 6. Zfp819 regulates the expression of endogenous retroviruses in ESCs. (A) qRT-PCR data showing the expression of *Zfp819* in control ESCs and two Zfp819 stable downregulation clones (Zfp819_DN1 and Zfp819_DN2). **(B-D)** Bar graph showing the expression of endogenous retroviruses in control ESCs as well as in two Zfp819 downregulated cells lines. SINE - small interspersed elements; LINE - long interspersed elements; IAP - Intracisternal A particle. **(E)** Schematic diagram showing the typical IAP retroviral element. LTR - long terminal repeat; U3 - U3 regulatory region; 5′ UTR - 5′ untranslated region; *gag* - group antigens; *pol* - polymerase. The locations of the primers used to analyze the IAP elements in ChIP experiments are indicated

with black bars on the bottom. **(F)** Chromatin prepared from ESCs transiently transfected with E2-Zfp819 was subjected to ChIP using either E2-tag specific antibodies or IgG (control) and amplified by PCR using the different IAP primers. The genomic regions of *Mest*, *Hoxa11*, *Actin*, and *Oct3/4* served as negative controls for Zfp819 binding. **(G)** Schematic representation of the luciferase vector containing the IAP U3 regulatory region (top), and the bar graph showing the luciferase activity in either absence or presence of Zfp819 (bottom).

known to bind DNA through their Znf motifs and recruit KAP1 through the interaction of KRAB and RBCC domains, respectively (Wolfe et al. 2000; Peng et al. 2002; Germain-Desprez et al. 2003). This interaction in turn leads to the recruitment of other repressor proteins and subsequent transcriptional inhibition (Urrutia 2003). The strong upregulation of IAP and LINE1 elements in Zfp819_DN cells led us to verify whether Zfp819 binds to these regions. The ChIP with E2-tag specific antibodies on chromatin of E2-Zfp819 transfected ESCs and subsequent PCR analysis revealed that Zfp819 binds to U3, 5′-UTR, and *pol* but not *gag* regions of IAP elements (Fig. 6F). We found no binding of Zfp819 to *Mest*, a known target of KAP1, suggesting the specific binding of Zfp819 to ERV elements (Fig. 6F). As a control, we performed PCR analysis for housekeeping gene as well as *Oct4* promoter regions and found no binding of Zfp819 (Fig. 6F). Further, we performed the luciferase assay with a construct containing the U3 region of IAP, which functions as a cis-regulatory element, and in either absence or presence of Zfp819 overexpression (Fig. 6G). The overexpression of Zfp819 showed ~75% reduction in the luciferase activity (Fig. 6G), confirming that Zfp819 functions in transcriptional silencing of retroviral elements.

Zfp819 depletion leads to enhanced DNA damage and impaired repair

The transcriptional activation of ERVs is known to cause the activation of DNA damage response, as they randomly integrate in the genome (Walsh et al. 1998; Maksakova et al. 2006). The dramatic upregulation of IAP and LINE1 elements in Zfp819_DN cells led us to hypothesize that these

cells might have enhanced DNA damage and subsequent activation of repair processes. To verify this hypothesis, we stained Zfp819_DN cells with γ-H2A.X, a marker of DNA double strand breaks, and identified several foci which are positive for γ-H2A.X, whereas control wildtype cells showed no or low levels of γ-H2A.X foci (Fig. 7B). These results were further confirmed by Western blot analysis, which showed high levels of γ-H2A.X in Zfp819_DN cells compared to low levels of γ-H2A.X in control cells (Fig. 7A). To check whether Zfp819-depleted cells are sensitive to DNA damaging agents, we further treated both Zfp819_DN and wild type cells with neocarzinostatin, a DNA damaging agent, and stained for γ-H2A.X at various time points (Fig. 7B). The wildtype cells showed an increase of γ-H2A.X foci by 1 h of treatment which gradually decreased by 24 h of recovery, suggesting an efficient repair and recovery (Fig. 7B). The Zfp819_DN cells also showed an increased number of γ-H2A.X foci at 3 h, but in contrast to the wildtype cells, these cells displayed a persistent staining for γ-H2A.X even at 48 h of recovery suggesting that DNA damage repair mechanisms are impaired in Zfp819-depleted cells. Moreover, the nuclei of Zfp819_DN cells appeared to be fragmented after 24 h of recovery indicating the activation of apoptosis in these cells (Fig. 7B).

Discussion

The Zinc finger protein family is the most abundant protein family in mammals (Lander et al., 2001). The members of this family are subdivided further according to the presence of additional conserved structural regions/motifs (Bellefroid et al. 1991; Rosati et al. 1991; Collins et al. 2001). The
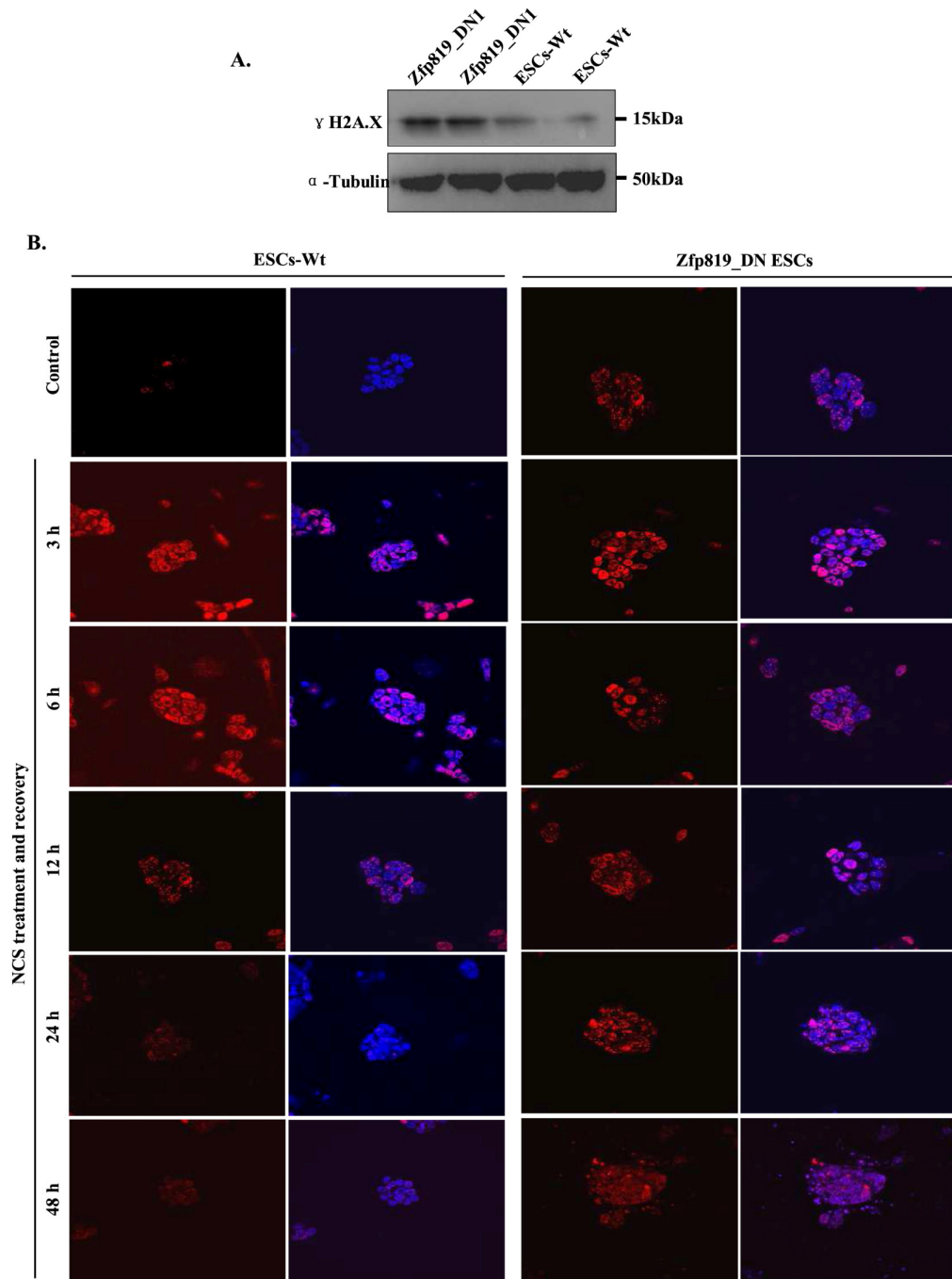
Figure 7. Zfp819 is essential for efficient repair of DNA damage in ESCs. (A) Western blots showing the expression of the DNA damage marker protein γH2A.X in control ESCs and Zfp819 stably downregulated ESCs (Zfp819_DN). **(B)** γH2A.X immunofluorescence staining showing the spontaneous DNA damage in uninduced Zfp819_DN cells, but not in uninduced ESCs-wt as well as the DNA damage in Zfp819_DN cells and ESCs that were treated with the DNA damaging agent NCS and analyzed at various time points of recovery.

KRAB domain-containing zinc finger proteins belong to one such sub-family, account for almost one-third of all Zinc finger proteins and constitute the single largest class of transcription factors in the mammalian genome (Looman et al. 2002). Even after three decades of their initial discovery, the function of many KRAB-zinc finger proteins is still largely unknown. Our previous comparative transcriptome analysis between undifferentiated pluripotent stem cells and their differentiated counterparts led to the finding that Zfp819, a member of the KRAB-zinc finger proteins, is specifically expressed in pluripotent stem cells (Meyer et al. 2010). Since the function of Zfp819 is completely unknown, we set out in the present study to systematically analyze the Zfp819 expression pattern in pluripotent stem cells and adult mouse tissues and to determine its function. Our results show that Zfp819 is expressed preferentially in pluripotent stem cells and involved in transcriptional repression of ERVs and maintenance of genomic integrity.

We demonstrated that the C-terminal part of Zfp819 harboring several zinc finger motifs is the region responsible for its localization to the nucleus. However, we were unable to identify any well-known NLS in this region, highlighting the presence of novel or non-canonical NLS. Recently, Lee et al. (2006) identified three non-canonical NLS motifs named as PY-NLS, which are recognized by Kapβ2, a nucleocytoplasmic trafficking protein. We identified and characterized four PY-NLS motifs in Zfp819 all of which showed efficient nuclear localization of the tagged EGFP.

The open chromatin structure designated by DNA hypomethylation and the presence of active histone modifications is a hallmark of promoter regions of pluripotency-associated genes such as *Oct3/4*, *Sox2*, and *Nanog* (Meshorer and Misteli 2006; Mikkelsen et al. 2007; Doi et al. 2009). Our DNA methylation and histone modification analysis of the *Zfp819* promoter region revealed the presence of open chromatin suggesting that this novel gene might

play an essential role in pluripotent cells. It is known that the members of the core pluripotency transcription factor network such as Oct3/4, Sox2, and Nanog mutually cooperate to activate the expression of genes involved in maintenance of pluripotency (Loh et al. 2006; Chen et al. 2008; Kim et al. 2008; Marson et al. 2008). Our confirmation of Oct3/4, Sox2, and Nanog binding to the DCRM-2 of *Zfp819* further strengthens that the expression of this gene is in turn regulated by the pluripotency network. It is well known that many KRAB zinc finger proteins function as transcriptional repressors through their interaction with KAP1 (Friedman et al. 1996). The transcriptional regulator KAP1 is known to maintain the pluripotency and also required for the terminal differentiation of ESCs (Cammas et al. 2002; Cammas et al. 2004; Hu et al. 2009). Moreover, KAP1 was shown to be important for maintenance of genomic integrity of ESCs by constitutive repression of ERVs (Rowe et al. 2010). In line with these observations, we found a very high activation of ERVs upon Zfp819 downregulation and could show the interaction of Zfp819 with KAP1, suggesting that Zfp819 might recruit KAP1 to the subset of ERVs and maintain their repression in ESCs.

Recently, we identified Chd4 as one of the interaction partners of Zfp819 in a yeast two-hybrid screening experiments and subsequently confirmed this interaction in both *in vivo* and *in vitro* approaches (Zheng et al. 2012). Chd4 is a component of the nucleosome-remodeling and histone deacetylases (NuRD) complex, which also consists of Chd3, Hdac1/2, Mta1 and Mta2 (Lai and Wade 2011). The genetic depletion of Chd4 is known to impair the recruitment of DNA damage response proteins and result in a high sensitivity to DNA damaging agents (Larsen et al. 2010). Both KAP1 and Chd4 have been implicated in DNA damage repair, where they might provide a local chromatin configuration which is necessary for efficient signaling activities and repair (White et al. 2006; Li et al. 2007; Urquhart et al. 2011). It has been shown

that in response to genotoxic stress, ATM phosphorylates KAP1, which facilitates its co-localization with several DNA damage response proteins and induces chromatin decondensation (White et al. 2006; Hu et al. 2012). We observed a very high level of spontaneous DNA damage on the basis of high levels of γH2A.X in unstressed Zfp819-downregulated cells, indicating that these cells are inefficient in DNA damage repair or more prone to DNA damage. Then, we induced DNA damage in these cells and found the prolonged presence of DNA damage foci, supporting the assumption that Zfp819-downregulated cells are inefficient in repairing the damage. Similar to these results, Chd4-deficient cells were shown to have reduced DNA damage repair ability (Smeenk at al., 2010). Collectively, our current study thus highlights that Zfp819 is a novel component of KAP1- and Chd4-mediated genomic integrity maintenance mechanisms of ESCs.

Acknowledgements

References

Bellefroid EJ, Poncelet DA, Lecocq PJ, Revelant O, Martial JA (1991) The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. Proc Natl Acad Sci U S A 88(9): 3608-3612.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ et al. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125(2): 315-326.

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS et al. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122(6): 947-956.

Burnicka-Turek O, Kata A, Buyandelger B, Ebermann L, Kramann N et al. (2010) Pelota interacts with HAX1, EIF3G and SRPX and the resulting protein complexes are associated with the actin cytoskeleton. BMC Cell Biol 11: 28.

Cammas F, Herzog M, Lerouge T, Chambon P, Losson R (2004) Association of the transcriptional corepressor TIF1beta with heterochromatin protein 1 (HP1): an essential role for progression through differentiation. Genes Dev 18(17): 2147-2160.

Cammas F, Mark M, Dolle P, Dierich A, Chambon P et al. (2000) Mice lacking the transcriptional corepressor TIF1beta are defective in early postimplantation development. Development 127(13): 2955-2963.

Cammas F, Oulad-Abdelghani M, Vonesch JL, Huss-Garcia Y, Chambon P et al. (2002) Cell differentiation induces TIF1beta association with centromeric heterochromatin via an HP1 interaction. J Cell Sci 115(Pt 17): 3439-3448.

Chen X, Xu H, Yuan P, Fang F, Huss M et al. (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133(6): 1106-1117.

Collins T, Stone JR, Williams AJ (2001) All in the family: the BTB/POZ, KRAB, and SCAN domains. Mol Cell Biol 21(11): 3609-3615.

Dejosez M, Zwaka TP (2012) Pluripotency and nuclear reprogramming. Annu Rev Biochem 81: 737-765.

Doi A, Park IH, Wen B, Murakami P, Aryee MJ et al. (2009) Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nat Genet 41(12): 1350-1353.

Eiges R, Schuldiner M, Drukker M, Yanuka O, Itskovitz-Eldor J et al. (2001) Establishment of human embryonic stem cell-transfected clones

carrying a marker for undifferentiated cells. Curr Biol 11(7): 514-518.

Friedman JR, Fredericks WJ, Jensen DE, Speicher DW, Huang XP et al. (1996) KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev 10(16): 2067-2078.

Germain-Desprez D, Bazinet M, Bouvier M, Aubry M (2003) Oligomerization of transcriptional intermediary factor 1 regulators and interaction with ZNF74 nuclear matrix protein revealed by bioluminescence resonance energy transfer in living cells. J Biol Chem 278(25): 22367-22373.

Gietz RD, Schiestl RH, Willems AR, Woods RA (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast 11(4): 355-360.

Hu C, Zhang S, Gao X, Gao X, Xu X et al. (2012) Roles of Kruppel-associated Box (KRAB)-associated Co-repressor KAP1 Ser-473 Phosphorylation in DNA Damage Response. J Biol Chem 287(23): 18937-18952.

Hu G, Kim J, Xu Q, Leng Y, Orkin SH et al. (2009) A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes Dev 23(7): 837-848.

Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J et al. (2006) Dissecting self-renewal in stem cells with RNA interference. Nature 442(7102): 533-538.

Kim J, Chu J, Shen X, Wang J, Orkin SH (2008) An extended transcriptional network for pluripotency of embryonic stem cells. Cell 132(6): 1049-1061.

Lai AY, Wade PA (2011) Cancer biology and NuRD: a multifaceted chromatin remodelling complex. Nat Rev Cancer 11(8): 588-596.

Larsen DH, Poinsignon C, Gudjonsson T, Dinant C, Payne MR et al. (2010) The chromatin-remodeling factor CHD4 coordinates signaling and repair after DNA damage. J Cell Biol 190(5): 731-740.

Li X, Lin HH, Chen H, Xu X, Shih HM et al. (2010) SUMOylation of the transcriptional co-repressor KAP1 is regulated by the serine and threonine phosphatase PP1. Sci Signal 3(119): ra32.

Li X, Lee YK, Jeng JC, Yen Y, Schultz DC et al. (2007) Role for KAP1 serine 824 phosphorylation and sumoylation/desumoylation switch in regulating KAP1-mediated transcriptional repression. J Biol Chem 282(50): 36177-36189.

Loh YH, Wu Q, Chew JL, Vega VB, Zhang W et al. (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet 38(4): 431-440.

Looman C, Abrink M, Mark C, Hellman L (2002) KRAB zinc finger proteins: an analysis of the molecular mechanisms governing their increase in numbers and complexity during evolution. Mol Biol Evol 19(12): 2118-2130.

Maksakova IA, Romanish MT, Gagnier L, Dunn CA, van de Lagemaat LN et al. (2006) Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. PLoS Genet 2(1): e2.

Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T et al. (2008) Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell 134(3): 521-533.

Meshorer E (2007) Chromatin in embryonic stem cell neuronal differentiation. Histol Histopathol 22(3): 311-319.

Meshorer E, Misteli T (2006) Chromatin in pluripotent embryonic stem cells and differentiation. Nat Rev Mol Cell Biol 7(7): 540-546.

Meyer S, Nolte J, Opitz L, Salinas-Riester G, Engel W (2010) Pluripotent embryonic stem cells and multipotent adult germline stem cells reveal similar transcriptomes including pluripotency-related genes. Mol Hum Reprod 16(11): 846-855.

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E et al. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448(7153): 553-560.

Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M et al. (2003) The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 113(5): 631-642.

Ng HH, Surani MA (2011) The transcriptional and signalling networks of pluripotency. Nat Cell Biol 13(5): 490-496.

Nishii T, Oikawa Y, Ishida Y, Kawaichi M, Matsuda E (2012) CtBP-interacting BTB zinc finger protein (CIBZ) promotes proliferation and G1/S transition in embryonic stem cells via Nanog. J Biol Chem 287(15): 12417-12424.

Pan G, Tian S, Nie J, Yang C, Ruotti V et al. (2007) Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell 1(3): 299-312.

Peng H, Feldman I, Rauscher FJ, 3rd (2002) Hetero-oligomerization among the TIF family of RBCC/TRIM domain-containing nuclear cofactors: a potential mechanism for regulating the switch between coactivation and corepression. J Mol Biol 320(3): 629-644.

Peng H, Begg GE, Harper SL, Friedman JR, Speicher DW et al. (2000) Biochemical analysis of the Kruppel-associated box (KRAB) transcriptional repression domain. J Biol Chem 275(24): 18000-18010.

Rosati M, Marino M, Franze A, Tramontano A, Grimaldi G (1991) Members of the zinc finger protein gene family sharing a conserved N-terminal module. Nucleic Acids Res 19(20): 5661-5667.

Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S et al. (2010) KAP1 controls endogenous retroviruses in embryonic stem cells. Nature 463(7278): 237-240.

Seki Y, Kurisaki A, Watanabe-Susaki K, Nakajima Y, Nakanishi M et al. (2010) TIF1beta regulates the pluripotency of embryonic stem cells in a phosphorylation-dependent manner. Proc Natl Acad Sci U S A 107(24): 10926-10931.

Smorag L, Zheng Y, Nolte J, Zechner U, Engel W et al. (2012) MicroRNA signature in various cell types of mouse spermatogenesis: Evidence for stage-specifically expressed miRNA-221, -203, and -34b-5p mediated spermatogenesis regulation. Biol Cell.

Sridharan R, Tchieu J, Mason MJ, Yachechko R, Kuoy E et al. (2009) Role of the murine reprogramming factors in the induction of pluripotency. Cell 136(2): 364-377.

Urquhart AJ, Gatei M, Richard DJ, Khanna KK (2011) ATM mediated phosphorylation of CHD4 contributes to genome maintenance. Genome Integr 2(1): 1.

Urrutia R (2003) KRAB-containing zinc-finger repressor proteins. Genome Biol 4(10): 231.

Walsh CP, Chaillet JR, Bestor TH (1998) Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet 20(2): 116-117.

Wang ZX, Teh CH, Chan CM, Chu C, Rossbach M et al. (2008) The transcription factor Zfp281 controls embryonic stem cell pluripotency by direct activation and repression of target genes. Stem Cells 26(11): 2791-2799.

Wang ZX, Kueh JL, Teh CH, Rossbach M, Lim L et al. (2007) Zfp206 is a transcription factor that controls pluripotency of embryonic stem cells. Stem Cells 25(9): 2173-2182.

White DE, Negorev D, Peng II, Ivanov AV, Maul GG et al. (2006) KAP1, a novel substrate for PIKK family members, colocalizes with numerous damage response factors at DNA lesions. Cancer Res 66(24): 11594-11599.

Witzgall R, O'Leary E, Leaf A, Onaldi D, Bonventre JV (1994) The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins

mediates transcriptional repression. Proc Natl Acad Sci U S A 91(10): 4514-4518.

Wolf D, Goff SP (2009) Embryonic stem cells use ZFP809 to silence retroviral DNAs. Nature 458(7242): 1201-1204.

Wolfe SA, Nekludova L, Pabo CO (2000) DNA recognition by Cys2His2 zinc finger proteins. Annu Rev Biophys Biomol Struct 29: 183-212.

Yu HB, Kunarso G, Hong FH, Stanton LW (2009) Zfp206, Oct4, and Sox2 are integrated components of a transcriptional regulatory network in embryonic stem cells. J Biol Chem 284(45): 31327-31335.

Zechner U, Nolte J, Wolf M, Shirneshan K, Hajj NE et al. (2009) Comparative methylation profiles and telomerase biology of mouse multipotent adult germline stem cells and embryonic stem cells. Mol Hum Reprod 15(6): 345-353.

Zhang J, Tam WL, Tong GQ, Wu Q, Chan HY et al. (2006a) Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nat Cell Biol 8(10): 1114-1123.

Zhang JZ, Gao W, Yang HB, Zhang B, Zhu ZY et al. (2006b) Screening for genes essential for mouse embryonic stem cell self-renewal using a subtractive RNA interference library. Stem Cells 24(12): 2661-2668.

Zhang W, Walker E, Tamplin OJ, Rossant J, Stanford WL et al. (2006c) Zfp206 regulates ES

cell gene expression and differentiation. Nucleic Acids Res 34(17): 4780-4790.

Zhang X, Zhang J, Wang T, Esteban MA, Pei D (2008) Esrrb activates Oct4 transcription and sustains self-renewal and pluripotency in embryonic stem cells. J Biol Chem 283(51): 35825-35833.

Zheng Y, Tan X, Pyczek J, Nolte J, Pantakani DV et al. (2012) Generation and Characterization of Yeast Two-Hybrid cDNA Libraries Derived From Two Distinct Mouse Pluripotent Cell Types. Mol Biotechnol.

Ziv Y, Bielopolski D, Galanty Y, Lukas C, Taya Y et al. (2006) Chromatin relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. Nat Cell Biol 8(8): 870-876.

Zuo X, Sheng J, Lau HT, McDonald CM, Andrade M et al. (2011) Zinc finger protein ZFP57 requires its co-factor to recruit DNA methyltransferases and maintains DNA methylation imprint in embryonic stem cells via its transcriptional repression domain. J Biol Chem 287(3): 2107-2118
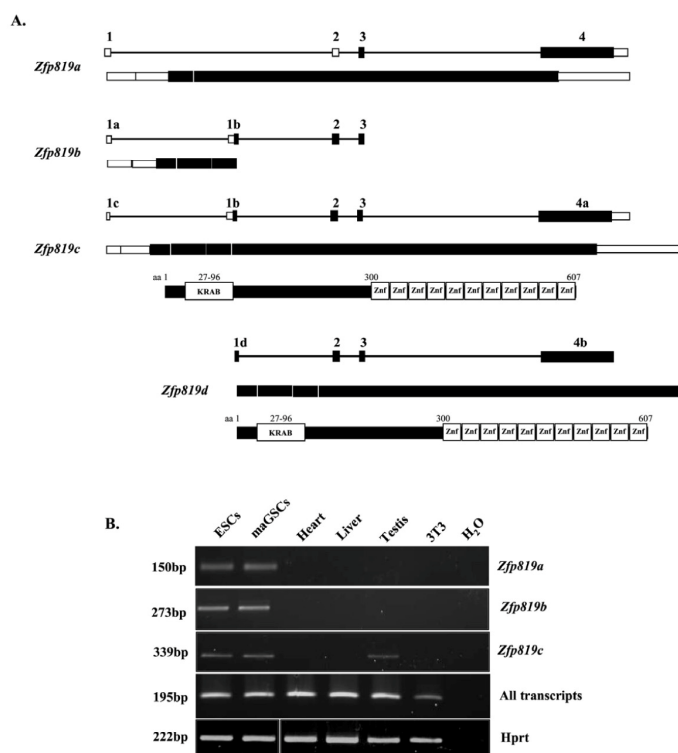
Figure S1. *Zfp189* **splice variants are specifically expressed in ESCs and testis. (A)** Schematic illustration of *Zfp819* transcript variants and their genomic organization as depicted in ENSEMBL database. *Zfp819* transcripts and two annotated protein structures are depicted below the each splice variant. **(B)** RT-PCR data showing the presence of *Zfp819* transcript variants in pluripotent cell types (ESCs and maGSCs) as well as in selected adult mouse tissues.

```
ZF1    302  CHICGKSFLQKSELSFHPGTDRGETHHG  329
ZF2    330  CPGCRKSLPSTSSLQARHGTHTKEKPYR  357
ZF3    358  CPDCGKSFSYASHLKIHLRIHTGERPYV  385
ZF4    386  CSDCGKAFSQKSVLTTHRRIHTGERPYT  413
ZF5    414  CNHCGKLFVYASDLKKHSRFHTGEKPYE  441
ZF6    442  CRDCGKLFNNKSHLPVHYRIHTGEKPYR  469
ZF7    470  CCDCGKSFRRKSHLKVHSRTHTGEKPYE  497
ZF8    498  CPDCGRAFSHSSVLSTHQRIHTGERPYL  525
ZF9    526  CSDCGKAMSSKAQLNEHRRIHTGEKPYV  553
ZF10   554  CSECGKAFSCRSSLQAHRRTHSREKPFV  581
ZF11   582  CHTCGKGFLHKSQLSSHQQTHTGENP    607
```

Figure S2. Zinc finger motifs of Zfp819 contain non-canonical nuclear localization signals. Multiple alignment of 11 zinc finger motifs (ZF1 to ZF11) of Zfp819 along with their respective amino acid number. The non-canonical nuclear localization signals (NLSs), PY-NLS motifs are highlighted. The PY residues of ZF are indicated in red. $R/K/HX_{(2-5)}PY$ NLS is underlined. $\Phi G/A/S\Phi\Phi X_{(11-13)}PY$ and basic enriched$_{(5-8)}X_{(8-10)}PY$ NLS sequences are highlighted in aqua and yellow color, respectively.

4 Discussion

Pluripotency is established and maintained by many pluripotent stem cell-specific transcription factors and their coordinated autoregulatory and feedforward transcriptional regulatory mechanisms (Loh et al., 2006; Kashyap et al., 2009). Recent genome-wide studies on pluripotent stem cells have uncovered several new players in the transcriptional circuitry that maintain the pluripotency and differentiation potential of these cells. In the present study, we developed two high-quality yeast two-hybrid (Y2H)-cDNA libraries, which can help in identification and construction of protein interactome of known and unknown pluripotent stem cell-specific protein(s). As a proof of concept, we have identified several interaction proteins of Zfp819, a novel pluripotent stem cell-specific factor, using one of these Y2H-cDNA libraries. The identification of Zfp819 interaction partners together with the molecular and functional studies revealed that *Zfp819* functions in the regulation of genomic stability of pluripotent cells by suppressing a subset of endogenous retroviruses (*ERVs*).

4.1 Methods for detection of protein-protein interactions

Many biological processes are accomplished by the function and regulation of multi-protein complexes in which one particular protein can be involved in various biological processes and might perform distinct functions. This extraordinary feat is achieved by formation of different protein complexes through interaction with various proteins. Identification of protein-protein interactions, the protein interactome, is therefore essential to understand the molecular function of any particular protein and its related biological processes. To this end, construction of protein interactomes has emerged as a new field to identify the proteins working in a concert to regulate various biological processes. Most human diseases account for the disruption of protein complexes and their functions due to mutations in either a single gene or in many genes (Lim et al., 2006; Lage et al., 2007; Vanunu et al., 2010). Therefore, protein interactomes have paved a way for medical researchers to seek for the protein interaction partners to dissect the pathological mechanism for many disease models (Oti et al., 2006; Lage et al., 2007).

A variety of high throughput methods have been applied in global investigations of the interaction and relationship between two or more proteins. Currently, there are several

methods (Table 4.1) which are available to identify the protein-protein interaction networks. The high sensitivity and specificity are two primary pursuits for researchers. However, each of these approaches has its own advantages and drawbacks regarding the identification of protein interactions. Among the available techniques, Y2H system

Table 4.1. Overview of various methods for identification of protein-protein interactions with a brief description of the method and the reference(s). (Table adapted and modified from Bruckner et al., 2009).

Method (Type)	Description	Reference
Yeast two-hybrid (Y2H) assay (*in vivo*)	The DNA binding domain and the activation domain of a transcription factor are separated and fused to two individual proteins. Only when two proteins interact with each other or come in close proximity leads to the transcriptional activation of reporter genes, thus the identification of protein interaction partner.	Walhout and Vidal., 2001
Pull-down assay (*in vitro*)	The recombinant protein immobilized on the resin will aid in identification of protein interaction partners after passing the whole cell/tissue protein extract through the resin. After several steps to eliminate the unspecific bindings, interaction proteins are eluted and identified by mass spectrometry analysis.	Jackson et al., 2002 Tanaka et al., 2006 Burklen et al., 2007
Immunoprecipitation coupled mass spectrometry (IP-MS) (*in vitro*)	IP-MS is based on protein immunoprecipitation (IP) method. The IPed fraction containing various proteins is subjected to peptide mass spectrometry analysis.	Ewing et al., 2007
Stable isotope labeling by amino acids in cells (SILAC) (*in vivo*)	This method is an improved technique from IP-MS. It is based on the incorporation of amino acids with substituted stable isotopic nuclei. SILAC can not only detect the potential interaction partners but also quantify their ratios.	Ong et al., 2002
Targeted releasable affinity probe (TRAP) (*in vivo*)	A targeted reversible affinity probe called TRAP can find and couple with the bait protein. Then the TRAP can hunt the proteins which are nearby. Subsequently, the prey will be pulled down and identified by mass spectrometry.	Yan et al., 2009

allows the most cost-effective and genomic-scale screening for protein-protein interactions with the potential to detect weak and transient interactions based on *in vivo* reporter systems (Fields and Song, 1989; Young, 1998). This approach relies on the activation of downstream reporter genes by the *GAL4* based system (Fields and Song, 1989; Young, 1998), in which the transcriptional activator *GAL4* is split into DNA-binding domain (BD) and activation domain (AD) and fused to bait and prey, respectively. The interaction between bait and prey proteins brings the *GAL4* domains into close proximity leading to the transcriptional activation of downstream reporter genes (Fields and Song, 1989).

The putative interaction partners identified by conventional Y2H methods are tend to produce false positive results at a low frequency, which are caused by the auto-activation of the prey protein. Therefore, a modified Y2H, the repressed transactivator (RTA) system, was developed in order to inhibit the transcriptional activation of the prey proteins (Table 4.2). Moreover, the classic Y2H system cannot be applied to study the interaction of membrane associated proteins. Although the truncated proteins can be used and applied for classic Y2H system, it can lead to protein misfolding and can provide a high rate of false negatives (Sugita et al., 1996; Niethammer et al., 1996; Borg et al., 1999). Therefore, the Son of Sevenless (SOS) and the Rat Sarcoma (RAS)-Recruitment Systems (SRS and RRS) were developed to use the RAS signaling activation as a readout, which is a cytosolic rather than the traditional transcriptional readout system (Table 4.2). In these two systems, the soluble bait protein is fused to mammalian SOS or RAS gene and its interaction with membrane localized prey proteins results in the activation of RAS signaling pathway and thereby in the survival of the cell. Taken together, the classical Y2H and its derivatives have shown powers by its methodological diversity and technical simplicity to rapidly screen and identify large amounts of reliable interaction partners of a protein.

Due to lack of a pluripotent stem cell-specific Y2H cDNA library the Y2H approach was not feasible to study the pluripotency-related protein interactome. Therefore, in the present study, we constructed high quality cDNA libraries of mouse ESCs and maGSCs in a *GAL4*-based Y2H vector system. Subsequently, to characterize each library, we used *Zfp819* and *Stra8* (Stimulated by retinoic acid) as baits to screen interaction proteins using ESCs and maGSC libraries, respectively. Consequently, several putative interaction proteins of our baits could be identified (Zheng et al., 2012).

Table 4.2. Summary of classic Y2H method and its derivatives for detection of protein interaction partners which are in different sub-cellular localizations. (Table adapted and modified from Bruckner et al., 2009).

Y2H method	Possible baits	Response	Cellular compartment	Reference
Classic Y2H system	Non-transactivating, proteins capable of entering nucleus	Transcriptional activation	Nucleus	Fields and Song, 1989
SOS recruitment system (SRS)	Transactivating, cytosolic proteins	Ras signalling	Membrane periphery	Aronheim et al.,1994
Split-ubiquitin system	Nuclear, membrane and cytosolic proteins	Uracil auxotrophy and 5-FoA resistance	Cytosol	Johnsson and Varshavsky, 1994
Membrane split-ubiquitin system (MbY2H)	Membrane proteins	Transcriptional activation	Membrane periphery	Stagljar et al., 1998
Ras recruitment system (RRS)	Transactivating, cytosolic proteins	Ras signalling	Membrane periphery	Broder et al., 1998
Dual bait system	Two non-transactivating proteins capable of entering nucleus	Transcriptional activation	Nucleus	Serebriiskii et al., 2002
RNA polymerase III based two-hybrid (Pol III)	Transactivating proteins (in the RNA polymerase II pathway)	Transcriptional activation	Nucleus	Petrascheck et al., 2001
Repressed transactivator system (RTA)	Transactivating, proteins capable of entering nucleus	Inhibition of transcription	Nucleus	Hirst et al., 2001
Reverse Ras recruitment system (rRRS)	Membrane proteins	Ras signalling	Membrane periphery	Hubsman et al., 2001
Split-Trp system	Cytosolic, membrane proteins	Trp1p activity	Cytosol	Tafelmeyer et al., 2004
Cytosolic split-ubiquitin system (cytoY2H)	Transactivating, cytosolic proteins	Transcriptional activation	ER membrane periphery	Mockli et al., 2007

Among the 17 identified interaction proteins of Zfp819, we recognized that several of these proteins such as Chd4, Ctc1, and H3f3b are implicated in the genome integrity maintenance mechanisms. *Chd4* is the core component of nucleosome-remodeling and histone deacetylases complex (NuRD), which is also composed of *Hdac1/2, Chd3, Mta1* and *Mta2* (Lai and Wade, 2011). The loss of *Chd4* function was found to result in early embryonic lethality and accumulation of DNA damage in cells (Larsen at al., 2010). *Ctc1* (CST telomere maintenance complex component 1) encodes a member of the mammalian homolog of yeast heterotrimeric CST telomeric capping complex. Mutations in human Ctc1 were shown to increase the spontaneous DNA damage (Surovtseva et al., 2009). *H3f3b* encodes a replacement histone variant which belongs to the histone 3 family (*H3.3A* and *H3.3B*) (Wu and Bonner, 1981; Brush et al., 1985; Wells et al., 1987) and was shown to function in genome stability and cell cycle control (De Luca et al., 2011). We could validate the interaction of Zfp819 with several of these interaction proteins using independent *in vitro* and *in vivo* methods such as co-immunoprecipitation (Co-IP), GST-pull down assay and co-localization studies. Collectively, the identification of Zfp819 interaction proteins directed us for the first time to study its possible role in transcriptional regulation and genomic integrity maintenance mechanisms. Similarly, we used maGSCs Y2H-cDNA library to screen for Stra8 interaction proteins (Zheng et al., 2012). Several of these putative interaction proteins are implicated in chromatin assembly/modification and transcriptional regulation processes highlighting that *Stra8* might function in these processes. Furthermore, we selected AT-rich interactive domain 4B (Arid4B) as a candidate and validated its interaction with Stra8 using Co-IP and co-localization studies (Zheng et al., 2012).

We generated two high quality Y2H-cDNA libraries from mouse ESCs and maGSCs and used them to identify protein interaction networks and to understand the molecular function of two pluripotency-related proteins, whose functions are unknown until now. Thus, we conclude that our pluripotent stem cell-specific Y2H cDNA libraries are powerful tools to study the protein interactomes of pluripotent stem cells.

4.2 KRAB zinc finger proteins function in pluripotent cells and in embryonic development

Zinc finger proteins are one of the most abundant protein families in eukaryotic genomes (Urrutia, 2003). About one third of these zinc finger proteins bear a KRAB domain which is responsible for protein-protein interactions, whereas the C-terminus Cys_2His_2-zinc finger motif facilitates the recognition of a specific DNA sequence thereby leading to either gene activation or repression (Witzgall et al., 1994; Looman et al., 2002). Several lines of evidence suggest that KRAB-zinc finger proteins play an important role in maintenance of pluripotency, differentiation potential, cell proliferation, and cell cycle control as well as embryonic development. As outlined in Table 4.3, KRAB-zinc finger proteins can function in a *Kap1* dependent manner as retrovirus repression factors (Wolf and Goff, 2009) or genomic methylation imprints maintenance factors in ESCs (Zuo et al., 2011). Abnormal expression of KRAB-zinc finger proteins was reported to interfere with the lineage commitment of ESCs (Perrotti et al., 1995). Additionally, KRAB-zinc finger proteins are also found to be important for embryonic development and organ formation (Costoya et al., 2004; Garcia-Garcia et al., 2008; Krebs et al., 2012). Another group of

Table 4.3. The list of Zinc finger proteins with a known function in pluripotent stem cells. The function of each zinc finger protein along with their reported expression pattern and the associated reference are indicated. * Stands for KRAB zinc finger protein. *ERV-* endogenous retrovirus elements; ICR- imprinting control region

Gene	Function	Expression in organism	References
Zfp809*	*ERV* repression	Ubiquitously expressed	Wolf and Goff, 2009
ZFP57*	ICR methylation regulation	Brain, heart, kidney	Li et al., 2008; Zuo et al., 2011
Zfp568*	Embryo development	Brain, spinal cord, heart, liver	Garcia-Garcia et al., 2008
MZF1*	ESC differentiation	Ubiquitously expressed	Perrotti et al., 1995
Zfp819*	*ERV* repression	Testis, ovary, heart	Present study
Zfp206	Pluripotency maintainance	Through embryogenesis and postnatal testis	Yu et al., 2009
Zfx	ESCs self-renewal	Kidney	Galan-Caridad et al., 2007
Zic3	Pluripotency maintainance	Ubiquitously expressed	Lim et al., 2007
Zfp42	Pluripotency maintainance	Placenta and testis	Masui et al., 2008
Sall4	Pluripotency regulation	Kidney, liver, heart, brain	Yang et al., 2008
Zfp296	Inhibition of differentiation	Kidney	Fischedick et al., 2012
Klf4	Pluripotency maintainance	Ubiquitously expressed	Bourillot and Savatier, 2010
Zfp281	Pluripotency maintainance	Ubiquitously expressed	Wang et al., 2008

KRAB-zinc finger proteins is essential for tumor suppression and they are implicated in regulation of cell proliferation, induction of apoptosis through inhibition of oncogene(s) expression (Huang et al., 2007; Cheng et al., 2010).

Although numerous studies have revealed the function of some of the KRAB-zinc finger proteins, the functions of a large number of KRAB-zinc finger proteins have to be further determined. In the present study (Tan et al., 2012, submitted), we identified *Zfp819*, a KRAB-zinc finger protein as a pluripotency-related gene. Interestingly, like many of the pluripotency-related genes, *Zfp819* was confirmed as a transcriptional target of *Oct3/4*, *Sox2*, and *Nanog* in our study. The high expression of *Zfp819* in various pluripotent cell types and the presence of transcriptionally active epigenetic marks at its promoter region suggest a possible crucial role of *Zfp819* in pluripotent cells.

4.3 Genome integrity maintenance in pluripotent stem cells

All eukaryotic organisms have evolved with an organized process of cell division, which enables the genetic information to be accurately copied and distributed to the daughter cells. The maintenance of genome integrity is essential not only for the health of the individual organism but also for the continued survival and perpetuation of particular species (Tichy, 2011). However, the genome is under constant attack from endogenous and exogenous reactive chemical substances as well as viral infections thus likely leading to DNA damage (Geacintov and Broyde, 2010).

Pluripotent stem cells are capable of giving rise to all the cell types in the entire organism. Any mutation or variation in stem cells could be fatal to the developing organism and even can affect the viability of the animal. In order to guard the genomic integrity, ESCs are highly sensitive to genotoxic reagents to repair damaged DNA or eliminate the defected cells, demonstrating notably lower spontaneous mutation frequency than somatic cells (Hong et al., 2006; Chuykin et al., 2008).

The formation of DNA double strand breaks (DSBs) is the most severe form of DNA damage. Once the DSBs have happened, the DNA damage sensor protein *ATM* is immediately activated (Momcilovic et al., 2010) and phosphorylates the histone protein, *H2AX*, at ser-139. The phosphorylation of *H2AX (γ-H2AX)* is known as an indicator of the presence of DSBs (Rogakou et al., 1998) and was found to accumulate at the foci where several DSBs repair proteins are recruited (Paull et al., 2000). Recent studies have

identified a number of proteins which are involved in the maintenance of genome integrity. In 2010, *Kap1* was found to play a role in genome protection through repression of *ERVs* (Rowe et al., 2010). Additionally, *ATM*-mediated phosphorylation of *Kap1* leads to the latter's co-localization with *γ-H2AX*, reflecting its correlation to DNA damage repair (White et al., 2006). Consistent with this, cells with loss of phosphorylated *Kap1* were shown to be hypersensitive to induced DNA damage and showed defects in chromatin decondensation, which is otherwise essential for repair process (Ziv et al., 2006). Similarly, the chromodomain helicase protein, *Chd4* was recently reported as a component of genome guardian machinery (Polo et al., 2010; Urquhart et al., 2011). The *Chd4* depletion gives rise to accumulation of spontaneous DNA damage as well as elevated sensitivity for irradiation (Larsen at al., 2010). More recently, Pan et al. (2012) identified *Chd4* as an important factor for controlling homologous recombination (HR) during the DNA damage repair process.

In our studies (Zheng et al., 2012; Tan et al., 2012, submitted) we identified Kap1 and Chd4 as interaction proteins of Zfp819 indicating that this novel KRAB-zinc finger protein may work in concert with *Kap1* and *Chd4* to maintain the genomic integrity in pluripotent stem cells. Furthermore, we also identified *Ctc1*, a component of telomere maintenance complex, and H3f3b, a histone variant, as potential interaction protein of Zfp819 (Zheng et al., 2012). Interestingly, these two interaction proteins were also implicated in regulating the genome stability and cell cycle control (Surovtseva et al., 2009; De Luca et al., 2011), further strengthening the possible role of *Zfp819* in genome integrity maintenance. In lines with these indications, we observed a very high level of spontaneous DNA damage in unstressed *Zfp819*-downregulated cells, indicating that *Zfp819* depleted cells are inefficient in DNA damage repair or more prone to DNA damage.

*ERV*s are another group of factors which pose threat to the genome stability of cells. As they can invade and duplicate within the host genome due to their transposition characteristic, *ERV*s are largely silenced. However, few elements of *ERV*s are expressed in placenta of human and other animals (Kalter et al., 1973; Levinson et al., 1978), lymphoid cells (Medstrand et al., 1992), mammary carcinoma cell lines (Patience et al., 1996) as well as in mouse germ cells and preimplantation embryos (Brulet et al., 1985; Poznanski et al., 1991; Kigami et al., 2003; Evsikov et al., 2004). Very recently, Macfarlan et al. (2012) found a small portion of ESCs/iPSCs which expresses high levels

of 2-cell stage specific-*ERV (MERVL-Gag)*. These 2C-like ESCs were shown to possess a potential to contribute not only to embryonic but also to extraembryonic tissues. This astonishing discovery suggested that the existence of *ERV*s is helpful for driving cell fate regulation in placental mammals.

Since the retrotransposition can affect gene expression and lead to genome instability and other phenotypes (Doolittle et al., 1982; Maksakova et al., 2006), these *ERV*s are highly repressed in ESCs (Maksakova et al., 2011). A variety of pathways have been employed to repress the expression of them and to minimize the harmful impact on the host genome (Walsh and Bestor, 1999; Pannell et al., 2000; Matsui et al., 2010). DNA methylation is one such mechanism which is required for the stable repression of genes including *ERV*s, but is dispensable for *ERV*s suppression in pluripotent stem cells (Okano et al., 1999). In ESCs it is likely that there is also another alternative pathway involving histone modifications to suppress *ERVs* (Leung and Lorincz, 2011). Intriguingly, the evolution of KRAB-zinc finger proteins along with the appearance of *ERVs*, together with the functional data suggest that KRAB-zinc finger proteins recognize *ERVs* and repress their expression in a *Kap1*-dependent manner (Emerson and Thomas, 2009; Wolf and Goff, 2009).

The protein structure of *Kap1* contains an N-terminal RBCC domain, a TIF1 signature sequence (TSS) and Heterochromatin protein 1 *(HP1)* binding domain in the centre with plant homeodomain (PHD) and bromodomain at the C-terminal (Iyengar and Farnham, 2011). The N-terminal RBCC domain is sufficient for interaction with KRAB-zinc fingers. This interaction contributes to the silencing of the specific target genes which are recognized by KRAB-zinc fingers (Peng et al., 2000). *Kap1* is thought to function as a co-repressor by assembling a complex with HP1 (Ryan et al., 1999). Further, PHD-bromodomain (PB) of *Kap1* is physically associated with three chromatin modifying enzymes: *ESET* (Schultz et al., 2002), whose H3K9 histone methyltransferase activity suggests that *KAP1* regulates transcription via changes in histone modifications at specific target sites. Further, the interaction of *Kap1* with *Chd3*, a subunit of NuRD transcription repression complex (Schultz et al., 2001), as well as with the *histone deacetylases (HDACs)* (Nielsen et al., 1999) links the histone deacetylation to *Kap1* based gene repression. Fig 4.1 shows how *Kap1* and KRAB zinc finger protein, *Zfp809*, regulate the repression of murine leukemia virus (*MLV*) in ESCs. Interestingly, we found very high expression levels of *ERV*s in *Zfp819*-downregulated cells indicating that

Zfp819 function in the suppression of *ERV*s (Tan et al., 2012, submitted). These observations are further strengthened by identifying a physical interaction between Zfp819 and Kap1/Chd4 (Tan et al., 2012, submitted; Zheng et al., 2012).
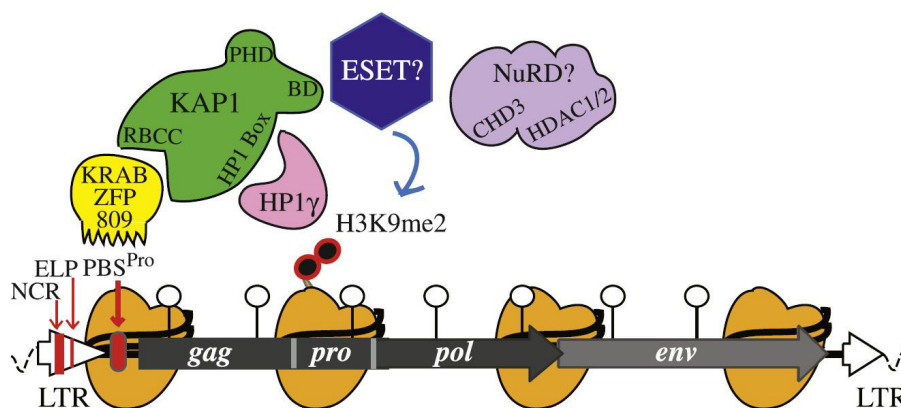

Figure 4.1 Mechanism of *ERV*s silencing by KRAB-zinc finger protein and Kap1 in embryonic stem cells. The gene expression of *murine leukemia virus (MLV)*, a type of *ERV*s, is regulated in *cis* by *Kap1* and KRAB-zinc finger protein complex. The primer binding site (PBS) is essential to recruit *Zfp809*. Subsequent interaction between *Zfp809* and *Kap1* leads to the interaction with *HP1* to form *ZFP809-KAP1-HP1*complex. This leads to the establishment of repression chromatin marks H3K9me2 and virus silencing (Wolf and Goff, 2009; Figure adapted from Rowe and Trono, 2011).

4.4 Generation of gain-of-function and loss-of-function mouse models for Zfp819

Genetically modified mice are commonly used animal models for demonstration of gene function *in vivo*. In the present study, we endeavored to develop both gain- and loss-of-function mouse models of *Zfp819* to better understand its role in embryonic development as well as in adult life. Firstly, we generated two *Zfp819* overexpression models in which one model expresses *Zfp819* constitutively whereas the other one is an inducible model (Fig.4.2). For the generation of constitutive model, the open reading frame (ORF) of *Zfp819* along with additional C-terminal E2-epitope tag was cloned downstream of human elongation factor 1 alpha (hEF1α) promoter (Fig.4.2A). For the generation of the inducible model, we took advantage of the Tet-On system in which the expression of a cloned gene can be temporarily controlled by doxycycline supplementation. The *Zfp819* ORF together with the FLAG-tag was cloned downstream of Tet responsive element (Tre) in a pTre vector system (Fig. 4.2B).
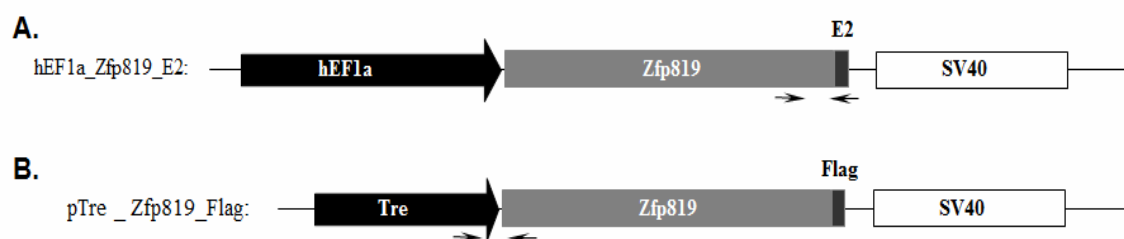
Figure 4.2 The scheme of two plasmid constructs used to generate *Zfp819* overexpression models. (A) The constitutive expression vector in which the E2-tag fused *Zfp819* will be expressed ubiquitously from highly active and constitutive promoter, hEF1α. **(B)** The schematic diagram showing the construct for producing Doxycycline inducible transgenic mouse line. An epitope Flag is tagged to the C- terminus of *Zfp819* ORF, which is under the control of the Tet responsive element (Tre). The arrows indicate the location for the genotyping primers.

The generated *Zfp819* overexpression constructs were used for pronuclear injection. Subsequently, we obtained several founders for each transgenic line. The germline transmission of founders was confirmed by the genotyping PCR of their progeny (Fig. 4.3). The generated pTre mice will be bred with already available Tet-On transgenic mice to establish a double transgenic mouse line, which will then be used to activate the expression of *Zfp819* using doxycycline. Currently, the analysis of transgene integration site and copy numbers is in progress. Afterwards, the transgenic lines with high and low expression levels will be established to study the role of *Zfp819* overexpression on mouse development.
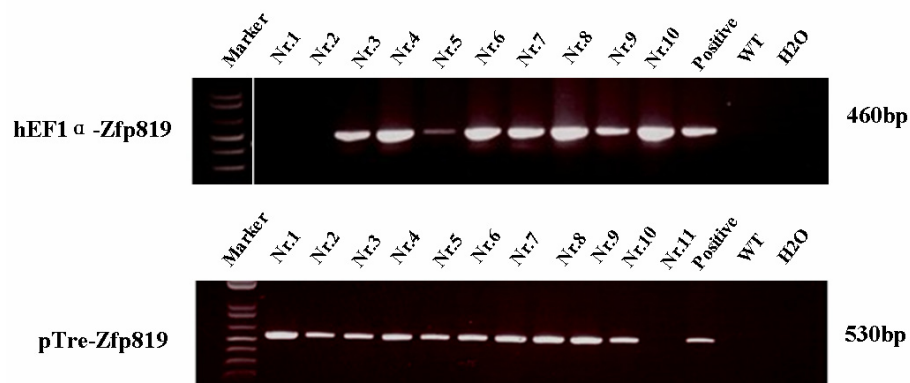
Figure 4.3 The genotyping-PCR results of *Zfp819* constitutive (top) and inducible (bottom) transgenic mouse live progeny. This is an example of genotyping showing that 8 out of 10 from one liter are positive for *hEF1α-Zfp819* construct, whereas10 out of 11 from one liter are positive for *pTre-Zfp819* construct.

Generation of loss-of-function/knockout (KO) mouse models holds great promise to study the function of a gene (s) of interest and its associated phenotypes. Towards this end, we have obtained *Zfp819* conditional KO ES cells from the International Knockout Mouse Consortium (IKMC). The blastocyst injections of these cells resulted in generation of high percentage chimeras. Although we tried to establish a homozygous

conditional KO mouse line using progeny of these chimeras, we were unable to obtain any offspring which are homozygous for the conditional allele. Subsequent analysis indicated that the conditional KO construct is integrated at the *En2* gene locus (data not shown). Then we began to generate the classical KO construct using bacterial artificial chromosome (BAC)-mediated recombination approach (Liu et al., 2003). The strategy for generating the *Zfp819* null allele is depicted in Fig. 4.4 A. Briefly, the BAC (CH29-485C23), which contains the complete genomic locus of *Zfp819*, was used for the BAC-mediated homologous recombination. The exons 2, 3 and 4 of *Zfp819* which code for KRAB domain were replaced by a neomycin (neo) resistance cassette using homologous recombination. We confirmed the integration of neomycin cassette into the *Zfp819* locus and the deletion of targeted exons 2, 3 and 4 in the BAC (Modified BAC) (Figure 4.4 B).

A.
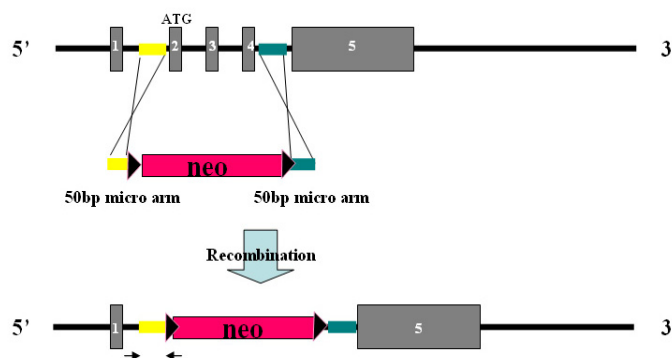
B.
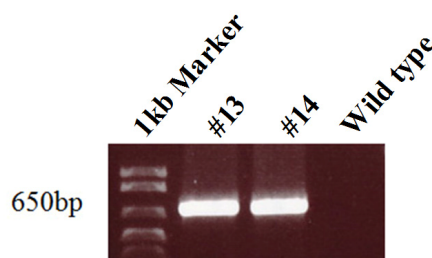
Figure 4.4. The strategy for generating *Zfp819* null allele. (A)The Neomycin cassette which is flanked by lox p sites was PCR amplified from vector pL452 using primers containing 50bp of each 5′ and 3′ micro arms (yellow and cyan rectangles) which are used for recombination. After homologous recombination, the targeting exons (exons 2, 3 and 4) are genetically replaced by neomycin resistance cassette (triangles). **(B)** The genotyping PCR result confirms the recombination of the Neo cassette and the deletion of exons 2, 3 and 4 in clones Nr.13 and Nr. 14. The locations of genotyping primers are indicated by arrows.

In order to retrieve a part of the modified BAC (neo cassette flanked by ~4-5 kb of homologous arms) into the targeting vector pL253, we PCR amplified ~400 bp of each 5′ and 3′ mini arms and cloned them into *Not*I/*Hin*dIII and *Hin*dIII/*Bam*HI restriction sites of pL253, respectively (Fig.4.5). Next, the *Hin*dIII linearized targeting vector will be used for retrieving the modified BAC into targeting vector. Upon successful retrieval, the final

targeting vector will be linearized with *Not*I and electroporated into ESCs. The successful homologous recombination into the *Zfp819* locus will be confirmed by Southern blot analysis as illustrated (Fig.4.6). The ESCs with correct homologous recombination will then be used to generate chimera and subsequent *Zfp819* KO mouse line.
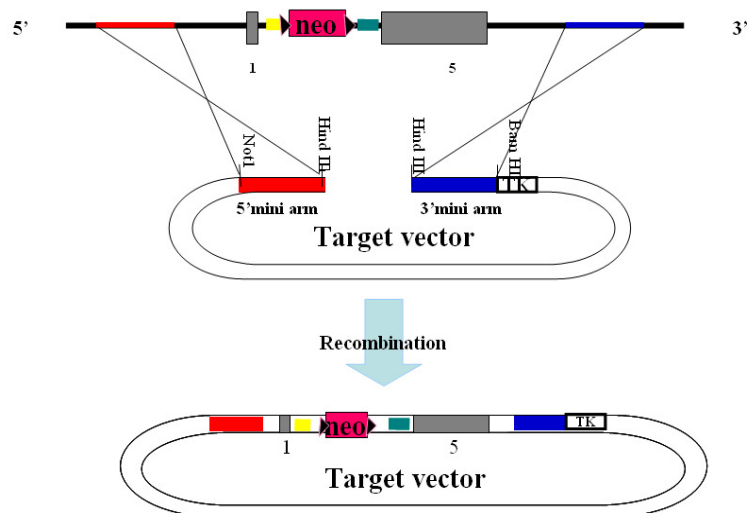

Figure 4.5. Generation of *Zfp819* knockout construct. The 5' and 3' homology mini arms are amplified by PCR and sub-cloned into TK vector (PL253) using restriction enzymes (5': *Not*I and *Hin*dIII, 3': *Hin*dIII and *Bam*HI). Subsequently the Neomycin cassette is inserted into TK vector by recombination.
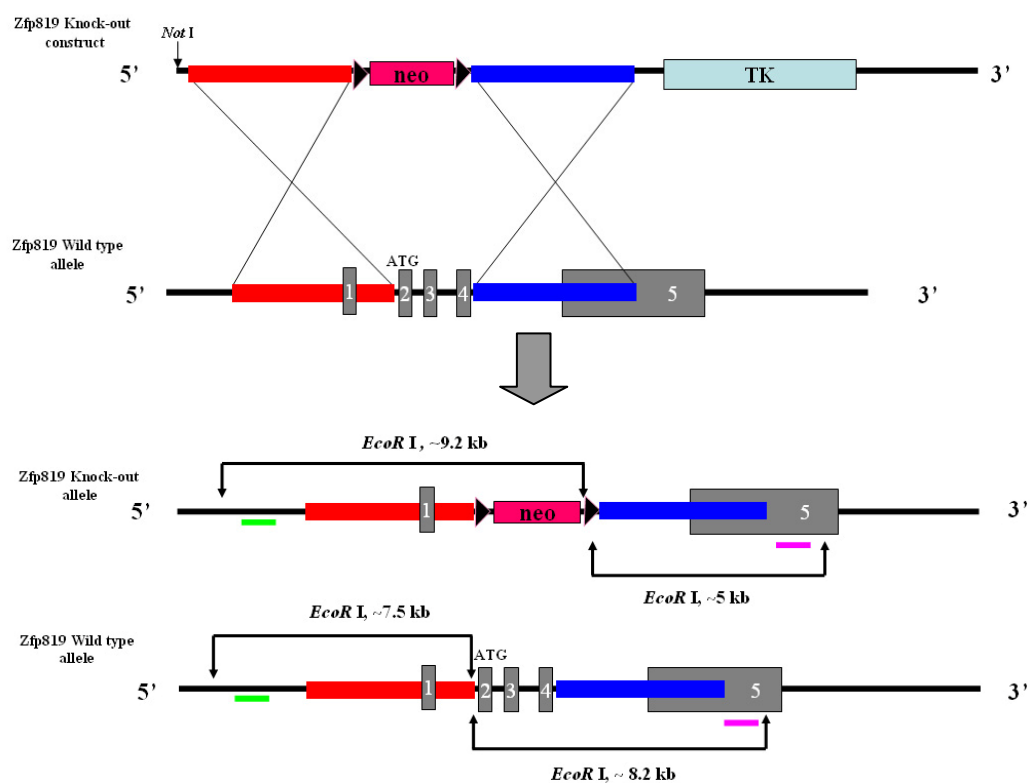
Figure 4.6. The scheme for the verification of homologous recombination in *Zfp819* locus. The targeting vector linearized by *Not*I will be electroporated in ESCs to generate the knockout allele. The Neo resistant ESC clones will be analyzed by Southern blot analysis for homologous recombination event of targeting construct into the *Zfp819* locus. The Southern blot analysis of recombinant clones with 5' external probe (green) should detect ~9.2 kb (KO) and ~7.5 kb (WT) bands upon *Eco*RI digestion. Similarly, the Southern blot analysis with 3' external probe (pink) should detect ~5 kb (KO) and ~8.2 kb (WT) bands upon *Eco*RI digestion.

4.5 Future plans and perspectives

The identification of *Ctc1* and *H3f3b* as putative interaction proteins of *Zfp819* is of particular interest, because both *Ctc1* and *H3f3b* are known to function in the maintenance of telomeres. Further studies aimed at characterization of *Zfp819* interaction with *Ctc1* and *H3f3b* and subsequent studies on the role of *Zfp819* in telomere maintenance will uncover its novel function in pluripotent stem cells.

In the present study, we found that *Zfp819* binds to *IAP* elements and suppresses their expression probably in a *Kap1*-dependent manner. Hence, it is interesting to identify the global *Zfp819* binding sites/targets using ChIP-Sequencing (ChIP-Seq) experiments. These studies will reveal the global function of *Zfp819* in pluripotent stem cells.

To further investigate the developmental role of *Zfp819*, in the present study we started with the generation of gain- and loss-of-function mouse models. The characterization of these mouse models will thus lead to the identification of *Zfp819* functions during embryonic development as well as in adulthood.

5 References

Abraham, R.T. (2001). Cell cycle checkpoint signaling through the ATM and ATR kinases. Genes Dev 15, 2177-2196.

Ankrum, J., and Karp, J.M. (2010). Mesenchymal stem cell therapy: Two steps forward, one step back. Trends Mol Med 16, 203-209.

Aronheim, A., Engelberg, D., Li, N., al-Alawi, N., Schlessinger, J., and Karin, M. (1994). Membrane targeting of the nucleotide exchange factor Sos is sufficient for activating the Ras signaling pathway. Cell 78, 949-961.

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.

Blanpain, C., and Fuchs, E. (2009). Epidermal homeostasis: a balancing act of stem cells in the skin. Nat Rev Mol Cell Biol 10, 207-217.

Borg, S., Podenphant, L., Jensen, T.J., and Poulsen, C. (1999). Plant cell growth and differentiation may involve GAP regulation of Rac activity. FEBS Lett 453, 341-345.

Bourillot, P.Y., and Savatier, P. (2010). Kruppel-like transcription factors and control of pluripotency. BMC Biol 8, 125.

Broder, Y.C., Katz, S., and Aronheim, A. (1998). The ras recruitment system, a novel approach to the study of protein-protein interactions. Curr Biol 8, 1121-1124.

Brons, I.G., Smithers, L.E., Trotter, M.W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., and Vallier, L. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature 448, 191-195.

Bruckner, A., Polge, C., Lentze, N., Auerbach, D., and Schlattner, U. (2009). Yeast two-hybrid, a powerful tool for systems biology. Int J Mol Sci 10, 2763-2788.

Brulet, P., Condamine, H., and Jacob, F. (1985). Spatial distribution of transcripts of the long repeated ETn sequence during early mouse embryogenesis. Proc Natl Acad Sci U S A 82, 2054-2058.

Brush, D., Dodgson, J.B., Choi, O.R., Stevens, P.W., and Engel, J.D. (1985). Replacement variant histone genes contain intervening sequences. Mol Cell Biol 5, 1307-1317.

Burklen, T.S., Hirschy, A., and Wallimann, T. (2007). Brain-type creatine kinase BB-CK interacts with the Golgi Matrix Protein GM130 in early prophase. Mol Cell Biochem 297, 53-64.

Cammas, F., Herzog, M., Lerouge, T., Chambon, P., and Losson, R. (2004). Association of the transcriptional corepressor TIF1beta with heterochromatin protein 1 (HP1): an essential role for progression through differentiation. Genes Dev 18, 2147-2160.

Cammas, F., Oulad-Abdelghani, M., Vonesch, J.L., Huss-Garcia, Y., Chambon, P., and Losson, R. (2002). Cell differentiation induces TIF1beta association with centromeric heterochromatin via an HP1 interaction. J Cell Sci 115, 3439-3448.

Cantz, T., Manns, M.P., and Ott, M. (2008). Stem cells in liver regeneration and therapy. Cell Tissue Res 331, 271-282.

Carlone, D.L., Lee, J.H., Young, S.R., Dobrota, E., Butler, J.S., Ruiz, J., and Skalnik, D.G. (2005). Reduced genomic cytosine methylation and defective cellular differentiation in embryonic stem cells lacking CpG binding protein. Mol Cell Biol 25, 4881-4891.

Cheng, Y., Geng, H., Cheng, S.H., Liang, P., Bai, Y., Li, J., Srivastava, G., Ng, M.H., Fukagawa, T., Wu, X., et al. (2010). KRAB zinc finger protein ZNF382 is a proapoptotic tumor suppressor that represses multiple oncogenes and is commonly silenced in multiple carcinomas. Cancer Res 70, 6516-6526.

Chuykin, I.A., Lianguzova, M.S., Pospelova, T.V., and Pospelov, V.A. (2008). Activation of DNA damage response signaling in mouse embryonic stem cells. Cell Cycle 7, 2922-2928.

Costoya, J.A., Hobbs, R.M., Barna, M., Cattoretti, G., Manova, K., Sukhwani, M., Orwig, K.E., Wolgemuth, D.J., and Pandolfi, P.P. (2004). Essential role of Plzf in maintenance of spermatogonial stem cells. Nat Genet 36, 653-659.

Deb-Rinker P., Ly D., Jezierski A., Sikorska M., Walker P. R. (2005)Sequential DNA methylation of the Nanog and Oct-4 upstream regions in human NT2 cells during neuronal differentiation. J. Biol. Chem 280:6257–6260

De Luca, P., Vazquez, E.S., Moiola, C.P., Zalazar, F., Cotignola, J., Gueron, G., Gardner, K., and De Siervi, A. (2011). BRCA1 loss induces GADD153-mediated doxorubicin resistance in prostate cancer. Mol Cancer Res 9, 1078-1090.

Dejosez, M., and Zwaka, T.P. (2012). Pluripotency and nuclear reprogramming. Annu Rev Biochem 81, 737-765.

Doolittle, R.F., Woodbury, N.W., and Jue, R.A. (1982). Ribosomal protein S1 is the product of a series of contiguous duplications. Biosci Rep 2, 405-412.

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D.J., Dash, C., Bazett-Jones, D.P., Le Grice, S., McKay, R.D., Buetow, K.H., et al. (2008). Global transcription in pluripotent embryonic stem cells. Cell Stem Cell 2, 437-447.

Elledge, S.J. (1996). Cell cycle checkpoints: preventing an identity crisis. Science 274, 1664-1672.

Emerson, R.O., and Thomas, J.H. (2009). Adaptive evolution in zinc finger transcription factors. PLoS Genet 5, e1000325.

Essers, J., Vermeulen, W., and Houtsmuller, A.B. (2006). DNA damage repair: anytime, anywhere? Curr Opin Cell Biol 18, 240-246.

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

Evsikov, A.V., de Vries, W.N., Peaston, A.E., Radford, E.E., Fancher, K.S., Chen, F.H., Blake, J.A., Bult, C.J., Latham, K.E., Solter, D., and Knowles, B.B. (2004). Systems biology of the 2-cell mouse embryo. Cytogenet Genome Res 105, 240-250.

Ewing, R.M., Chu, P., Elisma, F., Li, H., Taylor, P., Climie, S., McBroom-Cerajewski, L., Robinson, M.D., O'Connor, L., Li, M., et al. (2007). Large-scale mapping of human protein-protein interactions by mass spectrometry. Mol Syst Biol 3, 89.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature 340, 245-246.

Finch, B.W., and Ephrussi, B. (1967). Retention of multiple developmental potentialities by cells of a mouse testicular teratocarcinoma during prolonged culture *in vitro* and their extinction upon hybridization with cells of permanent lines. Proc Natl Acad Sci U S A 57, 615-621.

Fischedick, G., Klein, D.C., Wu, G., Esch, D., Hoing, S., Han, D.W., Reinhardt, P., Hergarten, K., Tapia, N., Scholer, H.R., and Sterneckert, J.L. (2012). Zfp296 is a novel, pluripotent-specific reprogramming factor. PLoS One 7, e34645.

Fluckiger, A.C., Marcy, G., Marchand, M., Negre, D., Cosset, F.L., Mitalipov, S., Wolf, D., Savatier, P., and Dehay, C. (2006). Cell cycle features of primate embryonic stem cells. Stem Cells 24, 547-556.

Friedman, J.R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G., and Rauscher, F.J., 3rd (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev 10, 2067-2078.

Fujiwara, T., Dunn, N.R., and Hogan, B.L. (2001). Bone morphogenetic protein 4 in the extraembryonic mesoderm is required for allantois development and the localization and survival of primordial germ cells in the mouse. Proc Natl Acad Sci U S A 98, 13739-13744.

Galan-Caridad, J.M., Harel, S., Arenzana, T.L., Hou, Z.E., Doetsch, F.K., Mirny, L.A., and Reizis, B. (2007). Zfx controls the self-renewal of embryonic and hematopoietic stem cells. Cell 129, 345-357.

Garcia-Garcia, M.J., Shibata, M., and Anderson, K.V. (2008). Chato, a KRAB zinc-finger protein, regulates convergent extension in the mouse embryo. Development 135, 3053-3062.

Geacintov, N.E., and Broyde, S. (2010).The Chemical Biology of DNA Damage.

Groner, A.C., Meylan, S., Ciuffi, A., Zangger, N., Ambrosini, G., Denervaud, N., Bucher, P., and Trono, D. (2010). KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. PLoS Genet 6, e1000869.

Gu, P., LeMenuet, D., Chung, A.C., Mancini, M., Wheeler, D.A., and Cooney, A.J. (2005). Orphan nuclear receptor GCNF is required for the repression of pluripotency genes during retinoic acid-induced embryonic stem cell differentiation. Mol Cell Biol 25, 8507-8519.

Guan, K., Nayernia, K., Maier, L.S., Wagner, S., Dressel, R., Lee, J.H., Nolte, J., Wolf, F., Li, M., Engel, W., and Hasenfuss, G. (2006). Pluripotency of spermatogonial stem cells from adult mouse testis. Nature 440, 1199-1203.

Harris, L.C. (2005). MDM2 splice variants and their therapeutic implications. Curr Cancer Drug Targets 5, 21-26.

Hirst, M., Ho, C., Sabourin, L., Rudnicki, M., Penn, L., and Sadowski, I. (2001). A two-hybrid system for transactivator bait proteins. Proc Natl Acad Sci U S A 98, 8726-8731.

Hong, Y., Cervantes, R.B., and Stambrook, P.J. (2006). DNA damage response and mutagenesis in mouse embryonic stem cells. Methods Mol Biol 329, 313-326.

Hong, Y., and Stambrook, P.J. (2004). Restoration of an absent G1 arrest and protection from apoptosis in embryonic stem cells after ionizing radiation. Proc Natl Acad Sci U S A 101, 14443-14448.

Howard, G., Eiges, R., Gaudet, F., Jaenisch, R., and Eden, A. (2008). Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. Oncogene 27, 404-408.

Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S.H., and Elledge, S.J. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes Dev 23, 837-848.

Huang, C., Jia, Y., Yang, S., Chen, B., Sun, H., Shen, F., and Wang, Y. (2007). Characterization of ZNF23, a KRAB-containing protein that is downregulated in human cancers and inhibits cell cycle progression. Exp Cell Res 313, 254-263.

Hubsman, M., Yudkovsky, G., and Aronheim, A. (2001). A novel approach for the identification of protein-protein interaction with integral membrane proteins. Nucleic Acids Res 29, E18.

Iyengar, S., and Farnham, P.J. (2011). KAP1 protein: an enigmatic master regulator of the genome. J Biol Chem 286, 26267-26276.

Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature 416, 556-560.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science 293, 1074-1080.

Johnsson, N., and Varshavsky, A. (1994). Split ubiquitin as a sensor of protein interactions *in vivo*. Proc Natl Acad Sci U S A 91, 10340-10344.

Kahan, B.W., and Ephrussi, B. (1970). Developmental potentialities of clonal *in vitro* cultures of mouse testicular teratoma. J Natl Cancer Inst 44, 1015-1036.

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

Kalter, S.S., Herberling, R.L., and Ratner, J.J. (1973). EBV antibody in monkeys and apes. Bibl Haematol 39, 871-875.

Kanatsu-Shinohara, M., Inoue, K., Lee, J., Yoshimoto, M., Ogonuki, N., Miki, H., Baba, S., Kato, T., Kazuki, Y., Toyokuni, S., et al. (2004). Generation of pluripotent stem cells from neonatal mouse testis. Cell 119, 1001-1012.

Kashyap, V., Rezende, N.C., Scotland, K.B., Shaffer, S.M., Persson, J.L., Gudas, L.J., and Mongan, N.P. (2009). Regulation of stem cell pluripotency and differentiation involves a mutual regulatory circuit of the NANOG, OCT4, and SOX2 pluripotency transcription factors with polycomb repressive complexes and stem cell microRNAs. Stem Cells Dev 18, 1093-1108.

Kigami, D., Minami, N., Takayama, H., and Imai, H. (2003). MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. Biol Reprod 68, 651-654.

Klug, A. (1999). Zinc finger peptides for the regulation of gene expression. J Mol Biol 293, 215-218.

Krebs, C.J., Schultz, D.C., and Robins, D.M. (2012). The KRAB Zinc Finger Protein RSL1 Regulates Sex- and Tissue-Specific Promoter Methylation and Dynamic Hormone-Responsive Chromatin Configuration. Mol Cell Biol 32, 3732-3742.

Kuo, W.L., Meyn, R.E., and Haidle, C.W. (1984). Neocarzinostatin-mediated DNA damage and repair in wild-type and repair-deficient Chinese hamster ovary cells. Cancer Res 44, 1748-1751.

Lagarkova, M.A., Shutova, M.V., Bogomazova, A.N., Vassina, E.M., Glazov, E.A., Zhang, P., Rizvanov, A.A., Chestkov, I.V., and Kiselev, S.L. (2010). Induction of pluripotency in human endothelial cells resets epigenetic profile on genome scale. Cell Cycle 9, 937-946.

Lage, K., Karlberg, E.O., Storling, Z.M., Olason, P.I., Pedersen, A.G., Rigina, O., Hinsby, A.M., Tumer, Z., Pociot, F., Tommerup, N., et al. (2007). A human phenome-interactome network of protein complexes implicated in genetic disorders. Nat Biotechnol 25, 309-316.

Lai, A.Y., and Wade, P.A. (2011). Cancer biology and NuRD: a multifaceted chromatin remodelling complex. Nat Rev Cancer 11, 588-596.

Lakso, M., Pichel, J.G., Gorman, J.R., Sauer, B., Okamoto, Y., Lee, E., Alt, F.W., and Westphal, H. (1996). Efficient *in vivo* manipulation of mouse genomic sequences at the zygote stage. Proc Natl Acad Sci U S A 93, 5860-5865.

Lamprecht, B., Bonifer, C., and Mathas, S. (2010). Repeat-element driven activation of proto-oncogenes in human malignancies. Cell Cycle 9, 4276-4281.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature 409, 860-921.

Larsen, D.H., Poinsignon, C., Gudjonsson, T., Dinant, C., Payne, M.R., Hari, F.J., Rendtlew Danielsen, J.M., Menard, P., Sand, J.C., Stucki, M., et al. (2010). The chromatin-remodeling factor CHD4 coordinates signaling and repair after DNA damage. J Cell Biol 190, 731-740.

Lawson, K. A., Dunn, N. R., Roelen, B. A., Zeinstra, L. M., Davis, A. M., Wright, C. V., Korving, J. P. and Hogan, B. L. (1999). Bmp4 is required for the generation of primordial germ cells in the mouse embryo. Genes Dev. 13, 424-436

Leung, D.C., and Lorincz, M.C. (2011). Silencing of endogenous retroviruses: when and why do histone marks predominate? Trends Biochem Sci 37, 127-133.

Levinson, D. J., Darrow, C. N., Klein, E. B., Levinson, M. A., and McKee, B. (1978). The Seasons of a Man's Life.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69, 915-926.

Li, L., Sun, L., Gao, F., Jiang, J., Yang, Y., Li, C., Gu, J., Wei, Z., Yang, A., Lu, R., et al. (2010). Stk40 links the pluripotency factor Oct4 to the Erk/MAPK pathway and controls extraembryonic endoderm differentiation. Proc Natl Acad Sci U S A 107, 1402-1407.

Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P., and Ferguson-Smith, A.C. (2008). A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. Dev Cell 15, 547-557.

Lim, L.S., Loh, Y.H., Zhang, W., Li, Y., Chen, X., Wang, Y., Bakre, M., Ng, H.H., and Stanton, L.W. (2007). Zic3 is required for maintenance of pluripotency in embryonic stem cells. Mol Biol Cell 18, 1348-1358.

Lim, U., Subar, A.F., Mouw, T., Hartge, P., Morton, L.M., Stolzenberg-Solomon, R., Campbell, D., Hollenbeck, A.R., and Schatzkin, A. (2006). Consumption of

aspartame-containing beverages and incidence of hematopoietic and brain malignancies. Cancer Epidemiol Biomarkers Prev 15, 1654-1659.

Lin, T., Chao, C., Saito, S., Mazur, S.J., Murphy, M.E., Appella, E., and Xu, Y. (2005). p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. Nat Cell Biol 7, 165-171.

Liu, P., Jenkins, N.A., and Copeland, N.G. (2003). A highly efficient recombineering-based method for generating conditional knockout mutations. Genome Res 13, 476-484.

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

Looman, C., Abrink, M., Mark, C., and Hellman, L. (2002). KRAB zinc finger proteins: an analysis of the molecular mechanisms governing their increase in numbers and complexity during evolution. Mol Biol Evol 19, 2118-2130.

Lower, J. (1996). Risk assessment and limitations of evaluation studies. Dev Biol Stand 88, 109.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature 487, 57-63.

Maksakova, I.A., Goyal, P., Bullwinkel, J., Brown, J.P., Bilenky, M., Mager, D.L., Singh, P.B., and Lorincz, M.C. (2011). H3K9me3-binding proteins are dispensable for SETDB1/H3K9me3-dependent retroviral silencing. Epigenetics Chromatin 4, 12.

Maksakova, I.A., Romanish, M.T., Gagnier, L., Dunn, C.A., van de Lagemaat, L.N., and Mager, D.L. (2006). Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. PLoS Genet 2, e2.

Margolin, J.F., Friedman, J.R., Meyer, W.K., Vissing, H., Thiesen, H.J., and Rauscher, F.J., 3rd (1994). Kruppel-associated boxes are potent transcriptional repression domains. Proc Natl Acad Sci U S A 91, 4509-4513.

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.

Masui, S., Ohtsuka, S., Yagi, R., Takahashi, K., Ko, M.S., and Niwa, H. (2008). Rex1/Zfp42 is dispensable for pluripotency in mouse ES cells. BMC Dev Biol 8, 45.

Matsui, T., Leung, D., Miyashita, H., Maksakova, I.A., Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M.C., and Shinkai, Y. (2010). Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature 464, 927-931.

Matsui, Y., Zsebo, K., and Hogan, B.L. (1992). Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. Cell 70, 841-847.

Medstrand, P., Lindeskog, M., and Blomberg, J. (1992). Expression of human endogenous retroviral sequences in peripheral blood mononuclear cells of healthy individuals. J Gen Virol 73 (Pt 9), 2463-2466.

Meshorer, E., and Misteli, T. (2006). Chromatin in pluripotent embryonic stem cells and differentiation. Nat Rev Mol Cell Biol 7, 540-546.

Meyer, S., Nolte, J., Opitz, L., Salinas-Riester, G., and Engel, W. (2010). Pluripotent embryonic stem cells and multipotent adult germline stem cells reveal similar transcriptomes including pluripotency-related genes. Mol Hum Reprod 16 (11):846-55

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

Mockli, N., Deplazes, A., Hassa, P.O., Zhang, Z., Peter, M., Hottiger, M.O., Stagljar, I., and Auerbach, D. (2007). Yeast split-ubiquitin-based cytosolic screening system to detect interactions between transcriptionally active proteins. Biotechniques 42, 725-730.

Momcilovic, O., Knobloch, L., Fornsaglio, J., Varum, S., Easley, C., and Schatten, G. (2010). DNA damage responses in human induced pluripotent stem cells and embryonic stem cells. PLoS One 5, e13410.

Mortusewicz, O., Roth, W., Li, N., Cardoso, M.C., Meisterernst, M., and Leonhardt, H. (2008). Recruitment of RNA polymerase II cofactor PC4 to DNA damage sites. J Cell Biol 183, 769-776.

Narlikar, G.J., Fan, H.Y., and Kingston, R.E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. Cell 108, 475-487.

Nielsen, A.L., Ortiz, J.A., You, J., Oulad-Abdelghani, M., Khechumian, R., Gansmuller, A., Chambon, P., and Losson, R. (1999). Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. EMBO J 18, 6385-6395.

Niethammer, M., Kim, E., and Sheng, M. (1996). Interaction between the C terminus of NMDA receptor subunits and multiple members of the PSD-95 family of membrane-associated guanylate kinases. J Neurosci 16, 2157-2163.

Nishii, T., Oikawa, Y., Ishida, Y., Kawaichi, M., and Matsuda, E. (2012). CtBP-interacting BTB zinc finger protein (CIBZ) promotes proliferation and G1/S transition in embryonic stem cells via Nanog. J Biol Chem 287, 12417-12424.

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

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99, 247-257.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1, 376-386.

Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. Cell 132, 631-644.

Oti, M., Snel, B., Huynen, M.A., and Brunner, H.G. (2006). Predicting disease genes using protein-protein interactions. J Med Genet 43, 691-698.

Pan, M.R., Hsieh, H.J., Dai, H., Hung, W.C., Li, K., Peng, G., and Lin, S.Y. (2012). Chromodomain helicase DNA-binding protein 4 (CHD4) regulates homologous recombination DNA repair, and its deficiency sensitizes cells to poly(ADP-ribose) polymerase (PARP) inhibitor treatment. J Biol Chem 287, 6764-6772.

Pannell, D., Osborne, C.S., Yao, S., Sukonnik, T., Pasceri, P., Karaiskakis, A., Okano, M., Li, E., Lipshitz, H.D., and Ellis, J. (2000). Retrovirus vector silencing is de novo methylase independent and marked by a repressive histone code. EMBO J 19, 5884-5894.

Patience, C., Simpson, G.R., Colletta, A.A., Welch, H.M., Weiss, R.A., and Boyd, M.T. (1996). Human endogenous retrovirus expression and reverse transcriptase activity in the T47D mammary carcinoma cell line. J Virol 70, 2654-2657.

Paull, T.T., Rogakou, E.P., Yamazaki, V., Kirchgessner, C.U., Gellert, M., and Bonner, W.M. (2000). A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. Curr Biol 10, 886-895.

Peaston, A.E., Evsikov, A.V., Graber, J.H., de Vries, W.N., Holbrook, A.E., Solter, D., and Knowles, B.B. (2004). Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. Development Cell 7, 597-606.

Peng, H., Begg, G.E., Schultz, D.C., Friedman, J.R., Jensen, D.E., Speicher, D.W., and Rauscher, F.J., 3rd (2000). Reconstitution of the KRAB-KAP-1 repressor complex: a model system for defining the molecular anatomy of RING-B box-coiled-coil domain-mediated protein-protein interactions. J Mol Biol 295, 1139-1162.

Perrotti, D., Melotti, P., Skorski, T., Casella, I., Peschle, C., and Calabretta, B. (1995). Overexpression of the zinc finger protein MZF1 inhibits hematopoietic development from embryonic stem cells: correlation with negative regulation of CD34 and c-myb promoter activity. Mol Cell Biol 15, 6075-6087.

Petrascheck, M., Castagna, F., and Barberis, A. (2001). Two-hybrid selection assay to identify proteins interacting with polymerase II transcription factors and regulators. Biotechniques 30, 296-298, 300, 302.

Polo, S.E., Kaidi, A., Baskcomb, L., Galanty, Y., and Jackson, S.P. (2010). Regulation of DNA-damage responses and cell-cycle progression by the chromatin remodelling factor CHD4. EMBO J 29, 3130-3139.

Poznanski, A.A., and Calarco, P.G. (1991). The expression of intracisternal A particle genes in the preimplantation mouse embryo. Dev Biol 143, 271-281.

Prindull, G., Prindull, B., and Meulen, N. (1978). Haematopoietic stem cells (CFUc) in human cord blood. Acta Paediatr Scand 67, 413-416.

Reynolds, B.A., and Weiss, S. (1992). Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. Science 255, 1707-1710.

Rogakou, E.P., Pilch, D.R., Orr, A.H., Ivanova, V.S., and Bonner, W.M. (1998). DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. J Biol Chem 273, 5858-5868.

Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P.V., Layard-Liesching, H., Verp, S., Marquis, J., et al. (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. Nature 463, 237-240.

Rowe, H.M., and Trono, D. (2011). Dynamic control of endogenous retroviruses during development. Virology 411, 273-287.

Ryan, R.F., Schultz, D.C., Ayyanathan, K., Singh, P.B., Friedman, J.R., Fredericks, W.J., and Rauscher, F.J., 3rd (1999). KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. Mol Cell Biol 19, 4366-4378.

Sato, N., Kondo, M., and Arai, K. (2006). The orphan nuclear receptor GCNF recruits DNA methyltransferase for Oct-3/4 silencing. Biochem Biophys Res Commun 344, 845-851.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev 16, 919-932.

Schultz, D.C., Friedman, J.R., and Rauscher, F.J., 3rd (2001). Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. Genes Dev 15, 428-443.

Scotland, K.B., Chen, S.M., Sylvester, R., and Gudas, L.J. (2009). Analysis of Rex1 (Zfp42) Function in Embryonic Stem Cell Differentiation. Developmental Dynamics 238, 1863-1877.

Serebriiskii, I.G., Mitina, O., Pugacheva, E.N., Benevolenskaya, E., Kotova, E., Toby, G.G., Khazak, V., Kaelin, W.G., Chernoff, J., and Golemis, E.A. (2002). Detection of peptides, proteins, and drugs that selectively interact with protein targets. Genome Res 12, 1785-1791.

Shi, S., Bartold, P.M., Miura, M., Seo, B.M., Robey, P.G., and Gronthos, S. (2005). The efficacy of mesenchymal stem cells to regenerate and repair dental structures. Orthod Craniofac Res 8, 191-199.

Shibata, M., Blauvelt, K.E., Liem, K.F., Jr., and Garcia-Garcia, M.J. (2011). TRIM28 is required by the mouse KRAB domain protein ZFP568 to control convergent extension and morphogenesis of extra-embryonic tissues. Development 138, 5333-5343.

Shibata, M., and Garcia-Garcia, M.J. (2011). The mouse KRAB zinc-finger protein CHATO is required in embryonic-derived tissues to control yolk sac and placenta morphogenesis. Dev Biol 349, 331-341.

Short, S.C., Bourne, S., Martindale, C., Woodcock, M., and Jackson, S.P. (2005). DNA damage responses at low radiation doses. Radiat Res 164, 292-302.

Smeenk, L., and Lohrum, M. (2010). Behind the scenes: unravelling the molecular mechanisms of p53 target gene selectivity (Review). Int J Oncol 37, 1061-1070.

Sripathy, S.P., Stevens, J., and Schultz, D.C. (2006). The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. Mol Cell Biol 26, 8623-8638.

Stagljar, I., Korostensky, C., Johnsson, N., and te Heesen, S. (1998). A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins *in vivo*. Proc Natl Acad Sci U S A 95, 5187-5192.

Stambrook, P.J. (2007). An ageing question: do embryonic stem cells protect their genomes? Mech Ageing Dev 128, 31-35.

Sugita, M., Jackman, R.M., van Donselaar, E., Behar, S.M., Rogers, R.A., Peters, P.J., Brenner, M.B., and Porcelli, S.A. (1996). Cytoplasmic tail-dependent localization of CD1b antigen-presenting molecules to MIICs. Science 273, 349-352.

Surovtseva, Y.V., Churikov, D., Boltz, K.A., Song, X., Lamb, J.C., Warrington, R., Leehy, K., Heacock, M., Price, C.M., and Shippen, D.E. (2009). Conserved telomere maintenance component 1 interacts with STN1 and maintains chromosome ends in higher eukaryotes. Mol Cell 36, 207-218.

Suwinska, A., Czolowska, R., Ozdzenski, W., and Tarkowski, A.K. (2008). Blastomeres of the mouse embryo lose totipotency after the fifth cleavage division:

expression of Cdx2 and Oct4 and developmental potential of inner and outer blastomeres of 16- and 32-cell embryos. Dev Biol 322, 133-144.

Sylvester, K.G., and Longaker, M.T. (2004). Stem cells: review and update. Arch Surg 139, 93-99.

Tan, X. Y., Xu, X.B., Zechner, U., Nolte, J., Engel, W., Pantakani, D.V. (2012). Zfp819, a novl KRAB-zinc finger protein, interacts with KAP1 and function in genomic integrity maintenance of mouse embryonic stem cell. (Submitted)

Tafelmeyer, P., Johnsson, N., and Johnsson, K. (2004). Transforming a (beta/alpha)8--barrel enzyme into a split-protein sensor through directed evolution. Chem Biol 11, 681-689.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131, 861-872.

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, H., Katoh, H., and Negishi, M. (2006). Pragmin, a novel effector of Rnd2 GTPase, stimulates RhoA activity. J Biol Chem 281, 10355-10364.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature 448, 196-199.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science 282, 1145-1147.

Tian, C., Xing, G., Xie, P., Lu, K., Nie, J., Wang, J., Li, L., Gao, M., Zhang, L., and He, F. (2009). KRAB-type zinc-finger protein Apak specifically regulates p53-dependent apoptosis. Nat Cell Biol 11, 580-591.

Tichy, E.D. (2011). Mechanisms maintaining genomic integrity in embryonic stem cells and induced pluripotent stem cells. Exp Biol Med (Maywood) 236, 987-996.

Torres-Ramos, C.A., Prakash, S., and Prakash, L. (2002). Requirement of RAD5 and MMS2 for postreplication repair of UV-damaged DNA in Saccharomyces cerevisiae. Mol Cell Biol 22, 2419-2426.

Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., and Niwa, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. Development 135, 909-918.

Urnovitz, H.B., and Murphy, W.H. (1996). Human endogenous retroviruses: nature, occurrence, and clinical implications in human disease. Clin Microbiol Rev 9, 72-99.

Urquhart, A.J., Gatei, M., Richard, D.J., and Khanna, K.K. (2011). ATM mediated phosphorylation of CHD4 contributes to genome maintenance. Genome Integr 2, 1.

Urrutia, R. (2003). KRAB-containing zinc-finger repressor proteins. Genome Biol 4, 231.

Vanunu, O., Magger, O., Ruppin, E., Shlomi, T., and Sharan, R. (2010). Associating genes and protein complexes with disease via network propagation. PLoS Comput Biol 6, e1000641.

Velkey, J.M., and O'Shea, K.S. (2003). Oct4 RNA interference induces trophectoderm differentiation in mouse embryonic stem cells. Genesis 37, 18-24.

Vooijs, M., Jonkers, J., and Berns, A. (2001). A highly efficient ligand-regulated Cre recombinase mouse line shows that LoxP recombination is position dependent. EMBO Rep 2, 292-297.

Walhout, A.J., and Vidal, M. (2001). High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. Methods 24, 297-306.

Walsh, C.P., and Bestor, T.H. (1999). Cytosine methylation and mammalian development. Genes Dev 13, 26-34.

Wang, Z.X., Kueh, J.L.L., Teh, C.H.L., Rossbach, M., Lim, L., Li, P., Wong, K.Y., Lufkin, T., Robson, P., and Stanton, L.W. (2007). Zfp206 is a transcription factor that controls pluripotency of embryonic stem cells. Stem Cells 25, 2173-2182.

Wang, Z.X., Teh, C.H., Chan, C.M., Chu, C., Rossbach, M., Kunarso, G., Allapitchay, T.B., Wong, K.Y., and Stanton, L.W. (2008). The transcription factor Zfp281 controls embryonic stem cell pluripotency by direct activation and repression of target genes. Stem Cells 26, 2791-2799.

Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., et al. (2002). Initial

sequencing and comparative analysis of the mouse genome. Nature 420, 520-562.

Wells, D., Hoffman, D., and Kedes, L. (1987). Unusual structure, evolutionary conservation of non-coding sequences and numerous pseudogenes characterize the human H3.3 histone multigene family. Nucleic Acids Res 15, 2871-2889.

White, D.E., Negorev, D., Peng, H., Ivanov, A.V., Maul, G.G., and Rauscher, F.J., 3rd (2006). KAP1, a novel substrate for PIKK family members, colocalizes with numerous damage response factors at DNA lesions. Cancer Res 66, 11594-11599.

Witzgall, R., O'Leary, E., Leaf, A., Onaldi, D., and Bonventre, J.V. (1994). The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. Proc Natl Acad Sci U S A 91, 4514-4518.

Wolf, D., and Goff, S.P. (2009). Embryonic stem cells use ZFP809 to silence retroviral DNAs. Nature 458, 1201-1204.

Wolfe, S.A., Nekludova, L., and Pabo, C.O. (2000). DNA recognition by Cys2His2 zinc finger proteins. Annu Rev Biophys Biomol Struct 29, 183-212.

Wu, R. S., Bonner, W. M.(1981). Separation of basal histone synthesis from S-phase histone synthesis in dividing cells. Cell 27, 321-330

Xu, C., Rosler, E., Jiang, J., Lebkowski, J.S., Gold, J.D., O'Sullivan, C., Delavan-Boorsma, K., Mok, M., Bronstein, A., and Carpenter, M.K. (2005). Basic fibroblast growth factor supports undifferentiated human embryonic stem cell growth without conditioned medium. Stem Cells 23, 315-323.

Yan, P., Wang, T., Newton, G.J., Knyushko, T.V., Xiong, Y., Bigelow, D.J., Squier, T.C., and Mayer, M.U. (2009). A targeted releasable affinity probe (TRAP) for *in vivo* photocrosslinking. Chembiochem 10, 1507-1518.

Yang, J., Chai, L., Gao, C., Fowles, T.C., Alipio, Z., Dang, H., Xu, D., Fink, L.M., Ward, D.C., and Ma, Y. (2008). SALL4 is a key regulator of survival and apoptosis in human leukemic cells. Blood 112, 805-813.

Yang, J., Yu, Y., Hamrick, H.E., and Duerksen-Hughes, P.J. (2003). ATM, ATR and DNA-PK: initiators of the cellular genotoxic stress responses. Carcinogenesis 24, 1571-1580.

Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell 115, 281-292.

Yoshida, T., Hazan, I., Zhang, J., Ng, S.Y., Naito, T., Snippert, H.J., Heller, E.J., Qi, X., Lawton, L.N., Williams, C.J., and Georgopoulos, K. (2008). The role of the chromatin remodeler Mi-2beta in hematopoietic stem cell self-renewal and multilineage differentiation. Genes Dev 22, 1174-1189.

Young, K.H. (1998). Yeast two-hybrid: so many interactions, (in) so little time. Biol Reprod 58, 302-311.

Yu, H.B., Kunarso, G., Hong, F.H., and Stanton, L.W. (2009). Zfp206, Oct4, and Sox2 are integrated components of a transcriptional regulatory network in embryonic stem cells. J Biol Chem 284, 31327-31335.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science 318, 1917-1920.

Zechner, U., Nolte, J., Wolf, M., Shirneshan, K., Hajj, N.E., Weise, D., Kaltwasser, B., Zovoilis, A., Haaf, T., and Engel, W. (2009). Comparative methylation profiles and telomerase biology of mouse multipotent adult germline stem cells and embryonic stem cells. Mol Hum Reprod 15, 345-353.

Zhang, H. and Bradley, A. (1996). Mice deficient for BMP2 are nonviable and have defects in amnion/chorion and cardiac development. Development 122, 2977-2986

Zhang, W., Walker, E., Tamplin, O.J., Rossant, J., Stanford, W.L., and Hughes, T.R. (2006b). Zfp206 regulates ES cell gene expression and differentiation. Nucleic Acids Research 34, 4780-4790.

Zheng, Y., Tan, X., Pyczek, J., Nolte, J., Pantakani, D.V., and Engel, W. (2012). Generation and characterization of Yeast Two-Hybrid cDNA Libraries Derived From Two Distinct Mouse Pluripotent Cell Types. Mol Biotechnol. DOI 10.1007/s12033-012-9561-4

Ziv, Y., Bielopolski, D., Galanty, Y., Lukas, C., Taya, Y., Schultz, D.C., Lukas, J., Bekker-Jensen, S., Bartek, J., and Shiloh, Y. (2006). Chromatin relaxation in

response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. Nat Cell Biol 8, 870-876.

Zuo, C., Liang, S., Wang, Z., Li, H., Zheng, W., and Ma, W. (2009). Enriching protein-protein and functional interaction networks in human embryonic stem cells. Int J Mol Med 23, 811-819.

Zuo, X., Sheng, J., Lau, H.T., McDonald, C.M., Andrade, M., Cullen, D.E., Bell, F.T., Iacovino, M., Kyba, M., Xu, G., and Li, X. (2011). Zinc finger protein ZFP57 requires its co-factor to recruit DNA methyltransferases and maintains DNA methylation imprint in embryonic stem cells via its transcriptional repression domain. J Biol Chem 287, 2107-2118.

6 Acknowledgments

7 Curriculum Vitae

Personal details:

Name: Xiaoying Tan

Date of birth: 19.02.1984

Place of birth: Sichuan (China)

Nationality: Chinese

Address: Albrecht-Thaer-Weg 22h, 02-03, Goettingen

E-Mail: mytanxy@gmail.com

Education:

Sep.1990-July 1996:   Primary school (Sichuan, China)

Sep.1996-July 1999:   Secondary school (Sichuan, China)

Sep.1999-July 2002:   High school (Sichuan, China)

Sep.2002-July 2006:   Bachelor study (Shaanxi, China)

Sep.2006-July 2009:   Master study with the title: "Study on Ovine homogeneity
                         and Goat-Ovine heterogeneity cloned embryos *in vitro*"
                         Center of Stem Cell Engineering & Technology,
                         Northwest A＆F University,
                         Yangling, China
                         Supervisor: Prof. Zhongying Dou

Since Sep.2009:   PhD study with the title: "Functional analysis of Zfp819 in
                         pluripotency and embryonic development "
                         Institute of Human Genetics, Georg-August-University,
                         Göttingen, Germany
                         Supervisor: Prof. Dr. med. Dr. h. c. Wolfgang Engel