Regulation Of Retroviral Silencing In Different Cell Types

Cheng Wang

Submitted in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2015

ABSTRACT

Regulation of Retroviral Silencing in Different Cell Types

Cheng Wang

The replication of Moloney Murine Leukemia Virus (MoMLV or MLV) is restricted in mouse embryonic stem (ES) and embryonic carcinoma (EC) cells, but not in differentiated cells. The restriction is mediated by the primer binding site (PBS) of proviral DNA of MLV. A restriction complex can bind to the PBS of MLV and block the transcription of viral genes. Two major components of the PBS-mediated silencing complex were identified in our lab, ZFP809 and Trim28. ZFP809 contains two conserved domains, a zinc finger domain responsible for DNA binding and a KRAB box recruiting Trim28, and hence other transcription repressors, such as HP1γ and ESET. A protein called L1td1 was identified during the purification of PBS-mediated restriction complex. L1td1 is a stem cell specific protein but little is known about the function of L1td1.

In differentiated cells, the replication of MLV is not restricted. Overexpression of ZFP809 in differentiated cells is sufficient to re-establish the PBS mediated restriction. However, data from various expression libraries shows that the mRNA levels of ZFP809 in stem cells and differentiated cells are approximately the same, which indicating that there is some post-transcriptional mechanism negatively regulating the protein levels of ZFP809 in differentiated cells. To study the post-translational regulation of ZFP809 may help us understand how retroviral restriction is regulated in different cell types. Here we

found that the down-regulation of ZFP809 proteins is due to the rapid degradation of protein but not on mRNA. The protein of ZFP809 is degraded rapidly in differentiated cells but not in stem cells. The last 50 amino acids, as well as the lysine residue within the peptide, are important for the turnover of ZFP809 protein in differentiated cells. The drug MG132 can stabilize the ZFP809 protein in differentiated and in vivo ubiquitination assay show that ZFP809 is heavily ubiquinated in differentiated cells, suggesting that ZFP809 is degraded through the ubiquitin-dependent proteasomal pathway. Interestingly, the protein Trim28, which is an essential factor for in the silencing complex, can promote the degradation of ZFP809. Mutations with the lysine residue mutated to alanine or abolished the interaction between Trim28 are less ubiquitinated. A small drug, MLN4924, which is the neddylation inhibitor, stabilizes ZFP809 in differentiated cells. Overall, these observations suggest that, during the differentiation of mouse stem cells, ZFP809 protein is eliminated by the proteasomal system, which leads to the loss of restriction of MLV in differentiated cells.

In addition, we studied the role of L1td1 in retroviral silencing. Knockdown or knockout of L1td1 partially relieves the restriction of MLV replication. Immunoprecipitation and pulldown assays show that L1td1 might interact with Trim28 and ZFP809 bridging by Trim28. In summary, L1td1 might interact with the essential factors of silencing complex and help the silencing of MLV in stem cells.

Proteins of the nucleosome remodeling deacetylase (NuRD) complex were also identified during purification of the restriction. The NuRD complex is shown to be involved in the transcriptional repression. However, depletion of single subunits of the NuRD complex does not affect the PBS-mediated retroviral restriction in mouse EC cells.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

Chapter Three

ACKNOWLEDGEMENTS

First, I would like to thank my advisor Steve Goff, who has provided me strong guidance and support for my research and enough freedom to figure out things by myself. I always remember his Santa-like smile whenever I went to his office with good or bad results. I have learned a lot from him not only about how to do research but also to be a good mentor and human being.

I also thank Robert Hogg who trained me as a rotation student in the lab and became a faculty in NIH to train more students. Bobby thinks smartly and precisely in designing and performing experiments. Special thanks to Sharon Schlesinger and Yiping Zhu for giving me help and suggestions for my research projects and Oya to read my dissertation.

When I first joined the lab, I was lucky to work with a group of smart and kind people who helped me set up in the lab – Jason Rodriguez, Robert Hogg, Sharon Schlesinger, Zhaozhu Qiu, Gary Wang, Sedef Onal, Gloria Arriagada and Lisa Green. During my research in the lab, some interesting people joined the lab and provided helpful insights – Yiping Zhu, Angela Erazo, Mojgan Naghavi, Yosef Sabo, Andreia Lee, Oya Cingoz, Daniel Griffin, Marina Ermakova and Michael Metzger. I also thank Kenia de los Santos, our lab manager to keep our lab running smoothly and Martha de los Santos and Helen Wang for technical support. It was so fortunate to work, discuss, play and laugh with those people in our lab.

I also owe many thanks to Lawrence Chasin and Vincent Racaniello to be my research committee and supervise my research progress. I also thank Songtao Jia and David Volsky to be my thesis committee and provide valuable insights for my dissertation.

Life is a balance of work, friends and family. I thank my wushu and calligraphy teachers to teach the beauty and philosophy of Chinese culture. As a foreign student studying in US, a lot of friends have become my new family members. They gave me lots of support and suggestions. Most importantly, they are always available when I was demoralized and needed them.

Last, but never least, I want to thank my family – my Dad, Mom and younger sister. I thank my parents who took great courage against the "one child" policy of China to bring me a younger sister. I cannot count the number of sacrifices they made to give me all of the opportunities in life. Few words cannot count how much I appreciate them.

.

Chapter 1

Introduction

The aim of the work described in this thesis is to increase our knowledge of the replication of a prototypical retroviruses, Moloney Murine Leukemia Virus (MLV) and its regulation in different types of cells.

1.1.Retroviruses

Retroviruses are a large family of enveloped, single-stranded and positive-sense animal viruses that contain a genome of diploid RNA. The hallmark of these viruses, which distinguishes retroviruses from other RNA viruses (e.g. polio viruses and influenza viruses) is the reverse transcription of single-stranded RNA to double-stranded DNA and integration of the viral DNA. Upon entry into the host cells, the single-stranded RNA genome is reverse transcribed into a double-stranded DNA copy, which will be transported into the nuclei and integrated into the chromosome of host cells. The integrated viral DNA, the provirus, thus becomes a permanent genetic element of newly infected cells and serves as template to make new viral RNAs and proteins to assemble new viral particles.

Many retrovirus-related pathogenesis have been documented. Some retroviruses that do not carry a transforming gene, are still able to induce leukemia or lymphoma after a long term of latency (Anson, 2004). For example, MLV causes T-cell leukemia in infected mice by insertion of viral DNA in the vicinity of a cellular oncogene, where the strong enhancer function of viral promoter drives the abnormal expression of the oncogene (Lazo et al., 1990).

Rapid tumor formation and mortality can be initiated by acute transforming retroviruses (Shackleford and Varmus, 1988). The genome of some retroviruses, like Rous

Sarcoma Virus, have acquired a transforming version of a cellular gene variant, v-src, which is delivered to newly infected cells to initiate tumor formation (Rosenberg and Jolicoeur, 1997). The human T-cell leukemia virus, HTLV, is also able to induce T-cell leukemia in infected people after a long period of latency. However, the tumor development of HTLV infected persons does not result from transduction or activation of oncogenes. The mechanism of transformation by HTLV-1 remains unknown (Matsuoka, 2003).

The most notable human retrovirus is human immunodeficiency virus type 1 (HIV-1), which is the agent responsible for the development of acquired immunodeficiency syndrome (AIDS). Infection by HIV-1 promotes eventual loss of CD-4 positive immune cells, then immunodeficiency, culminating in severe opportunistic infections (Stevenson, 2003). According to the report by the Joint United Nations program on HIV/AIDS 2013, there were 35.3 million people living with HIV and 2.3 million of deaths caused by AIDS in 2012.

Although most of the integration of retrovirus causes genetic mutations in the host genome, not all infections lead to disease *per se*. The host development and survival are not affected by infection with some retroviruses, such as spumaviruses (Meiering and Linial, 2001).

1.2.Moloney Murine Leukemia Virus

Moloney Murine Leukemia Virus was discovered as an agent that induces T-cell leukemia in mice in the 1960's (Moloney, 1960). Since MLV has many experimental

advantages, it has served as workhorse for studies of retroviral replication. Wild-type MLV is not cytopathic, allowing study of long term virus-host interactions. Since high titers of viral particles can be generated in many rodent cells, many stages of viral replication can be easily monitored. Monoclonal and polyclonal antibodies against viral proteins are available in academic labs or commercially, and there are many sensitive and rapid tests to probe for the presence of replication-competent viruses. The basic mechanisms are representative of many other known retroviruses, and as a result, many findings are relative to other pathogens like HIV-1.

1.3.MLV genome organization and virion structure

The retroviral genome is composed of two identical single-stranded RNAs. The genome of MLV is 8.2 kilobases. Its organization is depicted in figure 1 (Petropoulos, 1997). The 5' end of the genomic RNA starts with the repeat region (R) and the unique 5' region (U5). Adjacent to the U5 is the primer binding site (PBS), which serves as the binding site for a specific tRNA, which acts as a primer for the initiation of reverse transcription. Following the PBS is the packaging signal psi, ψ, which mediates encapsidation of the viral genome into the virus particle. The two open reading frames contain three genes, *gag*, *pol* and *env*. The viral structural protein-coding genes *gag* and *pol* lie in the same open reading frame and are separated by a stop codon. The *env* gene, which encodes the envelop glycoprotein, is positioned downstream of *gag-pol* and with some overlap with *pol* gene. Splice donor and acceptor sites (SD and SA respectively) are utilized for RNA splicing to generate *env* RNA. The polypurine tract (PPT) near the 3'

LTR is a primer for the start of plus-strand DNA synthesis during reverse transcription. The 3' region of viral RNA contains U3 and R. The viral RNAs contain poly(A) at the 3' end.

The RNA genome of retrovirus is converted by reverse transcription into the proviral DNA, which will be integrated into the host genome. The proviral DNA contains a complete LTR, consisting of U3, R and U5 regions, at both ends. The U3 region at the 5' end, which contains the promoter and enhancer responsible for the initiation of transcription of viral genome, is copied from the 3' end of retroviral RNA during the reverse transcription process. The transcribed RNA is capped at its 5' end with m7G5'ppp5'Gmp (7-methylguanosine). The viral genes *gag* and *pol* are expressed from the unspliced RNA. While *gag* protein is synthesized directly, the expression of *pol* protein is regulated by a pseudoknot RNA structure which recodes the stop codon and extends the translation (Felsenstein and Goff, 1988; Houck-Loomis et al., 2011). The *pol* gene is only expressed at a low level (2%-10%) compared to the gag gene, and is only expressed in the form of a Gag-Pol fusion protein. The *env* gene is expressed from a separate, spliced RNA.

After the assembly and release from the host cells, the viral particles are cleaved into the mature form by the viral protease. The maturation process includes the condensation of the proteolytic proteins in the viral core (Figure 2), as observed by electron microscopy(Nisole and Saib, 2004). During the maturation process, the MLV Gag

protein is cleaved into several proteins: matrix (MA), p12, capsid (CA) and nucleocapsid (NC). Each of them plays an important role in the retroviral replication cycle.

The products of the *gag* gene are involved in viral assembly and also in the early phase of viral infection. The pol gene encodes three enzymes, protease (PR), reverse transcriptase (RT) and integrase (IN), which are all essential for the replication of virus. The transmembrane (TM) and surface (SU) proteins are encoded by the *env* gene, and are responsible for the recognition of host cells and viral entry (Figure 3) (Coffin et al., 1997).

The matrix (MA) protein, located in the interior face of the lipid bilayer of the virion, contains a myristoylation site and a hydrophobic region which are critical for membrane association (Granowitz and Goff, 1994; Rein et al., 1986). Some regions of MA are also required for virion budding and envelope capture (Goff, 2007; Jorgensen et al., 1992).

The MLV p12 protein is a structural protein of the virion. Some mutations of p12 hamper virion release, indicating that the p12 domain acts in the late stages in the context of the Gag precursor (Sabo et al., 2008; Yuan et al., 1999). Additional mutations in this protein specifically affect the early stage of MLV replication (Crawford and Goff, 1984; Yueh and Goff, 2003). Recently studies show that p12 is part of the preintegration complex (PIC), and affect the transport of the PIC to nuclei and the integration of viral DNA (Prizan-Ravid et al., 2010). However the exact function of p12 protein is still unclear.

The capsid (CA) protein is closely packed to form the virion core. Consequently, many mutations in the protein block the formation of the retroviral virion (Schwartzberg et al., 1984). Many gag mutations in CA also affect the early stages of viral infection (Alin

and Goff, 1996; Auerbach et al., 2003). The only highly conserved domain among gag proteins resides in the CA domain, and is termed the Major Homology Region (MHR) (Craven et al., 1995). CA is recognized by several host restriction factors (Wolf and Goff, 2008), such as Trim5 alpha (Perron et al., 2004) and Fv1 (Pincus et al., 1971), and hence determines the tropism of retroviruses.

The nucleocapsid (NC) protein is a short, basic protein that coats the viral RNA in the mature virion. NC interacts with the ψ domain of genomic RNA through a zinc binding Cys-His box and packages the viral RNA into the virion. Thus, mutations in NC domain impair the encapsulation of viral RNA and can dramatically decrease the production of viral particles (Meric and Goff, 1989; Rein et al., 1994). Some other changes in the NC domain, specifically in the Cys-His box, have been shown to affect the reverse transcription of genomic RNA (Gonsky et al., 2001).

The Pol portion of the Gag-Pol polyprotein is processed into three enzymes: protease, reverse transcriptase and integrase. After assembly, the viral particles undergo a maturation process mediated by PR to form infectious viruses (Vogt, 1996). There are morphological and biochemical changes occurring during the process (Figure 2). PR is an aspartyl protease (Toh et al., 1985) and is responsible for processing and maturation of all viral proteins except the envelope glycoprotein (Crawford and Goff, 1985). The envelope glycoprotein is processed into SU and TM proteins by cellular proteinases. PR may also participate in cleavage of cellular protein during viral infection (Shoeman et al., 1993).

The reverse transcriptase (RT) is responsible for reverse transcribing the genomic RNA into viral DNA. It is the hallmark of retroviruses. RT mediates the complex process of reverse transcription that results in the conversion of single-stranded RNA to double-stranded DNA through two enzyme activities: a DNA polymerase active on RNA and DNA template and an RNase H able to digest RNA in the form of an RNA-DNA hybrid (Tanese and Goff, 1988; Telesnitsky and Goff, 1997). RT also has been shown to interact with eukaryotic release factors to modulate the expression of Gag-Pol protein (Orlova et al., 2003).

The integrase (IN) catalyzes the insertion of double-stranded proviral DNA into the host genome. Generally, the integration is not sequence specific, but MLV integration shows a bias in favor of transcriptional start regions (Mitchell et al., 2004).

The *env* gene encodes two proteins SU and TM, which stay on the surface of virion and serve the critical function of binding to host cell and mediate the fusion of virion and cell membrane. The SU serves to bind to the target molecules of host cell and TM serves both to anchor the virion to host cell surface and mediate the fusion during entry. SU and TM are synthesized together, then modified and folded in the ER and cleaved in Golgi. MLV SU but not TM is glycosylated.

A structural representation of a retroviral virion is shown in Figure 2 panel B.

1.4.Retroviral Replication Cycle

The retroviral life cycle can be divided into two phases: the early phase (steps 1-5, figure 3) and the late phase (steps 6-16, figure 3) (Flint, 2000). The early stage of

retroviral replication cycle occurs between entry and integration. The late stage starts from transcription of proviral DNA to maturation of virion.

1.4.1. Early events

The infection process begins with a receptor-dependent attachment of the virus to the surface of targeted cell (Pizzato et al., 1999; Sharma et al., 2000). The interaction between glycoprotein of viruses and the cell surface receptor facilitates the fusion of retroviruses and host cells. The uptake of MLV is mediated by the murine cationic amino-acid transporter (mCAT) (Wang et al., 1991). A human homolog of mCAT is different significantly from mouse mCAT, which explains the inability of MLV to infect human cells (Albritton et al., 1993; Yoshimoto et al., 1993). Amphotropic MLV uses a different receptor, PiT-2, to infect both rodent and non-rodent cells (Miller and Miller, 1994; van Zeijl et al., 1994). In the case of HIV-1, the receptor is CD4 (Berger et al., 1999), but the fusion process needs other co-receptors, such as CXCR4 and CCR5 (Bobardt et al., 2003).

Binding of the envelope glycoprotein of MLV to the receptor and fusion of the viral and host membranes results in the delivery of the viral core into the cell (Albritton et al., 1993; Kizhatil and Albritton, 1997). The interaction between SU and host receptor triggers the conformational changes of TM, leading to the insertion of TM molecule into cell membrane. Lipid rafts have been shown to create an optimal environment for fusion reaction (Lu and Silver, 2000). In some cases, binding of the receptor can trigger internalization of the viral particles before membrane fusion (Katen et al., 2001).

Following fusion, the viral core is disassembled in a process called uncoating. It is the least well understood process of infection. Mutations in Gag proteins, such as MA, CA and NC can block the infection before reverse transcription (Alin and Goff, 1996; Crawford and Goff, 1984; Reicin et al., 1995). The cytoskeleton network has also shown to be important for the entry and uncoating of viral core (Kizhatil and Albritton, 1997).

Reverse transcription is initiated in the cytoplasm (Zhang et al., 2000). The viral reverse transcriptase converts the single-stranded RNA genome to double-stranded DNA. The reverse transcription begins with elongation of a cellular proline tRNA serving as a primer(Palmer et al., 2007), and is finished by a series of transcription and RNA/DNA hybridization and degradation processes (Figure 4(Flint, 2000)) . Two template transfers of short DNA intermediates are required for the complete synthesis of double-stranded DNA (Telesnitsky and Goff, 1997) . The viral DNA is associated with viral proteins and other cellular factors to form the pre-integration complex (PIC)(Fassati and Goff, 1999; Risco et al., 1995).

The exact components of PIC are unclear. For MLV, viral proteins, CA, RT, IN and p12 have been reported to be contained in the PIC (Bowerman et al., 1989; Elis et al., 2012; Fassati and Goff, 1999; Prizan-Ravid et al., 2010). The manner in which PIC travels to the nucleus is also not entirely characterized. MLV PIC has been shown to associate with the microtubule network which indicates that PIC might be transported along the microtubules in cytoplasm (McDonald et al., 2002). The nuclear entry of MLV needs the nuclear membrane to be broken down (Lewis and Emerman, 1994; Roe et al., 1993). Cells

arrested in G1/G0 or G2 phase are not susceptible to MLV infection (Miller et al., 1990; Varmus et al., 1977), and the viral core only has access to the nucleoplasm when the nuclear membrane is broken down during mitosis. The p12 protein has been shown to tether the PIC to mitotic chromosomes (Elis et al., 2012).

In the PIC, several nucleotides at the 3' end of both strand of viral DNA are cleaved off by IN, in an reaction called 3'-end processing. Once inside the nucleus, the PIC targets the host DNA and the DNA with 3'-recessed end attacks the phosphate of cellular DNA, which is called strand transfer. The integration process is completed when DNA repair enzymes fill the gaps between viral DNA and host DNA, which ensures the viral DNA becomes a permanent residence of the host cell genome. Viral DNAs that are not integrated into host genome will form circular DNAs, which serve as a marker of nuclear entry (Gianni et al., 1975; Roth et al., 1990).

1.4.2. Late events

The early events result in the permanent integration of the viral DNA into the host chromosome, forming the provirus. This DNA serves as the template to express genomic RNA and messenger RNA and produce viral protein to assemble new virus for next round of infection. The late events include transcription of the provirus, splicing and transport of viral RNAs, translation to make precursor proteins, packaging of viral genome and assembly of virion, budding and maturation of viruses (Figure 3).

The provirus employs host transcription machinery to make viral RNAs. The U3 region contains the promoter and enhancer for transcription. The RNA is 5' capped and

then transcribed and finally cleaved and polyadenylated at the R-5 border of 3' LTR, which results in an unspliced genomic RNA that is suitable for packaging into viral particles.

Most cellular mRNAs undergo complete splicing before being transported into cytoplasm. Transport to the cytoplasm is prevented before splicing is completed. So retroviruses must develop some mechanism to export unspliced or intron-containing RNAs into cytoplasm, which serve as genomic RNA for packaging or mRNA for making the Gag and Gag-Pol precursor proteins (Pollard and Malim, 1998). Unspliced RNA is believed to contain constitutive export elements (CTE) to promote the export (Bray et al., 1994). For MLV, the CTE is thought to lie in the *gag* region (King et al., 1998). The intron-containing mRNAs of HIV-1 contain a cis-acting element, called the Rev response element (RRE), which is responsible for the export of unspliced mRNAs.

The exported mRNAs of MLV have three destinations: to be translated to make Gag and Gag-Pol proteins,to be spliced to make Env glycoprotein, and to be packaged into virion as genomic RNA. The RNA is incorporated into virions by interactions between the packaging element, Psi, and the NC domain of Gag (Bacharach and Goff, 1998; Berkowitz et al., 1996; Berkowitz et al., 1993; Berkowitz et al., 1995).

In the case of MLV, the *gag* and *gag-pol* coding sequences are in the same reading frame, and are separated by a UAG stop codon (Shinnick et al., 1981; Yoshinaka et al., 1985) . The translation of the RNA results in formation of the Gag protein and approximately 10-20 fold lower levels of the Gag-Pol protein (Hatfield et al., 1992) . A

precise ratio of Gag:Gag-Pro-Pol protein is essential for assembly of retroviral virion (Felsenstein and Goff, 1988). A retroviral recoding sequence (Houck-Loomis et al., 2011) and interaction of RT and eRF1 (Orlova et al., 2003) participate in the suppression of translation termination at the Gag and Gag-pol boundary.

The Env glycoprotein is expressed from a spliced mRNA. The precursor protein is cleaved at the leader sequence and transferred to ER (Abeijon and Hirschberg, 1990). In the ER, the Env protein is glycosylated (Hirschberg and Snider, 1987), folded, presumably with the help of chaperones and oligomerized, which is required for the stable expression of the protein (Tucker et al., 1991). The Env protein is exported to the Golgi apparatus and cleaved into two domains, SU and TM (Dong, Dubay et al. 1992), and then transported to cellular surface via the host vesicular secretory pathway (Selig et al., 1999).

The virion assembly process begins once all viral precursor proteins, Gag, Gag-Pro-Pol and Env, are synthesized. The assembly of MLV and HIV-1 takes place at the plasma membrane (Swanstrom and Wills, 1997). The major player of assembly is the Gag precursor polyprotein. Once the Gag precursor protein arrives at the plasma membrane, it is involved in three interactions, Gag-lipid, Gag-RNA and Gag-Gag, to promote the assembly of viral progeny. The MA domain is essential for the Gag-lipid interaction (Facke et al., 1993; Rowling et al., 1994). Myristylation of Gag has been shown to promote Gag-lipid interaction (Rein et al., 1986). The NC domain of Gag is important for the Gag-RNA and Gag-gag interaction (Krishnan and Pueppke, 1998; Zhang et al., 1996).

Retroviral budding usually recruits cellular factors involved in creating vesicles in a late endosomal compartment called the multivesicular body (MVB), indicating that the budding process might be analogous to MVB process (Garrus et al., 2001; Patnaik et al., 2000). Retroviruses can also bud into the endosomal membrane and exit at the sites of cell-cell contact (Sherer et al., 2003). Mutational analysis revealed that a proline-rich motif Pro-Pro-Pro-Tyr (PPPY) of p12 is required for efficient budding (Yuan et al., 2000; Yuan et al., 1999). Mutations in the PPPY motif result in mutant particles that fail to detach from the virus-expressing cell and remain attached to the plasma membrane (Gottlinger et al., 1991; Huang et al., 1995). After budding, the immature virion is processed by PR to form infectious virus. Mutant PR will result in immature viral particles that are not infectious (Katoh et al., 1985).

1.5.Host restriction factors

Retroviruses are very successful pathogens that infect virtually all branches of life. Host cells have evolved mechanisms that block or negatively regulate the replication of retroviruses under tremendous selection pressure. These mechanisms function at almost every step of retroviral replication (Figure 5) (Goff, 2004; Wolf and Goff, 2008). However, retroviruses, in turn, have evolved mechanisms to inactivate or overcome the host block of replication. By way of introduction, several factors, APOBECs, Trim5α, ZAP, Trim28, Tetherin and SLFN11 will be discussed briefly.

The role of the APOBEC proteins in restricting HIV-1 was discovered during studies of the viral infectivity factor (Vif) , one of the accessory proteins. The replication

of HIV-1 in "permissive" cell lines, such as CEM-SS and SupT1, is Vif independent but in non-permissive cell lines, such as primary CD4+ T-cells and monocyte-derived macrophages, is dependent on Vif (Fisher et al., 1987; Gabuzda et al., 1992; Sakai et al., 1993; Sova and Volsky, 1993; Strebel et al., 1987). Vif-minus virions produced in non-permissive cells are poorly infectious, but Vif-minus virions produced in permissive cells are able to infect both permissive and non-permissive cells (Gabuzda et al., 1992; Sakai et al., 1993; Strebel et al., 1987), which indicates that there is an antiviral factor in non-permissive cells that Vif can neutralize (Madani and Kabat, 1998; Simon et al., 1998). The factor was identified by screening transcripts from non-permissive cells eventually leading to the isolation of APOBEC3G, a cytidine deaminase enzyme (Sheehy et al., 2002). APOBEC3G has been shown to be recruited into HIV-1 virion by interaction with NC domain (Alce and Popik, 2004; Cen et al., 2004) and to introduce C to T mutations in the minus strand of viral DNA, producing G to A mutations in the plus strand (Harris et al., 2003; Mangeat et al., 2003; Zhang et al., 2003), which disrupt the replication of the virus. The block is overcome by Vif-mediated proteasomal degradation of APOBEC3G (Conticello et al., 2003; Liu et al., 2004; Marin et al., 2003). The replication of MLV is not blocked by APOBEC3G, which is believed to be inactivated by MLV PR, although it was shown to be incorporated into MLV virions (Abudu et al., 2006; Douaisi et al., 2004; Mariani et al., 2003).

The Trim5α protein contains three domains: a RING domain, one or two B-boxes, and a coiled-coiled domain (Nisole et al., 2005). The RING domain is usually found in E3

ligases, and there is evidence suggesting that Trim5α is an E3 ligase (Diaz-Griffero et al., 2006; Xu et al., 2003). The B-boxes are interaction domains, which determine the substrate specificity of the E3 ligase activity (Massiah et al., 2007; Massiah et al., 2006). The coiled-coiled domain is responsible for the homo- and hetromultimerization of Trim proteins (Peng et al., 2000; Reymond et al., 2001). There is a B30.2 domain in the C terminus of Trim5α which binds to the CA domain of retroviruses and determines the restriction specificity (Hatziioannou et al., 2004; Nakayama et al., 2005; Perez-Caballero et al., 2005; Perron et al., 2004; Sebastian and Luban, 2005; Stremlau et al., 2006; Stremlau et al., 2005). Species variation in the B30.2 domain of Trim5α leads to their different ability to restrict HIV-1 and other retroviruses (Sawyer et al., 2005; Schaller et al., 2007; Stremlau et al., 2004). For example, rhesus macaque Trim5α restricts HIV-1, and N-tropic MLV but not SIV_{mac}, whereas human Trim5α blocks N-tropic MLV but not B-tropic MLV, HIV-1 and SIV_{mac} (Hatziioannou et al., 2004; Perez-Caballero et al., 2005; Sawyer et al., 2005; Song et al., 2005; Stremlau et al., 2004; Yap et al., 2004). The mechanism that Trim5α uses to restrict retroviruses is still to be fully illuminated. Trim5a is likely to block early steps of the replication cycle of retrovirus, specifically after viral entry and before reverse transcription (Anderson et al., 2006; Himathongkham and Luciw, 1996; Shibata et al., 1995; Wu et al., 2006). The effects might be due to the E3 ligases activity which leads to the proteasomal degradation of CA and other viral proteins (Anderson et al., 2006; Wu et al., 2006). However, it has been shown that Trim5α also mediates proteasomal-independent degradation of viral proteins (Chatterji et al., 2006).

The zinc finger antiviral protein (ZAP) was identified in a screen of cells overexpressing a cDNA library in Rat2 cells and searching for dominant-acting factors restricting viral infection (Gao et al., 2002). The ZAP protein contains four CCCH-type zinc fingers, which are found in a small group of RNA binding proteins known as the tristetraprolin (TTP) tandem zinc finger (TZF) family (Lai et al., 2000). It has been shown that members of TZF family bind specifically to the AU-rich element (ARE) of 3' untranslated region of several cytokine RNAs and lead to their degradation (Carballo et al., 1998, 2000; Lai et al., 1999). This immediately suggested that ZAP might similarly bind RNA elements of retroviruses. It was indeed found that ZAP binds to the 3'LTR of MLV and blocks the accumulation of viral RNA in cytoplasm (Gao et al., 2002; Guo et al., 2004). Mutations in CCCH zinc fingers, which abolish the interaction between ZAP and viral RNA, lead to the loss of antiviral activity of ZAP (Gao et al., 2002). ZAP has also been shown to block the replication of other viruses, such as Sindbis virus (Bick et al., 2003; Guo et al., 2004; Zhang et al., 2007) and Ebola virus (Muller et al., 2007). ZAP interacts with components of the RNA exosome and mediates the degradation of viral mRNAs (Guo et al., 2007). A recent study showed that ZAP also interferes with the function of translational initiation factors, leading to the translational repression of viral mRNA, which surprisingly is required for the degradation of viral mRNA (Zhu et al., 2012).

Besides Trim5α, another Trim family protein, Trim28, was shown to be involved in the transcriptional silencing of retroelements, including exogenous and endogenous

retroviruses, in stem cells (Fasching et al., 2015; Rowe et al., 2010; Wolf and Goff, 2007). The major mechanism of action of Trim28 in retroviral silencing is through its ability to interact with Krab-box containing zinc finger proteins and many transcriptional repression factors (Iyengar and Farnham, 2011). Details of retroviral replication in stem cells will be discussed in the following section. Recently TRIM28Trim28 has also been shown to interact with acetylated Integrase of HIV-1 and mediate its deacetylation through interfering with HDAC1, which prevents the integration of HIV-1 (Allouch et al., 2011).

Like APOBEC3G, Tetherin was identified as a restriction factor by studying another HIV-1 accessary protein, Vpu (Neil et al., 2008; Van Damme et al., 2008). Tetherin is only effective against Vpu-minus HIV-1 virions (Varthakavi et al., 2003). Expressing tetherin in cells producing Vpu-minus virions results in the accumulation of fully formed and infectious virions bound to the extracellular leaflet of outer cell membrane (Neil et al., 2007). Tetherin is upregulated upon stimulation of interferon alpha (IFNα), and specifically inhibits Vpu-minus HIV-1 virions but not wild-type ones (Neil et al., 2008). Tetherin is a membrane protein with unique topology, including an N-terminal cytoplasmic tail, a membrane-spanning helix, an extracellular coiled-coiled domain and a C-terminal glycosylphosphatidylinositol (GPI) anchor (Kupzig et al., 2003; Rollason et al., 2007). It is hypothesized that tetherin bridges an interaction between the cellular membrane and the viral envelope with one of its two lipid-bilayer anchoring domains (Neil et al., 2007). The mechanism by which Vpu agonizes tetherin is not fully understood.

It has been shown that expressing Vpu could down-regulate tetherin from extracellular membranes through a proteasomal independent manner (Van Damme et al., 2008).

Schlafen 11 (SLFN11) is novel retroviral restriction factor that was discovered recently (Li et al., 2012). In mammals, the induction of interferon pathway is an important part of antiviral activity upon viral infection. SLFN gene is a subset of interferon-stimulated genes (ISGs) and induced by the interferon regulatory factor 3 (IRF3) pathway after pathogen infection (Sohn et al., 2007). SLFN11, which specifically abrogates the production of retroviruses, does not affect the early steps of retroviral replication, including reverse transcription, integration and transcription. It interacts with tRNAs and counteracts changes of tRNA pools elicited by presence of retroviruses, which disrupts the codon-usage of retroviruses and impair the production of viral proteins (Li et al., 2012).

In summary, the innate immunity of retroviral infection takes place almost in every step of viral replication. Correspondingly, retroviruses have evolved different kinds of mechanisms to evade the restriction from host factors for successful infection.

1.6.Activities of retroviruses and retrotransposons in different cell types

Retroviruses that infect germ cells or early embryos can introduce their DNA into the germ line and thereby become endogenous retroviruses (Wicker et al., 2007). About 8% of human genome and 10% of mouse genome are made up of endogenous retroviruses (Lander et al., 2001; McCarthy and McDonald, 2004).

Retrotransposons are genetic elements that can replicate themselves in a genome through RNA intermediates. They are very similar to retroviruses except that

retrotransposons do not bud from cells and infect other cells. Retrotransposons can induce mutation in genomes by insertion. In fact, 42% of human genome is composed of retrotransposons (Lander et al., 2001). Most of the retrotransposons are not active due to the accumulation of mutations. The replication of retrotransposons is tightly regulated during development (Figure 6). There are two types of retrotransposons, those with long terminal repeats (LTR) or without long terminal repeat (non-LTR) (Figure 7). LTR retrotransposons have LTRs ranging in size from 100 bps to 5 kbps. Non-LTR retrotransposons are composed of two subgroups, long interspersed elements (LINEs) and short interspersed elements (SINEs) (Singer, 1982).

The replication patterns of retroviruses and retrotransposons in stem cells and differentiated cells are different. LINEs or L1s make up 17% of human genome and only 0.1%-0.2% of those are capable of mobility (Brouha et al., 2003; Cordaux and Batzer, 2009; Doucet et al., 2010). L1 codes two proteins, ORF1 and ORF2, which is flanked by two untranslated regions (UTR) (Cordaux and Batzer, 2009). ORF1 is a RNA binding protein and ORF2 has reverse transcriptase and endonuclease activities, which are similar to the Gag and Pol proteins of retroviruses respectively (Ohshima and Okada, 2005). SINEs, composing 11% of human genome, do not encode function reverse transcriptase, and its mobility is dependent on other retroelements (Cordaux and Batzer, 2009; Santangelo et al., 2007).

Endogenous retroviruses (ERV) are the remnant of previous retroviral infection (Rahm et al., 2011). For most ERVs, extinction is their final fate but some of the ERVs are

still capable of mobility and replication. In mouse genome, there are several active ERVs, and in contrast, most of ERVs in the human genome are defective (Stocking and Kozak, 2008). The expression patterns of ERVs seem to be cell-type specific or developmental stage dependent (Figure 7) (Dupressoir and Heidmann, 1996; Spence et al., 1989; Taruscio and Mantovani, 2004). In 2-cell embryos, many transcripts are initiated from the LTR of ERVs, which might cause the stem cell potency fluctuation along with the activity of endogenous retroviruses (Macfarlan et al., 2012). A subgroup of human ERVs, the HERV-H family, was shown to be important for stem cell renewal (Wang et al., 2014b). ERV activities are also detectable in germ cells but later silenced during differentiation (Matzke et al., 2000; Peaston et al., 2004; Slotkin and Martienssen, 2007). It has been demonstrated that the expression of ERVs is regulated by many cellular factors, including Trim28 (Rowe et al., 2010), ESET, DNA methylation enzymes, histone acetylation enzymes and small RNAs (Rowe and Trono, 2011).

The frequency of L1 activities in different cell types is different. L1 RNA and proteins are predominantly found in germ cells and infrequently in differentiated tissues (Branciforte and Martin, 1994; Trelogan and Martin, 1995). In vivo studies using mouse models have shown that L1 retrotransposition activities are high in germ cells, in early embryo development and some somatic cells (Martin and Branciforte, 1993; Muotri et al., 2005; Ostertag et al., 2002). In vitro L1 retrotransposition assays have suggest that L1 mobilizes in a variety of human and rodent transformed cells (Moran et al., 1996), human embryonic stem cells (Garcia-Perez et al., 2007) and neuron progenitor cells (Muotri et al.,

2005) but relatively rarely in human fibroblasts (Shi et al., 2007) and embryonic carcinoma cells (Garcia-Perez et al., 2010). L1 can serve as a mutagen which result in diseases in mammals (Beck et al., 2011; Kazazian et al., 1988) or as a genomic diversifier, which is important for generating genetic diversity and plasticity of neurons and perhaps for broadening the spectrum of behavior phenotypes (Singer et al., 2010).

Similar to ERVs, exogenous retroviruses are tightly controlled by cellular machineries. MLV replicates efficiently in most dividing cells but is blocked in mouse ES and EC cells (Barklis et al., 1986; Teich et al., 1977). The infected viruses can integrate into the host genome, but transcription from the viral LTR is restricted by trans-acting repressive factors (Akgun et al., 1991; Flanagan et al., 1989). The major target of these repressive factors is the repressor binding site (RBS), which overlap with the PBS of proviral DNA (Feuer et al., 1989; Loh et al., 1990). The RBS acts in an orientation and position-independent manner (Feuer et al., 1989; Petersen et al., 1991). Our lab identified two key components of the repression complex, ZFP809 and Trim28. ZFP809 is a KRAB box containing zinc-finger protein which binds directly to the RBS and initiates the silencing (Wolf and Goff, 2007, 2009). Expression of ZFP809 alone in differentiated cells is sufficient to establish the silencing. It recruits Trim28 and other transcription repressors to silence the transcription of viral genes through placing epigenetic marks on the viral chromatin (Schlesinger and Goff, 2015; Wolf et al., 2008). It has been shown that there are other factors in the silencing complex to mediate full repression of viral transcription (Wang et al., 2014a).

1.7.Dissertation Road map

Retroviral replication is silenced in ES and EC cells by a trans-acting silencing complex, but is not silenced in most differentiated cells. In Chapter 2, the mechanism of regulation of retroviral silencing in differentiated cells will be discussed.

Since ZFP809 is the crucial factor that initiates the silencing of MLV, the mRNA level of ZFP809 was examined by comparing the expression profiles in available databases and as determined by semi-quantitative PCR. Interestingly, the mRNA levels of ZFP809 in differentiated cells and embryonic cells are comparable. Protein degradation assays showed that ZFP809 is degraded rapidly in differentiated cells. We noted the existence of two forms of ZFP809: a longer, full-length protein and a shorter, C-terminally truncated protein ZFP(1-353). A single lysine residue in the C-terminal region of the full-length ZFP809 protein is important for the turnover of ZFP809. The degradation of ZFP809 is subject to the ubiquitin-dependent proteasomal pathway. The C-terminal truncation form of ZFP809, ZFP809(1-353), is somewhat more stable in differentiated cells. Overall, our observations suggest that rapid depletion of ZFP809 proteins by the proteasomal system in differentiated cells result in the loss of retroviral repression in those cells.

In Chapter 3, we studied the function of a novel protein, L1td1, on retroviral silencing in embryonic cells. L1td1 was identified in the set of proteins that interact with Trim28 during purification of the silencing complex. Immunoprecipitation assays showed

that L1td1 interacts with both ZFP809 and Trim28. Depletion of L1td1 using shRNAs partially relieves the restriction from RBS.

An overall discussion of regulation of retroviral silencing in different types of cells and future direction of the work will be presented in Chapter 4.

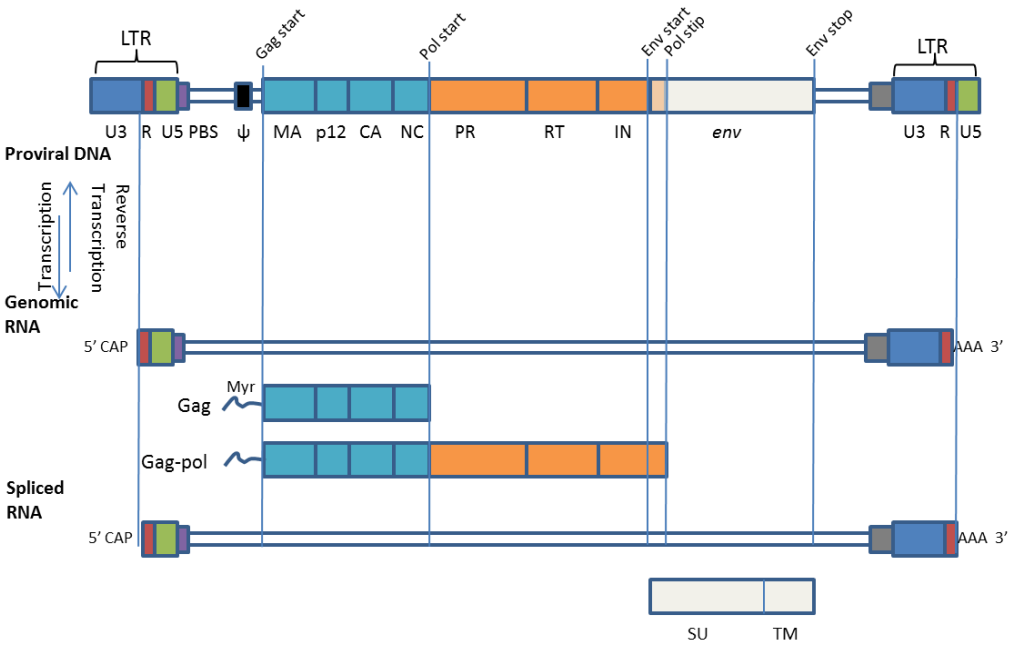

Figure 1 Organization of the Moloney murine leukemia virus (MLV) genome. The *gag* gene encodes structural proteins, *pol* encoding enzymes and *env* encoding the glycoprotein. The 5' LTR promoter drives transcription of both full length genomic and spliced mRNAs. Poly protein precussors *gag* and *gag-pol* are synthesized form the unspliced mRNA and *env* protein from spliced mRNA. The Gag and Gag-Pol polyproteins are cleaved and matured by the viral protease, while Env is processed by cellular proteases. The Gag precusser is cleaved into matrix (MA), p12, capsid (CA) and nucleocapsid (NC); Gag-pol, protease (PR), reverse transcriptase (RT) and integrase (IN); Env, surface (SU) and transmembrane (TM) components of the receptor (adapted from (Petropoulos, 1997)).

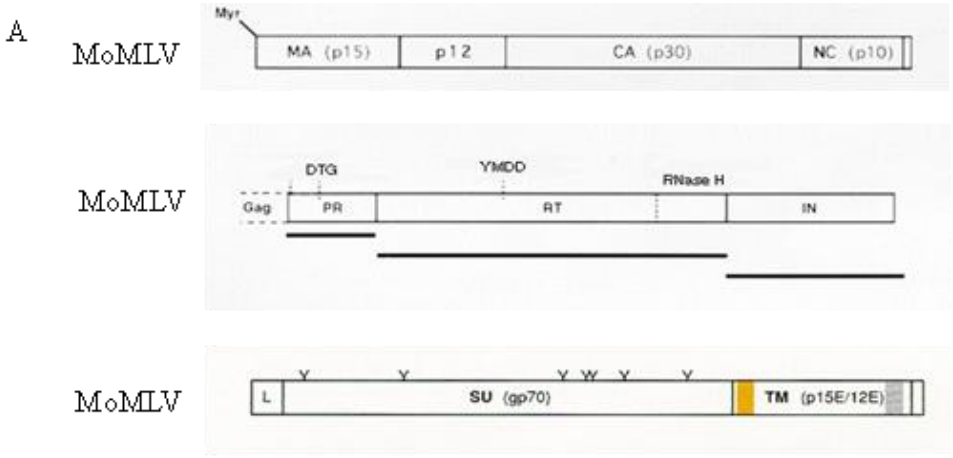

Figure 2 The organization of MLV genome and virion.
A) Viral proteins encoded by viral genes. The *gag* gene encodes MA, p12, CA and NC; the *pol* gene encodes PR, RT and IN and the *env* gene encodes SU and TM. B) Condensation of MLV. The maturation process of MLV was observed by the electron microscopy. C) The structural scheme of retroviral virion (Vogt, 1997).

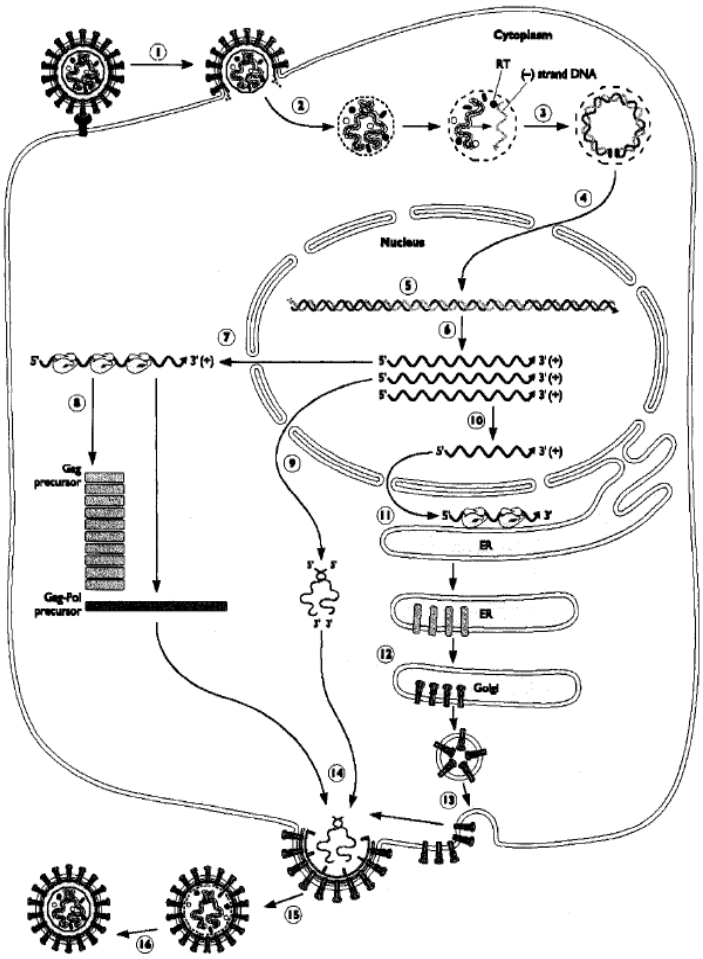

Figure 3 Graphical representation of retroviral replication cycle.
The early phase contains processes that before establishment of integrated proviral DNA, including ① attachment and entry, ② uncoating, ③reverse transcription, ④ transportation and ⑤ integration. Late stage begins with ⑥ viral transcription and continues with ⑦ ⑨ ⑩nuclear export of spliced and unspliced viral RNAs, ⑧ ⑪translation of gag precursor and env genes, ⑫ modification of env proteins, ⑬ ⑭ packaging and assembly, ⑮ release and ⑯ maturation of new viral particles(Flint, 2000).

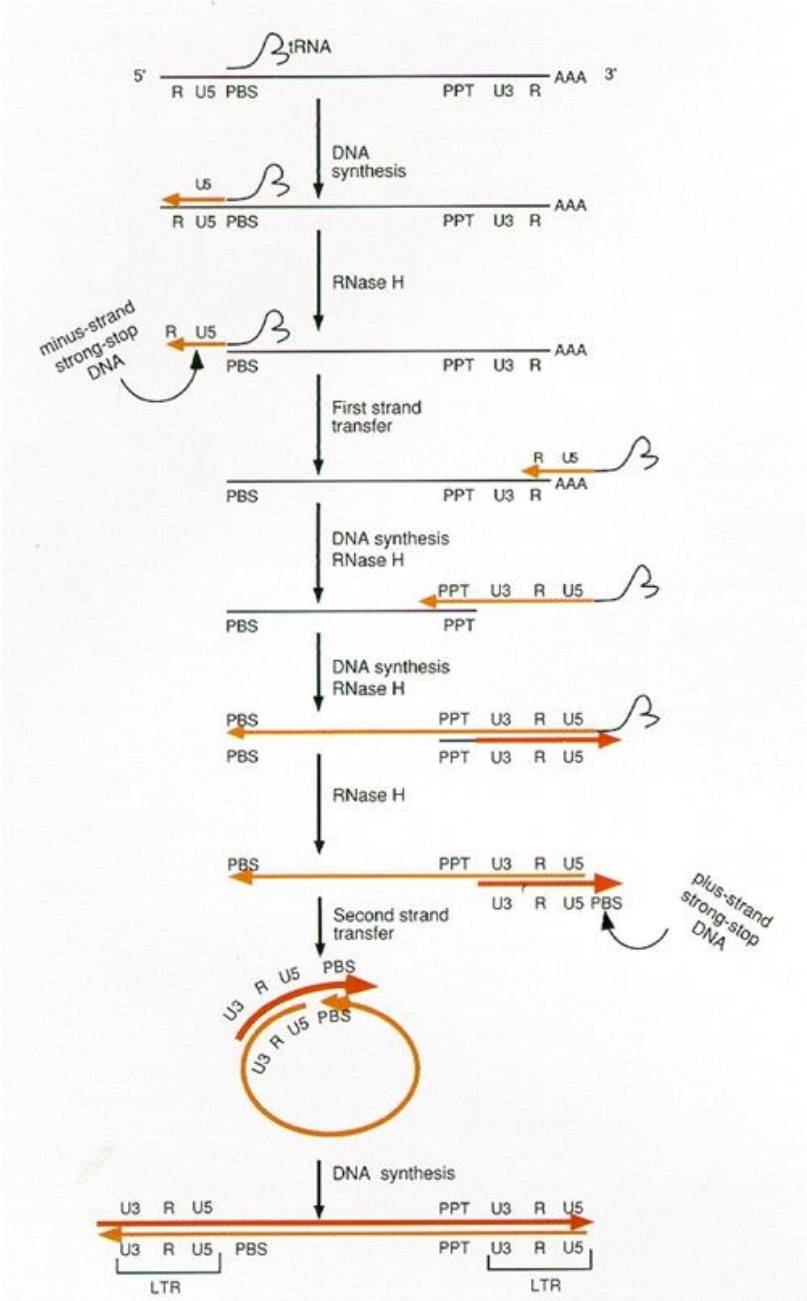

Figure 4 Steps of reverse transcription.
Reverse transcription is initiated by the Proline tRNA binding to the PBS. A series of Rnase H and DNA synthesase activities are undergoing to convert the single-stranded RNA to double-stranded DNA (Telesnitsky and Goff, 1997).

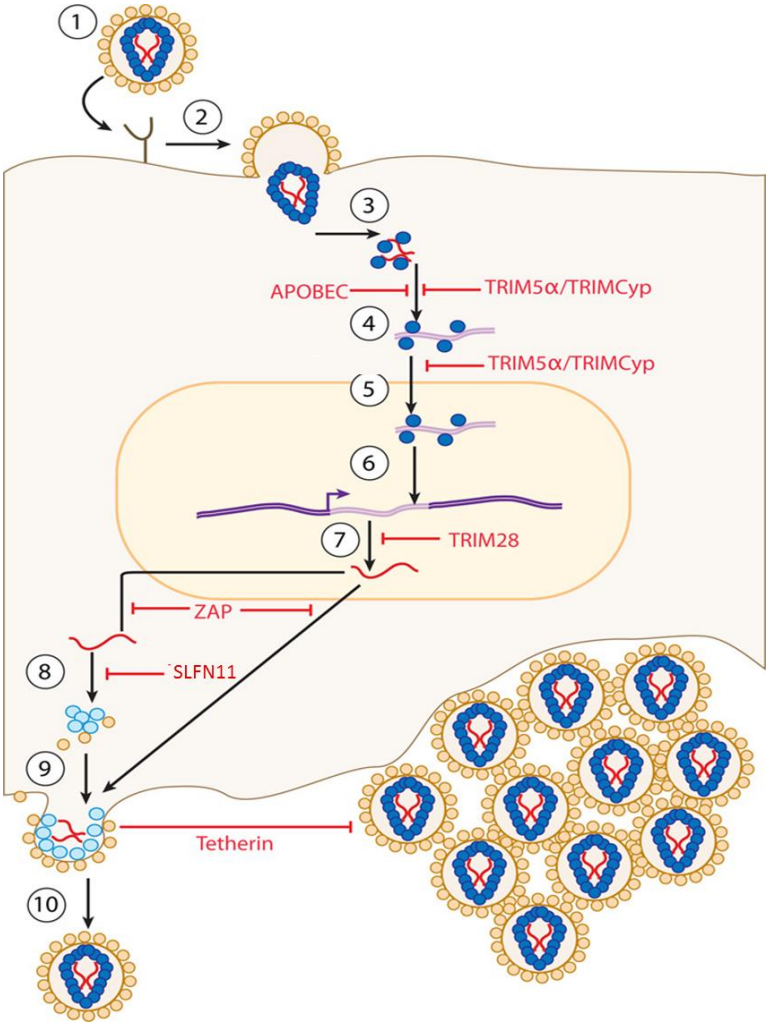

Figure 5 Host restriction factors block different steps of retroviral replication. APOBEC involves in the reverse transcription step. Trim5α inhibits reverse transcription or nuclear entry. Trim28 blocks transcription. ZAP interacts with viral RNAs. SLFN11 inhibit the translation of proteins and tetherin inhibits budding (Modified from (Wolf and Goff, 2008)).

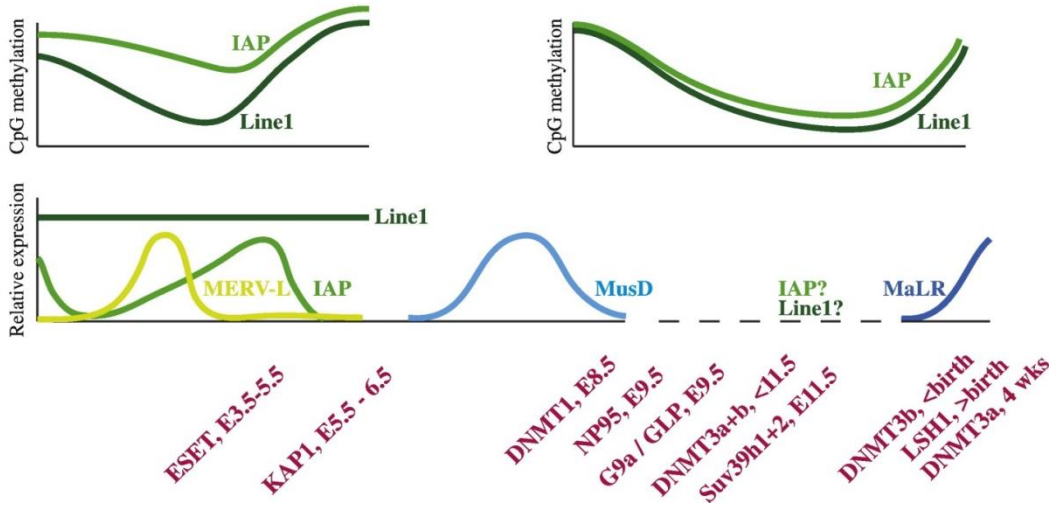

Figure 6 Expression and methylation pattern of endogenous retrotransposons. DNA methylation of IAPs reach their lowest level of around 62% at E3.5 stage, still considerably higher than Line1s that drop to around 25% methylation. There is another peak of demethylation at stage of E11.5. The middle figures show the peak activities of different retrotransposons in different development stage. The factors at the bottom are known to be important in ERV regulation are shown along with the time at which their knockout is lethal (Rowe and Trono, 2011).

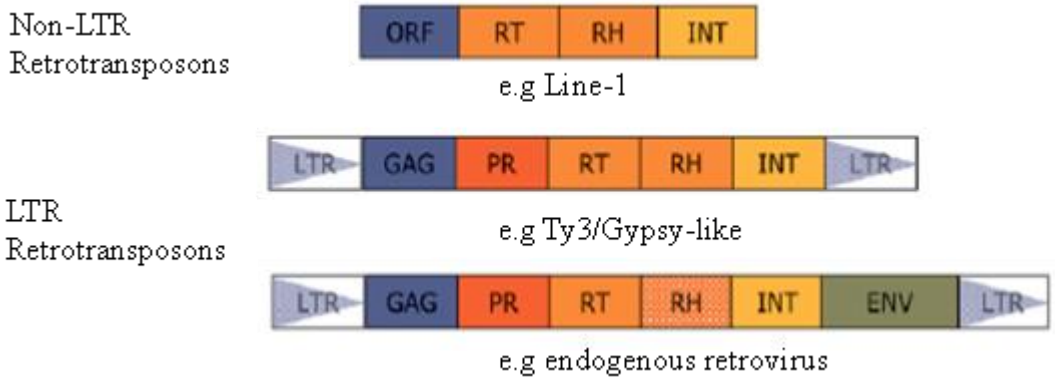

Figure 7 Structure scheme of retrotransposons.
This figure is adapted from (Wicker et al., 2007). Non-LTR retrotransposons encodes proteins retains the structure (ORF) and enzymatic functions (reverse transcriptase RT, Rnase H RH and integrase INT). There are two types of LTR retrotransposons, one with Env gene such as endogenous retrovirus and the other is Ty3 like retrotransposons (PR stands for protease).

1.8.References

Abeijon, C., and Hirschberg, C.B. (1990). Topography of initiation of N-glycosylation reactions. The Journal of biological chemistry *265*, 14691-14695.

Abudu, A., Takaori-Kondo, A., Izumi, T., Shirakawa, K., Kobayashi, M., Sasada, A., Fukunaga, K., and Uchiyama, T. (2006). Murine retrovirus escapes from murine APOBEC3 via two distinct novel mechanisms. Current biology : CB *16*, 1565-1570.

Akgun, E., Ziegler, M., and Grez, M. (1991). Determinants of retrovirus gene expression in embryonal carcinoma cells. Journal of virology *65*, 382-388.

Albritton, L.M., Kim, J.W., Tseng, L., and Cunningham, J.M. (1993). Envelope-binding domain in the cationic amino acid transporter determines the host range of ecotropic murine retroviruses. Journal of virology *67*, 2091-2096.

Alce, T.M., and Popik, W. (2004). APOBEC3G is incorporated into virus-like particles by a direct interaction with HIV-1 Gag nucleocapsid protein. The Journal of biological chemistry *279*, 34083-34086.

Alin, K., and Goff, S.P. (1996). Amino acid substitutions in the CA protein of Moloney murine leukemia virus that block early events in infection. Virology *222*, 339-351.

Allouch, A., Di Primio, C., Alpi, E., Lusic, M., Arosio, D., Giacca, M., and Cereseto, A. (2011). The TRIM family protein KAP1 inhibits HIV-1 integration. Cell host & microbe *9*, 484-495.

Anderson, J.L., Campbell, E.M., Wu, X., Vandegraaff, N., Engelman, A., and Hope, T.J. (2006). Proteasome inhibition reveals that a functional preintegration complex intermediate can be generated during restriction by diverse TRIM5 proteins. Journal of virology *80*, 9754-9760.

Anson, D.S. (2004). The use of retroviral vectors for gene therapy-what are the risks? A review of retroviral pathogenesis and its relevance to retroviral vector-mediated gene delivery. Genetic vaccines and therapy *2*, 9.

Auerbach, M.R., Shu, C., Kaplan, A., and Singh, I.R. (2003). Functional characterization of a portion of the Moloney murine leukemia virus gag gene by genetic footprinting. Proceedings of the National Academy of Sciences of the United States of America *100*, 11678-11683.

Bacharach, E., and Goff, S.P. (1998). Binding of the human immunodeficiency virus type 1 Gag protein to the viral RNA encapsidation signal in the yeast three-hybrid system. Journal of virology *72*, 6944-6949.

Barklis, E., Mulligan, R.C., and Jaenisch, R. (1986). Chromosomal position or virus mutation permits retrovirus expression in embryonal carcinoma cells. Cell *47*, 391-399.

Beck, C.R., Garcia-Perez, J.L., Badge, R.M., and Moran, J.V. (2011). LINE-1 elements in structural variation and disease. Annual review of genomics and human genetics *12*, 187-215.

Berger, E.A., Murphy, P.M., and Farber, J.M. (1999). Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. Annual review of immunology *17*, 657-700.

Berkowitz, R., Fisher, J., and Goff, S.P. (1996). RNA packaging. Current topics in microbiology and immunology *214*, 177-218.

Berkowitz, R.D., Luban, J., and Goff, S.P. (1993). Specific binding of human immunodeficiency virus type 1 gag polyprotein and nucleocapsid protein to viral RNAs detected by RNA mobility shift assays. Journal of virology *67*, 7190-7200.

Berkowitz, R.D., Ohagen, A., Hoglund, S., and Goff, S.P. (1995). Retroviral nucleocapsid domains mediate the specific recognition of genomic viral RNAs by chimeric Gag polyproteins during RNA packaging in vivo. Journal of virology *69*, 6445-6456.

Bick, M.J., Carroll, J.W., Gao, G., Goff, S.P., Rice, C.M., and MacDonald, M.R. (2003). Expression of the zinc-finger antiviral protein inhibits alphavirus replication. Journal of virology *77*, 11555-11562.

Bobardt, M.D., Saphire, A.C., Hung, H.C., Yu, X., Van der Schueren, B., Zhang, Z., David, G., and Gallay, P.A. (2003). Syndecan captures, protects, and transmits HIV to T lymphocytes. Immunity *18*, 27-39.

Bowerman, B., Brown, P.O., Bishop, J.M., and Varmus, H.E. (1989). A nucleoprotein complex mediates the integration of retroviral DNA. Genes & development *3*, 469-478.

Branciforte, D., and Martin, S.L. (1994). Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. Molecular and cellular biology *14*, 2584-2592.

Bray, M., Prasad, S., Dubay, J.W., Hunter, E., Jeang, K.T., Rekosh, D., and Hammarskjold, M.L. (1994). A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. Proceedings of the National Academy of Sciences of the United States of America *91*, 1256-1260.

Brouha, B., Schustak, J., Badge, R.M., Lutz-Prigge, S., Farley, A.H., Moran, J.V., and Kazazian, H.H., Jr. (2003). Hot L1s account for the bulk of retrotransposition in the human population. Proceedings of the National Academy of Sciences of the United States of America *100*, 5280-5285.

Carballo, E., Lai, W.S., and Blackshear, P.J. (1998). Feedback inhibition of macrophage tumor necrosis factor-alpha production by tristetraprolin. Science *281*, 1001-1005.

Carballo, E., Lai, W.S., and Blackshear, P.J. (2000). Evidence that tristetraprolin is a physiological regulator of granulocyte-macrophage colony-stimulating factor messenger RNA deadenylation and stability. Blood *95*, 1891-1899.

Cen, S., Guo, F., Niu, M., Saadatmand, J., Deflassieux, J., and Kleiman, L. (2004). The interaction between HIV-1 Gag and APOBEC3G. The Journal of biological chemistry *279*, 33177-33184.

Chatterji, U., Bobardt, M.D., Gaskill, P., Sheeter, D., Fox, H., and Gallay, P.A. (2006). Trim5alpha accelerates degradation of cytosolic capsid associated with productive HIV-1 entry. The Journal of biological chemistry *281*, 37025-37033.

Coffin, J.M., Hughes, S.H., and Varmus, H.E. (1997). The Interactions of Retroviruses and their Hosts. In Retroviruses, J.M. Coffin, S.H. Hughes, and H.E. Varmus, eds. (Cold Spring Harbor (NY)).

Conticello, S.G., Harris, R.S., and Neuberger, M.S. (2003). The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. Current biology : CB *13*, 2009-2013.

Cordaux, R., and Batzer, M.A. (2009). The impact of retrotransposons on human genome evolution. Nature reviews Genetics *10*, 691-703.

Craven, R.C., Leure-duPree, A.E., Weldon, R.A., Jr., and Wills, J.W. (1995). Genetic analysis of the major homology region of the Rous sarcoma virus Gag protein. Journal of virology *69*, 4213-4227.

Crawford, S., and Goff, S.P. (1984). Mutations in gag proteins P12 and P15 of Moloney murine leukemia virus block early stages of infection. Journal of virology *49*, 909-917.

Crawford, S., and Goff, S.P. (1985). A deletion mutation in the 5' part of the pol gene of Moloney murine leukemia virus blocks proteolytic processing of the gag and pol polyproteins. Journal of virology *53*, 899-907.

Diaz-Griffero, F., Li, X., Javanbakht, H., Song, B., Welikala, S., Stremlau, M., and Sodroski, J. (2006). Rapid turnover and polyubiquitylation of the retroviral restriction factor TRIM5. Virology *349*, 300-315.

Douaisi, M., Dussart, S., Courcoul, M., Bessou, G., Vigne, R., and Decroly, E. (2004). HIV-1 and MLV Gag proteins are sufficient to recruit APOBEC3G into virus-like particles. Biochemical and biophysical research communications *321*, 566-573.

Doucet, A.J., Hulme, A.E., Sahinovic, E., Kulpa, D.A., Moldovan, J.B., Kopera, H.C., Athanikar, J.N., Hasnaoui, M., Bucheton, A., Moran, J.V., *et al.* (2010). Characterization of LINE-1 ribonucleoprotein particles. PLoS genetics *6*.

Dupressoir, A., and Heidmann, T. (1996). Germ line-specific expression of intracisternal A-particle retrotransposons in transgenic mice. Molecular and cellular biology *16*, 4495-4503.

Elis, E., Ehrlich, M., Prizan-Ravid, A., Laham-Karam, N., and Bacharach, E. (2012). p12 tethers the murine leukemia virus pre-integration complex to mitotic chromosomes. PLoS pathogens *8*, e1003103.

Facke, M., Janetzko, A., Shoeman, R.L., and Krausslich, H.G. (1993). A large deletion in the matrix domain of the human immunodeficiency virus gag gene redirects virus particle assembly from the plasma membrane to the endoplasmic reticulum. Journal of virology *67*, 4972-4980.

Fasching, L., Kapopoulou, A., Sachdeva, R., Petri, R., Jonsson, M.E., Manne, C., Turelli, P., Jern, P., Cammas, F., Trono, D., *et al.* (2015). TRIM28 Represses Transcription of Endogenous Retroviruses in Neural Progenitor Cells. Cell reports *10*, 20-28.

Fassati, A., and Goff, S.P. (1999). Characterization of intracellular reverse transcription complexes of Moloney murine leukemia virus. Journal of virology *73*, 8919-8925.

Felsenstein, K.M., and Goff, S.P. (1988). Expression of the gag-pol fusion protein of Moloney murine leukemia virus without gag protein does not induce virion formation or proteolytic processing. Journal of virology *62*, 2179-2182.

Feuer, G., Taketo, M., Hanecak, R.C., and Fan, H. (1989). Two blocks in Moloney murine leukemia virus expression in undifferentiated F9 embryonal carcinoma cells as determined by transient expression assays. Journal of virology *63*, 2317-2324.

Fisher, A.G., Ensoli, B., Ivanoff, L., Chamberlain, M., Petteway, S., Ratner, L., Gallo, R.C., and Wong-Staal, F. (1987). The sor gene of HIV-1 is required for efficient virus transmission in vitro. Science *237*, 888-893.

Flanagan, J.R., Krieg, A.M., Max, E.E., and Khan, A.S. (1989). Negative control region at the 5' end of murine leukemia virus long terminal repeats. Molecular and cellular biology *9*, 739-746.

Flint, S.J., Enquist, L.W., Krug, R.M., Racaniello, V.R., and Skalka, A.M. (2000). Principles of Virology: Molecular Biology, Pathogenesis, and Control (Washington, D.C.: ASM Press).

Gabuzda, D.H., Lawrence, K., Langhoff, E., Terwilliger, E., Dorfman, T., Haseltine, W.A., and Sodroski, J. (1992). Role of vif in replication of human immunodeficiency virus type 1 in CD4+ T lymphocytes. Journal of virology *66*, 6489-6495.

Gao, G., Guo, X., and Goff, S.P. (2002). Inhibition of retroviral RNA production by ZAP, a CCCH-type zinc finger protein. Science *297*, 1703-1706.

Garcia-Perez, J.L., Marchetto, M.C., Muotri, A.R., Coufal, N.G., Gage, F.H., O'Shea, K.S., and Moran, J.V. (2007). LINE-1 retrotransposition in human embryonic stem cells. Human molecular genetics *16*, 1569-1577.

Garcia-Perez, J.L., Morell, M., Scheys, J.O., Kulpa, D.A., Morell, S., Carter, C.C., Hammer, G.D., Collins, K.L., O'Shea, K.S., Menendez, P., *et al.* (2010). Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. Nature *466*, 769-773.

Garrus, J.E., von Schwedler, U.K., Pornillos, O.W., Morham, S.G., Zavitz, K.H., Wang, H.E., Wettstein, D.A., Stray, K.M., Cote, M., Rich, R.L., *et al.* (2001). Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. Cell *107*, 55-65.

Gianni, A.M., Smotkin, D., and Weinberg, R.A. (1975). Murine leukemia virus: detection of unintegrated double-stranded DNA forms of the provirus. Proceedings of the National Academy of Sciences of the United States of America *72*, 447-451.

Goff, S.P. (2004). Genetic control of retrovirus susceptibility in mammalian cells. Annual review of genetics *38*, 61-85.

Goff, S.P. (2007). Host factors exploited by retroviruses. Nature reviews Microbiology *5*, 253-263.

Gonsky, J., Bacharach, E., and Goff, S.P. (2001). Identification of residues of the Moloney murine leukemia virus nucleocapsid critical for viral DNA synthesis in vivo. Journal of virology *75*, 2616-2626.

Gottlinger, H.G., Dorfman, T., Sodroski, J.G., and Haseltine, W.A. (1991). Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. Proceedings of the National Academy of Sciences of the United States of America *88*, 3195-3199.

Granowitz, C., and Goff, S.P. (1994). Substitution mutations affecting a small region of the Moloney murine leukemia virus MA gag protein block assembly and release of virion particles. Virology *205*, 336-344.

Guo, X., Carroll, J.W., Macdonald, M.R., Goff, S.P., and Gao, G. (2004). The zinc finger antiviral protein directly binds to specific viral mRNAs through the CCCH zinc finger motifs. Journal of virology *78*, 12781-12787.

Guo, X., Ma, J., Sun, J., and Gao, G. (2007). The zinc-finger antiviral protein recruits the RNA processing exosome to degrade the target mRNA. Proceedings of the National Academy of Sciences of the United States of America *104*, 151-156.

Harris, R.S., Bishop, K.N., Sheehy, A.M., Craig, H.M., Petersen-Mahrt, S.K., Watt, I.N., Neuberger, M.S., and Malim, M.H. (2003). DNA deamination mediates innate immunity to retroviral infection. Cell *113*, 803-809.

Hatfield, D.L., Levin, J.G., Rein, A., and Oroszlan, S. (1992). Translational suppression in retroviral gene expression. Advances in virus research *41*, 193-239.

Hatziioannou, T., Perez-Caballero, D., Yang, A., Cowan, S., and Bieniasz, P.D. (2004). Retrovirus resistance factors Ref1 and Lv1 are species-specific variants of TRIM5alpha. Proceedings of the National Academy of Sciences of the United States of America *101*, 10774-10779.

Himathongkham, S., and Luciw, P.A. (1996). Restriction of HIV-1 (subtype B) replication at the entry step in rhesus macaque cells. Virology *219*, 485-488.

Hirschberg, C.B., and Snider, M.D. (1987). Topography of glycosylation in the rough endoplasmic reticulum and Golgi apparatus. Annual review of biochemistry *56*, 63-87.

Houck-Loomis, B., Durney, M.A., Salguero, C., Shankar, N., Nagle, J.M., Goff, S.P., and D'Souza, V.M. (2011). An equilibrium-dependent retroviral mRNA switch regulates translational recoding. Nature *480*, 561-564.

Huang, M., Orenstein, J.M., Martin, M.A., and Freed, E.O. (1995). p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. Journal of virology *69*, 6810-6818.

Iyengar, S., and Farnham, P.J. (2011). KAP1 protein: an enigmatic master regulator of the genome. The Journal of biological chemistry *286*, 26267-26276.

Jorgensen, E.C., Pedersen, F.S., and Jorgensen, P. (1992). Matrix protein of Akv murine leukemia virus: genetic mapping of regions essential for particle formation. Journal of virology *66*, 4479-4487.

Katen, L.J., Januszeski, M.M., Anderson, W.F., Hasenkrug, K.J., and Evans, L.H. (2001). Infectious entry by amphotropic as well as ecotropic murine leukemia viruses occurs through an endocytic pathway. Journal of virology *75*, 5018-5026.

Katoh, I., Yoshinaka, Y., Rein, A., Shibuya, M., Odaka, T., and Oroszlan, S. (1985). Murine leukemia virus maturation: protease region required for conversion from "immature" to "mature" core form and for virus infectivity. Virology *145*, 280-292.

Kazazian, H.H., Jr., Wong, C., Youssoufian, H., Scott, A.F., Phillips, D.G., and Antonarakis, S.E. (1988). Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature *332*, 164-166.

King, J.A., Bridger, J.M., Gounari, F., Lichter, P., Schulz, T.F., Schirrmacher, V., and Khazaie, K. (1998). The extended packaging sequence of MoMLV contains a constitutive mRNA nuclear export function. FEBS letters *434*, 367-371.

Kizhatil, K., and Albritton, L.M. (1997). Requirements for different components of the host cell cytoskeleton distinguish ecotropic murine leukemia virus entry via endocytosis from entry via surface fusion. Journal of virology *71*, 7145-7156.

Krishnan, H.B., and Pueppke, S.G. (1998). Genetic characterization of a mutant of Sinorhizobium fredii strain USDA208 with enhanced competitive ability for nodulation of soybean, Glycine max (L.) Merr. FEMS microbiology letters *165*, 215-220.

Kupzig, S., Korolchuk, V., Rollason, R., Sugden, A., Wilde, A., and Banting, G. (2003). Bst-2/HM1.24 is a raft-associated apical membrane protein with an unusual topology. Traffic *4*, 694-709.

Lai, W.S., Carballo, E., Strum, J.R., Kennington, E.A., Phillips, R.S., and Blackshear, P.J. (1999). Evidence that tristetraprolin binds to AU-rich elements and promotes the deadenylation and destabilization of tumor necrosis factor alpha mRNA. Molecular and cellular biology *19*, 4311-4323.

Lai, W.S., Carballo, E., Thorn, J.M., Kennington, E.A., and Blackshear, P.J. (2000). Interactions of CCCH zinc finger proteins with mRNA. Binding of tristetraprolin-related zinc finger proteins to Au-rich elements and destabilization of mRNA. The Journal of biological chemistry *275*, 17827-17837.

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.

Lazo, P.A., Lee, J.S., and Tsichlis, P.N. (1990). Long-distance activation of the Myc protooncogene by provirus insertion in Mlvi-1 or Mlvi-4 in rat T-cell lymphomas. Proceedings of the National Academy of Sciences of the United States of America *87*, 170-173.

Lewis, P.F., and Emerman, M. (1994). Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. Journal of virology *68*, 510-516.

Li, M., Kao, E., Gao, X., Sandig, H., Limmer, K., Pavon-Eternod, M., Jones, T.E., Landry, S., Pan, T., Weitzman, M.D., *et al.* (2012). Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11. Nature *491*, 125-128.

Liu, B., Yu, X., Luo, K., Yu, Y., and Yu, X.F. (2004). Influence of primate lentiviral Vif and proteasome inhibitors on human immunodeficiency virus type 1 virion packaging of APOBEC3G. Journal of virology *78*, 2072-2081.

Loh, T.P., Sievert, L.L., and Scott, R.W. (1990). Evidence for a stem cell-specific repressor of Moloney murine leukemia virus expression in embryonal carcinoma cells. Molecular and cellular biology *10*, 4045-4057.

Lu, X., and Silver, J. (2000). Ecotropic murine leukemia virus receptor is physically associated with caveolin and membrane rafts. Virology *276*, 251-258.

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.

Madani, N., and Kabat, D. (1998). An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral Vif protein. Journal of virology *72*, 10251-10255.

Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., and Trono, D. (2003). Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature *424*, 99-103.

Mariani, R., Chen, D., Schrofelbauer, B., Navarro, F., Konig, R., Bollman, B., Munk, C., Nymark-McMahon, H., and Landau, N.R. (2003). Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. Cell *114*, 21-31.

Marin, M., Rose, K.M., Kozak, S.L., and Kabat, D. (2003). HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. Nature medicine *9*, 1398-1403.

Martin, S.L., and Branciforte, D. (1993). Synchronous expression of LINE-1 RNA and protein in mouse embryonal carcinoma cells. Molecular and cellular biology *13*, 5383-5392.

Massiah, M.A., Matts, J.A., Short, K.M., Simmons, B.N., Singireddy, S., Yi, Z., and Cox, T.C. (2007). Solution structure of the MID1 B-box2 CHC(D/C)C(2)H(2) zinc-binding domain: insights into an evolutionarily conserved RING fold. Journal of molecular biology *369*, 1-10.

Massiah, M.A., Simmons, B.N., Short, K.M., and Cox, T.C. (2006). Solution structure of the RBCC/TRIM B-box1 domain of human MID1: B-box with a RING. Journal of molecular biology *358*, 532-545.

Matsuoka, M. (2003). Human T-cell leukemia virus type I and adult T-cell leukemia. Oncogene *22*, 5131-5140.

Matzke, M.A., Mette, M.F., and Matzke, A.J. (2000). Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. Plant molecular biology *43*, 401-415.

McCarthy, E.M., and McDonald, J.F. (2004). Long terminal repeat retrotransposons of Mus musculus. Genome biology *5*, R14.

McDonald, D., Vodicka, M.A., Lucero, G., Svitkina, T.M., Borisy, G.G., Emerman, M., and Hope, T.J. (2002). Visualization of the intracellular behavior of HIV in living cells. The Journal of cell biology *159*, 441-452.

Meiering, C.D., and Linial, M.L. (2001). Historical perspective of foamy virus epidemiology and infection. Clinical microbiology reviews *14*, 165-176.

Meric, C., and Goff, S.P. (1989). Characterization of Moloney murine leukemia virus mutants with single-amino-acid substitutions in the Cys-His box of the nucleocapsid protein. Journal of virology *63*, 1558-1568.

Miller, D.G., Adam, M.A., and Miller, A.D. (1990). Gene transfer by retrovirus vectors occurs only in cells that are actively replicating at the time of infection. Molecular and cellular biology *10*, 4239-4242.

Miller, D.G., and Miller, A.D. (1994). A family of retroviruses that utilize related phosphate transporters for cell entry. Journal of virology *68*, 8270-8276.

Mitchell, R.S., Beitzel, B.F., Schroder, A.R., Shinn, P., Chen, H., Berry, C.C., Ecker, J.R., and Bushman, F.D. (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS biology *2*, E234.

Moloney, J.B. (1960). Biological studies on a lymphoid-leukemia virus extracted from sarcoma 37. I. Origin and introductory investigations. Journal of the National Cancer Institute *24*, 933-951.

Moran, J.V., Holmes, S.E., Naas, T.P., DeBerardinis, R.J., Boeke, J.D., and Kazazian, H.H., Jr. (1996). High frequency retrotransposition in cultured mammalian cells. Cell *87*, 917-927.

Muller, S., Moller, P., Bick, M.J., Wurr, S., Becker, S., Gunther, S., and Kummerer, B.M. (2007). Inhibition of filovirus replication by the zinc finger antiviral protein. Journal of virology *81*, 2391-2400.

Muotri, A.R., Chu, V.T., Marchetto, M.C., Deng, W., Moran, J.V., and Gage, F.H. (2005). Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature *435*, 903-910.

Nakayama, E.E., Miyoshi, H., Nagai, Y., and Shioda, T. (2005). A specific region of 37 amino acid residues in the SPRY (B30.2) domain of African green monkey TRIM5alpha determines species-specific restriction of simian immunodeficiency virus SIVmac infection. Journal of virology *79*, 8870-8877.

Neil, S.J., Sandrin, V., Sundquist, W.I., and Bieniasz, P.D. (2007). An interferon-alpha-induced tethering mechanism inhibits HIV-1 and Ebola virus particle release but is counteracted by the HIV-1 Vpu protein. Cell host & microbe *2*, 193-203.

Neil, S.J., Zang, T., and Bieniasz, P.D. (2008). Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature *451*, 425-430.

Nisole, S., and Saib, A. (2004). Early steps of retrovirus replicative cycle. Retrovirology *1*, 9.

Nisole, S., Stoye, J.P., and Saib, A. (2005). TRIM family proteins: retroviral restriction and antiviral defence. Nature reviews Microbiology *3*, 799-808.

Ohshima, K., and Okada, N. (2005). SINEs and LINEs: symbionts of eukaryotic genomes with a common tail. Cytogenetic and genome research *110*, 475-490.

Orlova, M., Yueh, A., Leung, J., and Goff, S.P. (2003). Reverse transcriptase of Moloney murine leukemia virus binds to eukaryotic release factor 1 to modulate suppression of translational termination. Cell *115*, 319-331.

Ostertag, E.M., DeBerardinis, R.J., Goodier, J.L., Zhang, Y., Yang, N., Gerton, G.L., and Kazazian, H.H., Jr. (2002). A mouse model of human L1 retrotransposition. Nature genetics *32*, 655-660.

Palmer, M.T., Kirkman, R., Kosloff, B.R., Eipers, P.G., and Morrow, C.D. (2007). tRNA isoacceptor preference prior to retrovirus Gag-Pol junction links primer selection and viral translation. Journal of virology *81*, 4397-4404.

Patnaik, A., Chau, V., and Wills, J.W. (2000). Ubiquitin is part of the retrovirus budding machinery. Proceedings of the National Academy of Sciences of the United States of America *97*, 13069-13074.

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. Developmental 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. Journal of molecular biology *295*, 1139-1162.

Perez-Caballero, D., Hatziioannou, T., Yang, A., Cowan, S., and Bieniasz, P.D. (2005). Human tripartite motif 5alpha domains responsible for retrovirus restriction activity and specificity. Journal of virology *79*, 8969-8978.

Perron, M.J., Stremlau, M., Song, B., Ulm, W., Mulligan, R.C., and Sodroski, J. (2004). TRIM5alpha mediates the postentry block to N-tropic murine leukemia viruses in human cells. Proceedings of the National Academy of Sciences of the United States of America *101*, 11827-11832.

Petersen, R., Kempler, G., and Barklis, E. (1991). A stem cell-specific silencer in the primer-binding site of a retrovirus. Molecular and cellular biology *11*, 1214-1221.

Petropoulos, C.J. (1997). Retroviral Taxonomy, Protein Structures, Sequences and Genetic

Maps, in Retroviruses (ed. J.M. Coffin, S.H. Hughes, and H.E. Varmus), (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Pincus, T., Rowe, W.P., and Lilly, F. (1971). A major genetic locus affecting resistance to infection with murine leukemia viruses. II. Apparent identity to a major locus described for resistance to friend murine leukemia virus. The Journal of experimental medicine *133*, 1234-1241.

Pizzato, M., Marlow, S.A., Blair, E.D., and Takeuchi, Y. (1999). Initial binding of murine leukemia virus particles to cells does not require specific Env-receptor interaction. Journal of virology *73*, 8599-8611.

Pollard, V.W., and Malim, M.H. (1998). The HIV-1 Rev protein. Annual review of microbiology *52*, 491-532.

Prizan-Ravid, A., Elis, E., Laham-Karam, N., Selig, S., Ehrlich, M., and Bacharach, E. (2010). The Gag cleavage product, p12, is a functional constituent of the murine leukemia virus pre-integration complex. PLoS pathogens *6*, e1001183.

Rahm, N., Yap, M., Snoeck, J., Zoete, V., Munoz, M., Radespiel, U., Zimmermann, E., Michielin, O., Stoye, J.P., Ciuffi, A., *et al.* (2011). Unique spectrum of activity of prosimian TRIM5alpha against exogenous and endogenous retroviruses. Journal of virology *85*, 4173-4183.

Reicin, A.S., Paik, S., Berkowitz, R.D., Luban, J., Lowy, I., and Goff, S.P. (1995). Linker insertion mutations in the human immunodeficiency virus type 1 gag gene: effects on virion particle assembly, release, and infectivity. Journal of virology *69*, 642-650.

Rein, A., Harvin, D.P., Mirro, J., Ernst, S.M., and Gorelick, R.J. (1994). Evidence that a central domain of nucleocapsid protein is required for RNA packaging in murine leukemia virus. Journal of virology *68*, 6124-6129.

Rein, A., McClure, M.R., Rice, N.R., Luftig, R.B., and Schultz, A.M. (1986). Myristylation site in Pr65gag is essential for virus particle formation by Moloney murine leukemia virus. Proceedings of the National Academy of Sciences of the United States of America *83*, 7246-7250.

Reymond, A., Meroni, G., Fantozzi, A., Merla, G., Cairo, S., Luzi, L., Riganelli, D., Zanaria, E., Messali, S., Cainarca, S., *et al.* (2001). The tripartite motif family identifies cell compartments. The EMBO journal *20*, 2140-2151.

Risco, C., Menendez-Arias, L., Copeland, T.D., Pinto da Silva, P., and Oroszlan, S. (1995). Intracellular transport of the murine leukemia virus during acute infection of NIH 3T3 cells: nuclear import of nucleocapsid protein and integrase. Journal of cell science *108 (Pt 9)*, 3039-3050.

Roe, T., Reynolds, T.C., Yu, G., and Brown, P.O. (1993). Integration of murine leukemia virus DNA depends on mitosis. The EMBO journal *12*, 2099-2108.

Rollason, R., Korolchuk, V., Hamilton, C., Schu, P., and Banting, G. (2007). Clathrin-mediated endocytosis of a lipid-raft-associated protein is mediated through a dual tyrosine motif. Journal of cell science *120*, 3850-3858.

Rosenberg, N., and Jolicoeur, P. (1997). Retroviral Pathogenesis. In Retroviruses, J.M. Coffin, S.H. Hughes, and H.E. Varmus, eds. (Cold Spring Harbor (NY)).

Roth, M.J., Schwartzberg, P., Tanese, N., and Goff, S.P. (1990). Analysis of mutations in the integration function of Moloney murine leukemia virus: effects on DNA binding and cutting. Journal of virology *64*, 4709-4717.

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.

Rowling, P.J., McLaughlin, S.H., Pollock, G.S., and Freedman, R.B. (1994). A single purification procedure for the major resident proteins of the ER lumen: endoplasmin, BiP, calreticulin and protein disulfide isomerase. Protein expression and purification *5*, 331-336.

Sabo, Y., Laham-Karam, N., and Bacharach, E. (2008). Basal budding and replication of the murine leukemia virus are independent of the gag L domains. Journal of virology *82*, 9770-9775.

Sakai, H., Shibata, R., Sakuragi, J., Sakuragi, S., Kawamura, M., and Adachi, A. (1993). Cell-dependent requirement of human immunodeficiency virus type 1 Vif protein for maturation of virus particles. Journal of virology *67*, 1663-1666.

Santangelo, A.M., de Souza, F.S., Franchini, L.F., Bumaschny, V.F., Low, M.J., and Rubinstein, M. (2007). Ancient exaptation of a CORE-SINE retroposon into a highly conserved mammalian neuronal enhancer of the proopiomelanocortin gene. PLoS genetics *3*, 1813-1826.

Sawyer, S.L., Wu, L.I., Emerman, M., and Malik, H.S. (2005). Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. Proceedings of the National Academy of Sciences of the United States of America *102*, 2832-2837.

Schaller, T., Hue, S., and Towers, G.J. (2007). An active TRIM5 protein in rabbits indicates a common antiviral ancestor for mammalian TRIM5 proteins. Journal of virology *81*, 11713-11721.

Schlesinger, S., and Goff, S.P. (2015). Retroviral Transcriptional Regulation and Embryonic Stem Cells: War and Peace. Molecular and cellular biology *35*, 770-777.

Schwartzberg, P., Colicelli, J., Gordon, M.L., and Goff, S.P. (1984). Mutations in the gag gene of Moloney murine leukemia virus: effects on production of virions and reverse transcriptase. Journal of virology *49*, 918-924.

Sebastian, S., and Luban, J. (2005). TRIM5alpha selectively binds a restriction-sensitive retroviral capsid. Retrovirology *2*, 40.

Selig, L., Pages, J.C., Tanchou, V., Preveral, S., Berlioz-Torrent, C., Liu, L.X., Erdtmann, L., Darlix, J., Benarous, R., and Benichou, S. (1999). Interaction with the p6 domain of the gag precursor mediates incorporation into virions of Vpr and Vpx proteins from primate lentiviruses. Journal of virology *73*, 592-600.

Shackleford, G.M., and Varmus, H.E. (1988). Construction of a clonable, infectious, and tumorigenic mouse mammary tumor virus provirus and a derivative genetic vector. Proceedings of the National Academy of Sciences of the United States of America *85*, 9655-9659.

Sharma, S., Miyanohara, A., and Friedmann, T. (2000). Separable mechanisms of attachment and cell uptake during retrovirus infection. Journal of virology *74*, 10790-10795.

Sheehy, A.M., Gaddis, N.C., Choi, J.D., and Malim, M.H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature *418*, 646-650.

Sherer, N.M., Lehmann, M.J., Jimenez-Soto, L.F., Ingmundson, A., Horner, S.M., Cicchetti, G., Allen, P.G., Pypaert, M., Cunningham, J.M., and Mothes, W. (2003). Visualization of retroviral replication in living cells reveals budding into multivesicular bodies. Traffic *4*, 785-801.

Shi, X., Seluanov, A., and Gorbunova, V. (2007). Cell divisions are required for L1 retrotransposition. Molecular and cellular biology *27*, 1264-1270.

Shibata, R., Sakai, H., Kawamura, M., Tokunaga, K., and Adachi, A. (1995). Early replication block of human immunodeficiency virus type 1 in monkey cells. The Journal of general virology *76 (Pt 11),* 2723-2730.

Shinnick, T.M., Lerner, R.A., and Sutcliffe, J.G. (1981). Nucleotide sequence of Moloney murine leukaemia virus. Nature *293,* 543-548.

Shoeman, R.L., Sachse, C., Honer, B., Mothes, E., Kaufmann, M., and Traub, P. (1993). Cleavage of human and mouse cytoskeletal and sarcomeric proteins by human immunodeficiency virus type 1 protease. Actin, desmin, myosin, and tropomyosin. The American journal of pathology *142,* 221-230.

Simon, J.H., Gaddis, N.C., Fouchier, R.A., and Malim, M.H. (1998). Evidence for a newly discovered cellular anti-HIV-1 phenotype. Nature medicine *4,* 1397-1400.

Singer, M.F. (1982). SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. Cell *28,* 433-434.

Singer, T., McConnell, M.J., Marchetto, M.C., Coufal, N.G., and Gage, F.H. (2010). LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? Trends in neurosciences *33,* 345-354.

Slotkin, R.K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. Nature reviews Genetics *8,* 272-285.

Sohn, W.J., Kim, D., Lee, K.W., Kim, M.S., Kwon, S., Lee, Y., Kim, D.S., and Kwon, H.J. (2007). Novel transcriptional regulation of the schlafen-2 gene in macrophages in response to TLR-triggered stimulation. Molecular immunology *44,* 3273-3282.

Song, B., Javanbakht, H., Perron, M., Park, D.H., Stremlau, M., and Sodroski, J. (2005). Retrovirus restriction by TRIM5alpha variants from Old World and New World primates. Journal of virology *79,* 3930-3937.

Sova, P., and Volsky, D.J. (1993). Efficiency of viral DNA synthesis during infection of permissive and nonpermissive cells with vif-negative human immunodeficiency virus type 1. Journal of virology *67*, 6322-6326.

Spence, S.E., Gilbert, D.J., Swing, D.A., Copeland, N.G., and Jenkins, N.A. (1989). Spontaneous germ line virus infection and retroviral insertional mutagenesis in eighteen transgenic Srev lines of mice. Molecular and cellular biology *9*, 177-184.

Stevenson, M. (2003). HIV-1 pathogenesis. Nature medicine *9*, 853-860.

Stocking, C., and Kozak, C.A. (2008). Murine endogenous retroviruses. Cellular and molecular life sciences : CMLS *65*, 3383-3398.

Strebel, K., Daugherty, D., Clouse, K., Cohen, D., Folks, T., and Martin, M.A. (1987). The HIV 'A' (sor) gene product is essential for virus infectivity. Nature *328*, 728-730.

Stremlau, M., Owens, C.M., Perron, M.J., Kiessling, M., Autissier, P., and Sodroski, J. (2004). The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. Nature *427*, 848-853.

Stremlau, M., Perron, M., Lee, M., Li, Y., Song, B., Javanbakht, H., Diaz-Griffero, F., Anderson, D.J., Sundquist, W.I., and Sodroski, J. (2006). Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. Proceedings of the National Academy of Sciences of the United States of America *103*, 5514-5519.

Stremlau, M., Perron, M., Welikala, S., and Sodroski, J. (2005). Species-specific variation in the B30.2(SPRY) domain of TRIM5alpha determines the potency of human immunodeficiency virus restriction. Journal of virology *79*, 3139-3145.

Swanstrom, R., and Wills, J.W. (1997). Synthesis, Assembly, and Processing of Viral Proteins. In Retroviruses, J.M. Coffin, S.H. Hughes, and H.E. Varmus, eds. (Cold Spring Harbor (NY)).

Tanese, N., and Goff, S.P. (1988). Domain structure of the Moloney murine leukemia virus reverse transcriptase: mutational analysis and separate expression of the DNA polymerase and RNase H activities. Proceedings of the National Academy of Sciences of the United States of America *85*, 1777-1781.

Taruscio, D., and Mantovani, A. (2004). Factors regulating endogenous retroviral sequences in human and mouse. Cytogenetic and genome research *105*, 351-362.

Teich, N.M., Weiss, R.A., Martin, G.R., and Lowy, D.R. (1977). Virus infection of murine teratocarcinoma stem cell lines. Cell *12*, 973-982.

Telesnitsky, A., and Goff, S.P. (1997). Reverse Transcriptase and the Generation of Retroviral DNA. In Retroviruses, J.M. Coffin, S.H. Hughes, and H.E. Varmus, eds. (Cold Spring Harbor (NY)).

Toh, H., Kikuno, R., Hayashida, H., Miyata, T., Kugimiya, W., Inouye, S., Yuki, S., and Saigo, K. (1985). Close structural resemblance between putative polymerase of a Drosophila transposable genetic element 17.6 and pol gene product of Moloney murine leukaemia virus. The EMBO journal *4*, 1267-1272.

Trelogan, S.A., and Martin, S.L. (1995). Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. Proceedings of the National Academy of Sciences of the United States of America *92*, 1520-1524.

Tucker, S.P., Srinivas, R.V., and Compans, R.W. (1991). Molecular domains involved in oligomerization of the Friend murine leukemia virus envelope glycoprotein. Virology *185*, 710-720.

Van Damme, N., Goff, D., Katsura, C., Jorgenson, R.L., Mitchell, R., Johnson, M.C., Stephens, E.B., and Guatelli, J. (2008). The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. Cell host & microbe *3*, 245-252.

van Zeijl, M., Johann, S.V., Closs, E., Cunningham, J., Eddy, R., Shows, T.B., and O'Hara, B. (1994). A human amphotropic retrovirus receptor is a second member of the gibbon ape

leukemia virus receptor family. Proceedings of the National Academy of Sciences of the United States of America *91*, 1168-1172.

Varmus, H.E., Padgett, T., Heasley, S., Simon, G., and Bishop, J.M. (1977). Cellular functions are required for the synthesis and integration of avian sarcoma virus-specific DNA. Cell *11*, 307-319.

Varthakavi, V., Smith, R.M., Bour, S.P., Strebel, K., and Spearman, P. (2003). Viral protein U counteracts a human host cell restriction that inhibits HIV-1 particle production. Proceedings of the National Academy of Sciences of the United States of America *100*, 15154-15159.

Vogt, V.M. (1996). Proteolytic processing and particle maturation. Current topics in microbiology and immunology *214*, 95-131.

Vogt, V.M. (1997). Retroviral Virions and Genomes. In Retroviruses, J.M. Coffin, S.H. Hughes, and H.E. Varmus, eds. (Cold Spring Harbor (NY)).

Wang, G.Z., Wolf, D., and Goff, S.P. (2014a). EBP1, a novel host factor involved in primer binding site-dependent restriction of moloney murine leukemia virus in embryonic cells. Journal of virology *88*, 1825-1829.

Wang, H., Kavanaugh, M.P., North, R.A., and Kabat, D. (1991). Cell-surface receptor for ecotropic murine retroviruses is a basic amino-acid transporter. Nature *352*, 729-731.

Wang, J., Xie, G., Singh, M., Ghanbarian, A.T., Rasko, T., Szvetnik, A., Cai, H., Besser, D., Prigione, A., Fuchs, N.V., *et al.* (2014b). Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. Nature *516*, 405-409.

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., *et al.* (2007). A unified classification system for eukaryotic transposable elements. Nature reviews Genetics *8*, 973-982.

Wolf, D., Cammas, F., Losson, R., and Goff, S.P. (2008). Primer binding site-dependent restriction of murine leukemia virus requires HP1 binding by TRIM28. Journal of virology *82*, 4675-4679.

Wolf, D., and Goff, S.P. (2007). TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. Cell *131*, 46-57.

Wolf, D., and Goff, S.P. (2008). Host restriction factors blocking retroviral replication. Annual review of genetics *42*, 143-163.

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

Wu, X., Anderson, J.L., Campbell, E.M., Joseph, A.M., and Hope, T.J. (2006). Proteasome inhibitors uncouple rhesus TRIM5alpha restriction of HIV-1 reverse transcription and infection. Proceedings of the National Academy of Sciences of the United States of America *103*, 7465-7470.

Xu, L., Yang, L., Moitra, P.K., Hashimoto, K., Rallabhandi, P., Kaul, S., Meroni, G., Jensen, J.P., Weissman, A.M., and D'Arpa, P. (2003). BTBD1 and BTBD2 colocalize to cytoplasmic bodies with the RBCC/tripartite motif protein, TRIM5delta. Experimental cell research *288*, 84-93.

Yap, M.W., Nisole, S., Lynch, C., and Stoye, J.P. (2004). Trim5alpha protein restricts both HIV-1 and murine leukemia virus. Proceedings of the National Academy of Sciences of the United States of America *101*, 10786-10791.

Yoshimoto, T., Yoshimoto, E., and Meruelo, D. (1993). Identification of amino acid residues critical for infection with ecotropic murine leukemia retrovirus. Journal of virology *67*, 1310-1314.

Yoshinaka, Y., Katoh, I., Copeland, T.D., and Oroszlan, S. (1985). Murine leukemia virus protease is encoded by the gag-pol gene and is synthesized through suppression of an amber termination codon. Proceedings of the National Academy of Sciences of the United States of America *82*, 1618-1622.

Yuan, B., Campbell, S., Bacharach, E., Rein, A., and Goff, S.P. (2000). Infectivity of Moloney murine leukemia virus defective in late assembly events is restored by late assembly domains of other retroviruses. Journal of virology *74*, 7250-7260.

Yuan, B., Li, X., and Goff, S.P. (1999). Mutations altering the moloney murine leukemia virus p12 Gag protein affect virion production and early events of the virus life cycle. The EMBO journal *18*, 4700-4710.

Yueh, A., and Goff, S.P. (2003). Phosphorylated serine residues and an arginine-rich domain of the moloney murine leukemia virus p12 protein are required for early events of viral infection. Journal of virology *77*, 1820-1829.

Zhang, H., Dornadula, G., Orenstein, J., and Pomerantz, R.J. (2000). Morphologic changes in human immunodeficiency virus type 1 virions secondary to intravirion reverse transcription: evidence indicating that reverse transcription may not take place within the intact viral core. Journal of human virology *3*, 165-172.

Zhang, H., Yang, B., Pomerantz, R.J., Zhang, C., Arunachalam, S.C., and Gao, L. (2003). The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. Nature *424*, 94-98.

Zhang, W.H., Hockley, D.J., Nermut, M.V., Morikawa, Y., and Jones, I.M. (1996). Gag-Gag interactions in the C-terminal domain of human immunodeficiency virus type 1 p24 capsid antigen are essential for Gag particle assembly. The Journal of general virology *77 (Pt 4)*, 743-751.

Zhang, Y., Burke, C.W., Ryman, K.D., and Klimstra, W.B. (2007). Identification and characterization of interferon-induced proteins that inhibit alphavirus replication. Journal of virology *81*, 11246-11255.

Zhu, Y., Wang, X., Goff, S.P., and Gao, G. (2012). Translational repression precedes and is required for ZAP-mediated mRNA decay. The EMBO journal *31*, 4236-4246.

Chapter 2

Proteasomal regulation of the ZFP809 retroviral repressor

2.1 Abstract

The replication of Moloney Murine Leukemia Virus (MLV) is strongly suppressed in mouse embryonic stem cells (ES) and carcinoma cells (EC) but not in differentiated cells. Proviral DNAs are specifically silenced in embryonic cells by a large complex containing Trim28 (also known as KAP-1), HP1 (heterochromatin protein 1), and chromatin modifying enzymes that repress transcription of the viral DNA. This silencing complex is targeted to a short DNA sequence element, which overlaps with Primer Binding Site (PBS) of MLV, by the DNA sequence-specific zinc finger protein ZFP809, which is specifically expressed in embryonic cells. The ectopic expression of ZFP809 in differentiated cells is sufficient to induce silencing, suggesting that ZFP809 is the key embryonic-specific component of the silencing machinery. To determine the mechanism of regulation of ZFP809, we examined the levels of ZFP809 mRNA and protein. Though the levels of ZFP809 mRNA were similar in embryonic and differentiated cells, ZFP809 protein was found to be stable in embryonic cells and highly unstable in differentiated cells. The protein was stabilized by the proteasome inhibitor MG132 and accumulated heavy modification by addition of polyubiquitin chains. A short sequence of amino acids at the C-terminus of ZFP809, including a single lysine residue, was required and sufficient to mediate the rapid turnover of the protein. Depletion of Trim28 stabilizes ZFP809 in differentiated cells. Abolishing the interaction between Trim28 and ZFP809 by introducing mutation in ZFP809 also stabilizes the protein, which indicating that Trim28 promotes the degradation of ZFP809 in differentiated cells. The neddylation inhibitor MLN4924 also stabilizes ZFP809 protein in differentiated cells.

2.2 Introduction

Replication of murine leukemia viruses is blocked in murine embryonic stem (ES) cells and other primitive stem cells (Teich et al., 1977; Wolf and Goff, 2007). Infection results in normal entry, reverse transcription and integration of proviral DNA, but the DNA is then silenced transcriptionally. A specific DNA sequence utilized to prime viral DNA synthesis, the Primer Binding Site complementary to the proline tRNA (PBSpro), is present on those viruses that are most strongly and rapidly silenced. A large protein complex is specifically expressed in ES cells and binds to the PBS sequences to mediate the transcriptional repression of proviral DNA (Barklis et al., 1986; Petersen et al., 1991; Wolf and Goff, 2008; Yamauchi et al., 1995). Work from our lab and others have shown that this silencing complex contains the transcriptional repressor Trim28, also known as KAP-1 or TIF1b. The complex is tethered to viral DNA by ZFP809, a zinc finger protein with sequence-specific DNA binding activity (Wolf et al., 2008; Wolf and Goff, 2007, 2009). ZFP809 binds directly to the PBSpro of proviral DNA, initiating the silencing (Wolf and Goff, 2009).

Trim28 contains several conserved domains, including a RBCC (Ring finger, two B-box zinc fingers, and a coiled coil) domain, a central TIF1 signature sequence (TSS) domain and a C-terminal combination plant homeodomain (PHD) and bromo domain (Friedman et al., 1996; Iyengar and Farnham, 2011). The RBCC domain binds as a homotrimer to a KRAB domain and encapsulates it to a protease resistant core (Peng et al., 2007). After tethering to a specific DNA region by KRAB-ZNFs, Trim28 recruits a number of other factors including

the NuRD histone deacetylase complex, the histone H3K9 methyltransferase ESET, and HP1 to silence the specific DNA area (Le Douarin et al., 1996; Schultz et al., 2002; Schultz et al., 2001). Trim28 also plays an important role in the protein modification pathway. The Ring domain containing proteins is a large family of E3 ligases (Deshaies and Joazeiro, 2009; Saurin et al., 1996). MAGE-C2-Trim28 was shown to target p53 for proteasome-dependent degradation (Doyle et al., 2010). The PHD domain of Trim28 functions as an intracellular E3 ligase that sumolyates adjacent Bromo domain which is essential for the interaction between Trim28 and ESET methyltransferase (Ivanov et al., 2007; Zeng et al., 2008).

ZFP809 contains two domains, a KRAB box in the N-terminal which is responsible for the interaction with Trim28 and a zinc finger domain containing seven zinc fingers and having the DNA binding activity (Figure 1A). While most of the components of the silencing complex are expressed ubiquitously, ZFP809 is expressed specifically in ES cells and not in differentiated cells. Ectopic expression of ZFP809 alone in differentiated cells is sufficient to establish the silencing in those cells (Wolf and Goff, 2009), suggesting that the lack of restriction of MLV in differentiated cells is largely due to the reduced amount of ZFP809 protein. The basis of the cell-type specific expression of ZFP809 is not known. ES cells have unusual transcriptional profiles, and a natural supposition would be that the ZFP809 gene is specifically transcribed in ES cells. But the cell-type expression of proteins such as ZFP809 could be effected by regulation at any of several steps: at the level of transcription(Sproul et al., 2005), post-transcriptional mRNA processing or stability, mRNA translation (Chen and

Rajewsky, 2007), or post-translational protein modification or stability (Haas et al., 1982). To test this idea, we have directly examined the synthesis and stability of the ZFP809 protein. We find that the down regulation of ZFP09 in differentiated cells is not due to the reduction of ZFP809 mRNA but to changes in protein stability. The full-length ZFP809 protein is relatively stable in ES cells but is degraded rapidly in differentiated cells. Mutagenesis analysis and protein ubiquitination assays indicate that the degradation is mediated through the ubiquitin-dependent proteosomal pathway. A lysine residue in the C-terminal domain is essential for the turnover of ZFP809. Depletion of Trim28 or abolishing the interaction between Trim28 and ZFP809 also stabilizes ZFP809 in differentiated cells, which indicates that Trim28 promotes the degradation of ZFP809 in those cells.

2.3 Methods

2.3.1 Cell culture and transfection

293A cells are human embryonic kidney cells. 293T cells are human embryonic kidney cells that stably express simian virus 40 (SV40) large T-Antigen. NIH3T3 cells are mouse embryonic fibroblast cells. Rat2 cells are rat fibroblast cells. 293A, 293T, NIH3T3 and Rat2 cell lines were maintained in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum, 100 UI/ml penicillin and 100mg/ml streptomycin. Mouse ES cells, E14, were cultured on the MEF layer in the medium of DMEM with 20% defined FBS, nonessential amino acids, 2uM sodium pyruvate, 100 UI/ml penicillin, 100 mg/ml streptomycin and 1000 units of LIF (Stem Cell Technologies). E14 cells were transferred to

plates coated with 0.1% gelatin for transfection purpose. DNAs were transfected into cells using polyethylenimine reagent (Akinc et al., 2005) or Lipofectamine 2000 according to the manufacturer's instruction. All cells were cultured at 37°C in 5% CO_2.

2.3.2 Plasmids

The plasmids of pcDNA3.1 3X Flag ZFP809(FL) or ZFP809(1-353) were described before (Wolf and Goff, 2009). Fragments with 3X Flag ZFP809(FL) or ZFP809(1-353) were swiped to pLVX IRES mCherry vector using the SnaBI and XhoI sites. 3X Flag ZFP809(FL) were cloned into pLVX EF1α IRES puro vector using the EcoRI and BamHI restriction sites. Renilla luciferase gene was cloned into pcDNA3.1 3XFlag vector using BamHI and HindIII restriction site. The 5'UTR, cDNA sequence and 3'UTR of ZFP809 mRNA were inserted between HindIII and NotI after the stop codon of the Renilla luciferase gene in pcDNA3.1 Renilla luciferase vector. Plasmids of pcDNA3 Flag p53 and pCI-his-Ubiquitin were purchased from Addgene. Mutations of ZFP809 were introduced using overlap extension PCR (Higuchi et al., 1988) and SLIC cloning (Li and Elledge, 2012), which replaced the last lysine residue (K391A) of ZFP809 to alanine or residues(D8V9 to AA). The C terminus of ZFP809 was cloned into pEGFP C1 vector using EcoRI and BamHI restriction site to generate pEGFP ZFP809 C50. While cloning of pcDNA3.1 3XFlag Renilla luciferase, a SalI right before the stop codon of Renilla luciferase gene for inducing ZFP809 C50 in the C terminal of Renilla luciferase. The oligo sequences for introducing mutations are listed below.

Oligonucleotides for cloning pcDNA3.1 3XFlag Renilla luciferase

Forward primer CATCATCACCATCACCATaccggtATG ACTTCGAAAGTTTATGATC

Reverse primer

TTCCAGTCTTTGAAATCTCAaccggtTTAGTCGACTTGTTCATTTTTGAGAACTCGC

Oligonucleotides for cloning pcDNA3.1 3XFlag Renilla luciferase ZFP809 5UTR, CDS or 3UTR

5UTR forward primer atgc AAGCTT GGGGTGGATTTAGAAAGTTCAGA

5UTR reverse primer agct GCGGCCGC ATGATGGGACTGGGAGCGTC

CDS forward primer atgc AAGCTT CATGGGGTTGGTGTCCTTT

CDS reverse primer agct GCGGCCGC TCAGAGCCCACCCCCAAT

3UTR forward primer atgc AAGCTT TGGGCTCTGAGATTTCAAAGAC

3UTR reverse primer agct GCGGCCGC TTTTTTTTTTTTTTTTTTTTTTTTTTTT

Oligonucleotides for cloning pcDNA3.1 3XFlag ZFP809(DV to AA)

Forward primer (restriction site Nhe1) AGGGAGACCCAAGCTGGCTA

DVR CACGGCTGCAGCCTCAAAGGA

DVF TCCTTTGAGGCTGCAGCCGTG

Reverse primer (restriction site Xho1) AACGGGCCCTCTAGACTCGA

Oligonucleotides for cloning pcDNA3.1 3X Flag ZFP809(K391A)

Forward primer (restriction site BamH1) TTGGTACCGAGCTCGGATC

K391AR ATGGCATTATGGACTTTGTGGCAGCGTGGAGCAGGTGTATCAT

K391AF ATGATACACCTGCTCCACGCTGCCACAAAGTCCATAATGCCAT

Reverse primer (restriction site Xho1) AACGGGCCCTCTAGACTCGA

Oligonucleotides for generating pEGFP ZFP809 C50

Forward primer tagc GAATTC CAAGAGGTAGTTTTCCAACAGATCA

Reverse primer tcga GGATCC TCAGAGCCCACCCCCAATGATA

2.3.3 Digital analysis of ZFP809 mRNA levels

The RNA expression data of ZFP809 was obtained from the RNA expression data published by Meissner Group (Mikkelsen et al., 2008) and the online database BioGPS. The cell lines, B cell, MEF (mouse embryonic fibroblast), and ES (embryonic stem) were generated from ROSA26–M2rtTA mice (Beard et al., 2006). IPS (induced pluripotent stem) cells were induced from MEFs. Expression data were generated from polyA RNA using GeneChip Mouse Genome 430 2.0 Arrays (Affymetrix). Absolute values were normalized by Robust Multi-Array (RMA) and processed using Genepattern (http://www.broad.mit.edu/cancer/software/genepattern/) (Mikkelsen et al., 2008). ZFP809 expression data in mouse tissues of NK, T-cell CD4+, T-cell CD8+, bone marrow and ES cells were selected in BioGPS database.

2.3.4 Analysis of ZFP809 mRNA level by quantitative PCR

Total RNA of E14, F9 and NIH3T3 cells was extracted using TRIZOL reagent (Invitrogen). 4 µg of total RNA per cell line were used to produce cDNA using random hexamers and SuperScript III kit (Invitrogen). 2 µl of each cDNA were used for quantitative

PCR analysis of ZFP809 and GAPDH mRNA levels. The expression of ZFP809 was normalized to the level of GAPDH.

2.3.5 Colony formation assay

NIH3T3 cells were transfected with pcDNA3.1 3X Flag, pcDNA3.1 3C Flag ZFP809(FL) or ZFP809(1-353). 48 hours after transfection, cells were selected in G418 (1ug/ml) and counted after 14 days of selection.

2.3.6 Protein degradation assay

All reagents were purchased from Sigma and dissolved in DMSO. 24 hours after transfecting plasmids into 293T, 293A or E14 (mouse ES) cells, 50μg/ml of cycloheximide (CHX) was added to stop translation (Schneider-Poetsch, Ju et al. 2010). Cells were lysed by addition of Laemmli buffer at different time points after drugs were added: 0, 10min, 30min, 1h, 2h, 3h and subjected to immunoblotting with anti-Flag antibody for measuring the amount of ZFP809 proteins. ES cells were treated with Trichostatin A (TSA, 500nM) and Forskolin (FSK, 0.1mM) for 4 hours before addition of CHX. Drugs of MG132 (30μM), Chloroquine (100μM), MLN4924 (1mM) were added with CHX for indicated purposed.

2.3.7 In vivo ubiquitination assay

293T cells were transfected with DNAs expressing ZFP809(FL) only or both ZFP809(FL) and his-tagged ubiquitin. 24 hours after transfection, MG132 was added to stop protein degradation for 4 hours. Cells were lysed in buffer A (6 M guanidinium-HCl, 0.1 M

Na2HPO4/NaH2PO4, 0.01 M Tris–HCl pH 8.0, 5 mM imidazole, 10 mM β-mercaptoethanol) and the cell lysates were incubated with Ni^{2+}-NTA beads (Qiagen) for 4 h at room temperature. The beads were washed with buffer A, B (8 M urea, 0.1 M Na_2PO_4/NaH_2PO_4, 0.01 M Tris–HCl pH 8.0, 10 mM β-mercaptoethanol), C (8 M urea, 0.1 M Na_2PO_4/NaH_2PO_4, 0.01 M Tris–HCl pH 6.3, 10 mM β-mercaptoethanol), and bound proteins were eluted with Laemmli buffer (modified from (Wang et al., 2005)). Samples were analyzed by western blotting with anti-Flag antibody (Sigma).

2.3.8 Western blotting

The proteins were resolved by SDS-polyacrylamide gel electrophoresis (PAGE). After transfer to nitrocellulose membranes, the blots were probed with mouse anti Flag monoclonal antibody (Sigma), mouse anti mCherry monoclonal antibody (abcam), mouse anti TRIM28 monoclonal antibody (abcam), mouse anti Tubulin monoclonal antibody (abcam) or mouse anti beta-actin monoclonal antibody. Secondary antibody was ECL anti mouse IgG, horseradish peroxidase linked whole antiboday (GE healthcare). Membranes were developed using ECL reagents (Life Technologies).

2.3.9 Luciferase assay

293T cells seeded in a 6-well plate were transfected with empty vector, Renilla luciferase, and Renilla luciferase with ZFP809 mRNA elements. Transfection was carried out using polyethylenimine (PEI) protocol in which the constructs were added at a 3 (Renilla luciferase) : 1 (firefly luciferase, transfection efficiency control) ratio. Cells were lysed 48-

hours post transfection with Passive lysis buffer (Promega) and the luciferase activity was measured using a Dual-Glo luciferase assay system (Promega) in a Veritas Microplate luminometer. Renilla luciferase data were normalized to Firefly luciferase readings in each well.

2.3.10 Knockdown of TRIM28

For the transient knockdown of Trim28, 293A cells were transfected with siRNA specifically targeting Trim28 or a non-targeting siRNA (Thermo Scientific). Cells were transfected with Lipofectamine2000 (Invitrogen) using a 2-day transfection protocol followed by 24-hours or 48-hours of rest to gain maximum knockdown efficiency. Cells were collected following the resting period and analyzed by western blotting.

2.3.11 Protein interaction assay

Same amount plasmids of pcDNA3 3X Flag, pcDNA3 3Xflag ZFP809(FL) or pcDNA3 3X Flag ZFP809(DV-AA) were transfected into 293T cells. 24 hours after transfection, cells were lysed in IPH buffer. Cell lysates were incubated with anti-Flag magnetic beads (sigma). After 4 hours of incubation, beads were washed with modified IPH buffer (the concentration of detergent was increased to 1%) and eluted with Laemmli buffer. The elution was probed with western blotting.

2.4 Results

2.4.1 Expression of mRNAs and proteins of ZFP809 in different cells types

The transcriptional silencing complex binding to the PBSpro of the murine leukemia viruses is only present in nuclear extracts of ES and other primitive cell types, and not in differentiated cells. The key ES-cell specific component of this complex is ZFP809. Expression of ZFP809 alone in differentiated cells are sufficient to re-establish the restriction, which indicates that lack of ZFP809 protein in differentiated cells result in the loss of restriction. Surprisingly, RNA expression databases indicate that the mRNA levels of ZFP809 are roughly comparable among murine ES, IPS and differentiated cells (Figure 8A) (Mikkelsen et al., 2008). ZFP809 mRNA levels were not specifically higher in ES cells, but sometimes were modestly lower in ES cells than in other tissues (Figure 8B) (Wu et al., 2009) .

To confirm these whole-genome expression data, we directly assessed the ZFP809 mRNA levels in different cell types by real time PCR. The mRNA levels of ZFP809 in mouse ES cells (E14), EC cells (F9) and fibroblast cells (NIH3T3) were all comparable (Figure 8C), suggesting that the expression was not controlled by differential transcription, nor differential stability of the mRNA.

Our initial efforts to establish cell lines stably expressing the full-length ZFP809(FL) were not immediately successful (Wolf and Goff, 2009), Transformation of differentiated

cells with DNA constructs expressing the full-length ZFP809 cDNA resulted in the isolation of cell lines expressing only very low levels of the protein. Screening large numbers of colonies arising after transformation yielded a few that expressed high levels, but these lines were found to express a truncated form. Then we cloned a short-form of the gene, ZFP809(1-353), lacking the C-terminal 50 amino acids and retaining the Zinc fingers and Krab box (Wolf and Goff, 2009). Ecotopic expression of ZFP809(1-353) in vectors carrying neomycin resistance gene resulted in much more colonies than that of ZFP809(FL) after G418 selection (Figure 9). Overall, the observations suggest that full-length ZFP809 may be toxic to differentiated cells.

To confirm these observations, we transformed 293T cells with the same amount of DNAs encoding epitope flag-tagged versions of the full-length cDNA (ZFP809(FL)) or the truncated form (ZFP809(1-353)), and examined the levels of expressed protein by Western blotting (Figure 10A). The levels of the full-length protein were significantly lower than that of the truncated form (Figure 10A). To rule out the possibility that the difference in expression was due to inconsistency of transfection efficiency or mRNA stability, both forms of ZFP809 cDNA were cloned into a vector expressing a single bicistronic mRNA encoding both ZFP809 and an mCherry reporter translated from an IRES element. 293T or NIH3T3 cells were transformed with the DNAs, and the levels of the expressed proteins were assessed by Western blotting. The full-length ZFP809 was expressed at much lower levels than ZFP809 (1-353), while the levels of mCherry protein were approximately the same in both

293T and NIH3T3 cells (Figure 10B). These results suggest that full-length ZFP809 protein is significant down-regulated by translational or post-translational mechanisms.

2.4.2 Rapid degradation of full-length ZFP809 protein in differentiated cells

Introducing a heterologous posttransciptional or translational regulator in the form of an RNA element to the mRNA of a reporter gene could establish the corresponding regulation of the reporter gene (Ha et al., 1996). To test whether sequences of ZFP809 mRNA would regulate the expression of other genes, we sub-cloned the 5'UTR, 3'UTR and cDNA sequences of ZFP809 mRNA and a scrambled sequence after the stop codon of Renilla luciferase reporter gene to make vectors Rluc, Rluc-5UTR, Rluc-CDS, Rluc-3UTR and Rluc-scram (Figure 11A). 48 hours after transfecting same amount of those vectors with a Firefly luciferase vector as a control into 293T cells, the luciferase units were measured by the Promega dual-luciferase reporter assay system. Renilla luciferase units were normalized to firefly luciferase units to generate relative Renilla luciferase unit (RLU). RLU of different DNAs was approximately the same, which suggested that neither post-transcriptional RNA regulation nor translation were likely points of regulation (Figure 11B).

To test if the low levels of full-length ZFP809 protein in differentiated cells were due to protein degradation, protein degradation assays were performed. Flag-tagged ZFP809(FL) and ZFP809(1-353) were expressed by transfection into 293T cells and 24 hours after transfection, cycloheximide (CHX) was added to block translation (Schneider-Poetsch et al., 2010). Cells were lysed at various times after addition of CHX (0, 10min, 30min, 1 hour, 2

hours and 3hours). Then the levels of ZFP809 were assessed using Western blotting. The blotting showed that levels of full-length ZFP809 protein dropped rapidly upon inhibition of translation, while the levels of ZFP809(1-353) remained relatively constant in those cells (Figure 12A). In order to calculate the half-life of ZFP809(FL), the intensity of each band was figured out by ImageJ. Exponential regression of the band intensity showed that the half-life of ZFP809 full-length protein was about 45min (Figure 12B).

Due to the low transfection efficiency of mouse ES cells and rapid silencing of CMV promoter in ES cells (Chung et al., 2002), the expression of ZFP809 driven by CMV promoter was undetectable after transfection. As a result, two drugs, Trichostatin A (TSA) (Spenger et al., 2004) and forskolin (FSK) (Keller et al., 2007) were added to release the restriction of CMV promoter in mouse ES cells when performing protein degradation assays in ES cells. TSA is a fungal antibiotic selective inhibiting mammalian histone deacetylase enzymes which act as transcription repressors (Vanhaecke et al., 2004). FSK activates the enzyme adenylyl cyclase and increases intracellular levels of cAMP, which stimulates the expression of promoter enhancer CREB (Wilkinson and Akrigg, 1992). EF1α promoter was shown to work efficiently in ES cells and hence was used to drive the expression of ZFP809 (Hong et al., 2007). The protein degradation assay showed that full-length ZFP809 proteins are stable in E14 ES cells (Figure 13). Since expression of ZFP809 in differentiated cells alone could restore the restriction, loss of silencing of MLV in differentiated cells might be due to the low abundance of ZFP809 protein, which is rapidly degraded in those cells.

Attaching small unstable peptides to proteins may induce the turnover and degradation of tagged proteins (Li et al., 1998). We showed that the short peptide in the C terminal is essential for the turnover of ZFP809. To test if this peptide can destabilize other proteins, we tagged the peptide of ZFP809 C terminus (ZFP809 C50) to the C terminal of EGFP and Renilla luciferase. Protein degradation assay showed that ZFP809 C50 tagged-GFP is stable in differentiated cells and luciferase units of Renilla C50 did not reduce much after addition of CHX. Those observations showed that the peptide alone is not sufficient to drive the turnover of those proteins (Figure 14A and B).

2.4.3 Proteasomal degradation of ZFP809

The intracellular degradation of proteins might be achieved in two major cellular pathways, lysosome and proteasome (Peters et al., 1994; Settembre et al., 2013). Lysosome is capable of breaking down all kinds of biomolecules, including proteins, nucleic acids and cell debris. Lysosomes are important organelles of protein degradation, particularly under stress conditions and most of the proteolysis of cytosolic proteins that occurs in lysosomes is relatively nonspecific (Knop et al., 1993). A lysosomotropic agent, chloroquine (CQ), can accumulate inside the lysosome and inhibit lysosomal protein degradation (Shintani and Klionsky, 2004). The main function of proteasomes is to selectively turn over and degrade unwanted or damaged proteins through proteolysis (Etlinger and Goldberg, 1977; Hochstrasser, 1996). Proteasomes count for the 80% of protein degradation in cells. The 26S proteasomes contain one 20S core subunit and two 19S regulatory cap subunits (Peters et al.,

1994). MG132 is a potent, membrane-permeable proteasome inhibitor (Rock et al., 1994). To find out through which pathway ZFP809 is degraded, MG132, CQ or DMSO was added with CHX to test if they are capable of stabilizing ZFP809 in differentiated cells. Immunoblotting of samples collected at different time points after adding drug showed that MG132 alone, but not CQ can stabilize the full-length protein of ZFP809 while block of translation in differentiated cells, which indicates that full-length ZFP809 is degraded through the proteosomal pathway, not the lysosomal pathway (Figure 15A and B). MG132 also stabilizes ZFP809 in a rodent cell line Rat2 (Figure 15C).

2.4.4 Proteasomal degradation of ZFP809 is ubiquitin-dependent

Degradation of proteins by proteasome is in two different manners, ubiquitin-dependent and ubiquitin-independent (Rape and Jentsch, 2002; Thrower et al., 2000). Ubiquitin-dependent pathway, with covalent link of Ubiquitin proteins to a lysine residue of the substrate, requires the coordination of three enzymes, the ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2) and ubiquitin ligase (E3). The ubiquitin ligase recognizes the substrate and transfer the ubiquitin from E2 to the substrate. Therefore, E3 confers the substrate specificity of the proteosomal system (Risseeuw et al., 2003). Proteins with intrinsically unstructured regions, might be targeted for ubiquitin-independent degradation (Asher et al., 2006). We previously showed that the 50 amino acids in the C-terminal of ZFP809 are responsible for the instability of ZFP809 (Figure 12A). Deletion of the 50 amino acids dramatically stabilizes ZFP809 in differentiated cells (Figure 12A). There

is only one lysine residue in this region. To test if the only lysine residue in C terminal of ZFP809 is important for the turnover of ZFP809, We mutated the lysine residue to alanine to make the construct ZFP809(K391A). Protein degradation assay showed that ZFP809(K391A) is stable in differentiated cells (Figure 16A and B), which indicates a role for the lysine residue on the turnover of ZFP809.

Proteins undergoing ubiquitin-mediated degradation are attached multiple copies of ubiquitin, which shows up in the form of a ladder when probed by western blotting(Maki et al., 1996). A plasmid expressing His-tagged ubiquitin (Young et al., 2011) was co-transfected with vectors expressing flag-tagged ZFP809(FL) or p53 (Yu et al., 1999) into 293T cells. 24 hours after transfection, MG132 was added to stop degradation for 4 hours. Then cells were lysed and the lysates were incubated with Ni-NTA beads (Medina et al., 2000). The bound proteins were eluted and subjected to western blotting using anti-Flag antibody. Flag-tagged ZFP809 and p53 were detected in lanes added with His-tagged ubiquitin. The precipitated ZFP809 and p53 have larger molecular weights than original Flag-tagged ones, which demonstrates modification by ubiquitin proteins (Figure 17).

In summary, ZFP809 is ubiquitinated in differentiated cells and the lysine residue in the C-terminal is essential for the turnover of ZFP809, suggesting that full-length ZFP809 may be degraded through the ubiquitin-dependent pathway. The ZFP809 mutation, ZFP809(K391A) are relatively stable in differentiated cells, suggesting that the lysine residue, K391 is important for the turnover of ZFP809. Full length ZFP809 is heavily ubiquitinated in

differentiated cells. To assess the ubiquitination status of ZFP809 FL and mutants, the in vivo ubiquitination assay was performed with vectors ZFP809 FL, ZFP809(K391A) and ZFP809(1-353). Western blotting showed that ZFP809(K391A) is also polyubiquitinated but the ubiquitination level is lower than that of ZFP809 FL (Figure 18), which consists with the observation that those mutations are comparatively stable than ZFP809 FL in differentiated cells. Interestingly, ZFP809(1-353) is mainly monoubiquitinated (Figure 18). Monoubiquitination is not a signal for protein degradation, which explains that ZFP809(1-353) is very stable in differentiated cells.

2.4.5 Trim28 promotes the degradation of ZFP809

Full-length ZFP809 protein is relatively stable in stem cells but not in differentiated cells, which indicates that there may be a factor in differentiated cells that could specifically target ZFP809 to the ubiquitin-dependent protein degradation. Usually the specificity of substrate selection is determined by the E3 ligases. Trim28, which binds to the KRAB box of ZFP809, is a RING domain containing proteins and has been shown to serve as an E3 ligase (Liang et al., 2011). RING domain containing proteins, are a large family of E3 ligases and usually associate Cullin proteins to form the Cullin-Ring ligases (CRL) complexes (Bennett et al., 2010). Cullin proteins serve as a scaffold for protein degradation. Neddylation of Cullin proteins results in structural change on those proteins, which allow the transfer of ubiquitin proteins from E2, E3 ligases to substrates (Rabut and Peter, 2008). To test the role of Trim28 in the degradation of ZFP809 in differentiated cells, we used siRNA to knock down Trim28

in 293A cells (Figure 19A). Protein degradation assays showed that ZFP809 (FL) was stabilized in Trim28 depleted cells by siRNA (Figure 19B), which suggests that Trim28 plays a role in ZFP809 degradation.

It has been previously shown that substitution of the amino acids DV with alanine residues in KRAB box of zinc finger proteins would abolish the interaction between KRAB box and Trim28 (Peng et al., 2000). To investigate the relevance of ZFP809-Trim28 interaction and ZFP809 degradation, we mutated the DV residues in the KRAB box of ZFP809 to make ZFP809(DV-AA). Exogenous expressed ZFP809 in 293T cells successfully pulled down Trim28 but not ZFP809(DV-AA) (Figure 20A), which indicates the interaction between Trim28 and ZFP809 has been abolished. Protein degradation assay showed that ZFP809(DV-AA) degraded in a much slower rate than wild-type ZFP809 (Figure 20B and C). The ubiquitination level of ZFP809(DV-AA) is much less than that of ZFP809(FL), which explains that ZFP809(DV-AA) is relatively stable over ZFP809(FL) (Figure 18). Those observations suggest that interaction between ZFP809 and Trim28 is essential for the degradation of ZFP809.

Overall, depletion of Trim28 or abolishing the interaction between Trim28 and ZFP809 stabilized full-length ZFP809 in differentiated cells, suggesting that Trim28 can promote the degradation of ZFP809 in differentiated cells.

2.4.6 Neddylation inhibitor MLN4924 stabilizes ZFP809

Trim28 is a RING domain containing protein, a large protein family that have been shown to assemble with multi-subunit Cullin proteins to form Cullin-Ring E3 ubiquitin ligase (CRL) complexes (Sarikas et al., 2011). Neddylation activates the CRL complexes and promotes their protein degradation activities (Merlet et al., 2009). A small drug MLN4924 was shown to block the neddylation activity of nedd8, which is required for the activation of cullin proteins to target proteins for degradation (Garcia et al., 2014; Godbersen et al., 2014). To test if the neddylation is important for ZFP809 degradation, we performed protein degradation assay with the neddylation inhibitor MLN4924. MLN4924 or DMSO was added with CHX in the experiment. Immunoblotting showed that MLN4924, but not DMSO, stabilized the ZFP809(FL) in differentiated cells (Figure 21). As MLN4924 blocks the neddylation activities, the stabilization of ZFP809 by MLN4924 indicates that neddylation is important for the degradation of ZFP809, suggesting that CRL may be involved in the degradation of ZFP809.

2.5 Discussion

ES cells and other primitive cell types exhibit unusual mRNA profiles, and much of their specific patterns of gene expression can be attributed to unusual chromatin modifications and transcriptional activity of their genomic DNAs. However, these cells also exhibit unusual patterns of global protein modifications including ubiquitinylation (Sutovsky et al., 2001), and indeed the very maintenance of their pluripotency requires unusual profiles

of ubiquitinylation and proteosomal degradation (Buckley et al., 2012) (van der Stoop et al., 2008) (for review, see (Naujokat and Saric, 2007)). In this study, we found that the ZFP809 protein, the key component of a retroviral silencing complex, is regulated by the proteasomal protein degradation system.

ZFP809 protein is expressed in ES cells but not in most differentiated cells and is responsible for the rapid and efficient silencing of exogenous (Wolf and Goff, 2009) and endogenous (Wolf et al., 2015) retrovirus DNAs in ES cells. Ectopic expression of ZFP809 in murine or human differentiated cells is sufficient to establish silencing of MLV proviruses, suggesting that it is the limiting factor and that all other components of the silencing machinery are present in most cells. Surprisingly, the levels of ZFP809 mRNAs are comparable In ES cells and differentiated cells. We found that the full-length ZFP809 protein is subject to ubiquitin-mediated degradation in most cells, but is uniquely stable in ES and EC cells. Thus, unlike many targets of ubiquitinylation and degradation in ES cells, ZFP809 is protected from or resistant to such regulatory events in these cells.

A curious feature of the ZFP809 gene is that its transcription results in the formation of two alternatively spliced mRNAs that encode two forms of the protein, a full-length and a C-terminally truncated form. Both mRNAs are expressed in ES and differentiated cells, with the mRNA for the full-length protein being by far the more abundant in both cases. Expression of cDNAs revealed that the full-length protein is the regulated form, selectively stable in ES cells and degraded elsewhere. The C-terminal 50 residues, and lysine391 in this sequence in

particular, are required for degradation. It is unclear why the two forms are made; both forms of the protein retain the KRAB domain for interaction with Trim28, and the zinc fingers required for DNA binding, and the short form seem to have full functionality in silencing. Why low levels of the truncated form would be expressed constitutively is not apparent.

Another surprising aspect of the system is the involvement of Trim28 in the regulation of the levels of ZFP809. The major function of the interaction between ZFP809 and Trim28 is to bring Trim28 and all its associated factors to the target DNA to mediate the silencing through chromatin and DNA modification. The interaction would be expected to result in the stabilization of ZFP809. But in differentiated cells, the interaction is actually involved in the destabilization of ZFP809. The results suggest that Trim28 is playing a major role in keeping ZFP809 levels low in differentiated cells. Trim28 may be acting as the E3 ligase directing the ubiquitinylation of ZFP809, or it may be acting indirectly, but depleting Trim28 or preventing the interaction with ZFP809 significantly reduces the rate of ZFP809 degradation. How its destabilizing activity is restricted to differentiated cells, and is prevented from functioning in ES cells, remains unclear.

The addition of MG132 stabilizes ZFP809 and results in the accumulation of heavily polyubiquitinylated protein, strongly suggesting regulation by ubiquitin-dependent proteasomal degradation. Tests of mutant forms of the protein for ubiquitin content suggest that eliminating the interaction with Trim28 by mutation (in ZFP809(DV-AA)), removing lysine391 by mutation (in K391A), or removing the C-terminus region (in ZFP809(1-353))

all reduce the levels of polyubiquitin many fold, though some of the modification remains. The continued modification of K391A and the truncated ZFP809(1-353) indicates that other lysines are serving as secondary targets for ubiquitinylation when K391 is not present. The truncated protein seems to be largely monoubiquitinylated, and as such would not be recognized by the proteosome for degradation.

Many ring-finger proteins such as Trim28 function as E3 ligases in concert with one of the cullin proteins to target proteins for degradation. The cullins in turn often require modification by addition of Nedd proteins, neddylation, for their functionality (Merlet et al., 2009). The stabilization of ZFP809 by addition of the neddylation inhibitor MLN4924 suggests that a cullin protein is likely involved in Trim28-mediated degradation of ZFP809. It is not known which cullin or cullins are involved.

ZFP809 is not the only protein involved in MLV silencing that is tightly regulated in ES cells, and indeed the levels and activity of Trim28 are also subject to complex regulation. The PHD domain of Trim28 functions as an intramolecular E3 ligase that sumoylates the adjacent Trim28 bromodomain, and this modification is required to recruit the SETDB1 histone methylase to effect transcriptional silencing (Ivanov et al., 2007). Phosphorylation of Trim28 is also required for its function in transcription repression (Li et al., 2007). The active Trim28, once phosphorylated and sumoylated, has a shorter half-life and is targeted for degradation by RNF4 (Kuo et al., 2014). This pattern of behaviors suggests that after silencing a specific chromatin region, Trim28 function is subsequently turned off. The continued

presence of Trim28 may not be required to maintain the silencing of MLV in stem cells. Once silenced, the proviruses cannot be activated during differentiation, because the restriction complex has already placed many silencing marks on the proviral DNA (Yao et al., 2004).

Retroviral DNAs are subject to profound regulation in ES cells and during their differentiation. This may be in part to prevent genotoxic expression of exogenous and endogenous proviruses, but it is also likely a reflection of the fact that viral transcriptional elements have been coopted by the host for the regulation of cellular genes important for maintenance of the pluripotent state and for appropriately timed gene expression during their differentiation (Robbez-Masson and Rowe, 2015; Rowe et al., 2013). While much of this regulation is at the level of transcription, the observations of ZFP809 here suggest that much of this regulation may be controlled at a higher level by protein turnover. A full characterization of "stemness" – the pleuripotent state – will thus likely require a detailed knowledge of the ubiquitin ligase substrates, and more generally the machinery that adds and reads out post-translational protein modifications.

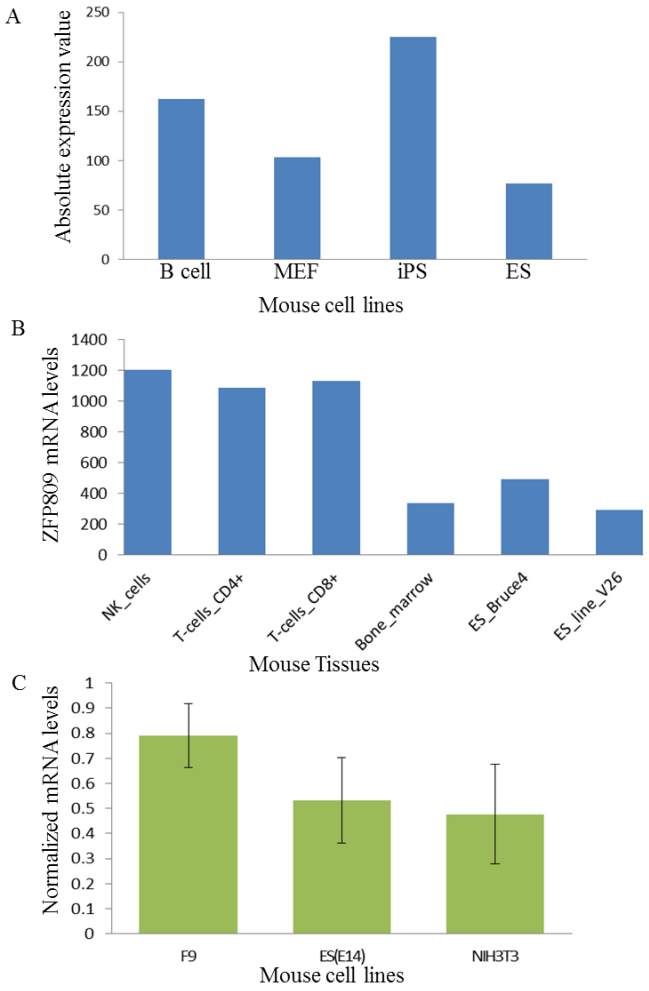

Figure 8 The mRNA expression levels of ZFP809.
A) The cell lines, B cell, MEF (mouse embryonic fibroblast), and ES (embryonic stem) were generated from ROSA26–M2rtTA mice (Beard et al., 2006). IPS (induced pluripotent stem) cells were induced from MEFs. Expression data were generated from polyA RNA using GeneChip Mouse Genome 430 2.0 Arrays (Affymetrix). Absolute values were normalized by Robust Multi-Array (RMA) and processed using Genepattern (Mikkelsen et al., 2008). B) The mRNA expression levels of ZFP809 in selected mouse tissues from BioGPS database. C) Real time-quantitative PCR measure of ZFP809 mRNA in mouse EC (F9), ES (E14) and fibroblast cells (NIH3T3). The expression level of ZFP809 was normalized to GAPDH

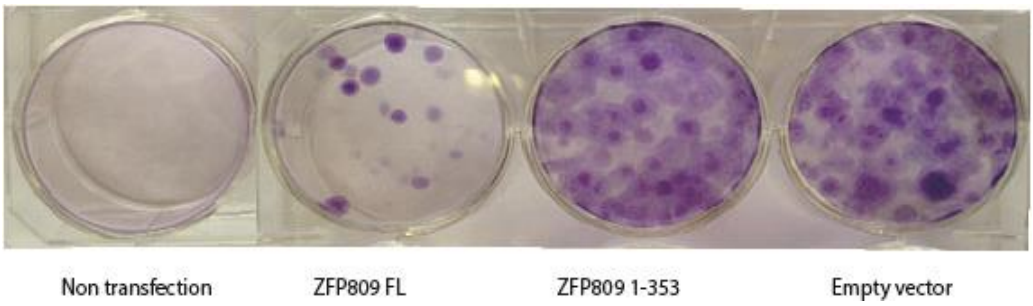

Figure 9 Expressing ZFP809 (FL) results in less colonies.
ZFP809(FL) and ZFP809(1-3530 were cloned to pcDNA3 3XFlag vector. 2ug of empty vector or vectors expressing ZFP809(FL) or ZFP809(1-353) were transfected into NIH3T3 cells. 48 hours after transfection, cells were selected with 1mg/ml G418 for 2 weeks.

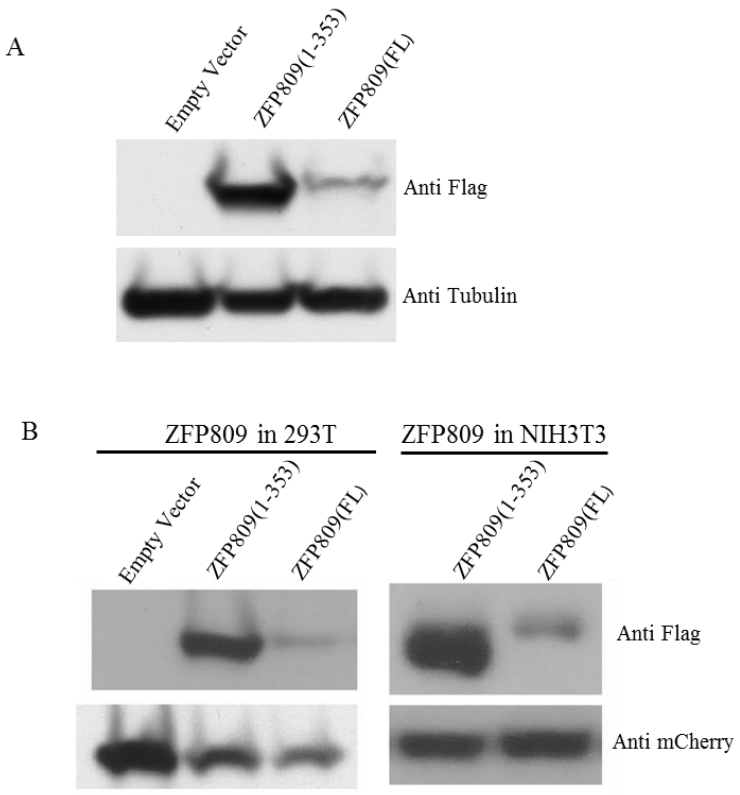

Figure 10 Expression of ZFP809 proteins in differentiated cells. A) 2 ug vectors of pcDNA3 3X Flag empty vector, ZFP809(1-353) or ZFP809(FL) were transfected into 293T cells. 24 hours after transfection, cells were lysed and subject to immuoblotting with anti-Flag antibody. B) 2ug of vector of pLVX 3XFlag ZFP809(FL) or ZFP809(1-353) IRES mCherry were transfected into 293T or NIH3T3 cells. Samples were analyzed by immunoblotting with anti-Flag and anti-mCherry protein.

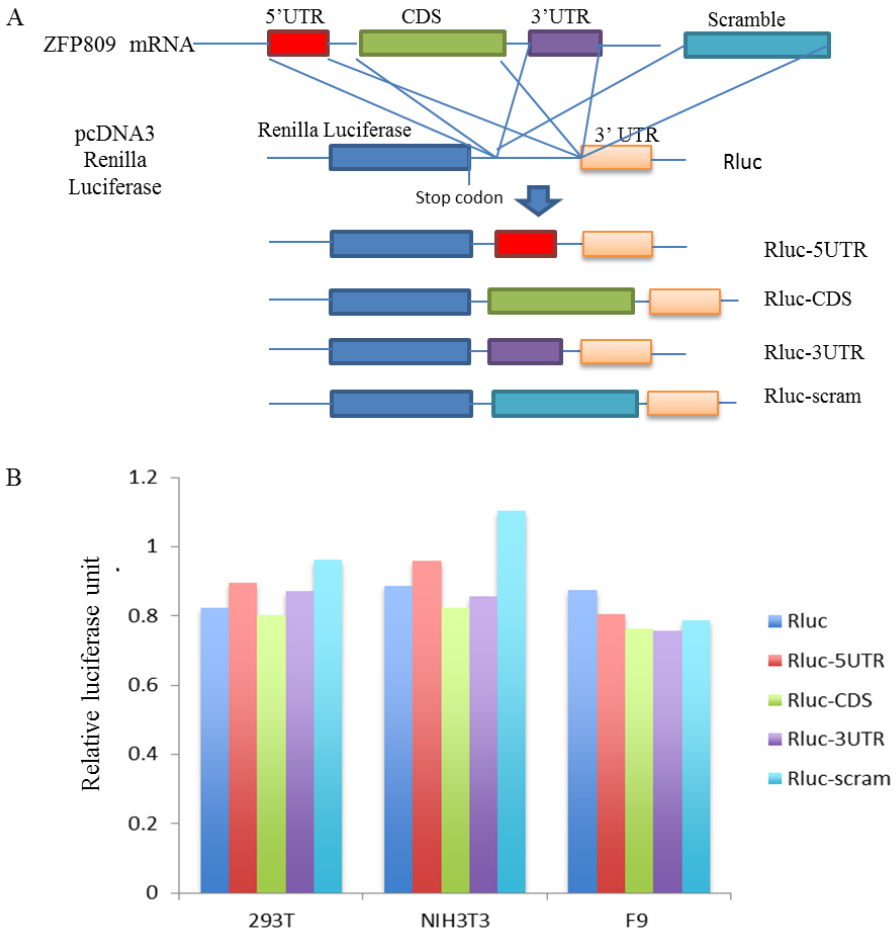

Figure 11 Elements of ZFP809 mRNA do not regulate luciferase expression. A) The 3'UTR, CDS, 5'UTR or a scramble sequence were cloned into pcDNA3 Renilla luciferase vector after the Renilla luciferase stop codon. B) Vectors were transfected into cells with a control of Firefly luciferase. Renilla luciferase units are normalized to firefly luciferase units to calculate the relative luciferase unit.

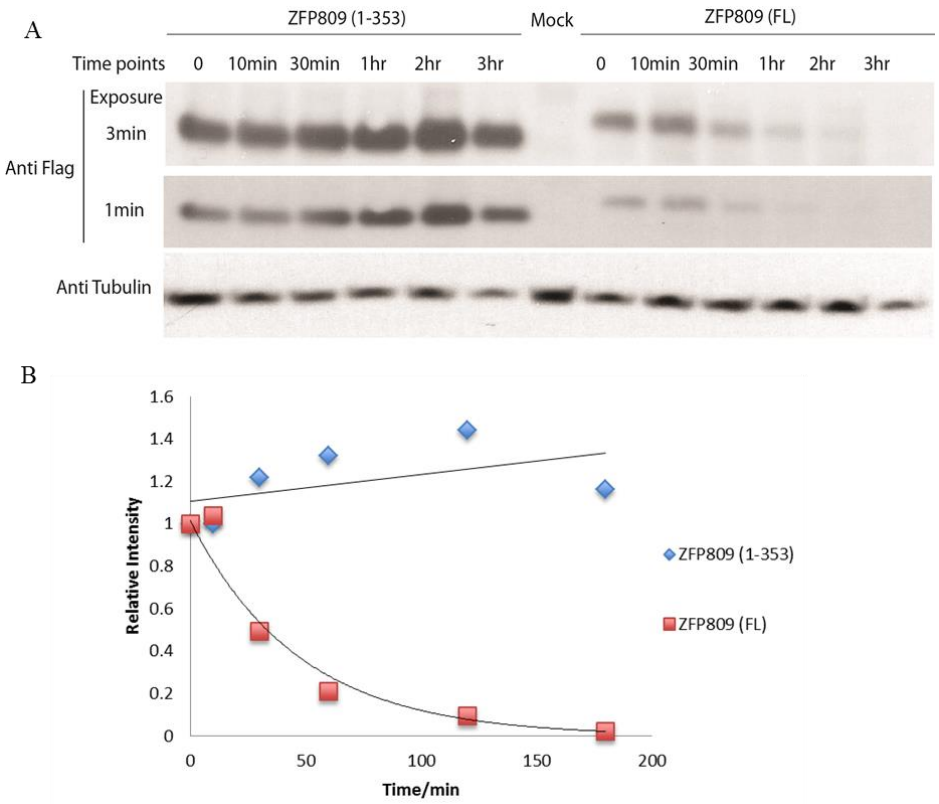

Figure 12 ZFP809(FL) is rapidly degraded in 293T cells.
A) Protein degradation assay. Plasmids expressing flag-tagged ZFP809(FL) or ZFP809(1-353) were transfected into 293T cells. 24 hours after transfection, CHX were added with a final concentration 50ug/ml to stop translation. Cell lysates were collected after 0, 10min, 30min, 1h, 2h and 3h after CHX addition and probed by western blotting. B) The intensity of each band in figure 12A was measured by ImageJ. Relative intensity was calculated by the intensity of each band dividing the intensity of band at time 0. The degradation curve of ZFP809(FL) was calculated by exponential regression.

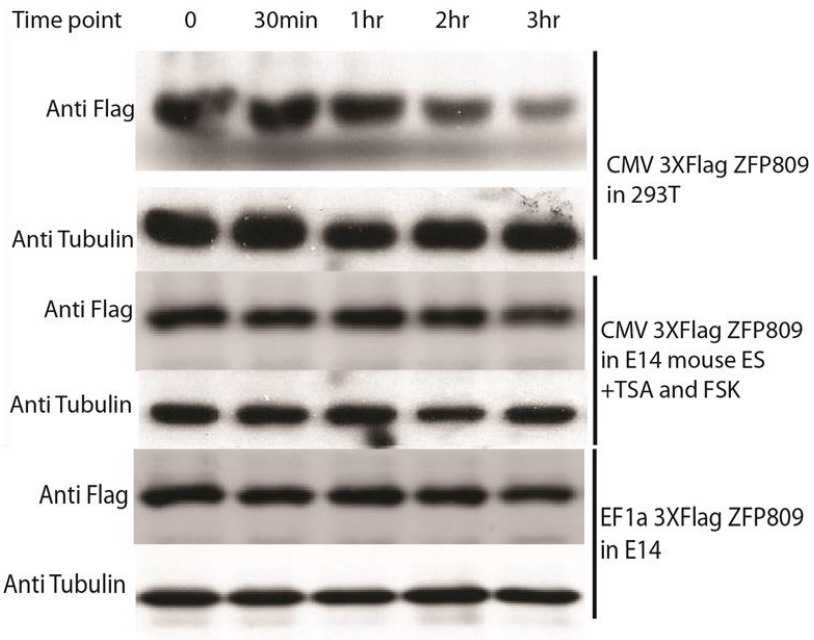

Figure 13 ZFP809 is stable in mouse stem cells.
Protein degradation assay was performed in E14 cells. 24 hours after transfecting pcDNA3 CMV 3XFlag ZFP809(FL) into ES cells, cells were treated in 500nM of Trichostatin A (TSA) and 0.1mM of forskolin (FSK) for 4 hours before CHX was added. ES cells transfected with pLVX EF1α 3XFlag ZFP809(FL) were not treated with TSA and FSK. Cell lysates were collected at the indicated time points after CHX addition and subjected to immuno-blotting.

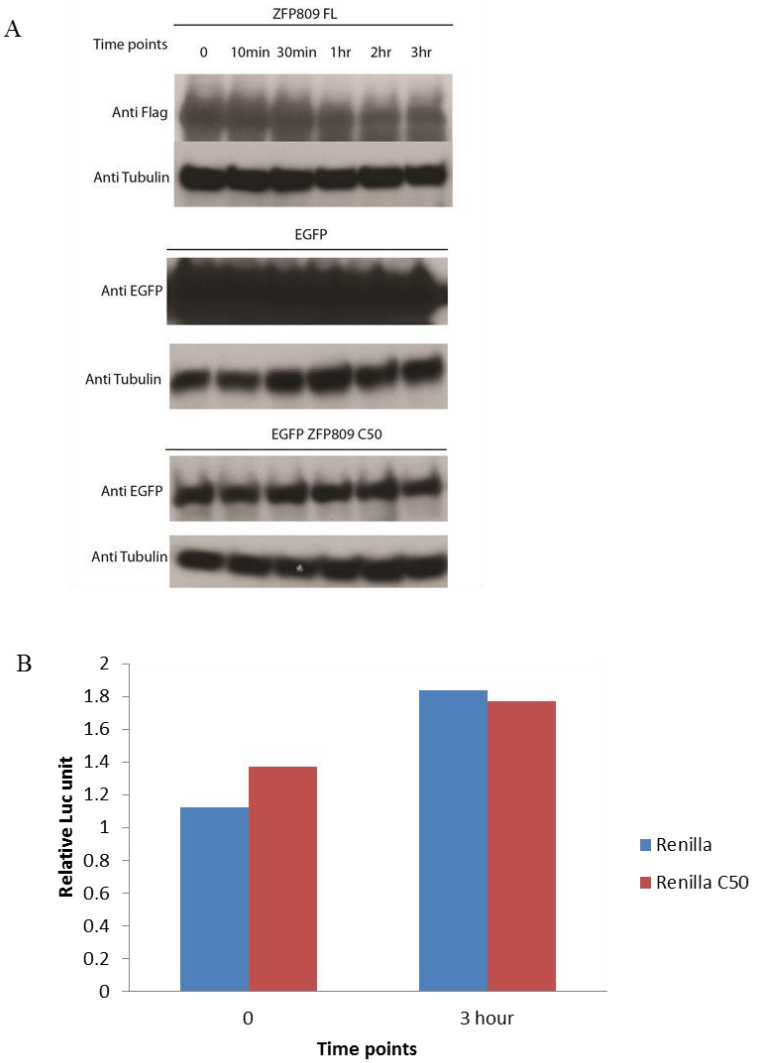

Figure 14 C terminus of ZFP809 is not sufficient to induece turnover of other proteins.
A) C terminus of ZFP809 was appended to EGFP to make EGFP-C50. DNAs expressing ZFP809(FL), EGFP and EGFP-C50 were transfected into 293T cells for protein degradation assay. B) Renilla luciferase was tagged with C terminus of ZFP809 to make Renilla C50. Renilla luciferase and Renilla C50 were transfected into 293T cells with Firefly luciferase as contol. Cycloheximide was added to stop translation for 3 hour. Cells were lysed and tested for luciferase units. Renilla luciferase units were normalized to firefly luciferase units.

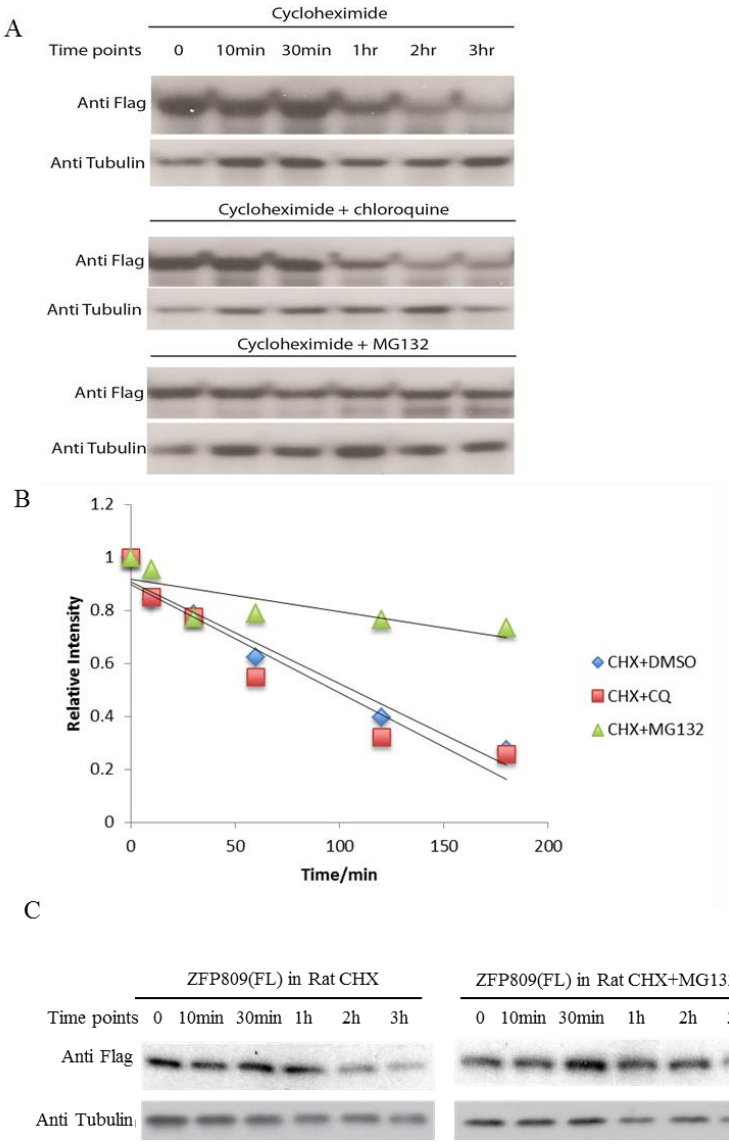

Figure 15 MG132 stabilized ZFP809(FL) in 293T and Rat2 cells. A) MG132(30uM), chloroquine(100uM) or DMSO was added with CHX to perform protein degradation assay . Cells lysates were collected at indicated time points and subjected to immuno-blotting with anti-Flag and anti-Tubulin antibodies. B) The intensity of each band in figure 15A was measured by ImageJ. Relative intensity was calculated by the intensity of each band dividing the intensity of band at time 0. C) Protein degradation assays were performed in Rat2 cells with MG132.

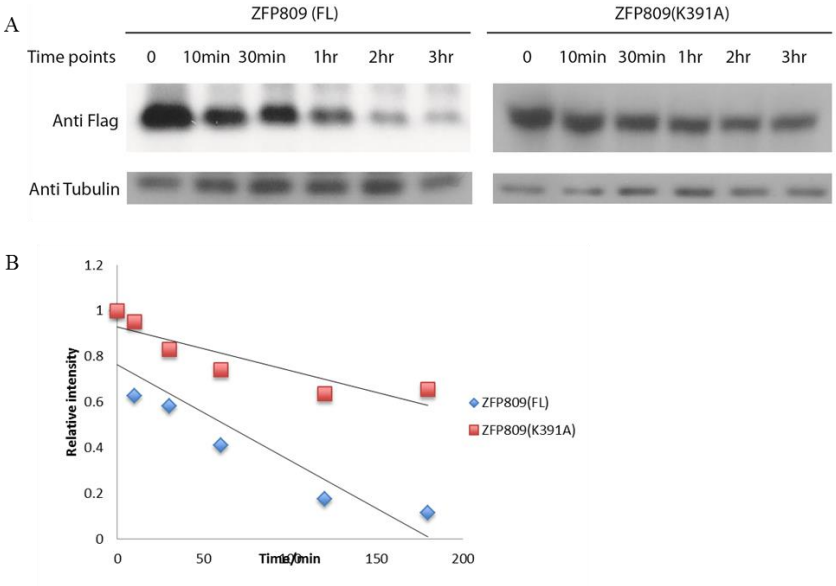

Figure 16 Mutating of lysine 391 to alanine stabilized ZFP809. A) DNAs expressing flag-tagged ZFP809(FL) and ZFP809(K391A) were transfected into 293T cells to perform protein degradation assay. B) The intensity of each band in figure 16A was measured by ImageJ. Relative intensity was calculated by the intensity of each band dividing the intensity of band at time point 0.

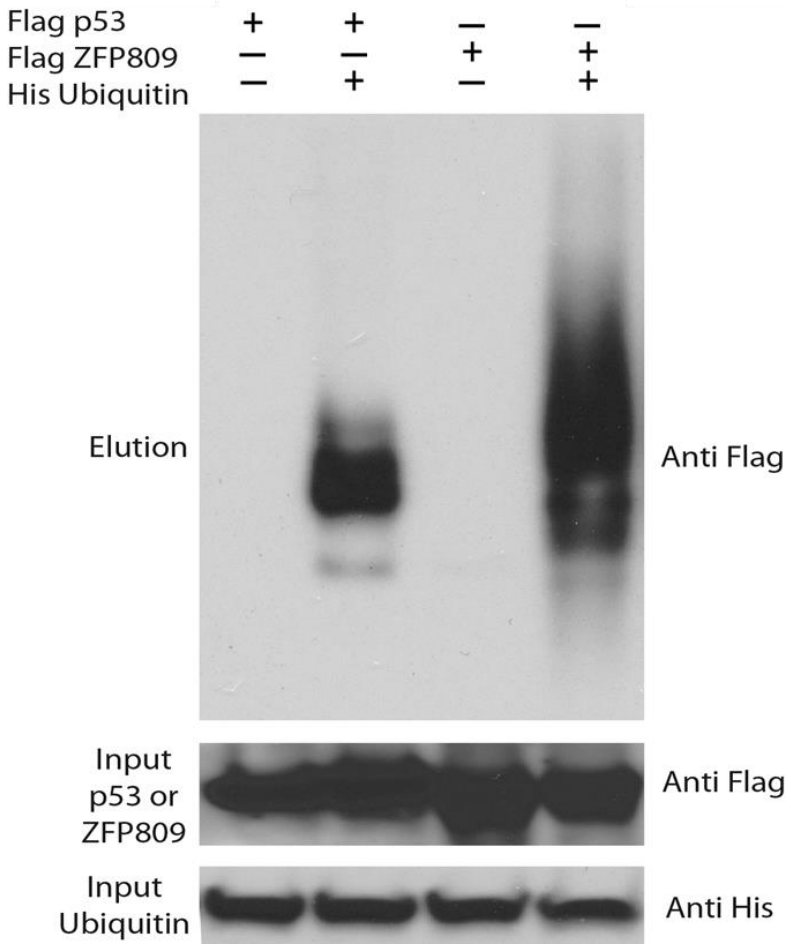

Figure 17 ZFP809 is heavily ubiquitinated.
Vectors expressing flag-tagged p53 or ZFP809(FL) were co-transfected into 293T cells with plasmid express his-ubiquitin. 24 hours after transfection, MG132 was added to stop the protein degradation for 4 hours and cells were lysed in buffer A and sonicated. Cell lysates was incubated with Ni-NTA agarose for 4 hours. Then agarose was washed with buffer A, buffer B and buffer C. Bound proteins were subjected to immuno-blotting. 10ul of p53 elution and 5ul of ZFP809 elution were loaded for SDSPAGE.

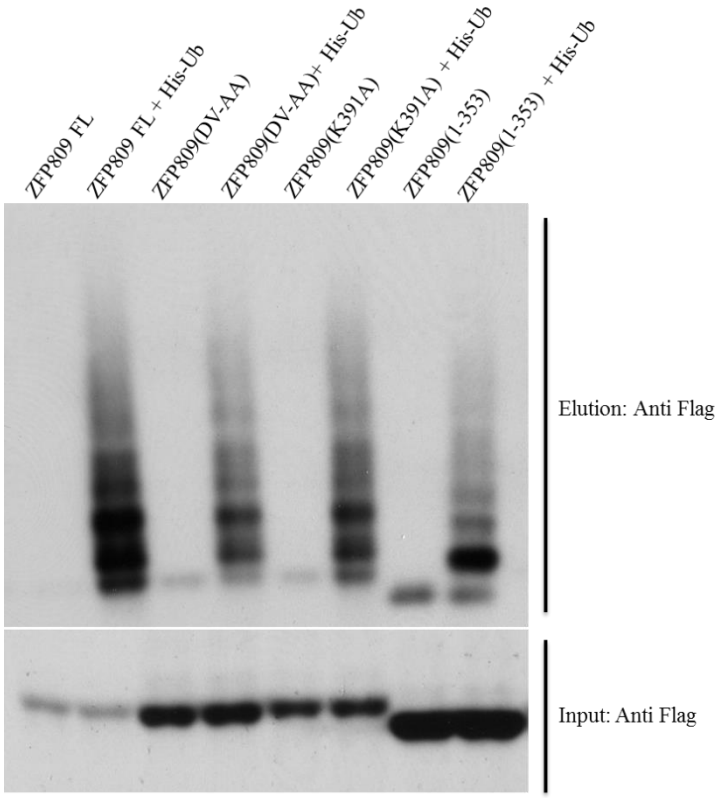

Figure 18 ZFP809(1-353) , ZFP809(K391A) and ZFP809(DV-AA) are ubiquitinated in 293T cells.
In vivo ubiquitination assay was performed as described in Figure 17 with DNAs expressing flag-tagged ZFP809(FL), ZFP809(1-353) , ZFP809(K391A) or ZFP809(DV-AA).

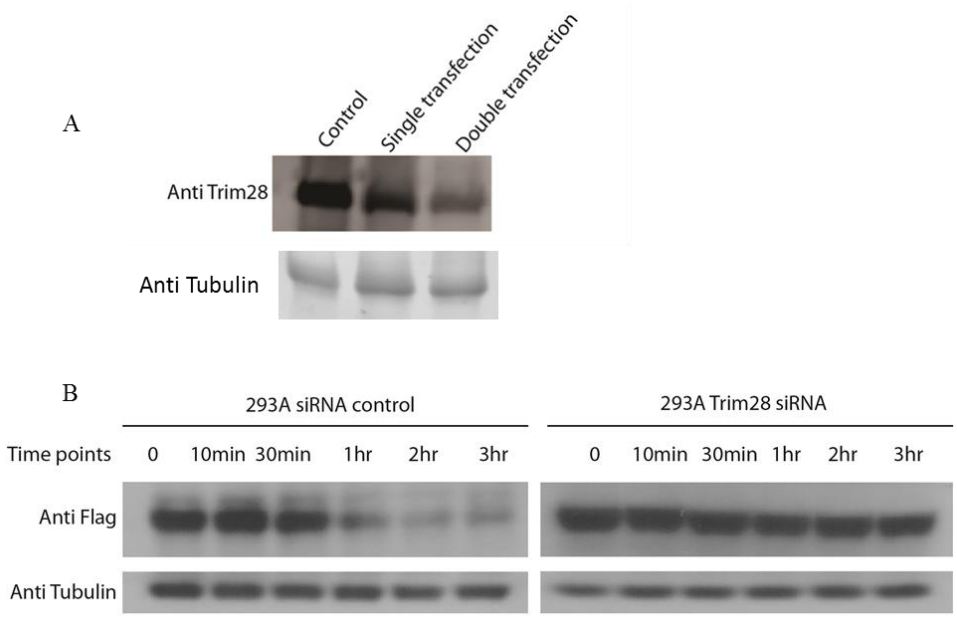

Figure 19 Trim28 promotes ZFP809 degradation.
A) Trim28 was depleted by siRNA. 293A cells were transfected with control siRNA and Trim28 siRNA twice. Cells were analyzed by immunoblotting with anti-Trim28 antibody. B) ZFP809(FL) was transfected into Trim28 depleted or control cells for protein degradation assays.

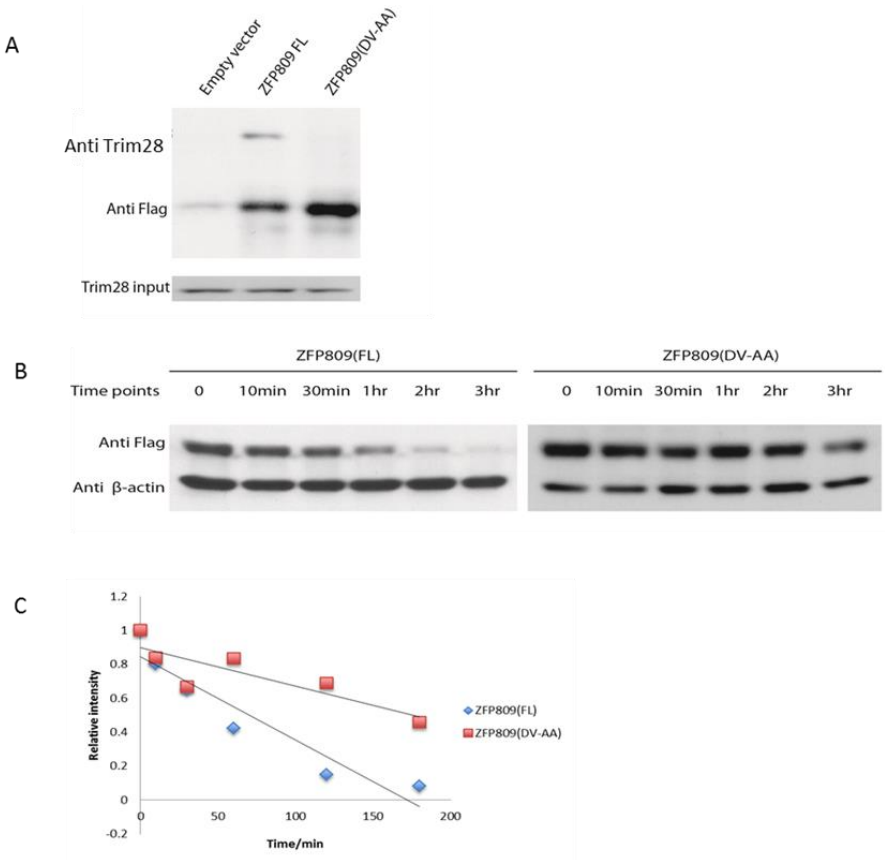

Figure 20 ZFP809(DV to AA) is relatively stable in 293T cells. A) Vectors expressing flag-tagged ZFP809(FL), or ZFP809(DV-AA) or empty vector were transfected into 293T cells. 24 hours after transfection, cells were lysed in IPH buffer. Cell lysates were incubated with anti-Flag magnetic beads overnight. Bound proteins were probed by western blotting. B) Protein degradation assay with DNAs expressing flag-tagged ZFP809(FL) and ZFP809(DV-AA). C) The intensity of each band in figure 20B was measured by ImageJ. Relative intensity was calculated by the intensity of each band dividing the intensity of band at time point 0.

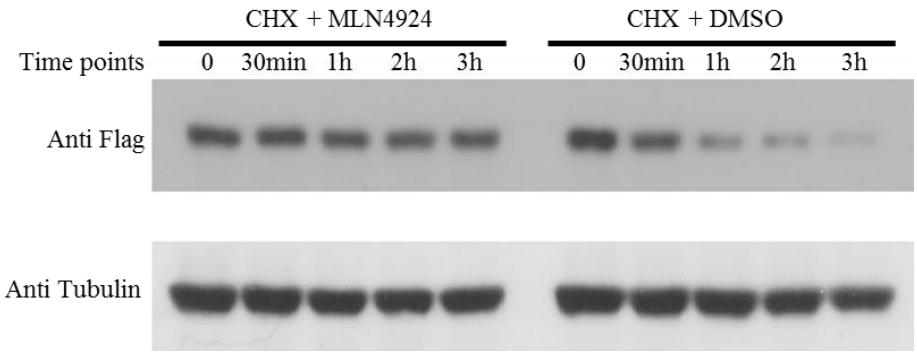

Figure 21 Neddylation inhibitor MLN4924 stabilized ZFP809(FL). MLN4924 (1uM) or DMSO was added with CHX. Cell lysates were collected at indicated time points and subjected to immuno-blotting.

2.6 References

Akinc, A., Thomas, M., Klibanov, A.M., and Langer, R. (2005). Exploring polyethylenimine-mediated DNA transfection and the proton sponge hypothesis. The journal of gene medicine *7*, 657-663.

Asher, G., Reuven, N., and Shaul, Y. (2006). 20S proteasomes and protein degradation "by default". BioEssays : news and reviews in molecular, cellular and developmental biology *28*, 844-849.

Barklis, E., Mulligan, R.C., and Jaenisch, R. (1986). Chromosomal position or virus mutation permits retrovirus expression in embryonal carcinoma cells. Cell *47*, 391-399.

Barry, M., and Fruh, K. (2006). Viral modulators of cullin RING ubiquitin ligases: culling the host defense. Science's STKE : signal transduction knowledge environment *2006*, pe21.

Beard, C., Hochedlinger, K., Plath, K., Wutz, A., and Jaenisch, R. (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis *44*, 23-28.

Bennett, E.J., Rush, J., Gygi, S.P., and Harper, J.W. (2010). Dynamics of cullin-RING ubiquitin ligase network revealed by systematic quantitative proteomics. Cell *143*, 951-965.

Buckley, S.M., Aranda-Orgilles, B., Strikoudis, A., Apostolou, E., Loizou, E., Moran-Crusio, K., Farnsworth, C.L., Koller, A.A., Dasgupta, R., Silva, J.C., *et al.* (2012). Regulation of pluripotency and cellular reprogramming by the ubiquitin-proteasome system. Cell stem cell *11*, 783-798.

Chen, K., and Rajewsky, N. (2007). The evolution of gene regulation by transcription factors and microRNAs. Nature reviews Genetics *8*, 93-103.

Chung, S., Andersson, T., Sonntag, K.C., Bjorklund, L., Isacson, O., and Kim, K.S. (2002). Analysis of different promoter systems for efficient transgene expression in mouse embryonic stem cell lines. Stem cells *20*, 139-145.

Deshaies, R.J., and Joazeiro, C.A. (2009). RING domain E3 ubiquitin ligases. Annual review of biochemistry *78*, 399-434.

Doyle, J.M., Gao, J., Wang, J., Yang, M., and Potts, P.R. (2010). MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases. Molecular cell *39*, 963-974.

Etlinger, J.D., and Goldberg, A.L. (1977). A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. Proceedings of the National Academy of Sciences of the United States of America *74*, 54-58.

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 & development *10*, 2067-2078.

Garcia, K., Blank, J.L., Bouck, D.C., Liu, X.J., Sappal, D.S., Hather, G., Cosmopoulos, K., Thomas, M.P., Kuranda, M., Pickard, M.D., *et al.* (2014). Nedd8-Activating Enzyme Inhibitor MLN4924 Provides Synergy with Mitomycin C through Interactions with ATR, BRCA1/BRCA2 and Chromatin Dynamics Pathways. Molecular cancer therapeutics.

Godbersen, J.C., Humphries, L.A., Danilova, O.V., Kebbekus, P.E., Brown, J.R., Eastman, A., and Danilov, A.V. (2014). The Nedd8-Activating Enzyme Inhibitor MLN4924 Thwarts Microenvironment-Driven NF-kappaB Activation and Induces Apoptosis in Chronic Lymphocytic Leukemia B Cells. Clinical cancer research : an official journal of the American Association for Cancer Research *20*, 1576-1589.

Ha, I., Wightman, B., and Ruvkun, G. (1996). A bulged lin-4/lin-14 RNA duplex is sufficient for Caenorhabditis elegans lin-14 temporal gradient formation. Genes & development *10*, 3041-3050.

Haas, A.L., Warms, J.V., Hershko, A., and Rose, I.A. (1982). Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. The Journal of biological chemistry *257*, 2543-2548.

Higuchi, R., Krummel, B., and Saiki, R.K. (1988). A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic acids research *16*, 7351-7367.

Hochstrasser, M. (1996). Ubiquitin-dependent protein degradation. Annual review of genetics *30*, 405-439.

Hong, S., Hwang, D.Y., Yoon, S., Isacson, O., Ramezani, A., Hawley, R.G., and Kim, K.S. (2007). Functional analysis of various promoters in lentiviral vectors at different stages of in vitro differentiation of mouse embryonic stem cells. Molecular therapy : the journal of the American Society of Gene Therapy *15*, 1630-1639.

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

Iyengar, S., and Farnham, P.J. (2011). KAP1 protein: an enigmatic master regulator of the genome. The Journal of biological chemistry *286*, 26267-26276.

Keller, M.J., Wu, A.W., Andrews, J.I., McGonagill, P.W., Tibesar, E.E., and Meier, J.L. (2007). Reversal of human cytomegalovirus major immediate-early enhancer/promoter silencing in quiescently infected cells via the cyclic AMP signaling pathway. Journal of virology *81*, 6669-6681.

Knop, M., Schiffer, H.H., Rupp, S., and Wolf, D.H. (1993). Vacuolar/lysosomal proteolysis: proteases, substrates, mechanisms. Current opinion in cell biology *5*, 990-996.

Kuo, C.Y., Li, X., Kong, X.Q., Luo, C., Chang, C.C., Chung, Y., Shih, H.M., Li, K.K., and Ann, D.K. (2014). An Arginine-rich Motif of Ring Finger Protein 4 (RNF4) Oversees the Recruitment and Degradation of the Phosphorylated and SUMOylated Kruppel-associated Box Domain-associated Protein 1 (KAP1)/TRIM28 Protein during Genotoxic Stress. The Journal of biological chemistry *289*, 20757-20772.

Le Douarin, B., Nielsen, A.L., Garnier, J.M., Ichinose, H., Jeanmougin, F., Losson, R., and Chambon, P. (1996). A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. The EMBO journal *15*, 6701-6715.

Li, M.Z., and Elledge, S.J. (2012). SLIC: a method for sequence- and ligation-independent cloning. Methods in molecular biology *852*, 51-59.

Li, X., Lee, Y.K., Jeng, J.C., Yen, Y., Schultz, D.C., Shih, H.M., and Ann, D.K. (2007). Role for KAP1 serine 824 phosphorylation and sumoylation/desumoylation switch in regulating KAP1-mediated transcriptional repression. The Journal of biological chemistry *282*, 36177-36189.

Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., Huang, C.C., and Kain, S.R. (1998). Generation of destabilized green fluorescent protein as a transcription reporter. The Journal of biological chemistry *273*, 34970-34975.

Liang, Q., Deng, H., Li, X., Wu, X., Tang, Q., Chang, T.H., Peng, H., Rauscher, F.J., 3rd, Ozato, K., and Zhu, F. (2011). Tripartite motif-containing protein 28 is a small ubiquitin-related modifier E3 ligase and negative regulator of IFN regulatory factor 7. Journal of immunology *187*, 4754-4763.

Maki, C.G., Huibregtse, J.M., and Howley, P.M. (1996). In vivo ubiquitination and proteasome-mediated degradation of p53(1). Cancer research *56*, 2649-2654.

Malim, M.H. (2014). HIV: Ringside views. Nature *505*, 167-168.

Medina, D., Moskowitz, N., Khan, S., Christopher, S., and Germino, J. (2000). Rapid purification of protein complexes from mammalian cells. Nucleic acids research *28*, E61.

Merlet, J., Burger, J., Gomes, J.E., and Pintard, L. (2009). Regulation of cullin-RING E3 ubiquitin-ligases by neddylation and dimerization. Cellular and molecular life sciences : CMLS *66*, 1924-1938.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49-55.

Naujokat, C., and Saric, T. (2007). Concise review: role and function of the ubiquitin-proteasome system in mammalian stem and progenitor cells. Stem cells *25*, 2408-2418.

Patnaik, A., Chau, V., and Wills, J.W. (2000). Ubiquitin is part of the retrovirus budding machinery. Proceedings of the National Academy of Sciences of the United States of America *97*, 13069-13074.

Peng, H., Begg, G.E., Harper, S.L., Friedman, J.R., Speicher, D.W., and Rauscher, F.J., 3rd (2000). Biochemical analysis of the Kruppel-associated box (KRAB) transcriptional repression domain. The Journal of biological chemistry *275*, 18000-18010.

Peng, H., Gibson, L.C., Capili, A.D., Borden, K.L., Osborne, M.J., Harper, S.L., Speicher, D.W., Zhao, K., Marmorstein, R., Rock, T.A., *et al.* (2007). The structurally disordered KRAB repression domain is incorporated into a protease resistant core upon binding to KAP-1-RBCC domain. Journal of molecular biology *370*, 269-289.

Peters, J.M., Franke, W.W., and Kleinschmidt, J.A. (1994). Distinct 19 S and 20 S subcomplexes of the 26 S proteasome and their distribution in the nucleus and the cytoplasm. The Journal of biological chemistry *269*, 7709-7718.

Petersen, R., Kempler, G., and Barklis, E. (1991). A stem cell-specific silencer in the primer-binding site of a retrovirus. Molecular and cellular biology *11*, 1214-1221.

Rabut, G., and Peter, M. (2008). Function and regulation of protein neddylation. 'Protein modifications: beyond the usual suspects' review series. EMBO reports *9*, 969-976.

Rape, M., and Jentsch, S. (2002). Taking a bite: proteasomal protein processing. Nature cell biology *4*, E113-116.

Risseeuw, E.P., Daskalchuk, T.E., Banks, T.W., Liu, E., Cotelesage, J., Hellmann, H., Estelle, M., Somers, D.E., and Crosby, W.L. (2003). Protein interaction analysis of SCF ubiquitin E3 ligase subunits from Arabidopsis. The Plant journal : for cell and molecular biology *34*, 753-767.

Robbez-Masson, L., and Rowe, H.M. (2015). Retrotransposons shape species-specific embryonic stem cell gene expression. Retrovirology *12*, 45.

Rock, K.L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., and Goldberg, A.L. (1994). Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. Cell *78*, 761-771.

Rowe, H.M., Kapopoulou, A., Corsinotti, A., Fasching, L., Macfarlan, T.S., Tarabay, Y., Viville, S., Jakobsson, J., Pfaff, S.L., and Trono, D. (2013). TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells. Genome research *23*, 452-461.

Sarikas, A., Hartmann, T., and Pan, Z.Q. (2011). The cullin protein family. Genome biology *12*, 220.

Saurin, A.J., Borden, K.L., Boddy, M.N., and Freemont, P.S. (1996). Does this have a familiar RING? Trends in biochemical sciences *21*, 208-214.

Schneider-Poetsch, T., Ju, J., Eyler, D.E., Dang, Y., Bhat, S., Merrick, W.C., Green, R., Shen, B., and Liu, J.O. (2010). Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. Nature chemical biology *6*, 209-217.

Schubert, U., Ott, D.E., Chertova, E.N., Welker, R., Tessmer, U., Princiotta, M.F., Bennink, J.R., Krausslich, H.G., and Yewdell, J.W. (2000). Proteasome inhibition interferes with gag polyprotein processing, release, and maturation of HIV-1 and HIV-2. Proceedings of the National Academy of Sciences of the United States of America *97*, 13057-13062.

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 & development *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 & development *15*, 428-443.

Settembre, C., Fraldi, A., Medina, D.L., and Ballabio, A. (2013). Signals from the lysosome: a control centre for cellular clearance and energy metabolism. Nature reviews Molecular cell biology *14*, 283-296.

Shintani, T., and Klionsky, D.J. (2004). Autophagy in health and disease: a double-edged sword. Science *306*, 990-995.

Spenger, A., Ernst, W., Condreay, J.P., Kost, T.A., and Grabherr, R. (2004). Influence of promoter choice and trichostatin A treatment on expression of baculovirus delivered genes in mammalian cells. Protein expression and purification *38*, 17-23.

Sproul, D., Gilbert, N., and Bickmore, W.A. (2005). The role of chromatin structure in regulating the expression of clustered genes. Nature reviews Genetics *6*, 775-781.

Strack, B., Calistri, A., Accola, M.A., Palu, G., and Gottlinger, H.G. (2000). A role for ubiquitin ligase recruitment in retrovirus release. Proceedings of the National Academy of Sciences of the United States of America *97*, 13063-13068.

Sutovsky, P., Motlik, J., Neuber, E., Pavlok, A., Schatten, G., Palecek, J., Hyttel, P., Adebayo, O.T., Adwan, K., Alberio, R., *et al.* (2001). Accumulation of the proteolytic marker peptide ubiquitin in the trophoblast of mammalian blastocysts. Cloning and stem cells *3*, 157-161.

Teich, N.M., Weiss, R.A., Martin, G.R., and Lowy, D.R. (1977). Virus infection of murine teratocarcinoma stem cell lines. Cell *12*, 973-982.

Thrower, J.S., Hoffman, L., Rechsteiner, M., and Pickart, C.M. (2000). Recognition of the polyubiquitin proteolytic signal. The EMBO journal *19*, 94-102.

van der Stoop, P., Boutsma, E.A., Hulsman, D., Noback, S., Heimerikx, M., Kerkhoven, R.M., Voncken, J.W., Wessels, L.F., and van Lohuizen, M. (2008). Ubiquitin E3 ligase Ring1b/Rnf2 of polycomb repressive complex 1 contributes to stable maintenance of mouse embryonic stem cells. PloS one *3*, e2235.

Vanhaecke, T., Papeleu, P., Elaut, G., and Rogiers, V. (2004). Trichostatin A-like hydroxamate histone deacetylase inhibitors as therapeutic agents: toxicological point of view. Current medicinal chemistry *11*, 1629-1643.

Vogel, C., and Marcotte, E.M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nature reviews Genetics *13*, 227-232.

Wang, C., Ivanov, A., Chen, L., Fredericks, W.J., Seto, E., Rauscher, F.J., 3rd, and Chen, J. (2005). MDM2 interaction with nuclear corepressor KAP1 contributes to p53 inactivation. The EMBO journal *24*, 3279-3290.

Wilkinson, G.W., and Akrigg, A. (1992). Constitutive and enhanced expression from the CMV major IE promoter in a defective adenovirus vector. Nucleic acids research *20*, 2233-2239.

Wolf, D., Cammas, F., Losson, R., and Goff, S.P. (2008). Primer binding site-dependent restriction of murine leukemia virus requires HP1 binding by TRIM28. Journal of virology *82*, 4675-4679.

Wolf, D., and Goff, S.P. (2007). TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. Cell *131*, 46-57.

Wolf, D., and Goff, S.P. (2008). Host restriction factors blocking retroviral replication. Annual review of genetics *42*, 143-163.

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

Wolf, G., Yang, P., Fuchtbauer, A.C., Fuchtbauer, E.M., Silva, A.M., Park, C., Wu, W., Nielsen, A.L., Pedersen, F.S., and Macfarlan, T.S. (2015). The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses. Genes & development *29*, 538-554.

Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., Hodge, C.L., Haase, J., Janes, J., Huss, J.W., 3rd, *et al.* (2009). BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome biology *10*, R130.

Yamauchi, M., Freitag, B., Khan, C., Berwin, B., and Barklis, E. (1995). Stem cell factor binding to retrovirus primer binding site silencers. Journal of virology *69*, 1142-1149.

Yao, S., Sukonnik, T., Kean, T., Bharadwaj, R.R., Pasceri, P., and Ellis, J. (2004). Retrovirus silencing, variegation, extinction, and memory are controlled by a dynamic interplay of multiple epigenetic modifications. Molecular therapy : the journal of the American Society of Gene Therapy *10*, 27-36.

Young, J.A., Sermwittayawong, D., Kim, H.J., Nandu, S., An, N., Erdjument-Bromage, H., Tempst, P., Coscoy, L., and Winoto, A. (2011). Fas-associated death domain (FADD) and the E3 ubiquitin-protein ligase TRIM21 interact to negatively regulate virus-induced interferon production. The Journal of biological chemistry *286*, 6521-6531.

Yu, J., Zhang, L., Hwang, P.M., Rago, C., Kinzler, K.W., and Vogelstein, B. (1999). Identification and classification of p53-regulated genes. Proceedings of the National Academy of Sciences of the United States of America *96*, 14517-14522.

Zeng, L., Yap, K.L., Ivanov, A.V., Wang, X., Mujtaba, S., Plotnikova, O., Rauscher, F.J., 3rd, and Zhou, M.M. (2008). Structural insights into human KAP1 PHD finger-bromodomain and its role in gene silencing. Nature structural & molecular biology *15*, 626-633.

Chapter 3

Role of L1td1 in retroviral silencing

3.1 Abstract

It has been known for over thirty years that the replication of Moloney Murine Leukemia Virus (MLV or MLV) is restricted in mouse embryonic stem (ES) and embryonic carcinoma (EC) cells (Teich et al., 1977). The restriction is targeted to a short DNA sequence, primer binding site (PBS) of MLV. The restriction requires the association of a zinc finger protein ZFP809 to the PBS(Wolf and Goff, 2009), which recruits Trim28 and other transcription factor to block the transcription of retroviral genes. Here we found that a protein called L1td1/Ecat11 could associate with Trim28. Depletion of L1td1 by shRNA or knockout can partially relieve retroviral restriction occurring at PBS in embryonic carcinoma cells and stem cells.

3.2 Introduction

The replication of Moloney Murine Leukemia virus (MLV or MLV) is efficient in most permissive dividing cells, but is restricted in mouse embryonic stem (ES) and embryonic carcinoma (EC) cells(Barklis et al., 1986). The retroviral DNAs integrate into the host genome successfully after infection, but the transcription of retroviral genes from LTR is potently silenced (Franz et al., 1986; Linney et al., 1984). The presence of the so-called repressor binding site (RBS) is essential for the restriction (Hilberg et al., 1987; Linney et al., 1984), which shows overlap with the primer binding site (PBSpro) of retroviral DNA (Barklis et al., 1986; Feuer et al., 1989; Loh et al., 1987). The RBS alone is sufficient to induce the silencing in a position and orientation- independent manner in mouse ES and EC cells. A

single mutation in the RBS sequence is able to abolish the restriction (Loh et al., 1990; Petersen et al., 1991). On such mutation, termed PBS B2, consists of a single base change near the center of the PBS. Electrophoretic mobility shift assays (EMSA), using the RBS sequence as probe, demonstrated the presence of stem-cell specific factors that fulfill the requirements for the silencing complex (Loh et al., 1990; Petersen et al., 1991). The recruitment of trans-acting repressors to the RBS introduces chromatin modifications, and then initiates and maintains the silencing (Akgun et al., 1991; Flanagan et al., 1989; Niwa et al., 1983; Schlesinger et al., 2014; Tsukiyama et al., 1989). The key components of the trans-acting complex were identified by our lab as ZFP809 and Trim28 (Wolf and Goff, 2007, 2009). ZFP809 binds to the RBS and initiates the silencing of incoming retroviral DNA (Wolf and Goff, 2009). It recruits Trim28 by the KRAB box domain and then other transcription factors (Wolf and Goff, 2009), such as HP1gamma (Wolf et al., 2008) and ESET (Matsui et al., 2010) to silence the retroviral DNAs.

Trim28, also known as TRIM28 (KRAB-associated protein 1), is an unusual member of the Trim (tripartite motif-containing protein) family. Trim28 interacts with many proteins involved with transcriptional repression and silencing, such as the NuRD complex (Rowe et al., 2013), HP1 (Wolf et al., 2008), ESET (a histone H3 lysine 9 methyltransferase)(Le Douarin et al., 1996; Schultz et al., 2002; Schultz et al., 2001), and EBP-1 (Wang et al., 2014).

It is known as a co-repressor of the large family of KRAB (Krüppel-associated box) domain-containing zinc finger transcription factors (Friedman et al., 1996).

F9 nuclear extract was prepared and the activity was precipitated with 40% ammonium sulfate. The recovered precipitate was applied to a butyl sepharose hydrophobic interaction column and a Sephacryl S-500 gel filtration column. The peak elution was coimmunoprecipitated with anti-Trim28 antibody and bound proteins were subject to mass spectrometry analysis. L1td1 was an interesting protein identified with higher abundance than ZFP809 (table 1) (Wolf and Goff, 2009).

L1td1 (LINE1 type transposase domain containing 1), homologous to the ORF1 protein of LINE1, was initially identified as ECAT11 (ES cell-associated transcripts 11), in a digital differential display analysis (DDD) of the expressed sequence tag libraries among various mouse tissues and cell lines (Mitsui et al., 2003). L1TD1 is highly expressed in undifferentiated cells, and could be used as a marker for undifferentiated cells. It is also a downstream target of the differentiation gene Nanog (Wong et al., 2011). L1TD1 localizes in P-bodies and is required for the renew of human embryonic stem cell(Narva et al., 2012). Perhaps surprisingly, L1td1 is not required for the induction and maintenance of pluripotency in mouse (Iwabuchi et al., 2011).

In this work, we found out that L1td1 interacts with Trim28 and ZFP809 in endogenous level. Purified GST-tagged L1td1 and its truncations can also pull down

endogenous L1td1 in F9 cells. In vitro protein-protein interaction assay using MBP-tagged L1td1, GST-tagged ZFP809 and GST-taggedTrim28 showed that L1td1 interacts with Trim28, probably not ZFP809, indicating that the interaction between ZFP809 and L1td1 we observed might be bridged by Trim28. The interactions suggest one mechanism by which L1td1 helps the silencing of retrovirus DNAs in mouse ES and EC cells. Depletion of L1td1by shRNA in F9 cells or knockout in ES cells could partially relieve the restriction of MLV occurring at the PBS.

3.3 Methods

3.3.1 Cell culture

RAT2 and 293T cells were cultured in DMEM with 10% FBS and L-glutamine, penicillin and streptomycin. F9 cells were cultured in DMEM with 10% FBS, L-glutamine, penicillin and streptomycin. Plates were coated with 0.1% gelatin. Mouse L1td1 wildtype and knockout cells were kindly provided by Dr. Yamanaka in Tokyo University (Iwabuchi et al., 2011). Mouse ES cells were cultured on deactivated mouse MEFs with DMEM with 20% defined FBS, nonessential amino acids, 2uM sodium pyruvate, penicillin and streptomycin, and 1000 units of mouse leukemia inhibitory factor (LIF). Cells were transferred to plates coated with 0.1% gelatin when performing experiments. All cells were cultured at 37℃ in 5% CO_2.

3.3.2 Plasmids

The oligonucleotides, vectors and restriction sites for cloning GST-tagged ZFP809, Trim28 and MBP-tagged L1td1 are listed below. Oligonucleotides for shRNA-mediated knockdown of L1td1 were cloned into the pSuper Retro zeo vector using the BglII and XhoI cloning sites according to standard protocols (Oligoengine). The Moloney MLV vectors with neomycin or puromycin reporter were described before (Wolf and Goff, 2007).

Oligonucleotides for cloning pGEX2TPL ZFP809

BamH1 agtc GGATCC ACC ATGGGGTTGGTGTCCTTTGAG

EcoR1 agct AGATCT TCAGAATTCAGCTAGATCTTCAGAATTC

pGEX2TPL ZFP809(1-353)

BamH1 agtc GGATCC ACC ATGGGGTTGGTGTCCTTTGAG

EcoR1 agct GAA TTCTCA ACGCGTAAAGTACGTTACCCCTG

pGEX2TPL ZFP809 60-236

BamH1 atgc GGATCC TTTGGGCCATGGAGCCTAGC

EcoR1 agct GAATTC TTATGAATGGGTTCTTTCATGG

pGEX2TPL ZFP809 151-293

Xba1 atgc TCTAGA GATGGGAATCTCTATGAATG

EcoR1 atgc GAATTC TTAGCCCGTATGTTTCTTCTG

Oligonucleotides for cloningpMAL C2 L1td1

Xba1 GCTTGCCTGCAGGTCGACTCTAGACTAATGTATATTGTTCAACAGATCTTTC

EcoR1 ATCGAGGGAAGGATTTCAGAATTCTCGGGCGTGCAGTCCAA

Oligonucleotides for shRNA

L1td1 AGGAAGAAGTGAACAGAAA

Scrambled GGTGCCGTTTAGTAGTACA

Negative control CAGCAGCCATGTCCTGAAA

3.3.3 Co-immunoprecipitation of endogenous proteins

The preparation of F9 cell lysates was as described previously (Wolf and Goff, 2007). F9 cells were lysed in 750 µl IPH buffer (150 mM NaCl, 50 mM Tris-HCl, pH 8, 0.5% Nonidet P-40, 5 mM MgCl2 1 mM CaCl2 0.5 mM EDTA, + protease inhibitors [Complete, Roche]). The lysates were then treated with 100 U DNase (Roche) and RNase (Life Technologies) for 24 h at 4°C. Coimmunoprecipitations were performed using 5 µg of antibodies and a mixture of 20 µl protein A and 25 µl protein G Dynabeads (Invitrogen). Bound proteins were eluted with Laemmli buffer, displayed by PAGE and blotted, and the Western blots were probed with anti L1td1 and anti TRIM28 antibodies

3.3.4 GST pulldown assay

E.coli BL21 cells, transformed with vectors expressing GST tagged proteins, were cultured at 30^{o}C and induced with IPTG at OD 0.6 for 6 hrs. Cells were sonicated and spun down at 12000 x rpm for 30 min. Supernatants were collected and incubated with glutathione magnetic beads at 4^{o}C for 2 hrs or overnight. The beads, now coated with GST tagged proteins or GST, were then added to F9 EC cell lysates. F9 cells were lysed in IPH buffer (Wolf and Goff, 2009). After incubation, beads were washed and eluted with Laemmli buffer. The eluted proteins were analyzed by Western blotting with anti-GST or anti L1td1 antibodies.

3.3.5 Protein-protein interaction assay

E.coli BL21 cells, transformed with vectors pMAL C2 or pMAL C2 L1td1, were cultured at 30^{o}C and induced with IPTG at OD of 0.6 for 6 hrs. The preparation of MBP-L1td1 supernatant was according to the standard protocol form New England Biolabs. Cells were sonicated and spun down at 12000 x rpm for 30 min. Supernatants were collected and incubated with amylose beads at 4^{o}C for 2hrs or overnight. The beads coated with MBP or MBP-L1td1 were added to supernatant containing GST, GST-ZFP809, GST-Trim28, prepared from bacteria expressing the appropriate GST proteins. Bound proteins were subject to western blotting analysis using anti-GST and anti-MBP antibodies.

3.3.6 Generation of L1td1 stable knockdown F9 cell line

VSV-G-pseudotyped viral particles were generated by transfecting three plasmids —

pMD.G, pCMV Intron and vectors with shRNA oligos or reporter genes – into 293T cells. Supernatants were collected after 48-72 hrs and filtered through a 0.25 um filter. F9 cells were infected by the supernatant and put in 200 uM zeocin selection after 48 hrs for 7 days to generate the L1td1 knockdown cell lines. The knockdown was confirmed by western blotting with antiL1td1 antibody generated by Dr. Daniel Wolf in our lab.

3.3.7 Real time PCR

RNA was harvested from L1td1 wildtype and knockout ES cells utilizing TRIzol (invitrogen). cDNA was generated from RNA using SuperScript III Reverse Transcriptase (Invitrogen) and L1td1 and other gene levels were then quantified by Q-PCR (FastStart Universal SYBR Green Roche). Primers for Q-PCR are listed here (Left 5', right 3').

L1td1 forward AAACGAGCTTCGGGAAAGAT

L1td1 reverse GGAAGCTCAAACTGTCACTCG

GAPDH forward ACCTTTAGCCTTGCCCTTT

GAPDH reverse ACATCACCCCCATCACTCAT

Oct-4 forward GGCGTTCTCTTTGGAAAGGTGTTC

Oct-4 reverse CTCGAACCACATCCTTCTCT

Nanog forward AACCAAAGGATGAAGTGCAAGCGG

Nanog reverse TCCAAGTTGGGTTGGTCCAAGTCT

Endogenous MLV forward TAGATGGAGCCTACCAAGCTCTCAA

Endogenous MLV reverse AGAGGTATGGTTGGAATAAGTA

IAP forward AAGCAGCAATCACCCACTTTGG

IAP reverse CAATCATTAGATGYGGCTGCCAAG

Line forward AGAGGTATGGTTGGAATAAGTA

Line reverse CCATGTTTAGCGCTTCCTTC

MusD forward GTGCTAACCCAACGCTGGTTC

MusD reverse CTCTGGCCTGAAACAACTCCTG

3.3.8 Genomic PCR

Mouse ES genomic DNA was prepared from the protocol in (Truett et al., 2000). The genotyping PCR primers are from (Iwabuchi et al., 2011). The PCR cycle is 95^{o}C for 1min, 65oC for 30s, 72^{o}C for 1min with 35 repeats.

3.3.9 Viral transduction assay

VSVG pseudotyped viral particles were generated by transfecting three plasmid—pMD.G, pCMV Intron and vectors with neomycin or puromycin-resistance reporter genes – into 293T cells. L1td1 knockout or knockdown cells were infected by VSV-G pseudotyped MLV viral particles with neomycin or puromycin reporter. 48hrs after infection, cells were selected in 1mg/ml neomycin for 10-14 days or 1ug/ml puromycin for 7 days. Colonies were fixed and counted.

3.4 Results

3.4.1 L1td1 is homologous to ORF-1 of L1

L1td1 (LINE1 type transposase domain containing 1), homologous to the ORF1 protein of LINE1, was identified as ECAT11 (ES cell-associated transcripts 11), in a digital differential display analysis (DDD) of the expressed sequence tag libraries among various mouse tissues and cell lines (Mitsui et al., 2003). L1TD1 protein is enriched in undifferentiated cells, and could be used as a marker for undifferentiated cells. The predicted open reading frame of L1td1 shows a similarity to ORF-1 of Line1, a very abundant non-LTR retrotransposon in human and mouse genomes (Doucet et al., 2010; Singer, 1982). Sequence alignments of L1td1 to the NCBI database revealed the presence of two transposase 22 domains, which also have identity to the ORF-1 protein (Figure 22). The transposease 22 domain of ORF-1 is responsible for RNA binding (Martin, 2006), suggesting that L1td1 might have RNA binding ability. Homologs of L1td1 can be identified in human, rat and other species, but the overall sequence similarity is low (Figure 22).

3.4.2 L1td1 interacts with both Trim28 and ZFP809

The original identification of ZFP809 as the PBS-specific DNA binding activity involved several purification steps that enriched for the silencing complex. L1td1 was identified along with ZFP809 in the complexes immunoprecipitated with Trim28 specific antibody (Wolf and Goff, 2009), which indicates L1td1 may interact with Trim28 or ZFP809. To confirm the authenticity of the interactions of L1td1 with Trim28 or ZFP809, lysates from

an EC cell line F9, were subjected to immunoprecipitation with either anti-L1td1 antibody or anti-ZFP809 antibody. The bound proteins were resolved by SDS-PAGE and analyzed by immunoblotting with anti-Trim28 or anti-L1td1 antibodies. Endogenous Trim28 was detected in both anti-L1td1 and anti-ZFP809 immunoprecipitates but not the IgG control (Figure 23). This observation shows that Trim28 and L1td1 interact at their endogenous levels of expression. As previously reported, ZFP809 was able to bind Trim28. A weak band of L1td1 was detected in the anti-ZFP809 immunoprecipitates, which indicates that L1td1 might also interact with ZFP809. All of the interactions are resistant to depletion of nucleic acids by adding nucleases (Figure 23), suggesting that the interactions of L1td1-Trim28 and ZFP809 - Trim28 are not mediated by nucleic acids. Due to the poor quality of anti-ZFP809 antibody, and to the similarity of ZFP809 and IgG molecular weights, ZFP809 was not detected by immunoblotting (Figure 23).

To verify the interaction between ZFP809 and L1td1, ZFP809 truncations tagged with GST protein were made (Figure 24A). Glutathione beads coated with bacterially-expressed GST-tagged ZFP809 and its truncation derivatives were incubated with F9 cell lysates. The bound proteins were subjected to immunoblotting analysis. L1td1 was detected in all GST-ZFP809 precipitates but not the GST control, indicating that ZFP809 can interact with endogenous L1td1 (Figure 24B). The observation also indicates that the region from 151 to 236 of ZFP809 is important for the interaction of L1td1 and ZFP809. Overall, the observations suggest that L1td1 interacts with both Trim28 and ZFP809.

These precipitation assays showed that L1td1 interacts with Trim28 and ZFP809, but it is not sufficient to conclude that L1td1 interacts with them directly since ZFP809 binds to Trim28 endogenously. To test if L1td1 binds to Trim28 and ZFP809 directly, bacteria expression vectors of MBP-tagged L1td1 and GST-tagged Trim28 and ZFP809 were made. MBP-L1td1 expressed in bacteria was bound to amylose beads and incubated with supernatants containing GST-tagged Trim28 or ZFP809. Bound proteins were subjected to immuno-blotting analysis using anti-GST antibody. The blotting showed that MBP-L1td1 could pull down GST-Trim28 but not GST-ZFP809. MBP and GST tags didn't pulldown other proteins or each other (Figure 25). The observation indicates that L1td1 interact with Trim28 directly. The interaction between ZFP809 and L1td1 we observed before might be bridged by Trim28.

3.4.3 Depletion of L1td1 partially relieves the MLV restriction in EC and ES cells

To investigate the functional relevance of the Trim28-ZFP809-L1td1 for PBS-mediated silencing, we used short hairpin RNA (shRNA, a modified siRNA system) to knock down the L1td1 levels in F9 cells. Successful depletion of L1td1 was achieved (Figure 26A). As negative control, cell lines expressing scrambled or negative shRNA were generated. To quantify the level of PBS-mediated repression, we made use of transducing MLV reporter viral particles expressing neomycin resistance. We compared the titer of a wild-type MLV (MLV-WT) reporter utilizing the wild-type PBS to an MLV-B2 reporter utilizing a mutant PBS (B2) escaping PBS-mediated silencing. The two viruses were used to infect the various

cell lines outlined in Figure 26B and cells were then selected with G418 for 2 weeks before colony counting. The ratio of the number of B2 colonies to WT colonies represents the degree of the fold of PBS-mediated silencing. As expected, in wild-type cell line, or cell lines expressing scrambled and negative, a B2/WT ratio close to 25 was observed (Figure 26B). In contrast, the B2/WT ration in L1td1-depleted cell line was around 6 fold. Thus, the L1td1 depletion in F9 cells caused a roughly 4-fold reduction in PBS-mediated restriction, indicating the functional importance of L1td1 in the ZFP809-TRIM28 complex. To verify the pluripotency of F9 cell upon depletion of L1td1, we tested the mRNA levels of two well-known markers of stem cells, Oct4 and Nanog. Semi-quantitative RT-PCR showed that depletion of L1td1 does not affect the expression level of those factors (Figure 26C), indicating the depletion of L1td1 may not disturb the pluripotency of F9 cells.

We also analyzed L1td1 knockout cell lines derived from C57BL/6 mouse, and kindly provided by Dr. Yamanaka from Kyoto University (Iwabuchi et al., 2011). To confirm the deletion of L1td1 gene, primers covering the L1td1 gene region were designed for genomic PCR. WT L1td1 will give a band of 17.2 Kb and knockout will give a band of 14.3 kb. Genomic PCR showed that L1td1 had been successfully deleted (Figure 27A). Viral transduction assays were performed using MLV particles pseudotyped with VSVG protein, transducing the neomycin or puromycin resistance genes and utilizing either the PBS pro or gln variants. The PBS gln is a variant PBS that is not recognized by ZFP809. In both of two L1td1 knockout cell lines, the PBS-mediated silencing is relieved approximately 4 fold,

which is consistent with the observations in the knockdown cell line (Figure 27B). Collectively, the observations suggestion that depletion of L1td1 partially relieves the PBS mediated restriction of retroviruses.

3.4.4 Depletion of L1td1 doesn't affect endogenous retroviruses

The previous results suggested that depletion of L1td1 partially relieved the PBS-mediated restriction of exogenous MLV, and indicates its participation in the RBS silencing complex. It has shown recently that ZFP809 is required for the initiation of epigenetic modification of endogenous retroviral silencing (Wolf et al., 2015). To measure the effects of L1td1 on the activities of endogenous retroelements, we used primers for quantitating the RNA levels of several endogenous retroelements, endogenous MLVs, IAPs, LINEs and MusD (Herniou et al., 1998; Matsui et al., 2010). Intracisternal A particles (IAP) and MusD are retrovirus-like retrotransposons (Aota et al., 1987; Ribet et al., 2007). The expression levels of endogenous retrolelements were assessed by reverse transcription-quantitative PCR of L1td1 wildtype and knockout ES cells. The levels of mRNA were normalized to that of UBC control. The RNA levels of E-MLV, IAP, MusD and Line were comparable between the wildtype and knockout cells, indicating that L1td1 might not affect the transcription of those families of retrotransposons (Figure 28).

3.5 Discussion

In both human and mouse, L1td1 is highly expressed in undifferentiated stem cells, but not in differentiated cell. The expression of L1td1 increases in the process of generating iPS

cells and decreases upon differentiation (Wong et al., 2011). In humans, L1TD1 is essential for the maintenance of pluripotency. Depletion of L1TD1 results in a down-regulation of OCT4 and NANOG(Emani et al., 2015). Its expression is tightly regulated by binding of those stem cell factors to the promoter region (Narva et al., 2012). Analysis of proteins bound to L1td1 in human embryonic stem cells reveals that L1td1 is part of the pluripotency interactome network of OCT4, SOX2, and NANOG, and playing an important role in RNA splicing, translation, protein traffic, and degradation, bridging nuclear and cytoplasmic regulation (Emani et al., 2015). However, L1td1 perhaps is dispensable for the induction and maintenance of pluripotency in mouse (Iwabuchi et al., 2011; Mitsui et al., 2003). Evolutionary analysis of L1TD1 found that it originated in the common ancestor of placental animals, and might have been retained for a role of genome defense against transposable element in embryonic cells. L1TD1, rapidly evolving in rodents and primates, has been lost in a subset of mammals or been selected for a function transition to pluripotency maintenance as in humans (McLaughlin et al., 2014). Our observation of the role of L1td1 in retroviral silencing in rodent cells suggests that L1td1 retains partial function of genome defense during evolution.

In summary, the binding of ZFP809 to MLV PBS ensures the capability of mouse embryonic cells to silence MLV. ZFP809, tethered to the proviral DNA, serves as a coordinator to recruit the silencing complex, including Trim28, HP1 and others, thereby transcriptionally deactivates the incoming viruses. In the report, the finding of L1td1 linking

to ZFP809-Trim28 silencing complex expands our understanding of retroviral silencing in mouse embryonic cells. Depletion of L1td1 in embryonic cells partially relieves the PBS mediated retroviral silencing, which indicates a function of L1td1, a stem-cell specific factor. However, the mechanism of how L1td1 promotes retroviral silencing is still not understood. Dr. Bestor's group recently found that L1td1 interacts with the DNA methyltransferase 3-like protein, Dnmt3L (not published). Dnmt3L tethers to the N terminus of histone H3 and recruit Dnmt3A, another DNA methytransferase (Ooi et al., 2007; Zhang et al., 2010). Dnmt3A and Dnmt3B are required for methylation of endogenous retroelements and to establish de novo methylation of newly integrated retroviruses, which is important for the initiation and maintenance of retroviral silencing (Kao et al., 2014; Liang et al., 2002; Okano et al., 1999). It is plausible that L1td1 initiates or promotes the recruitment of Dnmt3L to the newly integrated retroviral DNA, and thereby installs long-term methylation markers on the provirus. Further study will be necessary to address the relevance of L1td1 and methylation of newly integrated retroviral DNA.

Knockdown of L1td1 partially relieves the restriction occurring at the RBS, suggesting that it targets mainly PBSpro containing exogenous or endogenous retroviruses. The endogenous retrotransposons ERV, IAP and MusD utilize PBS gln, phe and lys respectively (Baust et al., 2003; Ono and Ohishi, 1983; Stoye and Coffin, 1987) and the Line elements have no PBS (Hardies et al., 2000). That might explain the reason that depletion of L1td1 doesn't affect their RNA levels. Primers targeting only PBSpro containing

retrotransposons now are available (Wolf et al., 2015) and could be used to measure the effects of L1td1 on endogenous retroviruses.

Table 1 Proteins identified by mass spectrometry.
Sample enriched for silencing complex was immunoprecipitated with anti-Trim28 antibody and IgG control. Bound proteins were probed by SDS-PAGE and displayed with coomassie blue. Bands present exclusively in anti-Trim28 immunoprecipitates were subjected to mass spectrometry.

Mouse Protein Identified	# Unique Peptides identified
LINE-1 type transposase domain containing 1	14
zinc finger protein 568	9
ATP-dependent RNA helicase DDX3X	6
zinc finger protein 809	5
Eukaryotic translation elongation factor 1 alpha 1	4
elongation factor 1-alpha	3

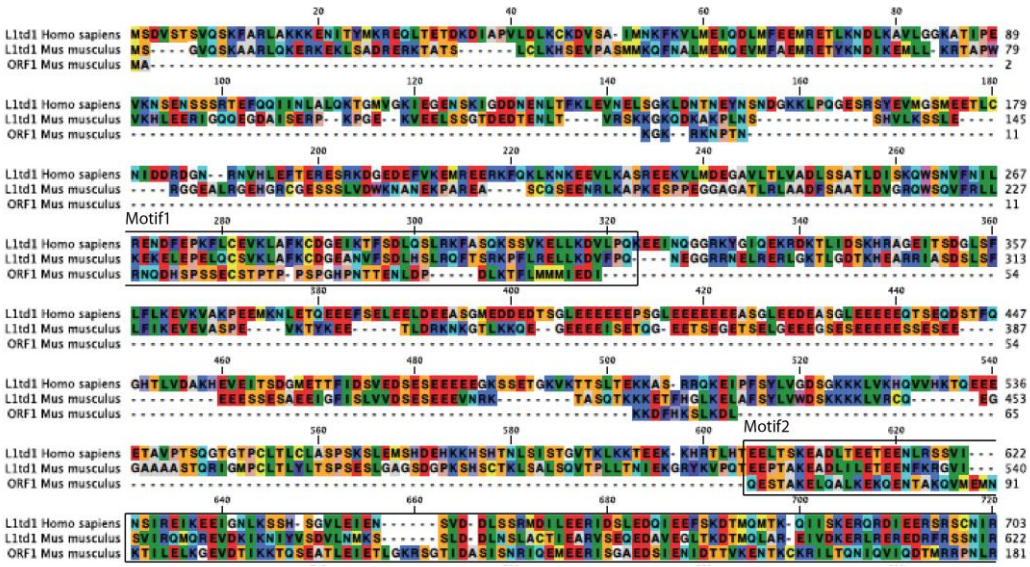

Figure 22 Protein sequence alignment of human L1TD1, mouse L1td1 and ORF1.

The sequences of human L1TD1, mouse L1td1 and ORF1 were aligned using CLCbio software. Motif 1 and 2 are two homologous domains among those three proteins. The homology of human and mouse L1td1 is around 30%.

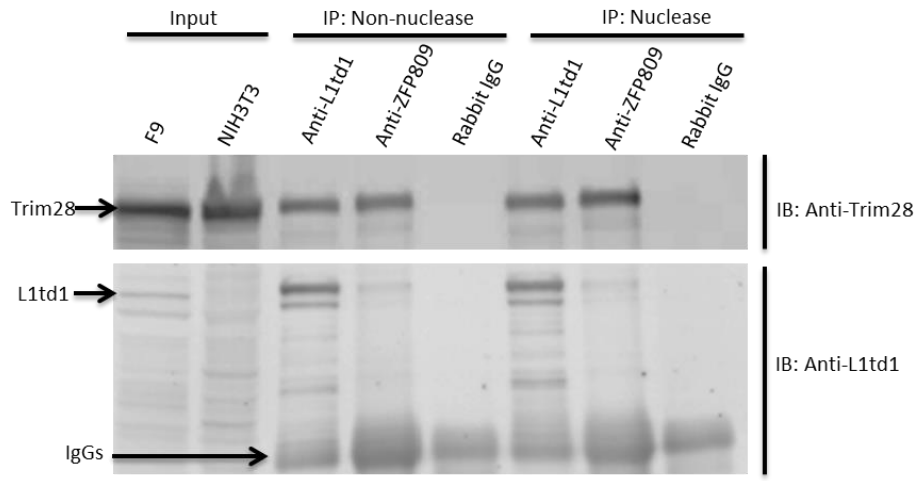

Figure 23 L1td1 interacts with both endogenous Trim28 and ZFP809. Total cell lysates of F9 cells were prepared as described previously (**Wolf and Goff, 2007**). Cells lysates were incubated with IgA and IgG beads coated with anti-L1td1 or anti-ZFP809 antibodies in 4^{o}C overnight. DNase I (1U, invitrogen) and RNase A(100 units, Qiagen) were added to indicated samples. Bound proteins were analyzed by western blotting using anti-Trim28 (Abcam) and anti-L1td1 antibodies as indicated on the right.

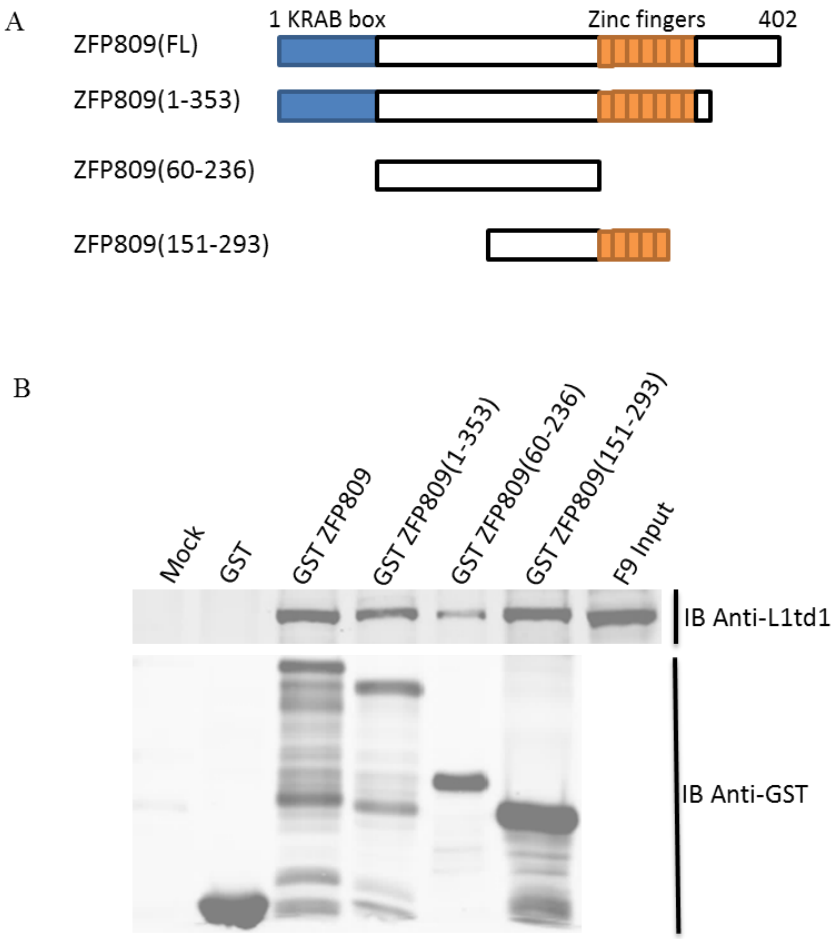

Figure 24 ZFP809 interacts with endogenous L1td1.
A) ZFP809 contains two conserved domains, KRAB box and zinc fingers. Several truncations were made as indicated and appended to GST tag. B) GST-tagged ZFP809 proteins pulled down endogenous L1td1. Bacterial expressed GST tagged ZFP809 proteins were bound to glutathione beads. Cell lysates of F9 cells were incubated with the beads in 4oC for overnight. Bound proteins were subject to 10% SDS-PAGE and western blotting with anti-L1td1 and anti-GST antibodies.

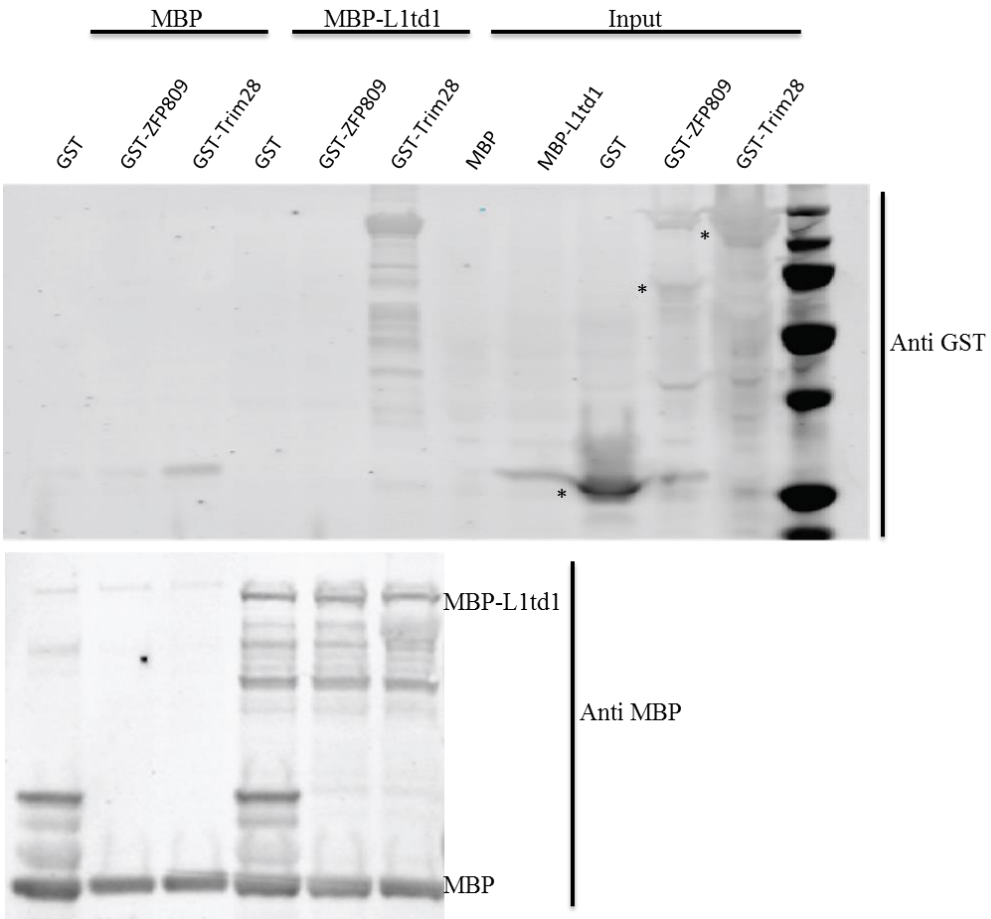

Figure 25 In vitro protein-protein interaction assay showed L1td1 interacts with Trim28 but not ZFP809.
Amylose beads were coated with bacterial expressed MBP or MBP-L1td1 and then incubated with supernatants containing GST, GST-ZFP809 or GST-Trim28. Bound proteins were eluted with Laemmli buffer and subjected to immunoblotting analysis with anti-GST and anti-MBP antibodies.

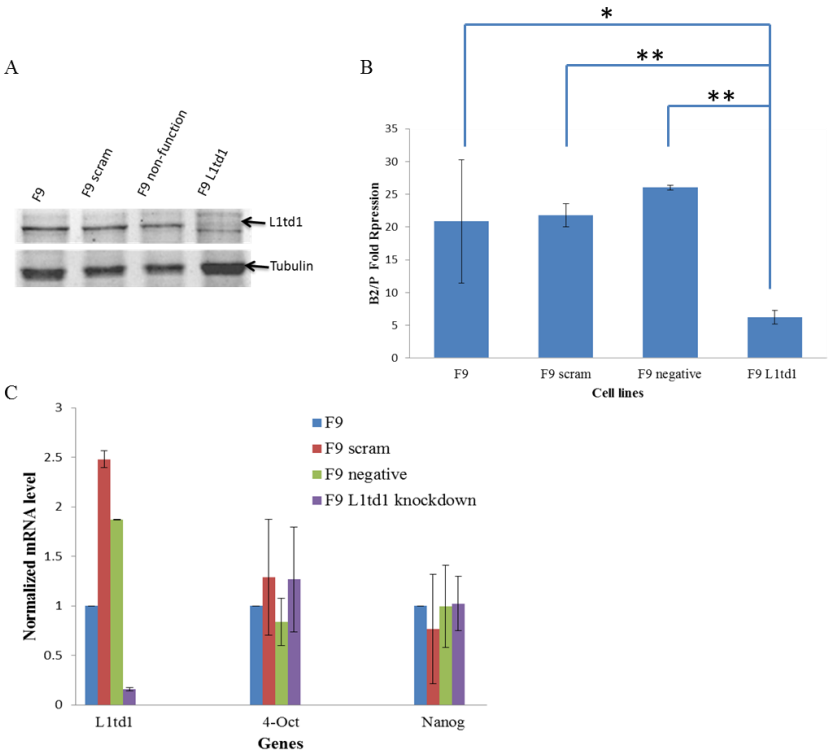

Figure 26 Depletion of L1td1 in F9 cells partially relieves the PBS-mediated retroviral silencing.

A) Total cell lysates of F9, F9 cell lines expressing scrambled (F9 scram), negative (F9 negative) and L1td1 knockdown (F9 L1td1 knockdown) shRNAs were subjected to immunoblotting with anti-L1td1 and anti-Tubulin antibodies (Abcam). L1td1 knockdown shRNAs were cloned into pSuperRetro Zeo vector (Oligoengine). The sequences are: scramble, GGTGCCGTTTAGTAGTACA; non-function, CAGCAGCCATGTCCTGAAA; L1td1 knockdown: AGGAAGAAGTGAACAGAAA. B, B) cells were infected with vesicular stomatitis virus glycoprotein (VSV G)-pseudotyped MLV particles containing either MLV-PBS$_{pro}$-Neo or MLV-PBS$_{B2}$-Neo constructs. Reporter gene expression in each cell line was assessed by colony counting after G418 selection. Graph shows the ratio of B2/Pro infection efficiency. Results shown are means ± standard errors of the means (SEMs) from three independent experiments performed in Duplicate. Student's *t* test was used for statistical analysis. * denotes $P < 0.05$, and ** $P < 0.01$. C), Total mRNA were prepared from cells indicated in the figure using TriZol reagent. Quantitative PCR, testing the relative expression level of L1td1 and two stem cell genes, Oct4 and Nanog, were performed using Gapdh and CycB as control. Gene expression levels were normalized to the control gene. The graph shows the ratio of gene expression level to those in F9 cells.

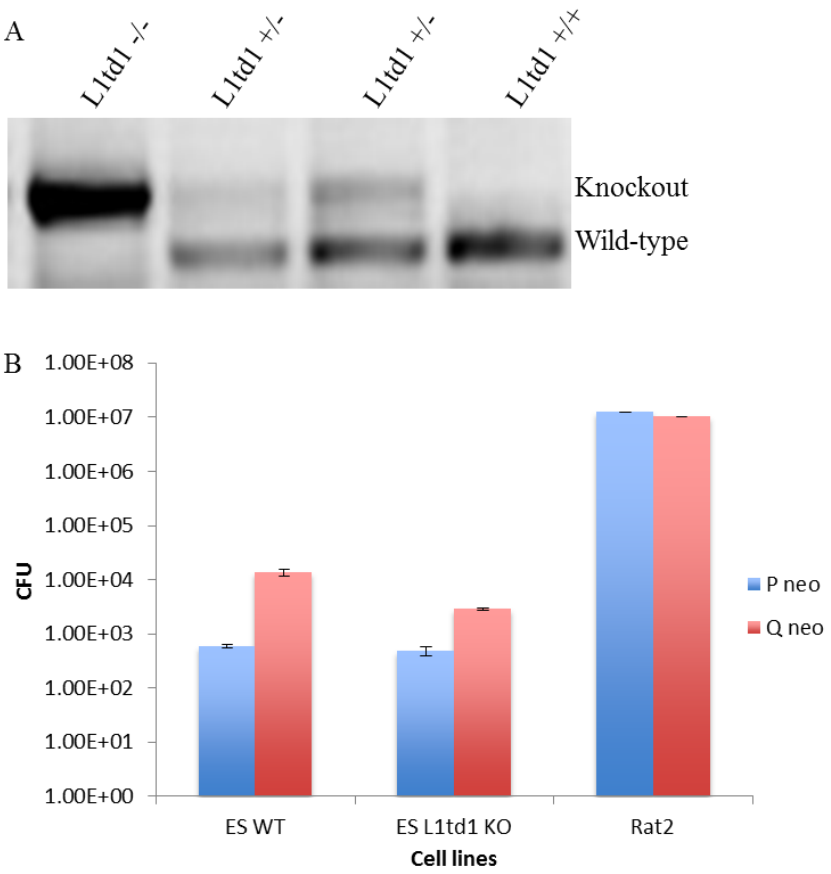

Figure 27 Knockout of L1td1 relieves the restriction in mouse ES cells. A) Genomic PCR showed successful deletion of L1td1. The bands representing knockout and wildtype L1td1 are indicated (**Iwabuchi et al., 2011**). B) Viral transduction assay using pseudotyped PBSpro or PBSGln transducing a neomycin resistance gene in L1td1 wild-type or knockout ES cells. Cells were selected in G418 for two weeks after infection. The normalized colony formation unit of P/B2 is about 20 fold in L1td1 WT cells and 5 fold in L1td1 null cells.

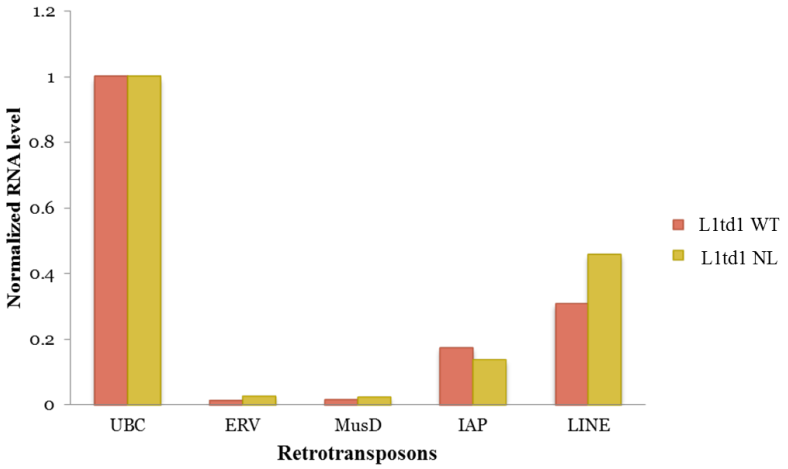

Figure 28 Knockout of L1td1 does not affect the RNA levels of several retrotransposons.
Total RNAs were prepared from L1td1 wildtype (L1td1 WT) and knockout (L1td1 NL) ES cells. Quantitative real time PCR was performed to assess the mRNA level of Endogenous retrovirus (ERV), MusD, IAP and LINE. The mRNA levels were normalized to UBC gene.

3.6 References

Akgun, E., Ziegler, M., and Grez, M. (1991). Determinants of retrovirus gene expression in embryonal carcinoma cells. Journal of virology *65*, 382-388.

Aota, S., Gojobori, T., Shigesada, K., Ozeki, H., and Ikemura, T. (1987). Nucleotide sequence and molecular evolution of mouse retrovirus-like IAP elements. Gene *56*, 1-12.

Barklis, E., Mulligan, R.C., and Jaenisch, R. (1986). Chromosomal position or virus mutation permits retrovirus expression in embryonal carcinoma cells. Cell *47*, 391-399.

Baust, C., Gagnier, L., Baillie, G.J., Harris, M.J., Juriloff, D.M., and Mager, D.L. (2003). Structure and expression of mobile ETnII retroelements and their coding-competent MusD relatives in the mouse. Journal of virology *77*, 11448-11458.

Doucet, A.J., Hulme, A.E., Sahinovic, E., Kulpa, D.A., Moldovan, J.B., Kopera, H.C., Athanikar, J.N., Hasnaoui, M., Bucheton, A., Moran, J.V., *et al.* (2010). Characterization of LINE-1 ribonucleoprotein particles. PLoS genetics *6*.

Emani, M.R., Narva, E., Stubb, A., Chakroborty, D., Viitala, M., Rokka, A., Rahkonen, N., Moulder, R., Denessiouk, K., Trokovic, R., *et al.* (2015). The L1TD1 Protein Interactome Reveals the Importance of Post-transcriptional Regulation in Human Pluripotency. Stem cell reports *4*, 519-528.

Feuer, G., Taketo, M., Hanecak, R.C., and Fan, H. (1989). Two blocks in Moloney murine leukemia virus expression in undifferentiated F9 embryonal carcinoma cells as determined by transient expression assays. Journal of virology *63*, 2317-2324.

Flanagan, J.R., Krieg, A.M., Max, E.E., and Khan, A.S. (1989). Negative control region at the 5' end of murine leukemia virus long terminal repeats. Molecular and cellular biology *9*, 739-746.

Franz, T., Hilberg, F., Seliger, B., Stocking, C., and Ostertag, W. (1986). Retroviral mutants efficiently expressed in embryonal carcinoma cells. Proceedings of the National Academy of Sciences of the United States of America *83*, 3292-3296.

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 & development *10*, 2067-2078.

Herniou, E., Martin, J., Miller, K., Cook, J., Wilkinson, M., and Tristem, M. (1998). Retroviral diversity and distribution in vertebrates. Journal of virology *72*, 5955-5966.

Hilberg, F., Stocking, C., Ostertag, W., and Grez, M. (1987). Functional analysis of a retroviral host-range mutant: altered long terminal repeat sequences allow expression in embryonal carcinoma cells. Proceedings of the National Academy of Sciences of the United States of America *84*, 5232-5236.

Iwabuchi, K.A., Yamakawa, T., Sato, Y., Ichisaka, T., Takahashi, K., Okita, K., and Yamanaka, S. (2011). ECAT11/L1td1 is enriched in ESCs and rapidly activated during iPSC generation, but it is dispensable for the maintenance and induction of pluripotency. PloS one *6*, e20461.

Kao, T.H., Liao, H.F., Wolf, D., Tai, K.Y., Chuang, C.Y., Lee, H.S., Kuo, H.C., Hata, K., Zhang, X., Cheng, X., *et al.* (2014). Ectopic DNMT3L triggers assembly of a repressive complex for retroviral silencing in somatic cells. Journal of virology *88*, 10680-10695.

Le Douarin, B., Nielsen, A.L., Garnier, J.M., Ichinose, H., Jeanmougin, F., Losson, R., and Chambon, P. (1996). A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. The EMBO journal *15*, 6701-6715.

Liang, G., Chan, M.F., Tomigahara, Y., Tsai, Y.C., Gonzales, F.A., Li, E., Laird, P.W., and Jones, P.A. (2002). Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. Molecular and cellular biology *22*, 480-491.

Linney, E., Davis, B., Overhauser, J., Chao, E., and Fan, H. (1984). Non-function of a Moloney murine leukaemia virus regulatory sequence in F9 embryonal carcinoma cells. Nature *308*, 470-472.

Loh, T.P., Sievert, L.L., and Scott, R.W. (1987). Proviral sequences that restrict retroviral expression in mouse embryonal carcinoma cells. Molecular and cellular biology *7*, 3775-3784.

Loh, T.P., Sievert, L.L., and Scott, R.W. (1990). Evidence for a stem cell-specific repressor of Moloney murine leukemia virus expression in embryonal carcinoma cells. Molecular and cellular biology *10*, 4045-4057.

Martin, S.L. (2006). The ORF1 protein encoded by LINE-1: structure and function during L1 retrotransposition. Journal of biomedicine & biotechnology *2006*, 45621.

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.

McLaughlin, R.N., Jr., Young, J.M., Yang, L., Neme, R., Wichman, H.A., and Malik, H.S. (2014). Positive selection and multiple losses of the LINE-1-derived L1TD1 gene in mammals suggest a dual role in genome defense and pluripotency. PLoS genetics *10*, e1004531.

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.

Narva, E., Rahkonen, N., Emani, M.R., Lund, R., Pursiheimo, J.P., Nasti, J., Autio, R., Rasool, O., Denessiouk, K., Lahdesmaki, H., *et al.* (2012). RNA-binding protein L1TD1 interacts with LIN28 via RNA and is required for human embryonic stem cell self-renewal and cancer cell proliferation. Stem cells *30*, 452-460.

Niwa, O., Yokota, Y., Ishida, H., and Sugahara, T. (1983). Independent mechanisms involved in suppression of the Moloney leukemia virus genome during differentiation of murine teratocarcinoma cells. Cell *32*, 1105-1113.

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.

Ono, M., and Ohishi, H. (1983). Long terminal repeat sequences of intracisternal A particle genes in the Syrian hamster genome: identification of tRNAPhe as a putative primer tRNA. Nucleic acids research *11*, 7169-7179.

Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D., *et al.* (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714-717.

Petersen, R., Kempler, G., and Barklis, E. (1991). A stem cell-specific silencer in the primer-binding site of a retrovirus. Molecular and cellular biology *11*, 1214-1221.

Ribet, D., Harper, F., Dewannieux, M., Pierron, G., and Heidmann, T. (2007). Murine MusD retrotransposon: structure and molecular evolution of an "intracellularized" retrovirus. Journal of virology *81*, 1888-1898.

Rowe, H.M., Kapopoulou, A., Corsinotti, A., Fasching, L., Macfarlan, T.S., Tarabay, Y., Viville, S., Jakobsson, J., Pfaff, S.L., and Trono, D. (2013). TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells. Genome research *23*, 452-461.

Schlesinger, S., Meshorer, E., and Goff, S.P. (2014). Asynchronous transcriptional silencing of individual retroviral genomes in embryonic cells. Retrovirology *11*, 31.

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 & development *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 & development *15*, 428-443.

Singer, M.F. (1982). SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. Cell *28*, 433-434.

Stoye, J.P., and Coffin, J.M. (1987). The four classes of endogenous murine leukemia virus: structural relationships and potential for recombination. Journal of virology *61*, 2659-2669.

Teich, N.M., Weiss, R.A., Martin, G.R., and Lowy, D.R. (1977). Virus infection of murine teratocarcinoma stem cell lines. Cell *12*, 973-982.

Truett, G.E., Heeger, P., Mynatt, R.L., Truett, A.A., Walker, J.A., and Warman, M.L. (2000). Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). BioTechniques *29*, 52, 54.

Tsukiyama, T., Niwa, O., and Yokoro, K. (1989). Mechanism of suppression of the long terminal repeat of Moloney leukemia virus in mouse embryonal carcinoma cells. Molecular and cellular biology *9*, 4670-4676.

Wang, G.Z., Wolf, D., and Goff, S.P. (2014). EBP1, a novel host factor involved in primer binding site-dependent restriction of moloney murine leukemia virus in embryonic cells. Journal of virology *88*, 1825-1829.

Wolf, D., Cammas, F., Losson, R., and Goff, S.P. (2008). Primer binding site-dependent restriction of murine leukemia virus requires HP1 binding by TRIM28. Journal of virology *82*, 4675-4679.

Wolf, D., and Goff, S.P. (2007). TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. Cell *131*, 46-57.

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

Wolf, G., Yang, P., Fuchtbauer, A.C., Fuchtbauer, E.M., Silva, A.M., Park, C., Wu, W., Nielsen, A.L., Pedersen, F.S., and Macfarlan, T.S. (2015). The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses. Genes & development *29*, 538-554.

Wong, R.C., Ibrahim, A., Fong, H., Thompson, N., Lock, L.F., and Donovan, P.J. (2011). L1TD1 is a marker for undifferentiated human embryonic stem cells. PloS one *6*, e19355.

Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W., *et al.* (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. Nucleic acids research *38*, 4246-4253.

Addendum to Chapter 3

The role of NuRD complexes in retroviral silencing

Text

It has been known for many years that the replication of Moloney Murine Leukemia Virus (MoMLV or MLV) is restricted in mouse embryonic stem (ES) and embryonic carcinoma (EC) cells (Barklis et al., 1986). Retroviral infection results in successful integration of proviral DNA, but the transcription of the provirus is blocked (Franz et al., 1986; Linney et al., 1984). The presence of so-called repressor binding site (RBS) is essential for the restriction (Hilberg et al., 1987; Linney et al., 1984), which shows overlap with the primer binding site (PBSpro) of retroviral DNA (Barklis et al., 1986; Feuer et al., 1989; Loh et al., 1987). Electrophoretic mobility shift assays (EMSA), using the RBS sequence as probe, demonstrated the presence of a stem-cell specific silencing complex (Loh et al., 1990; Petersen et al., 1991). The recruitment of trans-acting silencing complex to the RBS introduces chromatin modifications, which initiates and maintains the silencing (Akgun et al., 1991; Flanagan et al., 1989; Niwa et al., 1983; Schlesinger et al., 2014; Tsukiyama et al., 1989).

To purify the silencing complex, EMSA was used to monitor its presence. F9 nuclear extracts were precipitated with 40% ammonium sulfate. The recovered precipitate was applied to a butyl sepharose hydrophobic interaction column, and then a Sephacryl S-500 gel filtration column. The RBS binding complex was eluted from the gel filtration columns in a broad peak, indicative of an apparent size ranging from about 2 to 0.7 MDa, and centered on approximately 1.3 MDa (Wolf and Goff, 2007). Purified fractions were probed by EMSA and the shifted band defining the complex was excised from the gel. The proteins associated with the DNA were analyzed by mass spectroscopy.

The key components of the trans-acting complex were identified by our lab as ZFP809 and

Trim28 (Wolf and Goff, 2007, 2009). ZFP809 binds to the RBS and initiates the silencing of incoming retroviral DNA (Wolf and Goff, 2009). It recruits Trim28 by the KRAB box domain and then other transcription factors (Wolf and Goff, 2009), such as ESET (Matsui et al., 2010) and HP1gamma (Wolf et al., 2008) to silence the retroviral DNAs.

The sum of molecular weights of ZFP809, Trim28, ESET and HP1gamma is less than 0.7MDa, which means that there are other factors involving in the silencing complex. Further efforts to identify those factors revealed that EBP1 (Wang et al., 2014) and L1td1 might be part of the silencing complex. One interesting protein family identified in the initial mass spectrometry was nucleosome remodeling deacetylase (NuRD) complex (Table 2). NuRD complex, identified in 1998 (Tong et al., 1998), has been shown to regulate chromatin organization, gene transcription, genomic stability and developmental signaling (Allen et al., 2013; Basta and Rauchman, 2015; Ramirez and Hagman, 2009).

NuRD complex remains the only chromatin remodeler with ATPase and histone deacetylase functions (Wade et al., 1998; Xue et al., 1998; Zhang et al., 1998). it includes enzymatic subunits of HDAC1/2 (histone deacetylation complex) and CHD3/4 (chromo domain helicase DNA-binding protein) and non-enzymatic proteins, MBD2/3 (methyl-CpG-binding domain), RBBP7/4 (retinoblastoma-binding proteins), MTA1/2/3 (metastasis-associated) and p66α/β (Allen et al., 2013). HDAC1/2 (histone deacetylase) remove acetyl groups (O=C-CH3) from a ε-N-acetyl lysine amino acid on histones, allowing the histones to wrap DNAs more tightly (Choudhary et al., 2009). The chromo domain helicase DNA-binding protein 3 (CHD3) and CHD4, utilize energy derived from hydrolysis of ATP for DNA sliding and repositioning of nucleosomes (Mansfield et al., 2011). They contain tandem plant homeodomain fingers (PHDs), which bind to Lys9 (H3K9me3) and displace

HP1γ from pericentric sites (Musselman et al., 2009), but recruit other components of NuRD complex for transcription repression (Mansfield et al., 2011). MBD2 and MBD3 are members of the methyl cytosine-guanosine (CpG)-binding domain (MBD) family proteins. MBD2 and MBD3, which bind robustly to methylated DNA, are required for the complex formation and gene repression (Hendrich and Bird, 1998; Le Guezennec et al., 2006; Wade et al., 1999). The retinoblastoma-binding proteins RBBP7 and RBBP4 have histone chaperone activity (Murzina et al., 2008; Qian et al., 1993). Metastasis-associated gene (MTA) proteins, which are transcription co-regulators, were shown to be involved in breast cancer invasion and metastasis (Toh et al., 1994). The p66α and p66β are transcriptional repressor proteins (Feng et al., 2002).

It has been shown that Trim28 recruits CHD3 and CHD4 to the target for transcription silencing (Schultz et al., 2001). The presence of NuRD complex proteins in the shifted band strongly indicates that NuRD proteins may involve in the retroviral silencing. NuRD complex has also been shown to be recruited by and co-localize with Tet1 (Ten-eleven translocation methylcytosine dioxygenase 1) and 5hmC (5-hydroxymethylcytosine) (Yildirim et al., 2011). Tet1 catalyzes the conversion of the modified DNA base 5-methylcytosine (5-mC) to 5-hydroxymethylcytosine (5-hmC) (Tahiliani et al., 2009), which has been proposed as the initial step of active DNA demethylation in mammals (Shen and Zhang, 2012).

To test if NuRD complex and Tet family proteins are involved in the retroviral silencing, we cloned small hairpin RNA (shRNA) sequences, which have been shown to specifically knock down Mbd3 (Zhu et al., 2009), Mta1 (Jiang et al., 2011), Mta2 (Fu et al., 2011) and HDAC2(Harms and Chen, 2007), Tet1(Freudenberg et al., 2012) and Tet2 (Koh et al., 2011) into pLKO.1 vector. VSV-G

pseudotyped viruses, containing those shRNAs were used to infect F9 cells to generate stable knock down cell lines. Real-time quantitative PCR showed successful depletion of those proteins (Table 3). However, viral transduction assays (described in Chapter 2) using neomycin resistant gene as a reporter showed that none of those proteins affect the PBS-mediated retroviral silencing (Table 3).

Though depletion of NuRD complex proteins (Mbd3, Mta1, Mta2 and HDAC2) and Tet family proteins (Tet1 and Tet2) in F9 cells doesn't change the restriction fold of P/B2, it is not sufficient to conclude that NuRD complex and Tet proteins are involved in the retrovirus silencing. One of the reasons may be the redundancy of those proteins. Protein in families of Mbd, Mta, HDAC and Tet, have high similarity in sequences and functions to each other. The function of depleted protein may be complemented by other members in the protein family. All of the homologs should be depleted to test their roles in retroviral silencing. Other assays such as ChIP sequencing could be used to access the presence of NuRD complex and Tet proteins in retroviral DNAs.

Table 2 Proteins identified by mass spectrometry.
F9 nuclear extract was purified for enriching MLV silencing complex. After last step of purification, samples were subjected to electrophoretic mobility shift assay and the shifted bands were removed from the gel and processed by mass spectrometry(**Wolf and Goff, 2009**).

96-WT 4-7-05	
Protein Identified	No. of Peptides Matched
tripartite motif protein 28	9
Mcm6	7
ankyrin repeat and FYVE domain containing 1	8
MTA1	6
polyadenylate binding protein 2	5
Mcm6	5
Mcm4	6
putative eukaryotic translation initiation factor 3	3
chaperonin subunit 8	4
T-complex protein 1	2
heat shock protein 27	3

96-MUT 4-7-05	
Protein Identified	No. of Peptides Matched
ankyrin repeat and FYVE domain containing 1	8
tripartite motif protein 28	6
MTA1	4
MTA3	4
Mcm7	5
Mcm6	2
Mcm4	3
Mcm deficinet 3	2
retinoblastoma binding protein 7	2
Mcm defieict 2 mitotin	2
GDP-mannose phosphorylase A	2

97-WT 4-7-05	
Protein Identified	No. of Peptides Matched
alpha actinin 4	19
filamin B	18
tripartite motif protein 28	10
alpha actinin 1	7
valosin-containing protein	7
minichromosome maintenance protein 7 (Mcm7)	2
Mcm2	1
Mcm6	3
retinoblastoma binding protein 7 (CAF-1 subunit C)	2
ankyrin repeat and FYVE domain containing 1	1
metastasis-associated 1 (MTA1)	3

97-MUT 4-7-05	
Protein Identified	No. of Peptides Matched
alpha actinin 4	22
tripartite motif protein 28	12
Mcm6	2
Mcm4	2
casein kinase II	2
Mcm deficient 2 mitotin	3
Mcm deficient 5	3
ankyrin repeat and FYVE domain containing 1	1
RAN GTPase activating protein 1	3
histone deacetylase 2	2
COPS5	1
Mta3	2

Table 3 Small hairpin RNA depletion of components of NuRD complex and Tet proteins does not affect the PBS-mediated retroviral silencing.
The small hairpin RNA (shRNA) sequences for knockdown indicated proteins were listed below. Those sequences were cloned into pLKO.1 vector and transduced into F9 cells by VSVG pseudotyped viruses. The levels of knockdown were verified by quantitative real time PCR. The restriction fold was calculated from viral transduction assay using pseudotyped PBSpro or PBS-B2 viruses described in Chapter 3.

Genes	shRNA sequences	Knockdown (By RTPCR)	Restrction fold
Scrambled	GCGAAGTGCATTGTGTGGC	Non	25
Mbd3	GATGAATAAGAGTCGCCAG	Yes	20
	AGCCTTCATGGTGACAGAT	No	21
Mta2	GGTTGTCTGTCTTTTCCGG	Yes	30
Hdac2	CCAATGAGTTGCCATATAA	Yes	15
Mta1	GAAATATGGTGGCTTGAAA	Yes	23
	GGACATATTGGAAGAAATA	Yes	22
Tet1	AACTTGCATCCACGATTA	Yes	18
	GAATTACAGTTGTTACGGA	No	24
Tet2	GAAAGCAGCTCGAAAGCGT	Yes	20
	ACTACTAACTCCACCCTAA	Yes	21

References

Akgun, E., Ziegler, M., and Grez, M. (1991). Determinants of retrovirus gene expression in embryonal carcinoma cells. Journal of virology *65*, 382-388.

Allen, H.F., Wade, P.A., and Kutateladze, T.G. (2013). The NuRD architecture. Cellular and molecular life sciences : CMLS *70*, 3513-3524.

Barklis, E., Mulligan, R.C., and Jaenisch, R. (1986). Chromosomal position or virus mutation permits retrovirus expression in embryonal carcinoma cells. Cell *47*, 391-399.

Basta, J., and Rauchman, M. (2015). The nucleosome remodeling and deacetylase complex in development and disease. Translational research : the journal of laboratory and clinical medicine *165*, 36-47.

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V., and Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science *325*, 834-840.

Feng, Q., Cao, R., Xia, L., Erdjument-Bromage, H., Tempst, P., and Zhang, Y. (2002). Identification and functional characterization of the p66/p68 components of the MeCP1 complex. Molecular and cellular biology *22*, 536-546.

Feuer, G., Taketo, M., Hanecak, R.C., and Fan, H. (1989). Two blocks in Moloney murine leukemia virus expression in undifferentiated F9 embryonal carcinoma cells as determined by transient expression assays. Journal of virology *63*, 2317-2324.

Flanagan, J.R., Krieg, A.M., Max, E.E., and Khan, A.S. (1989). Negative control region at the 5' end of murine leukemia virus long terminal repeats. Molecular and cellular biology *9*, 739-746.

Franz, T., Hilberg, F., Seliger, B., Stocking, C., and Ostertag, W. (1986). Retroviral mutants efficiently expressed in embryonal carcinoma cells. Proceedings of the National Academy of Sciences of the United States of America *83*, 3292-3296.

Freudenberg, J.M., Ghosh, S., Lackford, B.L., Yellaboina, S., Zheng, X., Li, R., Cuddapah, S., Wade, P.A., Hu, G., and Jothi, R. (2012). Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. Nucleic acids research *40*, 3364-3377.

Fu, J., Qin, L., He, T., Qin, J., Hong, J., Wong, J., Liao, L., and Xu, J. (2011). The TWIST/Mi2/NuRD protein complex and its essential role in cancer metastasis. Cell research *21*, 275-289.

Harms, K.L., and Chen, X. (2007). Histone deacetylase 2 modulates p53 transcriptional activities through regulation of p53-DNA binding activity. Cancer research *67*, 3145-3152.

Hendrich, B., and Bird, A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. Molecular and cellular biology *18*, 6538-6547.

Hilberg, F., Stocking, C., Ostertag, W., and Grez, M. (1987). Functional analysis of a retroviral host-range mutant: altered long terminal repeat sequences allow expression in embryonal carcinoma cells. Proceedings of the National Academy of Sciences of the United States of America *84*, 5232-5236.

Jiang, Q., Zhang, H., and Zhang, P. (2011). ShRNA-mediated gene silencing of MTA1 influenced on protein expression of ER alpha, MMP-9, CyclinD1 and invasiveness, proliferation in breast cancer cell lines MDA-MB-231 and MCF-7 in vitro. Journal of experimental & clinical cancer research : CR *30*, 60.

Koh, K.P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C.A., Mostoslavsky, G., *et al.* (2011). Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell stem cell *8*, 200-213.

Le Guezennec, X., Vermeulen, M., Brinkman, A.B., Hoeijmakers, W.A., Cohen, A., Lasonder, E., and Stunnenberg, H.G. (2006). MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. Molecular and cellular biology *26*, 843-851.

Linney, E., Davis, B., Overhauser, J., Chao, E., and Fan, H. (1984). Non-function of a Moloney murine leukaemia virus regulatory sequence in F9 embryonal carcinoma cells. Nature *308*, 470-472.

Loh, T.P., Sievert, L.L., and Scott, R.W. (1987). Proviral sequences that restrict retroviral expression in mouse embryonal carcinoma cells. Molecular and cellular biology *7*, 3775-3784.

Loh, T.P., Sievert, L.L., and Scott, R.W. (1990). Evidence for a stem cell-specific repressor of Moloney murine leukemia virus expression in embryonal carcinoma cells. Molecular and cellular biology *10*, 4045-4057.

Mansfield, R.E., Musselman, C.A., Kwan, A.H., Oliver, S.S., Garske, A.L., Davrazou, F., Denu, J.M., Kutateladze, T.G., and Mackay, J.P. (2011). Plant homeodomain (PHD) fingers of CHD4 are histone H3-binding modules with preference for unmodified H3K4 and methylated H3K9. The Journal of biological chemistry *286*, 11779-11791.

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.

Murzina, N.V., Pei, X.Y., Zhang, W., Sparkes, M., Vicente-Garcia, J., Pratap, J.V., McLaughlin, S.H., Ben-Shahar, T.R., Verreault, A., Luisi, B.F., *et al.* (2008). Structural basis for the recognition of histone H4 by the histone-chaperone RbAp46. Structure *16*, 1077-1085.

Musselman, C.A., Mansfield, R.E., Garske, A.L., Davrazou, F., Kwan, A.H., Oliver, S.S., O'Leary, H., Denu, J.M., Mackay, J.P., and Kutateladze, T.G. (2009). Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications. The Biochemical journal *423*, 179-187.

Niwa, O., Yokota, Y., Ishida, H., and Sugahara, T. (1983). Independent mechanisms involved in suppression of the Moloney leukemia virus genome during differentiation of murine teratocarcinoma cells. Cell *32*, 1105-1113.

Petersen, R., Kempler, G., and Barklis, E. (1991). A stem cell-specific silencer in the primer-binding site of a retrovirus. Molecular and cellular biology *11*, 1214-1221.

Qian, Y.W., Wang, Y.C., Hollingsworth, R.E., Jr., Jones, D., Ling, N., and Lee, E.Y. (1993). A retinoblastoma-binding protein related to a negative regulator of Ras in yeast. Nature *364*, 648-652.

Ramirez, J., and Hagman, J. (2009). The Mi-2/NuRD complex: a critical epigenetic regulator of hematopoietic development, differentiation and cancer. Epigenetics : official journal of the DNA Methylation Society *4*, 532-536.

Schlesinger, S., Meshorer, E., and Goff, S.P. (2014). Asynchronous transcriptional silencing of individual retroviral genomes in embryonic cells. Retrovirology *11*, 31.

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 & development *15*, 428-443.

Shen, L., and Zhang, Y. (2012). Enzymatic analysis of Tet proteins: key enzymes in the metabolism of DNA methylation. Methods in enzymology *512*, 93-105.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., *et al.* (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930-935.

Toh, Y., Pencil, S.D., and Nicolson, G.L. (1994). A novel candidate metastasis-associated gene, mta1, differentially expressed in highly metastatic mammary adenocarcinoma cell lines. cDNA cloning, expression, and protein analyses. The Journal of biological chemistry *269*, 22958-22963.

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

Tsukiyama, T., Niwa, O., and Yokoro, K. (1989). Mechanism of suppression of the long terminal repeat of Moloney leukemia virus in mouse embryonal carcinoma cells. Molecular and cellular biology *9*, 4670-4676.

Wade, P.A., Gegonne, A., Jones, P.L., Ballestar, E., Aubry, F., and Wolffe, A.P. (1999). Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. Nature genetics *23*, 62-66.

Wade, P.A., Jones, P.L., Vermaak, D., and Wolffe, A.P. (1998). A multiple subunit Mi-2 histone deacetylase from Xenopus laevis cofractionates with an associated Snf2 superfamily ATPase. Current biology : CB *8*, 843-846.

Wang, G.Z., Wolf, D., and Goff, S.P. (2014). EBP1, a novel host factor involved in primer binding site-dependent restriction of moloney murine leukemia virus in embryonic cells. Journal of virology *88*, 1825-1829.

Wolf, D., Cammas, F., Losson, R., and Goff, S.P. (2008). Primer binding site-dependent restriction of murine leukemia virus requires HP1 binding by TRIM28. Journal of virology *82*, 4675-4679.

Wolf, D., and Goff, S.P. (2007). TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. Cell *131*, 46-57.

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

Xue, Y., Wong, J., Moreno, G.T., Young, M.K., Cote, J., and Wang, W. (1998). NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. Molecular cell *2*, 851-861.

Yildirim, O., Li, R., Hung, J.H., Chen, P.B., Dong, X., Ee, L.S., Weng, Z., Rando, O.J., and Fazzio, T.G. (2011). Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. Cell *147*, 1498-1510.

Zhang, Y., LeRoy, G., Seelig, H.P., Lane, W.S., and Reinberg, D. (1998). The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. Cell *95*, 279-289.

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

Chapter 4

Discussion

Studies of regulation of retroviral replication in stem cells and differentiated cells will provide us insights as to how organisms restrict exogenous viral infection to protect their genome integrity, and yet also allow moderate retroviral activities to develop necessary diversity for adapting to environmental changes. The primary focus of this thesis has been to investigate how mouse embryonic cells restrict retroviral infection and how the restriction is lost during differentiation. In Chapter 2, we demonstrated that the essential restriction factor of MLV ZFP809 is subject to proteasomal degradation in differentiated cells, but is relatively stable in stem cells.. In Chapter 3, we characterized the role of another stem cell specific protein, L1td1, in retroviral silencing in embryonic cells. In this chapter, the results will be reviewed briefly and explored in broader context. In addition, future directions will be discussed.

4.1 Proteasomal regulation of ZFP809, a retroviral repressor

ZFP809 was identified as the zinc finger protein binding to the PBS of proviral DNA and initiating the retroviral silencing in stem cells (Wolf and Goff, 2009). ZFP809 protein is not expressed in differentiated cells (Wolf and Goff, 2009), but the levels of ZFP809 mRNA are approximately the same in stem cells and differentiated cells. which indicates that some post-transcriptional mechanism is involved in the depletion of ZFP809

proteins. In Chapter 2, our observations showed that ZFP809 is degraded rapidly in differentiated cells but stable in stem cells. The degradation is mediated by the proteasomal pathway and very likely ubiquitin-dependent. The 50 amino acids and the lysine residue in the C-terminal of the protein are essential for the turnover of ZFP809. Surprisingly, Trim28 protein, which is recruited by ZFP809 and responsible for formation of silencing complexes, can promote the degradation of ZFP809 in differentiated cells. Using small drug MLN4924 also revealed that neddylation, perhaps by a cullin family protein might be involved in the degradation of ZFP809.

It has been known for many years that the transcription of MLV DNA is silenced in murine stem cells but not in differentiated cells. A zinc finger protein, ZFP809 binds directly to the PBS of MLV proviral DNA and initiates the silencing. Ectopic expression of ZFP809 alone in murine or human differentiated cells is sufficient to establish the silencing. In murine stem and differentiated cells, the mRNA levels of ZFP809 are comparable. We discovered that ZFP809 proteins are turned over rapidly in differentiated cells but are stable in stem cells, which may explain why the replication of MLV is restricted in stem cells but not in differentiated cells. After tethering to the PBS, ZFP809 recruits Trim28 through KRAB box and then other transcription repressors to block the expression of viral genes. Though trim28 is an essential component of the silencing complex in stem cells, it promotes the degradation of ZFP809 in differentiated cells. Overall the silencing machinery of MLV is subtly regulated.

In general, around 40% of variation in protein concentration can be explained by their mRNA abundance (Vogel and Marcotte, 2012). In many other cases, regulatory processes, occurring after transcription, such as RNA processing, translational regulation, and protein degradation, are controlling the abundances of proteins. In this work, we found that the key component of retroviral silencing complex, ZFP809, is regulated by the proteasomal protein degradation system. After translation, ZFP809 is subjected to the ubiquitin-mediated degradation. Post-translational regulation via ubiquitin proteasome system (UPS) plays an important role in retroviral replication. Ubiquitin is involved in the traffic and processing of Gag polyprotein, and in the budding of retroviral particles (Patnaik et al., 2000; Schubert et al., 2000; Strack et al., 2000). HIV subverts the Cullin-Ring finger Ubquitin ligases (CRLs) to overcome some restriction factors and optimize its replication and spreading (Barry and Fruh, 2006; Malim, 2014). Here we discovered that the UPS degrades the key restriction factor blocking the transcription of MLV in embryonic stem cells and so allows its replication in differentiated cells. Depletion of ZFP809 might also relieve the restriction of those endogenous retroelements using PBSpro.

Ubiquitin proteins are transferred to targets through three ligases: E1, E2 and E3. The specificity of substrates selection usually depends on the E3 ligases. Ring domain, F-box and HECT domain containing proteins can serve as E3 ligases(Deshaies and Joazeiro, 2009). TRIM28, which containing a Ring domain, is a potential E3 ligase(Iyengar and Farnham, 2011). Trim28 can promote the degradation of ZFP809 but it cannot be concluded that Trim28 serves as an E3 ligase for ZFP809. Usually Ring-finger proteins

function with cullin proteins as the Cullin-Ring finger complex to target proteins for degradation. The PHD domain of TRIM28 functions as an intramolecular E3 ligase that sumoylates the adjacent TRIM28 bromo domain, which recruits SETDB1, a transcriptional repressor (Ivanov et al., 2007). Another protein modification, phosphorylation is also important for its function in transcription repression (Li et al., 2007). The phosphorylated and sumoylated Trim28, has a shorter half-life and is targeted for degradation by RNF4 (Kuo et al., 2014). It seems that after silencing the specific chromatin region, Trim28 and its associated transcription repressors undergo protein degradation. The silenced MLV in stem cells cannot be activated during differentiation, because the restriction complex has already placed many silencing marks on the proviral DNA(Yao et al., 2004). Once the silencing is done, Trim28, ZFP809 and other transcription repressors might be subject to degradation. Future studies aimed at identifying the E3 ligase of ZFP809 and the mechanism of ZFP809 degradation will give us a more detailed view as to how the UPS regulates retroviral replication.

ZFP809 not only is required for the initiation of silencing of exogenous retroviruses, but also for the silencing of endogenous retroelements using PBSpro sequence (Wolf et al., 2015). Investigation of how ZFP809 protein is regulated in different cell types will also extend our understanding of how endogenous retroviruses are regulated in those cells. There are several classes of experiments that would allow us to further delineate the fine-tuned regulation of ZFP809 and retroviruses it tethers to.

Trim28 is a RING domain containing factor, which has been shown to have E3 ligase activity (Chen et al., 2011; Ivanov et al., 2007). Abolishing the interaction between Trim28 and ZFP809 stabilized ZFP809 in differentiated cells. Taken together with the observation that depletion of Trim28 using siRNA also stabilized ZFP809 in those cells, this result suggests that Trim28 promotes the degradation of ZFP809. We cannot yet conclude that Trim28 is the specific E3 ligase that links ubiquitin proteins to ZFP809 and targets the protein to proteasomes for degradation. In vitro ubiquitination assay could be employed to demonstrate if Trim28 directly ligates ubiquitin to ZFP809 (Zhao et al., 2012). Specificity and modification problems should be considered when doing in vitro experiments. Mutated Trim28 without E3 ligase activity could be introduced into Trim28 depleted cells to test its role in ZFP809 degradation.

Ring domain-containing proteins often couple with cullin family proteins to form Cullin-RING ligases (CRL) for protein degradation (Deshaies and Joazeiro, 2009). Cullin proteins serve as scaffold for recruiting E2, E3 ligases and substrate (Bosu and Kipreos, 2008). The neddylation, resulting in structural change of cullin proteins, is important for the interactions of E2, E3 and substrate on Cullin scaffold (Merlet et al., 2009). The neddylation inhibitor MLN4924 stabilized ZFP809 in differentiated cells, indicating potential involvement of neddylation, perhaps a cullin protein in ZFP809 degradation. There are seven Cullin proteins in mammals (Hua and Vierstra, 2011). Analyzing the expression profile of cullin proteins to identify differentiated cell-specific cullins would narrow down the targets. Then using siRNAs to knock down cullin proteins and testing the

stability of ZFP809 upon cullin depletion will allow us to characterize which cullin protein is responsible for ZFP809 degradation. Since MLN4924 inhibits neddylation globally, there might be other pathways deactivated by the drug and indirectly affecting the degradation of ZFP809.

Biochemical purification could also be employed to identify the E3 ligase for ZFP809 directly. Affinity purification of E3 ligases bound to immobilized substrates have been shown before (Reiss and Hershko, 1990). ZFP809 could be used as a bait to enrich the E3 ligases. As we showed above, full-length ZFP809 protein is stable in stem cells and a truncated ZFP809 protein, ZFP809(1-353) is not turned over in differentiated cells. Bound proteins from stem cells and ZFP809(1-353) could be used as controls. The interactions between E3 ligase and substrate are relatively weak and transient. Methods such as protein-protein linkers that could stabilize protein interactions should be used during the purification.

In addition to affinity purification, libraries for screening of E3 ligases are available. In order to create a clear phenotype for screening, a construct with mCherry IRES GFP-tagged ZFP809 could be made. The construct could be introduced into cells expressing the siRNA library by either viral transduction or transfection. High ratio of GFP/mCherry is an indication of stabilizing of ZFP809 by the specific siRNAs.

As we observed above, Trim28 promotes ZFP809 degradation in differentiated cells, however, the interaction between Trim28 and ZFP809 are present in stem cells as well. In stem cells, ZFP809 should be stable enough to silence the incoming retroviruses. It would

be interesting to reveal what mechanisms regulate the different stability of ZFP809 in different type of cells, which might lead to the discovery of the pathways that host cells use to regulate exogenous and endogenous viral activities. In order to identify factors involving in the signal of ZFP809 degradation, a genome-wide screening (CRISPR or shRNA) using the phenotype described in last paragraph could be performed. There are numerous studies on how the transcription of viral DNA is blocked, but there is little known about the upstream signal for initiating or removing the silencing. The screen would reveal factors involved in the signal pathway leading to degradation.

Retroviral replication is blocked in stem cells, but not in differentiated cells. To investigate in what stage of differentiation that host cells are starting to lose the restriction would give us a picture of when host cells need to protect the genome from exogenous infection and mutagenesis by endogenous retroviruses. Viral transduction or transfection assays with a reporter gene could be used to measure the restriction fold during the differentiation of stem cells. Then using high-throughput technology to profile the expression patterns of stem cells and the cells in pluripotent transition would reveal important transcriptional differences that might be essential for the regulation of exogenous and endogenous retroviruses.

Global ChIP-sequencing revealed that ZFP809 controls the replication of endogenous retroviruses (Wolf et al., 2015). Further studies of ZFP809 regulation would reveal how host cells respond to the exogenous and endogenous retroviral activities.

Most of retroelements inside mammalian genomes are tightly regulated. Most of them are controlled by the epigenetic machinery (Rowe and Trono, 2011). During the development of mammalian organisms, the methylation pattern of DNA are reprogrammed several times and the activities of most of the retroelements are hyperactive during the removal of methylation markers, which probably introduce the somatic mutations for genome expansion (Messerschmidt et al., 2014). For example, the non-LTR retrotransposon, Line-1 is extremely active in neuroprogenitor cells, which expands the neuronal diversity and is important for the development of neuron system (Thomas and Muotri, 2012). Depletion of ZFP809 using the ubiquitination system, which allows the retrotransposition of endogenous retroviruses using PBSpro, might also result in somatic mutations in the genome and contribute to the genetic diversity. But the exact function is not known yet.

4.2 Components of RBS silencing complex

L1td1, a stem cell-specific transcript, was identified during the purification of RBS silencing complex. Human L1td1 plays a crucial role in maintaining pluripotency of human stem cells (Narva et al., 2012), probably through their involvement in RNA biology (Emani et al., 2015). However, the mouse L1td1 seems to be dispensable for generation or maintenance of pluripotency (Iwabuchi et al., 2011). Our observation suggests that depletion of L1td1 in mouse embryonic cells partially relieves the restriction from PBS, by interacting with a key component of RBS silencing complex, Trim28.

The evolutionary analysis revealed that L1td1 genes of different species originated from a common ancestor and had genome defense activity. It lost or changed the function during evolution (McLaughlin et al., 2014). The role of L1td1 in retroviral silencing in mouse embryonic cells suggests that L1td1 may retain its function in mouse. L1td1 was also identified as a protein interacting with Dnmt3L, an important DNA methytransferase. The status of DNA methylation is positively associated with the expression of retroviral genes (Matsui et al., 2010). However, DNA methylation was proposed to maintain the long-term silencing of retroviral DNA but not initiation of the silencing (Cherry et al., 2000; Lorincz et al., 2000; Lorincz et al., 2002; Yao et al., 2004). The interaction between Dnmt3L and L1td1 might explain the minor effect of L1td1 in retroviral silencing, because depletion of de novo DNA methytransferase only increased the retroviral activity moderately (Laker et al., 1998; Niwa et al., 1983). To study the relevance of L1td1 and DNA methylation, protein-protein interaction assays should be done to confirm the interaction between Dnmt3L and L1td1. DNA methylation status of proviral DNA in L1td1 depleted cells could be monitored by bisulfate sequencing. There will be less DNA methylation in L1td1 depleted cells if L1td1 truly participates in the recruitment of DNA methylation machinery to proviral DNA.

L1td1 plays an important role in the stem cell biology in humans, probably in post-trasncriptional regulation or RNA biology (Emani et al., 2015). It has been demonstrated that retroelement activities are associated with the pluripotency of stem cells (Coufal et al., 2009; Macfarlan et al., 2012). It will be interesting to learn if L1td1 is involved in the

regulation of retroelements in human stem cells. L1td1 might not participate in the silencing of exogenous HIV and MLV since they are not likely to be silenced in human embryonic cells (Macia et al., 2011). As revealed by evolutionary analysis, L1td1 might have originated as defense against LINE-1 elements (McLaughlin et al., 2014). The activity of engineered Line-1 or expression level of Line-1 RNA in L1td1 depleted cell could be tested to investigate L1td1's role in Line-1 retrotransposition. Global profiling of the expression pattern of L1td1 depleted cells would reveal more retro-transposition activity change.

In addition to ZFP809, Trim28, HP1, EBP1 and L1td1, there are likely to be other factors in the silencing complex. Further identification of components in the silencing complex would allow us to understand the establishment of retroviral silencing comprehensively. The failure to detect restriction fold change in cell lines with depleting single proteins of NuRD complex might be the result of redundancy. Knockdown or knockout all of the homologous proteins should be achieved to test their effects on retroviral silencing. A global screening of factors that relieves the retroviral silencing could also mine out more participators.

4.3 References

Barry, M., and Fruh, K. (2006). Viral modulators of cullin RING ubiquitin ligases: culling the host defense. Science's STKE : signal transduction knowledge environment *2006*, pe21.

Bosu, D.R., and Kipreos, E.T. (2008). Cullin-RING ubiquitin ligases: global regulation and activation cycles. Cell division *3*, 7.

Chen, C.C., Chen, Y.Y., Tang, I.C., Liang, H.M., Lai, C.C., Chiou, J.M., and Yeh, K.C. (2011). Arabidopsis SUMO E3 ligase SIZ1 is involved in excess copper tolerance. Plant physiology *156*, 2225-2234.

Cherry, S.R., Biniszkiewicz, D., van Parijs, L., Baltimore, D., and Jaenisch, R. (2000). Retroviral expression in embryonic stem cells and hematopoietic stem cells. Molecular and cellular biology *20*, 7419-7426.

Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Yeo, G.W., Mu, Y., Lovci, M.T., Morell, M., O'Shea, K.S., Moran, J.V., and Gage, F.H. (2009). L1 retrotransposition in human neural progenitor cells. Nature *460*, 1127-1131.

Deshaies, R.J., and Joazeiro, C.A. (2009). RING domain E3 ubiquitin ligases. Annual review of biochemistry *78*, 399-434.

Emani, M.R., Narva, E., Stubb, A., Chakroborty, D., Viitala, M., Rokka, A., Rahkonen, N., Moulder, R., Denessiouk, K., Trokovic, R., *et al.* (2015). The L1TD1 Protein Interactome Reveals the Importance of Post-transcriptional Regulation in Human Pluripotency. Stem cell reports *4*, 519-528.

Hua, Z., and Vierstra, R.D. (2011). The cullin-RING ubiquitin-protein ligases. Annual review of plant biology *62*, 299-334.

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

Iwabuchi, K.A., Yamakawa, T., Sato, Y., Ichisaka, T., Takahashi, K., Okita, K., and Yamanaka, S. (2011). ECAT11/L1td1 is enriched in ESCs and rapidly activated during iPSC generation, but it is dispensable for the maintenance and induction of pluripotency. PloS one *6*, e20461.

Iyengar, S., and Farnham, P.J. (2011). KAP1 protein: an enigmatic master regulator of the genome. The Journal of biological chemistry *286*, 26267-26276.

Kuo, C.Y., Li, X., Kong, X.Q., Luo, C., Chang, C.C., Chung, Y., Shih, H.M., Li, K.K., and Ann, D.K. (2014). An Arginine-rich Motif of Ring Finger Protein 4 (RNF4) Oversees the Recruitment and Degradation of the Phosphorylated and SUMOylated Kruppel-associated Box Domain-associated Protein 1 (KAP1)/TRIM28 Protein during Genotoxic Stress. The Journal of biological chemistry *289*, 20757-20772.

Laker, C., Meyer, J., Schopen, A., Friel, J., Heberlein, C., Ostertag, W., and Stocking, C. (1998). Host cis-mediated extinction of a retrovirus permissive for expression in embryonal stem cells during differentiation. Journal of virology *72*, 339-348.

Li, X., Lee, Y.K., Jeng, J.C., Yen, Y., Schultz, D.C., Shih, H.M., and Ann, D.K. (2007). Role for KAP1 serine 824 phosphorylation and sumoylation/desumoylation switch in regulating KAP1-mediated transcriptional repression. The Journal of biological chemistry *282*, 36177-36189.

Lorincz, M.C., Schubeler, D., Goeke, S.C., Walters, M., Groudine, M., and Martin, D.I. (2000). Dynamic analysis of proviral induction and De Novo methylation: implications for a histone deacetylase-independent, methylation density-dependent mechanism of transcriptional repression. Molecular and cellular biology *20*, 842-850.

Lorincz, M.C., Schubeler, D., Hutchinson, S.R., Dickerson, D.R., and Groudine, M. (2002). DNA methylation density influences the stability of an epigenetic imprint and Dnmt3a/b-independent de novo methylation. Molecular and cellular biology *22*, 7572-7580.

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.

Macia, A., Munoz-Lopez, M., Cortes, J.L., Hastings, R.K., Morell, S., Lucena-Aguilar, G., Marchal, J.A., Badge, R.M., and Garcia-Perez, J.L. (2011). Epigenetic control of retrotransposon expression in human embryonic stem cells. Molecular and cellular biology *31*, 300-316.

Malim, M.H. (2014). HIV: Ringside views. Nature *505*, 167-168.

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.

McLaughlin, R.N., Jr., Young, J.M., Yang, L., Neme, R., Wichman, H.A., and Malik, H.S. (2014). Positive selection and multiple losses of the LINE-1-derived L1TD1 gene in mammals suggest a dual role in genome defense and pluripotency. PLoS genetics *10*, e1004531.

Merlet, J., Burger, J., Gomes, J.E., and Pintard, L. (2009). Regulation of cullin-RING E3 ubiquitin-ligases by neddylation and dimerization. Cellular and molecular life sciences : CMLS *66*, 1924-1938.

Messerschmidt, D.M., Knowles, B.B., and Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes & development *28*, 812-828.

Narva, E., Rahkonen, N., Emani, M.R., Lund, R., Pursiheimo, J.P., Nasti, J., Autio, R., Rasool, O., Denessiouk, K., Lahdesmaki, H., *et al.* (2012). RNA-binding protein L1TD1 interacts with LIN28 via RNA and is required for human embryonic stem cell self-renewal and cancer cell proliferation. Stem cells *30*, 452-460.

Niwa, O., Yokota, Y., Ishida, H., and Sugahara, T. (1983). Independent mechanisms involved in suppression of the Moloney leukemia virus genome during differentiation of murine teratocarcinoma cells. Cell *32*, 1105-1113.

Patnaik, A., Chau, V., and Wills, J.W. (2000). Ubiquitin is part of the retrovirus budding machinery. Proceedings of the National Academy of Sciences of the United States of America *97*, 13069-13074.

Reiss, Y., and Hershko, A. (1990). Affinity purification of ubiquitin-protein ligase on immobilized protein substrates. Evidence for the existence of separate NH2-terminal binding sites on a single enzyme. The Journal of biological chemistry *265*, 3685-3690.

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

Schubert, U., Ott, D.E., Chertova, E.N., Welker, R., Tessmer, U., Princiotta, M.F., Bennink, J.R., Krausslich, H.G., and Yewdell, J.W. (2000). Proteasome inhibition interferes with gag polyprotein processing, release, and maturation of HIV-1 and HIV-2. Proceedings of the National Academy of Sciences of the United States of America *97*, 13057-13062.

Strack, B., Calistri, A., Accola, M.A., Palu, G., and Gottlinger, H.G. (2000). A role for ubiquitin ligase recruitment in retrovirus release. Proceedings of the National Academy of Sciences of the United States of America *97*, 13063-13068.

Thomas, C.A., and Muotri, A.R. (2012). LINE-1: creators of neuronal diversity. Frontiers in bioscience *4*, 1663-1668.

Vogel, C., and Marcotte, E.M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nature reviews Genetics *13*, 227-232.

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

Wolf, G., Yang, P., Fuchtbauer, A.C., Fuchtbauer, E.M., Silva, A.M., Park, C., Wu, W., Nielsen, A.L., Pedersen, F.S., and Macfarlan, T.S. (2015). The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses. Genes & development *29*, 538-554.

Yao, S., Sukonnik, T., Kean, T., Bharadwaj, R.R., Pasceri, P., and Ellis, J. (2004). Retrovirus silencing, variegation, extinction, and memory are controlled by a dynamic interplay of multiple epigenetic modifications. Molecular therapy : the journal of the American Society of Gene Therapy *10*, 27-36.

Zhao, Q., Liu, L., and Xie, Q. (2012). In vitro protein ubiquitination assay. Methods in molecular biology *876*, 163-172.