UNIVERSITY OF CALIFORNIA, SAN DIEGO

Host Cell Factors Involved in Retrovirus Replication and Disease Pathogenesis

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Biology

by

Shannon Beth Seidel

Committee in Charge,
Professor John Young, Chair
Professor Steve Briggs
Professor Michael David
Professor John Guatelli
Professor Clodagh O'Shea

2012

The dissertation of Shannon Beth Seidel is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2012

DEDICATION

I wish to dedicate this thesis to my family for their un-ending support throughout my life and particularly during the years I have spent in graduate school.
With the Seidels around, laughter abounds.

TABLE OF CONTENTS

SIGNATURE PAGE iii
DEDICATION iv
TABLE OF CONTENTS v
LIST OF TABLES vii
LIST OF FIGURES viii
ACKNOWLEDGEMENTS ix
VITA xiii
PUBLICATIONS xiv
ABSTRACT OF THE DISSERTATION xv
Chapter 1. Introduction 1
1.1 Introduction 1
1.2 Retroviruses 4
1.2.1 Cellular Factors Involved in Retrovirus Replication 7
1.2.2 Murine Leukemia Virus 10
1.2.2.1 MLV Transcription 12
1.2.2.2 MLV Pathogenesis 13
1.2.3 Human Immunodeficiency Virus 16
1.2.3.1 HIV Transcription 18
1.2.3.2 HIV Pathogenesis 19
1.2.3.3 HIV Current Treatments 21
1.3 Role of cellular transcription factors in cellular tropism and disease pathogenesis 23
1.4 Zinc Finger Proteins 25
1.4.1 Zinc Finger Proteins In Retrovirus Replication 26
1.4.2 ZASC1 27
1.5 Contribution to the Field 32
Chapter 2. Identification and Characterization of Host Cell Factors in Retrovirus Replication 33
2.1 Background 33
2.2 Contribution to the Work 34
2.3 Results 36
2.4 Conclusions 49
Chapter 3. ZASC1 Knockout Mice Exhibit an Early Bone Marrow-specific Defect in Murine Leukemia Virus Replication 52
3.1 Abstract 52
3.2 Introduction 53
3.3 Results 54
3.4 Discussion 70
3.5 Methods 78
Chapter 4. Conclusions and Future Directions 90
Appendix 1. Generating the ZASC1$^{-/-}$ Mouse 95
Appendix 2. Additional Characterization of ZASC1 107

A2.1 Background 107
A2.2 Results 108
A2.3 Discussion 119
A2.4 Methods 121
References 125

LIST OF TABLES

Table 3.1 Statistical Analysis of MLV Infection at Early Time Points in the Spleen, Thymus and Bone Marrow. 61
Table 3.2 Primers used to design the ZASC1 Targeting Vector 79
Table 3.3 Primers used to generate southern blot probes 86
Table A1.1 Results of ES Cell injection and Chimera Breeding 103

LIST OF FIGURES

Figure 1.1 Retrovirus Genome Structure....6
Figure 1.2 Retroviral Life Cycle....9
Figure 1.3 HIV-1 and Moloney MLV Promoters....15
Figure 1.4 Retroviral Reverse Transcription....11
Figure 1.5 ZASC1 and Zinc Finger Proteins....29
Figure 1.6 Conservation of ZASC1 binding sites in MLV....31
Figure 3.1. Generating the ZASC1-/- mouse model....56
Figure 3.1. (continued)....57
Figure 3.2 Time course of Mo-MuLV infection *in vivo*....60
Figure 3.3 Mo-MuLV infection is delayed in ZASC1-/- mice at an early time point. 62
Figure 3.3 (continued)....63
Figure 3.4 ZASC1 is not required for normal differentiation of spenocytes. 66
Figure 3.5 ZASC1 is not required for development of thymocytes. 67
Figure 3.6 ZASC1 is not required for T-cell activation and proliferation. 69
Figure 3.7 Flowchart of Hematopoietic Development....71
Figure 3.8 ZASC1-/- mice have increased LK cell compartments in the bone marrow....73
Figure 3.8 (continued)....74
Figure 3.9 No defect in Mo-MuLV pathogenesis in ZASC1-/- mice. 75
Figure 3.10 MLV Infections in Primary MEFs....82
Figure A1.1 In vitro cre recombinase assay....96
Figure A1.2 Southern Blot at the downstream arm of the ZASC1 Targeting Vector 98
Figure A1.3 Southern Blot at the Upstream Arm of the ZASC1 Targeting Vector 100
Figure A1.3 (continued)....101
Figure A1.4 ZASC1 Conditional Knockout Mouse Strains Generated 106
Figure A2.1 Screening Jurkat cell clones....110
Figure A2.2 Growth Curve in A549 Cells with ZASC1 shRNA....111
Figure A2.3 Activation of T-cells – Antibody Titration....114
Figure A2.4 Cytokine expression in activated T-cells....115
Figure A2.5 MLV-luc (VSV-G) Infection in Primary CD4+ T-cells....117
Figure A2.6 MLV-luc (VSV-G) Infection in Primary Bone Marrow....118

ACKNOWLEDGEMENTS

First and foremost, I offer my deepest thanks to my thesis advisor, John Young for his guidance, financial support and trust as I endeavored into a new area of research for the lab. I appreciate the time and effort he has put into helping me become a more independent scientist, a better writer and a more effective communicator. I will always be grateful for his mentorship.

I wish to thank my thesis committee for contributing their valuable time and experience to my project and scientific development. Michael David, Clodagh O'Shea, Steve Briggs and John Guatelli have all provided thoughtful suggestions over the years and have helped to steer me in a more productive direction. In particular, thank you to Clodagh O'Shea for providing additional individual mentorship and encouragement over the years.

I would like to thank my co-authors from the publication in *Cell*: Renate Konig, Yingyao Zhou, Tracy Diamond, Ghislain Bonamy, Jeffrey Irelan, Chih-yuan Chiang, Buu Tu, Paul De Jesus, Caroline Lilley, Amanda Opaluch, Jeremy Caldwell, Matt Weitzman, Kelli Kuhen, Sourav Bandyopadhyay, Trey Ideker, Anthony Orth, Loren Miraglia, Rick Bushman, Sumit Chanda, and in particular Daniel Elleder who served as a mentor to me in the Young lab. Chapter 2 is a direct reprint of our publication, Koenig R, Zhou Y, Elleder D, Diamond T, Bonamy G, Irelan J, Chiang C, Tu B, De Jesus P, Lilley C, Seidel S, Opaluch A, Caldwell J, Weitzman M, Kuhen K, Bandyopadhyay S, Ideker T, Orth A, Miragila L, Bushman F, Young J, Chanda S. 2008. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* 135: 49–60.

To the other members of the Young lab past and present I have many debts of gratitude. In particular, I would like to thank Patti Ryan who served as the senior graduate student for most of my time in the lab and also became a wonderful friend. It is hard to imagine navigating life in the lab without her support and advice. I appreciate the guidance of mini-PI Jeff Murry as well as the numerous discussions of podcasts and politics. I'm also thankful to Melissa Rodgers and Rose-Marie Pilpa for providing everything from experimental advice to fashion tips. Thanks to John Naughton for keeping the lab running and designing beautiful figures, some of which appear in this thesis. Thanks to Taneashia Morrell for organizing numerous lab events and to the other excellent scientists that I have been lucky enough to work with in the Young lab. John Marlett, Suchita Bhattacharyya, Jonah Rainey, Heather Scobie, Brooke Baker, Shilpi Sharma, Arturo Diaz, Justine Swann and Sebastien Landry generously shared their expertise when needed and helped to create a great environment in the lab.

I am grateful to all of the collaborators both at The Salk Institutes and at other institutions around the country who helped in my efforts to study ZASC1 *in vivo* including: Kuo-Fen Lee, Mahru An, Matt Leblanc, Hung Fan, Jim Bruce, and Paul Ahlquist. Additionally, I greatly appreciate Hung Fan, Jim Bruce, Chassidy Johnson, Ye Zheng, Erica Stone and Steve Hedrick for generously sharing reagents for this project.

The Salk Institute has been a wonderful place to complete my graduate studies. I am grateful to the Animal Resources Department who housed the mice

used in my ZASC1 projects. In particular, I wish to convey many thanks to Maria Encizo and Sarah Chacon the technicians who cared for these animals daily. I also appreciate the help of the veterinary technicians Amanda Chambers, Liliana Vera and Caitlin Utschig. Additionally, I want to thank the individuals of the CCMI, Microarray/qPCR, Transgenic Mouse, and Biophotonics core facilities. Each of these groups contributed to my project. I would also like to thank the members of the Development Office who helped me to obtain funding and who facilitated interactions with financial supporters of The Salk Institute.

Within my graduate program at UCSD there are a number of individuals whose support made my studies possible. Thanks to Cathy Pugh, Tom Tomp, Jim Kadonaga and Gabriele Wienhausen. In particular, I wish to acknowledge Gabriele who has been an important advocate for graduate students. In the greater UCSD community, I wish to thank Rhonda Hackshaw and Jim Carter for providing support through CAPS and Rosalind Streichler and Beth Simon for providing opportunities through the Center for Teaching Development.

I am grateful for the financial support provided by the Cellular and Molecular Genetics Training Grant awarded to the University of California, San Diego by the National Institutes of Health. Funding from the CMG training grant supported the first three years of my graduate career. I also wish to acknowledge the support of the H.A. and Mary K. Chapman Charitable Trust, the Bruce J. Heim Foundation, the Timken-Sturgis Foundation and the Rose Hills Foundation.

My interest in Biology was ignited in college when I first worked in a research setting. I am grateful to Ed Jahngen at the University of Massachusetts

Lowell, and Paul Nieman at Fred Hutchinson Cancer Research Center for giving me the opportunity to work their labs and for encouraging me to attend graduate school.

A number of people enhanced my graduate school experience outside of the lab. I want to thank the members of my book club and my soccer team, Salsa Verde as well as my little sister, Joanna Melendez.

I truly appreciate the friendship of my graduate student colleagues who have supported, commiserated and celebrated every step of this journey with me. I have run thousands of miles with my running crew Julia Claggett and Lindsay Schwarz over the past 7 years. During those runs we have talked about scientific endeavors as well as those outside of the lab. Many thanks to my other classmates and friends in San Diego including Colleen Barefield, Mary Noe, Angela Parrish, Adam Pare, Laurel Seaholm, Justin Davison, Rachel Smith, Erin Soto, Seema Lakdawala, Britt Flaherty, Christene Banda and numerous others.

Lastly, and most importantly, I am grateful to my family. It is hard to find words to express how appreciative I am of my parents, David and Ruth Seidel who have loved and supported me unconditionally throughout my life. I also want to thank my brothers, Luke and Daniel Seidel, who have made me laugh even in the most difficult times. Many thanks as well to my sister-in-law Robyn Seidel, my grandma Florence Seidel, my sister-by-choice Sara McCleskey, and other extended family who have been particularly supportive.

VITA

Education and Research Experience

UC San Diego 2005 – 2012 Ph.D. in Biology
Dissertation: Host Cell Factors Involved in Retrovirus Replication and Disease Pathogenesis
Advisor: Dr. John AT Young, Salk Institute

University of Massachusetts, Lowell 2001 – 2005 B.S. in Biology, summa cum laude

Fred Hutchinson Cancer Research Center Summer, 2004
Undergraduate Research Project: Exploring the role of Nr13, a Bcl2 anti-apoptotic family member in zebrafish development Advisor: Dr. Paul Neiman

Teaching Experience

Guest Teacher, Canyon Crest Academy 2011
Teaching Assistant Consultant, UCSD 2008-2010
Consultant, NSF Socrates Fellows, UCSD 2008-2009
Teaching Assistant, Immunology Lecture, UCSD 2008
Teaching Assistant, Molecular Biology Section - California State 2007
Laboratory Teaching Assistant: Recombinant DNA Techniques, UCSD 2006

Outreach Experience

Team Member, Bio Education and Art for Science Innovation (BioEASI) 2009-2010
Exhibitor, R.H. Fleet Science Center. 2008; 2009
Instructor, Salk Institute Mobile Science Laboratory 2006-2010
Presenter, Salk Institute High School Science Day 2008-2010
Speaker, Salk Institute Development Office 2010
Activity Leader, BioBridge 2008
Activity Assistant, Expanding Your Horizons 2007; 2008
Big Sister, Big Sister League of San Diego 2007 to Present

University Service

UCSD Student Representative, Graduate Committee, Division of Biology 2009-2010
Biology Representative, Graduate Student Association 2009-2010
Graduate Recruitment Committee, Division of Biology
- Student Chair 2007
- CommitteeMember 2006

PUBLICATIONS

[Manuscript in preparation] Seidel S, Bruce J, Leblanc M, Lee K, Ahlquist P, Young J. ZASC1 Knockout Mice Exhibit an Early Bone Marrow-specific Defect in Murine Leukemia Virus Replication.

Koenig R, Zhou Y, Elleder D, Diamond T, Bonamy G, Irelan J, Chiang C, Tu B, De Jesus P, Lilley C, Seidel S, Opaluch A, Caldwell J, Weitzman M, Kuhen K, Bandyopadhyay S, Ideker T, Orth A, Miragila L, Bushman F, Young J, Chanda S. 2008. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* 135: 49–60.

ABSTRACT OF THE DISSERTATION

Host Cell Factors Involved in Retrovirus Replication and Disease Pathogenesis

by

Shannon Beth Seidel

Doctor of Philosophy in Biology

University of California, San Diego, 2012

Professor John Young, Chair

Retroviruses are relevant human pathogens affecting the lives of tens of millions of people globally each year and resulting in nearly 2 million deaths annually. Understanding the process of retrovirus infection has wide benefits including the identification of therapeutic targets to treat HIV/AIDS and better understanding of cellular pathways disrupted in diseases, such as cancer and AIDS. Here, I first present my role in a collaborative project that led to the identification of over 200 cellular factors with potentially important roles in the

early stages of retrovirus infection. We identified these factors through a genome-wide siRNA screen and through characterization of interactions between host and virus proteins. Additionally, I present my work on ZASC1, a host cell transcription factor that binds to the promoter regions of both HIV-1 and MLV and regulates virus gene expression in cultured established cell lines. I generated a ZASC1 knockout mouse model, showing that this gene is non-essential to development and reproduction. This mouse model was used to characterize the role of ZASC1 in retrovirus replication and disease pathogenesis. I showed that ZASC1 influences myeloid cell differentiation in this compartment and that this transcription factor is required for efficient early Mo-MuLV infection in the bone marrow compartment. Through studies of Mo-MuLV pathogenesis in the ZASC1 knockout mouse model, I show that despite this early defect, tumorigenesis is not regulated by ZASC1.

Chapter 1. Introduction

1.1 Introduction

Though viruses have been recognized as biological entities capable of causing disease for well over a century, they have been mutating and coevolving with their host organisms much longer. Viral pathogens commandeer the molecular machinery within the cells of their host and utilize this machinery for the purpose of viral replication. In addition to using cellular factors to help complete the virus life cycle, many viruses have developed specific strategies to evade or dampen host immune defenses.

Although a great deal of research has been done to better understand the process of virus infection, most has focused specifically on the viral proteins and their functions, while much less is known about the requirement of host cell proteins during virus infection. Research to help understand the cellular machinery used by viruses is essential to designing therapeutics that target the host cell proteins viruses rely upon. This is a new paradigm in antiviral research. Most traditional therapies target virus factors often leading to high mutation rates and rapid selection of resistant mutants. In the case of HIV-1, over 200 viral mutations have been linked to antiretroviral drug resistance (reviewed in [1]). The identification of druggable host cell factors required for HIV-1 replication will

provide new targets for antiretroviral drugs that presumably will be more difficult for the virus evade.

In addition to providing host protection, the study of host-virus interactions can lead to important insights into the cellular processes that viruses exploit. Viruses have evolved the ability to target proteins involved in crucial cellular pathways including endocytosis, transcription, nuclear export, cytoplasmic trafficking, and immune signaling. Many of these same pathways are commonly disrupted in other diseases such as cancer and autoimmunity. Studying viruses will not only further our understanding of their normal cellular function, but will also provide new insights into other diseases which affect them.

Viruses have become nearly ubiquitous tools for basic molecular and cell biology research. Retroviruses in particular are also important for gene delivery in research labs and therapeutic settings. In research settings, the ability of viruses to efficiently deliver genetic material into cells has been harnessed as delivery systems to insert genes or short hairpin RNAs into cultured cells and animals. In the clinical setting, viral delivery systems are being used for gene therapy. One of the major challenges of gene therapy is the ability to regulate the expression of the targeted genes, preventing the virus promoter from being silenced once it is integrated in the cell [2]. In addition to benefits to gene therapy, transcription is a major determinant of retrovirus disease pathogenesis [3-7]. Indeed, studies show that transcription factors regulate disease pathogenesis of many retroviruses . It is therefore important to identify and characterize the mechanisms exploited by viruses to transcribe their genomes.

Work presented in this thesis examines the role of host cell factors in the process of retrovirus replication. Studies described in Chapter 2 address efforts to identify novel cellular factors involved in HIV-1 replication through systems biology approaches. Studies in Chapter 3 and Appendix 2 characterize the role of ZASC1 in MLV replication and pathogenesis in a mouse model system.

1.2 Retroviruses

Retroviruses are important pathogens in humans and animals where they cause a wide range of diseases including cancer and Acquired Immunodeficiency Syndrome (AIDS). Peyton Rous discovered the first member of the retrovirus family in 1911. By injecting a cell free filtrate made from chicken sarcomas into healthy chickens, Rous showed that cancer could be transmissible [8]. Because the cell extract had been passed through a filter so fine that bacteria could not pass through, the infectious tumor-inducing agent had to be a virus. Since the initial discovery of this virus family, the number of retroviruses and the variety of host species have grown to include several important human pathogens including Human Immunodeficiency Viruses (HIV-1 and HIV-2) and the Human T-cell Lymphotropic Viruses (HTLV-I and HTLV-II). HTLV was the first human retrovirus identified and, like many of the zoonotic retroviruses, is oncogenic [9, 10]. HIV is the most well known member of the retrovirus family and will be discussed in more detail in section 1.2.2.

Retroviruses are RNA viruses that carry two copies of their non-segmented, plus-strand RNA genome within each enveloped viral particle. The common feature of members in this family is the ability to reverse-transcribe their RNA genome into a double-stranded DNA copy that can be integrated into the host cell genome. This process requires the enzymatic function of a viral-encoded reverse transcriptase (RT) [11, 12].

According to the International Committee on Taxonomy of Viruses, there are now 7 genera of exogenous retroviruses including the Alpha-, Beta-, Gamma-, Delta-, and Epsilon- retroviruses along with the Lentiviruses and Spumaviruses. All retroviruses have in common a minimum of three genes: *gag, pol* and *env*, which code for the virion's structural proteins, enzymatic proteins and envelope glycoprotein respectively (Figure 1.1). Retroviruses from the Alpha-, Beta-, Epsilon-and Gamma-retrovirus genera contain only these three genes and are considered simple retroviruses. Delta-retroviruses along with Spumaviruses and Lentiviruses are considered complex retroviruses because they contain additional accessory genes that further enhance their fitness.

The retroviral life cycle (Figure 1.2) begins with the virus binding to receptors and entering the cell, either by direct fusion with cell membrane or internalization by endocytosis followed by fusion with the endosomal membrane [13, 14]. Following fusion, the viral core is deposited in the cytoplasm and a poorly understood process of uncoating occurs in which the viral RT enzyme, packaged into the virion with the viral genome, generates a double-stranded DNA copy (Figure 1.3). The partially uncoated viral core containing the double-stranded DNA copy, termed the pre-integration complex (PIC), is then transported to the nucleus where the viral integrase enzyme works with cellular factors to integrate the viral DNA into the host chromatin.

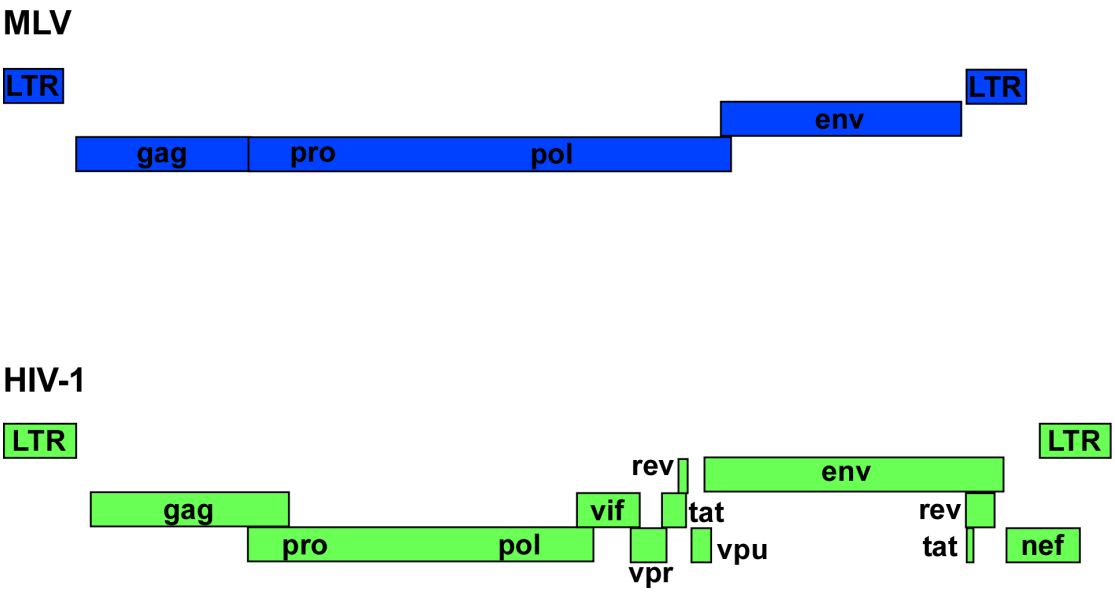


Figure 1.1 Retrovirus Genome Structure

All retroviruses share three common genes; gag, pol and env. Top Panel: MLV shown in blue contains only these three genes. Gag, pro and pol are all in the same reading frame and env is shifted by 2 base pairs. Bottom Panel: The HIV genome is shown in green. HIV is a complex retrovirus and contains the accessory proteins vif, vpr, rev, tat, vpu, and nef in addition to gag, pol and env.

1.2.1 Cellular Factors Involved in Retrovirus Replication

A variety of host cell factors interact with retroviruses at each stage of the viral life cycle beginning with binding of the cellular receptor. Cellular receptors for many mammalian retroviruses have been identified and those for MLV and HIV will be discussed below in section 1.2.2 and 1.2.3 respectively. Once the virus enters the host cell, there are a number of steps in the virus life cycle that require host cell factors. For example, IQGAPs, which are regulators of the cytoskeleton, have been shown to aid a post-entry, pre-RT step in MLV replication through an interaction with the viral Matrix protein [15]. Cyclophilin A has been shown to interact with HIV *gag* and regulate HIV-1 infectivity, however, the mechanism by which it does this is still unclear [16, 17].

Once the retroviral pre-integration complex has been translocated into the nucleus, the host factor LEDGF/p75 has been shown to interact with HIV-1 integrase in order to help facilitate proviral integration [18]. Many cellular factors are required for transcription from the provirus and these factors will be discussed in greater detail in sections 1.2.2.1 and 1.2.3.1. In the later stages of the virus life cycle, the ESCRT complex has been shown to interact with HIV-*gag* and promote virus release from the plasma membrane [19-22]. Although a number of host factors involved in retroviral replication have been identified and characterized, many others are yet to be uncovered.

When my thesis work began, only a handful of cellular factors had been identified as regulating retrovirus replication. However, RNA interference

technology was rapidly improving making it feasible to perform large-scale forward genetic screens in mammalian cells. Work in Chapter 2 will discuss my efforts in collaboration with other members of the Young Lab, the Chanda Lab and the Bushman Lab to identify cellular factors involved in retrovirus replication using a genome wide siRNA screen approach [23]. This work resulted in the identification of over 200 cellular proteins in a number of different cellular pathways with potential roles in the early stages of retrovirus infection. Several other labs also published screens to identify cellular factors involved in retrovirus infection around the same time and the results of this work have generated a number of cellular protein networks required for retrovirus infection [24-26].

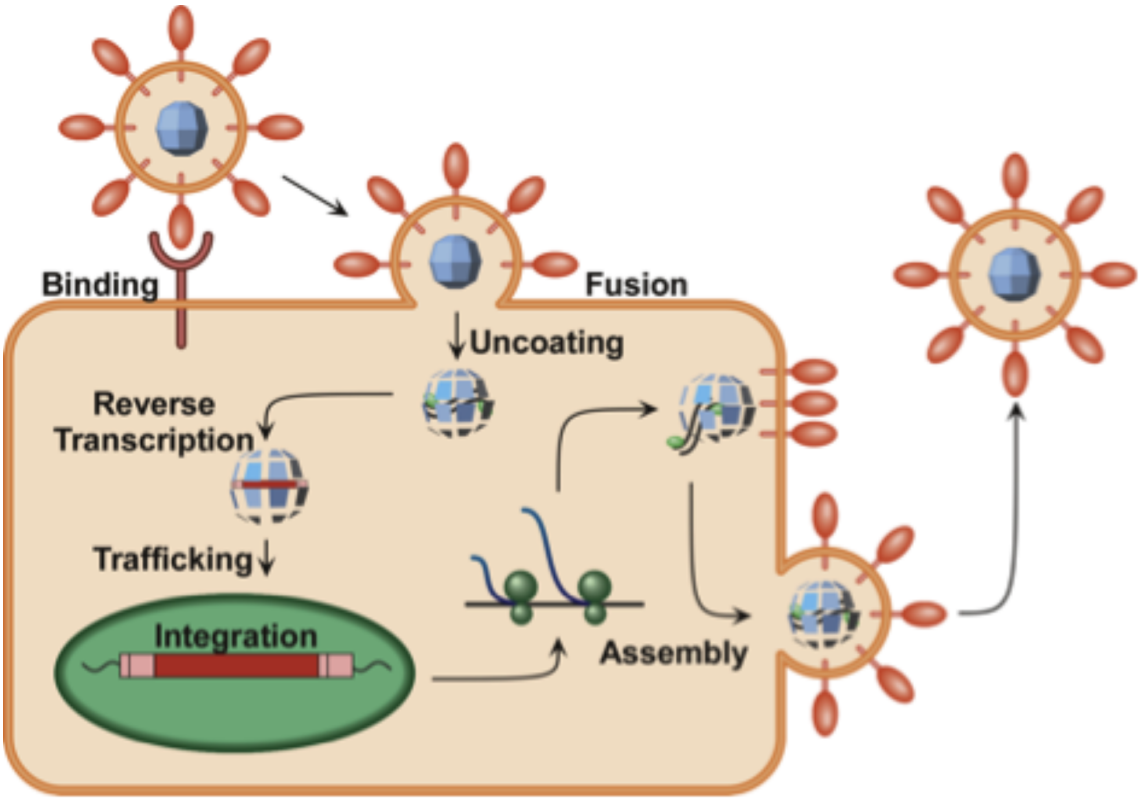


Figure 1.2 Retroviral Life Cycle

The retrovirus life cycle begins when the viral envelope glycoprotein binds to its cellular receptor. The virus envelope then fuses with the cell membrane either directly at the cell surface or through fusion with an endosomal membrane. The viral core is deposited into the cytoplasm inducing reverse transcription. The ” pre-integration complex” (PIC) including the viral DNA is then translocated to the nucleus and integrated into the host cell chromatin. Virus gene expression occurs in a similar manner to cellular transcription generating transcripts to produce viral proteins and complete genome RNAs. Nascent viral particles are assembled and bud from the membrane. Finally, the viral protease cleaves the Gag-Pol fusion proteins into their mature protein form, completing the viral life cycle.

1.2.2 Murine Leukemia Virus

Murine Leukemia Virus is a well-studied non-acute, oncogenic gamma-retrovirus that can be transmitted both exogenously and endogenously. MLV is one of the simplest members of the retrovirus family, encoding only the three genes common to all retroviruses; gag, pol and env. MLV pathogenesis has been studied now for over 60 years, yielding important discoveries in virology and cell biology.

MLV subtypes can be categorized as ecotropic, amphotropic, xenotropic, or polytropic based on the species they are capable of infecting. There are also endogenous retroviruses that are heritably integrated into the host genome and typically have inactivating mutations. Infectivity of exogenous MLVs is dependent on multiple-transmembrane receptors for entry into host cells. The cationic amino acid transporter (CAT-1) is the receptor for ecotropic MLV [27]; Xenotropic and polytropic MLV entry is facilitated by XPR1 [28, 29]; and amphotropic MLV entry is dependent upon by Pit1, a type III sodium-dependent phosphate transporter [30, 31]. These receptors are widely expressed in many different tissues, however, MLVs cannot infect non-dividing cells and are presumed to gain access to chromatin during mitosis. Therefore they typically replicate within cells of the lymphoid and myeloid lineages that are actively dividing [32].

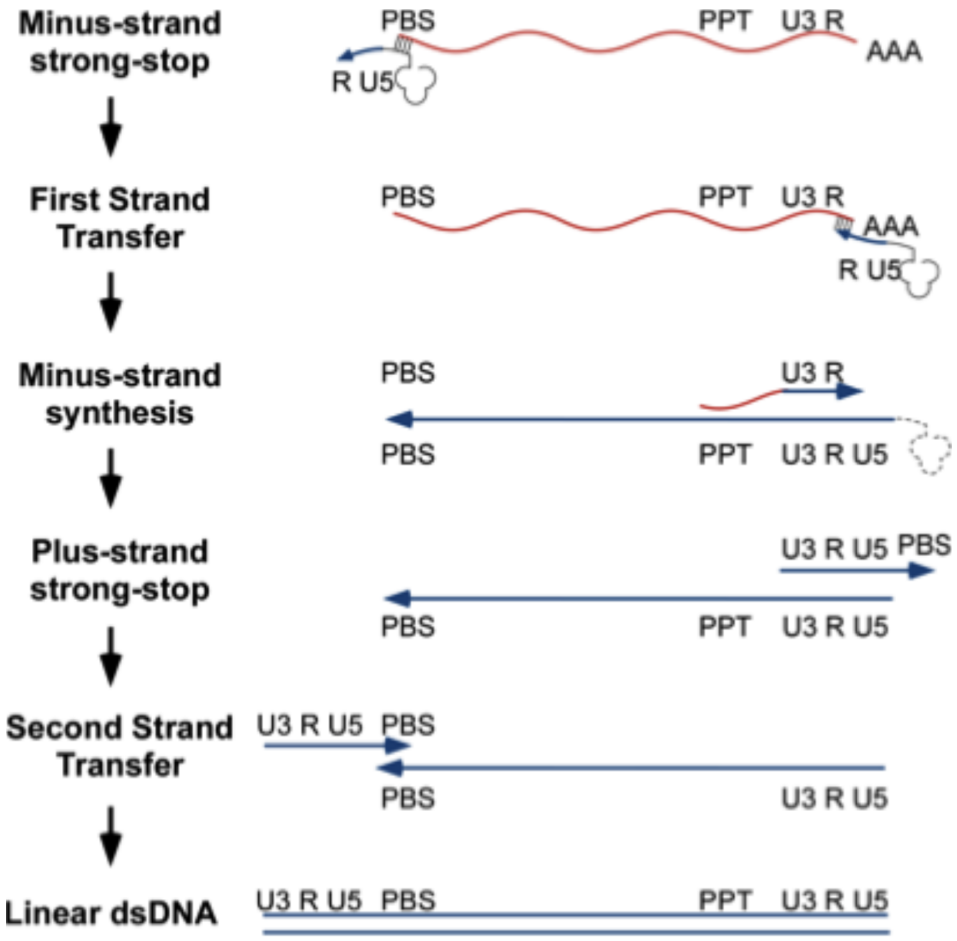


Figure 1.3 Retroviral Reverse Transcription

Reverse transcription of the viral RNA (red) into a linear double-stranded DNA copy (blue) is required for integration into the host cell chromatin. A specific tRNA is pre-bound to the primer binding site (PBS) near the 5'end of the viral genome. DNA nucleotides are added in a reverse transcriptase dependent manner to the tRNA primer generating the minus-strand strong-stop DNA (-sssDNA). The first of two strand transfers allows the newly generated DNA/RNA to move from the repeat sequence "R" at the 5' end to the R sequence at the 3' end of the RNA. The Rnase H function of RT is required for degradation of the viral RNA and cellular tRNA allowing strand transfers to occur. The polypurine track within the viral RNA is somewhat resistant to Rnase H degradation so it remains as a primer that allows for the synthesis of the plus-strand strong-stop DNA (+sssDNA). Next the second strand transfer occurs: the PBS of the minus strand and plus strand anneal and RT synthesizes the rest of the double stranded DNA (dsDNA) completing the reverse transcription process.

1.2.2.1 MLV Transcription

Transcription from the MLV provirus, like cellular gene transcription, is regulated by a complex set of cellular factors. Many cellular factors regulate transcription through binding sites in the U3-enhancer region within the long terminal repeat (LTR) region of the viral genome (Figure 1.4). The U3-enhancer region of the Moloney MLV (Mo-MuLV) LTR is made up of two 75bp-repeats, which contain numerous DNA binding sites of conserved sequences that can bind a variety of transcriptional regulators including core binding factor (CBF), mammalian type-C retrovirus enhancer factor (MCREF-1), E-twenty six family members ETS-1 and ETS-2, and GA-binding protein (GABP) [33-36]. Typically these transcription factors function as transcriptional enhancers, increasing levels of the Mo-MuLV transcripts. Studies have shown that many of the DNA binding sites for these transcription factors are overlapping and often they show cooperativity in their functions as well [37]. The role of transcription factors in cell tropism and disease phenotype will be discussed further in section 1.3.

1.2.2.2 MLV Pathogenesis

MLV pathogenesis varies in both latency and disease specificity depending on the strain of virus and genetic background of the host organism [38]. In mice, Friend MLV commonly causes erythroleukemia, however Graffi MLV typically induces myeloid leukemia. Mo-MuLV typically leads to thymic lymphoma by transforming immature T-cells, however, Mo-MuLV can also transduce B-cells and cause either T-cell or B-cell multicentric lymphoma that does not affect the thymus [39].

To study Mo-MuLV pathogenesis, inbred laboratory mice are typically infected with Mo-MuLV through intra peritoneal (i.p.) or intra venous (i.v.) injection 1-3 days after birth and have a 100% incidence of tumor development. Mice become viremic approximately 2-4 weeks post infection. However, they do not begin to develop tumors until 3-6 months post infection [40].

There are a number of pre-leukemic events that typically occur before leukemia/lymphoma is induced including splenic hyperplasia and the production of mink cell focus-inducing (MCF) viruses [41, 42]. MCF viruses are generated by the recombination of an endogenous polytropic virus with Mo-MuLV. These recombinant viruses are found in most tumors induced by exogenous MLVs and the absence of MCF virus production is associated with resistance to MLV induced leukemias. In a study by Hung Fan and colleagues, Mo+PyF101, an Mo-MuLV variant with an enhancer from the F101 strain of murine polyomavirus inserted in the U3 region was used to study pre-leukemic events. Their work

demonstrated that establishing early infection in the bone marrow is essential for activation of MCF viruses and efficient leukemogenesis [43].

It is thought that non-acute transforming retroviruses require multiple integrations to activate oncogenes and disrupt tumor suppressors in order to induce a tumor. Studies in transgenic mice that overexpress oncogenes that are often activated in Mo-MuLV induced tumors have demonstrated a more rapid induction of tumors in the presence of Mo-MuLV [44]. This is thought to be true because the presence of an activated oncogene greatly increases likelihood that a second oncogene will be activated by the virus within the same cell [40].

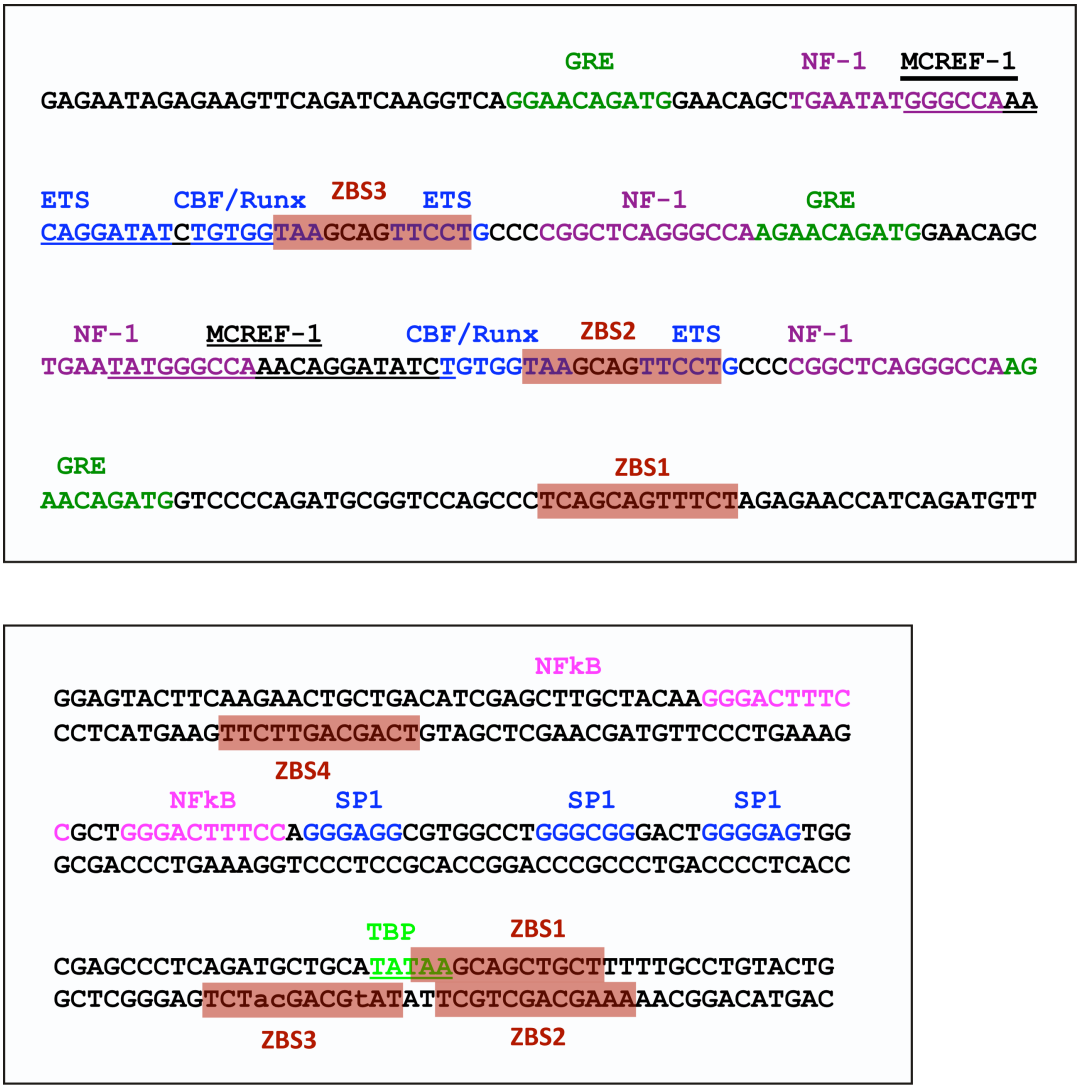


Figure 1.4 HIV-1 and Moloney MLV Promoters

Retroviral promoters are highly conserved and densely packed with transcription factor binding sites. Top Panel: Binding sites in the Mo-MuLV promoter including the three known ZASC1 binding sites [45]. Bottom Panel: 4 potential ZASC1 binding sites have been identified in the HIV promoter including one pair that is partially palandromic (Bruce, unpublished). Additional known transcription factor binding sites are highlighted. Although the ZASC1 binding sites in the HIV genome are primarily located adjacent to the TATA box, the MLV ZASC1 binding sites are located further upstream within the viral genome.

1.2.3 Human Immunodeficiency Virus

HIV is the most widely studied human retrovirus because it is the causative agent of Acquired Immunedeficiency Syndrome (AIDS), a disease that has affected over 65 million people and killed more than 25 million worldwide [46]. Geographically, the highest proportions of people living with HIV are in Sub-Saharan Africa, although there are millions of infected individuals in Asia, the Americas and Europe as well [46]. There is no cure for individuals infected with HIV, however a myriad of antiretroviral drugs have been developed over the past 30 years since the initial observation of AIDS in the early 1980s [47]. The benefits and challenges of current antiretroviral drugs will be discussed further in Section 1.2.3.1.

Since its discovery, much has been learned about HIV-1 and why this virus is so difficult to overcome. Challenges include sequence diversity, viral integration and the establishment of reservoirs of latently infected cells. The infidelity of reverse transcriptase will be discussed below. HIV-1infects cells within the immune system, integrating into host cell chromatin thereby linking removal of the virus to destruction of cells that fight virus infection. Additionally, infection of immune cells allows HIV-1 to establish reservoirs in long-lived cell types.

The host cell factors required for HIV-1 entry into cells have been identified and characterized. CD4 is the primary HIV receptor and main determinant of cell tropism [48, 49]. CD4 binds the viral envelope glycoprotein,

gp120 [50], on the viral envelope causing a conformational change promoting gp120 interaction with one of the primary HIV-1 co-receptors: the 7-transmembrane chemokine receptors CCR5 or CXCR4 [51-57]. Although additional co-receptors have been described it is believed that they have a limited role *in vivo* . CCR5 is the primary co-receptor for transmission of the virus [58], however, the virus may evolve to produce CXCR4-tropic virus in infected individuals [56].

HIV has developed highly effective mechanisms to evade immune defenses. Accessory proteins nef, vif, vpu, and vpr, encoded by HIV-1, have the ability to target specific innate and adaptive immune functions. For example, vif recruits a Culin3 ligase that targets APOBEC3G for ubiquitination and subsequent proteasomal degradation [59-63]. Another cellular restriction factor, BST-2 causes viral particles to be trapped at the cellular membrane, however, this function is counteracted by vpu also through a proteasomal dependent manner [64-67]. Vpr and nef cause cell cycle arrest and down regulation of immune cell surface markers such as MHC and CD4 allowing infected cells to evade the adaptive immune response [68-79].

Virus replication is a highly mutagenic process, facilitating the emergence of drug-resistant virus variants as well as those that can evade vaccine-mediated approaches. HIV-1 reverse transcriptase is a highly error prone enzyme, introducing 3.4×10^{-5} mutations per base pair per cycle [80]. With as many as 10^{10} virions produced in a single patient per day, this genetic instability leads to production of many replication competent viral particles with mutations which can

then escape the neutralizing antibody responses [81]. Current HIV treatments and issues of virus resistance will be discussed below in sections 1.2.3.2 and 1.2.3.3.

1.2.3.1 HIV Transcription

As discussed above, HIV is a lentivirus, a genus of complex retroviruses capable of infecting both non-dividing and dividing cells. Similar to other retroviruses, HIV-1 requires cellular transcription factors to recruit RNA Polymerase II to the viral promoter. Figure 1.4 shows the binding sites of cellular transcription factors known to interact with the virus promoter and regulate transcription including ZASC1, which will be discussed further in section 1.3.2. NF-kB has multiple HIV-1 binding sites in the HIV-1 promoter. Both NF-κB and NFAT are important for acute HIV-1 infection as well as reactivation from latency [82]. Sp1 modulates HIV-1 transcription by binding to the core promoter in the LTR [83] through coordination with the TATA binding protein (TBP) and TFIID [84, 85].

In addition to host cell factors, HIV-1 also requires a viral accessory protein in order to be efficiently transcribed. When RNA Pol II initiates transcription from the HIV-1 promoter it is in a weakly processive form and requires several phosphorylation events by the pTEF-b complex to efficiently transcribe the HIV-1 provirus. The transactivator of transcription (TAT) is a virally encoded protein, expressed early after integration of the provirus. TAT was

initially shown to have a role in overcoming a transcriptional elongation block [86-88]. TAT binds and recruits the pTEF-b complex, including CDK9 and CyclinT1, to the trans-activation response (TAR) element, an RNA hairpin that forms in the 5' end of the partially transcribed viral RNA [89]. CDK9 that has been recruited to the TAR element hyper-phosphorylates the C-terminal domain (CTD) of RNA Pol II, thereby converting RNA Pol II to a highly processive form that can efficiently generate and elongated HIV-1 transcript [90]. Until recently, it was thought that TAT was recruited to the HIV-1 promoter after the production of the TAR-element, eliminating any role for TAT in the initiation of the basal transcription complex. However, D'Orso and Frankel provide evidence that a TAT/inactivepTEF-b complex is first recruited to the HIV-1 core promoter in a TAR-independent manner prior to transfer of the Tat/active pTEFb complex to the TAR element [91, 92].

1.2.3.2 HIV Pathogenesis

There are a number of different factors that contribute to HIV-1 pathogenesis including host genetic determinants, virological factors, and host immune function. Some patients infected with the virus rapidly develop AIDS while others, because they maintain high CD4+ T-cell counts even in the absence of anti-retroviral therapy, are considered long term non-progressors (LTNP). There is a great deal of research into how these patients are able to resist and tolerate the virus. In some LTNP, the virus has been attenuated

through mutations. Other LTNP express certain HLA class I molecules that are able to present a specific subset of viral antigens to cytotoxic T-cells, helping to prime the immune system to identify and destroy infected cells [93-95]. These factors can all contribute to the LTNP phenotype, however, no single mechanism seems to be responsible for the delayed onset of AIDS in LTNP.

Individuals homozygous for the CCR5-Δ32 deletion in the HIV coreceptor CCR5, are highly resistant to HIV infection. The prevalence of the Δ32 mutation varies by geographical location with highest rates in parts of Europe and western Asia [96]. In 2007, an HIV/AIDS patient suffering from leukemia was given a bone marrow transplant from an individual with the CCR5-Δ32 mutation and discontinued antiretroviral treatment. As of 2011 the patient remains virus free effectively demonstrating this as a potential cure for HIV-1 infection [97, 98]. While this is an important finding it is unfortunately not a feasible method for effectively ending the worldwide HIV-1 epidemic. Bone marrow transplant is an expensive procedure that requires blood type matching and, because the prevalence of CCR5-Δ32 homozygous individuals is so low, it would be virtually impossible to treat more than a small subset of patients by this procedure.

Typically, individuals infected with HIV-1 reach a viremic state approximately 6-weeks post-infection. This is followed by a steady decrease in CD4+ T-cells. Although patients may suffer from flu-like symptoms immediately after infection, there are no distinguishable symptoms of HIV/AIDS until later stages of disease. Once CD4+ T-cell populations have dropped below 200 per

cubic millimeter of blood, patients typically succumb to opportunistic infections such as pneumonia and tuberculosis [99].

1.2.3.3 HIV Current Treatments

The best treatment options currently rely upon inhibiting viral replication in the host. This therapy, known as highly active antiretroviral therapy (HAART) uses a combination of drugs that generally block viral reverse transcription and protease function [100, 101]. Clinical trials and case studies have demonstrated the efficacy of this treatment approach [102, 103].

Despite being effective for many individuals, there are several down sides to HAART. The drugs are expensive and may lead to the emergence of drug-resistant mutant forms of the virus [104, 105]. Furthermore, long-term use of these antiretroviral drugs has been associated with cardiovascular disease and hepatotoxicity [106]. Because the virus is able to adapt so quickly to these treatments, cellular host factors required for viral replication may be better targets for therapeutic intervention. This new paradigm aimed at determining the contributions of host factors in retrovirus replication has become the research focus of many labs.

The first therapeutic result of this approach entered the market in 2007 when Maraviroc, an antiretroviral drug targeting the CCR5 co-receptor was approved by the FDA for use in treating HIV-infected patients. Though this drug is promising, there are caveats to its success. *In vitro* studies show that resistant

mutants can form albeit after more than a month in culture [107, 108]. Remarkably, some mutations allow the virus to adapt to recognize the drug-bound coreceptor as opposed to the normal physiological form [109]. Patient trials have resulted in a number of highly variable virus mutations [110], making the mechanism of resistance more difficult to determine. More work is needed to understand how best to utilize Maraviroc in HAART and to determine if other cellular factors are better suited as therapeutic targets.

Another issue with HAART therapy is that it requires a life-long commitment to antiretroviral drugs. Virus loads are suppressed to very low levels in HAART patients, however, latent virus reservoirs remain and are reactivated in less than a week if treatment is interrupted [111].

1.3 Role of cellular transcription factors in cellular tropism and disease pathogenesis

Cellular transcription factors are important regulators of retroviral gene expression from the provirus and are important determinants of cell tropism and disease pathogenesis [2, 5, 112-114]. In its simplest form, increased transcription can affect disease pathogenesis by increasing the odds of a new provirus insertion near a proto-oncogene. Additional mechanisms have been shown to play a larger role. For example, in the case of Mo-MuLV, removal of a single copy of the 75bp direct repeat from the U3' region was shown to increase latency [5]. Minimal mutations disrupting the NF-1-, ets-, or core- binding sites have also been shown to increase latency as well [115]. As described above, Mo-MuLV typically causes T-cell lymphoma and Friend-MuLV induces erythroleukemia. By generating chimeras between these two viruses Li, *et al.* mapped the region of disease specificity to a small number of base pairs including part of the direct repeat and a small sequence beyond this. Lewis *et al.* investigated Mo-MuLV pathogenesis by making a series of mutations in the sequence of the MLV U3-enhancer region that binds core binding factor (CBF). Specific mutations in this site increased the binding affinity for CBF, and resulted in increased instances of erythroleukemia and decreased average latency [116]. Interestingly, there were other high affinity CBF binding site mutations, which increased average latency. These mutations were in the CBF binding site regions, which overlap with ZASC1 binding sites (ZBS). The ZASC1 binding sites will be discussed further in section

1.4.1. The Mo-MuLV U3-enhancer region with known transcriptional regulator binding sites is shown in Figure 1.4.

HIV-1 tropism and pathogenesis are also in part regulated by transcription. HIV-1 preferentially replicates in activated CD4+ T-cells, however, the cytopathic effect of this virus or cytotoxic effect of CD8+ T-cells typically kills virus-producing cells within a few days [117-119]. Virus transcription is primarily dependent on host cell transcription factors such as NF-κB, Sp1 and NFAT [120]. Infection of resting T-cells, activated cells that become central memory T-cells soon after infection as well as macrophages and dendritic cells leads to a long-lived latent reservoir of virus that can be activated by cytokines or other stimuli [104, 121, 122].

1.4 Zinc Finger Proteins

Zinc finger proteins represent one of the largest protein families in eukaryotes accounting for 3% of all human genes and 2.3% of mouse genes [123]. There are several types of zinc finger motifs that use a combination of cysteine and histidine residues to coordinate one or more zinc ions. The C_2H_2 zinc finger motif is considered the classic fold and is commonly found in mammalian transcription factors where it mediates the interaction with DNA. In the C_2H_2 zinc finger motif, the zinc ion is coordinated by 2 cysteine residues present in the alpha helical region and two histidine residues within the beta sheet (Figure 1.5). Particular amino acid side chains within the alpha helical region bind in the major groove of the DNA molecule and determine the binding specificity. Each individual finger can bind approximately 3 bases so the majority of zinc finger proteins contain several of these motifs which results in increased binding specificity, although there are examples of proteins with as few as 1 and as many as 37 zinc finger motifs [124-127]. In addition to binding DNA, zinc fingers can also bind RNA and proteins.

In recent years strategies have been developed to rationally design zinc finger proteins for both research and therapeutic uses [123]. By linking zinc fingers to transcriptional activators or repressors, genes can be selectively up- or down-regulated. Furthermore, zinc fingers have been designed and fused to nucleases in order to increase the efficiency of homologous recombination targeting specific genes. In fact, this has been done to specifically delete CCR5

from hematopoietic stem cells [128]. Further characterization of zinc finger proteins, their binding sites and their functions is important in helping better understand how members of this large family of proteins functions in the cell.

1.4.1 Zinc Finger Proteins In Retrovirus Replication

Based on the number of zinc finger proteins encoded in the human and mouse genome, it is not surprising that this protein family plays a role in retrovirus replication. Several zinc finger containing proteins with a variety of functions have been implicated in retrovirus infection. ZAP, an antiviral zinc finger protein can inhibit MLV reverse transcription and integration by inhibiting the accumulation of viral RNAs in the cytoplasm. This effect is mediated by a sequence located within the viral 3'UTR [129]. A second example is that of ZFP809, a 7-zinc finger containing protein expressed in ESCs that binds to the MLV PBS and recruits Trim28 to transcriptionally silence the virus [130]. Retroviruses also interact with a number of positive acting zinc finger proteins. Sp1 is a transcription factor described above in section 1.2.3.1, which is critical for HIV-1 transcription.

1.4.2 ZASC1

Zinc finger protein amplified in esophageal squamous carcinoma (ZASC1) is a 485 amino acid C_2H_2 zinc finger protein that is ubiquitously expressed in both mouse and human tissues [131]. This gene is highly conserved in mammals, birds, reptiles and amphibians, including over 93% sequence identity between the mouse and human forms of the protein. Little is known about the function of ZASC1, but sequence analysis has shown that it contains 9-zinc fingers at the C-terminus, which is split into two groups of 4- and 5- fingers (Figure 1.5A). ZASC1 has been shown to play a role in translocating alpha-N-Catenin to the nucleus through a binding site in zinc fingers 6-8 [132]. ZASC1 is also over-expressed in esophageal squamous cell carcinoma [133]. In 2011, it was shown that increased gene copy number of ZASC1 is associated with recurrent oral squamous cell carcinoma (OSCC). In addition, the authors demonstrated that knocking down expression of ZASC1 using siRNA causes a decrease in growth and colony formation in the OSCC cell line SAS [134]. ZASC1 is also associated with the histone acetyltransferase CREB-binding protein [135] an important transcriptional co-activator [136].

ZASC1 was previously described as a transcriptional repressor but a collaborative study between the Young lab and the Ahlquist lab at the University of Wisconsin has demonstrated that it acts as an activator of retroviral gene expression [45, 132]. ZASC1 was first linked to retroviral replication in a screen to identify cellular factors affecting the early stages of MLV replication [137]. In

the screen, the MLV based pRET vector [138] containing the neomycin phosphotransferase gene and an mRNA instability element was packaged and pseudotyped with the vesicular stomatitis virus glycoprotein (VSV-G) and subsequently used to transduce the functionally haploid Chinese hamster ovary (CHO-K1) cell line in order to randomly disrupt the expression of the active allele of single genes by insertional mutagenesis. G418 resistant cells that were enriched for pRET insertions within introns of cellular genes were subjected to repeated rounds of infection with a Mo-MuLV reporter virus in order to identify cells that were resistant to MLV infection. Resistant cells were single cell cloned to produce mutant clonal cell lines. The focus of Chapter 3 of this thesis is the disrupted gene identified in the insertional mutagenesis cell line 1 (IM1), which showed approximately a 10-fold decrease in reporter virus expression. Subsequently, the ability of a VSV-G pseudotyped HIV-1 vector to infect the IM1 cell line was tested. There is an approximate 5-fold decrease in reporter virus expression when transcription was driven by the HIV-1 LTR. The IM1 cell line was characterized and a single pRET insertion was identified in the promoter region of ZASC1 as described below.

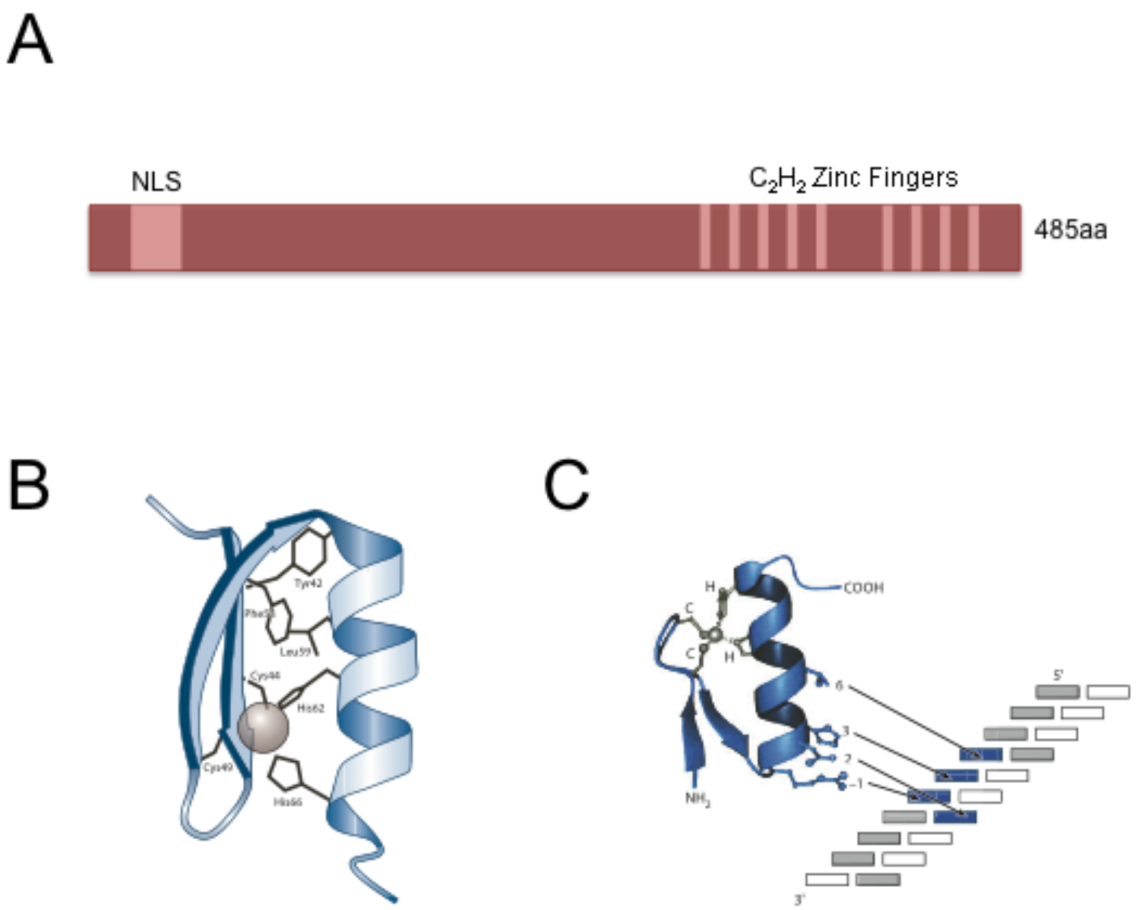


Figure 1.5 ZASC1 and Zinc Finger Proteins

A.) Schematic representation of ZASC1. ZASC1 is a 485 amino acid protein with a nuclear localization signal (NLS) near the 5' end. The 3' end encodes the 9 Cys_2His_2 zinc finger motifs. B.) A structural schematic of a classic C_2H_2 zinc finger motif. First published [139] C.) A different view of a single zinc-finger motif with specific DNA contacts. Amino acids at helical positions -1, 3 and 6 make contact with the coding strand and helical position 2 contacts the non-coding strand. First published [140]. Part B and C of this figure reprinted from [123] with permission from the publisher.

To ensure that the resistance to retrovirus infection in the IM1 cell line was due to the disruption of this gene, complementation experiments were performed using the human ZASC1 cDNA. Expression of the human ZASC1 protein in the IM1 cell line was able to fully complement the MLV infection defect [45]. In order to determine what stage of the virus life cycle was affected, qPCR analysis of viral DNA products was used to determine that ZASC1 functioned at a post-integration step such as transcription. Electrophoretic mobility shift assays (EMSA) revealed that ZASC1 binds to 3 regions of the MLV U3 enhancer (highlighted in Figure 1.4), indicating that it may have a role in transcriptional regulation. Mutational analysis of the three ZASC1 binding sites (ZBS) demonstrated that ZBS1 is the most important for MLV replication [45]. This is also the most well-conserved of the three known ZBS among all of the MLV strains (Fig 1.6). Homology searches of other retroviral promoters revealed 4 putative HIV-1 ZASC1 binding sites. EMSA and ChIP have since confirmed that ZASC1 binds to the HIV-1 promoter (Bruce, unpublished).

Over the past few years, the mechanism of how ZASC1 functions at the HIV promoter has been further characterized. Work from Jim Bruce has shown that ZASC1 interacts with pTEF-b and HIV-1 TAT, stimulating TAT transactivation of the virus. He has also shown that ZASC1 is required for efficient HIV-1 transcriptional elongation and that ZASC1 is required to recruit both pTEF-b and TAT to the HIV-1 core promoter in a TAR-independent manner (Bruce, personal communication).

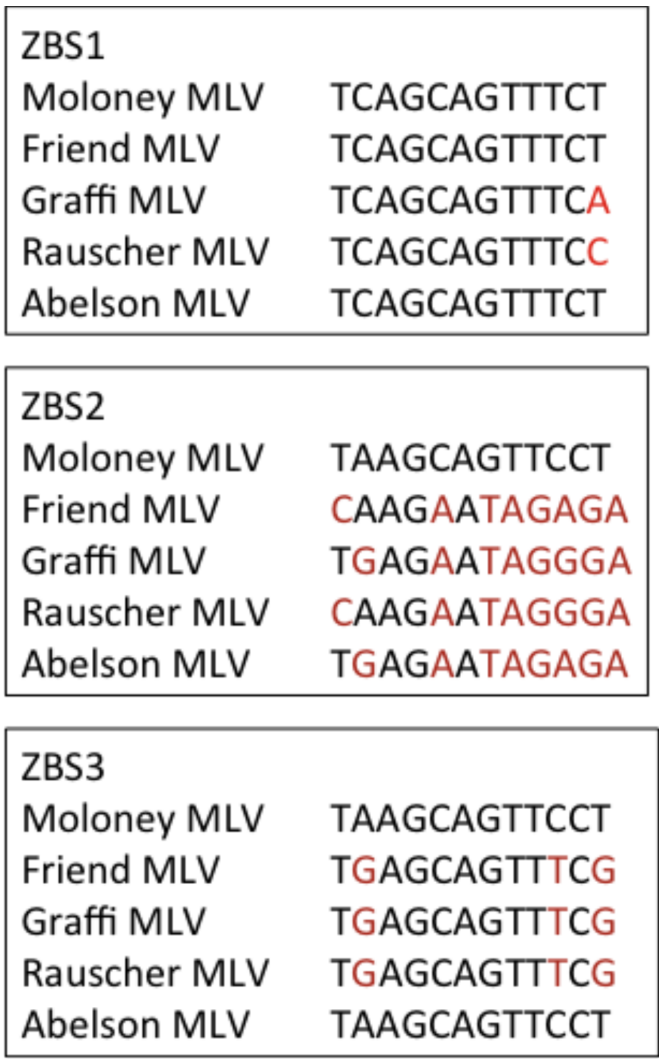


Figure 1.6 Conservation of ZASC1 binding sites in MLV

The U3 region of different MLV strains are highly conserved, however, some variations in transcription factor binding sites exist. 3 ZASC1 binding sites have been identified in Mo-MuLV. Above are the equivalent sequences to ZBS1, ZBS2 and ZBS3 in several common MLV strains.

1.5 Contribution to the Field

Host cell factors are required for retrovirus infection and understanding these mechanisms is important for our knowledge of host/pathogen interactions. Work in Chapter 2 will describe my contribution to a large-scale collaborative effort to identify and characterize novel cellular factors involved in retrovirus replication. Additionally in Chapter 3, I describe the generation of a ZASC1 knockout mouse. This work was an important first step in understanding the role this ubiquitously expressed and highly conserved transcription factor plays in the host organism. Additionally, the generation of this knockout mouse provided an important tool to further study the role of ZASC1 in retroviral infection and disease pathogenesis.

Chapter 2. Identification and Characterization of Host Cell Factors in Retrovirus Replication

2.1 Background

The HIV genome is only about 10 kb in length and encodes just 9 open reading frames. Despite its small size, this virus is able to successfully execute all of the steps required for successful replication, evade host immunity, and dramatically alter the function of its host cell. The virus can accomplish these tasks by interacting with numerous factors of the host cell, exploiting them to execute the critical aspects of its life cycle. Previous work has identified a multitude of factors, primarily through targeted studies on interactions with specific viral proteins or through detailed dissection of specific parts of the viral life cycle. Several examples are described in more detail in Section 1.2.1. By aggregating the knowledge gained from these studies, several investigators have begun to create a more complete picture of virus-host interactions, but it is clear that there are many gaps in our knowledge of these interactions [141].

The development of systems biology approaches has facilitated a new way to investigate virus-host interactions. Instead of starting from a specific viral protein or step in the life cycle, we can now use systematic genome wide approaches to probe the effects of the virus on the host cell. This was begun using microarrays to identify the ways in which the virus altered host transcription

(reviewed in [142]), but this method is limited because it is difficult to use this information to deduce the functional consequences of altered transcription. In this respect, genome wide siRNA approaches have a distinct advantage because they systematically determine which host factors are functionally required for viral replication. By mapping the defects caused by depletion of a required factor to a specific step in the viral life cycle, we are able to construct a much more global picture of virus-host interactions and identify broad networks of proteins that are involved in these interactions.

In collaboration with a team led by Renate König, John Young and Sumit Chanda, we conducted an unbiased systematic genome-wide siRNA screen to identify cellular factors involved in the early stages of the lentiviral life cycle, including uncoating, reverse transcription, nuclear import, and gene expression. We prioritized those factors that were expressed in relevant cell types and were previously shown to interact with viral proteins or complexes that bound viral proteins. Using this approach, we were able to identify several complexes that affect early steps of HIV replication.

2.2 Contribution to the Work

As mentioned above, this genome-wide siRNA screen was the collaborative effort of a number of individual labs at multiple institutions. I worked closely with Daniel Elleder, a post-doctoral fellow in the Young lab, to lead the efforts to stage the blocks to HIV-1 infection from siRNA knockdown of required

host cell factors. We used qPCR-based methods to measure viral DNA products at specific times post infection. Briefly, 293T cells were transfected with siRNAs against genes validated from the initial screen. Forty-eight hours post transfection, transfected cells were challenged with HIV(VSV-G) and 12 or 24 hours post infection total cellular DNA was collected. Quantitative PCR was used to measure early RT products and late RT products. 2-LTR circle formation (a signal of nuclear import) and integrated provirus DNA were measured by members of the Bushman Lab.

In addition to the qPCR staging, I contributed to the validation of genes identified in the initial screen. Individual siRNAs were picked from master plates to independently validate those targeting interesting host factors. Screening assays were repeated in a larger well format with siRNAs from the original screen and in some cases new siRNAs targeting the same genes. Additionally, I contributed to morphological screening to identify cytotoxic siRNAs.

2.3 Results

Published manuscript - Koenig R, Zhou Y, Elleder D, Diamond T, Bonamy G, Irelan J, Chiang C, Tu B, De Jesus P, Lilley C, Seidel S, Opaluch A, Caldwell J, Weitzman M, Kuhen K, Bandyopadhyay S, Ideker T, Orth A, Miragila L, Bushman F, Young J, Chanda S. 2008. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* 135: 49–60.

Cell

Global Analysis of Host-Pathogen Interactions that Regulate Early-Stage HIV-1 Replication

Renate König,[1] Yingyao Zhou,[6,8] Daniel Elleder,[2,8] Tracy L. Diamond,[5,8] Ghislain M.C. Bonamy,[6,7] Jeffrey T. Irelan,[6] Chih-yuan Chiang,[6] Buu P. Tu,[6] Paul D. De Jesus,[1] Caroline E. Lilley,[3] Shannon Seidel,[2] Amanda M. Opaluch,[1] Jeremy S. Caldwell,[6] Matthew D. Weitzman,[3] Kelli L. Kuhen,[6] Sourav Bandyopadhyay,[4] Trey Ideker,[4] Anthony P. Orth,[6] Loren J. Miraglia,[6] Frederic D. Bushman,[5] John A. Young,[2,*] and Sumit K. Chanda[1,*]

[1]Infectious & Inflammatory Disease Center, Burnham Institute for Medical Research, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA
[2]Infectious Disease Laboratory
[3]Laboratory of Genetics
The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA
[4]Department of Bioengineering, University of California, San Diego, La Jolla, CA 92037, USA
[5]Department of Microbiology, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA
[6]The Genomics Institute of the Novartis Research Foundation, 10675 John J. Hopkins Drive, San Diego, CA 92121, USA
[7]Hudson-Alpha Institute for Biotechnology, 127 Holmes Avenue, Huntsville, AL 35801, USA
[8]These authors contributed equally to this work
*Correspondence: jyoung@salk.edu (J.A.Y.), schanda@burnham.org (S.K.C.)
DOI 10.1016/j.cell.2008.07.032

SUMMARY

Human Immunodeficiency Viruses (HIV-1 and HIV-2) rely upon host-encoded proteins to facilitate their replication. Here, we combined genome-wide siRNA analyses with interrogation of human interactome databases to assemble a host-pathogen biochemical network containing 213 confirmed host cellular factors and 11 HIV-1-encoded proteins. Protein complexes that regulate ubiquitin conjugation, proteolysis, DNA-damage response, and RNA splicing were identified as important modulators of early-stage HIV-1 infection. Additionally, over 40 new factors were shown to specifically influence the initiation and/or kinetics of HIV-1 DNA synthesis, including cytoskeletal regulatory proteins, modulators of post-translational modification, and nucleic acid-binding proteins. Finally, 15 proteins with diverse functional roles, including nuclear transport, prostaglandin synthesis, ubiquitination, and transcription, were found to influence nuclear import or viral DNA integration. Taken together, the multiscale approach described here has uncovered multiprotein virus-host interactions that likely act in concert to facilitate the early steps of HIV-1 infection.

INTRODUCTION

Over the course of the last several decades, a number of host cell proteins that influence the early steps of retroviral replication have been defined (Goff, 2007). However, it is likely that many other cellular factors and processes are exploited by the virus during these stages, which involves uncoating steps that give rise to an active reverse transcription complex (RTC), movement of the viral preintegration complex (PIC) to the cell nucleus, and then integration of the linear viral DNA into a host cell chromosome to generate the provirus (Nisole and Saib, 2004). Recently, a genome-wide siRNA analysis revealed over 250 host cellular factors that influence HIV-1 infection (Brass et al., 2008). Notably, only a small fraction of these factors was proposed to influence the early stages of HIV-1 replication, making it likely that additional cellular factors that regulate these steps remain to be identified. Here, we present a genome-wide analysis of virus-host interactions affecting the early steps of HIV-1 infection.

RESULTS

In this study, we combined several genome-wide analytical methods to characterize the host factors required in the early steps of HIV infection. We performed genome-wide RNAi screens for genes required for infection by utilizing a single-cycle HIV-1 reporter virus engineered to encode luciferase and bearing the Vesicular Stomatitis Virus Glycoprotein (VSV-G) on its surface to facilitate efficient infection (Figure 1A). As controls, parallel screens were performed with other viral vectors encoding luciferase: (1) Moloney murine leukemia virus (MuLV) vector pseudotyped with VSV-G, and (2) an adeno-associated virus (AAV) vector (Figure S1 available online). Prior to infection, human 293T cells were transfected with an arrayed genome-wide siRNA library, which targets ~20,000 human genes. Typically, six unique siRNAs were used to interrogate each gene, with two siRNAs targeting the same gene arrayed in a single well (3 wells/gene)

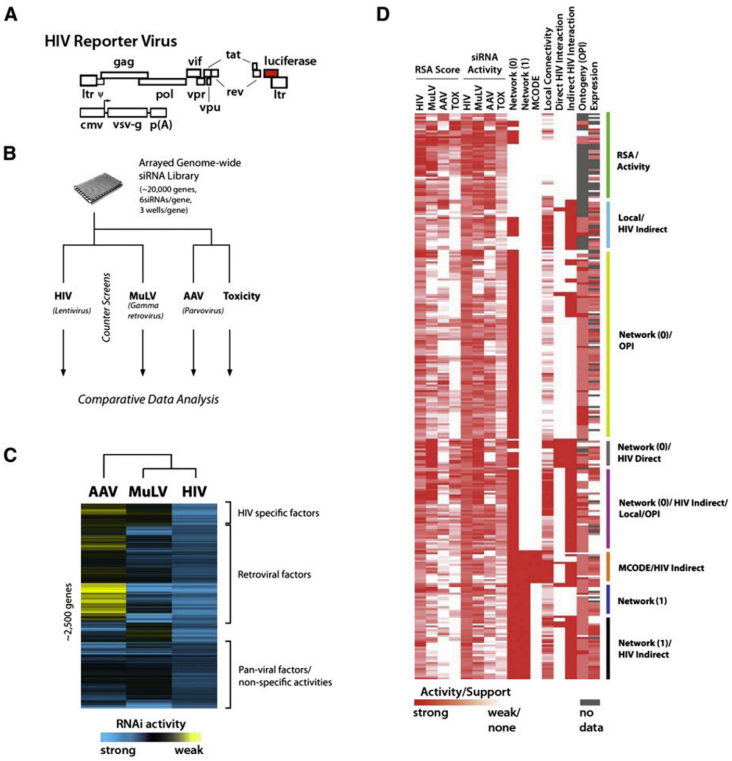


Figure 1. Integrative Analysis of HIV-Host Interactions

(A) To monitor the early stages of HIV replication, we carried out infections by using a single cycle, VSV-G-pseudotyped, HIV-1 vector encoding luciferase (pNL43R^{-}E^{-} luc (VSV-G)).

(B) Human 293T cells were transfected with an arrayed genome-wide siRNA library and then challenged with HIV (VSV-G) encoding the luciferase gene. Infection was monitored by measuring luciferase activity. Additionally, counterscreens were performed to identify those host factors affecting gammaretrovirus (MuLV) or parvovirus (AAV) infection, as well as cellular viability (toxicity).

(C) Inhibition data for each gene were normalized, scaled, and enriched for genes that have at least two siRNAs supporting their activity (RSA analysis). Comparative analysis of all viral screens, after filtering for toxicity, was visualized through hierarchical clustering.

(D) To facilitate the validation of biologically relevant functional and biochemical activities, we employed a multiscale approach to select candidate genes for further study. In addition to identifying genes based on either gene (RSA) or single siRNA activity in the genome-wide assay, an "evidence score" was compiled for each gene based upon a variety of other supporting criteria. Specifically, proteins, which were encompassed in statistically significant host-protein interaction networks (Network 0, Network 1, MCODE), or had significant network connectivity with other putative host factors (local connectivity) or HIV-encoded proteins (HIV Direct/Indirect interaction), were given additional consideration. Furthermore, genes identified through the genome-wide assay, which were members of overrepresented functional groups (ontology/OPI), or had coincident expression with CD4/CXCR4 or CD4/CXCR5 were also given additional weight. The plot depicts the activities and support evidence of 295 genes for which activities were subsequently confirmed by two or more siRNAs (Table S3).

(Figure 1B). The toxicity associated with each pair of siRNAs was also measured by assaying viable cell numbers (Figure 1B).

Since the primary screen was executed in an arrayed format, we were able to employ Redundant siRNA Analysis (RSA) to identify genes that were significantly inhibited by at least two independent siRNAs, many of which possessed reproducible, but moderate, activities (45% or greater reduction of HIV infectivity) (Konig et al., 2007). This analysis revealed that a significant fraction of host proteins required by HIV-1 is also required by MuLV (80%) (Figure 1C). These proteins included a number of factors that are already known to affect retroviral replication (Brass et al., 2008; Goff, 2007) (Table S1).

To characterize the interactions between HIV-1 and these host cell proteins, we developed an integrative approach that considered additional lines of functional, biochemical, or transcriptional data (collectively called "evidence score") for candidate gene verification. This approach is based on the assumption that a cellular factor is more likely to be a proximal regulator of viral replication if its activity is supported by multiple independent lines of evidence.

Using a yeast two-hybrid (Y2H) human protein-protein interaction database (Hynet) (Mukherji et al., 2006), we identified an extended interactome network, which contained 2458 putative host cellular factors affecting HIV infection that form 120,211 direct or indirect interactions (Network 0; data not shown). This network map was further refined by removing protein clusters, which influenced infection by the AAV vector or were associated with cellular toxicity (Network 1; $p < 0.001$) (Figure S2A). Additionally, we applied a graph theoretic clustering algorithm (MCODE) to Network 1 to identify regions of locally enhanced connectivity (Figures S2B–S2G) (Bader and Hogue, 2003). Finally, since the Hynet Y2H data set employed in these analyses did not contain virus-host protein-protein interactions, we integrated protein interaction relationships contained in the NIAID HIV-1, Human Protein Interaction Database (http://www.ncbi.nlm.nih.gov/RefSeq/HIVInteractions/), as well as newly generated HIV/host Y2H interaction data in which each of the HIV-encoded proteins was tested individually against a library of human proteins (see Experimental Procedures). The resulting host-pathogen interaction map between Network 0 proteins and HIV-encoded proteins was determined to be of high statistical significance (direct interactions, $p = 0.005$; indirect interactions, $p = 10^{-103}$, see Supplemental Data).

In an effort to identify those genes most relevant to HIV-1 infection, the mRNA expression profiles of the identified host genes, extracted from the GNF Expression database (http://symatlas.gnf.org), were also correlated to those of the viral receptor CD4 and coreceptors CXCR4 and CCR5. The majority of host proteins in Network 0 displayed strong expression patterns in tissues of lymphoid and neuronal origin (Figure S3). In a subsequent analysis for each candidate host factor, we calculated correlation coefficients with CD4/CXCR4 and CD4/CCR5 coexpression and incorporated this value into the evidence score (see Experimental Procedures). Finally, to correlate gene functions, processes, components, or domains with gene activities across the four assays previously described (siRNA screens against the three viral vectors, and cytotoxicity), we employed an ontogeny-based pattern identification algorithm (OPI) and identified over 100 enriched activity signatures across the four genome-wide screens (Figure S4; Table S2). Since OPI activity profiles are supported by multiple functionally related genes and are statistically reliable ($p < 0.05$), they were also considered as favorable evidence for the elucidation of physiologically relevant host cellular factors.

We combined these lines of evidence into a decision matrix to prioritize those factors that are likely regulators of HIV-1 replication. Criteria included gene-based (RSA score) and siRNA-based activity, gene expression signatures, gene ontology data (OPI), cellular protein-protein interaction (using Network 0, Network 1, MCODE, and local connectivity), and virus-host interaction data (HIV Direct/Indirect) (Figures 1D and S2; Table S2; data not shown). This enabled us to identify ~800 strongly supported genes based on evidence scores; 295 of these genes have been confirmed to inhibit HIV (VSV-G) infection by more than 45% with at least 2 siRNAs while not significantly affecting cellular viability (Table S3). Most confirmed factors showed broad expression in cell types relevant to HIV pathogenesis (Figure S3), and, in addition, 75 of the confirmed factors share highly coincident expression patterns with CD4/CXCR4 or CD4/CCR5 (Figures 2A and 3).

Next, we reanalyzed these confirmed factors to identify overrepresented biological annotations or protein families based on gene ontology or Interpro (IPR) databases, respectively ($p < 0.01$; Figure S5; Table S4). This analysis elucidated 176 statistically enriched biological terms, including over 25 functional processes or protein domains that influence HIV (VSV-G) infection (Figure 2B). Since functional overrepresentation was one of the criteria used to select genes for further study, many of these classifications reflected functional categories identified in the OPI analysis conducted across the four genome-wide screens (Figure S4; Table S2). Among others, a collection of genes involved in DNA repair and nucleoside diphosphate (NUDIX) hydrolase activity as well as members of the Tripartite Motif (TRIM) family of proteins were implicated as factors likely to be important for early HIV (VSV-G) infection (Figure 2C).

To visualize the global topography of HIV-host interactions, we reconstructed a network based upon the 295 HIV host cellular factors confirmed in secondary assays. Using Hynet, as well as additional protein-protein interaction databases (see Experimental Procedures), we assessed the network connectivity of these confirmed factors and additional host and HIV-encoded proteins (Figure 3A, inset; Figure S6; Table S5). This host-pathogen interaction network contained 213 functionally validated and 169 predicted HIV host cellular nodes, which were internally connected via 2291 binary protein interactions, and 318 interactions to HIV-encoded proteins ($p < 3 \times 10^{-6}$). To identify densely connected local network neighborhoods, we executed a Molecular Complex Detection (MCODE) analysis, which revealed several subnetworks with high local network connectivity (Figure 3A). Strikingly, the functional activities of these protein clusters reflect classifications identified by using the GO and OPI analysis, including the ubiquitin-proteasome pathway, DNA transcription and repair, and nucleic acid binding, further underscoring their importance in the direct regulation of early-stage HIV replication (Figures 3B–3I).

Quantitative PCR analysis of viral DNA products was performed to determine the effects of siRNAs on specific steps of early HIV-1 replication (Figure 4A, left panel). To analyze the accumulation of reverse transcription products, cells transfected with siRNAs targeting confirmed host factors were infected with the VSV-G-pseudotyped HIV-1 vector, and total DNA was extracted for analysis at either 12 hr or 24 hr postinfection. A reduction in viral DNA indicates that the targeted genes were acting either to promote synthesis or inhibit degradation. Control tests performed with inactivated virus or with an inhibitor of reverse transcription (AZT) showed strong inhibition at the expected steps of virus replication (Figure 4B, top panel).

Analysis of the siRNA activities indicates that factors affecting reverse transcription can be segregated into at least two distinct functional classes, those that are absolutely required for viral DNA synthesis (AKAP13 to SPEN) and those that appear to affect the kinetics of viral DNA synthesis (MED7 to SUMO2) (Figure 4A, right panel; Table S6). A third set of proteins was found to affect HIV-1 uncoating or reverse transcription (AP1G2 to TRIAD3), but, based upon statistical criteria, could not be placed into either functional class (Figure 4A, right panel). We hypothesize that factors that fall within the first group influence a step that is

Cell

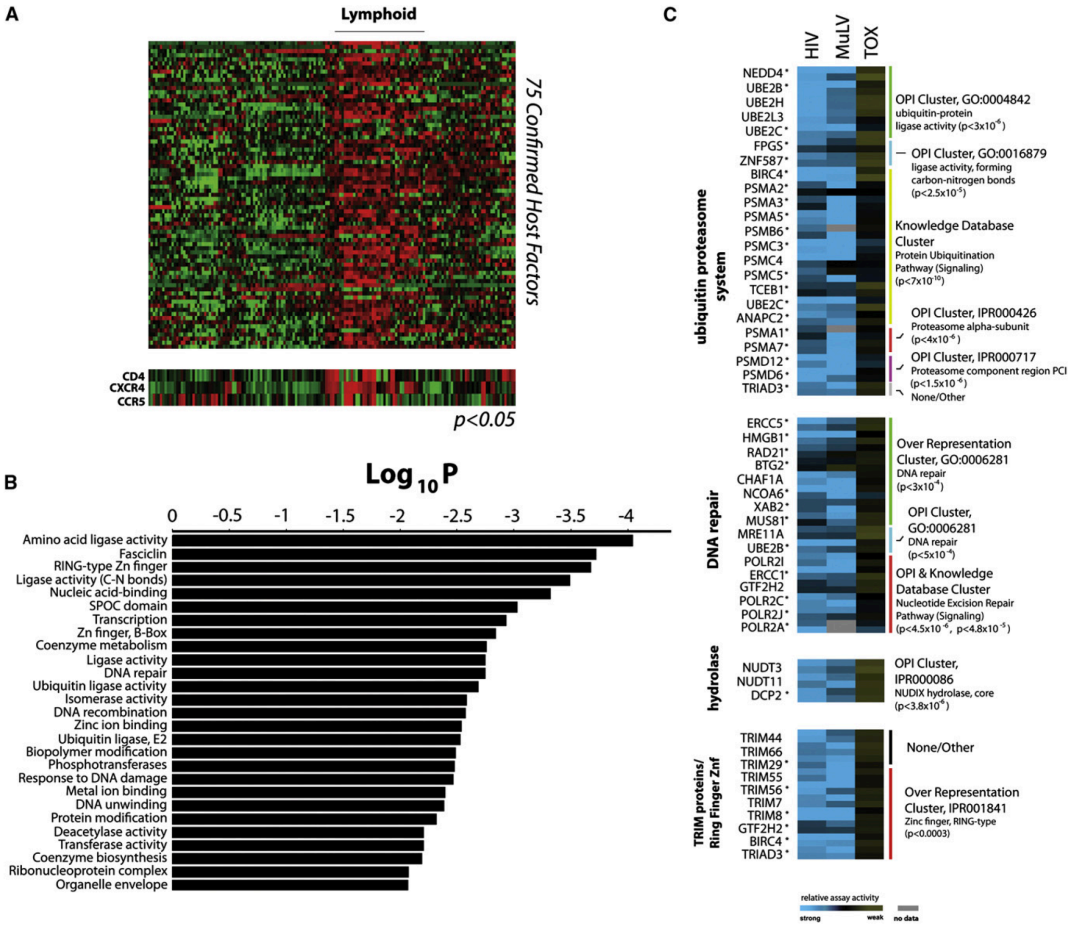


Figure 2. Characterization of Confirmed Factors Required for Infection by the VSV-G-Pseudotyped HIV-1 Vector

(A) Tissue expression of 75 confirmed host cellular factors with statistically coincident expression profiles to HIV receptor/coreceptors CD4/CXCR4 or CD4/CXCR5 ($p < 0.05$, standard chi-square test, $R_{CD4,CXCR4} \geq 0.18$, $R_{CD4,CCR5} \geq 0.13$). Tissues and cell types are shown in the x axis, while genes are displayed on the y axis, using a continuum of green (low expression) to red (high expression). mRNA transcriptional signatures were extracted from the Genomics Institute of the Novartis Research Foundation expression database (http://symatlas.gnf.org).

(B) Statistically significant overrepresentation of functional classes and protein families based upon gene ontology (GO) and interpro (IPR) domain mapping of 295 confirmed host factors required for HIV infection (Table S4). P values were calculated by an accumulated hypergeometric distribution function (Zar, 1999).

(C) Relative activities of confirmed genes, each represented by two active siRNAs, across HIV, MuLV, and cytotoxicity assays (TOX) are shown from strong (blue) to weak (yellow) (Figure S4; Table S2). Functional classes are derived from an ontogeny-based pattern identification algorithm (OPI & Knowledge Database Clusters) or GO/IPR overrepresentation analysis of primary screening data. All measurements represent the mean of at least four assays, and activities not tested or unconfirmed are depicted in gray. To identify those factors most likely to influence the intracellular steps of HIV-1 replication, a subgroup of genes was further assayed and confirmed to inhibit infection by an HIV-1 vector pseudotyped with the 10A1 MuLV envelope protein, a viral envelope protein that directs pH-independent viral entry at the plasma membrane. These genes are indicated with an asterisk (*) (also see Table S3).

important for capsid uncoating or initiation of reverse transcription, or else inhibit degradation of the viral DNA after synthesis. Conversely, we predict that the siRNAs, which alter the kinetics of reverse transcription, target cellular cofactors that facilitate, but are not essential to, the process (Figure 4A). Additional studies will be required to determine if delayed viral DNA synthesis in the absence of these genes produces PICs that are competent for integration. However, we anticipate that these kinetic factors are likely to play critical roles in maintaining viral fitness in vivo.

Additional PCR-based assays were used to identify those factors that play roles in either nuclear import of the viral PIC or viral DNA integration. The amount of 2-LTR circular forms of viral

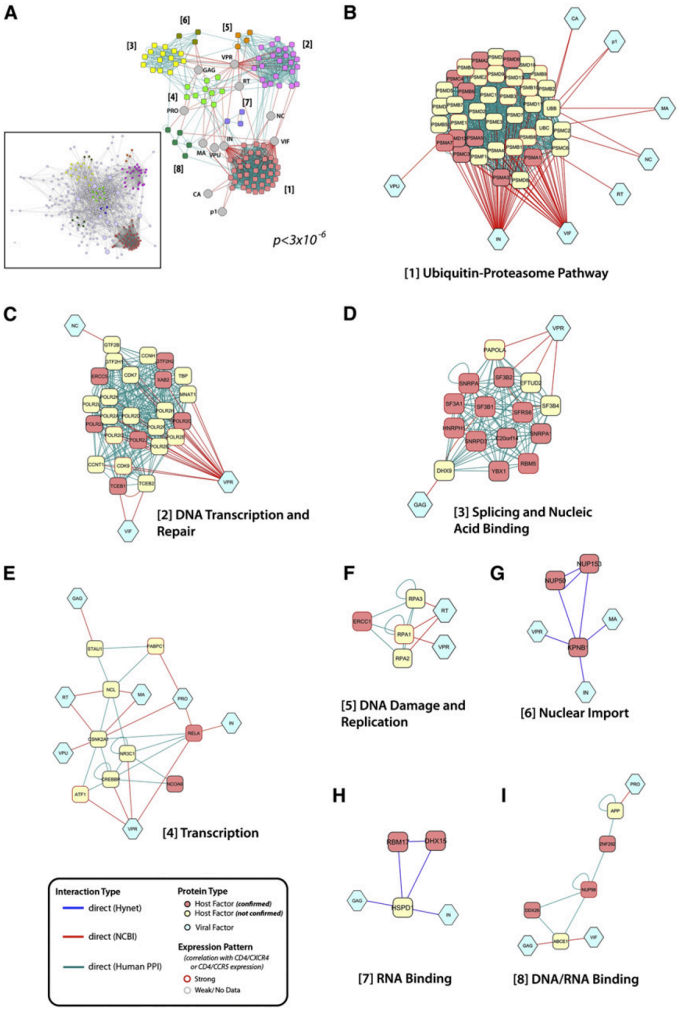


Figure 3. Network Topology of HIV-Host Protein Interactions

(A–I) The interaction network was elucidated based upon protein-protein binding data derived from the Hynet yeast two-hybrid (Y2H) database (blue connectors) and additional various human protein interaction databases (green connectors). Furthermore, connections to HIV-encoded proteins (light blue) were completed by incorporation of data from the HIV-1, Human Protein Interaction Database (NIAID) (red connectors). (A, inset) The resulting network contained 2291 direct protein interactions, connecting 213 confirmed host cellular factors, 11 HIV proteins, and 169 additional human proteins, which interact directly with at least two confirmed HIV host factors and one HIV-encoded protein (Figure S6). Using permutation testing, the density of protein interactions in this network was found to be significantly enriched ($p < 3 \times 10^{-6}$). (A) To identify potential molecular complexes, this network was analyzed for highly connected local network modules (MCODE). (B–I) Importantly, a number of these densely connected areas formed distinct functional subgroups, suggesting that they represent multiprotein complexes, which directly interact with viral factors to facilitate HIV replication. HIV-encoded proteins were abbreviated as follows: CA, Capsid; GAG, gag polyprotein; MA, Matrix; VPR, Vpr; NC, Nucleocapsid; IN, Integrase; VIF, Vif; VPU, Vpu; RT, Reverse Transcriptase; PRO, Protease; P1, p1.

DNA was quantified and served as markers of blocks either to nuclear import (reduction in 2-LTR circles) or to integration (accumulation of 2-LTR circles). In addition, viral DNA integration was quantified by taking advantage of the amplification of HIV DNA sequences adjacent to cellular Alu repeats (Butler et al., 2001) (Figure 5A). For control purposes, parallel infections were performed in the presence of an HIV-1 integrase inhibitor (Hazuda et al., 2000).

A total of 15 genes were identified that, when targeted by cognate siRNAs, significantly reduced viral DNA integration ($p < 0.001$) (Figure 5A; Table S7). The activities of all functional siRNAs targeting the same gene were combined to classify these factors into two categories based on the quantitative PCR assays. In the first category, RNAi knockdowns resulted in a reduction in integration and an increase in 2-LTR circles, indicating effects on viral DNA integration itself (ANAPC2 to IK) (Figure 5A). In the second category, little or no reduction in 2-LTR circles was seen, indicating a block to nuclear import (TNPO3 to NUP153) (Figure 5A). One factor, CHADL/LOC150356, could initially not be definitely placed, but further analysis suggests that it is likely involved in regulating viral DNA integration (Figure S7).

Several of the identified factors were members of the nucleocytoplasmic transport machinery. Unexpectedly, inhibition of these factors affected HIV-1 infection at two different steps. NUP153 and RANBP2 seemed to be involved in nuclear import of the PIC (Figure 5B, upper panel). In contrast, the nuclear import protein KPNB1 (importin β-1) as well as NUP98 seemed to be required for viral integration. (Figure 5B, middle panel). These results suggest that the karyopherin and the nucleoporin can regulate viral DNA integration at a postnuclear entry step; however, additional studies will be helpful in determining whether they act directly on viral nucleoprotein complexes or regulate a host cellular factor required for integration.

Four out of six of the cellular factors implicated in nuclear import, including NUP358/RANBP2, NUP153, and the importin-β-family protein Transportin-SR2/TNPO3, were selective for HIV (Figure 5A). TNPO3 was identified by Brass et al. (2008) as being

Cell

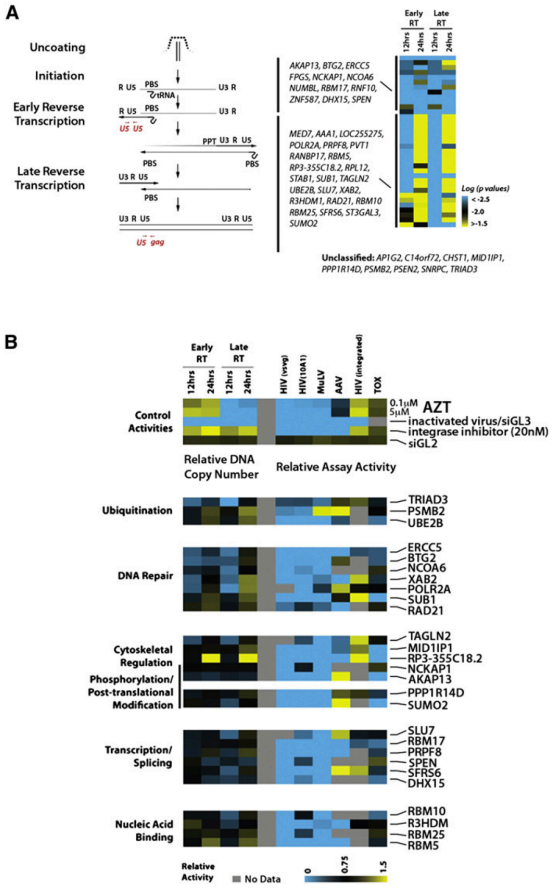


Figure 4. Host Factors Important for HIV-1 Reverse Transcription

(A) 293T cells were transfected with siRNAs targeting indicated genes and, after 48 hr, infected with pNL43R+E-luc(VSV-G). Real-time quantitative PCR amplification analysis on DNA extracts was performed at the indicated time points to assess for the amount of early and late viral RT products. The left panel depicts a schematic of the reverse transcription process and the corresponding primer/probe sets utilized for the quantitative assay (red). Multiple experiments for one of more active siRNAs targeting the same gene were considered to determine statistical significance compared to negative controls (Wilcoxon signed-rank test) (right panel). A pairwise comparison across the assays (data not shown) was also performed, and then the factors were categorized into two functional groups (indicated by the black bars). The black bars in the left panel indicate the potential steps during reverse transcription in which the two separate groups of factors are proposed to act. Additional factors that influenced reverse transcription and could not be statistically segregated into either of these two groups are indicated (unclassified).

(B) As in (A), reverse transcription was monitored with 293T cells that had been transfected with siRNAs targeting indicated host mRNAs (lower panel) or with a negative control siRNA GL2 (upper panel). For control purposes, assays were also performed with the HIV-1 RT inhibitor AZT, an HIV-1 integrase inhibitor, or with heat-inactivated virus. (Top panel) Quantitative PCR values were normalized to an internal control gene, and then to the control siRNA GL2 values. The right section of the heatmap depicts the effects of gene inhibition on various cellular assays, including infection by VSV-G or 10A1-pseudotyped HIV-1 vector, VSV-G-pseudotyped MuLV vector, or by AAV, as indicated, as well as on HIV-1 LTR-mediated transcription (HIV integrated) and on cellular viability (TOX), as described in Figure 1. Luciferase signal values were also normalized to control siRNA GL2. The median relative activity is depicted in a continuum of strong (blue) to weak (yellow). Only the genes that, upon depletion, significantly alter the initiation or rate of reverse transcription are shown ($p < 0.001$, Student's t test).

important for HIV-1 infection, and here we demonstrate its involvement in nuclear import of the PIC (Figure 5A). In contrast, all genes that influence integration also affected infection by MuLV (Figure 5A). Thus, the ability of the lentivirus HIV-1 to replicate in nondividing cells (Lewis and Emerman, 1994), as compared with MuLV, is reflected in the differences of host factors needed for nuclear import (De Rijck et al., 2007). Interestingly, the HIV PIC was unable to integrate into the host genome independently of nuclear translocation in cycling cells (postnuclear envelope breakdown), suggesting that HIV nuclear import and integration may be functionally coupled processes (see Discussion).

DISCUSSION

Here, we describe a genome-wide assay with an arrayed siRNA library to identify genes required for early stages of HIV infection. Inhibition of the gene function of putative host factors may induce cellular responses that destabilize the viral cDNA, accelerate turnover, or otherwise indirectly interfere with viral replication. To identify those proteins that are more likely to be direct regulators of the viral life cycle, and thus be critical for HIV-1 pathogenesis, we developed further selection criteria, which took into consideration a number of additional, and statistically significant, lines of evidence, including protein interactions, mRNA expression, and gene ontology (Figure 1D; Supplemental Data). Importantly, since this approach does not only rely on identifying host factors based solely on ranking siRNA activities, we were able to more effectively mine the genetic data sets to identify those factors more likely to be immediate regulators of HIV-1 infection, and establish a network of host-pathogen interactions that coordinates HIV-1 infection. Retrospective analysis showed that evidence scores of confirmed genes are, on average, 40% higher than those that did not confirm ($p = 9.7 \times 10^{-15}$), indicating that our approach enriched for biologically relevant activities (Table S3; data not shown).

Further validation studies have identified over 40 host factors that regulate capsid uncoating and reverse transcription steps of early HIV-1 replication (Figure 4). Additionally, we have elucidated 15 cellular factors that facilitate nuclear entry of the HIV-1 PIC and integration of proviral DNA (Figure 5). Importantly, only genes that regulated infection with HIV-1 virus pseudotyped with both VSV-G and 10A1 envelopes were considered for further analysis, thus excluding factors that may regulate

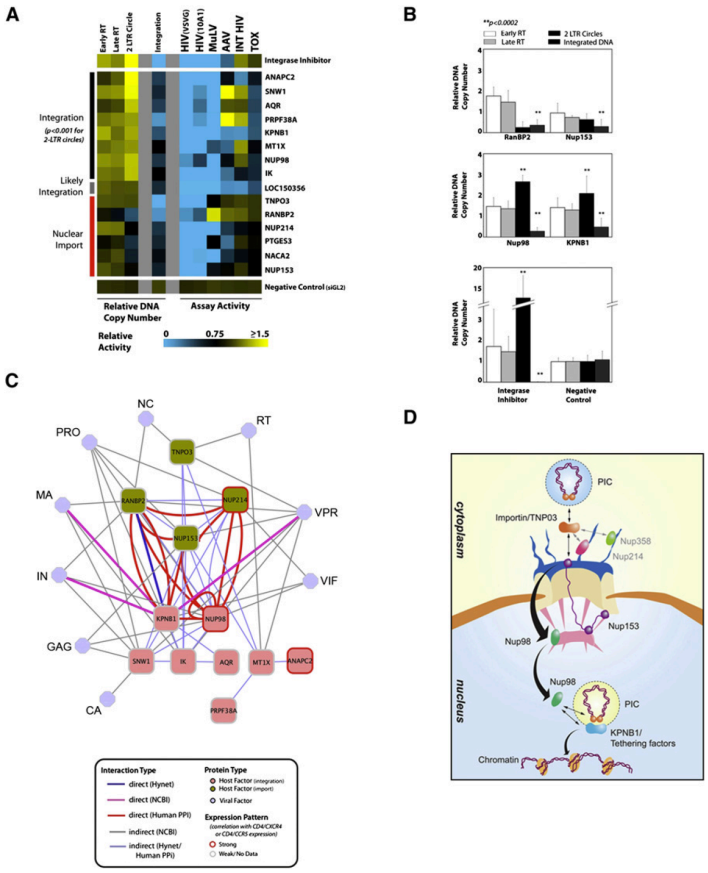


Figure 5. Host Factors Important for Nuclear Import of HIV-1 PICs and Viral DNA Integration

(A) 293T cells were transfected with siRNAs targeting indicated host factors or control siRNA (lower panel) or were treated with an HIV-1 integrase inhibitor and subsequently challenged with pNL43R$^+$E$^-$luc(VSV-G). Real-time quantitative PCR amplification analysis was then used to measure early and late RT products at 24 hr postinfection, as well as levels of 2-LTR circles and levels of integrated viral DNA at 24 and 48 hr postinfection, respectively. siRNA-transfected 293 T cells were also subjected to the same infection, proviral expression, and toxicity screens as described in Figure 4B (right panel of heatmap). The median of at least three replicate experiments for one or more active siRNAs targeting the same gene is shown in a blue-to-yellow continuum. Genes that, when inhibited, significantly impede viral integration are shown ($p < 0.01$, Student's t test). Segregation of factors into nuclear import, likely integration, and integration classes is primarily based on the level of 2-LTR circles and is also supported by additional statistical significance calculations based upon the collective activity across all PCR assays in comparison to positive (integrase inhibitor) and negative (siGL2) controls (Figure S7).

(B) Representative single siRNA activities of nuclear envelope/import-associated genes in quantitative PCR assays reflect the dual categorization of this class of proteins in both nuclear import (upper panel) and integration (middle panel). Relative DNA copy numbers of early RT, late RT, 2-LTR circle formation, and integrated DNA as well statistical significance for 2-LTR circles and integrated DNA are shown ($p < 0.0001$, Student's t test). Error bars represent standard deviations of the data.

(C) Biochemical relationships, based upon the network analysis shown in Figure 3, between proteins involved in integration (red) and nuclear import (green) and direct or indirect interactions among those proteins and with proteins encoded by HIV (blue) are depicted.

(D) A model for the molecular coupling between nuclear import of the viral PIC and viral DNA integration processes. Proteins were organized as predicted from the protein interaction data in (C) and are oriented on the basis of the quantitative PCR data (A and B).

endosomal function associated with VSV-G-mediated entry (see Figure 2C legend). Taken together, these studies indicate that host cellular factors are involved in a variety of different cellular processes that influence HIV-1 reverse transcription, nuclear import, and integration.

Comparison with Reported HIV Host Factors Identified through RNAi-Based Functional Screening

A recent genome-wide RNAi analysis by Brass et al. (2008) has identified ~284 genes as host cellular factors required for HIV replication. Comparison with the 295 confirmed genes presented here reveals a modest, but statistically significant, overlap of 13 genes ($p = 0.00021$). We speculate that this moderate concordance is largely due to differences in the analysis and experimental methodologies used. For example, the criteria Brass et al. (2008) employ to report host cellular factors are genes targeted by one or more siRNAs with activities > 2 standard deviations from the mean. Since ~154 of the genes in the Brass et al. study were supported by the activity of only a single siRNA, it is likely that a fraction of these reported host factors represent false-positive readouts due to off-target RNAi activity (Echeverri et al., 2006).

If we apply the criteria used by Brass et al. to the data presented here, we can identify 60 genes that the two RNAi studies have in common ($p = 0.024$; Table S8). Further reinforcing the parallels, we find that, based upon protein network analysis, an additional 64 genes reported by Brass et al. directly interact with a confirmed gene in our study ($p = 0.019$; Table S8). In our study, we have also prioritized gene activities not only based upon siRNA activity, but also based on comparative activities in additional screens, HIV-host protein interaction data, as well as gene expression and ontology analysis. We anticipate that this approach enabled us to enrich for the most relevant host cellular factors that promote HIV infection, but this also likely

contributed to the observed differences between the two host factor data sets.

Several experimental differences must also be considered when comparing these studies (also see Table S8). Two independent RNAi libraries, constructed by using separate design criteria and arrayed in different formats, were used to conduct these screens. Also, we have employed a single-cycle, replication-defective HIV vector pseudotyped with a VSV-G envelope in HEK293 cells and measured viral infection at a 24 hr time point. Brass et al. (2008) monitored the replication of a wild-type X4 strain of HIV in CD4/CXCR4-expressing HeLa cells over the course of 48 hr. Thus, we would not be able to identify factors involved in CD4-mediated viral entry, as well as host molecules that regulate late-stage HIV replication. In contrast, however, we monitored single-cycle virus infectivity at an early (24 hr) time point, which likely enabled us to elucidate a more comprehensive set of host factors specifically involved in the early stages of replication, including uncoating, reverse transcription, and integration. These also encompassed proteins that regulate the kinetics of these processes (Figure 4A). Thus, whereas false-positive activities are an inherent part of large-scale analyses, it is likely that variations in both experimental and data analysis techniques can largely account for the differences between our results and those reported by Brass et al. (2008).

The Role of Cytoskeletal Proteins in Early HIV Replication

The actin cytoskeleton was previously implicated in regulating the initiation of HIV-1 reverse transcription (Bukrinskaya et al., 1998) as well as in the movement of intracellular viral nucleoprotein complexes (NPCs) (Arhel et al., 2006). Consistently, the present study has revealed important roles in the earliest steps of HIV-1 infection for AKAP13, a RhoA-specific guanine nucleotide exchange factor (GEF) that regulates actin stress fiber formation; for NCKAP1, which associates with WAVE proteins that regulate actin nucleation/organization; and for TAGLN-2, a putative actin crosslinking/gelling protein (Table 1). Microtubules, previously shown to be involved in the intracellular movement of HIV-1 RTCs (Arhel et al., 2006; Naghavi et al., 2007), may also play a regulatory role in reverse transcription, since viral DNA levels were influenced by RP3-355C18.2, a predicted tubulin tyrosine-ligase, and MID1IP1, involved in bundling and stabilizing microtubules (Table 1). MID1IP1 was also identified recently by Elledge and colleagues (Brass et al., 2008), but its role in HIV-1 replication was not defined.

Involvement of the DNA-Damage Repair Pathway in HIV Uncoating/Reverse Transcription

Cellular DNA repair machinery has been implicated in playing roles in viral DNA integration and in the completion of viral DNA synthesis after integration (Goff, 2007). An unanticipated finding here was that proteins involved in DNA-damage response and repair also influence the initiation of reverse transcription and the accumulation of HIV-1 DNA products prior to integration (Table 1). Moreover, we have found two locally dense networks of proteins containing host factors that participate in DNA repair (Figure 3C, DNA transcription/repair; Figure 3F, DNA damage/replication). Both clusters contain multiple confirmed factors, including POLR2A, XAB2, and ERCC5, which have been mapped to early steps in the viral life cycle (Table 1). The viral interface for this host-pathogen interaction is mediated by Vpr, a component of the RTC/PIC that was recently being linked to the DNA-damage response pathway (Schrofelbauer et al., 2007). Several other DNA repair proteins, including MUS81, ERCC1, and MRE11, involved in nucleotide excision repair were also implicated in the early events of HIV-1 infection (Figure 2C).

Nucleic Acid-Binding Proteins Participate in the Early Stages of HIV Replication

Retroviral reverse transcription presumably involves the unwinding of RNA-RNA, RNA-DNA, and DNA-DNA strands, suggesting, that one or more cellular helicase may participate in viral DNA synthesis. Our results indicate that DHX15, a putative ATP-dependent RNA helicase that plays a role in pre-mRNA splicing, is important for early viral DNA synthesis in target cells (Table 1). Other nucleic acid-binding proteins, including RBM17, were also implicated in regulating HIV-1 DNA synthesis (Table 1). DHX15 and RBM17 are contained in a densely connected network (Figure 3H), supporting their status as proximal regulators of HIV-1 replication. In addition, several factors involved in transcription or mRNA splicing that regulate viral uncoating or reverse transcription were also identified (Table 1). Correspondingly, two MCODE protein complexes, which regulate transcription, splicing, and nucleic acid binding, were identified through our network analysis (Figures 3D and 3E). The latter cluster consists of 11 confirmed factors, of which 2 were required for reverse transcription, suggesting that these proteins collectively coordinate either the uncoating or DNA synthesis steps of HIV replication. It remains to be determined whether these host proteins act by interacting directly with the viral nucleic acids, or whether they regulate other host factors required for HIV infection.

The Ubiquitin-Proteasome Pathway Is Required for Capsid Uncoating and Viral DNA Synthesis

The ubiquitin-proteasome pathway has previously been associated with the early steps of HIV replication, where it acts negatively to destroy incoming viral replication complexes (Butler et al., 2002; Schwartz et al., 1998). Our studies have also revealed a positive role for this pathway. The ubiquitin ligases UBE2B (RAD6) and TRIAD3, as well as the proteasome component PSMB2, were each important for HIV-1 reverse transcription (Table 1). Network analysis also revealed that the viral Integrase and Vif proteins have multiple interactions with a cluster of proteins that function in the ubiquitin-proteasome pathway (Figure 3B), indicating that these viral factors may play a structural role in the HIV RTC to recruit the proteosomal machinery and facilitate uncoating or reverse transcription.

Posttranslational Modifications in Early HIV Replication

Previously, the cAMP-dependent protein kinase (PKA) was implicated in regulating early steps of HIV-1 replication (Cartier et al., 2003). AKAP13 is a PKA scaffold protein (Table 1) and as such may be involved in mediating PKA-dependent regulation of HIV-1 reverse transcription. Protein dephosphorylation events may also act to regulate HIV-1 reverse transcription,

Table 1. Selected siRNA Phenotypes in Early Steps of HIV-1 Replication

Gene Names	Functional Category[a]	Proposed Role in HIV Replication	Replication Block
ERCC5, BTG2, NCOA6	DNA-damage repair	DNA repair during RT	RT initiation
DHX15	Nucleic acid binding	Putative RNA helicase	RT initiation
RBM17	Nucleic acid binding (RNA splicing)	RNA binding	RT initiation
AKAP13	Regulation of cytoskeleton	PKA scaffold/actin organization	RT initiation
NCKAP1	Regulation of cytoskeleton	Actin organization	RT initiation
SPEN	Signal transduction	Wnt-signaling regulation	RT initiation
NUMBL	Signal transduction	Notch-signaling regulation	RT initiation
POLR2A, SUB1, XAB2, RAD21	DNA-damage repair	DNA repair during RT	RT kinetics or viral DNA stability
R3HDM1	Nucleic acid binding	Single-stranded RNA/DNA binding	RT kinetics or viral DNA stability
RBM5, RBM10, RBM25	Nucleic acid binding (RNA splicing)	RNA binding	RT kinetics or viral DNA stability
SFRS6, PRPF8, SLU7	Nucleic acid binding (RNA splicing)	RNA binding	RT kinetics or viral DNA stability
RANBP17	Nuclear Import	RNA binding	RT kinetics or viral DNA stability
SUMO2	Posttranslational modification	SUMOylation	RT kinetics or viral DNA stability
RP3-355C18.2	Regulation of cytoskeleton	Putative tubulin tyrosine ligase	RT kinetics or viral DNA stability
TAGLN2	Regulation of cytoskeleton	Actin organization	RT kinetics or viral DNA stability
UBE2B	Ubiquitin/proteasome	Ubiquitin ligase	RT kinetics or viral DNA stability
SNRPC	Nucleic acid binding (RNA splicing)	RNA binding	RT (unclassified)
PPP1R14D	Posttranslational modification	Phosphatase activity	RT (unclassified)
MID1IP1	Regulation of cytoskeleton	Microtubule organization	RT (unclassified)
PSEN2	Signal transduction	Notch signaling	RT (unclassified)
TRIAD3	Ubiquitin/proteasome	Ubiquitin ligase	RT (unclassified)
PSMB2	Ubiquitin/proteasome	Proteasome component	RT (unclassified)
TNPO3	Nuclear import	Translocation of PIC	Nuclear import
RANBP2	Nuclear import/posttranslation modification	Translocation of PIC	Nuclear import
NUP153	Nuclear pore component	Translocation of PIC	Nuclear import
NUP214	Nuclear pore component	Translocation of PIC	Nuclear import
PTGES3	Prostaglandin E synthesis	Translocation of PIC	Nuclear import
KPNB1	Nuclear import	Integration coupling factors	Integration
NUP98	Nuclear pore component	Integration coupling factors	Integration
ANAPC2	Anaphase promotion/ubiquitin ligase activity	Putative tethering factors	Integration
PRPF38A	Nucleic acid binding	Putative tethering factors	Integration
SNW1	Transcriptional regulation	Putative tethering factors	Integration
AQR	DNA repair	Putative integration cofactor	Integration

[a] References are contained in Table S9.

as indicated by the importance of protein phosphatase 1 regulatory subunit 14D (PPP1R14D), a negative regulator of the catalytic subunit of the serine/threonine protein phosphatase 1 (Table 1).

SUMOylation events have been proposed to be important in the early steps of MuLV infection (Yueh et al., 2006). We have found that SUMO-2, one of three small, ubiquitin-related modifier proteins, is important during the late stage of HIV-1 (and MuLV) reverse transcription (Table 1). In addition, RANBP2, a SUMO-1 E3 ligase that is a component of the cytoplasmic filaments of the nuclear pore complex, was required for nuclear import of the HIV-1 DNA (Table 1), perhaps through the sumoylation of viral proteins in the PIC or host factors required in the PIC. Since RANBP2 influenced HIV, but not MuLV, infection (Figure 4B), different SUMO conjugating systems may be important for each of these two viruses.

Host Factors Required for Nuclear Import of the Viral Preintegration Complex

The mechanism of HIV nuclear import is controversial, with multiple proteins and nucleic acids proposed to play a role (Suzuki and Craigie, 2007). Our studies, combined with those of Brass et al. (2008), indicate the involvement of NUP153, RANBP2, and TNPO3 as factors involved in HIV-1 PIC import (Table 1). We have also uncovered roles for NUP214; the nascent polypeptide-associated complex alpha subunit 2, NACA2; and

prostaglandin E synthase, PTGES3 (Table 1 and Figure 5). The potential role of prostaglandins in HIV-1 nuclear import is particularly intriguing, because these factors are already known to regulate the import of other types of nuclear cargo (Gomez et al., 2005; Malki et al., 2005), and they may represent a new therapeutic target for HIV-1 infection.

Host Proteins and Viral DNA Integration

Our studies have also revealed several cellular factors important for HIV-1 DNA integration. The first is ANAPC2, a component of the anaphase-promoting complex that promotes the metaphase-anaphase transition. ANAPC2 is a cullin protein with a role in ubiquitin ligase activity (Table 1), suggesting that ubiquitin modification and/or proteolysis of virus PIC components may play a regulatory role during the integration stage of viral DNA replication. Recent reports have suggested that passage through mitosis may promote HIV as well as MLV DNA integration (Mannioui et al., 2004; Roe et al., 1993), providing a possible role for ANAPC2. The second factor, SNW1, is a transcriptional coactivator that associates with a cyclophilin-like protein, peptidyl-prolyl isomerase-like 1 (PPIL1) (Table 1). The structure of PPIL1 resembles that of other members of the cyclophilin family, in particular Cyclophilin A (Xu et al., 2006), suggesting that cyclophilin proteins might be involved in regulating both late (integration) and early (uncoating) (Luban, 2007) replication steps. A third factor, aquarius or AQR, also regulates HIV-1 DNA integration (Figure 5A). This factor associates with XAB2, a protein involved in transcription-coupled repair (Kuraoka et al., 2008), suggesting that this DNA repair pathway might also play a role during viral DNA integration.

A Model for Import-Coupled Integration

Protein networks that are implicated in nuclear import and integration of the HIV PIC are shown in Figure 5C. This network was constructed based upon known or experimentally determined interactions between host factors and HIV-encoded proteins, and proteins are organized and oriented on the basis of the quantitative PCR mapping data (Figure 5A). Based upon these data, we propose a model wherein nuclear import of the PIC and proviral DNA integration are molecularly coupled events mediated by nucleoporins, karyopherin, and putative tethering factors (Figure 5D). Specifically, we hypothesize that nuclear transport of the viral PIC through the nuclear pore complex is mediated by soluble transport receptors, such as TNPO3, and nuclear pore components (Stewart, 2007). Next, the PIC cargo is transferred consecutively to phenylalanine-glycine (FG) repeat domains of variant nucleoporins, for example NUP358/RANBP2. Subsequently, in this model, the cargo is delivered to the FG repeats of NUP214 (anchored to the cytoplasmic ring) and NUP153 (distal ring). The FG repeats of both NUP214 and NUP153 are highly flexible and can either reach down to the nuclear basket (NUP214) or toward the cytoplasmic periphery of the central pore (NUP153) (Paulillo et al., 2005), thereby possibly acting to accelerate PIC translocation. As an alternative, we cannot rule out that RNAi silencing of nuclear import machinery components may be acting indirectly by altering the localization of protein(s) required for HIV nuclear import.

Upon traversing the nuclear pore, we hypothesize that the PIC is then released into the nucleus through interaction with the FG-containing NUP98 located near the nuclear basket, which dynamically associates with and dissociates from the nuclear pore (Griffis et al., 2004). Our data indicate that NUP98 is essential for viral integration, suggesting that NUP98 likely directs the viral PIC from the nuclear pore to the proximity of the chromatin. The intranuclear mobility of NUP98 has previously been linked to active transcription sites, possibly through direct interactions with the transcriptional machinery or with newly produced transcripts and RNP complexes (Griffis et al., 2004). Additionally, the PIC is anchored to chromatin through potential "tethering" factors. Targeted integration into active transcription units is mediated by PSIP1/LEDGF/p75 (Ciuffi and Bushman, 2006) (and references therein). Additional factors may help direct integration to transcriptionally active regions, as suggested by the finding that histone posttranslational modifications associated with expression positively correlate with HIV integration frequency (Wang et al., 2007). We have identified several proteins that are required for viral integration and are linked to transcription, signaling, and splicing (ANAPC2, MT1X, SNW1, IK, PRPF38A, and AQR). We hypothesize that these proteins might act to target the PIC to the chromatin and/or act as enzymatic cofactors for DNA integration. Our data indicate that karyopherin KPNB1 is also required for viral integration, suggesting that it may either cooperate with NUP98 to direct the PIC to transcriptionally active chromatin or, alternatively, could serve as an intermediate to couple the viral PIC to the tethering factors. These data support a model in which components of nuclear import machinery facilitate both HIV nuclear entry and integration, and they suggest that the coupling of these steps is required for the establishment of the HIV provirus.

Taken together, this integrative approach toward genome-wide host-pathogen interaction analysis has revealed cellular factors that coordinately regulate the early steps of HIV viral replication, including those required for reverse transcription and proviral establishment. To our knowledge, altogether new pathways, such as the possible involvement of Notch signaling in reverse transcription and prostaglandins in nuclear import, are suggested for the first time. An understanding of their role in cell types targeted by HIV during infection, such as helper T cells, macrophages, and dendritic and microglial cells, will be critical for characterizing the contribution of these factors toward disease progression. The development of small molecules that modulate the activity of these proteins may provide novel strategies for the treatment of HIV/AIDS, particularly since the inhibition of early stages in the viral life cycle has already proven to be therapeutically effective.

EXPERIMENTAL PROCEDURES

For detailed Experimental Procedures, see the Supplemental Data.

High-Throughput and Reconfirmation siRNA Screens

Genome-wide siRNA libraries targeting 19,628 human genes (2 siRNAs/well, ~3 wells/gene) were screened in duplicate or triplicate runs by using a high-throughput robotic system to identify siRNAs that influence the infection of human 293T cells by either a VSV-G-pseudotyped HIV-1 reporter virus encoding luciferase, by a VSV-G-pseudotyped MLV vector encoding luciferase,

or by an AAV vector encoding luciferase. In addition, the library was counterscreened to identify siRNAs that influence cell viability. siRNAs for reconfirmation were individually rearrayed in 384 wells in duplicate and were screened in at least 4 independent experiments and were analyzed as described in the Supplemental Data.

Expression Vectors and Reporter Viruses

For the HIV and MLV screens, pNL43-Luc-E$^-$R$^+$ (HIV-1 wild-type Δenv, encoding firefly luciferase GL3) vector was used to generate VSV-G-pseudotyped lentiviral supernatant, and the MLV-based retroviral vector pVGIP3, a derivative of pBabe, was used to generate a Moloney-based virus. For the AAV screen, GL3 luciferase was introduced into an ITR-flanked AAV expression vector to generate AAV particles encoding luciferase. For secondary confirmation assays, an amphotropic 10A1 envelope-pseudotyped HIV virus was generated.

Secondary Confirmation Analysis

Genes subject to staging of their effects in the HIV life cycle were also screened in parallel for their activities on HIV-1 (VSV-G), HIV-1 (10A1), MLV (VSV-G), and AAV reporter viruses as well on HIV-1 LTR-mediated transcription and cell viability in 96-well-based assays.

Viral DNA Quantitation by Real-Time PCR

HEK293T cells were transfected with siRNAs and were infected with HIV-1 (VSV-G) 2 days later. Total DNA was isolated at 12 , 24, and 48 hr postinfection (hpi). For control purposes, a reverse transcription inhibitor or integrase inhibitor was added at the time of infection. DNA was subject to real-time PCR to specifically quantify HIV Early reverse transcription (RT) products, Late RT products, 2-LTR circle forms, and Proviral DNA by using Alu-based nested PCR. To normalize the amount of DNA in each PCR assay, the copy number of the cellular gene porphobilinogen deaminase (PBGD) was quantified. With each experiment, a standard curve was generated.

Yeast Two-Hybrid Screen with HIV Proteins

A yeast two-hybrid (Y2H) screen was performed by using the entire HIV HXB2 proteome as bait and screening against a human leukocyte cDNA library cloned into the GAL4 expression vector.

Data Sets and Network Analysis

Multiple human protein-protein interaction databases, such as Y2H databases and other curated literature-based protein-protein interaction databases (Bind, HPRD, MINT, Reactome), were incorporated in the network analysis. In addition, two human-HIV interaction databases (the NIAID HIV-1, Human Protein Interaction Database and the Y2H in-house database) and the GNF Tissue Atlas Gene Expression database were used in this study. Detailed methodology of the multiscale strategy for hit selection and network analysis can be found in Supplemental Data.

SUPPLEMENTAL DATA

Supplemental Data include Supplemental Experimental Procedures, eight figures, and nineteen tables and are available with this article online at http://www.cell.com/cgi/content/full/135/1/49/DC1/.

ACKNOWLEDGMENTS

Tracy L. Diamond is a Leukemia and Lymphoma Society Fellow (Grant #5217-06). We thank U. O'Doherty and J.J. Yu for technical advice with the Alu-PCR assay and Zhang J. for excellent technical assistance; Walter Tian and Eric Lader (Qiagen GMBGH) for facilitating the acquisition of reagents used for this study; and John Hogenesch for critical feedback on the manuscript. This work was supported by a grant from the United States National Institutes of Health (1 RO1 AI072645-01 to J.A.Y. and RO1 AI052845 to F.D.B.) and the University of Pennsylvania Center for AIDS Research.

Received: May 1, 2008
Revised: June 12, 2008
Accepted: July 18, 2008
Published: October 2, 2008

REFERENCES

Arhel, N., Genovesio, A., Kim, K.A., Miko, S., Perret, E., Olivo-Marin, J.C., Shorte, S., and Charneau, P. (2006). Quantitative four-dimensional tracking of cytoplasmic and nuclear HIV-1 complexes. Nat. Methods *3*, 817–824.

Bader, G.D., and Hogue, C.W. (2003). An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics *4*, 2.

Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Xavier, R.J., Lieberman, J., and Elledge, S.J. (2008). Identification of host proteins required for HIV infection through a functional genomic screen. Science *319*, 817–824.

Bukrinskaya, A., Brichacek, B., Mann, A., and Stevenson, M. (1998). Establishment of a functional human immunodeficiency virus type 1 (HIV-1) reverse transcription complex involves the cytoskeleton. J. Exp. Med. *188*, 2113–2125.

Butler, S.L., Hansen, M.S., and Bushman, F.D. (2001). A quantitative assay for HIV DNA integration in vivo. Nat. Med. *7*, 631–634.

Butler, S.L., Johnson, E.P., and Bushman, F.D. (2002). Human immunodeficiency virus cDNA metabolism: notable stability of two-long terminal repeat circles. J. Virol. *76*, 3739–3747.

Cartier, C., Hemonnot, B., Gay, B., Bardy, M., Sanchiz, C., Devaux, C., and Briant, L. (2003). Active cAMP-dependent protein kinase incorporated within highly purified HIV-1 particles is required for viral infectivity and interacts with viral capsid protein. J. Biol. Chem. *278*, 35211–35219.

Ciuffi, A., and Bushman, F.D. (2006). Retroviral DNA integration: HIV and the role of LEDGF/p75. Trends Genet. *22*, 388–395.

De Rijck, J., Vandekerckhove, L., Christ, F., and Debyser, Z. (2007). Lentiviral nuclear import: a complex interplay between virus and host. Bioessays *29*, 441–451.

Echeverri, C.J., Beachy, P.A., Baum, B., Boutros, M., Buchholz, F., Chanda, S.K., Downward, J., Ellenberg, J., Fraser, A.G., Hacohen, N., et al. (2006). Minimizing the risk of reporting false positives in large-scale RNAi screens. Nat. Methods *3*, 777–779.

Goff, S.P. (2007). Host factors exploited by retroviruses. Nat. Rev. Microbiol. *5*, 253–263.

Gomez, P.F., Pillinger, M.H., Attur, M., Marjanovic, N., Dave, M., Park, J., Bingham, C.O., 3rd, Al-Mussawir, H., and Abramson, S.B. (2005). Resolution of inflammation: prostaglandin E2 dissociates nuclear trafficking of individual NF-κB subunits (p65, p50) in stimulated rheumatoid synovial fibroblasts. J. Immunol. *175*, 6924–6930.

Griffis, E.R., Craige, B., Dimaano, C., Ullman, K.S., and Powers, M.A. (2004). Distinct functional domains within nucleoporins Nup153 and Nup98 mediate transcription-dependent mobility. Mol. Biol. Cell *15*, 1991–2002.

Hazuda, D.J., Felock, P., Witmer, M., Wolfe, A., Stillmock, K., Grobler, J.A., Espeseth, A., Gabryelski, L., Schleif, W., Blau, C., et al. (2000). Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. Science *287*, 646–650.

Konig, R., Chiang, C.Y., Tu, B.P., Yan, S.F., DeJesus, P.D., Romero, A., Bergauer, T., Orth, A., Krueger, U., Zhou, Y., et al. (2007). A probability-based approach for the analysis of large-scale RNAi screens. Nat. Methods *4*, 847–849.

Kuraoka, I., Ito, S., Wada, T., Hayashida, M., Lee, L., Saijo, M., Nakatsu, Y., Matsumoto, M., Matsunaga, T., Handa, H., et al. (2008). Isolation of XAB2 complex involved in pre-mRNA splicing, transcription, and transcription-coupled repair. J. Biol. Chem. *283*, 940–950.

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

Luban, J. (2007). Cyclophilin A, TRIM5, and resistance to human immunodeficiency virus type 1 infection. J. Virol. *81*, 1054–1061.

Malki, S., Nef, S., Notarnicola, C., Thevenet, L., Gasca, S., Mejean, C., Berta, P., Poulat, F., and Boizet-Bonhoure, B. (2005). Prostaglandin D2 induces nuclear import of the sex-determining factor SOX9 via its cAMP-PKA phosphorylation. EMBO J. *24*, 1798–1809.

Mannioui, A., Schiffer, C., Felix, N., Nelson, E., Brussel, A., Sonigo, P., Gluckman, J.C., and Canque, B. (2004). Cell cycle regulation of human immunodeficiency virus type 1 integration in T cells: antagonistic effects of nuclear envelope breakdown and chromatin condensation. Virology *329*, 77–88.

Mukherji, M., Bell, R., Supekova, L., Wang, Y., Orth, A.P., Batalov, S., Miraglia, L., Huesken, D., Lange, J., Martin, C., et al. (2006). Genome-wide functional analysis of human cell-cycle regulators. Proc. Natl. Acad. Sci. USA *103*, 14819–14824.

Naghavi, M.H., Valente, S., Hatziioannou, T., de Los Santos, K., Wen, Y., Mott, C., Gundersen, G.G., and Goff, S.P. (2007). Moesin regulates stable microtubule formation and limits retroviral infection in cultured cells. EMBO J. *26*, 41–52.

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

Paulillo, S.M., Phillips, E.M., Koser, J., Sauder, U., Ullman, K.S., Powers, M.A., and Fahrenkrog, B. (2005). Nucleoporin domain topology is linked to the transport status of the nuclear pore complex. J. Mol. Biol. *351*, 784–798.

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

Schrofelbauer, B., Hakata, Y., and Landau, N.R. (2007). HIV-1 Vpr function is mediated by interaction with the damage-specific DNA-binding protein DDB1. Proc. Natl. Acad. Sci. USA *104*, 4130–4135.

Schwartz, O., Marechal, V., Friguet, B., Arenzana-Seisdedos, F., and Heard, J.M. (1998). Antiviral activity of the proteasome on incoming human immunodeficiency virus type 1. J. Virol. *72*, 3845–3850.

Stewart, M. (2007). Molecular mechanism of the nuclear protein import cycle. Nat. Rev. Mol. Cell Biol. *8*, 195–208.

Suzuki, Y., and Craigie, R. (2007). The road to chromatin - nuclear entry of retroviruses. Nat. Rev. Microbiol. *5*, 187–196.

Wang, G.P., Ciuffi, A., Leipzig, J., Berry, C.C., and Bushman, F.D. (2007). HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. Genome Res. *17*, 1186–1194.

Xu, C., Zhang, J., Huang, X., Sun, J., Xu, Y., Tang, Y., Wu, J., Shi, Y., Huang, Q., and Zhang, Q. (2006). Solution structure of human peptidyl prolyl isomerase-like protein 1 and insights into its interaction with SKIP. J. Biol. Chem. *281*, 15900–15908.

Yueh, A., Leung, J., Bhattacharyya, S., Perrone, L.A., de los Santos, K., Pu, S.Y., and Goff, S.P. (2006). Interaction of moloney murine leukemia virus capsid with Ubc9 and PIASy mediates SUMO-1 addition required early in infection. J. Virol. *80*, 342–352.

Zar, J. (1999). Biostatistical Analysis, Fourth Edition (Upper Saddle River, NJ: Prentice Hall).

Note Added in Proof

After acceptance of this manuscript, another study has also described a role for TNPO3 in HIV-1 nuclear import (Christ et al., 2008, Curr. Biol. *18*, 1192). That study demonstrated a direct interaction between TNPO3 and HIV-1 Integrase.

2.4 Conclusions

This publication was the result of a large multi-institutional effort that resulted in the identification of nearly 300 genes that influence the early stages of HIV-1 infection. Published protein-protein interactions were used to interrogate the siRNA screen results and reveal those cellular factors most likely to directly interact with the virus. This is an important next step in identifying and characterizing individual factors required for viral replication. If specific direct interactions can be identified between HIV and druggable genes, new therapies can be developed that target cellular factors rather than viral proteins.

Around the same time that our work was completed, two similar genome-wide siRNA screens also identified host cell factors involved in HIV replication [24, 26]. Brass *et al.* identified over 280 cellular proteins involved in HIV-1 replication, including the late stages of the life cycle using a CD4- CCR5-modified Hela cell line and a CCR5-tropic virus. Zhou *et al.* identified a similar number of factors using the same cell type and a similar virus, however, there were very few overlaps in the individual genes identified. Additionally, a genome wide shRNA screen for cellular factors involved in HIV-1 replication was completed in Jurkat cells (an immortalized CD4+ T-cell line) [143].

Despite sharing a common experimental question, there was a strikingly small degree of overlap in the specific factors identified in these RNAi-based screens [25]. This indicates that the specific methods used for each screen can dramatically affect the results. Each of these four published experiments used

different cell types, assay timing, viruses, and siRNA/shRNA libraries. However, comparison of multiple screens with a variety of different methods has allowed for the identification of the strongest acting factors, which are the best candidates for detailed mechanistic analysis. For example, TNPO3 was identified in all three of the genome wide siRNA screens [23, 24, 26] published in 2008. Follow up work on this factor has shown that it directly interacts with the HIV core and affects a step after nuclear import but before integration of the provirus [144-147].

Another advantage to having several independent screens is that it facilitates network-based analysis. Bushman and colleagues did a comprehensive analysis of these three genome-wide siRNA screens along with 6 other screens that utilized different approaches to uncover the roles of novel host factors in retrovirus infection [25]. Combined, these screens implicated over 5% of protein-coding human genes as possibly impacting HIV-1 infection. Instead of focusing on thousands of specific genes, they focus on networks or subsystems within the cell that the virus recruits to aid in the replication process. Subunits of the proteasome were prominently identified in the siRNA screens and although the mechanism is not yet clear, the proteasome is thought to play a role in the early stages of retroviral replication. Additionally the mediator complex and RNA Polymerase II were identified as important gene networks required for HIV-1 replication [25].

Work discussed in this chapter has demonstrated that the 9 open reading frames encoded by HIV have evolved the ability to facilitate an amazingly diverse network of interactions to meet its end goal of replication. This is consistent with

what we have already learned about HIV transcription, as the viral promoter is densely packed with cellular transcription factor binding sites that optimize viral expression in specific cell types and at specific times. Our results suggest that HIV is capable of similarly complex interactions not only at the transcriptional level, but also at several other steps in the life cycle such as uncoating or nuclear import. These results indicate that HIV has many different routes that will allow it to move through its life cycle under a variety of conditions.

This work may lead us to rethink the targeting of cellular factors in order to inhibit virus replication. Instead of targeting specific factors, strategic targeting of a combination of factors may be required. This work demonstrates the power genomic technologies resulting in a global view of the cellular factors required for HIV-1 infection. Future studies are now required to develop detailed mechanistic understanding of how particular proteins and networks are exploited for retrovirus replication.

Chapter 2, in full, is a reprint of the material as it appears in *Cell* 2008. Koenig R, Zhou Y, Elleder D, Diamond T, Bonamy G, Irelan J, Chiang C, Tu B, De Jesus P, Lilley C, Seidel S, Opaluch A, Caldwell J, Weitzman M, Kuhen K, Bandyopadhyay S, Ideker T, Orth A, Miragila L, Bushman F, Young J, Chanda S. 2008. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* 135: 49–60. The dissertation author was a co-author of this paper.

Chapter 3. ZASC1 Knockout Mice Exhibit an Early Bone Marrow-specific Defect in Murine Leukemia Virus Replication

3.1 Abstract

ZASC1 is a zinc finger-containing transcription factor that was previously shown to bind to specific DNA binding sites in the Moloney murine leukemia virus (Mo-MuLV) promoter and is required for efficient viral mRNA transcription [45]. To determine whether ZASC1 influences Mo-MuLV replication and virus disease pathogenesis *in vivo*, we generated a ZASC1-/- knockout mouse model. Mice lacking ZASC1 were born at the expected Mendelian ratio and showed no obvious physical or behavioral defects. Infection of neonatal ZASC1-deficient mice revealed that this cellular transcription factor is required for efficient early Mo-MuLV replication in the bone marrow, but not in the thymus or spleen. Analysis of bone marrow progenitor populations revealed a specific defect in myeloid differentiation in ZASC1-deficient mice, a result that is of considerable interest because osteoclasts derived from the myeloid lineage are amongst the first bone marrow cells infected by Mo-MuLV [148]. Despite this early replication defect, the absence of ZASC1 did not influence either the timing of tumor progression or the types of tumors that resulted from virus replication. These studies have identified a specific defect in myeloid cell differentiation in ZASC1-deficient mice, one that is associated with a block to early Mo-MuLV replication in the bone marrow compartment.

3.2 Introduction

Transcription of retroviral genomes involves numerous cellular transcription factors that bind to the unique 3' element (U3) enhancer element located in the viral promoter [33-35, 149]. These transcription factors can dictate viral cell-type tropism as well as the types of virus disease that are elicited [5, 115, 116]. For example, in the case of Mo-MuLV mutations in the core binding site primarily induce erythroleukemia instead of thymic lymphoma [115].

Previously, our group employed a forward genetic screen to identify ZASC1 (ZNF639) as a cellular transcription factor that binds three specific DNA sequences located within the U3 region of the Mo-MuLV genome [45]. ZASC1 is a 9-C_2H_2 zinc finger protein [133] that binds and translocates alpha-N-catenin to the nucleus [132] and also interacts with the CREB binding protein (CBP) [135]. We also showed that ZASC1 promotes viral gene expression in established cultured cell lines [45]. More recently we have found that ZASC1 is also required for efficient HIV-1 gene expression by a mechanism that is associated with recruitment of a viral Tat/cellular pTEFb complex to the viral core promoter (Bruce *et al.* in preparation).

To determine the role of ZASC1 in retroviral replication and disease pathogenesis *in vivo*, we have now generated a ZASC1$^{-/-}$ mouse strain. We report here that ZASC1-knockout mice exhibit an early replication-defect in bone marrow, one that is associated with a specific defect in myeloid cell differentiation.

3.3 Results

Generation and characterization of ZASC1$^{-/-}$ mice.

To address the potential role of ZASC1 in WT Mo-MuLV replication and disease pathogenesis *in vivo*, we generated a *ZASC1* conditional knockout mouse (Fig. 3.1A) introducing loxP sites flanking *ZASC1* exons 4-7 in order to delete the entire *ZASC1* open reading frame via Cre-mediated recombination. Puromycin (puro) and neomycin (neo) cassette selection cassettes were used in order to insert loxP sites. These casettes were subsequently removed by Flp-mediated F3 and Frt recombination in early breeding.

The targeting vector was inserted into 129/J embryonic stem cells and 294 G418 resistant colonies were selected. ES cell clones were screened for homologous recombination by southern blot analysis detecting a 5kb restriction fragment that was diagnostic of homologous recombination (Fig. 3.1B). 9 positive ES cell clones were subsequently screened using an internal probe at the upstream arm of the targeting vector. Of the 5 positive clones (data not shown) the two with the best ES cell morphology were chosen for injection into C57bl/6 blastocysts. Chimeric males were bred for germline transmission of the *ZASC1* allele and screened by coat color.

After germline transmission was achieved, mice were crossed to a ß-actin-Flp^{+} C57Bl/6 strain (Jackson Labs: strain 005703) in order to remove the antibiotic resistance cassettes and to generate ZASC1^{fl} mice. Finally, ZASC1^{fl}

mice were bred to a CMV-Cre C57Bl/6 strain (Jackson Labs strain 003465) to generate ZASC1$^{-/-}$ mice. Future generations of ZASC1$^{+/-}$ heterozygotes were bred to generate ZASC1$^{-/-}$ offspring. Quantitative PCR (qPCR) analysis of tail genomic DNA confirmed the complete deletion of exons 4-7 in ZASC1 from the ZASC1-/- animals (Fig. 3.1C). RT-PCR analysis of whole blood samples demonstrated that ZASC1 expression was also completely absent in ZASC1$^{-/-}$ mice (Fig. 3.1D). Deletion of ZASC1 did not cause any obvious physical or behavioral defects (data not shown). Furthermore, ZASC1$^{-/-}$ mice were born at normal Mendelian ratios and had similar weight and lifespan as compared with their WT littermates (Fig. 3.1E and data not shown).

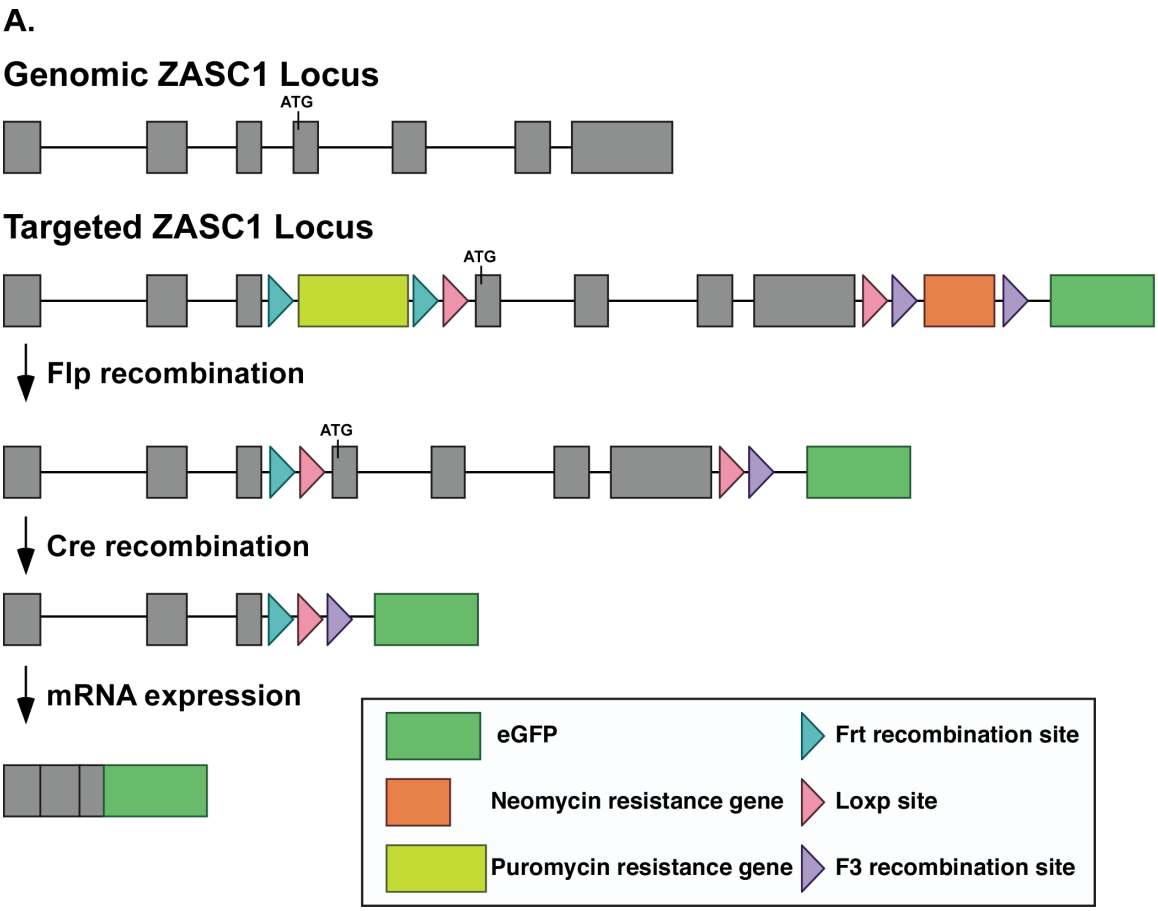
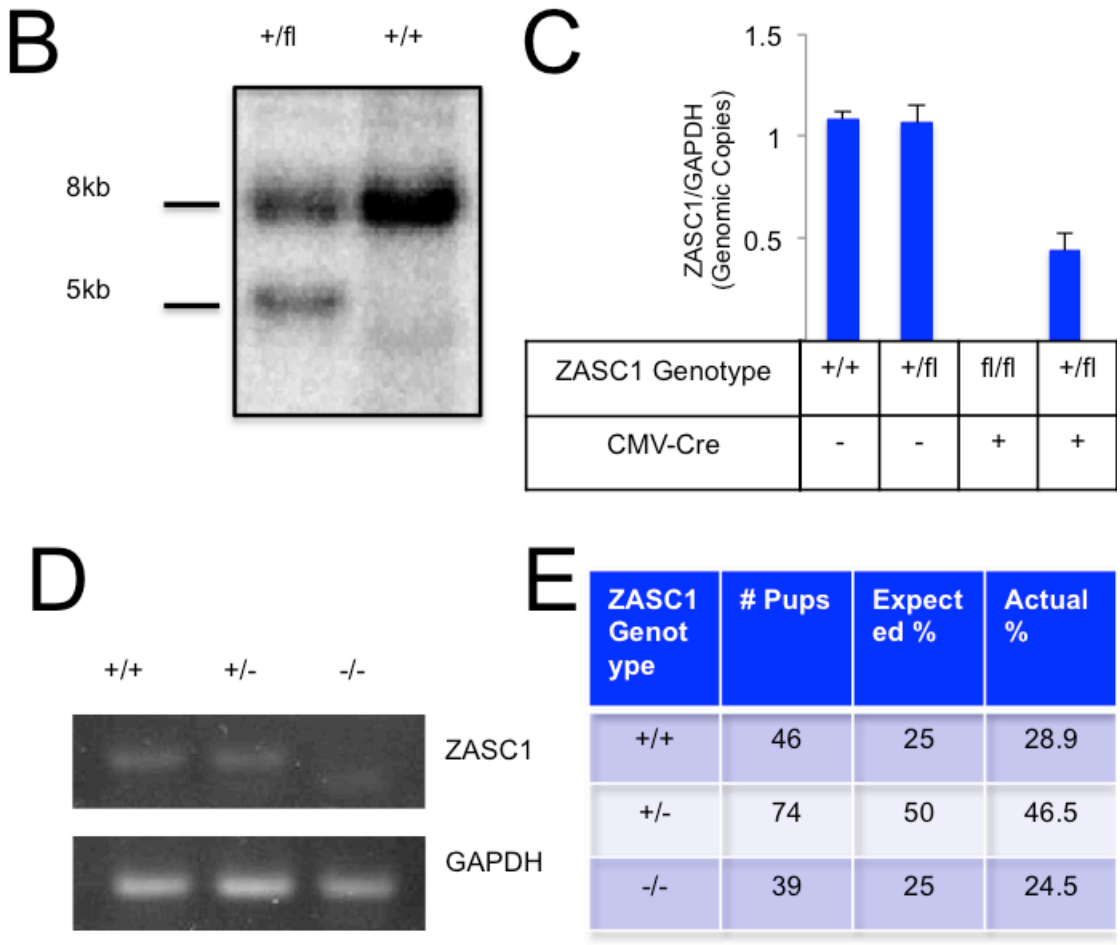


Figure 3.1 Generating the ZASC1$^{-/-}$ mouse model

A.) Schematic representation of the ZASC1 genomic locus (top) and ZASC1 locus after inserting loxp sites via cassettes containing puromycin and neomycin resistance genes. Grey boxes represent exons. Flow chart demonstrates the ZASC1 locus after Flp mediated recombination, deleting antibiotic resistance cassetes, Cre-mediated recombination, deleting the coding region of ZASC1 and splicing to generate an eGFP transcript from the native ZASC1 promoter. Efforts to visualize eGFP in ZASC1$^{-/-}$ cells were unsuccessful. Black line represents intronic or other sequence.



ZASC1 Genotype	# Pups	Expected %	Actual %
+/+	46	25	28.9
+/-	74	50	46.5
-/-	39	25	24.5

Figure 3.1 (continued)

B.) Southern blot of Sph1 digested DNA showing homologous recombination at the ZASC1 locus in genomic DNA isolated from 129J ES cells; the 8kb band is the ZASC1 genomic locus and the 5kb band is the floxed ZASC1 locus. C.) ZASC1 gene copy number measured by qPCR, normalized to GAPDH gene copy number, in animals with different ZASC1 genotypes. D.) ZASC1 RNA expression in whole blood cells from each ZASC1 genotype measured by RT-PCR. Lower band in ZASC1$^{-/-}$ lane represents primer-dimer. E.) Numbers and percentages of ZASC1$^{+/+}$, ZASC1$^{+/-}$ and ZASC1$^{-/-}$ births derived from ZASC1 $^{+/-}$ breeding pairs.

Early defect in Mo-MuLV replication in the bone marrow of ZASC1$^{-/-}$ mice.

To address the role of ZASC1 in early retrovirus infection *in vivo,* a preliminary time course of infection was performed in mixed background 129J and C57bl/6 mice with the ZASC1$^{+/+}$, ZASC1$^{+/-}$ and ZASC1$^{-/-}$ genotypes. Neonatal mice were inoculated p2-p3 by i.p. injection and viral titers were measured in the spleen, thymus and bone marrow using an established focal immunofluorescence assay [61] at various time points post infection (7, 10, or 13 days) (Fig. 2). Virus titers in the bone marrow of ZASC1$^{-/-}$ mice were significantly lower than those found in ZASC1$^{+/-}$ mice 10 days post infection (p-value = 0.03) (Table 3.1). A similar trend was noted in the spleens of ZASC1$^{-/-}$ mice at 10 days and in the bone marrow at 7 days after infection, but these differences were not significantly different from ZASC1$^{+/-}$ mice (p-value = 0.12 and 0.27, respectively).

To investigate this early defect in Mo-MuLV replication in the bone marrow in more detail, the study sample size was increased and bone marrow samples were harvested at 10 days post-infection. This analysis revealed a statistically significant 5.1-fold (p-value = 0.025) and 4.2-fold (p-value = 0.047) decrease in MLV replication in ZASC1$^{-/-}$ mice as compared to ZASC1$^{+/+}$ or ZASC1$^{+/-}$ mice respectively (Fig. 3.3A). These results confirmed that ZASC1 is required for efficient Mo-MuLV replication in the bone marrow early after infection.

Given the importance of ZASC1 for MLV replication in established cell lines and the fact that this gene is widely expressed in many different tissues [133] (including those screened for infection), it was surprising that the Mo-MuLV

replication defect was only found in the bone marrow compartment. To determine the bone marrow cell types that were associated with this replication defect at 10 days post-infection, bone marrow was pooled (typically 2-4 animals per group) by genotype (ZASC1$^{+/+}$, ZASC1$^{+/-}$, or ZASC1$^{-/-}$) and mature T-cells and B-cells were sorted from the bone marrow using CD90.2- or B220-coated magnetic beads followed by FIA analysis. Flow cytometry of a subset of cells in the complete bone marrow, T-cell, B-cell, and T/B-cell depleted samples confirmed cell purity and the efficiency of depletion (data not shown). The FIA results from each enriched cell population mirrored that seen with the complete bone marrow sample, with significantly lower Mo-MuLV titers in ZASC1$^{-/-}$ samples as compared with ZASC1$^{+/+}$ samples ($p<0.01$ for all populations) (Fig. 3.3B). Therefore, the defect seen with early Mo-MuLV replication in the bone marrow is a general one and is not confined to any one specific cell type.

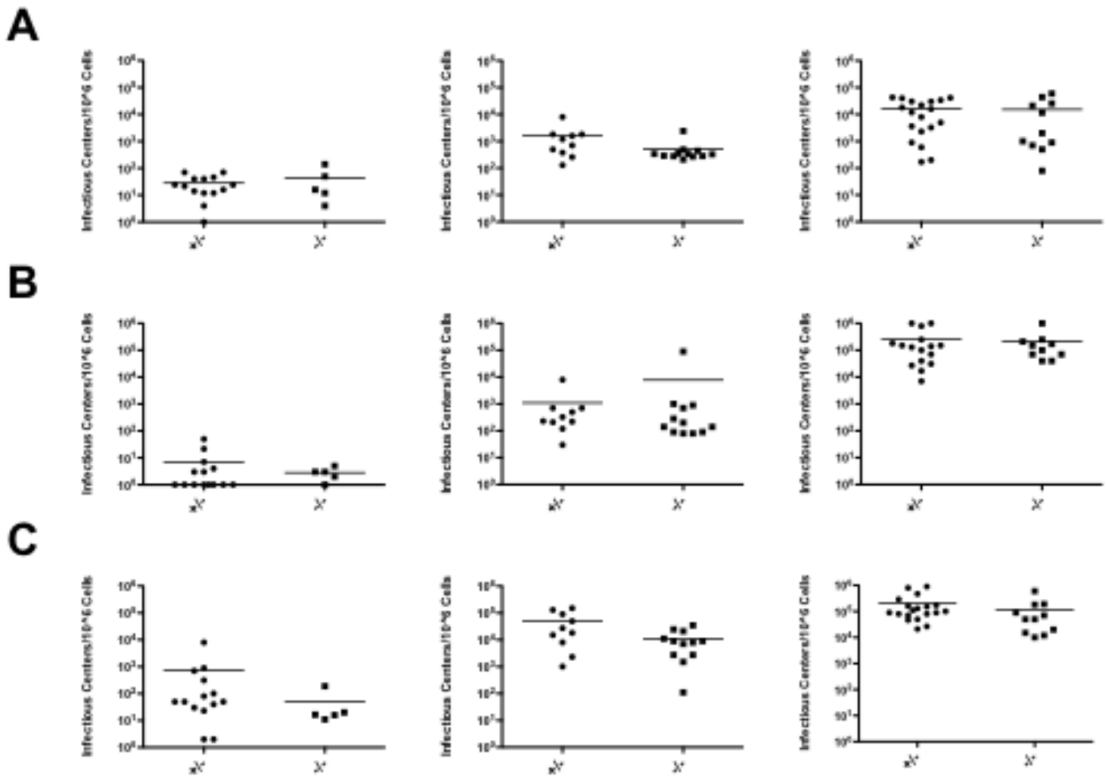


Figure 3.2 Time course of Mo-MuLV infection *in vivo*

Virus titers (infectious centers/10^6 cells) were measured from spleen, thymus and bone marrow of Mo-MuLV infected animals, collected at 7, 10 and 13 days post infection. Each dot represents one animal. ZASC1+/+ animals were not shown because the sample sizes were too small to draw meaningful conclusions. ZASC1+/- (squares) and ZASC1-/- (triangles). Mean average values indicated by horizontal bar. A.) Spleen B.) Thymus C.) Bone Marrow.

Table 3.1 Statistical Analysis of MLV Infection at Early Time Points in the Spleen, Thymus and Bone Marrow.

Fold change of MLV infections (shown in figure 3.2), comparing ZASC1$^{+/-}$ and ZASC1$^{-/-}$ mice in various organs at different time points post infection. P-values compare animals by genotype using a two-tailed t-test.

Organ	Days Post Infection	Fold Change	p-value
Spleen	7	1.1	0.64
	10	3.3	0.12
	13	1	0.99
Thymus	7	2.1	0.66
	10	0.14	0.4
	13	1.1	0.8
Bone Marrow	7	4.4	0.27
	10	2.8	0.03*
	13	1.6	0.4

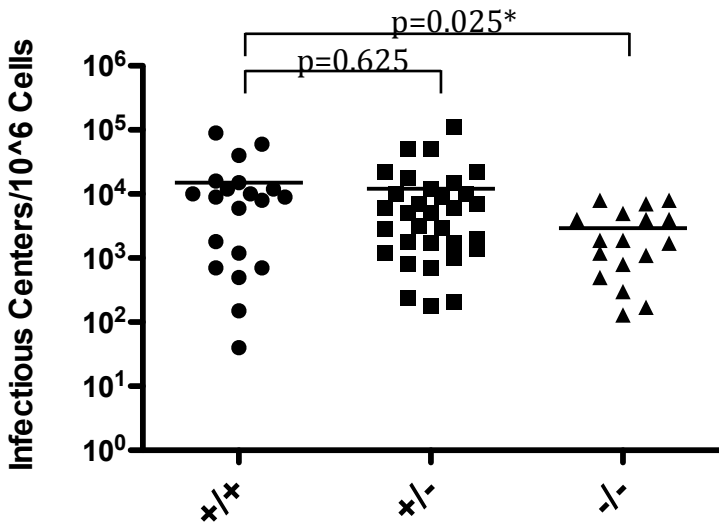


Figure 3.3 Mo-MuLV infection is delayed in ZASC1-/- mice at an early time point.

A.) The offspring of ZASC1$^{+/-}$ breeding pairs were inoculated i.p. with Mo-MuLV at p2-p3. Bone marrow from infected animals was harvested 10 days post infection along with tails to determine ZASC1 genotype. Virus infection was measured by a focal immunofluorescence assay to determine infectious centers per 10^6 cells. Each point represents a single animal ZASC1$^{+/+}$ (circles) ZASC1$^{+/-}$ (squares) and ZASC1$^{-/-}$ (triangles). Horizontal bars represent the mean average values. Statistically significant differences between different genotypes were measured by two-tailed t-test.

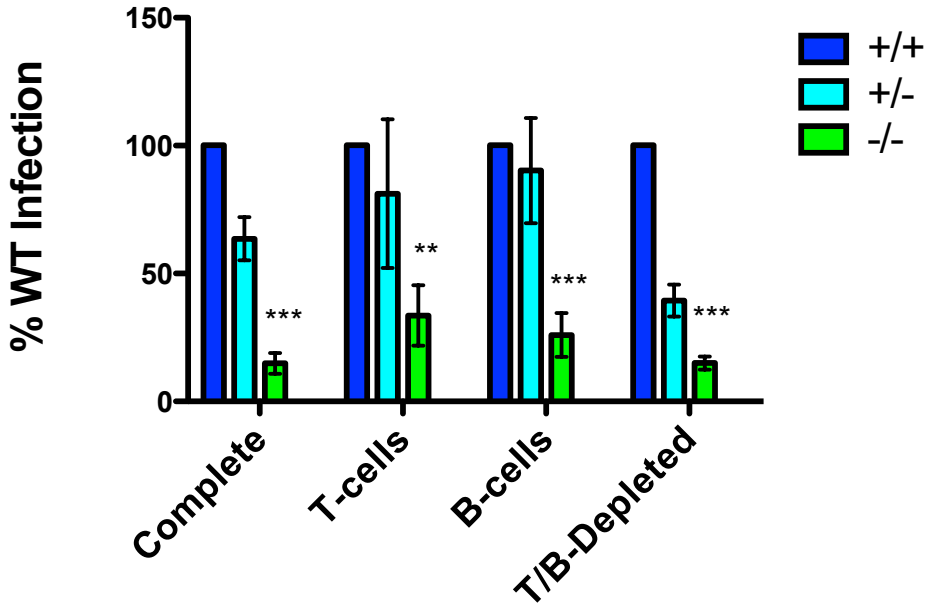


Figure 3.3 (continued)

B.) Reduced virus titers in various bone marrow cell types. Litters of pups from ZASC1 +/- breeding pairs were infected p2-p3 with Mo-MuLV and bone marrow was collected 10 days post infection. Mice were genotyped and bone marrow was pooled accordingly before magnetic cell sorting was used to purify T-cells and B-cells from the bone marrow. “Complete” is unsorted bone marrow, “T-cells” were isolated from bone marrow using α-CD90.2 conjugated Dynabeads. “B-cells” were isolated from bone marrow using α-B220 conjugated Dynabeads. Focal immunofluorescence assay was used to quantify infected cells in each population. Graph is normalized mean average levels of infection in WT cells. N=3 experiments. Error bars show SD. Statistically significant differences comparing +/+ to -/- using a two-way ANOVA are indicated: (**) $p<0.01$ and (***) $p<0.005$.

ZASC1$^{-/-}$ mice have a defect in myeloid cell differentiation in the bone marrow.

Since Mo-MuLV replication was reduced in T-cells and B-cells in the bone marrow of ZASC1$^{-/-}$ mice early after infection, we examined whether there was any obvious defect in the relative abundance of these cells in ZASC$^{-/-}$ mice. These studies revealed no deficiency in the proportion of T-cells or B-cells in the spleen and thymus of ZASC1-/- mice as compared to WT mice (Fig. 3.4 and 3.5). Moreover, T-cells derived from young ZASC1-/- mice showed no functional defect when assayed for activation in vitro (Fig. 3.6). Therefore, there was no obvious defect seen in T-cells or B-cells in ZASC1-/- mice.

Flow cytometric analysis was also used to examine the progenitor cell populations in the bone marrow. These cells were initially characterized by a lack of lineage specific markers (lin-) and expression of c-kit and sca-1 on their surface. The lin-sca+kit+ (LSK) population contained the earliest bone marrow progenitors: hematopoietic stem cells (HSCs) were identified by gating on CD105+ and CD105+ whereas multipotent progenitors (MPPs) were found in the CD105+ CD150- cell subset, as described in [150]. No differences were observed in HSC or MPP populations in ZASC1$^{+/+}$ and ZASC1$^{-/-}$ mice. The LK compartment (lin-sca-kit+) includes myeloid and erythroid progenitors (Fig. 3.7). We found that although HSC and MPP populations were comparable, there was a significant increase in the heterogenous lin^{-} ska-1^{-} c-kit^{+} (LK) compartment (Fig. 3.8). This compartment contains common myeloid progenitor cells and

downstream myeloid precursors with no long-term repopulation potential [151]. Interestingly, the myeloid progenitor cells within the bone marrow are amongst the first cells infected in the bone marrow by Mo-MuLV [148].

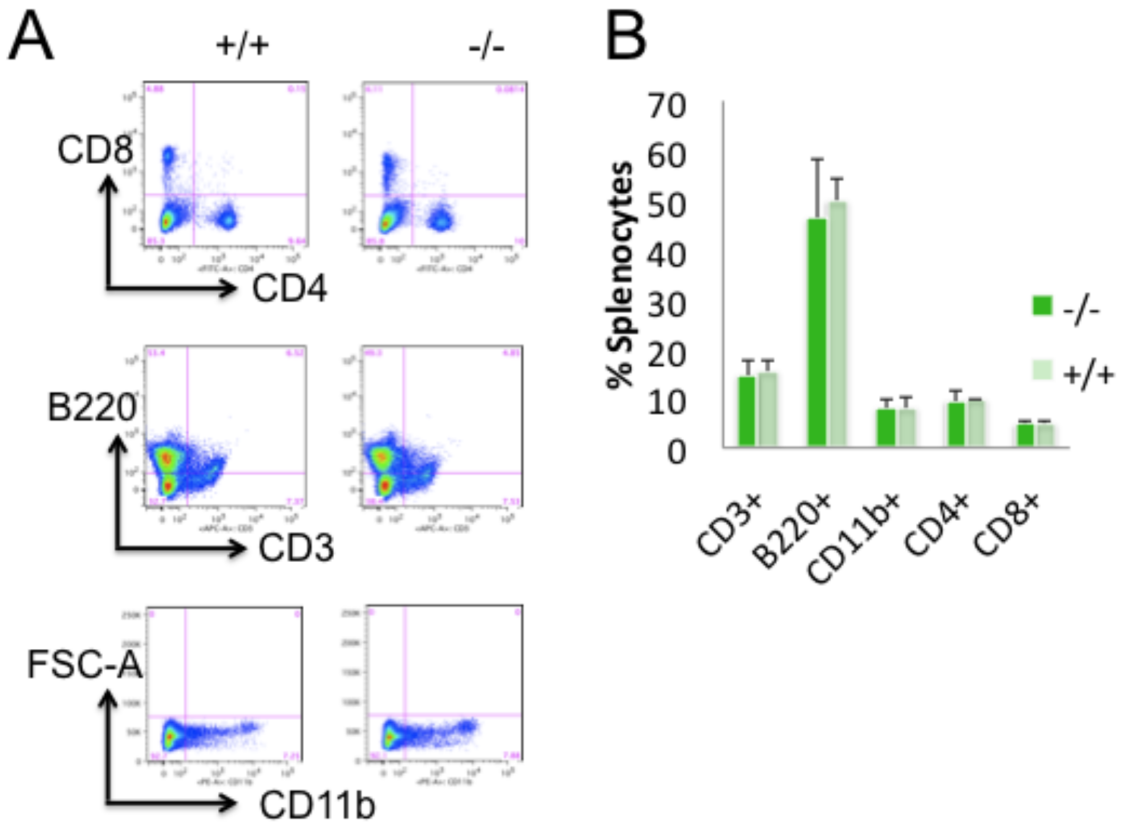


Figure 3.4 ZASC1 is not required for normal differentiation of spenocytes.

Spleens from ZASC1$^{+/+}$ and ZASC1$^{-/-}$ animals were harvested and stained using α-CD3, α-B220, α-CD11b, α-CD4 and α-CD8 fluorophore-conjugated antibodies and then analyzed by flow cytometry. A.) Example of gating for mature cell populations after FSC/SSC analysis to gate on live cells. Top panel shows T-cell subsets (CD4+ and CD8+). Middle panel shows B-cells (B220) and T-cells (CD3). Bottom panel shows monocytes/macrophages (CD11b). B. Graphical representation of the FACS plots shown in panel A. N=3 mice per group with littermate controls. Error bars = SD. No statistical significance was observed between the different animal genotypes by two-tailed t-test.

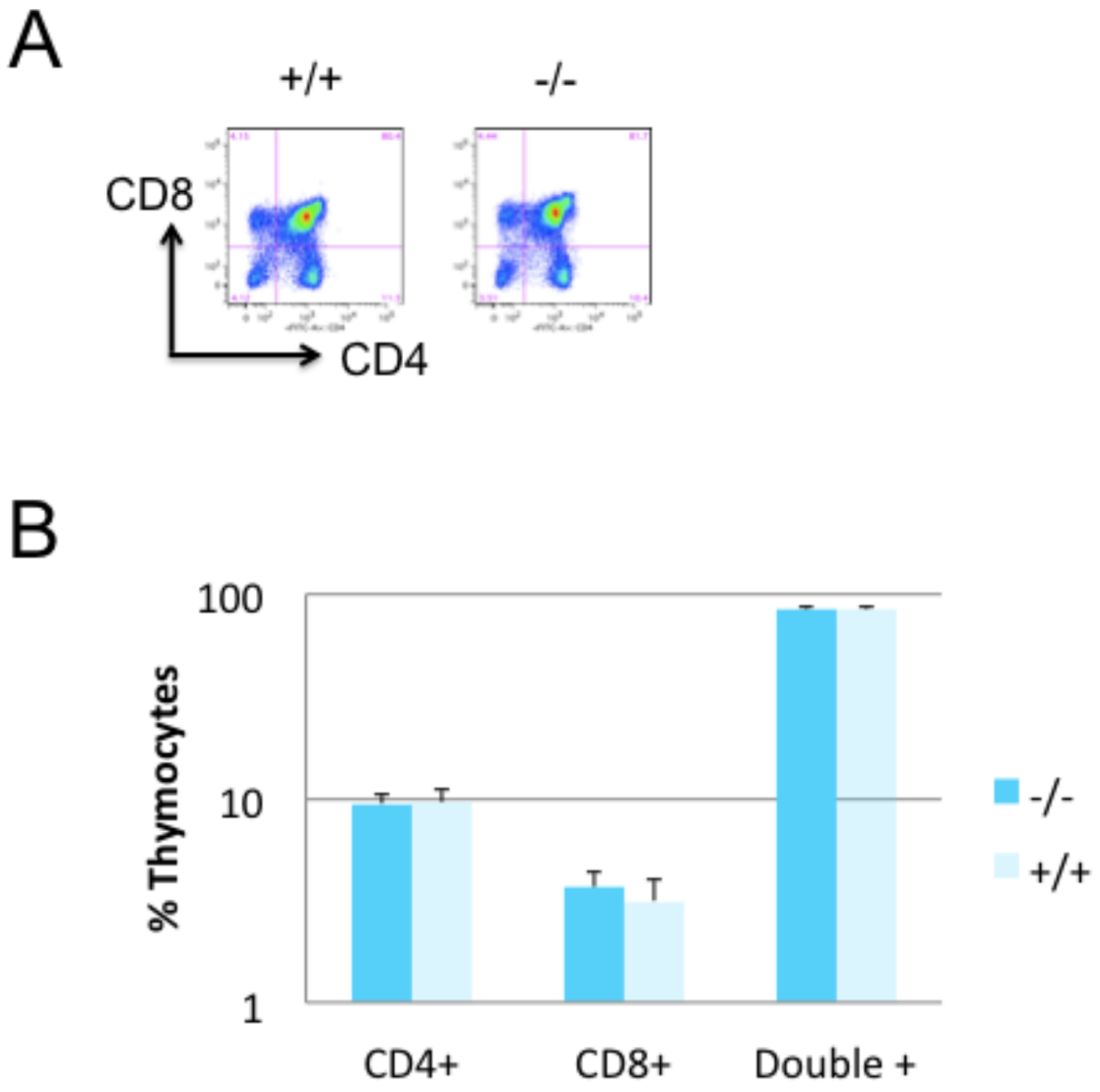


Figure 3.5 ZASC1 is not required for development of thymocytes.

Mouse thymus from ZASC1$^{+/+}$ and ZASC1$^{-/-}$ animals were harvested and stained using fluorophore-conjugated α-CD4 and α-CD8 antibodies and analyzed by flow cytometry. A.) Example of gating for CD4+, CD8+ and double positive thymocytes. B.) Graphical representation of thymocyte populations shown in panel A. N=3 mice per group with littermate controls. Error bars show SD. No statistical significance was observed between the different animal genotypes by two-tailed t-test.

ZASC1 does not affect MLV disease pathogenesis.

Since work from other labs had shown previously that bone marrow infection at early time points is critical to MLV pathogenesis [116], we wanted to address the role of ZASC1 in disease pathogenesis. Neonatal mice were infected with MLV at the p2-p3 stage and monitored for up to 1 year for tumorigenesis. Moribund animals were analyzed for tumor type, timing, and severity. These studies showed that ZASC1-deficient mice develop lethal disease at the same rate as wild type littermate controls (Fig. 3.9A). Also, immunostaining of tumor tissues with α-CD3 and α-CD79 antibodies, to indicate T-cell and B-cell lymphomas, respectively revealed that all tumors analyzed from $ZASC1^{+/+}$ $ZASC1^{+/-}$ and $ZASC1^{-/-}$ animals were derived from the T-cell lineage. Also, tumor localization was essentially the same in WT and ZASC1-deficient mice, with similar numbers of thymic and multicentric lymphomas (Fig. 3.9B). Therefore, despite the early Mo-MuLV replication defect in ZASC1-deficient animals, ZASC1 did not influence long-term virus replication or disease pathogenesis.

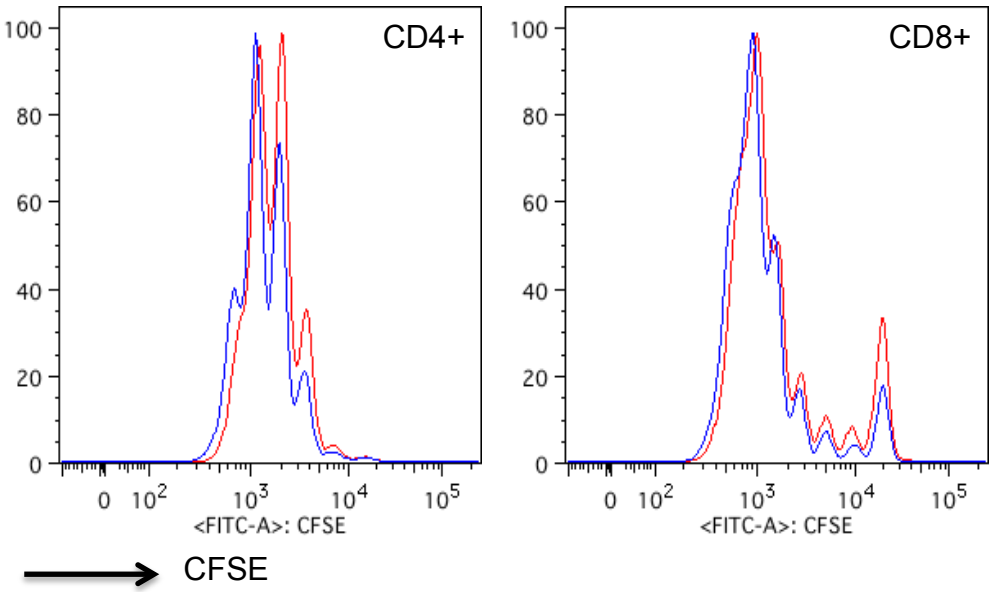


Figure 3.6 ZASC1 is not required for T-cell activation and proliferation.

CD4+ and CD8+ T-cells were isolated independently from mouse spleen and lymph nodes from 3-week-old ZASC1$^{+/+}$ and ZASC1$^{-/-}$ animals using α-CD4 and α-CD8 conjugated magnetic beads. Purified cells were labeled with CFSE and incubated in activating conditions for 4 days to induce T-cell activation and proliferation. CFSE intensity decreases by 50% with each round of cell division leading to peaks representing different levels of proliferation.

3.4 Discussion

In this study, we developed a gene knockout model to investigate the roles of *ZASC1*, a gene previously shown to be important for efficient Mo-MuLV replication in established cells [45], in normal murine development and in retroviral pathogenesis. Mice deficient in ZASC1 developed normally, had no fertility defects, and exhibited normal lymphocyte development and function. However, Mo-MuLV titers were specifically reduced in the bone marrow compartment at an early time point following infection. This defect was associated with a significant increase in a myeloid progenitor cell population (heterogenous lin^- $ska\text{-}1^-$ $c\text{-}kit^+$ or LK cells). The LK cell compartment includes the common myeloid progenitor (CMP), which eventually gives rise to megakaryocytes, erythrocytes, monocytes, granulocytes and osteoclasts. The bone marrow defect in virus replication was seen with whole bone marrow cell populations and with T-cell, B-cell, and T/B-cell depleted bone marrow samples. Our data also argue against an important role for ZASC1 in the timing or severity of MLV-induced disease pathogenesis since *ZASC1*-deficient mice developed tumors of a similar type and at similar rates as WT littermate controls.

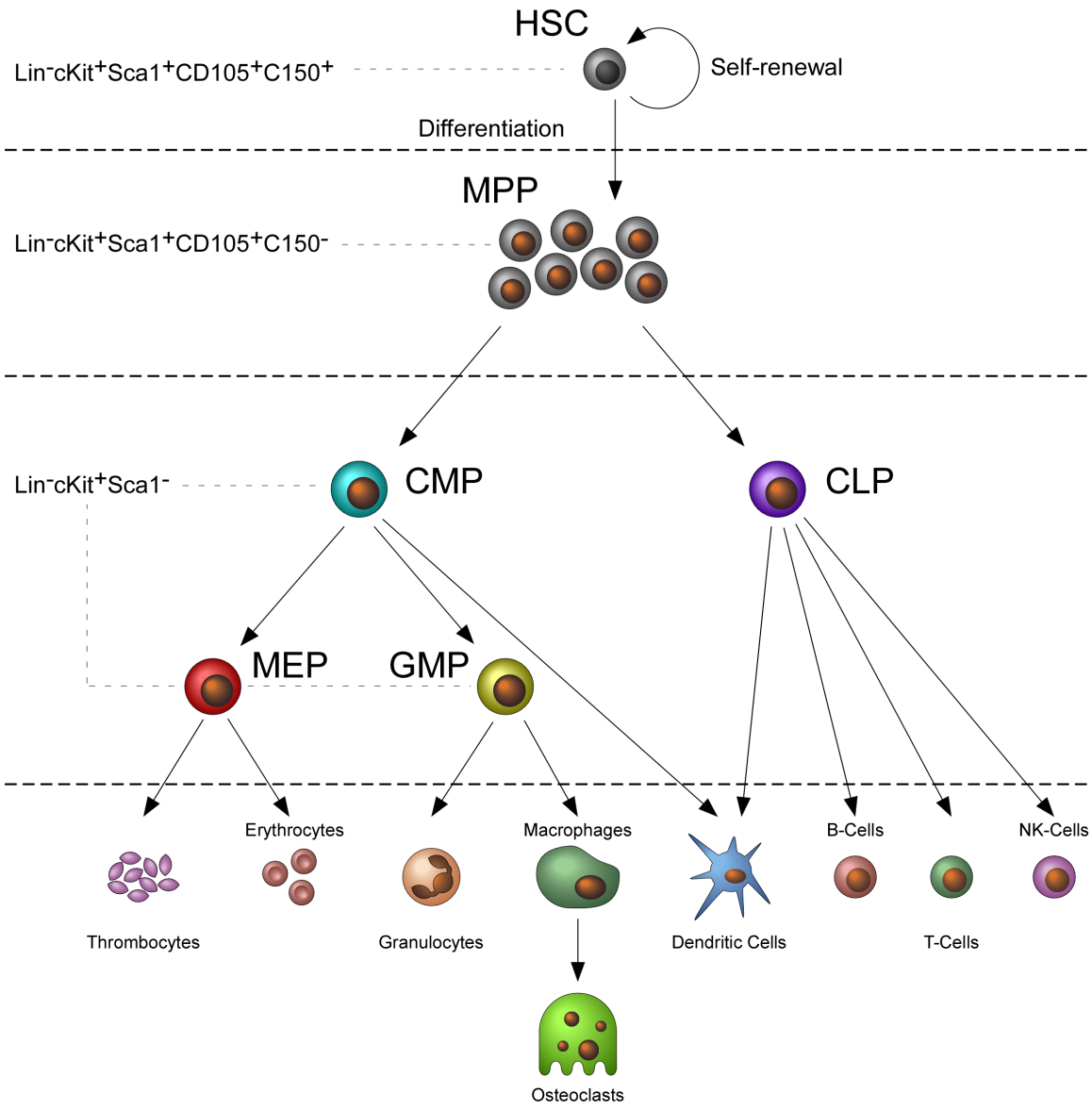


Figure 3.7 Flowchart of Hematopoietic Development

Hematopoietic stem cells (HSC) are the self-renewing cell type capable of repopulating the entire hematopoietic lineage. Flow chart shows hematopoietic development to various myeloid and lymphoid lineages. Dashed lines link cell types with the surface markers used to identify them by flow cytometry.

Previous studies had shown that early infection of the bone marrow correlates with leukemogenicity of MLV [43]. Furthermore, cells of the myeloid lineage make up the majority of cell types infected during early stage Mo-MuLV infection after i.p. inoculation [148]. Specifically osteoclasts and osteoclast precursors are amongst the first cells infected and these cell types are derived from the myeloid lineage [152]. Since myeloid differentiation is perturbed in the ZASC1$^{-/-}$ mice we speculate that this may account for the delay in Mo-MuLV replication that is seen in the bone marrow of these animals, i.e. fewer of these cells may be available to support the initial rounds of virus replication. According to that model, virus replication might “catch up” in ZASC1-deficient animals as the virus is amplified in other bone marrow cell types, perhaps explaining why there was no difference in virus titers seen at later time points and the fact that there was no obvious impact of ZASC1 deletion on MLV disease pathogenesis. Alternatively, there may be a block to efficient virus gene expression in ZASC1$^{-/-}$ bone marrow cells, limiting the initial virus spread through the tissue. However, such an effect would have to be “context-dependent” because the ZASC1-deficiency was associated with low virus titers in multiple bone marrow cell types, including purified T- and B-cells, and there was no obvious defect in T- or B-cell infections in either the thymus or spleen. Further studies are required to determine how ZASC1 regulates myeloid cell differentiation and how this factor impacts early Mo-MuLV replication *in vivo*.

A.

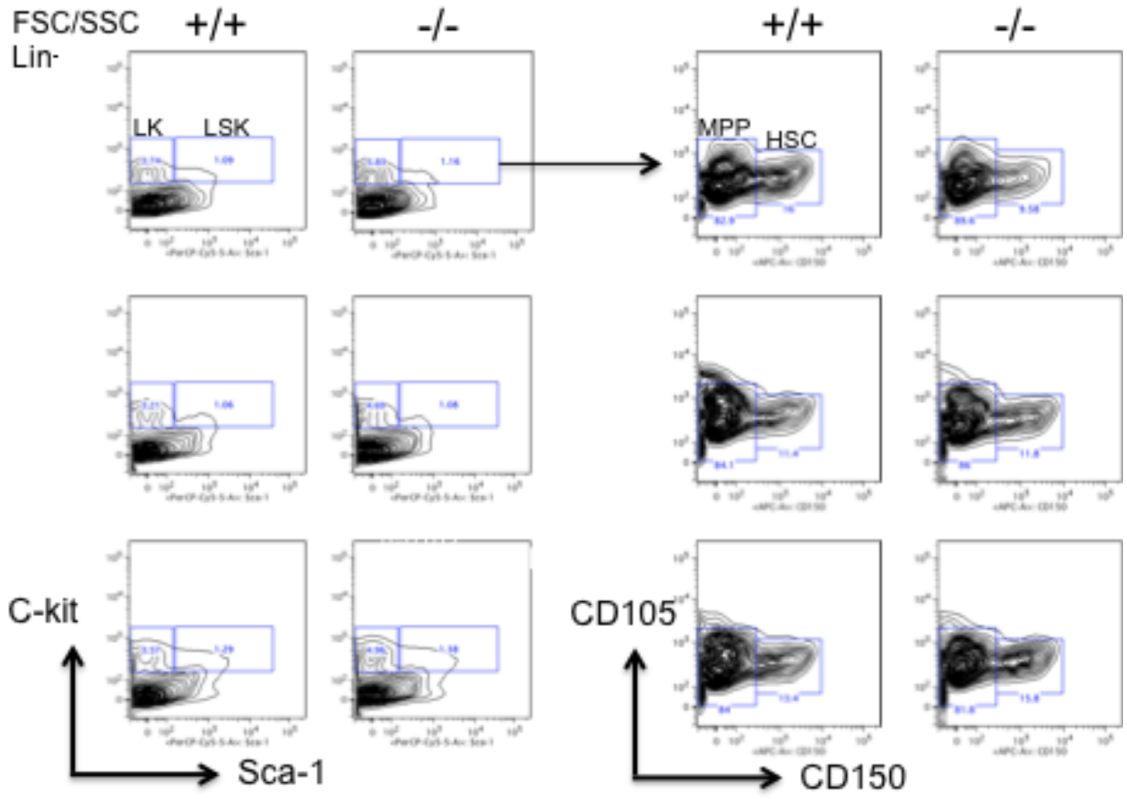


Figure 3.8 ZASC1$^{-/-}$ mice have increased LK cell compartments in the bone marrow.

A.) FACS plots gating hematopoietic progenitor populations in the bone marrow. Antibodies used: α-c-kit, α-sca-1, α-CD105, α-CD150, lineage markers (α-CD11b, α-Gr-1, α-Ter119, α-CD3, α-Ly6C, α-CD19, α-CD11c). LK (Lin$^-$ Sca-1$^-$ C-kit$^{+)}$, LSK (Lin$^-$ Sca-1$^+$ C-kit$^{+)}$, MPP (multipotent progenitor) and HSC (hematopoietic stem cell) populations within the bone marrow were gated and labeled as in the top left box.

B.

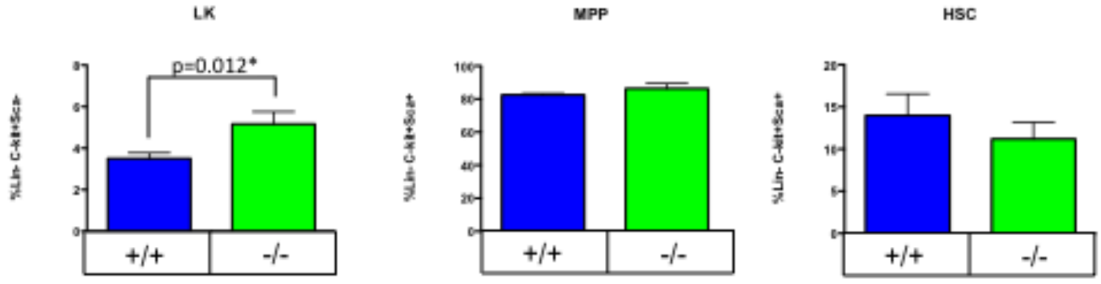
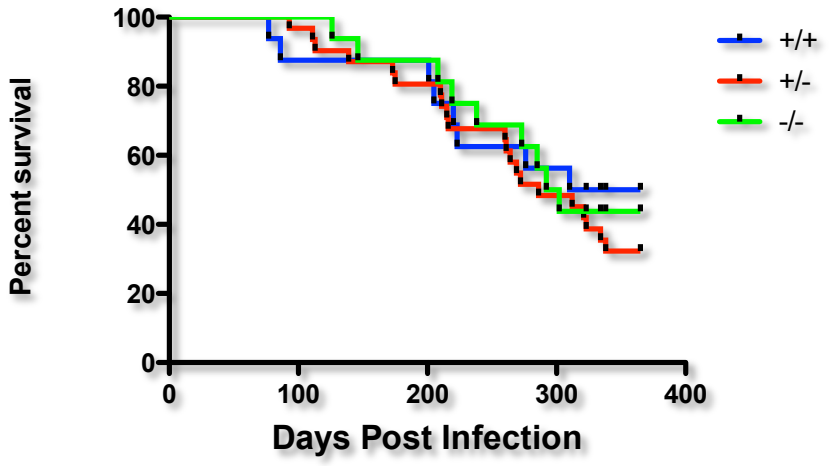


Figure 3.8 (continued)

B.) Graphical representation of the cell populations gated in each in plot in part A. A. N=3 mice per group. Shown are mean average values of cell percentages. Error bars indicate SD. Two-tailed T-test was used to measure statistically significant differences between ZASC1$^{+/+}$ and ZASC1$^{-/-}$ populations. Significant p-value is indicated.

A.



B.

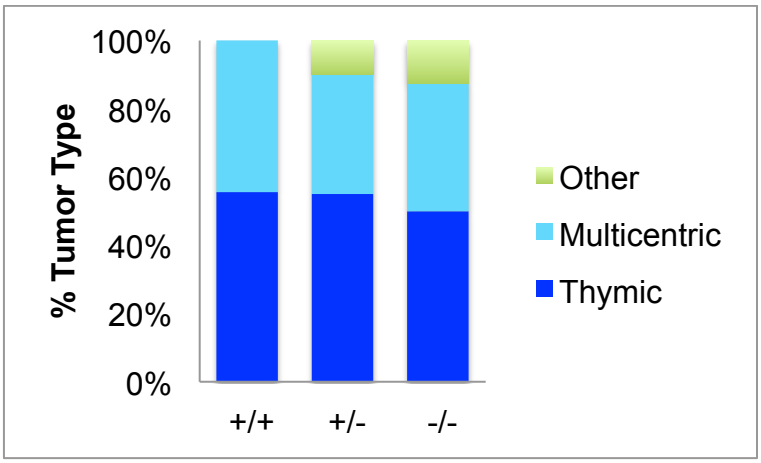


Figure 3.9 No defect in Mo-MuLV pathogenesis in ZASC1-/- mice.

A.) Kaplan-Meier plot comparing survival of MMLV infected animals of ZASC1$^{+/+}$ (Blue), ZASC1$^{+/-}$ (Red) and ZASC1$^{-/-}$ (Green) genotypes. Litters of mice from ZASC1+/- breeding pairs were infected p2-p3 and were allowed to develop disease. Animals that had not exhibited overt signs of disease were terminated 1 year post infection. ZASC1+/+ (N=15), ZASC1 +/- (N=31), ZASC1 -/-(N=15). B.) Stacked graph of tumor types by ZASC1 genotype.

Given the importance of ZASC1 for Mo-MuLV replication in established cell lines [45] it was surprising to us that deletion of this gene did not have more of a major impact on virus replication *in vivo*. This result could mean that ZASC1 is not important for MLV replication outside of the early events in the bone marrow. Alternatively, it could mean that there is another cellular factor that could substitute for ZASC1 in most other tissues. Although no other transcription factors are known to interact with the characterized ZASC1 binding site, it is possible that another transcription factor, perhaps another zinc finger protein shares this binding specificity.

Consistent with a possible redundant factor we found that WT Mo-MuLV replication was equivalent in WT and ZASC1-deficient mouse embryonic fibroblasts (MEFs). However, Mo-MuLV containing a mutation in ZBS1, the most important ZASC1 binding site in the viral genome [45], exhibited significantly less infection of ZASC1$^{-/-}$ MEFs, as compared with Wt virus (Fig. 3.10). That latter result indicates that ZBS1 is still important for virus replication in ZASC1-deficient cells, consistent with the notion that a redundant factor might interact with that site to modulate Mo-MuLV gene expression. Future studies will address the identification of a putative redundant factor to determine if another transcription factor can act at ZBS1.

The requirement for ZASC1 in myeloid cell differentiation is of interest. This factor may be required for differentiation downstream of the CMP into one or more derivative lineages. However, it is also possible that ZASC1 affects progenitor cell division or skews the myeloid lineage cell fate choice. Additional

studies will define the precise defect in bone marrow differentiation in the absence of ZASC1.

3.5 Methods

Generation of the ZASC1 conditional knockout (CKO) targeting construct

The Bacterial Artificial Chromosome (BAC) clone RP23-386L20 (BACP PAC Resources Online) was modified using homologous recombineering technology as described by Chan et al [153]. Briefly, targeting cassettes containing loxP sites were introduced in the intron between *ZASC1* exons 3 and 4 and downstream of exon 7 and the mRNA polyadenylation signal to facilitate deletion of the entire coding region of *ZASC1*. All plasmids and cassettes were used as described in Chan et al with the exception of modifications made to the Neo-Cassette to add in a splice acceptor-GFP-polyA sequence in an effort to facilitate GFP expression from the ZASC1 promoter following Cre-recombinase mediated deletion of ZASC1. BAC containing DH10B E. coli cells were transformed with pSIM18 [153] and selected on LB-Hygro plates (75µg/ml). A positive clone was prepared for insertion of the Bsd-cassette by recombineering. The Bsd-cassette was PCR-amplified using primers Bsd-F and Bsd-R shown in Table 3.2 and the product was electroporated into BAC containing DH10B E.coli cells. The transformants were selected on LB-Hygro-Bsd plates (75µg/ml each) and one colony was chosen after confirmation of Bsd-cassette integration.

Table 3.2 Primers used to design the ZASC1 Targeting Vector

Primer Name	Primer Sequence
Bsd-F (ss20f)	CTCTTTATGTAGACCAGGCTAGCCTTGAACTCTG CCTGTCTCAGCCTCTGCGACTCACTATAGGGCGAATTGGG
Bsd-R (ss20r)	GACCCTATCTCATAAGACAAAGCAAACACACACCTTTTAGTC CCAGAACCGCTATGACCATGATTACGCCAAGC
GFP-F(ss23f)	TAGGCGCGCCGCCACCATGGTGAGCAAGGGCGAG
GFP-R (ss23r)	GCTTAATTAATTACTTGTACAGCTCGTCCATGCC
SA-GFP-PA-F (ss24f)	CAGGATCCATCTGTAGGGCGCAGTAGTCCAG
SA-GFP-PA-R (ss24r)	TAGCGGCCGCGGATCGAGCCCCAGCTGGTTC
Neo-F (ss25f)	CCTTGTATGAAGTAGGCTATAGGAGTAGGGTACGTCTGTGT GGTGGGGGAGGAATTCTGTAAAACGACGGCCAGTG
Neo-R (ss25r)	GCCATCAGGCAAACACCTCACATAAAATAAATTATGAAAAA CTACTCCACCTCGAAATTAACCCTCACTAAAGGG
Retrieval-F (ss33f)	GTGAGATTCAGTTTGGTGACCTCATAATCTGTAGGCATTC TCCTCTTAGACATGGAGTTAGTGAGGGTTAATTATCGAC
Retrieval-R (ss33r)	GCCAAGGTTGAATTACCCAGGTTTGGACTACAGCCTAAA GCTCTGTCACACCACTGTAATACGACTCACTATAGGGAGG

The modified Neo-cassette was PCR amplified using primers Neo-F and Neo-R (Table 3.2) and electroporated into BAC-Bsd containing DH10B E. coli cells. The transformants were selected on LB-Hygro-Kan plates (75µg/ml hygromycin, 50µg/ml kanamycin). Next, the retrieval vector pL611 [153] was PCR-amplified using the Retrieval-F and Retrieval-R primers (Table 3.2) and the PCR-product was electroporated into BAC-Bsd-Neo containing DH10B E. coli. The transformants were selected on LB-Amp (50µg/ml) plates and colonies were checked for plasmid recovery of the targeted ZASC1 locus including a 6kb upstream homology arm and a 2kb downstream homology arm. After retrieval, the Bsd-cassette in pL611-Bsd-ZASC1-Neo-GFP was replaced with a lacZ reporter cassette (isolated from pL613 [153]) by digesting both vectors with I-Sce1 and I-Ceu homing endonucleases and the lacZ fragment was ligated into the retrieved plasmid. The integrity of the ZASC1 targeting vector was then confirmed by PCR amplification and DNA sequencing of cassette-junctions, as well as by restriction enzyme digestion.

Modification of Neo-cassette: GFP was amplified by PCR from pLEGFP using primers GFP-F and GFP-R shown in Table 3.2. The PCR product was then ligated into the Rosa26-SA-polyA plasmid (gift from Vale Lab, Salk Institute) after double digest with Asc1 and Pac1 restriction enzymes. The resulting Rosa26-SA-GFP-polyA DNA fragment was amplified using primers SA-GFP-PA-F and SA-GFP-PA-R (Table 3.2). This cassette along with the Neo-cassette from Chan et al. were digested with BamH1 and Not1 restriction enzymes in order to ligate the

SA-GFP-PA fragment into the Neo-cassette downstream of the second FRT recombination site.

Homologous recombination of the ZASC1 CKO targeting vector in ES cells

The ZASC1 CKO targeting vector was linearized with Xma1, purified, and resuspended in ddH2O at a final concentration of 660ng/ul. $2x10^7$ male 129/J5 ES cells were electroporated (BioRad GenePulser) with 40ug Xma1 linearized ZASC1 CKO targeting vector , plated with irradiated MEFs, then 24 hours later selected with medium containing 200ug/ml G418. G418 positive ES cell clones were picked for further analysis after 8 days of culture.

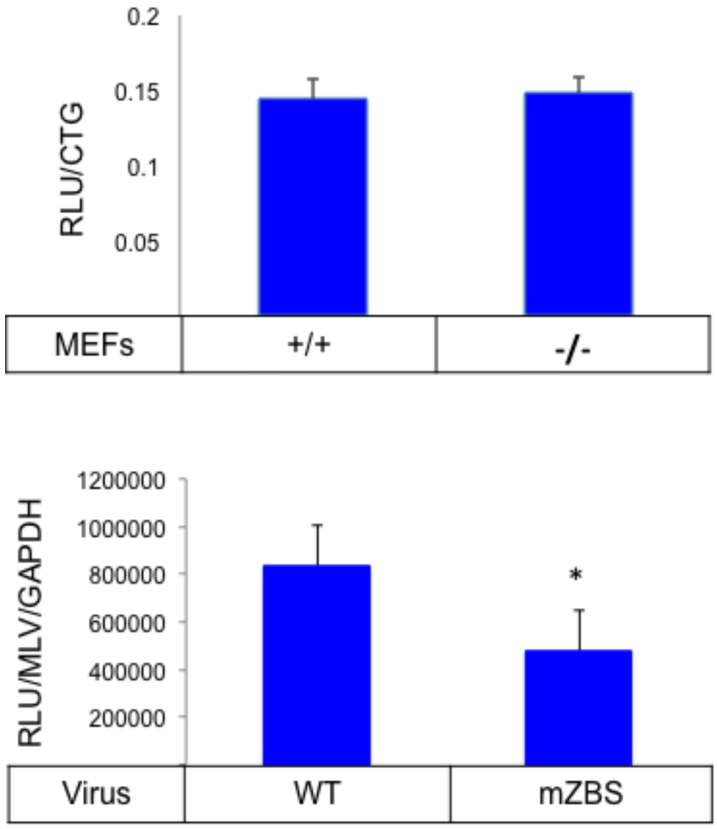


Figure 3.10 MLV Infections in Primary MEFs

MEF cells were generated from E13.5 embryos by making a single cell suspension from fibroblast tissues. Cells were then resuspended in media and pipetted up and down multiple times to create a single cell suspension. MEFs were then cultured in DMEM containing 10% FBS and pen/strep. A. An MMLV luciferase reporter virus was used to infect primary MEFs of the +/+ or -/- genotype. 30 hours post infection luciferase assay was used to measure virus infection and CTG was used to normalize for cell number. B. ZASC1-/- MEF cell line was challenged with either a WT MMLV reporter virus construct encoding luciferase or a mZBS version. Differences in virus titer were normalized by qPCR for viral DNA copies. Error bars are SD. Statistically significant differences were measured by two-tailed T-test. (*) $p<0.05$.

Southern blot analysis

15ug of genomic DNA derived from G418-resistant ES cells was digested using the Sph1 restriction enzyme (downstream arm analysis) or BamH1 restriction enzyme (upstream arm analysis) overnight at 37°C. The digested DNA samples were subjected to 0.8% agarose gel electrophoresis and then transferred to a charged nylon membrane (Millipore) overnight using a standard Southern blot transfer procedure . To monitor homologous recombination of the downstream and upstream arms, 32P-labeled DNA probes were prepared using the random hexamer kit, RediPrime II (GE Healthcare) following PCR amplification using primer set Short-F and Short-R, and primer set Long-F and Long-R (Table 3.3), respectively. The probes were incubated with the nylon membranes in Church buffer at 65°C overnight. The membrane was then washed once with low stringency buffer (2%SSC, 0.1%SDS) for 10 minutes at 60°C, then with high stringency buffer (0.2%SCC, 0.1%SDS) for 10 minutes, first at 60°C and then a second time at 42°C. After final wash, the membrane was exposed to BioMax Film (Kodak) with an enhancer screen at -80°C? for 1-7 days. Positive ES cell clones were recovered from frozen stocks and reconfirmed by southern blot analysis before they were injected into C57/Bl6 blastocysts. Microinjections were performed in the Salk Institute Transgenic Mouse Core Facility. Chimeric males (agouti) were bred with C57/Bl6 females and pups were screened by coat color. Agouti pups were screened by PCR genotyping to identify progeny with the targeted ZASC1 genomic locus using primers in GFP transgene. Progeny positive for the targeted ZASC1 genomic locus were bred

with aC57/Bl6 transgenic mouse expressing FLP recombinase from the human Beta-actin promoter (Jackson Laboratory Stock Number 005703). Offspring that show FLP recombination by genotype analysis (Transnetyx, Inc.) were then bred with one of three C57/Bl6 Cre Transgenics: Vav1-Cre (Jackson Laboratory Stock Number 008610), Esr1-Cre (Jackson Laboratory Stock Number 004682) or CMV-Cre (Jackson Laboratory Stock Number 006054).

Genotyping

Mouse tails were collected from 14 day old pups and genomic DNA was harvested by incubating over night in tail lysis buffer (100mM Tris pH8, 5mM EDTA, 2% SDS and 200mM NaCl +Proteinase K) at 55°C. DNA was purified by ethanol precipitation and dried pellets were resuspended in 500µL TE. 1ul diluted tail DNA samples were added to a 10ul qPCR reaction mixed that also included genotyping primers and FAST SYBR Green Master Mix (Applied Biosystems). Germline transmission of the ZASC1 knockout locus was detected with primer pair: 5'- TCAGTGCATCAGCATACTCAAG-3' and 5-CTGGTGGTTCTTCATCGGC-3' and then normalized to GAPDH using primer pair: '5- ACCCAGAAGACTGTGGATGG-3' and 5'-CACATTGGGGGTAGGAACAC-3'. qPCR analysis was performed using the Applied Biosystems ViiA 7 instrument. Genotyping for cre-recombinase, flp-recombinase and floxed ZASC 1 alleles was done through a professional genotyping service (Transnetyx).

Mo-MuLV production and titer determination

Virus was generated in extracellular supernatants from NIH 3T3-43D cells (Gift from Fan Lab, UC Irvine), filtered through a 0.45 micron filter, aliquoted and stored at -80°C. Viral titers were determined by Focal Immunofluorescence Assay (FIA) in the presence of polybrene as described previously [61]. Virus-infected cells were identified by staining with 538antibody supernatant (Fan lab, UC Irvine) [154], and an Alexa Fluor 488 Goat-anti-Mouse secondary IgG (Life Technologies). Plates were scanned and colonies counted using the Fuji FLA-5100 scanner.

The viral genome plasmid pCMMP-luciferase was used to generate the Mo-MuLV reporter viruses and mZBS-Mo-MuLV reporter virus, as described previously [45]. Virus-containing supernatants were was harvested 48hr post transfection and treated with DNase (40U/ml) before aliquoting and freezing at -80°C.

Table 3.3 Primers used to generate southern blot probes

Primer Name	Primer Sequence
Short-F	ATCAGGTGCATCCTCAGTCAC
Short-R	CAAGAATAGAAACCAGGTGTGC
Long-F	ATCTTGGCCATGGTAAGCTG
Long-R	GCGACTCACTATAGGGCGAAT

Flow cytometric analysis

Spleen and thymus samples were harvested from 4-week old, euthanized animals and ground into a single cell suspension through a 70uM cell strainer and incubated in ACK buffer to lyse red blood cells. Bone marrow samples were harvested from 10-day old, euthanized animals by crushing bones and treated as above with ACK buffer. After washing with medium, cells were resuspended in staining buffer (3% BSA in PBS) on ice and incubated with antibodies for 30 minutes before washing away unbound antibodies. For spleen analysis, α-CD3, α-B220, α-CD11b, α-CD4, and α-CD8 antibodies conjugated to fluorophores were used (eBiosciences). For thymus analysis, fluorescently labeled α-CD4 and α-CD8 antibodies were used (eBiosciences). For bone marrow, α-c-kit, α-sca-1, α-CD105, α-CD150, and lineage markers (α-CD11b, α-Gr-1, α-Ter119, α-CD3, α-Ly6C, α-CD19, α-CD11c) were used (eBiosciences and gift from Steve Hedrick Lab at UCSD). Samples were analyzed using the FACSAria III (BD) instrument and FlowJo software (Tree Star, Inc.).

In vivo infections

p2-p3 pups were infected by intraperitoneal-injection with 25μl Mo-MuLV ($8x10^5$ IU/ml). For early time points, animals were sacrificed and spleen, thymus, and bone marrow were collected in order to quantify virus infection by FIA, as described above. Briefly, single cell suspensions of each tissue were generated as described before and 10-fold serial dilutions of these cells were co-cultured with NIH3T3 cells until these cells reached confluency. Similar experiments were

conducted with T-cells, B-cells, and T/B-depleted samples that were purified from pooled 10-day mouse bone marrow samples using the Dynabeads FlowComp Pan T-, or Pan-B cell isolation kits (Life Technologies). The cells were then stained using antibodies described in [154].

Luciferase Assay

Chemiluminescent readout of infection was performed as described previously [137, 155]. 96-well plates were seeded with cells and incubated with virus. Spinnoculation (1hr at 1200xg) was used to help increase MOI. Then plates were incubated at 37°C for 24 or 48 hours before media was aspirated from adherent cells and Bright-Glo or CellTiter-Glo were added per manufacturers conditions. Cell lysates with luciferase reagents were then transferred to black 96-well plates and read using a Topcount NXT HTS (Perkin Elmer).

Tumor diagnosis and histopathological analysis

For tumor studies, 2-3 day old pups were infected as described above and were monitored until they developed disease. Tumors were diagnosed as described previously [61]. Mice were observed daily for signs of lethargy, hunched posture, and scruffy fur associated with Mo-MuLV-induced tumors. Moribund animals were sacrificed and necropsy was performed. Spleen, thymus, liver, and kidney tissues were weighed and pieces of these organs along

with lymph nodes were snap frozen and fixed in either 4%PFA in PBS (UBC) or 10% formaldehyde.

Fixed tissues were processed by Pacific Pathology, Inc. Fixed tissue samples were embedded in paraffin and 5um sections were cut and studied under light microscopy after hematoxilin-eosin (H&E) staining. Additional slides were prepared for immunofluorescence analysis by rehydrating slides, performing antigen retrieval using a sodium citrate buffer (composition) and staining with primary rabbit antibodies to CD3 (abcam) or CD79B (Santa Cruz Biotechnology) and with a secondary anti-rabbit antibody (Life Technologies).

Chapter 4. Conclusions and Future Directions

The data presented in this thesis demonstrate my contributions to our growing understanding of the role host cell factors play in retrovirus replication. Genome wide screens such as the one presented in Chapter 2 are a powerful way to look at virus replication on a global scale. Factors identified from this screen and others like it have now become the focus of individual research projects aimed at teasing apart the detailed mechanisms of virus-host interactions. The next crucial steps involve sorting out which factors are acting directly and which play indirect roles and many labs are actively pursuing this end. The Young lab is currently involved with this work as part of a multi-institutional group of investigators in the HIV Immune Networks Team (HINT).

ZASC1 was originally identified in a forward genetic screen and my work has demonstrated a key step in understanding how this transcription factor may be acting in an *in vivo* system. Previous studies from the Young lab and the Ahlquist lab had identified ZASC1 as a positive acting transcription factor that binds to and enhances HIV and MLV transcription [45]. This work was initiated in CHO-K1 cell lines and then characterized further in established human and mouse cell lines. I was interested in determining the role ZASC1 plays in disease pathogenesis in an animal model to determine the physiological relevance of this protein. In order to address this question, I designed and generated a ZASC1 knockout mouse model. Through this work I showed that although ZASC1 is

highly conserved and displays nearly ubiquitous expression patterns, the ZASC1 knockout displays no obvious phenotypic abnormalities. These mice are indistinguishable from littermate controls in regards to breeding and aging, however, an important future step in understanding the role of ZASC1 in normal cellular function includes a comprehensive characterization of this knockout mouse model.

Early infection studies using these ZASC1-deficient mice showed a very interesting defect in MLV replication specifically in the bone marrow compartment at an early time point post infection. No defects were identified in the spleen or thymus and by about two weeks, virus levels in ZASC1$^{-/-}$ mice rebounded. This replication defect could not be duplicated in cultured cells from the bone marrow indicating a niche specific effect. There are several possible models discussed in Chapter 3, however, I will try to elaborate on these models and suggest possible experiments to further our understanding of this observation.

First, there are two possible categories of effect that we must differentiate between. ZASC1 may act either directly at the virus promoter or indirectly through another mechanism. It is also possible that multiple mechanisms are at play. Studies in Chapter 3 demonstrated that although ZASC1 may be required for retrovirus replication in immortalized cell lines, its role in primary cells is more difficult to establish. It may be that ZASC1 is not required in primary murine cells, however, it is also possible that its role is more difficult to determine because a redundant factor acts at the ZBS, masking the knockout phenotype. Determining if the ZBS is occupied in the presence and absence of ZASC1 in primary cells

would be the first step to addressing the possibility of a redundant factor. This could be accomplished using a DNase footprinting assay. If the virus promoter is bound in the absence of ZASC1, it would be useful to identify the responsible factor and try knocking it down in primary ZASC1$^{-/-}$ cells to see if this removes the redundancy.

Alternatively, ZASC1 might act directly at the promoter in an activation dependent manner. Very little is known about ZASC1, in particular how this protein is regulated. The N-terminus has a known NLS and no other characterized protein domains, however, several phosphorylation sites have been identified and ZASC1 was shown to be phosphorylated after T-cell activation [156]. If this step is required for ZASC1 to effect retrovirus transcription, perhaps this process is highly regulated, silencing ZASC1 function in many organs. It would be interesting to compare the ZASC1 phosphorylation state in immortalized cell lines and primary cells of different mouse organs to determine if any pattern emerges. Collaborators in the Ahlquist lab are currently doing ChIP-seq studies with ZASC1 to identify additional cellular promoters that can bind ZASC1. This work will be valuable in understanding where and under what conditions ZASC1 regulates gene expression.

The hypothesis that ZASC1 may effect retrovirus replication through an indirect mechanism originated with the observation that young mice have an increase in the heterogenous compartment within the bone marrow contain the common myeloid progenitors. This is the only known defect in immune cell differentiation in the ZASC1 knockout mice and happens to be in a cell type that

is important for establishing early MLV infection after i.p. inoculation of virus [148]. In order to determine if we can attribute the infection phenotype to this differentiation defect, we must further characterize the LK compartment to see if there are specific downstream populations that are misrepresented in ZASC1-deficient mice. Eventually, if a population of cells is identified, it would be useful to infect mice, harvest the bone marrow at various times post infection from 1-10 days and purify the LK cell populations (or subsets) to quantify infection using $ZASC1^{+/+}$ and $ZASC1^{-/-}$ mice. At 10 days post infection, there seems to be a general defect but if this is a result of delayed early replication in LK compartment, it may be possible to identify an earlier more specific cell type defect. If an antibody can be used to identify infected cells by flow cytometry, it may be easier to determine if an indirect effect of ZASC1 on cell differentiation in the myeloid compartment causes the bone marrow specific virus replication defect.

Finally, although work in this thesis did not address it, future studies with the $ZASC1^{-/-}$ mice will be done to explore the possible role of ZASC1 in development and progression of squamous cell carcinoma. In collaboration with Reuben Shaw at the Salk Institute, I have generated a line of mice with $ZASC1^{fl/fl}$ $KRAS^{fl/+}$ and $LKB^{fl/fl}$ where upon inhalational challenge with Adeno-Cre virus ZASC1 and LKB are deleted and KRAS is activated leading to efficient induction of squamous cell and adenocarcinomas. If ZASC1 is a required oncogene for squamous cell carcinoma, we expect to see a shift in disease specificity of these mice, although, it is possible that this gene must be amplified and over-

expressed in order to see a phenotype. Additionally, our collaborators in Paul Ahlquist's lab are initiating studies of how ZASC1 may be involved in cervical carcinoma, a disease that impacts nearly half a million women annually.

Appendix 1. Generating the ZASC1$^{-/-}$ Mouse

This appendix contains additional assays done in the process of generating the ZASC1 conditional knockout mouse strains. The purpose of experiments and methodologies used will be briefly described with figures to follow.

In vitro cre assay

In order to confirm the recombination ability of the ZASC1 targeting vector before transforming mouse embryonic stem cells, Cre-induced recombination of the loxP sites was tested using an in vitro cre assay.

Purified Cre recombinase (NEB) was incubated with the ZASC1 targeting vector in an *in vitro* reaction to test the ability of Cre to catalyze the reaction excising the 6382bp band that includes the coding region of ZASC1. 1µl of the enzyme was added to a 20µl reaction mixture containing 1µg of the targeting vector (see manufacturers instructions). Both the standard circular plasmid and a linearized version were tested to ensure functional cre recombination. Additionally, control reactions without the enzyme were used as controls. Reactions were incubated over night at 37°C before being visualized on a 1% agarose gel stained with ethidium bromide. Figure A1.1 shows the presence of the expected 6382bp band that represents the excised loxP fragment in reactions containing Cre recombinase.

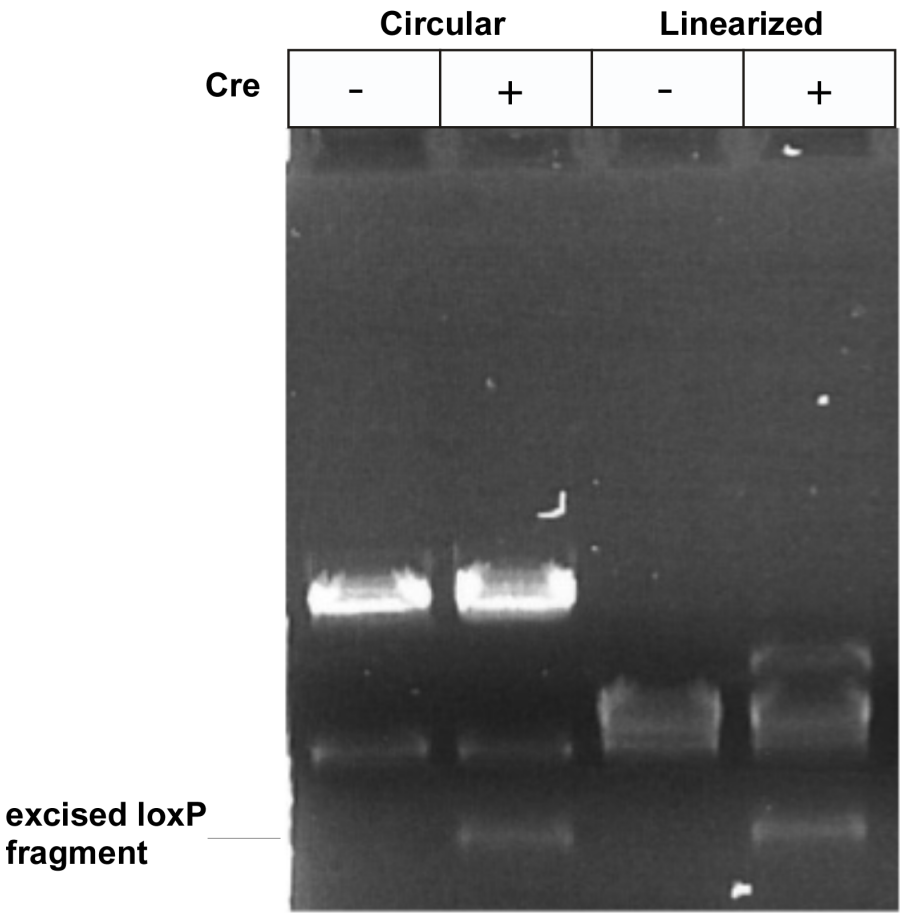


Figure A1.1 In vitro cre recombinase assay

The ZASC1 targeting vector was incubated with cre recombinase in order to test the effectiveness of the loxp sites for removal of the ZASC1 coding region. Both a linearized and a circularized version of the targeting vector were tested. Although efficiency of the in vitro cre reaction was <50% it clearly shows that in the presence of cre recombinase the expected 6382bp band is present.

Downstream Southern Blot

In order to identify embryonic stem cells with homologous recombination at the ZASC1 locus, a southern blot was done. To check homologous recombination at the downstream arm of the targeting vector, a probe within the genomic region downstream of the homologous arm was used. ES cell clones were subjected to G418 selection and resistant colonies were cultured for further analysis. The restriction enzyme Sph1 was used to digest genomic DNA before it was run on an agarose gel, transferred to a charged nylon membrane (Millipore) and blotted (refer to Chapter 3 for complete experimental details).

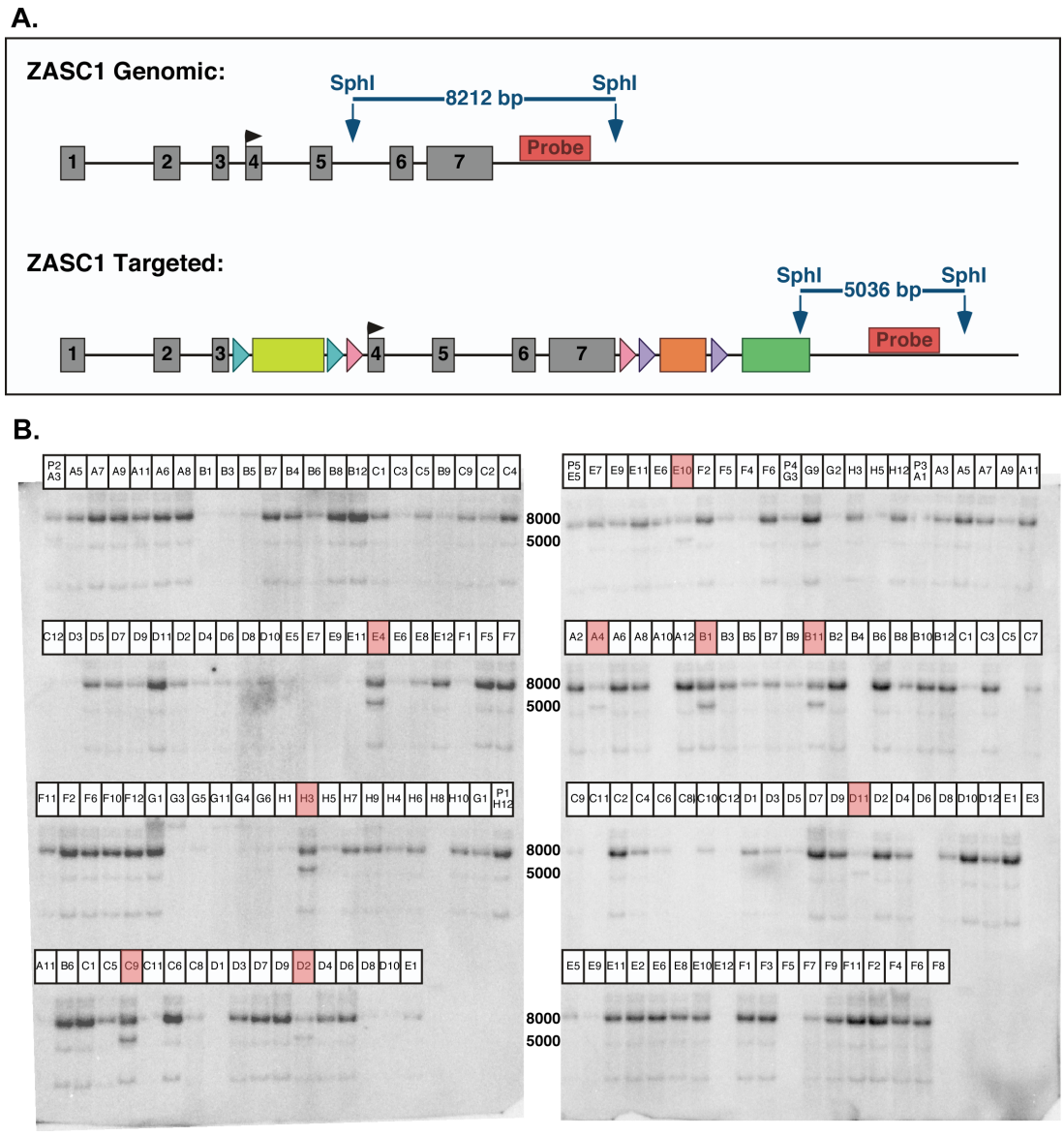


Figure A1.2 Southern Blot at the downstream arm of the ZASC1 Targeting Vector

A.) Schematic representation of the southern blot used to screen ES cell colonies for homologous recombination. ZASC1 exons (grey boxes), F3 recombination sites (blue triangles), puromycin cassette (yellow box), loxp recombination sites (pink triangles), FRT recombination sites (purple triangles), neomycin resistance cassette (orange box), and spice acceptor-eGFP (green box) are shown in the targeted locus. DNA was digested with Sph1 and probed as indicated above. B.) Two southern blots used to screen colonies. Red highlighted boxes show ES cell clones positive for homologous recombination. Positive colonies were indicated by equal intensity bands around 8kb (Genomic) and 5kb (Targeted).

Upstream Southern Blot

To ensure that both arms of the ZASC1 targeting vector were properly inserted into the genome via homologous recombination colonies that were positive in the downstream arm were then screened at the upstream arm. Genomic DNA was digested with the restriction enzyme BamH1 and then run on a 0.8% agarose gel. DNA was transferred to a charged nylon membrane (Millipore) and probed using an internal probe to the targeted vector that would generate a band at approximately 10,000bp. For complete methods, see experimental details in Chapter 3.

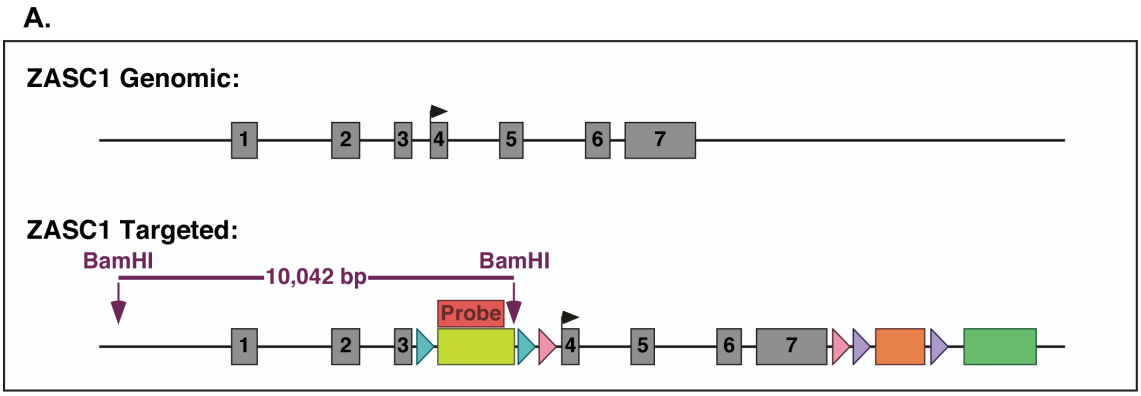


Figure A1.3 Southern Blot at the Upstream Arm of the ZASC1 Targeting Vector

A schematic of the southern blot designed to identify homologous recombination at the upstream arm of the ZASC1 targeting vector. The restriction enzyme BamH1 was used to digest genomic DNA and the radiolabled probe was internal to the targeting vector. ZASC1 exons (grey boxes), F3 recombination sites (blue triangles), puromycin cassette (yellow box), loxp recombination sites (pink triangles), FRT recombination sites (purple triangles), neomycin resistance cassette (orange box), and spice acceptor-eGFP (green box) are shown in the targeted locus.

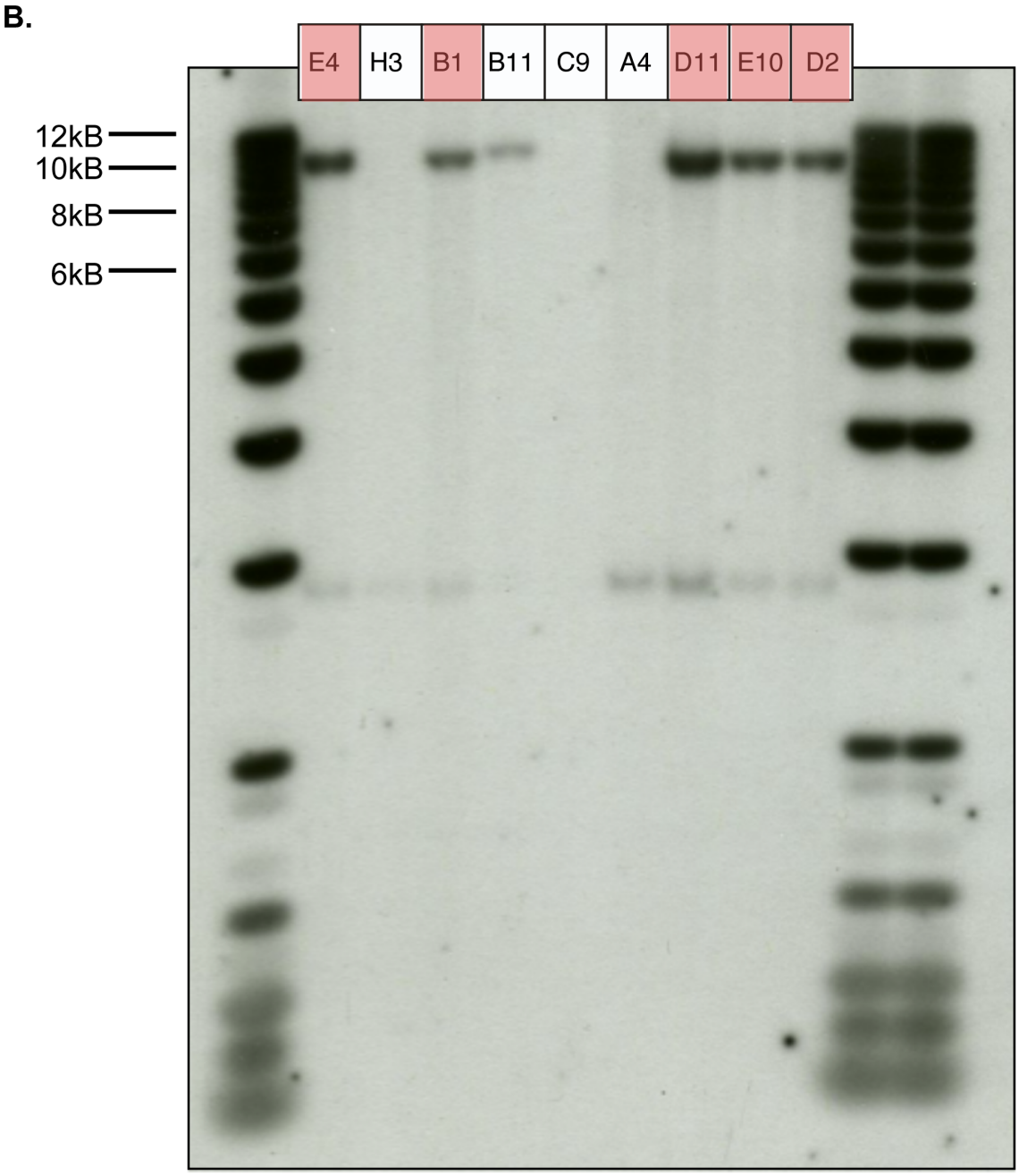


Figure A1.3 (continued)

9 colonies positive for homologous recombination at the downstream arm were tested for recombination at the upstream homology arm. 5 were positive and clones D2 and E10 were chosen for blastocyst injection. Red highlighted boxes show colonies positive for recombination at both homologous arms.

Breeding ZASC1^{fl} chimeras for germline recombination

In order to generate ZASC$^{-/-}$ mice, ES cells with homologous recombination at the ZASC1 locus were injected into the blastocysts of C57Bl/6 embryos to generate chimeras. Because the ES cell line used were male, males born with >20% chimerism by coat color were bred to WT C57Bl/6 females to screen for germline transmission of the ZASC1^{fl} allele. Table A1.1 shows breeding results from chimeric males.

Table A1.1 Results of ES Cell injection and Chimera Breeding

19 male chimeras were born between the two ES cell clones injected into mouse blastocysts. The %chimerism shows the approximated percentage of the coat of each animal that was agouti rather than black. Litters shows the number of females associated with males that gave birth to pups. Black and agouti show the total number of pups with each coat color born to a specific male. Germline transmission resulted from one male #3012 from the D2 ES cell clone.

Male	ES Clone	% Chimerism	Litters	Black	Agouti
3001	E10	20	Not Bred		
3002	E10	20	Not Bred		
3003	E10	0	Not Bred		
3004	E10	80	Sterile		
3005	E10	60	3	24	0
3006	E10	60	3	23	0
3007	E10	30	1	8	0
3008	E10	30	1	9	0
3009	E10	80	2	16	0
3010	E10	70	Sterile		
3011	E10	70	Sterile		
3012	D2	60	5	40	3
3013	D2	30	Sterile		
3014	D2	50	1	6	0
3015	D2	50	3	19	0

ZASC1 Conditional Knockout Strains

A conditional knockout targeting vector was generated because the role of ZASC1 in development was unknown. ZASC1 is widely expressed in many tissues and is highly conserved in mammals causing concern about the developmental role it might play. Our goal was to generate a model system in which to study retrovirus replication *in vivo* so to avoid possible complications in development, we generated 3 different strains of ZASC1 conditional knockout mice simultaneously.

The first was a hematopoietic specific ZASC1 knockout. This was generated by breeding the ZASC1$^{fl/+}$ 129/J and C57Bl/6 mixed background mice to a C57Bl/6 Vav1Cre strain (Jackson Labs Strain Number: 008610) While this strain generated an efficient ZASC1 deletion within the hematopoietic system (Fig. A1.5) it demonstrated leaky expression of Cre recombinase in germ line tissues generating complete germline knockouts. This strain was used for some preliminary studies of virus infection, however, it was discarded when it became clear that the ZASC1 germline knockout was viable and functional for all studies of virus infection.

A second group of ZASC1$^{fl/+}$ mice were bred into a tamoxifen-inducible Cre strain (Jackson Labs Strain Number: 004682) This strain was extensively tested both by in vivo tamoxifen injection and by culturing cells from the mouse with media containing tamoxifen. In no case was recombination observed (data not shown). Because of this the inducible cre strain was discarded.

Finally, a third set of ZASC1$^{fl/+}$ mice were bred to a CMV C57Bl/6 strain expressing CMV cre (Jackson Labs Strain Number: 003465). These mice eventually produced a germline ZASC1 knockout allele that was bred and used for the majority of experiments done in this thesis. These mice were also used to generate mouse embryonic fibroblast primary cell lines.

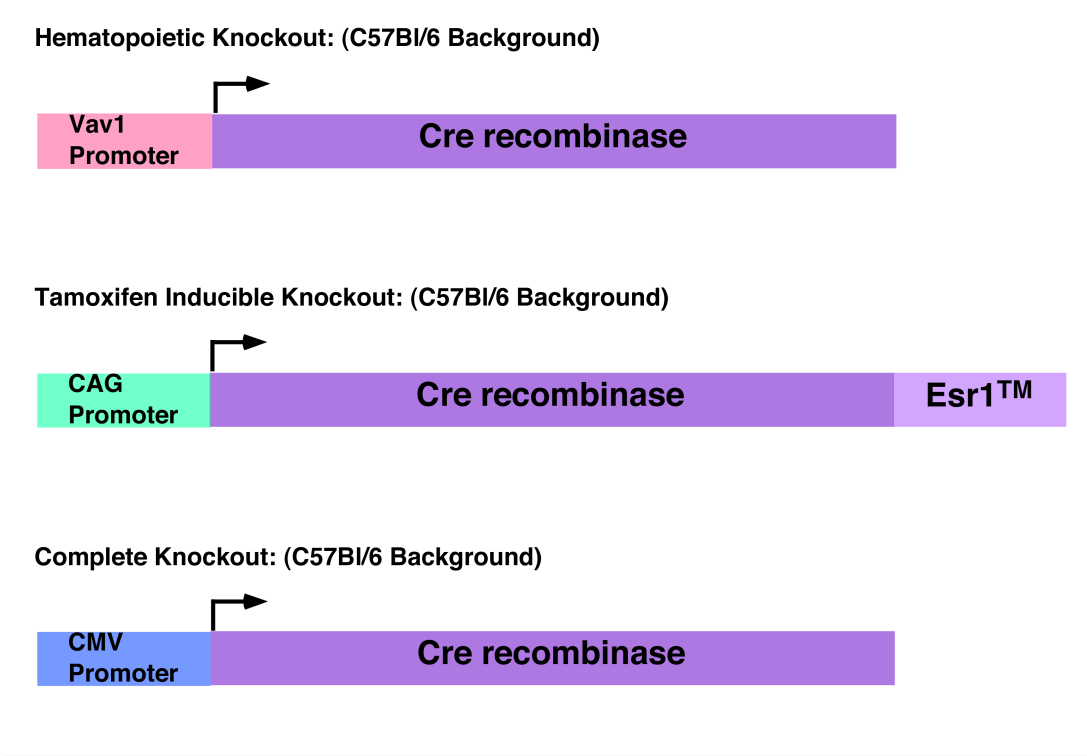


Figure A1.4 ZASC1 Conditional Knockout Mouse Strains Generated

Schematic representations of 3 cre recombinase genes bred into the ZASC1^{fl} 129J/C57Bl/6 mixed background. The hematopoietic specific cre was driven by the Vav1 promoter. The tamoxifen inducible cre was driven by a hybrid CAG promoter (CMV enhancer region combined with the chicken actin promoter). The cre recombinase gene is fused to a tamoxifen-inducible mutant of the estrogen receptor 1. The compete knockout was a cre recombinase gene driven by the ubiquitously expressed CMV promoter.

Appendix 2. Additional Characterization of ZASC1

This appendix contains experiments both in cultured cells and the ZASC1 mouse model aimed at elucidating the role of ZASC1 in retrovirus replication and normal cellular function.

A2.1 Background

Immune response to MLV typically includes both cell-based and humoral immune responses. Cytotoxic T-cells target infected cells and a neutralizing antibodies bind to the envelope glycoproteins on the surface of viral particles [157]. In the case of adult mice with fully developed immune systems, this response leads to efficient clearance of the virus. However, neonatal mice do not clear the virus when infected before the p3 stage. Neonates can therefore be used to study virus infection and disease pathogenesis *in vivo* because they are highly infectable and consistently develop disease. It is thought that virus is able to infect and spread to the thymus before the immune system is fully developed tolerizing the animal to infection. Because the immune system is so important for retrovirus infection and replication, we wanted to determine if ZASC1 played a role in any aspects of immune development or activation. Studies in this section along with data from Chapter 3 describe efforts to analyze immature and mature immune cells within the spleen, thymus and bone marrow. Additional experiments functionally test the immune system to determine if ZASC1 is required for a proper immune response.

The role of ZASC1 in normal cellular function is very poorly understood. Previous studies have identified ZASC1 as a transcription factor and implicated it in a variety of head and neck carcinomas, however, the virus promoter is the only established site of ZASC1 function. Because of the role ZASC1 plays in HIV-1 transcription, studies were initiated to identify cellular factors regulated by ZASC1. Although this work proved to be technically challenging, progress toward this end will be documented here.

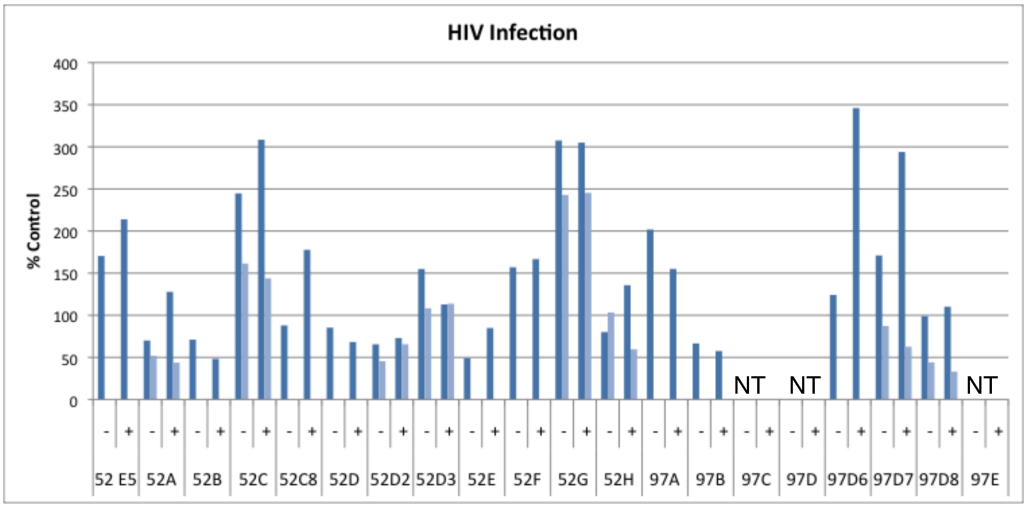
A2.2 Results

In order to address the mechanism of ZASC1 function at the viral promoter and in regulation of cellular factors, studies were established using cultured cell lines transduced with shRNA vectors to regulate ZASC1 expression. When these studies were initiated, the ZASC1 mouse model had not yet been generated so 3 constitutively-expressing shRNAs were independently cloned into a vector that also constitutively expressed GFP. These 3 shRNA constructs along with two dox-inducible shRNAs were transduced into Jurkat cells and shown to knockdown ZASC1 expression 50-90% (data not shown). Single cell clones were generated from pools of shRNA-transduced Jurkat cells using an antibiotic resistance marker within the constructs. Clones were grown in the presence of antibiotic selection and then screened for HIV infection and ZASC1 mRNA expression (Fig. A2.1 and data not shown). Dox-inducible shRNA clones were screened in the presence and absence of doxycycline.

Interestingly, there was a wide range of HIV infection in the clones, however, very few showed even a 50% reduction in HIV expression (Figure A2.1A). While many of the clones were dox-inducible to some degree, overall ZASC1 shRNA constructs had only a modest effect (Figure A2.1B), explaining the lack of phenotype with the virus. Jurkat clones were also generated using the 3 constitutive shRNAs (data not shown). shRNA56 exhibited the strongest effect on ZASC1 mRNA, however, none of the single cell clones grew in culture, indicating a potential growth defect in the absence of ZASC1.

In order to test if ZASC1 was required for cell viability, a growth curve after transfection with ZASC1 shRNAs or a control shRNA was completed. The results (Fig. A2.2A) show that A549 cells transduced with ZASC1 shRNA56 exhibit a strong growth defect as compared to control shRNA vectors or untransduced A549 cells. The other two shRNAs tested showed similar but less pronounced growth defects (data not shown). Figure A2.2B shows that ZASC1 mRNA knockdown correlates with the cell cycle defect. A549 cells were used for this analysis instead of Jurkat cells as a part of another project, beyond the scope of this thesis.

A.



B.

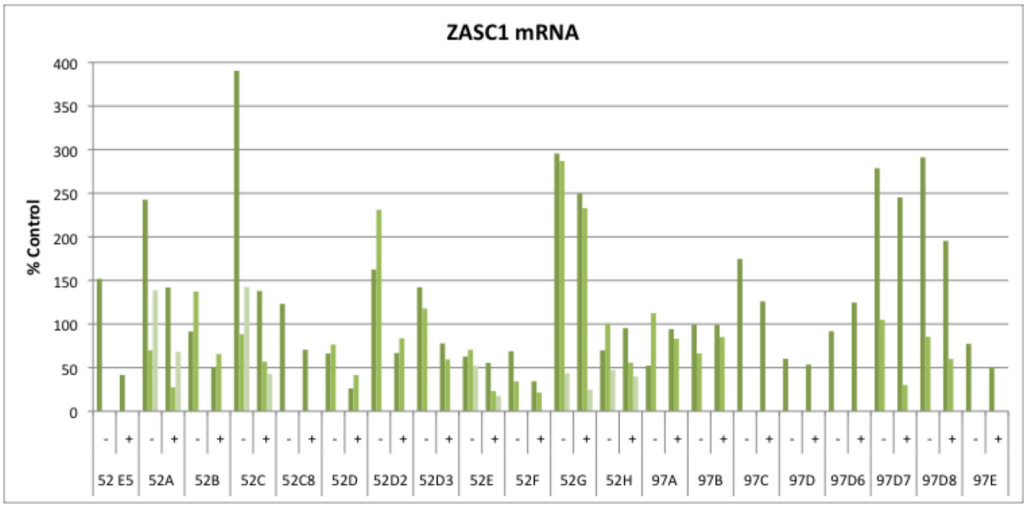
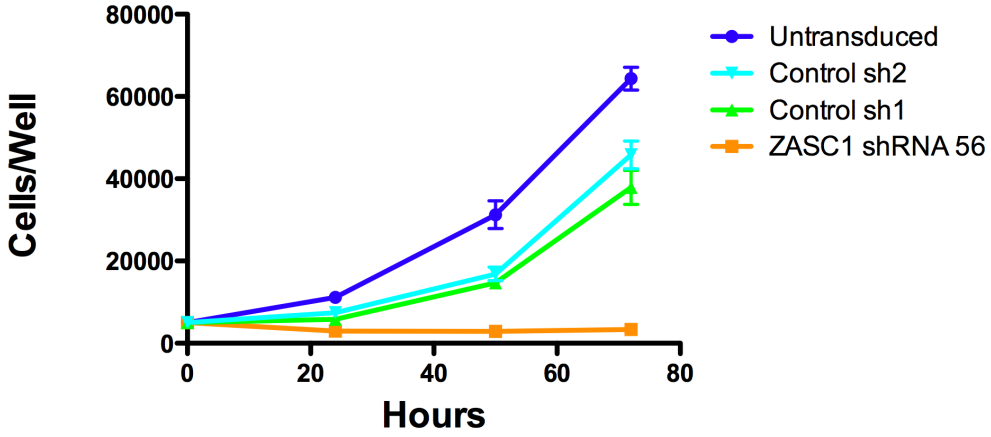


Figure A2.1 Screening Jurkat cell clones

Jurkats were transduced with ZASC1 shRNAs and then sorted and single cell cloned. Shown are dox-inducible clones. A.) Single cell clones were challenged with HIV-luc (VSV-G) and viral luciferase was measured and graphed as a percent control (untransduced). Dark and light blue bars represent independent experiments. Clones not tested for HIV infection are labeled NT. B.) ZASC1 mRNA was measured from single cell clones. Dark, medium and light green bars represent independent experiments. The presence of Dox is indicated by (-) and (+).

A



B

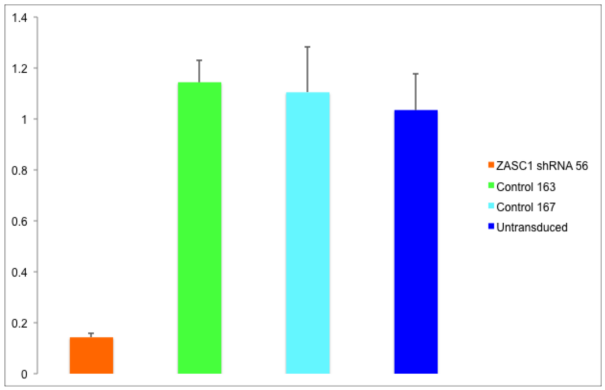


Figure A2.2 Growth Curve in A549 Cells with ZASC1 shRNA

A.) Growth curve in A549 cells after transduction with shRNA targeting ZASC1 or control scramble sequences. Cells were counted and plated 48 hours after transduction with shRNAs. At each time point, 3 independent wells were trypsinized and counted for each sample. Averages were plotted with error bars represented SD. B.) Relative mRNA levels were measured for each condition 96 hours post transduction. 3 independent samples were harvested, run in triplicate and normalized to GAPDH. Error bars represent SD.

Because the ZASC1 shRNAs exhibited such a strong growth defect we decided to focus on addressing the role of ZASC1 in normal cell function using the ZASC1 mouse model. In Chapter 3 we analyzed the bone marrow, spleen and thymus of ZASC1$^{-/-}$ mice to see if there were any defects in immune system development. We found a possible differentiation defect within the myeloid lineage in the bone marrow, however, no defects were found in spleen or thymus. In order to determine if ZASC1 effected functionality of immune cells, we first compared T-cell activation in ZASC1$^{+/+}$ and ZASC1$^{-/-}$ mice. T-cells were purified from the spleen and lymph nodes and labeled with CFSE, a cell-permeable dye that can be used to measure cell division. With each round of division, CFSE intensity decreases by 50% making it possible to follow proliferation through multiple rounds. T-cells were cultured in activating conditions for four days before being stained for cell surface receptors and analyzed by flow cytometry. Initially we used optimized activation conditions and found no defect in T-cell activation in ZASC1$^{-/-}$ mice (Chapter 3).

In order to observe potentially subtle defects in T-cell activation, purified T-cells labeled with CFSE were cultured with decreasing amounts of α-CD3 antibody creating suboptimal activation and proliferation conditions. The results in Figure A2.3 show that a subset of T-cells from the ZASC1$^{-/-}$ genotype have a decreased activation potential. We found that in younger mice (~3 weeks), T-cell activation and proliferation were similar for both genotypes but preliminary results from this assay suggest that older mice (~6 weeks) have some defect in T-cell

activation. This experiment was repeated once with similar results but requires additional testing to confirm.

In addition to monitoring T-cell activation and proliferation, we wanted to test whether activated ZASC1$^{-/-}$ T-cells were able to produce cytokines. Cytokine production is an important part of the adaptive immune response and provides a measure of T-cell functionality. Purified CD4+ and CD8+ T-cells from ZASC1$^{+/+}$ and ZASC1$^{-/-}$ mice were cultured in activating conditions for 4 days. Subsequently, cultures were stimulated with Ionamycin and treated with Golgi-Block to block cytokine export from the cell. IFNγ expression was observable in both CD4+ and CD8+ cultures for both mouse genotypes (Fig A2.4). TNF-α expression was also detected for CD8+ T-cells. Attempts to measure IL-4 production were unsuccessful in both ZASC1$^{+/+}$ and ZASC1$^{-/-}$ CD4+ T-cells due to technical difficulties.

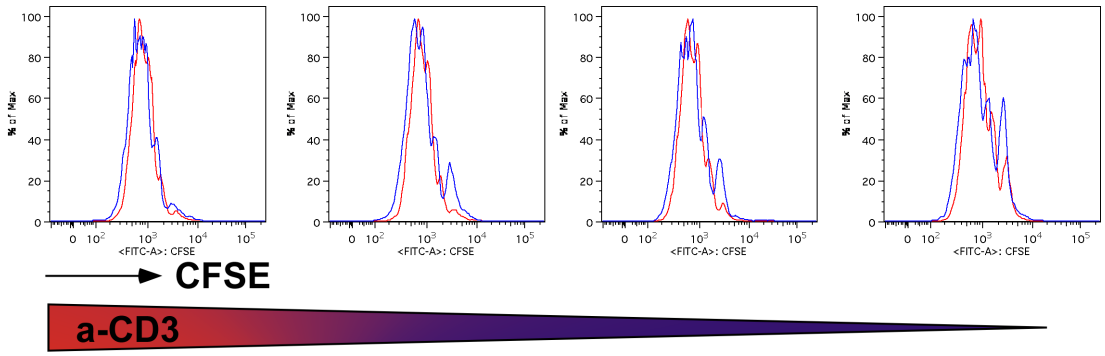


Figure A2.3 Activation of T-cells – Antibody Titration

CD4+ T-cells were harvested from mice, labeled with CFSE and activated in culture with α-CD28 and decreasing amounts of α-CD3. ZASC1$^{+/+}$ is shown in blue and ZASC1$^{-/-}$ is shown in red. N=2 independent experiment.

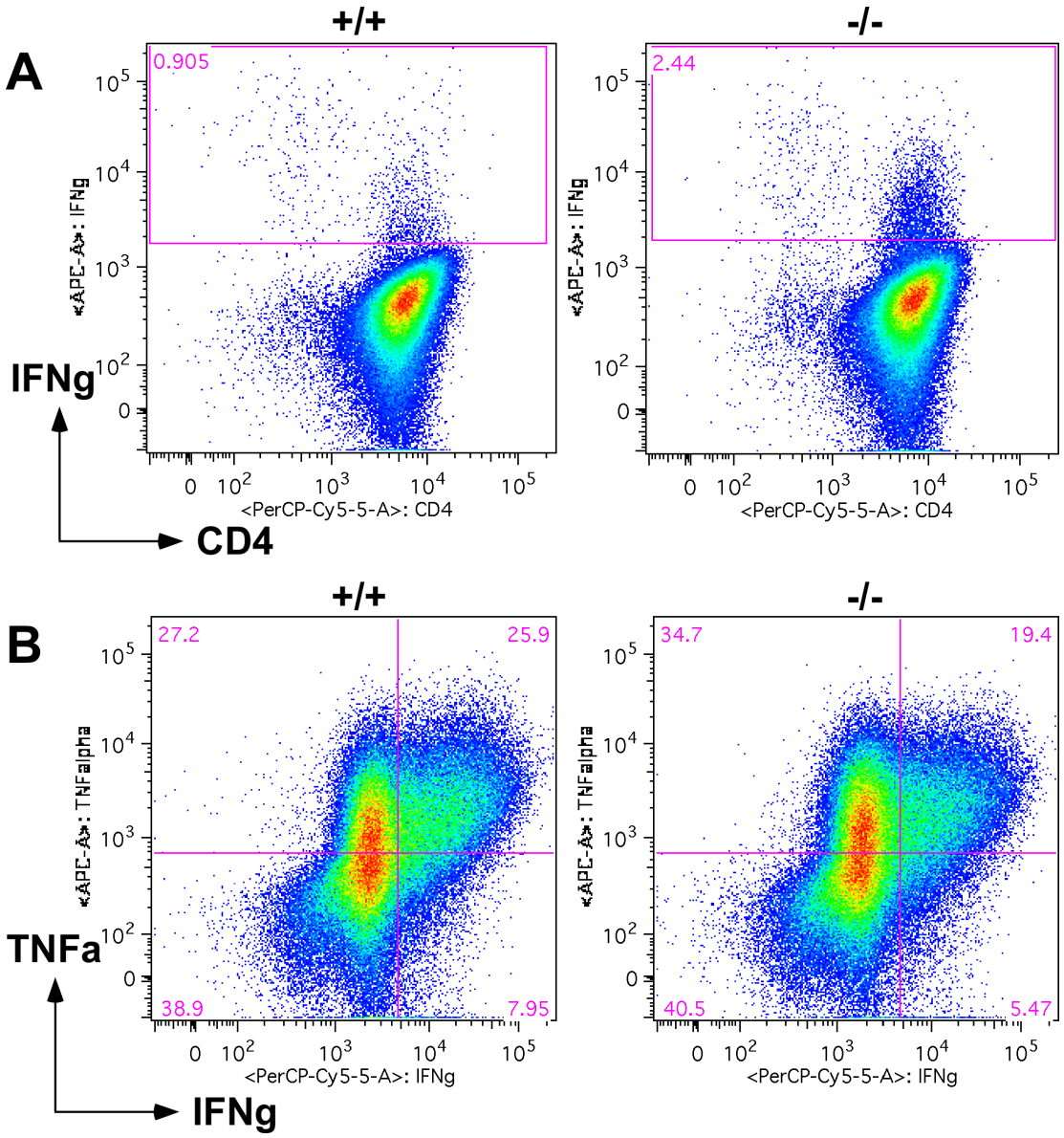


Figure A2.4 Cytokine expression in activated T-cells

CD4+ and CD8+ T-cells purified from ZASC1$^{+/+}$ and ZASC1$^{-/-}$ mice and activated in culture using α-CD3 and α-CD28 antibodies. 4 days after activation, cultures were stimulated and cytokine profiles were measured by flow cytometry after staining with α-TNFα, α-IFNγ antibodies. A.) CD4+ T-cells from ZASC1$^{+/+}$ (left) and ZASC1$^{-/-}$ (right) mice express similar levels of IFNγ. B.) CD8+ T-cells from ZASC1$^{+/+}$ (left) and ZASC1$^{-/-}$ (right) express both IFNγ and TNFα.

Next, we wanted to determine whether primary cells from the ZASC1 knockout mice would show a similar MLV infection phenotype to published studies using established cell lines treated with ZASC1 siRNA [45]. We examined ZASC1 function in primary cells from ZASC1$^{+/+}$ and ZASC1$^{-/-}$ mice. Surprisingly, we found that MLV infection in all ZASC1$^{-/-}$ primary cell types tested to be normal. MEFs (Chapter 3), CD4+ T-cells (Fig. A2.5) and whole bone marrow (Fig. A2.6) from ZASC1$^{+/+}$ and ZASC1$^{-/-}$ mice showed comparable levels of MLV-luc (VSV-G) gene expression. Follow up experiments to determine whether there was a possible redundant factor masking ZASC1 function were done and these results are discussed in Chapter 3.

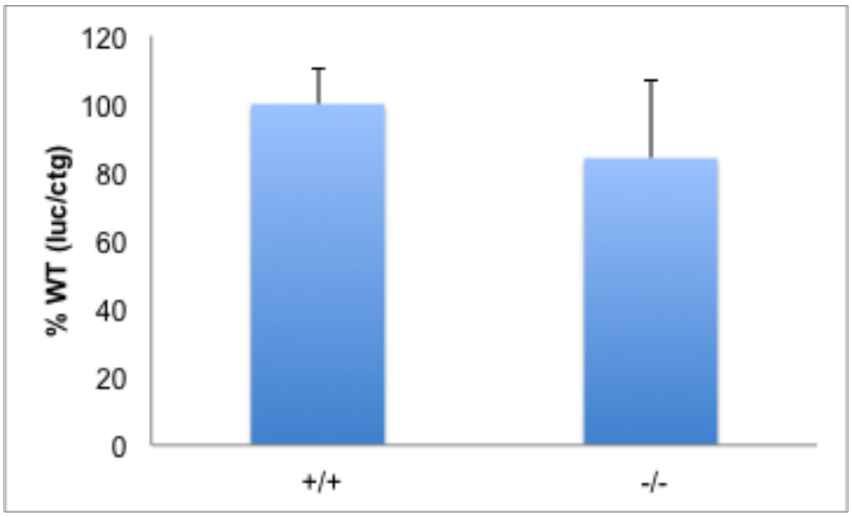


Figure A2.5 MLV-luc (VSV-G) Infection in Primary CD4+ T-cells

CD4+ T-cells purified from ZASC1$^{+/+}$ and ZASC1$^{-/-}$ mice were cultured in activating conditions using α-CD3 and α-CD28 along with IL-2 for 24 hours before virus challenge. MLV-luc (VSV-G) was added to T-cells in 96-well format. 30 hours post infection, viral luciferase was measured by BrightGlo (luc) and normalized to cell viability (CellTiterGlo). Error bars show SD. Graph is normalized to +/+ genotype. Graph shown is representative of three independent experiments.

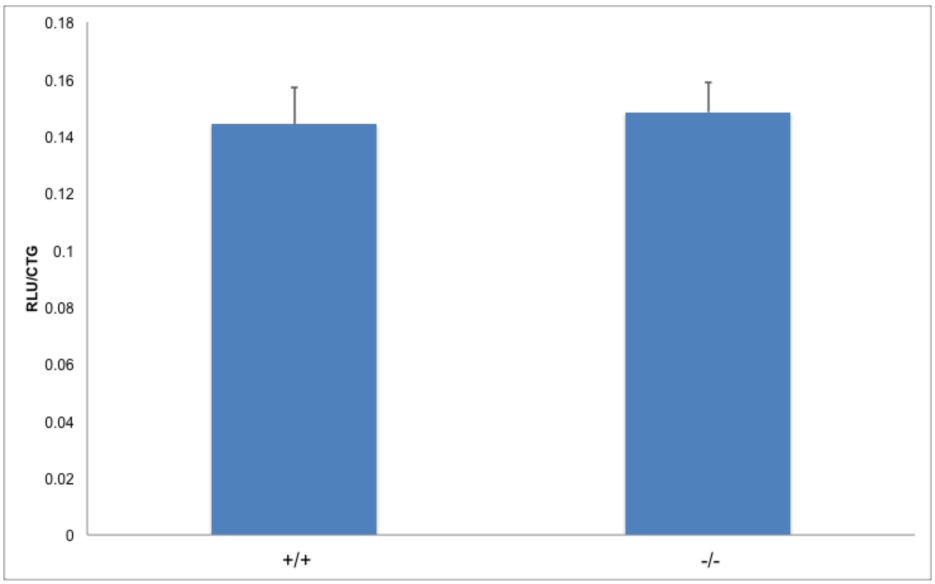


Figure A2.6 MLV-luc (VSV-G) Infection in Primary Bone Marrow

Bone marrow from 10 day old ZASC1$^{+/+}$ and ZASC1$^{-/-}$ mice was harvested and cultured for 24 hours before challenging with an MLV-luc (VSV-G) reporter virus. 30 hours post infection, viral luciferase was measured by BrightGlo (luc) and normalized to cell viability (ctg). Error bars show SD. Graph is normalized to +/+ genotype. Graph shown is representative of three independent experiments.

A2.3 Discussion

Although previous work has utilized ZASC1 siRNA and shRNA as a tool to study the mechanism of ZASC1 function in retrovirus replication, in my hands, strong knockdown of ZASC1 in immortalized cell lines caused clear growth defects. This created a number of challenges to studying ZASC1 in cultured cells. However, it also led to an interesting question about the role of ZASC1 in different cell types. If immortalized cells require ZASC1 function but most primary cells do not, it is possible that immortalized cells have selected for ZASC1 activity. Alternatively, normal cells might express a redundant factor that is down regulated in immortalized cells. Evidence for a redundant factor came from experiments in ZASC1$^{-/-}$ MEFs showing that an MLV virus with mZBS1 supported reduced levels of infection (Chapter 3). If the redundant factor is differentially expressed in primary versus immortalized cells, this may partially explain the difference in HIV and MLV infection studies performed in these different environments.

First, T-cell activation and proliferation were measured in culture to determine if ZASC1 plays any role in this process. T-cell activation under optimized conditions is functional, however, this is done with an artificially high level of α-CD3 antibody, which can mask more subtle activation defects. We show that titrating down the amount of α-CD3 results in a subtle T-cell activation defect in older mice. Although younger mice (~3 weeks of age) do not seem to have this defect, there is evidence that the immune system is not fully developed

[158]. Although it is clear from pathogenesis studies in Chapter 3 that any T-cell activation defect does not affect MLV pathogenesis, we cannot rule out other potential roles for this effect *in vivo*. Further studies are needed to characterize this defect in more detail. It has been shown that ZASC1 is phosphorylated after T-cell activation [156] so it is possible that ZASC1 plays a role in T-cell activation.

With the generation of the ZASC1 mouse model, we were able to purify and culture primary cells that provide a complete ZASC1 knockout in which to study the function of this transcription factor. We tested the role of ZASC1 in MLV infection in cultured primary cells. ZASC1 genotype had no effect on MLV infection in primary MEFs. Primary bone marrow and CD4+ T-cells similarly showed no difference in MLV infection based on ZASC1 genotype. Expression of ZASC1 in these tissues was confirmed by qRTPCR (data not shown). However, studies discussed in Chapter 3 have indicated that a redundant factor may be masking the infection phenotype in ZASC1$^{-/-}$ primary cells. Further studies are needed to determine if another cellular factor is binding to the ZBS and whether this interaction is competitive in the presence of ZASC1.

Based on this data it is not surprising that microarray studies using CD4+ T-cells and RNA-seq from primary MEFs revealed very few differences in global gene expression when comparing ZASC1$^{+/+}$ and ZASC1$^{-/-}$ cells (data not shown). Currently, Jim Bruce a co-member of the Young and Ahlquist labs is performing ChIP-Seq studies to identify more broadly ZASC1 regulated sites within the genome.

A2.4 Methods

Generating siRNAs targeting ZASC1

3 siRNAs were cloned individually with ZASC1 specific sequences using pSUPERsiGFP (Gift from Oded Singer, Verma Lab). Cassette was PCR amplified using T3 primer: 5'-CTCGAAATTAACCCTCACTAAAGGG-3' and one of 3 oligos containing the hairpin sequence: ss54: 5'-CTGTCTAGACAAAAATGAAGAACCGCCAGCTAAATCTCTTGAATTTAGCTGGCGGTTCTTCAGGGGATCTGTGGTCTCATACA-3' ss55: 5'-CTGTCTAGACAAAAA GAAAGTTTCTCTACCAATA TCTCTTGAA TATTGGTAGAGAAACTTTC GGGGATCTGTGGTCTCATACA-3' ss56: 5'-CTGTCTAGACAAAAA GGAGTAGACTTCACACAAA TCTCTTGAA TTTGTGTGAAGTCTACTCC GGGGATCTGTGGTCTCATACA-3' Fragment was purified by gel purification, digested with Xba1 and ligated into plenty (linearized with Xba1). Diagnostic digest with Ssp1was completed wherein a 1.6kb band for correct orientation, 1.4kb band for reverse orientation. Vectors with the correct orientation were sequenced using H1-F: 5'-TGGCAGGAAGATGGCTGTGA-3' to ensure the correct hairpin sequence was present.

Growth curve

Cells were seeded in 10cm plates and allowed to sit down overnight. Media was then aspirated and replaced with virus containing supernatant in a total volume of 4ml. Plates were incubated at 37°C for 6-8 hours before

additional media was added to a total volume of 10ml. Plates were incubated 48 hours in order to allow shRNAs time to knockdown targets. Plates were then trypsinized and cells were counted and replated in 96-well format (5000 cells per well). At 24-hour intervals, wells were trypsinized and total cell number was counted using a hemocytometer and trypan blue exclusion. 3 independent wells were counted for each condition per time point.

MLV infection in primary cells

Spleen and lymph nodes were purified from 4-week old mice. Organs were ground into a single cell suspension through a 70uM nylon mesh cell strainer and washed through using cell culture media (RPMI + 10% FBS, P/S, Sodium Pyruvate, βME). Cells were then spun down at 4°C 1200xg for 5 minutes and resuspended in MACS Buffer (1ml per mouse). Cell sorting was done using Dynabead FlowComp CD4+ T-cell isolation kit (Life Technologies) per manufacturers protocol. Once purified CD4+ T-cells were released from beads, resuspended in culture media at $1x10^6$ cells per ml and plated in α-CD3, α-CD28 coated 24-well plates with IL-2 supplemented culture media. 24 hours post plating, cells were counted and plated into 96-well plates (also antibody coated) with virus and were spinnoculated for 1 hour at RT (1200xg). Plates were incubated 24-30 hours at 37°C before being transferred to black plates, mixed with BrightGlo (MLV infected wells) or CellTiterGlo (uninfected wells) for luciferase and cell viability readout using a Topcount NXT HTS. All samples were plated in triplicate.

MLV infection in primary undifferentiated bone marrow

Bone marrow was harvested from 10-12 day old ZASC1$^{+/+}$ and ZASC1$^{-/-}$ leg bones. Marrow was dislodged from bones by twisting and crushing with a 5ml pipette tip and then run through a 70uM cell strainer into cell culture media (DMEM +10%FBS + P/S). Single cell suspension was spun down (1200xg) and resuspended in ACK lysis buffer for 2-3 minutes followed by a wash step and resuspension in cell culture media at 1x10^{6} cells per ml and incubated at 37°C overnight. Cells were then counted and plated for infection in 96-well plates (60,000 cells per well) in triplicate and virus was added. Plates were spinnoculated for 1 hour at RT (1200xg) and then incubated 30 hours at 37°C before Bright Glo and CellTiterGlo readout (described above).

In vitro T-cell activation assay

CD4+ and CD8+ T-cells were purified from mice 3-6 weeks of age using the Dynabead FlowComp positive isolation kits described earlier. Next, purified cells were spun down and resuspended in room temperature PBS with 5% FBS (900ul total volume) and mixed with CFSE in PBS with 5% FBS (100ul). 1mL staining mixtures were incubated for 5 minutes at room temperature in the dark and then 10ml of PBS with 5%FBS was added and cells were pelleted. A wash step was repeated for a second time before resuspending cells in culture media with IL-2 and plating on α-CD3 (2ug/ml), α-CD28 (2ug/ml) coated 24-well plates. 4 days after plating, cells were harvested and stained with α-CD4 or α-CD8

fluorophore conjugated antibodies and were analyzed by flow cytometry on the FACSAria III. In the titration assay, all conditions were the same except α-CD28 antibody concentration was held steady at 1ug/ml and α-CD3 concentration was serially diluted from 1ug/ml down to 0.2ug/ml. Activation and readout were the same as described above.

Cytokine expression after T-cell activation

T-cell purification and activation was completed as described above for the T-cell activation assay except in the absence of CFSE labeling. Plates were coated with α-CD3 (2ug/ml) and α-CD28 (2ug/ml) and 4 days after activation, cells were stimulated with PMA and Ionamycin for 4 hours by directly adding it to the media. 1 hour after adding Golgi Plug was added as well to inhibit secretion of cytokines into the media. Intracellular staining was performed by spinning down cells and resuspending in PBS on ice. Next cells were spun down again and resuspended in BD fixed permeablization buffer and incubated on ice for 20 minutes. After two rounds of PBS wash cells were resuspended in permeablization buffer and incubated on ice for 30 minutes with antibodies for staining. For staining of CD4+ T-cells α-CD4, α-IFNgamma, α-IL4 were used. For CD8+ T-cells, α-TNFα and α-IFNgamma were used. After staining, cells were washed in permeablization buffer and resuspended in PBS for FACS analysis.

References

1. Shafer, R.W. and J.M. Schapiro, *HIV-1 drug resistance mutations: an updated framework for the second decade of HAART.* AIDS reviews, 2008. **10**(2): p. 67-84.

2. Swindle, C.S. and C.A. Klug, *Mechanisms that regulate silencing of gene expression from retroviral vectors.* J Hematother Stem Cell Res, 2002. **11**(3): p. 449-56.

3. Ruddell, A., *Transcription regulatory elements of the avian retroviral long terminal repeat.* Virology, 1995. **206**(1): p. 1-7.

4. Michalides, R., E. Wagenaar, J. Hilkens, J. Hilgers, B. Groner, and N.E. Hynes, *Acquisition of proviral DNA of mouse mammary tumor virus in thymic leukemia cells from GR mice.* Journal of virology, 1982. **43**(3): p. 819-29.

5. Li, Y., E. Golemis, J.W. Hartley, and N. Hopkins, *Disease specificity of nondefective Friend and Moloney murine leukemia viruses is controlled by a small number of nucleotides.* J Virol, 1987. **61**(3): p. 693-700.

6. Hsu, C.L., C. Fabritius, and J. Dudley, *Mouse mammary tumor virus proviruses in T-cell lymphomas lack a negative regulatory element in the long terminal repeat.* Journal of virology, 1988. **62**(12): p. 4644-52.

7. Dudley, J. and R. Risser, *Amplification and novel locations of endogenous mouse mammary tumor virus genomes in mouse T-cell lymphomas.* Journal of virology, 1984. **49**(1): p. 92-101.

8. Rous, P., *A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells.* The Journal of experimental medicine, 1911. **13**(4): p. 397-411.

9. Poiesz, B.J., F.W. Ruscetti, A.F. Gazdar, P.A. Bunn, J.D. Minna, and R.C. Gallo, *Detection and isolation of type C retrovirus particles from fresh and*

cultured lymphocytes of a patient with cutaneous T-cell lymphoma. Proc Natl Acad Sci U S A, 1980. **77**(12): p. 7415-9.

10. Yoshida, M., I. Miyoshi, and Y. Hinuma, *Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease.* Proc Natl Acad Sci U S A, 1982. **79**(6): p. 2031-5.

11. Temin, H.M. and S. Mizutani, *RNA-dependent DNA polymerase in virions of Rous sarcoma virus.* Nature, 1970. **226**(5252): p. 1211-3.

12. Baltimore, D., *RNA-dependent DNA polymerase in virions of RNA tumour viruses.* Nature, 1970. **226**(5252): p. 1209-11.

13. Miyauchi, K., Y. Kim, O. Latinovic, V. Morozov, and G.B. Melikyan, *HIV enters cells via endocytosis and dynamin-dependent fusion with endosomes.* Cell, 2009. **137**(3): p. 433-44.

14. Barnard, R.J., S. Narayan, G. Dornadula, M.D. Miller, and J.A. Young, *Low pH is required for avian sarcoma and leukosis virus Env-dependent viral penetration into the cytosol and not for viral uncoating.* J Virol, 2004. **78**(19): p. 10433-41.

15. Leung, J., A. Yueh, F.S. Appah, Jr., B. Yuan, K. de los Santos, and S.P. Goff, *Interaction of Moloney murine leukemia virus matrix protein with IQGAP.* The EMBO journal, 2006. **25**(10): p. 2155-66.

16. Luban, J., K.L. Bossolt, E.K. Franke, G.V. Kalpana, and S.P. Goff, *Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B.* Cell, 1993. **73**(6): p. 1067-78.

17. Braaten, D. and J. Luban, *Cyclophilin A regulates HIV-1 infectivity, as demonstrated by gene targeting in human T cells.* The EMBO journal, 2001. **20**(6): p. 1300-9.

18. Ciuffi, A., M. Llano, E. Poeschla, C. Hoffmann, J. Leipzig, P. Shinn, J.R. Ecker, and F. Bushman, *A role for LEDGF/p75 in targeting HIV DNA integration.* Nat Med, 2005. **11**(12): p. 1287-9.

19. Maertens, G., P. Cherepanov, W. Pluymers, K. Busschots, E. De Clercq, Z. Debyser, and Y. Engelborghs, *LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells.* J Biol Chem, 2003. **278**(35): p. 33528-39.

20. Cherepanov, P., G. Maertens, P. Proost, B. Devreese, J. Van Beeumen, Y. Engelborghs, E. De Clercq, and Z. Debyser, *HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells.* J Biol Chem, 2003. **278**(1): p. 372-81.

21. Martin-Serrano, J., T. Zang, and P.D. Bieniasz, *Role of ESCRT-I in retroviral budding.* Journal of virology, 2003. **77**(8): p. 4794-804.

22. Usami, Y., S. Popov, E. Popova, M. Inoue, W. Weissenhorn, and G.G. H, *The ESCRT pathway and HIV-1 budding.* Biochemical Society transactions, 2009. **37**(Pt 1): p. 181-4.

23. Konig, R., Y. Zhou, D. Elleder, T.L. Diamond, G.M. Bonamy, J.T. Irelan, C.Y. Chiang, B.P. Tu, P.D. De Jesus, C.E. Lilley, S. Seidel, A.M. Opaluch, J.S. Caldwell, M.D. Weitzman, K.L. Kuhen, S. Bandyopadhyay, T. Ideker, A.P. Orth, L.J. Miraglia, F.D. Bushman, J.A. Young, and S.K. Chanda, *Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication.* Cell, 2008. **135**(1): p. 49-60.

24. Zhou, H., M. Xu, Q. Huang, A.T. Gates, X.D. Zhang, J.C. Castle, E. Stec, M. Ferrer, B. Strulovici, D.J. Hazuda, and A.S. Espeseth, *Genome-scale RNAi screen for host factors required for HIV replication.* Cell Host Microbe, 2008. **4**(5): p. 495-504.

25. Bushman, F.D., N. Malani, J. Fernandes, I. D'Orso, G. Cagney, T.L. Diamond, H. Zhou, D.J. Hazuda, A.S. Espeseth, R. Konig, S. Bandyopadhyay, T. Ideker, S.P. Goff, N.J. Krogan, A.D. Frankel, J.A. Young, and S.K. Chanda, *Host cell factors in HIV replication: meta-analysis of genome-wide studies.* PLoS pathogens, 2009. **5**(5): p. e1000437.

26. Brass, A.L., D.M. Dykxhoorn, Y. Benita, N. Yan, A. Engelman, R.J. Xavier, J. Lieberman, and S.J. Elledge, *Identification of host proteins required for HIV infection through a functional genomic screen.* Science, 2008. **319**(5865): p. 921-6.

27. Albritton, L.M., L. Tseng, D. Scadden, and J.M. Cunningham, *A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection.* Cell, 1989. **57**(4): p. 659-66.

28. Miller, D.G., R.H. Edwards, and A.D. Miller, *Cloning of the cellular receptor for amphotropic murine retroviruses reveals homology to that for gibbon ape leukemia virus.* Proc Natl Acad Sci U S A, 1994. **91**(1): p. 78-82.

29. Battini, J.L., J.E. Rasko, and A.D. Miller, *A human cell-surface receptor for xenotropic and polytropic murine leukemia viruses: possible role in G protein-coupled signal transduction.* Proc Natl Acad Sci U S A, 1999. **96**(4): p. 1385-90.

30. Wilson, C.A., K.B. Farrell, and M.V. Eiden, *Properties of a unique form of the murine amphotropic leukemia virus receptor expressed on hamster cells.* J Virol, 1994. **68**(12): p. 7697-703.

31. Miller, D.G. and A.D. Miller, *A family of retroviruses that utilize related phosphate transporters for cell entry.* J Virol, 1994. **68**(12): p. 8270-6.

32. Roe, T., T.C. Reynolds, G. Yu, and P.O. Brown, *Integration of murine leukemia virus DNA depends on mitosis.* Embo J, 1993. **12**(5): p. 2099-108.

33. Manley, N.R., M. O'Connell, W. Sun, N.A. Speck, and N. Hopkins, *Two factors that bind to highly conserved sequences in mammalian type C retroviral enhancers.* J Virol, 1993. **67**(4): p. 1967-75.

34. Sun, W., B.J. Graves, and N.A. Speck, *Transactivation of the Moloney murine leukemia virus and T-cell receptor beta-chain enhancers by cbf and ets requires intact binding sites for both proteins.* J Virol, 1995. **69**(8): p. 4941-9.

35. Sun, W., M. O'Connell, and N.A. Speck, *Characterization of a protein that binds multiple sequences in mammalian type C retrovirus enhancers.* J Virol, 1993. **67**(4): p. 1976-86.

36. LaMarco, K., C.C. Thompson, B.P. Byers, E.M. Walton, and S.L. McKnight, *Identification of Ets- and notch-related subunits in GA binding protein.* Science, 1991. **253**(5021): p. 789-92.

37. Wotton, D., J. Ghysdael, S. Wang, N.A. Speck, and M.J. Owen, *Cooperative binding of Ets-1 and core binding factor to DNA.* Mol Cell Biol, 1994. **14**(1): p. 840-50.

38. Moloney, J.B., *The Rodent Leukemias: Virus-Induced Murine Leukemias.* Annu Rev Med, 1964. **15**: p. 383-92.

39. Reddy, E.P., C.Y. Dunn, and S.A. Aaronson, *Different lymphoid cell targets by transformation by replication-competent Moloney and Rauscher mouse leukemia viruses.* Cell, 1980. **19**(3): p. 663-9.

40. Fan, H., *Leukemogenesis by Moloney murine leukemia virus: a multistep process.* Trends in microbiology, 1997. **5**(2): p. 74-82.

41. Davis, B., E. Linney, and H. Fan, *Suppression of leukaemia virus pathogenicity by polyoma virus enhancers.* Nature, 1985. **314**(6011): p. 550-3.

42. Brightman, B.K., B.R. Davis, and H. Fan, *Preleukemic hematopoietic hyperplasia induced by Moloney murine leukemia virus is an indirect consequence of viral infection.* J Virol, 1990. **64**(9): p. 4582-4.

43. Belli, B. and H. Fan, *The leukemogenic potential of an enhancer variant of Moloney murine leukemia virus varies with the route of inoculation.* J Virol, 1994. **68**(11): p. 6883-9.

44. van Lohuizen, M., S. Verbeek, B. Scheijen, E. Wientjens, H. van der Gulden, and A. Berns, *Identification of cooperating oncogenes in E mu-myc transgenic mice by provirus tagging.* Cell, 1991. **65**(5): p. 737-52.

45. Bruce, J.W., M. Hierl, J.A. Young, and P. Ahlquist, *Cellular transcription factor ZASC1 regulates murine leukemia virus transcription.* J Virol, 2010. **84**(15): p. 7473-83.

46. CDC, *The Global HIV/AIDS pandemic, 2006.* MMWR Morb Mortal Wkly Rep, 2006. **55**(31): p. 841-4.

47. CDC, *A Cluster of Kaposi's Sarcoma and Pneumocystis carinii Pneumonia among Homosexual Male Residents of Los Angeles and range Counties, California* MMWR Morb Mortal Wkly Rep, 1982(31): p. 305-307.

48. Klatzmann, D., E. Champagne, S. Chamaret, J. Gruest, D. Guetard, T. Hercend, J.C. Gluckman, and L. Montagnier, *T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV.* Nature, 1984. **312**(5996): p. 767-8.

49. Dalgleish, A.G., P.C. Beverley, P.R. Clapham, D.H. Crawford, M.F. Greaves, and R.A. Weiss, *The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus.* Nature, 1984. **312**(5996): p. 763-7.

50. McDougal, J.S., M.S. Kennedy, J.M. Sligh, S.P. Cort, A. Mawle, and J.K. Nicholson, *Binding of HTLV-III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule.* Science, 1986. **231**(4736): p. 382-5.

51. Alkhatib, G., C. Combadiere, C.C. Broder, Y. Feng, P.E. Kennedy, P.M. Murphy, and E.A. Berger, *CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1.* Science, 1996. **272**(5270): p. 1955-8.

52. Choe, H., M. Farzan, Y. Sun, N. Sullivan, B. Rollins, P.D. Ponath, L. Wu, C.R. Mackay, G. LaRosa, W. Newman, N. Gerard, C. Gerard, and J. Sodroski, *The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates.* Cell, 1996. **85**(7): p. 1135-48.

53. Deng, H., R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. Di Marzio, S. Marmon, R.E. Sutton, C.M. Hill, C.B. Davis, S.C. Peiper, T.J. Schall, D.R. Littman, and N.R. Landau, *Identification of a major co-receptor for primary isolates of HIV-1.* Nature, 1996. **381**(6584): p. 661-6.

54. Doranz, B.J., J. Rucker, Y. Yi, R.J. Smyth, M. Samson, S.C. Peiper, M. Parmentier, R.G. Collman, and R.W. Doms, *A dual-tropic primary HIV-1*

isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. Cell, 1996. **85**(7): p. 1149-58.

55. Dragic, T., V. Litwin, G.P. Allaway, S.R. Martin, Y. Huang, K.A. Nagashima, C. Cayanan, P.J. Maddon, R.A. Koup, J.P. Moore, and W.A. Paxton, *HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5.* Nature, 1996. **381**(6584): p. 667-73.

56. Feng, Y., C.C. Broder, P.E. Kennedy, and E.A. Berger, *HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor.* Science, 1996. **272**(5263): p. 872-7.

57. Oberlin, E., A. Amara, F. Bachelerie, C. Bessia, J.L. Virelizier, F. Arenzana-Seisdedos, O. Schwartz, J.M. Heard, I. Clark-Lewis, D.F. Legler, M. Loetscher, M. Baggiolini, and B. Moser, *The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1.* Nature, 1996. **382**(6594): p. 833-5.

58. Zhu, T., H. Mo, N. Wang, D.S. Nam, Y. Cao, R.A. Koup, and D.D. Ho, *Genotypic and phenotypic characterization of HIV-1 patients with primary infection.* Science, 1993. **261**(5125): p. 1179-81.

59. Soros, V.B., W. Yonemoto, and W.C. Greene, *Newly synthesized APOBEC3G is incorporated into HIV virions, inhibited by HIV RNA, and subsequently activated by RNase H.* PLoS Pathog, 2007. **3**(2): p. e15.

60. Sheehy, A.M., N.C. Gaddis, J.D. Choi, and M.H. Malim, *Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein.* Nature, 2002. **418**(6898): p. 646-50.

61. Low, A., C.M. Okeoma, N. Lovsin, M. de las Heras, T.H. Taylor, B.M. Peterlin, S.R. Ross, and H. Fan, *Enhanced replication and pathogenesis of Moloney murine leukemia virus in mice defective in the murine APOBEC3 gene.* Virology, 2009. **385**(2): p. 455-63.

62. Bogerd, H.P. and B.R. Cullen, *Single-stranded RNA facilitates nucleocapsid: APOBEC3G complex formation.* RNA, 2008. **14**(6): p. 1228-36.

63. Jager, S., D.Y. Kim, J.F. Hultquist, K. Shindo, R.S. LaRue, E. Kwon, M. Li, B.D. Anderson, L. Yen, D. Stanley, C. Mahon, J. Kane, K. Franks-Skiba, P. Cimermancic, A. Burlingame, A. Sali, C.S. Craik, R.S. Harris, J.D. Gross, and N.J. Krogan, *Vif hijacks CBF-beta to degrade APOBEC3G and promote HIV-1 infection.* Nature, 2012. **481**(7381): p. 371-5.

64. Varthakavi, V., R.M. Smith, S.P. Bour, K. Strebel, and P. Spearman, *Viral protein U counteracts a human host cell restriction that inhibits HIV-1 particle production.* Proc Natl Acad Sci U S A, 2003. **100**(25): p. 15154-9.

65. Van Damme, N., D. Goff, C. Katsura, R.L. Jorgenson, R. Mitchell, M.C. Johnson, E.B. Stephens, and J. Guatelli, *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, 2008. **3**(4): p. 245-52.

66. Neil, S.J., T. Zang, and P.D. Bieniasz, *Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu.* Nature, 2008. **451**(7177): p. 425-30.

67. Mangeat, B., G. Gers-Huber, M. Lehmann, M. Zufferey, J. Luban, and V. Piguet, *HIV-1 Vpu neutralizes the antiviral factor Tetherin/BST-2 by binding it and directing its beta-TrCP2-dependent degradation.* PLoS pathogens, 2009. **5**(9): p. e1000574.

68. Chaudhuri, R., O.W. Lindwasser, W.J. Smith, J.H. Hurley, and J.S. Bonifacino, *Downregulation of CD4 by human immunodeficiency virus type 1 Nef is dependent on clathrin and involves direct interaction of Nef with the AP2 clathrin adaptor.* J Virol, 2007. **81**(8): p. 3877-90.

69. Thoulouze, M.I., N. Sol-Foulon, F. Blanchet, A. Dautry-Varsat, O. Schwartz, and A. Alcover, *Human immunodeficiency virus type-1 infection impairs the formation of the immunological synapse.* Immunity, 2006. **24**(5): p. 547-61.

70. Binette, J., M. Dube, J. Mercier, D. Halawani, M. Latterich, and E.A. Cohen, *Requirements for the selective degradation of CD4 receptor molecules by the human immunodeficiency virus type 1 Vpu protein in the endoplasmic reticulum.* Retrovirology, 2007. **4**: p. 75.

71. Margottin, F., S.P. Bour, H. Durand, L. Selig, S. Benichou, V. Richard, D. Thomas, K. Strebel, and R. Benarous, *A novel human WD protein, h-beta TrCp, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif.* Mol Cell, 1998. **1**(4): p. 565-74.

72. Meusser, B. and T. Sommer, *Vpu-mediated degradation of CD4 reconstituted in yeast reveals mechanistic differences to cellular ER-associated protein degradation.* Mol Cell, 2004. **14**(2): p. 247-58.

73. Belzile, J.P., G. Duisit, N. Rougeau, J. Mercier, A. Finzi, and E.A. Cohen, *HIV-1 Vpr-mediated G2 arrest involves the DDB1-CUL4AVPRBP E3 ubiquitin ligase.* PLoS Pathog, 2007. **3**(7): p. e85.

74. DeHart, J.L., E.S. Zimmerman, O. Ardon, C.M. Monteiro-Filho, E.R. Arganaraz, and V. Planelles, *HIV-1 Vpr activates the G2 checkpoint through manipulation of the ubiquitin proteasome system.* Virol J, 2007. **4**: p. 57.

75. Hrecka, K., M. Gierszewska, S. Srivastava, L. Kozaczkiewicz, S.K. Swanson, L. Florens, M.P. Washburn, and J. Skowronski, *Lentiviral Vpr usurps Cul4-DDB1[VprBP] E3 ubiquitin ligase to modulate cell cycle.* Proc Natl Acad Sci U S A, 2007. **104**(28): p. 11778-83.

76. Le Rouzic, E., N. Belaidouni, E. Estrabaud, M. Morel, J.C. Rain, C. Transy, and F. Margottin-Goguet, *HIV1 Vpr arrests the cell cycle by recruiting DCAF1/VprBP, a receptor of the Cul4-DDB1 ubiquitin ligase.* Cell Cycle, 2007. **6**(2): p. 182-8.

77. Schrofelbauer, B., Y. Hakata, and N.R. Landau, *HIV-1 Vpr function is mediated by interaction with the damage-specific DNA-binding protein DDB1.* Proc Natl Acad Sci U S A, 2007. **104**(10): p. 4130-5.

78. Tan, L., E. Ehrlich, and X.F. Yu, *DDB1 and Cul4A are required for human immunodeficiency virus type 1 Vpr-induced G2 arrest.* J Virol, 2007. **81**(19): p. 10822-30.

79. Wen, X., K.M. Duus, T.D. Friedrich, and C.M. de Noronha, *The HIV1 protein Vpr acts to promote G2 cell cycle arrest by engaging a DDB1 and*

Cullin4A-containing ubiquitin ligase complex using VprBP/DCAF1 as an adaptor. J Biol Chem, 2007. **282**(37): p. 27046-57.

80. Mansky, L.M. and H.M. Temin, *Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase.* J Virol, 1995. **69**(8): p. 5087-94.

81. Perelson, A.S., A.U. Neumann, M. Markowitz, J.M. Leonard, and D.D. Ho, *HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time.* Science, 1996. **271**(5255): p. 1582-6.

82. Nabel, G. and D. Baltimore, *An inducible transcription factor activates expression of human immunodeficiency virus in T cells.* Nature, 1987. **326**(6114): p. 711-3.

83. Harrich, D., J. Garcia, F. Wu, R. Mitsuyasu, J. Gonazalez, and R. Gaynor, *Role of SP1-binding domains in in vivo transcriptional regulation of the human immunodeficiency virus type 1 long terminal repeat.* Journal of virology, 1989. **63**(6): p. 2585-91.

84. Chiang, C.M. and R.G. Roeder, *Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators.* Science, 1995. **267**(5197): p. 531-6.

85. Emili, A., J. Greenblatt, and C.J. Ingles, *Species-specific interaction of the glutamine-rich activation domains of Sp1 with the TATA box-binding protein.* Molecular and cellular biology, 1994. **14**(3): p. 1582-93.

86. Feinberg, M.B., D. Baltimore, and A.D. Frankel, *The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation.* Proc Natl Acad Sci U S A, 1991. **88**(9): p. 4045-9.

87. Laspia, M.F., A.P. Rice, and M.B. Mathews, *HIV-1 Tat protein increases transcriptional initiation and stabilizes elongation.* Cell, 1989. **59**(2): p. 283-92.

88. Garber, M.E., P. Wei, V.N. KewalRamani, T.P. Mayall, C.H. Herrmann, A.P. Rice, D.R. Littman, and K.A. Jones, *The interaction between HIV-1*

Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein. Genes Dev, 1998. **12**(22): p. 3512-27.

89. Wei, P., M.E. Garber, S.M. Fang, W.H. Fischer, and K.A. Jones, *A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA.* Cell, 1998. **92**(4): p. 451-62.

90. Kim, Y.K., C.F. Bourgeois, C. Isel, M.J. Churcher, and J. Karn, *Phosphorylation of the RNA polymerase II carboxyl-terminal domain by CDK9 is directly responsible for human immunodeficiency virus type 1 Tat-activated transcriptional elongation.* Mol Cell Biol, 2002. **22**(13): p. 4622-37.

91. Jeronimo, C., D. Forget, A. Bouchard, Q. Li, G. Chua, C. Poitras, C. Therien, D. Bergeron, S. Bourassa, J. Greenblatt, B. Chabot, G.G. Poirier, T.R. Hughes, M. Blanchette, D.H. Price, and B. Coulombe, *Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme.* Mol Cell, 2007. **27**(2): p. 262-74.

92. Krueger, B.J., C. Jeronimo, B.B. Roy, A. Bouchard, C. Barrandon, S.A. Byers, C.E. Searcey, J.J. Cooper, O. Bensaude, E.A. Cohen, B. Coulombe, and D.H. Price, *LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated.* Nucleic Acids Res, 2008. **36**(7): p. 2219-29.

93. Saez-Cirion, A., M. Sinet, S.Y. Shin, A. Urrutia, P. Versmisse, C. Lacabaratz, F. Boufassa, V. Avettand-Fenoel, C. Rouzioux, J.F. Delfraissy, F. Barre-Sinoussi, O. Lambotte, A. Venet, and G. Pancino, *Heterogeneity in HIV suppression by CD8 T cells from HIV controllers: association with Gag-specific CD8 T cell responses.* Journal of immunology, 2009. **182**(12): p. 7828-37.

94. Saez-Cirion, A., C. Lacabaratz, O. Lambotte, P. Versmisse, A. Urrutia, F. Boufassa, F. Barre-Sinoussi, J.F. Delfraissy, M. Sinet, G. Pancino, and A. Venet, *HIV controllers exhibit potent CD8 T cell capacity to suppress HIV infection ex vivo and peculiar cytotoxic T lymphocyte activation phenotype.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(16): p. 6776-81.

95. O'Connell, K.A., Y. Han, T.M. Williams, R.F. Siliciano, and J.N. Blankson, *Role of natural killer cells in a cohort of elite suppressors: low frequency of the protective KIR3DS1 allele and limited inhibition of human immunodeficiency virus type 1 replication in vitro.* Journal of virology, 2009. **83**(10): p. 5028-34.

96. Novembre, J., A.P. Galvani, and M. Slatkin, *The geographic spread of the CCR5 Delta32 HIV-resistance allele.* PLoS biology, 2005. **3**(11): p. e339.

97. Hutter, G., D. Nowak, M. Mossner, S. Ganepola, A. Mussig, K. Allers, T. Schneider, J. Hofmann, C. Kucherer, O. Blau, I.W. Blau, W.K. Hofmann, and E. Thiel, *Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation.* The New England journal of medicine, 2009. **360**(7): p. 692-8.

98. Allers, K., G. Hutter, J. Hofmann, C. Loddenkemper, K. Rieger, E. Thiel, and T. Schneider, *Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation.* Blood, 2011. **117**(10): p. 2791-9.

99. Fauci, A.S., G. Pantaleo, S. Stanley, and D. Weissman, *Immunopathogenic mechanisms of HIV infection.* Annals of internal medicine, 1996. **124**(7): p. 654-63.

100. Collier, A.C., R.W. Coombs, D.A. Schoenfeld, R. Bassett, A. Baruch, and L. Corey, *Combination therapy with zidovudine, didanosine and saquinavir.* Antiviral Res, 1996. **29**(1): p. 99.

101. Jacobsen, H., M. Haenggi, M. Ott, I.B. Duncan, M. Andreoni, S. Vella, and J. Mous, *Reduced sensitivity to saquinavir: an update on genotyping from phase I/II trials.* Antiviral Res, 1996. **29**(1): p. 95-7.

102. Hammer, S.M., K.E. Squires, M.D. Hughes, J.M. Grimes, L.M. Demeter, J.S. Currier, J.J. Eron, Jr., J.E. Feinberg, H.H. Balfour, Jr., L.R. Deyton, J.A. Chodakewitz, and M.A. Fischl, *A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team.* The New England journal of medicine, 1997. **337**(11): p. 725-33.

103. Gulick, R.M., J.W. Mellors, D. Havlir, J.J. Eron, C. Gonzalez, D. McMahon, D.D. Richman, F.T. Valentine, L. Jonas, A. Meibohm, E.A. Emini, and J.A. Chodakewitz, *Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy.* The New England journal of medicine, 1997. **337**(11): p. 734-9.

104. Finzi, D., M. Hermankova, T. Pierson, L.M. Carruth, C. Buck, R.E. Chaisson, T.C. Quinn, K. Chadwick, J. Margolick, R. Brookmeyer, J. Gallant, M. Markowitz, D.D. Ho, D.D. Richman, and R.F. Siliciano, *Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy.* Science, 1997. **278**(5341): p. 1295-300.

105. Grant, P.M. and A.R. Zolopa, *Optimal antiretroviral therapy: HIV-1 treatment strategies to avoid and overcome drug resistance.* Curr Opin Investig Drugs, 2010. **11**(8): p. 901-10.

106. Montessori, V., N. Press, M. Harris, L. Akagi, and J.S. Montaner, *Adverse effects of antiretroviral therapy for HIV infection.* CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne, 2004. **170**(2): p. 229-38.

107. Maeda, K., H. Nakata, Y. Koh, T. Miyakawa, H. Ogata, Y. Takaoka, S. Shibayama, K. Sagawa, D. Fukushima, J. Moravek, Y. Koyanagi, and H. Mitsuya, *Spirodiketopiperazine-based CCR5 inhibitor which preserves CC-chemokine/CCR5 interactions and exerts potent activity against R5 human immunodeficiency virus type 1 in vitro.* Journal of virology, 2004. **78**(16): p. 8654-62.

108. Westby, M., C. Smith-Burchnell, J. Mori, M. Lewis, M. Mosley, M. Stockdale, P. Dorr, G. Ciaramella, and M. Perros, *Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the CCR5 antagonist maraviroc utilize inhibitor-bound receptor for entry.* Journal of virology, 2007. **81**(5): p. 2359-71.

109. Trkola, A., S.E. Kuhmann, J.M. Strizki, E. Maxwell, T. Ketas, T. Morgan, P. Pugach, S. Xu, L. Wojcik, J. Tagat, A. Palani, S. Shapiro, J.W. Clader, S. McCombie, G.R. Reyes, B.M. Baroudy, and J.P. Moore, *HIV-1 escape from a small molecule, CCR5-specific entry inhibitor does not involve CXCR4 use.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(1): p. 395-400.

110. Mori J, M.M., Lewis M, et al., *Characterization of maraviroc resistance in patients failing treatment with CCR5-tropicvirus in MOTIVATE 1 and 2.* Antivir Ther., 2007. **12**(S12): p. [abstract 10].

111. Harrigan, P.R., M. Whaley, and J.S. Montaner, *Rate of HIV-1 RNA rebound upon stopping antiretroviral therapy.* Aids, 1999. **13**(8): p. F59-62.

112. DesGroseillers, L. and P. Jolicoeur, *The tandem direct repeats within the long terminal repeat of murine leukemia viruses are the primary determinant of their leukemogenic potential.* J Virol, 1984. **52**(3): p. 945-52.

113. DiFronzo, N.L. and C.A. Holland, *Sequence-specific and/or stereospecific constraints of the U3 enhancer elements of MCF 247-W are important for pathogenicity.* J Virol, 1999. **73**(1): p. 234-41.

114. Lenz, J., D. Celander, R.L. Crowther, R. Patarca, D.W. Perkins, and W.A. Haseltine, *Determination of the leukaemogenicity of a murine retrovirus by sequences within the long terminal repeat.* Nature, 1984. **308**(5958): p. 467-70.

115. Speck, N.A., B. Renjifo, E. Golemis, T.N. Fredrickson, J.W. Hartley, and N. Hopkins, *Mutation of the core or adjacent LVb elements of the Moloney murine leukemia virus enhancer alters disease specificity.* Genes Dev, 1990. **4**(2): p. 233-42.

116. Lewis, A.F., T. Stacy, W.R. Green, L. Taddesse-Heath, J.W. Hartley, and N.A. Speck, *Core-binding factor influences the disease specificity of Moloney murine leukemia virus.* J Virol, 1999. **73**(7): p. 5535-47.

117. Margolick, J.B., D.J. Volkman, T.M. Folks, and A.S. Fauci, *Amplification of HTLV-III/LAV infection by antigen-induced activation of T cells and direct suppression by virus of lymphocyte blastogenic responses.* Journal of immunology, 1987. **138**(6): p. 1719-23.

118. Ho, D.D., A.U. Neumann, A.S. Perelson, W. Chen, J.M. Leonard, and M. Markowitz, *Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection.* Nature, 1995. **373**(6510): p. 123-6.

119. Althaus, C.L. and R.J. De Boer, *Implications of CTL-mediated killing of HIV-infected cells during the non-productive stage of infection.* PLoS One, 2011. **6**(2): p. e16468.

120. Duh, E.J., W.J. Maury, T.M. Folks, A.S. Fauci, and A.B. Rabson, *Tumor necrosis factor alpha activates human immunodeficiency virus type 1 through induction of nuclear factor binding to the NF-kappa B sites in the long terminal repeat.* Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(15): p. 5974-8.

121. Chun, T.W., L. Stuyver, S.B. Mizell, L.A. Ehler, J.A. Mican, M. Baseler, A.L. Lloyd, M.A. Nowak, and A.S. Fauci, *Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(24): p. 13193-7.

122. Wong, J.K., M. Hezareh, H.F. Gunthard, D.V. Havlir, C.C. Ignacio, C.A. Spina, and D.D. Richman, *Recovery of replication-competent HIV despite prolonged suppression of plasma viremia.* Science, 1997. **278**(5341): p. 1291-5.

123. Klug, A., *The discovery of zinc fingers and their applications in gene regulation and genome manipulation.* Annu Rev Biochem, 2010. **79**: p. 213-31.

124. Omichinski, J.G., P.V. Pedone, G. Felsenfeld, A.M. Gronenborn, and G.M. Clore, *The solution structure of a specific GAGA factor-DNA complex reveals a modular binding mode.* Nat Struct Biol, 1997. **4**(2): p. 122-32.

125. Pedone, P.V., R. Ghirlando, G.M. Clore, A.M. Gronenborn, G. Felsenfeld, and J.G. Omichinski, *The single Cys2-His2 zinc finger domain of the GAGA protein flanked by basic residues is sufficient for high-affinity specific DNA binding.* Proc Natl Acad Sci U S A, 1996. **93**(7): p. 2822-6.

126. Ruiz i Altaba, A., H. Perry-O'Keefe, and D.A. Melton, *Xfin: an embryonic gene encoding a multifingered protein in Xenopus.* Embo J, 1987. **6**(10): p. 3065-70.

127. Tsai, R.Y. and R.R. Reed, *Cloning and functional characterization of Roaz, a zinc finger protein that interacts with O/E-1 to regulate gene expression: implications for olfactory neuronal development.* J Neurosci, 1997. **17**(11): p. 4159-69.

128. Holt, N., J. Wang, K. Kim, G. Friedman, X. Wang, V. Taupin, G.M. Crooks, D.B. Kohn, P.D. Gregory, M.C. Holmes, and P.M. Cannon, *Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo.* Nature biotechnology, 2010. **28**(8): p. 839-47.

129. Guo, X., J.W. Carroll, M.R. Macdonald, S.P. Goff, and G. Gao, *The zinc finger antiviral protein directly binds to specific viral mRNAs through the CCCH zinc finger motifs.* Journal of virology, 2004. **78**(23): p. 12781-7.

130. Wolf, D. and S.P. Goff, *Embryonic stem cells use ZFP809 to silence retroviral DNAs.* Nature, 2009. **458**(7242): p. 1201-4.

131. Su, A.I., M.P. Cooke, K.A. Ching, Y. Hakak, J.R. Walker, T. Wiltshire, A.P. Orth, R.G. Vega, L.M. Sapinoso, A. Moqrich, A. Patapoutian, G.M. Hampton, P.G. Schultz, and J.B. Hogenesch, *Large-scale analysis of the human and mouse transcriptomes.* Proc Natl Acad Sci U S A, 2002. **99**(7): p. 4465-70.

132. Bogaerts, S., A. Vanlandschoot, J. van Hengel, and F. van Roy, *Nuclear translocation of alphaN-catenin by the novel zinc finger transcriptional repressor ZASC1.* Exp Cell Res, 2005. **311**(1): p. 1-13.

133. Imoto, I., Y. Yuki, I. Sonoda, T. Ito, Y. Shimada, M. Imamura, and J. Inazawa, *Identification of ZASC1 encoding a Kruppel-like zinc finger protein as a novel target for 3q26 amplification in esophageal squamous cell carcinomas.* Cancer Res, 2003. **63**(18): p. 5691-6.

134. Chiang, W.F., P.S. Hung, S.Y. Liu, T.C. Yuan, K.W. Chang, Y.P. Chen, Y.C. Liu, and S.C. Lin, *Increase of ZASC1 gene copy number in recurrent oral carcinoma.* Oral Dis, 2011. **17**(1): p. 53-9.

135. Jung, S.Y., A. Malovannaya, J. Wei, B.W. O'Malley, and J. Qin, *Proteomic analysis of steady-state nuclear hormone receptor coactivator complexes.* Mol Endocrinol, 2005. **19**(10): p. 2451-65.

136. Vo, N. and R.H. Goodman, *CREB-binding protein and p300 in transcriptional regulation.* The Journal of biological chemistry, 2001. **276**(17): p. 13505-8.

137. Bruce, J.W., P. Ahlquist, and J.A. Young, *The host cell sulfonation pathway contributes to retroviral infection at a step coincident with provirus establishment.* PLoS Pathog, 2008. **4**(11): p. e1000207.

138. Ishida, Y. and P. Leder, *RET: a poly A-trap retrovirus vector for reversible disruption and expression monitoring of genes in living cells.* Nucleic Acids Res, 1999. **27**(24): p. e35.

139. Neuhaus, D., Y. Nakaseko, J.W. Schwabe, and A. Klug, *Solution structures of two zinc-finger domains from SWI5 obtained using two-dimensional 1H nuclear magnetic resonance spectroscopy. A zinc-finger structure with a third strand of beta-sheet.* Journal of molecular biology, 1992. **228**(2): p. 637-51.

140. Isalan, M., Y. Choo, and A. Klug, *Synergy between adjacent zinc fingers in sequence-specific DNA recognition.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(11): p. 5617-21.

141. Goff, S.P., *Host factors exploited by retroviruses.* Nature reviews. Microbiology, 2007. **5**(4): p. 253-63.

142. Giri, M.S., M. Nebozhyn, L. Showe, and L.J. Montaner, *Microarray data on gene modulation by HIV-1 in immune cells: 2000-2006.* Journal of leukocyte biology, 2006. **80**(5): p. 1031-43.

143. Yeung, M.L., L. Houzet, V.S. Yedavalli, and K.T. Jeang, *A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication.* J Biol Chem, 2009. **284**(29): p. 19463-73.

144. Krishnan, L., K.A. Matreyek, I. Oztop, K. Lee, C.H. Tipper, X. Li, M.J. Dar, V.N. Kewalramani, and A. Engelman, *The requirement for cellular transportin 3 (TNPO3 or TRN-SR2) during infection maps to human immunodeficiency virus type 1 capsid and not integrase.* Journal of virology, 2010. **84**(1): p. 397-406.

145. Levin, A., Z. Hayouka, A. Friedler, and A. Loyter, *Transportin 3 and importin alpha are required for effective nuclear import of HIV-1 integrase in virus-infected cells.* Nucleus, 2010. **1**(5): p. 422-31.

146. Ocwieja, K.E., T.L. Brady, K. Ronen, A. Huegel, S.L. Roth, T. Schaller, L.C. James, G.J. Towers, J.A. Young, S.K. Chanda, R. Konig, N. Malani, C.C. Berry, and F.D. Bushman, *HIV integration targeting: a pathway involving Transportin-3 and the nuclear pore protein RanBP2.* PLoS pathogens, 2011. **7**(3): p. e1001313.

147. Valle-Casuso, J.C., F. Di Nunzio, Y. Yang, N. Reszka, M. Lienlaf, N. Arhel, P. Perez, A.L. Brass, and F. Diaz-Griffero, *TNPO3 Is Required for HIV-1 Replication after Nuclear Import but prior to Integration and Binds the HIV-1 Core.* Journal of virology, 2012. **86**(10): p. 5931-6.

148. Okimoto, M.A. and H. Fan, *Identification of directly infected cells in the bone marrow of neonatal moloney murine leukemia virus-infected mice by use of a moloney murine leukemia virus-based vector.* Journal of virology, 1999. **73**(2): p. 1617-23.

149. Chao, S.H., J.R. Walker, S.K. Chanda, N.S. Gray, and J.S. Caldwell, *Identification of homeodomain proteins, PBX1 and PREP1, involved in the transcription of murine leukemia virus.* Mol Cell Biol, 2003. **23**(3): p. 831-41.

150. Welinder, E., R. Mansson, E.M. Mercer, D. Bryder, M. Sigvardsson, and C. Murre, *The transcription factors E2A and HEB act in concert to induce the expression of FOXO1 in the common lymphoid progenitor.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(42): p. 17402-7.

151. Akashi, K., D. Traver, T. Miyamoto, and I.L. Weissman, *A clonogenic common myeloid progenitor that gives rise to all myeloid lineages.* Nature, 2000. **404**(6774): p. 193-7.

152. Wada, T., T. Nakashima, N. Hiroshi, and J.M. Penninger, *RANKL-RANK signaling in osteoclastogenesis and bone disease.* Trends in molecular medicine, 2006. **12**(1): p. 17-25.

153. Chan, W., N. Costantino, R. Li, S.C. Lee, Q. Su, D. Melvin, D.L. Court, and P. Liu, *A recombineering based approach for high-throughput conditional knockout targeting vector construction.* Nucleic Acids Res, 2007. **35**(8): p. e64.

154. Lander, J.K., B. Chesebro, and H. Fan, *Appearance of mink cell focus-inducing recombinants during in vivo infection by moloney murine leukemia virus (M-MuLV) or the Mo+PyF101 M-MuLV enhancer variant: implications for sites of generation and roles in leukemogenesis.* Journal of virology, 1999. **73**(7): p. 5671-80.

155. Bruce, J.W., K.A. Bradley, P. Ahlquist, and J.A. Young, *Isolation of cell lines that show novel, murine leukemia virus-specific blocks to early steps of retroviral replication.* J Virol, 2005. **79**(20): p. 12969-78.

156. Mayya, V., D.H. Lundgren, S.I. Hwang, K. Rezaul, L. Wu, J.K. Eng, V. Rodionov, and D.K. Han, *Quantitative phosphoproteomic analysis of T cell receptor signaling reveals system-wide modulation of protein-protein interactions.* Science signaling, 2009. **2**(84): p. ra46.

157. Coffin JM, H.S., Varmus HE *Retroviruses*1997, NY: Cold Spring Harbor Laboratory Press.

158. Dakic, A., Q.X. Shao, A. D'Amico, M. O'Keeffe, W.F. Chen, K. Shortman, and L. Wu, *Development of the dendritic cell system during mouse ontogeny.* Journal of immunology, 2004. **172**(2): p. 1018-27.