1-1-2013

Interactions Between APOBEC3 and Murine Retroviruses: Mechanisms of Restriction and Drug Resistance

Alyssa Lea MacMillan
University of Pennsylvania, alyssahuegel@gmail.com

Interactions Between APOBEC3 and Murine Retroviruses: Mechanisms of Restriction and Drug Resistance

Abstract
APOBEC3 proteins are important for antiretroviral defense in mammals. The activity of these factors has been well characterized in vitro, identifying cytidine deamination as an active source of viral restriction leading to hypermutation of viral DNA synthesized during reverse transcription. These mutations can result in viral lethality via disruption of critical genes, but in some cases is insufficient to completely obstruct viral replication. This sublethal level of mutagenesis could aid in viral evolution. A cytidine deaminase-independent mechanism of restriction has also been identified, as catalytically inactive proteins are still able to inhibit infection in vitro. Murine retroviruses do not exhibit characteristics of hypermutation by mouse APOBEC3 in vivo. However, human APOBEC3G protein expressed in transgenic mice maintains antiviral restriction and actively deaminates viral genomes. The mechanism by which endogenous APOBEC3 proteins function is unclear.

The mouse provides a system amenable to studying the interaction of APOBEC3 and retroviral targets in vivo. Virions packaging endogenous protein were isolated from mice for analysis of APOBEC3 without a need for protein overexpression. Biochemical and molecular studies are possible using endogenous protein and viral nucleic acids. Additionally, the effect of APOBEC3-mediated viral mutagenesis and subsequent drug resistance can be modeled in this system. Human APOBEC3G transgenic mice infected with murine retroviruses and treated with an antiretroviral drug allows examination of natural levels of viral replication, APOBEC3 induced hypermutation, and potential viral escape.

Studies described herein explore mechanisms of APOBEC3-mediated restriction and drug resistance in vivo. We show that endogenous APOBEC3 protein is efficiently packaged into viral cores, and this protein maintains catalytic activity against artificial substrates. We recovered low levels of G-to-A mutations from natural reverse transcription products, although approximately five to ten fold lower than that thought to be necessary for efficient viral restriction. We show that inhibition of reverse transcription is the main mechanism of restriction in vivo, and can be targeted through virion-packaged or cell-associated protein.

Transgenically-expressed human APOBEC3G is instead able to heavily deaminate viral DNA, although frequently to sublethal levels. We assessed the effect of both murine APOBEC3 and APOBEC3G on viral replication in the presence and absence of an antiretroviral drug, and examined viruses for drug resistance mutations. APOBEC3G has a clear effect on the rate of viral mutagenesis in vivo, with the potential to induce drug resistance mutations.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Cell & Molecular Biology

First Advisor
Susan R. Ross

Keywords
cytidine deaminase, HIV, intrinsic immunity, restriction factor

Subject Categories
Virology

INTERACTIONS BETWEEN APOBEC3 AND MURINE RETROVIRUSES: MECHANISMS OF RESTRICTION AND DRUG RESISTANCE

Alyssa L. MacMillan

A DISSERTATION

in

Cell and Molecular Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2013

Supervisor of Dissertation

_____

Susan R. Ross, Ph.D.

Professor of Microbiology

Graduate Group Chairperson

_____

Daniel S. Kessler, Ph.D., Associate Professor of Cell and Developmental Biology

Dissertation Committee

Paul Bates, Ph.D., Professor of Microbiology

Rahul Kohli, M.D., Ph.D., Assistant Professor of Medicine

Sara Cherry, Ph.D., Associate Professor of Microbiology

James Hoxie, M.D., Professor of Medicine

ACKNOWLEDGMENT

A PhD degree is awarded to a single person, but the effort and support come from many. I owe significant thanks to my advisor, Susan Ross, for supporting me through many years in her laboratory. Her patience, expertise, and constructive criticism are much appreciated. Former lab member Chioma Okeoma was critical in the early stages of this project, and her kindness and intelligence continue to be extraordinarily valuable to me. I also owe much thanks to current postdoctoral lab members Christian Cuevas and Spiros Stavrou for answering my endless questions and for helpful discussions. Transgenic animal models used in this work would also not have been available without the incredibly hard work of Spiros. Kristin Blouch offered not only animal expertise and assistance, but also constant support and friendship, for which I am most grateful.

Of course non-lab support was also critical for my overall sanity and well-being. My parents, Marty and Carol Huegel, have been extremely supportive of my extended education, and I would not have made it this far without them. I owe the most to my two strongest pillars of support, my twin sister Julianne Huegel and my husband Matt MacMillan. I cannot thank them enough for their never ending devotion and care.

ABSTRACT

INTERACTIONS BETWEEN APOBEC3 AND MURINE RETROVIRUSES: MECHANISMS OF RESTRICTION AND DRUG RESISTANCE

Alyssa L. MacMillan

Susan R. Ross, Ph.D.

APOBEC3 proteins are important for antiretroviral defense in mammals. The activity of these factors has been well characterized *in vitro*, identifying cytidine deamination as an active source of viral restriction leading to hypermutation of viral DNA synthesized during reverse transcription. These mutations can result in viral lethality via disruption of critical genes, but in some cases is insufficient to completely obstruct viral replication. This sublethal level of mutagenesis could aid in viral evolution. A cytidine deaminase-independent mechanism of restriction has also been identified, as catalytically inactive proteins are still able to inhibit infection *in vitro*. Murine retroviruses do not exhibit characteristics of hypermutation by mouse APOBEC3 *in vivo*. However, human APOBEC3G protein expressed in transgenic mice maintains antiviral restriction and actively deaminates viral genomes. The mechanism by which endogenous APOBEC3 proteins function is unclear.

The mouse provides a system amenable to studying the interaction of APOBEC3 and retroviral targets *in vivo*. Virions packaging endogenous protein were isolated from mice for analysis of APOBEC3 without a need for protein overexpression. Biochemical and molecular studies are possible using endogenous protein and viral nucleic acids. Additionally, the effect of APOBEC3-mediated viral mutagenesis and subsequent drug resistance can be modeled in this system. Human APOBEC3G transgenic mice infected with murine retroviruses and treated with an antiretroviral drug allows examination of natural levels of viral replication, APOBEC3 induced hypermutation, and potential viral escape.

Studies described herein explore mechanisms of APOBEC3-mediated restriction and drug resistance *in vivo*. We show that endogenous APOBEC3 protein is efficiently packaged into

viral cores, and this protein maintains catalytic activity against artificial substrates. We recovered low levels of G-to-A mutations from natural reverse transcription products, although approximately five to ten fold lower than that thought to be necessary for efficient viral restriction. We show that inhibition of reverse transcription is the main mechanism of restriction *in vivo*, and can be targeted through virion-packaged or cell-associated protein.

Transgenically-expressed human APOBEC3G is instead able to heavily deaminate viral DNA, although frequently to sublethal levels. We assessed the effect of both murine APOBEC3 and APOBEC3G on viral replication in the presence and absence of an antiretroviral drug, and examined viruses for drug resistance mutations. APOBEC3G has a clear effect on the rate of viral mutagenesis *in vivo*, with the potential to induce drug resistance mutations.

TABLE OF CONTENTS

Abstract…….. iii

List of Tables…… ix

List of Figures…… x

Chapter 1: Introduction……1

Chapter 1.1: Retroviruses……..1

Chapter 1.2: The RNA Genome and Virus Structure…… 2

1.2.1 The viral gene *gag*……...2

1.2.2 The viral genes *pro* and *pol*…….. 4

1.2.3 The viral gene *env*……..5

1.2.4 The long terminal repeats……..6

1.2.5 Virus Structure……7

Chapter 1.3: Viral Replication Cycle…… 8

1.3.1. Viral Entry……8

1.3.2. Reverse Transcription…….. 10

1.3.3. Integration…….. 13

1.3.4. Transcription and Translation…… 14

1.3.5. Assembly and Budding……15

Chapter 1.4: Mouse Mammary Tumor Virus…….. 16

Chapter 1.5: Murine Leukemia Virus…….. 19

Chapter 1.6: Retroviral Restriction Factors…… 20

Chapter 1.6.1: Overview of Retroviral Restriction Factors…… 21

Chapter 1.6.2: HIV-1 and the discovery of APOBEC3…….. 21

Chapter 1.6.3: APOBEC3 Evolution……..26

Chapter 1.6.4: APOBEC3 Structure, Function, and Regulation……26

Chapter 1.6.5: Murine APOBEC3 and Murine Retroviruses…… 29

Chapter 1.6.6: APOBEC3 and Drug Resistance……34

Chapter 1.7: Specific Aims and Significance of Dissertation…….. 36

Chapter 2: APOBEC3 inhibition of MMTV infection: the role of cytidine deamination versus inhibition of reverse transcription……..38

Chapter 2.1: Abstract……. 38

Chapter 2.2: Introduction……... 38

Chapter 2.3: Materials and Methods…… 41

Chapter 2.3.1: Plasmids, Cells, and Viruses…… 41

Chapter 2.3.2: Mice……..42

Chapter 2.3.3: *In vitro* deamination……43

Chapter 2.3.4: *In vitro* reverse transcription…… 46

Chapter 2.3.5: Polymerase Chain Reaction…… 46

Chapter 2.3.6: Sequencing and Statistical Analysis……... 47

Chapter 2.4: Results……...48

Chapter 2.4.1: Both murine APOBEC3 variants are packaged into virion cores…….. 48

Chapter 2.4.2: Allelic differences in substrate specificity…….. 50

Chapter 2.4.3: APOBEC3 deamination of natural MMTV transcripts is low…….. 52

Chapter 2.4.4: APOBEC3 inhibits MMTV reverse transcription…… 57

Chapter 2.4.5: Restriction of reverse transcription is correlated to level of packaged APOBEC3…… 59

Chapter 2.4.6: Target cell expression of either APOBEC3 allele leads to virus restriction…… 61

Chapter 2.5: Discussion……63

Chapter 3: The effect of APOBEC3 on antiretroviral drug resistance *in vivo*…….. 67

Chapter 3.1: Abstract…… 67

Chapter 3.2: Introduction……... 68

Chapter 3.3: Materials and Methods……72

Chapter 3.3.1: Cells and Viruses……72

Chapter 3.3.2: *In vitro* infection and drug optimization…….. 73

Chapter 3.3.3: Mice and Sample Collection……74

Chapter 3.3.4: *In vivo* infection and drug optimization……75

Chapter 3.3.5: Polymerase Chain Reaction……76

Chapter 3.3.6: Titering of Plasma Viremia……77

Chapter 3.3.7: Sequencing…….. 77

Chapter 3.4: Results…….. 78

Chapter 3.4.1: Potential long term effects of mAPOBEC3 on MMTV…….. 78

Chapter 3.4.2: MMTV is restricted by therapeutic antiretrovirals *in vitro*…… 82

Chapter 3.4.3: Selection of antiretroviral drug and dosing for *in vivo* studies of viral evolution…….. 85

Chapter 3.4.4: The effects of antiretroviral treatment and APOBEC3 on MMTV replication……86

Chapter 3.4.5: Barriers to secondary analysis of exogenous MMTV variants……..90

Chapter 3.4.6: MLV is restricted by therapeutic antiretroviral *in vitro* and *in vivo*…… 92

Chapter 3.4.7: MLV is restricted and hypermutated in human APOBEC3G-expressing mice……97

Chapter 3.4.8: Antiretroviral treatment is more effective in APOBEC3G-expressing mice……98

Chapter 3.4.9: Human APOBEC3G affects MLV diversity *in vivo*……101

Chapter 3.5: Discussion……109

Chapter 4: Summary…… 114

Chapter 4.1: Discussion…… 114

Chapter 4.1.1: Antiviral mechanisms of endogenous murine APOBEC3 *in vivo*…….. 114

Chapter 4.1.2: APOBEC3-induced viral evolution of murine retroviruses *in vivo*…… 116

Chapter 4.2: Future Directions…… 117

REFERENCES…….. 121

LIST OF TABLES

CHAPTER ONE

Table 1.1 Retroviral Classifications…….. 18

Table 1.2 Murine APOBEC3 Alleles…… 40

CHAPTER TWO

Table 2.1 Oligomer substrate sequences for *in vitro* deamination assay…… 55

Table 2.2 Mutation analysis of MMTV reverse transcription products…… 66

CHAPTER THREE

Table 3.1 Retroviral Genome Composition…….. 81

Table 3.2 Human APOBEC3G deaminates M-MLV *in vivo*……... 108

Table 3.3 APOBEC3G affects MLV genomic evolution *in vivo*…… 113

Table 3.4 Viral clones assessed for amino acid mutations…… 114

LIST OF FIGURES

CHAPTER ONE

Figure 1.1. Basic structure of a retroviral genome…… 13

Figure 1.2. Basic structure of a retroviral virion……. 13

Figure 1.3 Retroviral Lifecycle.…… 19

Figure 1.4 Process of reverse transcription…… 22

Figure 1.5 HIVΔVif Permissive and Non-Permissive Cells…….. 35

Figure 1.6 Murine APOBEC3 Alleles…….. 41

CHAPTER TWO

Figure 2.1 APOBEC3 is packaged within MMTV viral cores *in vivo*……. 59

Figure 2.2 *In vivo* packaged APOBEC3 variants are enzymatically active and target different substrates……61

Figure 2.3 Packaged APOBEC3 edits MMTV reverse transcription products…… 64

Figure 2.4 Packaged APOBEC3 restricts MMTV reverse transcription……68

Figure 2.5 Restriction of reverse transcription correlates with the amount of packaged APOBEC3…….. 70

Figure 2.6 Target cell APOBEC3 can restrict reverse transcription of incoming virions and is independent of cytidine deamination……..72

CHAPTER THREE

Figure 3.1 MMTV passaged in APOBEC3-/- mice differs from virus passed in APOBEC3+/+ mice……. 91

Figure 3.2 Antiretrovirals effectively block MMTV *in vitro*…….. 94

Figure 3.3 The dual effect of APOBEC3 and AZT on MMTV replication *in vivo*…….. 99

Figure 3.4 M-MLV is susceptible to antiretroviral drugs *in vitro*…… 104

Figure 3.5 M-MLV is susceptible to AZT *in vivo*…….. 106

Figure 3.6 The antiviral effects of APOBEC3G and AZT are distinct early in infection, but become less apparent over time…… 109

Figure 3.7 APOBEC3G affects level of viral amino acid diversity *in vivo* …….. 116

Figure 3.8 MLV reverse transcriptase amino acid mutations in virions recovered from APOBEC3-/- and APOBEC3G-expressing mice……..118

Chapter 1- Introduction

1.1: Retroviruses

Retroviruses are a family of enveloped RNA viruses that are distinctive due to the replication of their RNA genome into DNA, which is then incorporated into the host cell genome. This integration event has occurred thousands of times over the course of mammalian evolution, leading to a significant portion of mammalian genomes being comprised of endogenous viral remnants. Most of these proviruses no longer encode for an active, infectious virus, but exogenous retroviruses still pose an important threat as human pathogens.

Retroviruses were first discovered in the early twentieth century, as investigators worked to find causes for specific neoplasias in animals that could be transmitted from one animal to another. Ellermann and Bang first showed that a form of chicken leukemia was transmissible, the cause of which was eventually found to be avian leukosis virus (ALV) (Ellerman and Bang, 1908). Soon after, Peyton Rous determined that sarcomas could be transferred between chickens by injection of cell filtrates; this virus was termed Rous sarcoma virus (Rous, 1910). Virus-induced tumors were then discovered in a number of other species, including mice, cats and cattle. The first human retroviral pathogen was human T-cell leukemia virus (HTLV-1), found in leukemic human T-cells after epidemiological studies suggested an infectious cause for a cluster of aggressive leukemias in small islands off of Japan (Poiesz, 1980). The most devastating human retrovirus was identified soon after, as epidemics of Acquired Immune Deficiency Syndrome arose in several countries around the world (Barre-Sinoussi, 1983). Now known as human immunodeficiency virus, HIV has been the most difficult challenge to the field of retrovirology since its discovery.

Retroviruses have a condensed core of electron-dense material, and are 80-120 nm in diameter (Coffin, 2002). An outer lipid envelope studded with viral glycoproteins surrounds these viruses. The viral genome is a single-stranded, positive sense RNA of between 7-12 kb in size, with each viral particle containing two copies (Coffin, 2002). This genome is transcribed into a double-stranded DNA intermediate through a process known as reverse transcription, and this

DNA is integrated into the host cell genome. The virus recruits host cell machinery to transcribe and translate viral proteins and full genomic RNAs for new viral particles (Coffin, 1996). Both introduction of highly expressed, virus-associated oncogenic genes or mutation of oncogenic host genes through integration can lead to transduction of the host cell, eventually leading to pathogenic neoplasias (Robinson, 1982).

1.2: The RNA Genome and Virus Structure

All retroviral genomes are organized around common core genes (see Figure 1.1). The four major coding domains *gag, pro, pol,* and *env* produce proteins necessary for the basic retroviral life cycle.

1.2.1 The viral gene *gag*

The *gag* gene contains the genetic information to produce internal structural proteins, including matrix, capsid, and nucleocapsid. These proteins are all proteolytically cleaved from a large polyprotein called Gag (Vogt and Eisenman, 1973). The amino-terminal portion of *gag* encodes for matrix (MA, membrane-associated). Matrix drives the interaction between the virus and the inner leaflet of cellular membranes during virion formation (Figure 1.2) (Hamard-Peron and Muriaux, 2011). Matrix proteins have a myristate added co-translationally. This 14-carbon fatty acid promotes association with membranes, and without it, viral assembly is blocked (Grand, 1989). Moving downstream of matrix on the *gag* gene, the next encoded protein is capsid. It is typically a very large antigenic protein, and through multimerization, forms a shell within the virion interior to encapsulate the genomic RNA and associated proteins. This capsid structure and its contents are considered the viral core. The true function of this capsid-created shell is unknown, but it likely serves to shield the genomic RNA until the virion is ready to undergo reverse transcription after infection of a new cell (Roos, et al., 2007). Capsid also serves a number of other functions, including recruitment of critical protein components into the assembling virion (Smith, et al., 1993).

Fig. 1.1. Basic structure of a retroviral genome. The retroviral genome encodes four major coding domains shown in gray, confined between regulatory 5' and 3' LTR segments. Both the 5' and the 3' end contain an identical sequence, R, shown in red. R is necessary for correct strand transfer during reverse transcription (see Chapter 1.3.2). U5 (orange) and U3 (green) are short unique sequences, which contain regulatory elements. The primer binding site (PBS; yellow) is an 18 bp segment complementary to the tRNA primer used by the virus to begin reverse transcription. Figure credit: Matt MacMillan

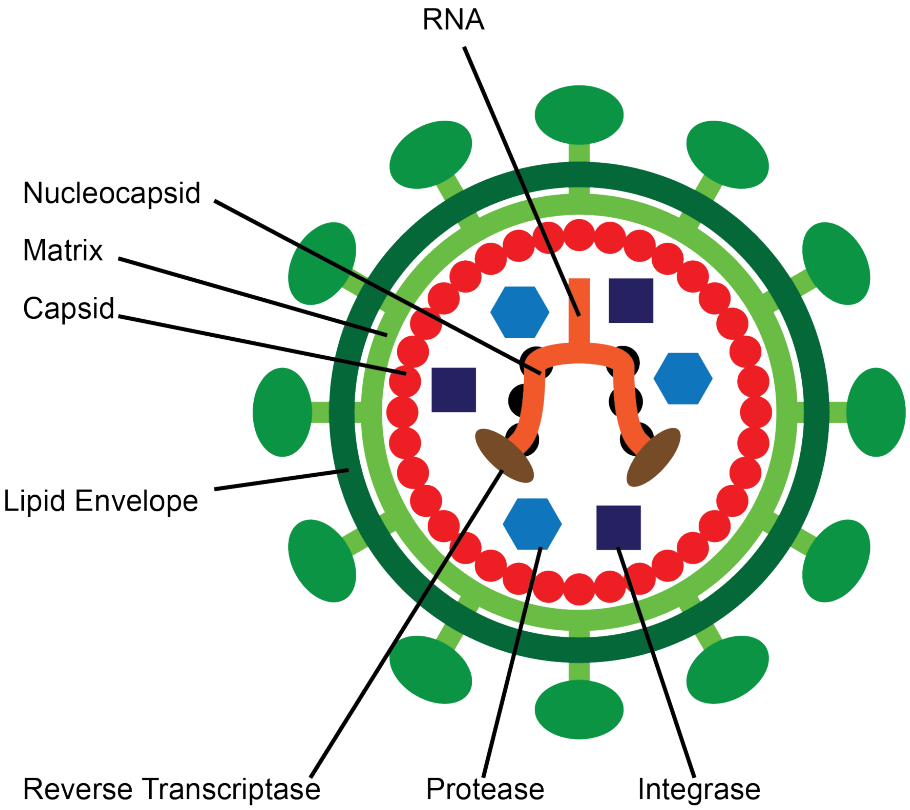


Fig. 1.2. Basic structure of a retroviral virion. The main virion components are the lipid envelope (dark green), RNA genome (orange), structural proteins (capsid, red; matrix, light green; and envelope glycoproteins, green), and nonstructural proteins (protease, light blue, integrase, dark blue; reverse transcriptase, brown). Figure credit: Matt MacMillan

The final critical component created from *gag* is nucleocapsid, encoded at the 3'-terminus of the gene. This small protein is found within the viral core created by the capsid shell, and is associated with the viral RNA (Davis and Rueckert, 1972; Meric, 1984). Nucleocapsid has a number of roles important for the viral life cycle, including assisting in RNA dimerization, promoting nucleic acid annealing, assisting with the early steps of reverse transcription, and stabilizing the newly formed viral DNA before integration (Freed, 1998). However, perhaps the most critical role is recruiting viral RNA into the virion through direct binding, using zinc-finger motifs (South and Summers, 1993). In some retroviruses, *gag* also encodes for one or two additional small proteins. These sequences can be located in between those encoding for the three main Gag proteins, and vary between viruses. Roles for these proteins are not well understood, although they are typically necessary for a complete viral lifecycle. Investigations continue to examine these additional factors, with recent work showing the murine leukemia virus protein p12 is critical for both core stability and tethering of the viral DNA to host chromosomes during integration (Wight, 2012; Elis, 2012).

1.2.2 The viral genes *pro* and *pol*

The genes *pro* and *pol* are adjacent to one another in the retroviral genome. Some viruses transcribe these genes as a single open reading frame, with cleavage of the polyprotein late in assembly (HIV), while others transcribe them from separate open reading frames (MMTV, MPMV). Still others have overlapping open reading frames with *gag* (HTLV) or an even larger polyprotein formed with Gag (MLV) (Coffin, 2002). Regardless of the organization, all retroviruses encode three important enzymes within *pro* and *pol*: protease, reverse transcriptase (RT), and integrase. Regulation of transcription and translation leads to significantly less enzyme production than Gag protein production (Coffin, 2002). Protease is required to cleave the large polyproteins into mature individual proteins. This activity occurs late in the viral lifecycle, immediately after budding (Dougherty and Semler, 1993). Viruses that are deficient in protease

will bud, but will not form infectious mature particles, thereby making protease a popular drug target (Lee, 2012). Protease functions through two apposed aspartate residues, each provided by two protease monomers, which together form a homodimer (Navia, 1989). This homodimer targets specific stretches of amino acids that are thought to be linkers between different domains of the larger polyprotein (Coffin, 2002).

Reverse transcriptase converts the viral ssRNA genome into dsDNA. The paradigm existing prior to studies of retroviral RT considered DNA the template for RNA, with the reverse reaction impossible. This was known as the "central dogma," coined by Francis Crick (Crick, 1970). The discovery of RT redefined this paradigm, and also helped shape modern molecular biology (Baltimore, 1970; Temin and Mizutani, 1970). RT exists as either a monomer or a heterodimer in different retroviral families, but all have a similar three-dimensional structure necessary to bind specifically to its nucleic acid substrate, and are homologous in the active site region (Goff, 1990). RT is a complex enzyme, working not only as a DNA polymerase using both RNA and DNA substrates, but also as a ribonuclease (Grandgenett, 1978). This activity is termed RNase H, based on its specificity for RNA in a RNA:DNA hybrid, and is necessary to complete reverse transcription (Moelling, 1971).

The final protein product of *pro* and *pol* is integrase, which is responsible for integrating viral DNA into the host genome. Integrase recognizes the ends of linear dsDNA after reverse transcriptase is complete, and characteristically cleaves two or three nucleotides from the 3' end of each strand. This exposes a CA dinucleotide end, which is covalently linked to the host genome (Coffin, 2002). Some retroviruses integrate randomly throughout the genome, while others preferentially integrate into actively transcribed genes (Mitchell, 2004). Integrase is also often targeted by antiretroviral drugs.

1.2.3 The viral gene *env*

In order to achieve entry, retroviruses must display glycoproteins on the virion surface, which pass through the lipid membrane and are exposed to target cell viral receptors. The main

function of these proteins is to enable entry into a host cell. These proteins are encoded by *env*, located at the 3' end of the genome. Unlike *gag*, *pro*, and *pol*, this gene is spliced from subgenomic RNA. It is also unique to retroviral proteins, as it must be processed for transport to the cell surface. These proteins are heavily glycosylated, primarily though *N*-linked glycosylation but also by *0*-linked glycosylation (Pinter and Honnen, 1988). The purpose for this glycosylation includes increased protein stability, more efficient protein folding, and protection from host enzymes and immune responses (Elder, 1986). Further processing is required as typically a single Env polyprotein is produced which is then proteolytically cleaved into two distinct proteins, termed SU (surface) and TM (transmembrane) (Coffin, 2002). This typically requires a cellular protease such as furin (Einfeld, 1996). SU contains the receptor binding domain, and helps instigate conformational changes in TM which leads to the active fusion of virus and cell. After cleavage, the two subunits continue to interact via disulfide or noncovalent bonds, depending on the virus. Envelope heterodimers form multimers, consisting of three SU-TM proteins, which provides a more rigid structure that appears as a spike by electron microscope (Coffin, 2002). Each of the SU units is able to independently bind to a host cell receptor, and multiple interactions may be necessary to facilitate entry (Layne, 1990). Structurally, the cytoplasmic tail of TM interacts with internal Gag proteins. This interaction mediates incorporation of the viral envelope proteins located on the cell surface with all the necessary internal viral components.

1.2.4. The long terminal repeats

The RNA genome of retroviruses is flanked by sequences required for replication, integration, and transcriptional regulation post-integration. The 5' end begins with an R sequence followed by U5 (Unique 5'), while the 3' end consists of a U3 (Unique 3') sequence and ends with an R sequence identical to that found at the amino-terminus. During reverse transcription, the RNA genome is copied in such a way that both ends of the DNA are flanked by a U3-R-U5 sequence repeat, termed the long terminal repeat. These regions range in size from 200 bp to over 1200 bp (Coffin, 2002). The 5' LTR serves as an RNA pol II promoter, and contains a

variety of regulatory sequences such as enhancers responsive to both viral and cellular activating factors (Krebs, 2001). Although the 3' LTR has an identical sequence to that of the 5' LTR, its location at the 3' end of the genome provides it a different role. Instead of acting as a promoter, the 3' LTR signals for transcriptional termination and polyadenylation. In certain retroviruses, the LTR also contains an open reading frame, which is transcribed via alternative splicing from the 5' to the 3' LTR (Korman, 1992).

1.2.5 Virus Structure

Retroviral particles consist of RNA, lipids, and proteins. Approximately 1-2% of the total viral content is RNA, two-thirds of which are viral genomic RNA, and one-third of which is transfer RNA (used to prime reverse transcription) and small host RNAs (Coffin, 2002). The lipid content accounts for 35% of the virus, and comes from host cell lipid membranes. However, retroviral lipid envelopes have a larger concentration of cholesterol and sphingomyelin. These components aid in temperature stability and viral assembly (Slosberg and Montelaro, 1982). All of the virally encoded proteins make up the bulk of the virion, with approximately 75% Gag-derivatives, 15-20% envelope proteins, and 5-10% non-structural enzyme and accessory proteins. Together, these produce a virion of approximately 80-120 nm in diameter. However, virion morphology varies greatly, and is used to characterize and classify viruses (Table 1.1).

Mature virions have a condensed core, which can be round, slightly angular, cone shaped or bar shaped (Coffin, 2002). The viral core contains two copies of the RNA genome, which are non-covalently joined together through RNA dimerization (Paillart, 2004). This RNA is coated with nucleocapsid proteins. Protease, reverse transcriptase, and integrase are also found within the viral core, which is composed of tightly packed capsid proteins surrounding this RNA-protein complex. A spherical layer of matrix proteins surrounds the viral core, and serves as the interior layer of the viral envelope, beneath the lipid bilayer envelope. Envelope glycoproteins stud the viral surface (Coffin, 2002).

Table 1.1 Retroviral Classification

Viral Group	Example	Family Name	Morphology	Genome
Avian sarcoma and leukosis	Rous sarcoma virus	Alpha-retrovirus	Central, spherical core (C particles)	Simple
B-type mammalian	Mouse mammary tumor virus	Beta-retrovirus	Irregular, spherical core (B particles)	Simple or complex
Murine leukemia viruses	Moloney MLV	Gamma-retrovirus	Central, spherical core	Simple
HTLV & bovine leukemia	Human T-cell leukemia virus	Delta-retrovirus	Central, spherical core	Complex
D-type virus	Mason-Pfizer monkey virus	Betaretrovirus	Cylindrical core "D particles"	Simple
Lentivirus	Human immunodeficiency virus	Lentivirus	Cone shaped core	Complex
Spumavirus	Foamy virus	Spumavirus	Central, spherical core	Complex

Chapter 1.3 Viral Replication Cycle

1.3.1 Viral Entry

The first step of the viral infection cycle is binding of the virus to a target cell and entry into the cytoplasm (Figure 1.3). Initial absorption onto the cell surface can be enhanced by additional host proteins or molecules, such as heparin sulfate proteoglycans, that are distinct from the viral receptor (Sharma, 2000). The surface envelope glycoproteins bind to specific host cell membrane components (the viral receptor) through high affinity binding via the receptor binding site. Each virus uses a very specific receptor, which determines the tropism, or range of cell types able to be infected with the virus. Binding of the receptor initiates a conformational change in the envelope protein subunits, which brings the viral and cellular membranes into very close proximity and mediates fusion of the two. The TM glycoprotein of each retrovirus contains the fusogenic peptide required for fusion. This peptide is positioned next to a domain similar to a leucine zipper that forms a coiled-coil structure. Because the envelope is oligomerized, the multiple coiled-coil domains form a stable structure that can insert the fusogenic peptide into the host cell membrane (Bentz and Mittal, 2000). Although the exact process is not clearly

understood, a fusion pore is created, the lipid membranes mix, and the viral core is internalized into the cell.

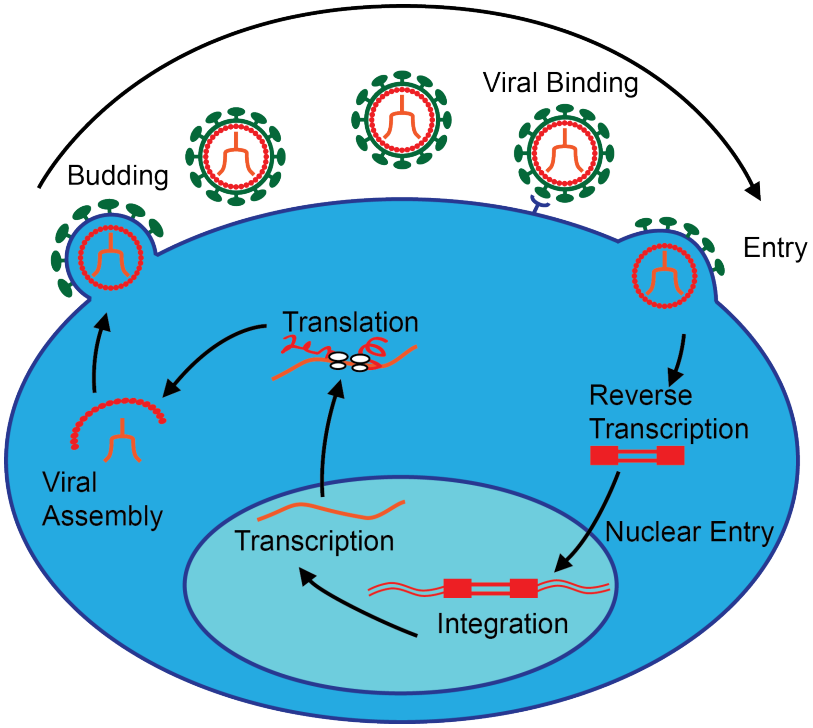


Fig. 1.3 Retroviral Lifecycle. MMTV completes a viral lifecycle similar to other retroviruses. Virion binds to a target cell via cellular receptor, and enters the cell via membrane fusion. The RNA genome is converted to dsDNA via reverse transcriptase, and targeted to cellular chromosomes for integration into the host DNA. Integrated viral proviruses are templates for new viral genomes, as well as for transcription of viral genes to produce new viral proteins. Viral assembly occurs at different sites of the cell, depending on the specific virus, and eventually produces new virions budding out of cellular membranes. Figure credit: Matt MacMillan

Viral internalization occurs quickly after initial fusion, typically within five minutes. Most retroviruses are thought to enter at the plasma membrane, while others appear to enter from an intracellular compartment after endocytosis, with fusion occurring between the virus and the endocytic vesicle (Nisole and Saib, 2004). Other retroviruses, including MMTV and amphotropic MLV require a drop in pH to catalyze fusion and thus enter cells from an acidic compartment (McClure, 1993; Ross, 2002). Some retroviruses, including HIV-1 and HTLV-1, may also enter a new cell through cell-to-cell infection. There is some evidence that the mechanism of specific virus entry may be cell-type dependent.

1.3.2. Reverse Transcription

After the core is released into the target cell cytoplasm, the RNA genome must be reverse transcribed into dsDNA. This process is linked to the disassembly of the capsid core, which exposes the reverse transcription complex (RTC) (Zhang, 2000). Disassembly and reverse transcription may occur simultaneously, but the exact timing and connection are poorly defined. However, disrupting reverse transcription slows capsid uncoating (Hulme, 2011). Some viruses seem to completely disassemble, while others remain relatively intact (Fassati and Goff, 2001). Some structure is necessary, as complete disassembly via detergent prevents DNA synthesis *in vitro* (Haseltine, 1976). However, all of the components necessary for reverse transcription are contained in the virion, as purified virions can still complete the process if dNTPs are supplied (Baltimore, 1970; Temin and Mizutani, 1970). Reverse transcription involves a number of steps, outlined below and in Figure 1.4.

1. tRNA is annealed to the primer binding site (PBS), an approximately 18 bp complementary region (Mak and Kleiman, 1997). Annealing occurs prior to or during virus assembly.
2. Minus strand DNA is synthesized from 5' to 3', ending at the 5' end of the genomic RNA. This new synthesized piece is called minus-strand strong-stop DNA, and is about 100-150 bp long (Varmus, 1978).

3. RT mediates RNase H degradation of RNA in duplex with strong-stop DNA, and first-strand transfer delivers the DNA to the 3' end of the genomic RNA, based on matching sequences present on both ends (R) (Tanese, 1991).
4. Minus-strand DNA synthesis resumes, and RNase H degradation continues as needed, skipping a short polypurine tract (PPT) upstream of U3.
5. The RNA left at the PPT serves as a primer for plus strand DNA synthesis to begin, continuing this time to the 3' end. This results in a second strong-stop DNA, called plus-strand strong stop (Finston and Champoux, 1984).
6. The tRNA primer is displaced by plus-strand synthesis and second strand transfer attaches the exposed plus-strand PBS to the minus strand PBS, creating a circular template (Boone and Skalka, 1981).
7. Both strands then complete synthesis, using each other as a template.

The primer tRNA is a critical host-derived factor, which is bound to the genomic RNA during viral assembly and maturation (Mak and Kleiman, 1997). Each retrovirus uses a specific tRNA, which is captured by the virus in the cytoplasm during viral assembly (Kelly, 2003). For example, Moloney murine leukemia virus uses $tRNA^{Pro}$, while HIV uses $tRNA^{Lys,3}$ (Mak and Kleiman, 1997). RT begins reverse transcription with the tRNA primer marking the PBS as the start site. Although each viral particle contains two copies of the viral genome, only one dsDNA will go on to integrate into the host genome (Coffin, 2002). During reverse transcription, template switching can occur between the two strands, leading to increased viral diversity if differences in the two copies exist. It is also important to note that RT does not have a proofreading function, and is therefore error-prone, which results in sequence variation amongst viral swarms. However, each RT has a slightly different rate of error. Analysis of SIV variation by infecting monkeys with a molecular clone estimated variation to be 9×10^{-3} at single nucleotide bases, with most mutations resulting in amino acid changes (Burns and Desrosiers, 1994). *In vitro* reactions using purified HIV-1 reverse transcriptase determined an error rate of 3×10^{-5} per target base pair per replication

cycle, while the same assay using M-MLV reverse transcriptase only resulted in 2 x 10^{-6} (Mansky, 1996; Varela-Echavarria, 1992).

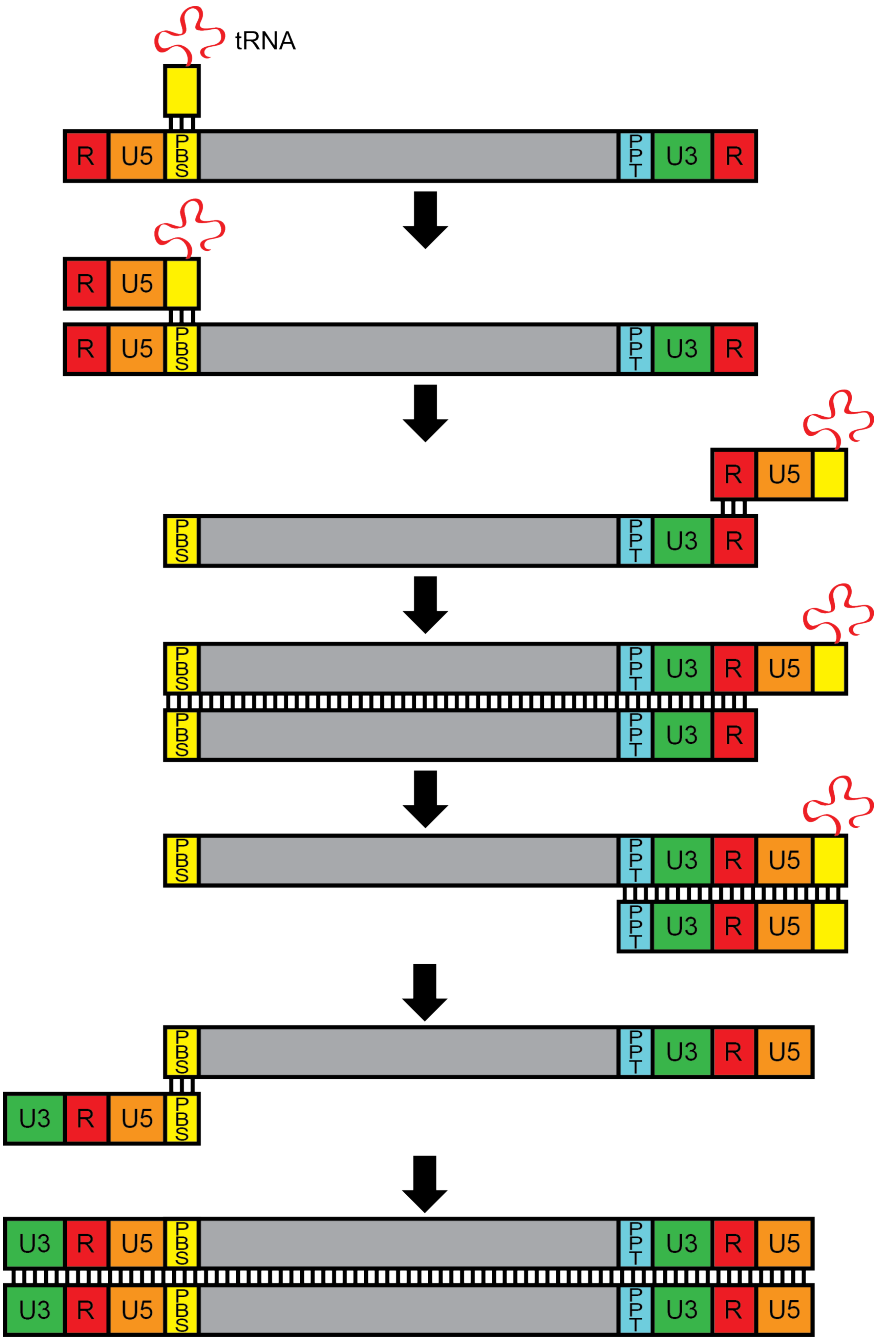


Fig. 1.4 Process of reverse transcription. Reverse transcription begins with DNA synthesis from a tRNA primer bound to primer binding site (PBS) towards the 5' end of the RNA genome. First-strand transfer allows continued synthesis using the 3' end of the RNA genome as a template. RNA in RNA:DNA hybrids is degraded by RNaseH. Second-strand transfer joins PBS

from the second strand to the complementary PBS on the first strand, allowing full extension of both strands. Figure credit: Matt MacMillan

1.3.3. Integration

The final product of correctly completed reverse transcription is a linear, blunt-ended dsDNA. While this is now ready for integration in the host genome, two alternatives can occur which prevent normal integration. First, integrase can mediate circularization of the complete DNA, known as a 2-LTR circle, which is thought to be a dead-end product that is unable to integrate into the genome. A separate circularization can also be caused by host ligases, similar to homologous recombination, which forms circles with a single LTR (Kulkosky, 1990).

Integration into the host genome is a defining characteristic of retroviruses. This process allows them to be stably maintained during division of cells, prevents degradation of the viral genome, and allows the cell to use host transcriptional machinery as the provirus acts essentially like a mammalian gene (Coffin, 2002). During reverse transcription, the viral nucleic acids and associated proteins undergo a transformation, which readies them for integration. This complex is called the pre-integration complex (PIC) (Delelis, 2010). One or more of these proteins typically helps hijack cellular machinery to actively transport the PIC to the nucleus (Naghavi and Goff, 2007). Some retroviruses must then await nuclear membrane breakdown during mitosis in order to access the viral genome. These viruses, including MLV and MMTV, cannot infect differentiated cells or cells in prolonged interphase. The PICs of HIV and other lentiviruses are actively imported into the nucleus, thereby surpassing the requirement for mitosis in some cell types, including macrophages (Bukrinsky, 1992).

Once the PIC is properly localized for integration at the chromatin, the enzymatic activities of integrase are required. First, as previously mentioned, the exonuclease activity acts on the 3' ends, removing two or more bases. Integration into the host genome uses a strand transfer process to simultaneously insert both ends of the linear DNA into host chromatin. Integrase utilizes nucleophilic attack using the 3'-OH of the processed strand to insert the DNA, removing 2-4 bases from the host DNA. Cellular factors then repair the dinucleotide overhang via polymerization and ligation (Delelis, 2008). Different retroviral integrases create characteristic duplications of host sequences at the insertion point, from 2-4 bp.

Once integrated, the provirus will be replicated prior to mitosis and passed on to daughter cells. The process is also inherently mutagenic, as integration disrupts the region of the chromatin it lies within. Insertion into a proto-oncogene can change the expression pattern of the gene, causing transformation and eventually lead to tumorigenesis (Robinson, 1982).

1.3.4 Transcription and Translation

Integration is necessary for transcription of retroviruses, which differs from other viruses that transcribe immediately from cytoplasmic viral RNA or DNA (Coffin, 2002). At this point in the viral lifecycle, the virus requires host cellular machinery in order to transcribe the provirus into RNA. RNA polymerase II and associated co-factors transcribe the provirus just as it would a cellular gene, beginning directly upstream of the R region of the 5' LTR. This region contains appropriate promoter sequences, including a core promoter element TATA box, which helps direct the cellular machinery (Laimins, 1984). Each viral promoter has unique enhancer sequences, thereby attracting different positive and negative regulators to the LTR. This can change the pattern of expression in different cell types and in different cellular conditions (Coffin, 2002). The rate of transcription can also be dependent on the site of integration, with placement near a cellular enhancer increasing transcription, or placement near an epigenetically silenced region decreasing or silencing transcription (Niwa, 1983). Conversely, the viral enhancers can influence the expression of cellular genes at the site of insertion.

Transcription terminates in the 3' LTR, at or near the end of the 3' R sequence, and is coupled to 3' end processing, which both cleaves and polyadenylates the transcript (Guntaka, 1993). These transcripts can also serve as templates for splice machinery, as splicing is required for expression of some viral genes, including *env*. Simple retroviruses will only produce unspliced and singly-spliced transcripts, but complex retroviruses can produce multiply spliced RNAs, as well as use alternative splice sites to produce variations on the same product (Coffin, 2002). Any spliced transcript will contain the same first exon, the U5 region of the 5' LTR, as well as the same sequence in the final exon, the U3 and R regions of the 3' LTR (McNally, 2008). The

unspliced transcript will either act as genomic RNAs for new virions, or as templates to translate Gag and Pro-Pol proteins.

Translation of viral genes begins similarly to that of cellular genes, with transport out of the nucleus. However, some viral proteins are translated from unspliced RNA, which is also necessary to be used as genomic RNA. Retroviruses have evolved mechanisms to allow for nuclear export of unspliced RNA. Simple retroviruses utilize *cis*-acting stem-loop structures called constitutive transport elements (CTE), which are bound by nuclear export factors to mediate export (Bray, 1994). Complex retroviruses, including HIV-1, typically encode for *trans*-acting protein factors that bind to viral RNA sequences and recruit other cellular export factors to transport this RNA/protein complex (Fischer, 1995; Harris, 2000).

Membrane-associated envelope proteins are translated within the endoplasmic reticulum, while other viral proteins such as Gag are translated within the cytoplasm. Viral RNAs, like cellular RNAs, are capped at their 5' end, which typically attracts translational machinery and allows ribosomes to bind and begin translation. However, it is hypothesized that some retroviruses, including HIV-1, may be able to navigate translation independently of ribosome scanning. The 5' untranslated region shows a significant degree of secondary structure, and AUG start codons downstream of the initial AUG codon are sometimes used (Yilmaz, 2006). Internal ribosome entry site (IRES) elements have been found within ALV, F-MLV, and HIV-2 (Balvay, 2007). Therefore, cap-dependent and cap-independent translation may both play important roles for retroviruses.

1.3.5 Assembly and Budding

After translation of the required proteins, assembly of new virions can occur. This typically takes place at the plasma membrane, but for both type B and type D retroviruses, can begin internally, with resulting immature virions assembling prior to associating with the plasma membrane or an internal cellular membrane. Localization of viral assembly is driven through association between core and membrane-associated envelope glycoproteins, and is dependent

on several factors, including cell localization properties of both envelope and core proteins (Sandrin and Cosset, 2006). For all retroviruses, a high concentration of Gag proteins must accumulate together in order to drive self-assembly. Gag-Gag interactions typically rely on specific regions of the nucleocapsid portion of the Gag polyprotein (Gheysen, 1989). Cellular factors also play an important role in mediating core assembly. For example, the ATPase ABCE1 is recruited to HIV-1 core assembly sites, and is required for HIV-1 assembly (Klein, 2007). Gag in immature cores also interacts with Env proteins expressed on the cell surface, ensuring its presence in the budding virion. However, these interactions are not required for membrane proteins to associate with viral particles, as envelope proteins from unrelated viruses can be used to produce pseudotype viruses with retroviral cores *in vitro* (Sandrin and Cosset, 2006).

RNA is also required for particle assembly. While viral like particles can form in the absence of specific viral RNA, the virus also selectively incorporates two copies of the full-length RNA genome (Campbell and Vogt, 1997; Johnson and Telesnitsky, 2010). As discussed previously, membrane association relies on the matrix region of the Gag polyprotein (see Chapter 1.2.1). Eventually, through exponential rates of recruitment and association, enough Gag particles will connect with the membrane and with each other to form a lattice, which begins the formation of a budding sphere (Carlson, 2010). While this lattice is important for creating the structure of virion, it is in an immature form. An immature virion is non-infectious. Protease is responsible for viral processing, cleavage of the polyproteins, and subsequent condensation of the viral core, a process resulting in mature virions (Dunn, 2002).

Chapter 1.4 Mouse Mammary Tumor Virus

Mouse Mammary Tumor Virus (MMTV) was the first mammalian retrovirus characterized. John Bittner first identified a milk-transmitted, microscopic agent as the cause of mammary carcinomas in mice (Bittner, 1936). Its identification has led to decades of research, with focuses on both virology and cancer, and the development of mouse models for breast cancer. It is characterized as a betaretrovirus, which consists of virions with Type B or Type D morphology

(see Table 1.1). Type B viruses are identified by an eccentric condensed core and prominent surface spikes. Unlike gammaretroviruses and lentiviruses, betaretroviruses are assembled intracellularly, and then traffic to the plasma membrane for budding (Coffin, 2002). MMTV is a complex retrovirus, encoding accessory genes as well as the basic genes required for retroviral replication. Accessory genes are non-structural, and typically help the virus evade the host immune response *in vivo.* The genome is approximately 9 kB in length, with notably long LTRs (approximately 1.3 kB compared to 400-800 bp of most other retrovirus) (Coffin, 2002).

Viral proteins are translated from five different transcripts. The full-length genomic RNA encodes Gag and Pol proteins, and an additional non-structural protein dUTPase (Dut). The *dut* gene is also carried by equine infectious anemia virus (EIAV), where it likely aids in replication in nondividing cells which carry a surplus of UTP (Elder, 1992). A 73 kD Env polyprotein is translated from a singly spliced transcript. This is cleaved by furin into SU (52kD) and TM (36 kD), both of which are glycosylated. The Env subunits facilitate entry into acidic endosomes. Unlike many retroviruses, MMTV entry is pH dependent (Ross, 2010). Two additional viral proteins are translated from additional splice variants. A double spliced mRNA encodes the regulator of export of MMTV (Rem), which is similar to the HIV accessory protein Rev. These proteins aid in transportation of unspliced viral RNA into the cytoplasm (Mullner, 2008). The other is encoded by an ORF in the 3'LTR, discussed below.

The MMTV LTRs define several important characteristics of MMTV. First, the ORF in U3 encodes a gene termed *sag*, producing the accessory protein Superantigen (Sag). The inclusion of this ORF contributes to the particularly long LTR, even though the R region of the LTR is the shortest among all retroviruses at 15 nucleotides (Coffin, 2002). Sag is necessary *in vivo* to boost early viral infection in target lymphocytes, resulting in increased viral spread (Held, 1993; Golovkina, 1998). The LTR is also home to a number of MMTV-specific transcription factor binding sites, which act within the 5' LTR after proviral integration. The transcription factors necessary are tissue-specific, allowing for expression in B and T cells, as well as mammary epithelial cells (Reuss and Coffin, 2000). Hormones including glucocorticoid and progesterone

regulate expression in the mammary gland because of hormone responsive elements in the LTR (Payvar, 1983).

Exogenous MMTV infection is transmitted vertically via cell-free virus expressed into the breast milk. The virus binds to transferrin receptor 1 (TfR1) on the cell surface of dendritic cells in the Peyer's patches of the intestine, and then spreads to T and B cells in this secondary lymphoid organ (Ross, 2002; Ross, 2008). Infected dendritic cells and B cells present the viral Sag protein on their surface major histocompatibility (MHC) class II proteins, and thereby activate T cell receptors on specific CD4+ T cells. This stimulates activation of the T cells, which causes them to proliferate locally, provide help to nearby B cells, and produce cytokines that attract additional cells of the immune system to the area. This attraction then provides a large pool of new target cells (Ross, 2010). Different strains of MMTV encode different Sag proteins, each of which binds to a different T-cell receptor based on variability in the C-terminus of the protein (Ross, 2008). This leads to strain-dependent activation of different subsets of CD4+ T cells, which can be monitored through flow cytometry. Eventually, activation of the targeted subset will lead to almost complete depletion (Ignatowicz, 1992). Without Sag-mediated activation, the virus will not replicate to levels necessary for viral spread to the mammary gland (Golovkina, 1998). This depletion also alters the immune response profile of the mouse, causing an infected mouse to be more or less susceptible to other pathogens (Ross, 2010).

The MMTV receptor, the murine transferrin transporter, TfR1, is highly expressed on activated lymphocytes and dividing mammary epithelial cells (Schulman, 1989; Brekelmans, 1994). This expression pattern helps to determine the tropism of the virus. Infected lymphocytes trafficking to the mammary gland deliver virus to this region. Mammary epithelial cells undergo division during both puberty and pregnancy, at which time cells are susceptible to infection by lymphocyte-derived virions. Because of cellular hormonal responses, viral transcription is highest during pregnancy and lactation, producing a large quantity of new virions that are shed into milk (Ross, 2008). Besides this virus expressed in milk, virus is completely cell-associated *in vivo*.

Mammary tumors are typically formed when MMTV integrates near a cellular proto-oncogene, as tissue-specific, hormonally responsive transcription of the virus can upregulate nearby genes (Ross, 2010). Time to tumor and tumor incidence mirrors overall viral load, as increased viral replication leads to more integration sites and higher likelihood of integration near an oncogene (Golovkina, 1993). The majority of MMTV-induced tumors are caused by integration near just a few families of genes, including *Wnt, Int, Hst,* and *Fgf* genes (Callahan and Smith, 2008).

Mice can also acquire MMTV through endogenous proviruses. Almost all inbred mice carry at least one endogenous Mtv, although more than 50 have been identified and some strains of mice carry up to ten (Cohen and Varmus, 1979). These proviruses are named *Mtv* loci and encode primarily defective viruses due to mutations acquired over the 20 million years they are estimated to have been in the mouse genome (Ross, 2008). However, the majority of endogenous loci encode a functional Sag protein, which can lead to activation and depletion of specific subsets of CD4+ T cells (Kappler, 1988). Endogenous viral transcripts can be packaged into exogenous virions and cause the formation of infectious recombinant viruses. This can lead to significant viral diversity, and cause increased viral spread due to activation of additional subsets of T-cells by the endogenous *Mtv* Sag (Golovkina, 1994).

Chapter 1.5 Murine Leukemia Virus

Murine leukemia viruses, while similar to MMTV in terms of basic retroviral structure and function, present a number of important differences. MLV was discovered over two decades after MMTV, but through similar experiments that showed that leukemia could be transmitted to a newborn mouse through a microscopic, filterable agent (Gross, 1957). It is characterized as a gammaretrovirus, and displays Type C viral morphology, which is characterized by a central spherical condensed inner core (See Table 1.1; Coffin, 2002). Unlike MMTV, MLV is a simple retrovirus, which does not encode accessory proteins outside of Gag/Pol/Env. MLV does not cause general pathogenic effects in infected cells, unless transformation occurs (Rein, 2011).

However, immune responses to infected cells in the nervous system can cause motor disease and encephalopathy (Zachary, 1992).

The gag gene of MLV has some important differences. As stated previously, an additional protein, p12, is produced from the Gag polyprotein. Additionally, some, but not all, murine leukemia viruses produce a secondary form of Gag, called glyco-Gag or $gPr80^{gag}$. This protein begins from a start site upstream of the normal Gag AUG initiation codon, adding an additional 264 bases onto the amino-terminus (Prats, 1989). Unlike the smaller, non-glycosylated Gag protein, glyco-Gag undergoes Golgi-processing and is not incorporated into budding virions. This form of Gag is dispensable *in vitro*, but is necessary to efficiently replicate in some strains of mice (Corbin, 1994). Recent research has shown that this protein functions to counteract the intrinsic immune factor, APOBEC3 (see Chapter 1.6.5) (Kolokithas, 2010).

MLV *env* is transcribed and translated into the Env polyprotein, which is glycosylated in the Golgi apparatus and cleaved by a furin-like protease to form gp70 SU and the p15E transmembrane TM. These proteins trimerize before relocating to the plasma membrane (Rein, 2011). Viral maturation includes an additional processing step of Env, in which a small cytoplasmic C-terminus portion of p15E is cleaved off. This cleavage is critical for fusion of the virus to a target cell (Freed and Risser, 1987). MLVs are polymorphic in terms of their cellular receptors. The receptor use determines host specificity or tropism, and is useful for classifying MLV into four main groups: ecotropic, amphotropic, xenotropic, and polytropic. Ecotropic MLVs are only able to infect mice, and do so using the mCAT-1 receptor, a cationic amino acid transporter (Albritton, 1989; Kim, 1991). Amphotropic MLVs have a wider tropism, infecting both mice and cells of many other species using Pit-2, a type III sodium phosphate cotransporter, as a receptor (Hartley and Rowe, 1976; Miller, 1994). Xenotropic MLVs are unable to infect mice, but can infect cells of other species, including human, cat, and rabbit cells. These viruses use XPR-1 as a receptor, which is a multiple-membrane-spanning protein with no identified function (Kozak, 2010). Finally, the polytropic MLVs have similarities to both amphotropic and xenotropic MLVs. They are able to infect some mouse cells and some cells of other species, similar to amphotropic

MLVs. However, they use XPR1 as their receptor, like xenotropic MLVs, but have a more limited host range (Kozak, 2010). Viruses of each classification are found in the genomes of mice as endogenous viruses. As with MMTV, transcripts of endogenous MLV can cause endogenous-exogenous recombinants in infected cells.

Although murine leukemia viruses were first named based on their ability to cause leukemias in mice, the retroviral group has been expanded to include many related mouse gammaretroviruses. The most well characterized viruses include Friend MLV (F-MLV), Moloney MLV (M-MLV), Rauscher MLV (R-MLV), and Abelson MLV (A-MLV). All of these cause slightly different phenotypes of disease and are named after the primary investigator responsible for the discovery (Coffin, 2002). Work described in subsequent chapters focuses on Moloney murine leukemia virus, which causes T-cell lymphomas due to insertional activation of proto-oncogenes (Dunn, 1961). Tissue tropism of the virus is carried in the 5' LTR, which carries *cis* elements including a direct-repeat enhancer responsible for promoting transcription and thus potential pathogenesis (Li, 1987). Different mouse strains have diverse susceptibilities to MLV-induced leukemia, ranging from 1-100%. Susceptible mice develop disease in 6 to 24 months (Shiff and Oliff, 1986). Mice can become infected through horizontal and vertical routes: either by postnatal exogenous infection or trans-uterine congenital exogenous infection. The majority of mother-to-pup transmission is thought to occur via breast milk (Duggan, 2006).

Chapter 1.6 Retroviral Restriction Factors

The host cells targeted by viruses have many means of defense. Both the innate and adaptive immune responses recognize invasion by pathogens, and then actively work to clear infection. The innate immune system is comprised both of cellular proteins and specific immune cells. Most cell types are able to recognize pattern-associated molecular patterns (PAMPs), molecular signatures of invading pathogens, and mount a quick response. Other cells, including leukocytes such as natural killer cells, mast cells, and eosinophils, as well as phagocytic cells including macrophages, neutrophils, and dendritic cells are professional innate immune cells

(Janeway, 2001). These cells recognize and bind common constituents of pathogens and will lead to inflammation and assist with viral or bacterial clearance. The adaptive immune response is induced secondarily, and requires highly specific pathogen recognition by T- and B-lymphocytes via antigen-presenting cells. This pathway takes longer to mount during an initial infection, but leads to a long-lasting defense generating immunological memory (Janeway, 2001).

Chapter 1.6.1 Overview of Retroviral Restriction Factors

Restriction factors are cellular proteins of the innate immune system that make up the intrinsic immune system that serve to block infection. Many of these factors have been found to specifically inhibit retroviruses. A retroviral restriction factor is identified by several major characteristics. First, it must directly cause a decrease in viral infectivity, and have a dominant effect. Targeted retroviruses must show evolution of a counter-defense mechanism, which is necessary for viral survival against a potent restriction factor. This defense and counter-defense must continuously evolve according to the Red Queen hypothesis, which proposes that constant adaptation and evolution is required for survival against changing opposing organisms (Van Valen, 1973). Therefore, the restriction factor will show signatures of rapid evolution, typically in an increased number of amino acid substitution mutations, often in regions of the protein that specifically interact with the virus. Finally, the expression regulation of a restriction factor is tightly controlled by the general anti-viral immune response of a cell, typically the interferon response (Neil and Bieniasz, 2009).

Restriction factors of HIV-1 are best studied, and include APOBEC3, BST-2/tetherin, and TRIM5. Genes responsible for conferring resistance to mouse retroviruses were discovered even before the onset of the HIV epidemic. Friend virus models provided information about genetic susceptibility to disease, with many of these genetic determinants encoding retrovirus-specific restriction factors. Susceptibility to F-MLV is limited to certain strains of mice, with the two major host ranges determined by a gene called *Fv-1* (Pincus, 1971). Susceptibility to different strains of MLV (either N- or B-tropic, named for their ability to infect NIH Swiss or Balb/c mice) depends on

the allele of *Fv-1* (either N- or B-type) encoded by a specific mouse strain. A third class of MLV strains termed NB-tropic can infect mouse cells regardless of the *Fv1* allele. Further research showed that the viral determinants for tropism map to the capsid protein, and the protein encoded by *Fv1* is related to endogenous retroviral Gag proteins (Bieniasz, 2003). However, the exact mechanism of restriction is unclear. Another genetic factor controlling F-MLV infection is *Rfv-3*. Mice expressing the *Rfv*3 recessive allele fail to mount neutralizing antibodies, and cannot control infection, while mice expressing the dominant allele recover from viremia (Chesebro and Wehrly, 1979). The molecular identity of *Rfv3* was not discovered until after the HIV-1 restriction factor APOBEC3 was identified in human cells, at which time *Rfv3* was found to encode murine APOBEC3 (Santiago, 2008).

Chapter 1.6.2 HIV-1 and the Discovery of APOBEC3

HIV-1 infectivity is inhibited by several restriction factors, which target different steps of the viral life cycle. HIV-1 is a complex retrovirus, carrying genes for six additional non-structural accessory proteins, one of which is named Vif (virion infectivity factor) (Fisher, 1987). Studies using molecular clones of HIV with the Vif gene deleted (HIVΔ*vif*) show markedly decreased infectivity in some cell types, thus providing Vif its name (Gabuzda, 1992**;** Sova and Volsky, 1993**;** von Schwedler, 1993). Restrictive cells are called non-permissive cells, and include normal HIV targets such as CD4+ T cells and macrophages. However, pseudotyped HIVΔ*vif* is able to infect to wild-type levels in other cells, termed permissive cells, which include non-hematopoietic cells such as HeLa (human cervical carcinoma cells), 293T (immortalized human kidney epithelial cells), and COS (immortalized monkey kidney fibroblasts) cells (Sova and Volsky, 1993). The nonpermissive phenotype is dominant over permissive in heterokaryons, which led to the hypothesis that nonpermissive cells expressed a restriction factor whose function is overcome by Vif **(**Madani and Kabat, 1998).

Interestingly, infected nonpermissive cells are able to support a single round of replication, producing high levels of virions comparable to permissive cells. The block to infection

occurs in the second round of infection, early after viral entry (see Figure 1.5A vs. 1.5B). This block is dependent on the producer cell type (nonpermissive vs. permissive), but independent of the target cell type (see Figure 1.5C). Gene expression screens of nonpermissive and permissive cells narrowed the list of potential candidates. Genetically related cell lines were used in a complementary DNA subtraction strategy using a nonpermissive T-cell line CEM and a permissive subclone of this cell line, CEM-SS (Sheehy, 2002). This method removes all cDNAs present in both cell types, and enriches sequences found solely in CEM cells. These cDNAs were screened in further RNA hybridization screens of other permissive cell lines, and resulted in the identification of a single gene, called CEM15 (Sheehy, 2002). When cloned and transfected into new cells, CEM15 could convert permissive cells into nonpermissive cells, and was later identified as a known member of the apolipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like family and named human APOBEC3G (A3G). These studies also showed that APOBEC3G was packaged into virions budding from non-permissive cells, which explained the block in the subsequent round of infection (Sheehy, 2002; Fig. 1.5).

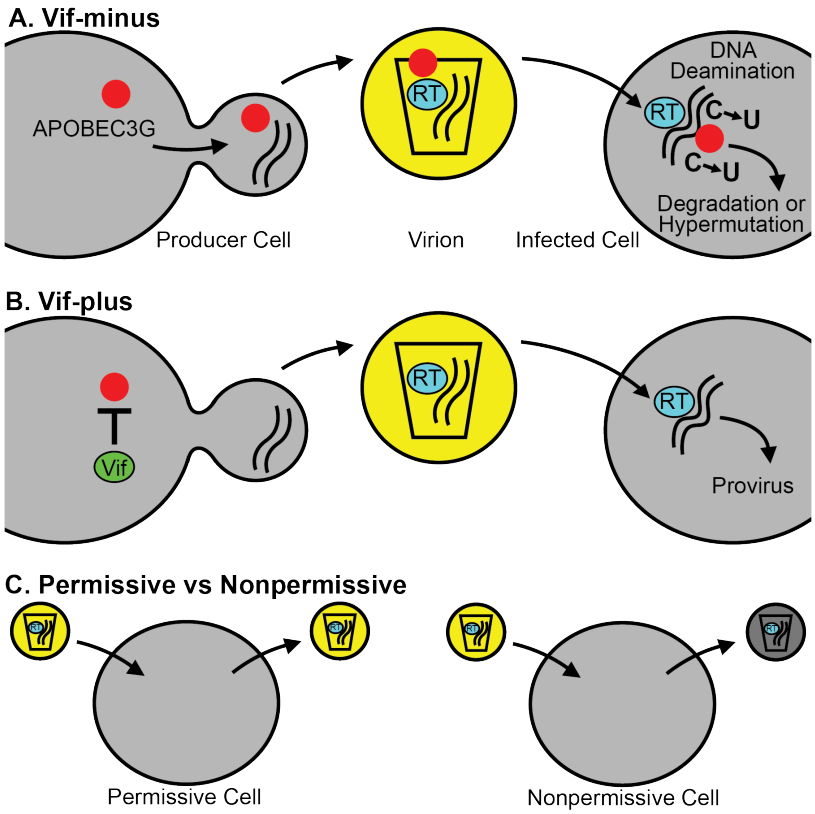


Fig. 1.5. Wild-type HIV and HIVΔVif in Permissive and Nonpermissive Cells. (A) HIVΔVif infection of APOBEC3-expressing cells is able to produce virions at wild-type levels. These virions contain APOBEC3G, which acts to restrict reverse transcription and deaminate newly formed viral DNA, preventing productive infection. (B) Cells infected with HIV containing Vif suffer degradation of cytoplasmic APOBEC3G, preventing incorporation into virions. Subsequent infection with these virions is not impeded. (C) Permissive cells do not contain APOBEC3, allowing production of infectious virions regardless of Vif expression. Nonpermissive cells contain APOBEC3, leading to production of noninfectious APOBEC3-containing HIVΔVif virions (shown in gray). Figure credit: Matt MacMillan

Chapter 1.6.3: APOBEC3 Evolution

Cytidine deaminases are found in a wide array of organisms, including yeast. However, the AID/APOBEC gene family is found only in higher organisms. Genes are clustered into AID, APOBEC1, APOBEC2, and APOBEC3 subfamilies based on nucleotide sequence surrounding the active site. Homologs of AID and APOBEC2 are found in chicken, frog, and bony fish (Conticello, 2005). Both of these loci are found in syntenic regions in chickens, mice, and humans. However, APOBEC1 and APOBEC3 evolved later, likely derived from duplication of AID, and are found only in mammals. APOBEC3 underwent further duplication early in primate evolution, producing seven bona fide APOBEC3 genes, designated APOBEC3A to APOBEC3H, all of which are clustered together on human chromosome 22 (Conticello, 2005). This cluster is syntenic to a region of chromosome 15 in mice, although lack of expansion provides only a single APOBEC3 gene in mice. Other mammalian species have varying numbers of APOBEC3 genes. For example, pigs have two APOBEC3 genes, cats have four, and horses have six (LaRue, 2008).

APOBEC3 genes in both primates and rodents have undergone evolutionary changes suggestive of positive selection, and it is hypothesized that host/virus interactions have been the driving force for rapid expansion and evolution. In humans, five of the APOBEC3 genes display some of the highest signals for positive selection in the entire genome, likely attributed to an ancient battle between viruses and hosts (Sawyer, 2004). The evolution of this family provides primates with a more varied antiviral response compared to mammals with fewer APOBEC3 genes. The human APOBEC3 genes also have antiviral activity on other viruses, including HTLV-1, Hepatitis B virus, and endogenous retroviruses, suggesting that many different viruses may have helped shape this evolution (Chiu and Greene, 2008).

1.6.4 APOBEC3 Structure, Function, and Regulation

Human APOBEC3 belongs to a family of cytidine deaminases, enzymes that are able to edit RNA and/or DNA (Teng, 1993; Harris, 2002). The most well characterized members of this

family are activation-induced deaminase (AID) and APOBEC1. AID is expressed in B cells and specifically edits the genome to promote somatic hypermutation and class switch hypermutation (Muramatsu, 2000). APOBEC1 is expressed in the gastrointestinal tract and edits a single site of apolipoprotein B mRNA to induce expression of a truncated version of the protein, which is important for regulation of lipid metabolism (Mehta, 2000). All of the family members share a similar domain structure. An alpha-helix is followed by a catalytic domain, a linker peptide, and a pseudoactive domain (Jarmuz, 2002).

Some family members contain a duplication of this structure, with two catalytic domains. Generally, a single domain is the active domain, while the other functions to bind substrate and aid packaging into the virion (Goila-Gur and Strebel, 2008). Human APOBEC3G's active domain is the C-terminal domain, whereas mouse APOBEC3 has an active N-terminal domain (Hakata and Landau, 2006). The catalytic domain is characterized by a conserved zinc-binding motif (Cys/His)-Xaa-Glu-Xaa_{23-28}-Pro-Cys-Xaa_{2-4}-Cys. The His and Cys residues are responsible for coordinating the Zn^{2+} and the Glu plays a role in proton transfer during hydrolytic deamination of a cytidine. This deamination occurs at the C4 position of the C base, which converts the cytidine to a uracil (Chiu and Greene, 2008).

Although cytidine deaminases can edit DNA or RNA, APOBEC3 selectively targets single-stranded DNA (Suspene, 2004). This is the direct means by which viral replication is targeted. APOBEC3 is packaged into virions and is closely associated with the viral core, with access to the reverse transcription complex during reverse transcription (Schafer, 2004). Reverse transcription first creates an RNA:DNA hybrid, but the RNA is degraded by reverse transcriptase (see Chapter 1.2.2). The RNA degradation is also critical for APOBEC3 activation, as APOBEC3 bound to RNA is not enzymatically active (Soros, 2007). The minus-strand viral DNA is thus the single stranded target of APOBEC3 (Yu, 2004).

Cytidine deamination of the minus-strand causes extensive dC-to-dU mutagenesis. DNA containing uracils is typically targeted for degradation by DNA base repair enzymes, but it is unclear whether this occurs in the viral context (Kaiser and Emerman, 2006). A dU in the minus

strand causes dA to be incorporated into the coding plus-strand, resulting in G to A mutations. Mutation rates vary, but can reach up to 10% of all dG residues (Chiu and Greene, 2008). The volume of mutagenesis of different regions of the genome is directly related to the reverse transcription process. DNA that is produced earlier is single stranded for a longer period of time, therefore more susceptible to APOBEC3 editing. This results in a 5' to 3' increasing gradient of mutation through the genome (Suspene, 2006). These mutations disrupt viral open reading frames and can translate into stop codons.

Cytidine deaminases prefer to target a cytidine within a specific sequence context. Examination of site targets can be used to distinguish the probable deaminase. This specificity maps to a loop of 9-11 amino acids, which targets the enzyme to the specific sequence. The loop is composed primarily of amino acids with basic, acidic, and aromatic side chains, which may help mediate specific interactions (Kohli, 2009). For example, APOBEC3G prefers to target 5'-CC-3', resulting in a GG-to-AG pattern on the coding strand (Mangeat, 2003). However, certain GG dinucleotides are "hotspots" for deamination over others, which is likely due to secondary structure of the ssDNA (Holtz, 2013).

APOBEC3 proteins also exert cytidine deaminase-independent antiviral activity. The mechanisms are unclear, but include impeding tRNA-mediated reverse transcription initiation, impairing plus-strand DNA transfer, and causing inefficient nuclear import and integration (Chiu and Greene, 2008). However, perhaps the most important deaminase-independent antiviral activity comes from inhibition of reverse transcription post-initiation. APOBEC3 with impaired active sites (via site directed targeting of necessary amino acids) are still packaged into virions, and still exhibit a strong antiviral capacity in the absence of G to A mutations (Newman, 2005). These mutant proteins can inhibit viral cDNA accumulation by stalling elongation of reverse transcription products (Bishop, 2008). This inhibition may be mediated by interactions with viral reverse transcriptase (Wang, 2012). The individual contributions of deamination and deaminase-independent activity are unclear.

APOBEC3 is interferon-inducible, and is a major component of this antiviral response for targeted retroviruses (Peng, 2007; Okeoma, 2009). Different APOBEC3 family members are differentially regulated, and this regulation leads to different expression levels in different cell types. APOBEC3G protein is primarily expressed in peripheral blood mononuclear cells and immune organs such as the spleen and thymus. APOBEC3A, APOBEC3C, and APOBEC3F are also highly expressed in PBMCs. Stimulation of naïve T-cells leads to induction of all of the APOBEC3 genes except for APOBEC3A. APOBEC3A is instead highly expressed in CD14+ cells, including macrophages and monocytes (Koning, 2009). Expression of all APOBEC3 proteins is also generally high in the lung, the ovary, the cervix, and adipose tissue, and low in the brain, esophagus, kidney, testes, and skeletal muscle in humans (Refsland, 2010). APOBEC3G is also highly inducible in the liver (Bonvin, 2006).

In order to be packaged into budding virions, APOBEC3 must be expressed in the cytoplasm. Highly expressed cytoplasmic APOBEC3 can also protect a cell by restricting incoming viruses, without specific packaging for activity in subsequent round of infection (Vetter and D'Aquila, 2009). However, some members of the APOBEC family are expressed in the nucleus (APOBEC3B), or pan-cellular (APOBEC3A). These members have significantly less activity against exogenous retroviruses (Chiu and Greene, 2008).

Chapter 1.6.5: Murine APOBEC3 and Murine Retroviruses

Multiple allelic variants of mouse APOBEC3 (mAPOBEC3) exist which differ in their ability to restrict murine retroviruses. The two main alleles are represented by inbred strains BALB/c and C57BL/6, and will be referred to as such ($APOBEC3^{BALB/C}$ and $APOBEC3^{C57BL/6}$). These two alleles differ in protein sequence, splicing pattern, and expression level, all of which may contribute to differences in viral resistance (see Table 1.2).

There are also fifteen polymorphic amino acids between the two APOBEC3 alleles (Figure 1.6). Genetic analysis from dozens of inbred and wild mouse species identified twenty amino acids that are under positive selection, eleven of which are in this polymorphic group

(Sanville, 2010). Several of these residues are found within two clusters in the active cytidine deaminase domain, and are likely involved in long-term genetic conflict with pathogen targets. These amino acids are directly involved in viral restriction, with the APOBEC3$^{C57BL/6}$-encoded residues creating a more restrictive enzyme. Genetic analysis also discovered a polymorphism in a splice branch selection site, and an inserted sequence in the more virus restrictive APOBEC3$^{C57BL/6}$ that acts as an enhancer (Okeoma, 2009). This insertion is a long terminal repeat from a xenotropic mouse gammaretrovirus, and drives elevated mRNA expression levels (Sanville, 2010). Additional post-transcriptional regulation influences translation of APOBEC3$^{BALB/c}$, as the inclusion of the fifth exon leads to decreased protein production (Li, 2012). This leads to significantly more APOBEC3 protein produced in animals carrying the APOBEC3$^{C57BL/6}$ allele, further contributing to increased viral restriction. Data included herein support this differential expression as the primary influence on the difference in restriction (see Chapter 2.4.5).

Table 1.2. Murine APOBEC3 Alleles

	APOBEC3$^{C57BL/6}$	**APOBEC3$^{BALB/c}$**
Exons	8	9
Polymorphic Residues	15	15
Prototype Strains	C57BL/6, NZB/N, RF/J	BALB/c, C3H/He, AKR, NIH Swiss
mRNA Expression	High	Low
Murine Retrovirus Restriction	High	Low

The full length *Apobec3* gene encodes nine potential exons. Sequence variations between the two alleles, specifically a difference in number of TCCT repeats in intron 4 and a single nucleotide difference in exon 5, determine the inclusion or exclusion of exon 5. APOBEC3$^{C57BL/6}$ excludes this exon, which translates into a slightly shorter protein of 49 kDa compared to the 52 kDa APOBEC3$^{BALB/c}$ (Li, 2012). Mice carrying the allele for APOBEC3$^{BALB/c}$

produce primarily the full-length transcript, but also produce a small proportion of transcripts lacking exon 5. Transcripts have also been found that lack both exon 2 and 5, but are very rare (Sanville, 2010; Okeoma, 2009; Takeda, 2008).

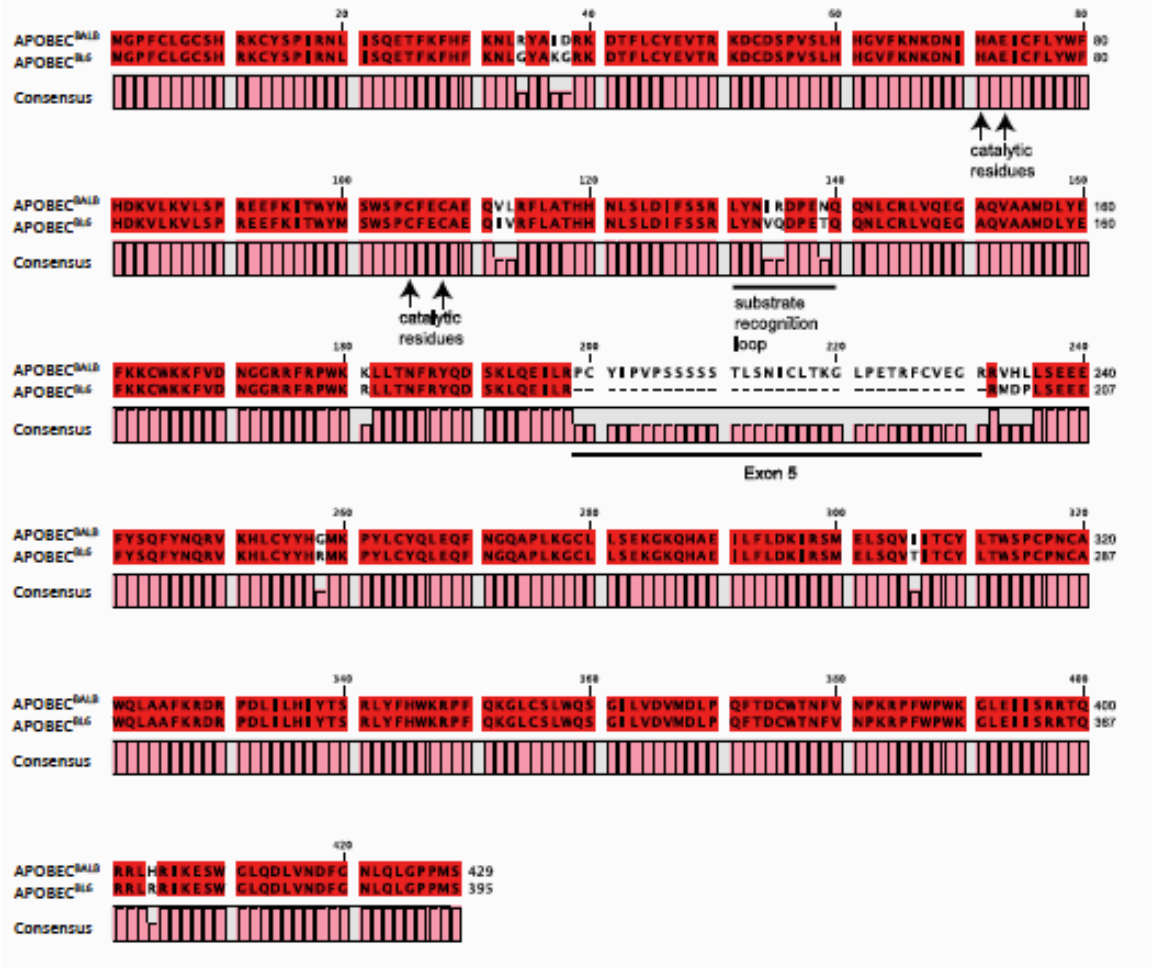


Figure 1.6 Murine APOBEC3 Alleles. Protein sequence alignment of the two allelic variants of mouse APOBEC3. Alleles are labeled by the representative strain encoding each. APOBEC^{BALB} retains a fifth exon whereas APOBEC$^{C57BL/6}$ does not due to a putative change in splice acceptor site (Sanville, 2010). Polymorphic amino acids do not directly affect residues critical for catalysis of cytidine deamination, but several amino acid differences in the substrate recognition loop potentially alter the affinity for cytidines in different DNA sequence contexts (Kohli, 2009). Figure credit: Rahul Kohli

Mouse APOBEC3 shows a strong inhibitory phenotype against HIV *in vitro*. This inhibition is resistant to Vif, and is associated with heavy G-to-A mutagenesis of the viral genome, at a rate of 1 mutation per 100 bp (Bishop, 2004; Hakata and Landau, 2006). These mutations display a pattern of 5'-TT/CC-3', giving mouse APOBEC3 a distinctive signature compared to human APOBEC3 preferred targets; human APOBEC3G prefers to deaminate at 5'-CCC'-3'; APOBEC3F prefers 5'-CTC-3', and APOBEC3B prefers 5'-C/GTC-3' (Bishop, 2004). The beteretrovirus mouse mammary tumor virus is inhibited by both human and murine versions of the protein, and served as the first model to demonstrate the antiviral activity of APOBEC3 *in vivo*, using APOBEC3-deficient knockout animals (Okeoma, 2007). Although MMTV is not completely restricted by APOBEC3, mice lacking the gene have increased viral loads, faster viral replication, and decreased time to mammary tumor formation. However, this restriction occurs in the absence of G-to-A mutagenesis *in vivo*, contrary to that seen in an HIV vector restricted by murine APOBEC3 *in vitro* (Okeoma, 2007). Mouse mammary tumor virus is also differentially restricted by different APOBEC3 alleles, with an identical phenotype compared to Friend MLV (Okeoma, 2009a). Mice expressing APOBEC3$^{C57BL/6}$ are thus more resistant to several classes of murine retroviruses compared to those expressing APOBEC3$^{BALB/c}$. This resistance appears to be directly correlated to levels of APOBEC3 expression, as inducing expression *in vivo* increases restriction in both APOBEC3$^{C57BL/6}$- and APOBEC3$^{BALB/c}$- expressing mice (Okeoma, 2009b).

Murine leukemia viruses are also inhibited by murine APOBEC3. This observation was first made before the discovery of APOBEC3, as a genetic factor influencing recovery from F-MLV (see Chapter 1.6.1). A single autosomal gene named *Rfv3* (Recovery from Friend virus 3) has long been known to influence neutralizing antibody responses and help control MLV viremia, but the molecular identity was not known until after the discovery of APOBEC3 (Chesebro and Wehrly, 1979). Mice are categorized as either resistant or susceptible to Friend virus, and this pattern matches the expression of allelic variants, with those expressing APOBEC3$^{BALB/c}$ characterized as susceptible (Santiago, 2008; Takeda, 2008). For MLV, APOBEC3 acts as an

intrinsic immune factor, but also affects the adaptive immune system. One theory suggests that restriction by APOBEC3 reduces virus-induced immune dysfunction, thereby promoting neutralizing antibody responses (Santiago, 2010).

APOBEC3's role in viral restriction extends to other murine retroviruses as well, including Moloney murine leukemia virus. Bone marrow-derived dendritic cells isolated from APOBEC3-/- mice support higher levels of infection than those isolated from mice expressing APOBEC3. This higher level of infection also leads to more rapid development of leukemia (Low, 2009). Another gammaretrovirus, the endogenous AKV virus found specifically in AKR mice, is also restricted by both alleles of APOBEC3 *in vitro*, under identical expression conditions. APOBEC3 proteins expressed via transfection also caused low-level deamination of AKV genomes, at a rate of approximately 2 mutations per kB (0.2%) (Langlois, 2009).

Evidence suggests that mouse APOBEC3 may have had mutagenic activity in the past, as some clades of endogenous murine leukemia viruses show significant G-to-A mutations consistent with the mouse APOBEC3 editing pattern (Jern, 2007). These mutations may have helped inactivate the viruses as they integrated into the mouse germline. Retrotransposons, particularly the IAPE element, are also potently restricted by mouse APOBEC3 *ex vivo*, and show low-level mutagenesis (0.1-0.5%) (Esnault, 2008). Mouse APOBEC3 does not deaminate present day exogenous murine leukemia viruses, although the protein is packaged into budding virions and causes inhibition of reverse transcription (Browne and Littman, 2008). Based on the ability for mouse APOBEC3 to restrict and mutate HIV *in vitro*, it was hypothesized that murine retroviruses have developed mechanisms to resist the mutagenic effects of their species-specific APOBEC3 (Rulli, 2008).

Human APOBEC3G is an incredibly potent antiviral factor of HIVΔVif, but is unable to restrict wild-type HIV. Murine APOBEC3 shows an intermediate phenotype in restricting murine retroviruses such as MMTV and MLV. This inability to fully restrict these viruses suggests a potential Vif-like factor. However, MLV does not encode for any accessory proteins, and the accessory proteins of MMTV have other defined roles in the viral lifecycle. However, MLV does

encode for an elongated form of Gag, known as glycosylated Gag (gGag) (see Chapter 1.5). Without this factor, MLV replicates inefficiently in cells (and mice) expressing mouse APOBEC3, but replicates to wild-type levels in APOBEC3-deficient cells (Kolokithas, 2010). This restriction can be either virion- or cell-associated, and gGag does not prevent packaging of the protein into the virion. However, viruses expressing gGag have more stable viral cores, which may protect the viral reverse transcription complex from APOBEC3 (Stavrou, 2013). Reverse transcription within these viruses proceeds unimpeded, while reverse transcription of gGag-mutant viruses is reduced. The mechanism by which gGag improves capsid stability and prevents APOBEC3 access is unknown.

Chapter 1.6.6: APOBEC3 and Drug Resistance

A defining feature of retroviruses is significant diversity and enormous evolutionary potential based partly on error-prone reverse transcriptase enzymes (approximately one mutation per genome per replication cycle) (Chiu and Greene, 2008). Other sources of viral mutation come from cellular RNA polymerase (during transcription of the genome) and genomic recombination, as reverse transcriptase can utilize both copies of the viral RNA packaged into the viral core (Rambaut, 2004). This rapid evolution is a particular problem in the case of HIV-1. This virus has a particularly high replication rate and maintains a large population size, which together with an error prone polymerase, allow the virus to quickly evolve to evade immune responses and develop resistance to antiviral drugs. This variation also makes creating an effective HIV vaccine more difficult. APOBEC3 is also a source of viral mutagenesis. The degree to which APOBEC3 is responsible for viral evolution is unknown, but it was early noted that G-to-A mutations account for a large percentage of drug-resistance mutations (Berkhout and de Ronde, 2004). Although APOBEC3 can cause lethal mutagenesis based on missense and nonsense codon changes, if APOBEC3 is restricted by Vif-mediated degradation, limited mutations may occur. Thus, APOBEC3 may not only serve as an antiretroviral factor, but also as a potential proviral factor by aiding viral evolution.

Several important experiments have provided evidence that APOBEC3-mediated mutagenesis can in fact aid viral evolution and lead to drug resistance. The majority of these experiments have utilized *in vitro* transfection and infection systems with human APOBEC3G. Using a wide range of APOBEC3 expression levels in single round infections, Sadler and colleagues found a number of integrated HIV proviruses with only a single APOBEC3 mutation (at least within the viral gene examined). Although APOBEC3G levels affected virus infectivity, low-level, non-lethal mutagenesis occurred both in the presence and absence of Vif (Sadler, 2010). *In silico* analysis also supports a role for APOBEC3G in HIV-1 evolution, through simulation of multiple rounds of infection using prediction models to map mutations based on calculated APOBEC3G target probabilities. These models also assessed G-to-A nonsynonymous and synonymous mutations in present-day HIV compared to ancestral genomes, and showed a higher proportion of nonsynonymous mutations compared to random controls, implying some selection for APOBEC3G-mediated mutations (Jern, 2009).

In vitro experiments have also directly tested the ability for APOBEC3G to cause drug resistance through viral mutation. The use of the nucleoside analog reverse transcriptase inhibitor 3TC often leads to drug resistance in patients based on a single amino acid substitution in the catalytic site of RT. This mutation is at position 184, changing the normally occurring methionine in the conserved YMDD motif to isoleucine, valine, or threonine (Boucher, 1993). Monotherapy with this drug leads to drug resistance within a few weeks in HIV-infected patients (Sarafianos, 1999). The M184I mutation is specifically caused by mutation from AUG to AUA, which also lies within the preferred editing context of APOBEC3G (5'-GG-3'). Wild-type virus grown *in vitro* in the presence of IC_{90} levels of 3TC rapidly developed resistance via the M184I mutation in cells engineered to express APOBEC3. The development of resistance did not occur in cells lacking APOBEC3 during the timeframe assessed, nor did the mutation become fixed in viral populations not grown with 3TC (Kim, 2010).

Natural variations in the Vif protein may largely affect APOBEC3 activity. A partially active or inactive Vif could profoundly impact viral sequence evolution within virally infected

individuals. Viruses isolated from HIV-infected patient samples show a large diversity in the *vif* coding region, with variation of up to 8% from the consensus sequence (Simon, 2005). When tested *in vitro*, these naturally occurring variants had variable activity against APOBEC3, with selected mutants failing to effectively block these enzymes. For example, Vif harboring a K22E mutation fails to restore HIV-1 infectivity in the presence of APOBEC3G (Simon, 2005).

Virus containing this mutant Vif was propagated in PBMCs *in vitro* for two weeks, and resulting viruses had over thirty times more mutations compared to wild-type viruses undergoing similar propagation. These mutations were characterized by a strong bias toward G-to-A mutations in the APOBEC3G editing context (Mulder, 2008). Over forty percent of recovered viral sequences also harbored the G-to-A mediated M184I 3TC-resistance mutation, compared to none in the wild-type viral population. However, all of the viruses harboring this drug resistance mutation were contained within replication-defective proviruses, based on premature stop codons also induced by APOBEC3. Infection with a viral quasi-species containing these resistance-containing replication-defective viruses together with wild-type sequences eventually led to drug-resistant, replication-competent viruses produced through viral recombination *in vitro* (Mulder, 2008). These experiments provide essential evidence that APOBEC3 can result in drug resistance, even in the complete absence of selective pressure by existing drug.

These experiments are corroborated by analysis of patient viral samples. A cohort of HIV-1 infected individuals failing antiretroviral therapy was compared to a cohort of treatment naive infected patients. The patients failing therapy typically have a higher rate of drug resistant viruses (Sethi, 2003). Virus was isolated from plasma, and protease, reverse transcriptase and Vif genes were amplified and sequenced. Mutation K22H in Vif was more frequently seen in viruses isolated from the cohort failing treatment. Of viruses harboring this Vif mutation, 72% had at least two drug-resistance-associated mutations in a GA or GG dinucleotide context, compared to 42% of all viruses with a wild-type K22 Vif (Fourati, 2010). Thus, a partially effective Vif is associated with APOBEC3-mediated drug resistance mutations *in vivo*.

Chapter 1.7: Specific Aims and Significance of Dissertation

APOBEC3 is an important antiviral factor, which specifically targets retroviruses. Since its discovery, APOBEC3 has been shown to restrict many different viruses, including important human pathogens such as HIV-1. The overall aim of this study is to understand how APOBEC3 functions *in vivo*. We used a murine model to address three main questions that are poorly understood to date.

Specific Aim 1: Determine the mechanism by which murine APOBEC3 restricts MMTV *in vivo.* We hypothesize that murine APOBEC3 is unable to create G-to-A mutations in MMTV reverse transcription products, but instead inhibits reverse transcription via cytidine deaminase-independent mechanisms. We used biochemical approaches to test the ability of endogenous murine APOBEC3 to catalyze cytidine deamination, and endogenous reverse transcription assays to assess mutagenesis and inhibition of MMTV reverse transcription products.

Specific Aim 2: Establish important differences between murine APOBEC3 alleles that lead to differences in viral restriction. We hypothesize that APOBEC3 expression level differences in mice expressing $APOBEC3^{BALB/c}$ or $APOBEC3^{C57BL/6}$ contribute to the associated difference in level of viral restriction. Biochemical approaches were used to determine potential target sequence specificity differences, and viruses isolated from heterozygous APOBEC3+/- mice utilized for assessing a dose effect of APOBEC3 on inhibition of reverse transcription.

Specific Aim 3: Assess the effect of APOBEC3 on viral evolution and drug resistance *in vivo*. We hypothesize that the mutagenic effect of APOBEC3 will increase the rate of nucleotide mutation, thus decreasing time to development of drug resistance. We utilized two systems to address this: MMTV-infected APOBEC3+/+ mice, as well as MLV-infected transgenic APOBEC3G-expressing mice. We measured viral replication over time in AZT-treated vs. untreated mice, and sequenced reverse transcriptase from MLV viruses at multiple time points post-infection. These sequences were analyzed for nucleotide and amino acid mutations.

Chapter 2: APOBEC3 inhibition of MMTV infection: the role of cytidine deamination versus inhibition of reverse transcription

Chapter 2.1: Abstract

The apolipoprotein B editing complex 3 (APOBEC3) family of proteins is a group of intrinsic anti-viral factors active against a number of retroviral pathogens including HIV in humans and mouse mammary tumor virus (MMTV) in mice. APOBEC3 restricts its viral targets through cytidine deamination of viral DNA during reverse transcription or via deaminase-independent means. Here, we used virions from the mammary tissue of MMTV-infected inbred wild type mice with different allelic APOBEC3 variants ($APOBEC3^{BALB}$ and $APOBEC3^{C57BL/6}$) and knockout mice to determine whether cytidine deamination was important for APOBEC3's anti-MMTV activity. First, using anti-murine APOBEC3 antisera, we showed that both APOBEC3 allelic variants are packaged into the cores of milk-borne virions produced *in vivo*. Next, using an *in vitro* deamination assay, we determined that virion-packaged APOBEC3 retains its deamination activity and that allelic differences in APOBEC3 affect the sequence specificity. In spite of this *in vitro* activity, cytidine deamination by virion-packaged APOBEC3 of MMTV early reverse transcription DNA occurred only at low levels. Instead, the major means by which *in vivo* virion-packaged APOBEC3 restricted virus was through inhibition of early reverse transcription, both in cell free virions and *in vitro* infection assays. Moreover, the different wild type alleles varied in their ability to inhibit this step. Our data suggest that while APOBEC3-mediated cytidine deamination of MMTV may occur, it is not the major means by which APOBEC3 restricts MMTV infection *in vivo*. This may reflect the long term co-existence of MMTV and APOBEC3 in mice.

Chapter 2.2: Introduction

Organisms adapt to infectious agents by developing protective responses and conversely, infectious agents develop adaptive countermeasures to these responses. Retroviruses are major causes of disease such as cancer and acquired immunodeficiency in animals and humans and are likely one of the infectious agents that put selective pressure on host evolution. Because of the frequent encounter of vertebrates with retroviruses, selection for host anti-viral defense systems is likely and indeed, various host restriction factors have been identified (Emerman and Malik, 2010). These include the TRIM proteins, Bst2/tetherin and APOBEC3 proteins (Malim and Bieniasz, 2012; see Chapter 1.6). Most of these anti-viral intrinsic restriction factors were identified through the discovery of viral gene products that counteract their actions.

Members of the *APOBEC3* gene family encode DNA and RNA editing enzymes. APOBEC3 proteins inhibit retroviral infection in target cells by deaminating deoxycytidine residues on the DNA minus strand following reverse transcription, inducing hypermutation in newly synthesized retroviral DNA. The various human APOBEC3 proteins have different preferred target sequences; for example, APOBEC3G-mediated deamination occurs more frequently at CCC residues, whereas APOBEC3F preferentially deaminates TCC sites (Bishop, 2004; Beale, 2004; Langlois, 2005). APOBEC3 proteins also inhibit replication by undefined cytidine deaminase (CDA)-independent mechanisms (Newman, 2005). A number of *in vitro* studies have suggested that APOBEC3 proteins in particles can inhibit binding of the $tRNA^{Lys3}$ primer to the viral RNA, RT-mediated elongation and accumulation of HIV-1 reverse transcription products or integration into the host genome, at least in tissue culture cells (Bishop, 2006; Iwatani, 2007; Li, 2007; Guo, 2007; Holmes, 2007a; Holmes, 2007b; Luo, 2007; Bishop, 2008). Although several studies have argued that CDA-independent inhibition is an artifact of APOBEC3 over-expression (Schumacher, 2008; Miyagi, 2007), other studies have argued that it is biologically relevant (Holmes, 2007; Bishop, 2008; Gillick, 2013) and recent work with mouse retroviruses indicates that CDA-independent inhibition occurs *in vivo* (see Chapter 2.4).

Clearly, viruses that persist in their hosts cannot be totally restricted by anti-viral host mechanisms. In support of this, we have shown that mouse APOBEC3 contributes to resistance to MMTV infection but does not totally restrict infection. MMTV is an endemic milk-borne retrovirus that entered mice ~10 - 20 million years ago and causes mammary carcinomas in female mice (Nandi and McGrath, 1973). Using APOBEC3 knockout mice, we provided the first demonstration that APOBEC3 proteins function *in vivo* by showing that APOBEC3-/- mice were more susceptible to MMTV infection compared to their wild type littermates; virus spread and tumorigenesis were more rapid and extensive in the knockout mice (Okeoma, 2007). Moreover, we found no evidence of cytidine deamination of the MMTV genome in wild type mice. Similarly, several studies have shown that APOBEC3 inhibits Moloney (M-MLV) and Friend (F-MLV) murine leukemia virus infection *in vivo* in the absence of mutations suggestive of cytidine deamination (Rulli, 2007; Browne and Littman, 2008; Takeda, 2008; Low, 2009). However, mouse APOBEC3 retains deaminase activity on HIV substrates, with a preference for TCC targets (Bishop, 2004) and has also been shown to have low activity on TTC residues in the endogenous murine leukemia virus AKV (AKV-MLV) (Langlois, 2009).

Interestingly, there are allelic differences in mouse *APOBEC3* among different inbred mouse strains that encode proteins with the ability to restrict MMTV, as well as F-MLV, to different extents (Takeda, 2008; Okeoma, 2009). In particular, C57BL/6 and BALB/c mice express different APOBEC3 variants, with 15 polymorphic amino acids in the proteins encoded in their genomes (Figure 1.6). Moreover, the BALB/c allele ($APOBEC3^{BALB}$) encodes a longer transcript and protein than that found in C57BL/6 mice ($APOBEC3^{C57BL/6}$) because it includes a fifth exon that is spliced out in the latter. Finally, $APOBEC3^{BALB}$ transcript and protein levels in BALB/c mice are significantly lower in various tissues compared to $APOBEC3^{C57BL/6}$ in C57BL/6 mice (Takeda, 2008; Okeoma, 2009; Miyazawa, 2008; Sanville, 2010; Li, 2012). The $\textit{APOBEC3}^{C57BL/6}$ allele encodes a more effective *in vivo* restriction factor of murine retroviruses. It is currently unclear as to whether the differential ability to restrict MLV or MMTV infection *in vivo*

is due to the coding region changes that affect its anti-viral activity or to expression level differences (Okeoma, 2009; Miyazawa, 2008; Li, 2012).

Here we show that APOBEC3 packaged into virions retains its deamination activity and deaminates MMTV reverse transcripts at a low level. Interestingly, the APOBEC3$^{C57BL/6}$ and APOBEC3^{BALB} variants packaged in MMTV virions have altered substrate preferences for deamination. In spite of the deamination activity, however, the major means by mouse APOBEC3 inhibits MMTV is by blocking early reverse transcription.

Chapter 2.3: Materials and Methods

Chapter 2.3.1 Plasmids, Cells, and Viruses

Plasmids expressing flag-tagged APOBEC3 (pFLAG-CMV vector) cloned from primary DNA of C57BL/6 and BALB/c mice and HA-tagged catalytically inactive APOBEC3$^{C57BL/6}$ (E73Q/E253Q; pCMVpA vector) were acquired as a gift from Masaaki Miyazawa. Primary cDNA samples were prepared from spleens, and full-length APOBEC3 was amplified using the olionucleotide primers 5'-GGGGTACCGCCGCCACCATGGGACCATTCTGTCTGGGATGCAGCCATCGC-3' AND 5'GGTCTAGACATCGGGGGTCCAAGCTGTAGGTTTCC-3' (Takeda, 2008). The pFLAG-CMV2 vector (Sigma-Aldrich Corp., St.. Louis, MO.) was digested with SalI and EcoRI enzymes, and the APOBEC3 cDNA was ligated into the vector. The catalytically inactive vector was created through site directed mutagenesis using the pFLAG-CMV2-APOBEC3$^{C57BL/6}$ plasmid as the template (Takeda, 2008).

293 cells stably expressing mouse transferrin receptor 1 were maintained in DMEM supplemented with 10% FBS, 100 U/ml penicillin, 100 ug/ml streptomycin and 100 ug/ml geneticin. These cells were originally produced by cotransfecting 293 cells with a plasmid containing the mouse transferrin receptor coding region in a CMV promoter-containing vector, as well as a plasmid containing a neomycin resistance gene under SV40 promoter control (Zhang, 2003). Cells were selected in G418 (100 ug/mL) and FACS sorted for transferrin receptor 1

expression. These cells were then stably transfected with APOBEC3 plasmids (described above) and a plasmid containing a puromycin resistance gene, followed by selection in puromycin (1 ug/mL). Expression was verified by SDS-PAGE and Western Blotting using anti-murine APOBEC3 (Okeoma, 2010), anti-Flag (Cell Signaling) and anti-HA (Abcam) antibodies.

MMTV virions were isolated from mammary tumors or breast milk of MMTV-infected mice. All steps of purification were performed on ice. First, tumors were cut into small pieces and homogenized via Dounce homogenizer in buffer with 10 mM Tris pH 8.0, 0.1M NaCl, and 1 mM EDTA. The homogenate was centrifuged at 66g for 5 minutes at 4°C, and the supernatant removed to a fresh 50 mL centrifuge tube (pellet discarded). The supernatant was further centrifuged at 12,400g for 10 minutes at 4°C. Further purification of the supernatant was achieved via ultracentrifugation at 21,000 rpm for 1 hour at 4°C. The pellet was dissolved in 2 mL PBS, which was layered over 0-60% sucrose gradient (sucrose dissolved in PBS), or over a 30% sucrose cushion, and then ultracentrifuged for 1 hour at 32,000rpm at 4°C. Pellets were resuspended in PBS (100 ul per gram of tumor tissue), and stored in aliquots at -80°C. Virions from breast milk were isolated with minor modifications. Stomachs were removed from 1-3 day old mouse pups of APOBEC3+/+ or APOBEC3-/- MMTV RIII-infected mothers. At least ten stomachs of the same genotype were combined, and purified as above starting with Dounce homogenization. Milk fat was skimmed from the top of the supernatant after the first two spins and discarded.

RNA was purified from these virions using the QIAGEN RNeasy Mini Kit (QIAGEN, Germantown, MD). Eight uL of virus was combined with 95 ul of PBS and 95 uL of RNA*later* RNA Stabilization Reagent (QIAGEN), then instructions per kit manual followed. RNA was eluted into 30 uL of RNase free water, and reverse transcribed using SuperScript III First-Strand Synthesis SuperMix, based on a modified M-MLV RT (Invitrogen Corp., Carlsbad, CA). Viral cDNA was then measured by quantitative PCR for normalization, using primers specific for the

MMTV RIII LTR (Forward: 5'-CGTGAAAGACTCGCCAGAGCTA-3'; Reverse: 5'-GAAGATCTTCAAGGGCAATGCCTTAA-3').

Viral cores were isolated by ultracentrifugation of purified virions through 10% sucrose with 5% Triton-X on a 30% sucrose cushion, at 32,000rpm for 1 hour at 4°C. Pellets were resuspended directly in 4x SDS-PAGE protein loading buffer (200 mM Tris pH 6.8, 400 mM DTT, 8% SDS, 0.4% bromophenol blue, 40% glycerol).

Chapter 2.3.2 Mice

MMTV(RIII)-infected APOBEC-/- mice and APOBEC+/+ mice were used for virus and organ harvesting. APOBEC3-/- mice were generated by gene trapping of intron 4 of the APOBEC3 gene (Okeoma, 2007). This trapping would theoretically create a fusion protein of the first four APOBEC3 exons and B-galactosidase, but no fusion protein was detected in tissues from these mice (Okeoma, 2007). APOBEC-/- and APOBEC+/+ mice were crossed onto the C57BL/6 background and thus wild type mice encode the APOBEC3$^{C57BL/6}$ allele. For genetic consistency, experiments used the two matched strains; however, wild type animals are referred to as C57BL/6 in the text to reflect their APOBEC3 allele. Mice were housed according to the policies of the Institutional Animal Care and Use Committee of the University of Pennsylvania. Heterozygous F1 mice were created by crossing APOBEC-/- males with MMTV(RIII)-infected C57BL/6 or BALB/c females. The MMTV-infected female F1 mice were then bred and virus was isolated from milk collected from their 1-3 day old pups, as described above.

Chapter 2.3.3 *In vitro* deamination

FAM-labelled 50mer oligonucleotide substrates with a TTC target site were incubated with serial (ten-fold) dilutions of tumor-derived virions. Virus was first lysed in buffer containing 10 ug/ml RNase A, 50 mM Tris pH 8.0, 40 mM KCl, 50 mM NaCl, 5 mM EDTA and 0.1% TritonX-100. Virus was then incubated with 1x UDG Buffer (NEB), 2.5 u UDG (NEB), 5 mM EDTA, and 100 nM substrate at 30°C for 4 hours. Samples were heated to 95°C for 20 minutes before

addition of formamide load buffer and 0.1 M NaOH to cleave abasic sites. Separation of products on a 20% acrylamide/TBE/Urea gel was imaged via direct fluorescence on a Typhoon 9410 Molecular Imager. Control deamination reactions were carried out with purified mouse APOBEC1 or APOBEC3, received from Rahul Kohli. These proteins are produced in and purified from BL21(DE3)-Star E. coli (Novagen), from pET41 vectors encoding the $APOBEC3^{BALB/C}$ gene downstream of an N-terminal maltose-binding protein (Nabel, 2012).

For targeting preference experiments, 60mer oligonucleotide substrates were labeled with ddUTP-FAM at their 3' end using Terminal dNTP Tranferase (NEB). Ten units of terminal transferase is combined with 0.5 mM ddUTP-FAM, 200 ng of substrate DNA, and 1.25 mM CoCl2, and incubated for 30 minutes at 37°C. The reaction is stopped by heating to 70°C for 10 minutes, and labeled substrate was purified using the QIAquick Nucleotide Removal Kit (QIAGEN Corp.). Substrate sequences are shown in Table 2.1. Determination of targeting preferences was then carried out by a two-step calculation (Kohli, 2009). Percent product formation was calculated for each sample, and values averaged for all sequences that contain the same nucleotide at either the -1 or -2 position. Then, the percent deamination of each nucleotide at the -1 or -2 position was calculated compared to the other nucleotide variants, for example for A at the -2, (ATC/(AmCC + AAC + AGC)) x 100 (Kohli, 2009).

Table 2.1 Oligomer substrate sequences for *in vitro* deamination assay.
Target substrates used for cytidine deamination targeting assay (1-17) (Fig. 2.2) contained 60 nucleotides and differed solely in the two nucleotides directly upstream of the single cytosine. Fluorescent ddUTP was added after synthesis. Substrates used for initial *in vitro* deamination assay (Fig. 2.2) contained 50 nucleotides and were fluorescently labeled upon synthesis. F: Enzo Life Sciences ddUTP-FAM

Target	Sequence
Template	GTGGTGTGGT TTGATGGTAT GXXCTTGGTG TGGTTGATAG TTGTGATGAA TTGTTTTATTF
AAC	GTGGTGTGGT TTGATGGTAT GAACTTGGTG TGGTTGATAG TTGTGATGAA TTGTTTTATTF
AmCC	GTGGTGTGGT TTGATGGTAT GAmCCTTGGT GTGGTTGATA GTTGTGATGA ATTGTTTTATTF
AGC	GTGGTGTGGT TTGATGGTAT GAGCTTGGTG TGGTTGATAG TTGTGATGAA TTGTTTTATTF
ATC	GTGGTGTGGT TTGATGGTAT GATCTTGGTG TGGTTGATAG TTGTGATGAA TTGTTTTATTF
mCAC	GTGGTGTGGT TTGATGGTAT GmCACTTGGT GTGGTTGATA GTTGTGATGA ATTGTTTTATTF
mCmCC	GTGGTGTGGT TTGATGGTAT GmCmCCTTGG TGTGGTTGAT AGTTGTGATG AATTGTTTTATTF
mCGC	GTGGTGTGGT TTGATGGTAT GmCGCTTGGT GTGGTTGATA GTTGTGATGA ATTGTTTTATTF
mCTC	GTGGTGTGGT TTGATGGTAT GmCTCTTGGT GTGGTTGATA GTTGTGATGA ATTGTTTTATTF
GAC	GTGGTGTGGT TTGATGGTAT GGACTTGGTG TGGTTGATAG TTGTGATGAA TTGTTTTATTF
GmCC	GTGGTGTGGT TTGATGGTAT GGmCCTTGGT GTGGTTGATA GTTGTGATGA ATTGTTTTATTF
GGC	GTGGTGTGGT TTGATGGTAT GGGCTTGGTG TGGTTGATAG TTGTGATGAA TTGTTTTATTF
GTC	GTGGTGTGGT TTGATGGTAT GGTCTTGGTG TGGTTGATAG TTGTGATGAA TTGTTTTATTF
TAC	GTGGTGTGGT TTGATGGTAT GTACTTGGTG TGGTTGATAG TTGTGATGAA TTGTTTTATTF
TmCC	GTGGTGTGGT TTGATGGTAT GTmCCTTGGT GTGGTTGATA GTTGTGATGA ATTGTTTTATTF
TGC	GTGGTGTGGT TTGATGGTAT GTGCTTGGTG TGGTTGATAG TTGTGATGAA TTGTTTTATTF
TTC	GTGGTGTGGT TTGATGGTAT GTTCTTGGTG TGGTTGATAG TTGTGATGAA TTGTTTTATTF
TTU	GTGGTGTGGT TTGATGGTAT GTTUTTGGTG TGGTTGATAG TTGTGATGAA TTGTTTTATTF
S50TTC	ATTATTATTA TTTTCATTTA TTTATTTATT TATGGTGTTT GGTGTGGTT/36-FAM/
S50TTU	ATTATTATTA TTTTUATTTA TTTATTTATT TATGGTGTTT GGTGTGGTT/36-FAM/

Chapter 2.3.4: *In vitro* reverse transcription

Virions isolated as described above were normalized to equal RNA levels, and incubated with PBS, 2.5 mM $MgCl_2$, 0.01% NP-40 and 1 mM dNTPs at 37°C. Reverse transcription products were taken from the reaction in aliquots at the specified time points, added to equal volumes of PBS with 40 ug/mL salmon sperm DNA (Stratagene) and frozen at -80°C (Bishop, 2008). Control reactions were incubated without dNTPs and in the presence of 2.5 mM EDTA. DNA was isolated from the samples using the DNeasy Blood & Tissue Kit (QIAGEN Corp.).

Intracellular reverse transcription products were produced by infecting 293T-mTrf1 cells with tumor-derived virions in the presence of 8 ug/ml polybrene. At the designated time points, media and virus were removed, cells were washed with PBS and cells were removed from the plate. Total DNA was purified using the DNeasy Blood & Tissue Kit (QIAGEN Corp.). To measure the effect of intracellular APOBEC3, the same protocol was followed using the 293T-mTrf1 cells stably transfected with $APOBEC3^{C57BL/6}$, $APOBEC3^{BALB/c}$ or mutant (E73Q/E253Q) expression vectors. Viral DNA was then quantified (see Chapter 2.3.5) or assessed for mutagenesis (see Chapter 2.3.6).

Chapter 2.3.5: Polymerase Chain Reaction

PCR reactions to amplify edited reverse transcription products were performed using a gradient cycler. Equal volumes of viral DNA were added to GoTaq DNA Polymerase (Promega), and either a 200 bp region of *env* or a 400 bp region of the 3' LTR was amplified with following primers. Env_Forward: 5'-GCCTCGAGCTAAGTAACACAG-3', Env_Reverse: 5'-TCAGGGGCCAATACAAAACTGGT-3', LTR_Forward: 5'-CGTGAAAGACTCGCCAGAGCTA-3', LTR_Reverse: 5'-GAAGATCTTCAAGGGCAATGCCTTAA-3'. A temperature gradient of 81-95°C was applied during an initial 5 min denaturing step, as well as the 1 min denaturing step of each subsequent cycle (35 total), followed by 1 min at 55°C and 1 min at 72°C. DNA was purified from a 1% agarose gel using QIAGEN QiaQuick Gel Extraction Kit (QIAGEN Corp.).

Quantitative PCR for strong stop DNA was performed in triplicate, using primers StrongStop_Forward: 5'-CGTGTGTTTGTGTCTGTTCG-3' and StrongStop_Reverse: 5'-GACCCTCTGGAAAGTGAAAGTCAAGG-3'. These primers amplify 90bp product from the MMTV RIII U5 region. Equal volumes of viral DNA was combined with Power SYBR Green PCR Master mix and amplified under standard conditions on an ABI Prism Model 7900HT.

Chapter 2.3.6: Sequencing and Statistical Analysis

Following PCR and DNA isolation, the *env* and LTR fragments were cloned into pCR2.1-TOPO vector, as directed (Invitrogen, Inc.). The cloning product was used to transform OneShot TOP10 chemically competent *E. coli* (Invitrogen, Inc.), and bacteria plated on LB plates containing 100 ug/mL ampicillin for selection. Individual colonies were picked and grown in LB + ampicillin culture shaking overnight at 37°C. Plasmid DNA was isolated from cultured bacteria using QIAprep Spin MiniPrep Kit (QIAGEN Corp.). The MMTV fragment within the isolated plasmids were then sequenced using a BigDye Terminator v3.1 Cycle Sequencing Kit from Applied Biosystems. Sequences were aligned using the ClustalW program, and G to A mutations were annotated by Hypermut (www.hiv.lanl.gov/content/sequence/HYPERMUT/hypermut.html). Statistical analysis was performed using the GraphPad/PRIZM software.

Chapter 2.4: Results

Chapter 2.4.1: Both murine APOBEC3 variants are packaged into virion cores

We first considered the possibility that APOBEC3 allelic variants could differ in the levels of protein incorporated into virus particles or in their localization within virions. Most APOBEC3 proteins are incorporated into virus particles during virion assembly in infected producer cells. Indeed, APOBEC3G binds to the HIV NC (Zheng, 2004) and mouse APOBEC3 to the MMTV NC (Okeoma, 2007), as well as viral RNA, suggesting that upon viral maturation, the restriction factor is concentrated within the viral core. However, APOBEC3A is packaged into virions, yet remains outside the viral core, preventing it from restricting HIV infection (Aguiar, 2008). If APOBEC3 was

excluded from the MMTV cores, this could result in an inability to access the reverse transcription complex and cause cytidine deamination. Moreover, the changes in protein sequence could differentially alter packaging of the full-length APOBEC3^{BALB}, thereby diminishing its ability to restrict infection.

To test this, we isolated virions from the mammary glands of MMTV-infected C57BL/6, APOBEC3 knockout (APOBEC3-/-), and BALB/c mice. We used reverse transcribed real-time quantitative PCR primers specific to viral RNA to normalize virion levels in the three different virion preparations. In addition, Western blot analysis for both MMTV and APOBEC3 proteins were performed to confirm equal protein loading, and to determine the relative level of packaged APOBEC3. APOBEC3 was detected in virions isolated from C57BL/6 and BALB/c but not the APOBEC3-/- mice (Figure 2.1). The presence or absence of the fifth exon was visible by a 4 kDa size difference. BALB/c-produced virions packaged primarily the full-length version of the protein, but also packaged a small quantity of the shorter version. C57BL/6-isolated virions solely packaged the isotype lacking exon 5 and contained a significantly higher level of packaged protein as compared to BALB/c virions. The level of virion-packaged protein reflected the intracellular levels of APOBEC3 mRNA in C57BL/6 and BALB/c mouse tissues, as previously described (Okeoma, 2010).

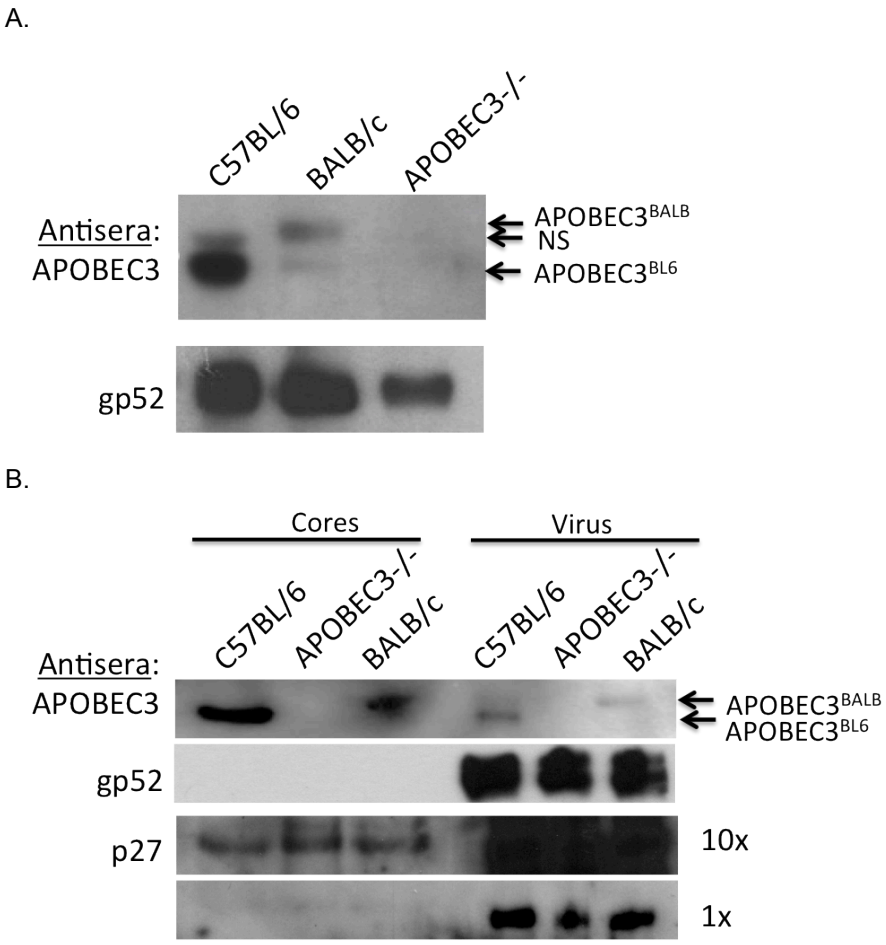


Figure 2.1 APOBEC3 is packaged within MMTV viral cores *in vivo*. (A) Virions were isolated from the mammary tissue of MMTV-infected C57BL/6, BALB, and APOBEC3-/- mice. Virus levels were normalized to viral RNA via RT-qPCR and equal amounts subjected to Western blot analysis using rabbit anti-mouse APOBEC3 antisera. The immunoblot was stripped and reprobed with goat anti-Env (gp52) antisera. C57BL/6 mice solely express and package a shorter form of the protein lacking a fifth exon. BALB mice express predominantly a protein including all exons (4 kD larger), but also express a small amount of protein lacking this exon. The anti-APOBEC3 antisera also detects a cross-reacting protein (non-specific, NS) of approximately 50kD that is present in some but not all virion preparations, as previously reported (Okeoma, 2010). (B) Core fractions and whole virus were assessed for APOBEC3 as described above, and the immunoblot was serially stripped and reprobed for MMTV Env and p27 as controls for removal of envelope and retention of viral cores. The p27 blot was exposed for five seconds and one minute to visualize protein in both cores and whole virus.

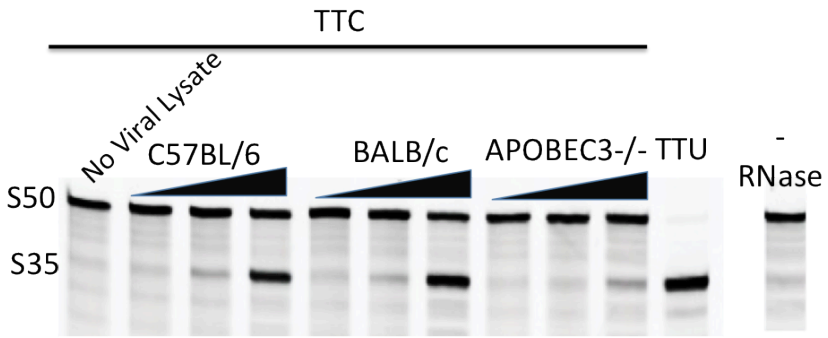
We next isolated viral cores by centrifugation of purified virus through 10% sucrose containing 5% Triton-X over 30% sucrose cushions. The detergent treatment removes the viral envelope and contaminating cellular membrane compartments, including exosomes, and allows viral cores to pellet. Purified viral cores were then used for Western blotting. Both the APOBEC3$^{C57BL/6}$ and APOBEC3^{BALB} were incorporated within viral cores. Importantly, there was substantially more APOBEC3$^{C57BL/6}$ than APOBEC3^{BALB} in cores (Figure 2.1), suggesting that the overall higher levels of APOBEC3$^{C57BL/6}$ in virions are not associated with mislocalized APOBEC3.

Chapter 2.4.2: Allelic differences in substrate specificity

A second possible explanation for the lack of cytidine deamination was that the APOBEC3 packaged into MMTV virions is rendered enzymatically inactive, perhaps by a viral protein. To determine if MMTV-packaged APOBEC3 retains its deaminase activity, we isolated virions from the mammary glands of MMTV-infected C57BL/6, APOBEC3 knockout, and BALB/c mice and examined the purified viral lysates for *in vitro* deaminase activity against a fluorescently tagged oligonucleotide substrate bearing the reported mouse APOBEC3 consensus target sequence (5'TTC3') (Langlois, 2009). Although C57BL/6-isolated virions packaged substantially more APOBEC3 than BALB/c virions (Figure 2.1), the protein packaged in BALB/c virions had equivalent activity to that packaged in BL/6 virions (Figure 2.2). Densitometry analysis measuring packaged APOBEC3 protein (Figure 2.1) and deamination on this substrate (Figure 2.2) revealed a 1.9 fold increase in activity of APOBEC3$^{BALB/c}$ compared to APOBEC3$^{C57BL/6}$ when normalized to protein content. In contrast, very little deamination activity was detected in particles isolated from knockout mice: the small amount of deamination (<5% of virions containing APOBEC3$^{C57BL/6}$ or APOBEC3$^{BALB/c}$) may be attributed to cellular contamination with other known mouse cytidine deaminases such as APOBEC1 (Petit, 2009). The deaminase activity required RNaseA treatment, demonstrating that like human APOBEC3G, mouse APOBEC3 is not enzymatically active when bound to RNA within the virion (Figure 2.2). These data suggested that both

APOBEC3$^{C57BL/6}$ and APOBEC3$^{BALB/c}$ are functional deaminases that retain enzymatic activity in virions.

A.



B.

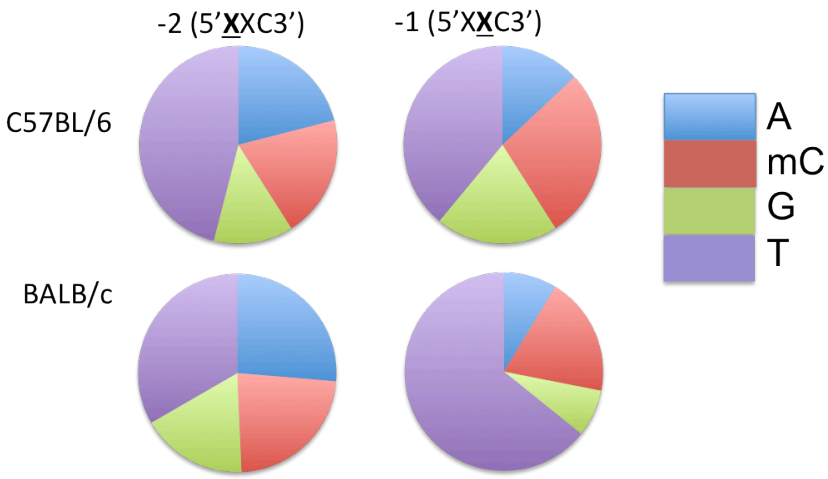


Figure 2.2 *In vivo* packaged APOBEC3 variants are enzymatically active and target different substrates. (A) Ten-fold increasing concentrations of virions isolated from C57BL/6, BALB, and APOBEC3-/- mice were lysed with 0.1% Triton-X in the presence or absence of RNase A and incubated with a 50mer single-stranded oligonucleotide (S50) containing a single cytosine and uracil DNA glycosylase, followed by cleavage with NaOH (see substrate sequences, Table 2.1). An oligonucleotide containing a uracil confirmed complete cleavage of uracilated products and acted as a size standard (uncleaved substrate at S50, cleavage product at S35). C57BL/6 virions were also lysed in the absence of RNaseA and subjected to otherwise identical conditions. (B) Representation of the two murine APOBEC3 variants. *= active site residue, vertical lines= location of amino acid differences, green= differences within substrate recognition loop, dotted line= deletion of exon 5. Numbering is relative to APOBEC3^{BALB} variant. (C) Virions isolated from C57BL/6 and BALB mice were incubated with 16 different substrates containing a single cytosine with variations of nucleotides in the -1 and -2 position to include each possible combination of A, G, T, or mC. Product to substrate ratios were calculated for 3 to 4 replicate experiments. Blue= A, Red= mC, Green = G, Purple= T. T-tests were used to verify significant differences between the APOBEC variants preference for G (higher preference by APOBEC3^{BALB},

$p<0.05$) and T (higher preference by $APOBEC3^{C57BL/6}$, $p=0.05$) at the -2 position and T at the -1 position (higher preference by $APOBEC3^{BALB}$, $p<0.05$).

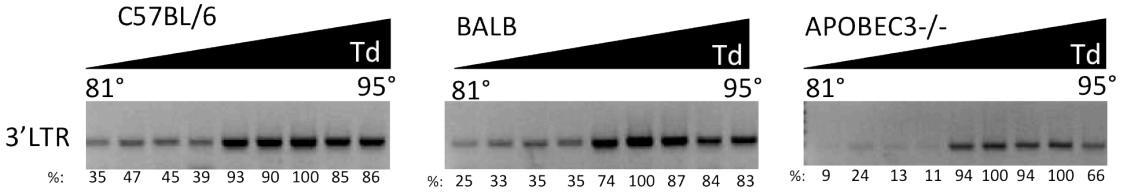
Different cellular cytidine deaminases can be distinguished from one another by their preferred editing contexts, based on the nucleotides surrounding the targeted cytosine. This preference has been mapped to a recognition loop of 9-11 amino acids that interacts with nucleotides upstream of the target site (Kohli, 2009). The two APOBEC3 isoforms differ in the presence of exon 5 and in addition have several amino acid differences in their other exons (Figure 2.2), including three polymorphisms in the putative substrate recognition loop which could influence the editing context. We thus examined if the nucleotide context of cytidine deaminase activity differed between each isoform. The *in vitro* deaminase assay was carried out with fluorescently labeled substrates bearing each of the sixteen possible 5'-XXC-3' combinations. To make the assay specific for deamination at the target cytosine, 5-methylcytosine (mC) was used in place of cytosine at the two upstream positions (Nabel, 2012). The assay was performed with APOBEC3 isolated from virions as described above, and substrate to product ratios measured (Figure 2.2). While APOBEC3$^{C57BL/6}$ exhibited a higher selectivity at the -2 position for the canonical TTC substrate, APOBEC3$^{BALB/c}$ had stronger preferences at the -1 position, with ATC as a favored substrate. These results imply a difference in activity that is likely due to a change in substrate recognition based on the allelic amino acid differences, not simply based on different levels of packaged protein.

Chapter 2.4.3: APOBEC3 deamination of natural MMTV transcripts is low

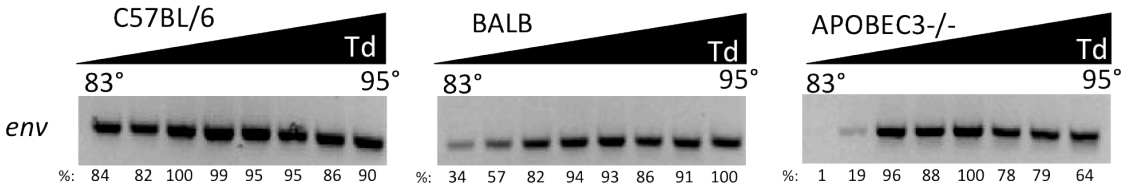
Although the results of the preceding sections demonstrate that APOBEC3 packaged into MMTV virions retained functional deaminase activity, we previously showed that integrated MMTV proviruses found in wild type mice showed no mutagenic hallmarks of cytidine deamination compared to knockout mice (Okeoma, 2007). However, uracil-containing reverse-transcribed DNA could be degraded in cells prior to integration into the host genome or prevented from integration (Russell, 2009), which would account for the absence of G to A mutations in integrated MMTV proviruses. We thus next tested whether *ex vivo* reverse-transcribed MMTV DNA made from viruses containing packaged APOBEC3 underwent cytidine deamination.

Purified virions from wild type, knockout and BALB/c mice were used in endogenous reverse transcription assays (EnRT) and DNA produced by these particles was subjected to differential DNA deamination 3DPCR using primers to two different regions of MMTV (3' long terminal repeat (LTR) and envelope (*env*) (Figure 2.3). PCR was performed using a gradient range of denaturing temperatures from 81-95°C. Amplification of products from all three types of virions was similar at melting temperatures from 85-95°C (3'LTR amplification shown in Figure 2.3). When the PCR was carried out at higher stringency (≤ 83°C melting temperature for 3' LTR, ≤ 85°C for *env*), reverse transcribed DNA from either BALB/c or BL/6 virions was more highly amplified than that from APOBEC3-/- virions, indicating a higher AT content in reverse transcribed DNA produced in the presence of either allele of APOBEC3.

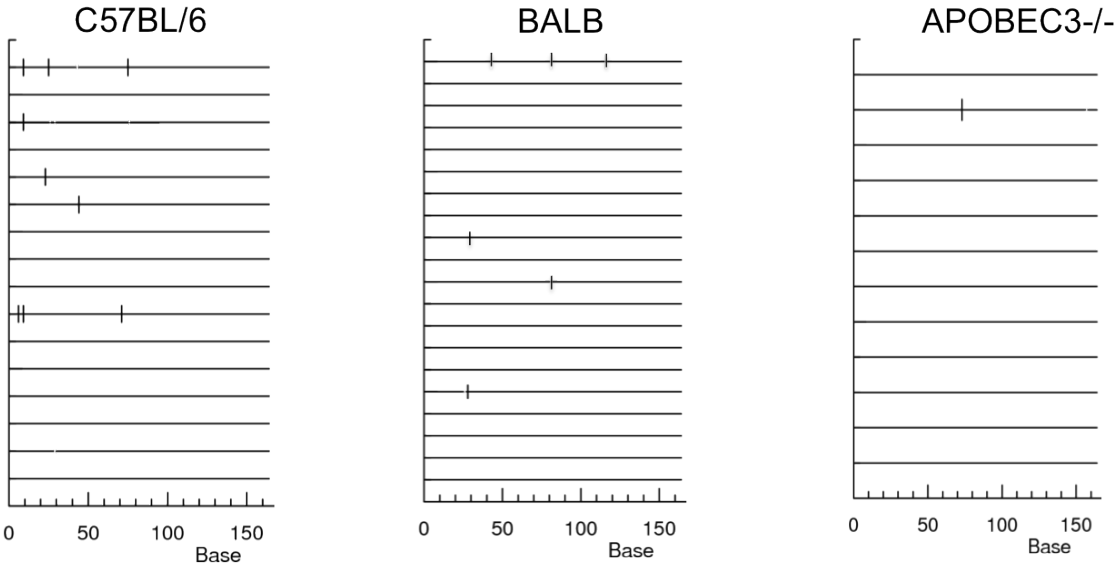
A.



B.



C.



D.

5'XXC3'

5'XXC3'	A	C	T	G
A		2	2	
C		1		
T	2	43	38	7
G		1		2

5'XXC3'

5'XXC3'	A	C	T	G
A	1	3	13	
C		1		
T	4	14	15	3
G				

5'XXC3'

5'XXC3'	A	C	T	G
A			1	
C			1	
T		3	3	
G				1

FIG. 2.3. Packaged APOBEC3 edits MMTV reverse transcription products. (A) 3DPCR amplification of MMTV viral DNA isolated from EnRT reactions using virions from C57BL/6, BALB, and APOBEC3-/- mice. A ~200 bp region from the 3'LTR was amplified using an increasing denaturing temperate gradient (Td) from 81° to 95°. Densitometry was performed, and density of each band normalized to 95° (labeled as 100%). (B) 3DPCR amplification of MMTV viral DNA isolated 6 hours after infection of 293T-mTfR1 cells with virions isolated from C57BL/6, BALB, and APOBEC-/- mice. A 480 bp region of *env* was amplified using an increasing denaturing gradient from 83° to 95°. (C) 3' LTR DNA isolated from low and high denaturation temperature PCR reactions shown in (A) were cloned and sequenced. These sequences were aligned and G to A mutations were annotated using HyperMut as compared to a reference MMTV RIII NCBI GenBank sequence (AF136898). A representative portion of these sequences is displayed. (D) All G to A mutations found in clonal sequences isolated from 3'LTR sequences (A) and *env* sequences (B) were analyzed for sequence context at the -1 and -2 nucleotide positions for the minus strand targeted cytosine.

To confirm these results in infection, we performed a similar experiment in 293T-mTfR1 cells (293T cells expressing the MMTV entry receptor transferrin receptor 1). The cells were infected with viruses isolated from the three strains of mice, and total cellular and viral DNA was isolated 6 hours after infection, at which time pre-integration reverse transcription products will predominate. This DNA was then amplified using 3DPCR. Similar to the *in vitro* reverse transcribed DNA, MMTV reverse transcription products produced in the presence of either APOBEC3$^{BALB/c}$ or APOBEC3$^{C57BL/6}$ was amplified using highly stringent (low denaturation temperature) conditions, whereas APOBEC3-/- reverse transcribed DNA was not amplified under these conditions (*env* amplification shown in Figure 2.3).

The PCR products generated in both cell-free and tissue culture infection conditions were cloned and sequenced. When the PCR was carried out at high stringency (95°C melting temperature), there were few G to A mutations in any of the viruses and there was no statistical difference in the rate of G to A mutations found in reverse transcription products of viral particles or in infected cells with or without packaged APOBEC3 (Table 2.2, $p=0.17$). However, low levels of cytidine deamination were detected in the APOBEC3$^{C57BL/6}$ and APOBEC3$^{BALB/c}$ reverse transcribed DNA produced in both the *in vitro* EnRT and the tissue culture infection, when amplified at low stringency (Figure 2.3; Table 2.2). This rate of deamination was statistically higher than that seen in products amplified at 95°C ($p<0.05$). Even in the less stringent conditions, which should enhance detection of G to A transversions, the overall rate of G to A

mutation (0.47% for APOBEC3$^{C57BL/6}$, 0.41% for APOBEC3$^{BALB/c}$) was significantly lower than that seen in viruses known to be disabled by APOBEC3 editing, such as HIV, which averages 7-8% G to A mutations *in vivo* (Armitage, 2012). There was no difference in the frequency of mutations seen in reverse transcription products isolated from APOBEC3$^{BALB/c}$- or APOBEC3$^{C57BL/6}$-containing particles (p=0.67). This further supports our hypothesis that hypermutation is not a significant source of viral restriction, as the more restrictive APOBEC3$^{C57BL/6}$ variant did not show higher rates of mutagenesis. Some G to A mutations were detected from DNA isolated from reactions using APOBEC3-/- viral particles at both low and high stringency PCR, but the very low number can be attributed to RT error (not statistically different from no G to A mutations; p=0.42).

Table 2.2 Mutation analysis of MMTV reverse transcription products. Sequences cloned from 3DPCR amplification products in (Figure 2.3) were aligned to their respective reference sequence. G to A versus all other mutations were recorded. Sequences from C57BL6-isolated viral DNA isolated using a low denaturing temperature contained significantly more mutations than APOBEC3-/- isolated DNA (p<0.05), but not significantly more mutations than BALB-isolated viral DNA (p=0.67). Deamination levels in viral DNA isolated from all three genotypes using a regular denaturing temperature were not statistically significantly different from one another.

Virus/Temperature (°C)	G to A	Other	# of Clones	Total (bp)	G to A Frequency
C57BL/6 95°	28	20	50	13236	0.21
C57BL/6 81-83°	73	24	44	15652	0.47
BALB/c 95°C	2	2	18	4050	0.05
BALB/c 81-83°	52	15	34	12690	0.41
APOBEC3-/- 95°	2	7	17	4726	0.04
APOBEC3-/- 81-83°	7	9	30	9720	0.07

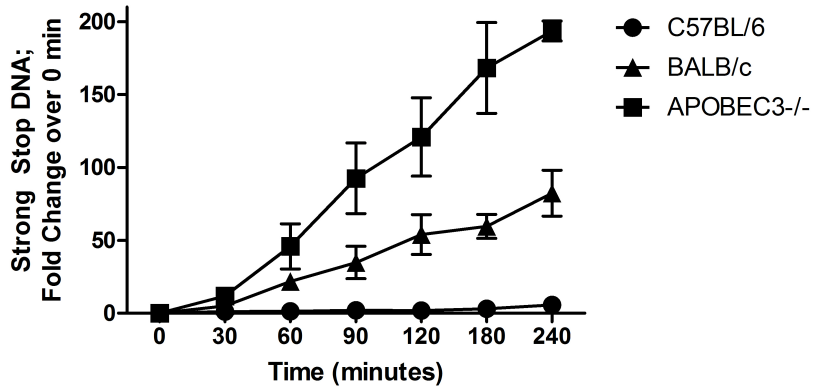
When the target nucleotide context was analyzed, a majority, but not all mutations were found in the mouse APOBEC3-preferred TXC context (92% for APOBEC3$^{C57BL/6}$ and 67% for APOBEC3$^{BALB/c}$) **(**Figure 2.3). As seen in the in vitro targeting assay APOBEC3$^{BALB/c}$ had a stronger preference for ATC, with 24% of all mutations falling within this context, compared to 2% for APOBEC3$^{C57BL/6}$. APOBEC3$^{C57BL/6}$ maintained a stronger preference that APOBEC3$^{BALB/c}$ for

T at the -2 position, further supporting our *in vitro* data (Figure 2.2). Within the *env* and LTR fragments sequenced, 22.9% of all cytosines were found in the preferred TXC context (2.6% TGC, 3.9% TAC, 9.2% TTC, 7.2% TCC). The most highly favored TTC (APOBEC3$^{C57BL/6}$) and ATC (APOBEC3$^{BALB/c}$) target sequences were present 14 and 15 times, respectively, in these genomic regions (of 152 total cytosines); thus, the low levels of deamination are not due to a lack of preferred substrate.

Chapter 2.4.4: APOBEC3 inhibits MMTV reverse transcription

These results suggested that while APOBEC3 packaged into MMTV virions retains its deaminase activity, this is unlikely to be the major mode of restriction since it occurs only at low levels. To determine if APOBEC3 restricts virus infection by other means, we next tested if the mouse protein restricted MMTV at an early step of reverse transcription. We utilized the DNA produced by EnRT reactions with virions isolated from C57BL/6, APOBEC3-/- and BALB/c mice, and performed RT-qPCR using primers specific to MMTV strong-stop DNA. We found that synthesis of this early reverse transcription product was strongly inhibited when the reaction was carried out with APOBEC3$^{C57BL/6}$-containing virions in comparison to APOBEC3-/- virions, while APOBEC3^{BALB} virions showed an intermediate phenotype (Figure 2.4).

A.



B.

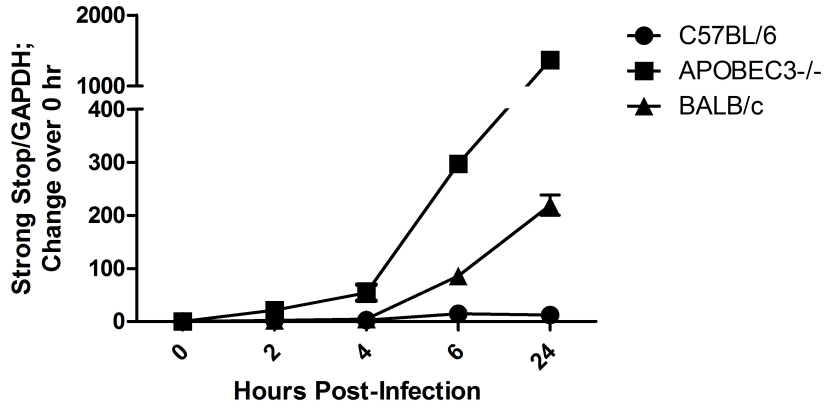


FIG. 2.4. Packaged APOBEC3 restricts MMTV reverse transcription. (A) Equivalent amounts of MMTV virions isolated from C57BL/6, BALB, and APOBEC3-/- mice were subjected to EnRT reactions. DNA was harvested at the indicated time points and strong stop DNA was quantified via qRT-PCR. Shown is the average of three independent experiments. (B) 293T-mTfR1 target cells were infected with MMTV virions isolated from C57BL6, BALB, and APOBEC3-/- mice. Total DNA was isolated from cells at the indicated times after infection, and strong stop DNA was quantified relative to the housekeeping gene GAPDH using qRT-PCR. Triplicates were performed for each time point and averaged; shown is a representative from three independent experiments.

To determine if reverse transcription was also inhibited during virus infection, the virions were used to infect 293T-mTfR1 cells. DNA was isolated at various time points after infection and subjected to RT-qPCR. Similar to what was seen in the *in vitro* EnRT reactions, strong-stop DNA synthesis was greatly reduced in cells infected with APOBEC3-containing virions compared to cells infected with knockout virions, particularly at later time points. Infection with APOBEC3$^{BALB/c}$ virions again showed an intermediate phenotype, suggesting that these virions are restricted compared to those produced in the absence of APOBEC3 but to a lesser extent than those produced in C57BL/6 mice (Figure 2.4). These data suggest that inhibition of early reverse transcription is the major means by which virion-packaged APOBEC3 inhibits MMTV infection.

Chapter 2.4.5: Restriction of reverse transcription is correlated to level of packaged APOBEC3

The differential ability of APOBEC3$^{C57BL/6}$ and APOBEC3^{BALB} to restrict MMTV infection could be due to the level of packaging or to the differences in the proteins. We used a genetic approach to determine if the level of restriction was affected by the level of packaged APOBEC3. We crossed MMTV-infected APOBEC3-/- females and C57BL/6 or BALB/c males, and generated F1 heterozygotes; these heterozygotes acquired the virus from their infected mothers (Figure 2.5). Virus was then isolated from the milk of the heterozygotes at their first pregnancy. Virions derived from both APOBEC3$^{C57BL/6}$/- and APOBEC3$^{BALB/c}$/- F1 mice packaged an intermediate amount of APOBEC3 compared to virions isolated from the parental homozygous strains of mice (Figure 2.5).

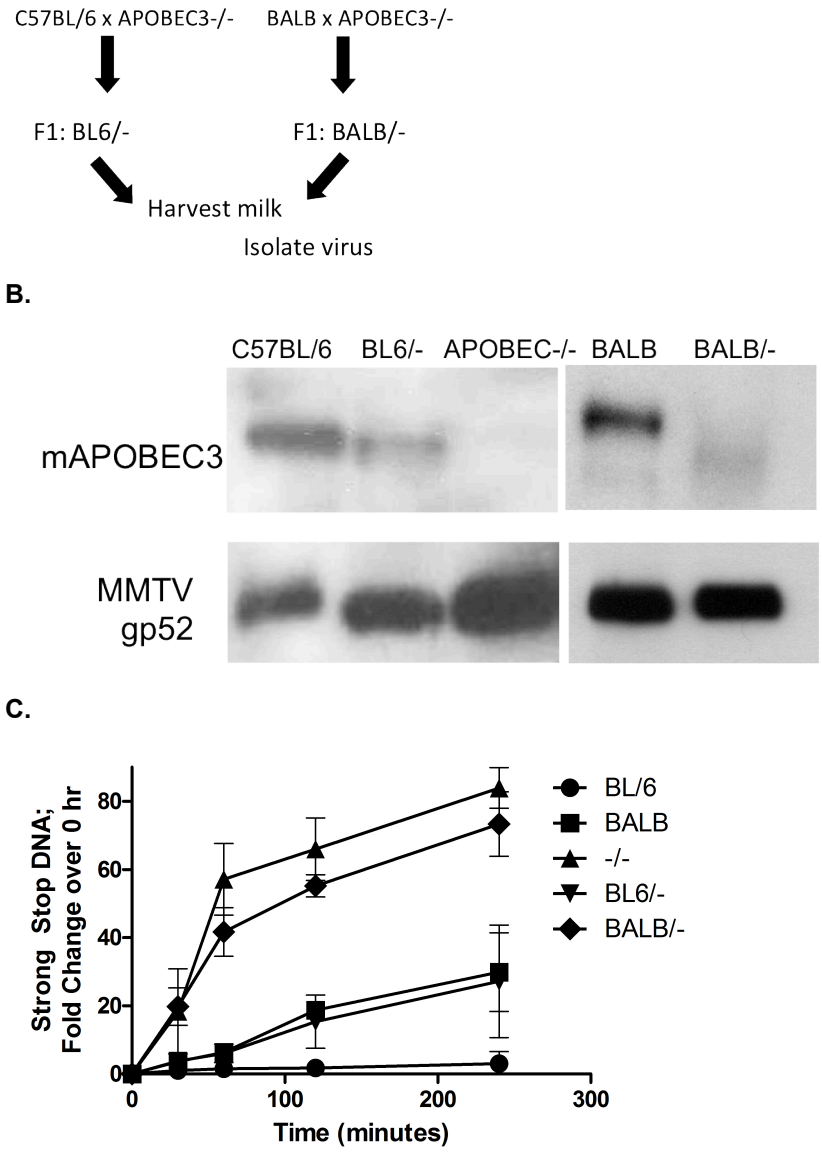


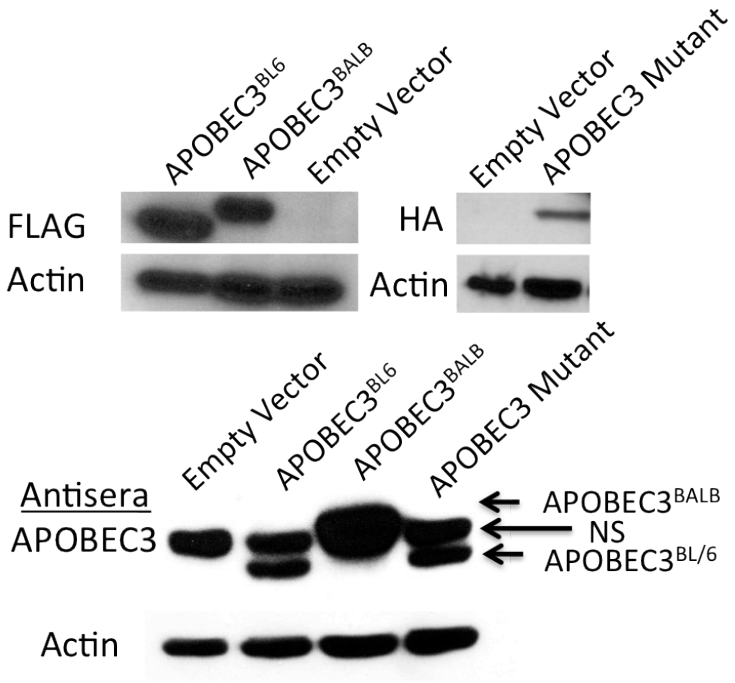
FIG. 2.5. Restriction of reverse transcription correlates with the amount of packaged APOBEC3. (A) MMTV-infected heterozygous F1 mice were obtained by mating APOBEC3-/- male mice to MMTV-infected wild-type C57BL/6 or BALB females. The female F1 mice were bred and MMTV was isolated from milk as described in Materials and Methods. (B) Virions obtained from homozygous, heterozygous, and APOBEC3-/- milk were subjected to Western blotting after normalization by RT-qPCR. Immunoblots were serially probed with rabbit anti-mouse APOBEC3 antisera and goat anti-gp52 antisera as a loading control. (C) Virions isolated from milk of the indicated mouse genotypes were used in endogenous reverse transcription reactions. Strong stop DNA was quantified from total viral DNA harvested at the indicated time points. Three independent reactions were averaged.

Virions isolated from both heterozygous crosses were then used in EnRT assays, and the level of strong-stop DNA was compared to that produced by virus isolated from homozygous C57BL/6, BALB, and APOBEC3-/- mice; the level of virus in each reaction was normalized by RNA and viral protein levels. Virions isolated from homozygous C57BL/6 mice showed the strongest effects of APOBEC3, reflected by low levels of viral transcripts, and homozygous APOBEC3-/- virions produced the highest level of strong-stop DNA (Figure 2.5). Virions from heterozygous animals produced an intermediate level of strong stop DNA between that seen with virus isolated from the parental strain and knockout animals (Figure 2.5). In addition, the amount of strong-stop DNA produced by the APOBEC3^{BALB}/- virions was higher than that seen with the APOBEC3$^{C57BL/6}$/- virions and was not statistically different than that seen with virions from APOBEC3-/- mice (Figure 2.5). This suggests that while alteration in the protein sequences in APOBEC3$^{C57BL/6}$ and APOBEC3^{BALB} affect the substrate specificity, the differential restriction by these proteins is largely due to the level of packaged protein.

Chapter 2.4.6: Target cell expression of either APOBEC3 allele leads to virus restriction

Virion-packaging of APOBEC3 is thought to be critical for viral restriction. However, intracellular cytoplasmic APOBEC3 is also believed to restrict incoming HIV virions (Gillick, 2013, Vetter, 2009; Berger, 2011; Koning, 2011). We also previously showed that cells transfected with APOBEC3 can restrict incoming MMTV viral particles and that APOBEC3 expression in target dendritic cells limits infection *in vivo* (Okeoma, 2009). To determine the intracellular effect of both natural variants of APOBEC3 on incoming *in vivo* produced virions, 293T-mTfR1 cells were stably transfected with Flag epitope-tagged BL/6 or BALB/c variants of APOBEC3 (Rulli, 2008). Western blot analysis using anti-Flag and anti-APOBEC3 antibodies demonstrated that the two variants were expressed at similar levels in each cell line (Figure 2.6).

A.



B.

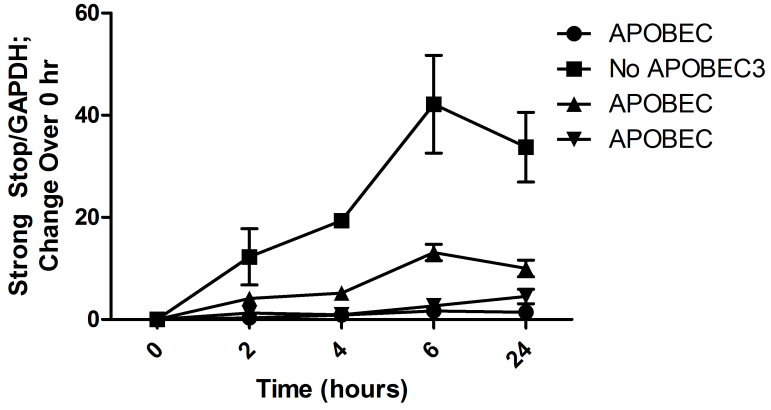


FIG. 2.6. Target cell APOBEC3 can restrict reverse transcription of incoming virions and is independent of cytidine deamination. (A) 293T-TfR1 cells were stably transfected with plasmids encoding APOBEC3^{BL6}, APOBEC3^{BALB}, APOBEC3^{mutant}, or an empty vector (no APOBEC3). Top panel: protein extracts were assessed for APOBEC3 content using their respective epitope tags (anti-FLAG for APOBEC3^{BL6} and APOBEC3^{BALB}; anti-HA for APOBEC3^{mutant}). Blots were stripped and reprobed for β-actin as a loading control. Bottom panel: cell lysates were also probed using anti-APOBEC3 antisera. A cross-reactive protein (NS) seen in all lysates is indicated. (B) Stably transfected cells were infected with MMTV isolated from APOBEC3-/- mice. Total cellular DNA was harvested at indicated time points after infection, and strong stop DNA was quantified in relation to GAPDH. Data representative of three independent experiments.

Virions produced from APOBEC-/- mice were used to infect each of the cell lines, total cellular DNA was collected at various time points after infection and early reverse transcription products were measured by quantitative PCR. Cells expressing either variant were similarly able to restrict early reverse transcription, with approximately five to ten fold less DNA produced in comparison to cells lacking APOBEC3 (Figure 2.6). The intermediate phenotype typical of virions packaging APOBEC3^{BALB} was attenuated in cells expressing this variant compared to cells expressing APOBEC3$^{C57BL/6}$, suggesting that the two natural variants may share similar restriction capacity when expressed at similar levels.

To further test the importance of cytidine deamination in restriction of MMTV, we performed a similar experiment using cells stably transfected with an HA-tagged APOBEC3$^{C57BL/6}$ cytidine deaminase mutant (E73Q/E253Q) (Figure 2.6). Cells expressing this variant were able to restrict reverse transcription of incoming MMTV equally well as the WT APOBEC3$^{C57BL/6}$ (Figure 2.6). Restriction of reverse transcription is thereby independent of cytidine deamination, and these data further support this mechanism as the main method of APOBEC3 mediated MMTV restriction in mice.

Chapter 2.5: Discussion

It is now well-established that mouse APOBEC3 limits the pathogenesis of a number of murine retroviruses *in vivo*, including F-MLV, M-MLV and MMTV (Okeoma, 2007; Takeda, 2008; Low, 2009; Santiago, 2008). Indeed, these studies, which compared infection and pathogenesis in wild type and APOBEC3 knockout mice, provided the ultimate proof that APOBEC3 proteins functioned as *in vivo* restriction factors. APOBEC3 restriction of murine viruses *in vivo* has also provided a model for understanding the natural selection of host restriction factors, as the two natural allelic variants of murine APOBEC3 show differential restriction against a number of viruses which also affects their pathogenicity in different inbred mouse strains (Takeda, 2008; Okeoma, 2009; Santiago, 2008).

However, the mechanism of APOBEC3-mediated restriction *in vivo* has been less well defined. The ability to deaminate cytidines within newly transcribed viral DNA certainly contributes to the effect of APOBEC3 against viruses such as HIV, creating an array of G to A mutations that can lead to missense and nonsense mutations thereby preventing productive infection. Moreover, HIV DNA isolated from patients bears characteristic signatures of APOBEC3-mediated deamination (Janini, 2001; Kieffer, 2005; Pace, 2006). In contrast, while *in vitro* transduced mouse APOBEC3 generates a high rate of G to A mutations on HIV in the presence or absence of Vif, *in vivo*, neither M-MLV nor MMTV show evidence of APOBEC3 mediated mutations, although viral restriction still occurs (Okeoma, 2007; Takeda, 2008).

Here, we show that murine APOBEC3 packaged within MMTV virions is enzymatically active and is able to catalyze cytidine deamination. Moreover, although there is abundant evidence that APOBEC3^{BALB} is less restrictive than APOBEC3$^{C57BL/6}$ *in vivo*, this is not due to diminished catalytic activity. We show that both allelic forms of the protein, produced and packaged *in vivo*, were capable of catalyzing this potentially antiviral reaction. In fact, APOBEC3^{BALB} showed an increase in activity on the 5'TTC3' substrate when normalized to protein input. However, when tested against a full array of deaminase substrates, APOBEC3^{BALB} did not show an increase in activity on all substrates. Notably, while both variants showed relatively promiscuous targeting when compared to APOBEC3G, the two alleles showed different patterns of preferred target sequences (Beale, 2004). Until this study, the mouse APOBEC3 preferred editing site was not well defined, as previous efforts have been inconsistent in APOBEC3 plasmid, allelic variant, and target substrate (Bishop, 2004; Langlois, 2009; Renard, 2010).

A difference in targeting between the allelic variants could be attributable to the presence of several polymorphisms localized to the substrate recognition loop that we have previously identified in the APOBEC enzyme family (Kohli, 2009). However, as with HIV where altering the targeting preference of APOBEC3G did not impact retroviral restriction, we did not find evidence

that the altered sequence preference contributed to differences in MMTV restriction in our efforts to sequence reverse transcript products (Kohli, 2010).

Indeed, while native APOBEC3 is able to catalyze cytidine deamination, further investigation of early reverse transcription products showed very low levels of G to A mutation, similar to what has been seen in integrated proviruses *in vivo* (Okeoma, 2007). Utilization of 3DPCR, a technique that maximizes detection and amplification of DNA with increased A/T content, allowed us to uncover reverse transcription products with slightly more G to A mutations. However, even under these highly selective conditions, rates of mutation remained a fraction of that seen in HIV, which can vary from 7 to 8% total G to A mutations *in vivo* (Armitage, 2012; Gandhi, 2008; Land, 2008).

Others have suggested that murine viruses may be able to block APOBEC3 deaminase activity. Browne and Littman showed that F-MLV was restricted by both forms of murine APOBEC3 *in vitro*, but had only 1 G to A mutation per kb of genome. In comparison, both APOBEC3 alleles were also able to restrict HIV, but produced 16 G to A mutations per kb of HIV genome (Browne, 2008). However, Petit et al. utilized 3DPCR to uncover a significantly higher rate of G to A mutations from F-MLV both *in vitro* and *in vivo* (Petit, 2009). Our MMTV data closely matches the F-MLV data from these two groups, suggesting similarities between the two murine viruses. Both show G to A mutation at a very low level, and may share a mechanism to block APOBEC3-induced deamination. Our data also showed that reverse transcription amplicons recovered from virions isolated from BALB/c or C57BL/6 mice had similar levels of editing. The large majority of the G to A mutations was found in the preferred deamination site 5'-GXA-3', with preference for GG and GA dinucleotides within those parameters. Our analysis further supports a difference in target sequence preference between the two APOBEC3 alleles, as targeting patterns between APOBEC3$^{C57BL/6}$ and APOBEC3^{BALB} mirrored what was seen in the *in vitro* assay. However, if deamination were a critical component of restriction, we would expect these two alleles to show different levels of editing, to mirror their different levels of overall restriction.

Since restriction of MMTV is clearly mediated by mouse APOBEC3, we also assessed whether reverse transcription was inhibited in the presence of the enzyme. This block to reverse transcription has been described in several systems, although primarily in viruses produced *in vitro* (Bishop, 2006; Iwatani, 2007; Holmes, 2007; Bishop, 2008). However, other studies suggest that this is an artifact of high levels of transfected protein and maintain that deamination is necessary for restriction (Miyagi, 2007). Our analysis of virions isolated *in vivo* allows an understanding of the activity of native protein packaged at endogenous levels. Our results clearly show that both alleles of APOBEC3 are able to prevent early reverse transcription when packaged within virions. The alleles show distinct differences in their ability to restrict in this manner, with packaged APOBEC3^{BALB} showing less restriction than the APOBEC3$^{C57BL/6}$ counterpart. We have also shown that APOBEC3 expressed within target cells is able to restrict reverse transcription of incoming MMTV virions that do not contain packaged APOBEC3. This suggests a possible dual role for APOBEC3- both within the virion and within target cells. It also provides evidence that the MMTV core and the RT complex is accessible to intracellular factors during reverse transcription, supporting recent studies showing that uncoating and reverse transcription are linked (Hulme, 2011; Roa, 2012). However, unlike with the *in vivo* packaged APOBEC3, we did not see as distinct a difference between the two allelic variants, which is likely due do the similar levels of expression within our transfected cells. An APOBEC3 mutant that is unable to catalyze reverse transcription was also able to efficiently prevent reverse transcription, confirming that this arm of restriction is truly independent of deamination. Taken together, our data suggest that the mechanism by which APOBEC3 restricts infection by MMTV, namely blocking reverse transcription, is the same when packaged or present in the target cell.

The two major APOBEC3 alleles differ in both sequence and level of expression *in vivo* and it has been suggested that positive selection by murine retroviruses may have played a role in their acquisition. While we showed that the two alleles show altered sequence substrate specificity, our data support the idea that it is the levels of packaged APOBEC3 rather than polymorphic amino acid differences in the APOBEC3$^{C57BL/6}$ and APOBEC3^{BALB} alleles that affect

the degree of reverse transcriptase inhibition. By creating heterozygous mice that express intermediate levels of APOBEC3 we generated MMTV virions that package half the parental levels of protein and showed that the amount of reverse transcription products seen correlated with level of packaged protein. While several *in vitro* studies have suggested that coding sequence differences between the two alleles contribute to the more restrictive phenotype of the APOBEC3$^{C57BL/6}$ allele, *in vivo* studies using C57BL/6 X BALB/c crosses have led to conflicting results (Takeda, 2008; Santiago, 2008; Li, 2012; Abudu, 2006; Santiago, 2010). Additionally, several studies suggested that a single copy of the APOBEC3$^{C57BL/6}$ allele is sufficient to confer full resistance to MLV infection *in vivo*, since there was no difference in F-MLV or AKV infection of APOBEC3+/+ and APOBEC3+/- mice (Takeda, 2008; Langlois, 2009). These studies may be confounded by multiple genes in the different genetic backgrounds that affect both infectivity and immune response (e.g. levels of neutralizing antibodies) in whole animal studies. Moreover, these studies examined late stages in infection, by which time restriction by the single copy of APOBEC3 in APOBEC3+/- mice may be dominant after many rounds of infection. Indeed, our studies with M-MLV indicated that at early times after infection, APOBEC+/- mice were intermediate in infection levels between knockouts and wild type mice, but succumbed to disease with the same kinetics as the knockouts (Low, 2009). Our studies here, which used virions to infect the same target cells lacking APOBEC3 proteins (293T-mTfR1) in single round infections, may more accurately reflect the direct effects of APOBEC3 on infection.

In sum, our data show that MMTV is susceptible to reverse transcriptase inhibition by APOBEC3, and is sensitive to levels of protein expressed and packaged within virions. Future work is required to determine the mechanism by which MMTV and other murine viruses avoid APOBEC3 mediated G to A mutations, as well as the mechanism by which all APOBEC3 proteins are able to prevent reverse transcription.

Chapter 3: The effect of APOBEC3 on antiretroviral drug resistance *in vivo*

Chapter 3.1: Abstract

The high rate of retroviral reverse transcription errors can result in rapid viral evolution. This is particularly important in pathogenic viruses, as mutations can lead to drug resistance or immune system evasion. Mutations can also arise via cellular factors, including APOBEC3. This antiviral family of proteins creates G-to-A mutations in retroviral DNA leading to replication-defective viruses. However, at a low level, these mutations could lead to increased viral evolution without affecting viral infectivity, potentially decreasing time to drug resistance. We have shown that mouse APOBEC3 generates a very low level of G-to-A mutations in the murine retrovirus MMTV. We assessed whether these mutations can lead to increased viral evolution in the presence of a strong evolutionary pressure. Treatment of MMTV infected animals with the antiretroviral AZT began early in infection. We tracked viral loads via PBMC infection levels in APOBEC3+/+ and APOBEC3-/- mice with or without treatment over a course of ten weeks. On average, virus in APOBEC3+/+ mice responded better at the beginning of treatment but these mice showed a rise in viral load over time, while virus in APOBEC3-/- mice responded only slightly, but maintained steady viral loads over time. These data suggest that APOBEC3 plays a role in viral evolution in response to drug treatment.

Additional experiments, including in-depth sequencing analysis and functional assays to confirm viral mutation and drug resistance, were not easily addressed with the MMTV-infected wild type and knockout mice. We thus next assessed similar questions using human APOBEC3G transgenic mice infected with M-MLV, with APOBEC3-/- mice used as controls. Animals were treated with AZT, and both integrated viral DNA and plasma virus RNA were isolated and measured at various time points over six weeks of infection. APOBEC3G-expressing mice had an increased response to AZT compared to their APOBEC3-/- littermates. A region of *pol* corresponding to the active site region of reverse transcriptase was isolated and sequenced from mice at two and six weeks post infection, and viral diversity was assessed within each group. We found that APOBEC3G led to increased viral diversity in the presence or absence of the drug. These data support a role for APOBEC3G in viral evolution and potential drug resistance.

Chapter 3.2: Introduction

APOBEC3 acts as an intrinsic immune factor against retroviral infection. This family of proteins is capable of a two-pronged antiviral attack, through inhibition of viral reverse transcription (see Chapter 2.4.4), as well as via cytidine deamination. APOBEC3 deaminates C residues on the minus strand of retroviral DNA during reverse transcription, resulting in G-to-A transitions on the coding strand. The effect of different human APOBEC3 proteins on HIV is highly variable, as the accessory protein Vif blocks APOBEC3 variants to different levels. Vif targets the APOBEC3 family for ubiquitination and proteosomal degradation by bringing APOBEC3 and an Elongin-BC-dependent ubiquitin ligase complex together (Yu, 2003). Degradation of APOBEC3 prevents its incorporation into budding virions, thus preventing any antiviral activity at the subsequent round of infection. Vif is able to neutralize APOBEC3D, APOBEC3F, and APOBEC3G, but not APOBEC3A and APOBEC3B. Also, APOBEC3F may be slightly less sensitive to Vif than APOBEC3G (Chiu and Greene, 2008). Of note, the low expression of APOBEC3B in HIV-infected cell types eliminates it as a threat to the virus. APOBEC3H may or may not be susceptible to Vif, depending on the haplotype of APOBEC3H encoded. This enzyme is highly polymorphic, with each haplotype possessing a variable level of anti-HIV activity, which is at least in part dependent on interactions with Vif (Ooms, 2013).

Normal variation in Vif can also cause different levels of anti-APOBEC3 activity. Vif variants isolated from HIV infected patients showed 3-9% divergence from a prototype *vif* sequence at the nucleotide level (Simon, 2005). Much of this nucleotide diversity results in amino acid changes, which can cause truncated proteins, folding defects, or mislocalization (Simon, 1999). These changes can also affect binding capacity to APOBEC3, or interactions with key mediators of Vif regulation and APOBEC3 degradation, including core-binding factor-beta (CBFβ) and Elongin B/C (Zhou, 2012). This disruption of regular Vif function prevents complete inactivation of APOBEC3.

If Vif fully blocks APOBEC3 from accessing the viral genome, no APOBEC3-mediated mutations will occur. However, if Vif is unable to cause complete degradation of APOBEC3, viral

packaging can occur, potentially resulting in significant levels of G-to-A mutations. These mutations are typically deleterious to the virus, preventing productive infection. However, HIV viruses with G-to-A mutations have been found in human plasma, suggesting that these mutations are not always lethal. These viruses may be the result of an incomplete block of APOBEC3 by different natural Vif variants (Simon, 2005). Thus, partial neutralization of APOBEC3 could result in sublethal mutagenesis, thereby promoting viral sequence diversity.

There is also evidence that even lethal levels of APOBEC3-mediated mutagenesis can affect viral evolution, through recombination of mutant virus and replication competent virus. Retroviruses package two copies of viral RNA into virions, and recombination of both copies commonly occurs during reverse transcription (Rambout, 2004). In a cell dually infected with a lethally mutated virus and a functional provirus, progeny viruses may carry one copy of each. Mutations that benefit the virus may be transferred from the mutant copy to the functional one, working as an additional factor to promote viral evolution. Several studies have shown that G-to-A mutations in HIV can result in drug resistance to many first line therapeutics. These APOBEC3-mediated drug resistance mutations have been transferred to drug-susceptible viruses via this recombination method in *in vitro* infection models (Mulder, 2008). However, this potentially pro-viral role of APOBEC3 is not well understood. Furthermore, the effect of APOBEC3 on long-term viral evolution has not been studied *in vivo*.

G-to-A induced drug resistance mutations may stem from sublethal APOBEC3 mutagenesis, as viruses with defective Vif proteins are more likely to contain these drug resistance mutations (Fourati, 2010). It has been estimated that approximately 20% of all known drug-resistance mutations are the result of G-to-A substitutions, with A-to-G transitions also common at approximately similar levels (Berkhout and de Ronde, 2004). This pattern of mutations reflects the general mutational bias seen in HIV-1 (Berkhout, 2001). The proportion of these mutations stemming from APOBEC3, error-prone reverse transcriptase, or RNA transcription machinery is difficult to estimate, and is highly dependent on variantions in Vif and

reverse transcriptase. Thus, the specific contribution of APOBEC3 to the development of HIV drug resistance *in vivo* has been difficult to ascertain.

Mice have a single APOBEC3 gene, and the natural murine pathogen, Mouse Mammary Tumor Virus, has no known Vif-like factor. Murine APOBEC3 is able to restrict MMTV, but does not induce significant G-to-A mutations *in vivo* (Okeoma, 2007; Chapter 2.4.3). However, both HIV and MMTV share a similar bias in their nucleotide composition, with high percentage A content. Interestingly, despite the rapid evolution of HIV throughout its relatively short lifespan in humans, this nucleotide composition has remained stable (van der Kuyl and Berkhout, 2012). This high A content is a general property of the lentiviral family to which HIV belongs, but is not consistent in all other retroviruses; for example, the deltaretrovirus HTLV-1 has relatively low A content (Table 3.1). The adenine-enriched HIV and MMTV genomes could be the result of long-term selective pressure by APOBEC3. This selective pressure may also continue to shape viral genomes, aiding the virus in overcoming evolutionary pressures such as antiretroviral drugs.

Table 3.1 Retroviral Genome Composition
The genomes of both the human retrovirus HIV and mouse retrovirus MMTV are not constructed from equal amounts of each of the four nucleotides. These genomes are enriched in adenine, and depleted of cytosine. This skewing of nucleotide content is specific to these viruses, as the human retrovirus HTLV-1 and the murine retrovirus M-MLV have their own unique nucleotide content patterns (van der Kuyl and Berkhout, 2012).

Virus	%A	%G	%T/U	%C
HIV	36	24	22	18
MMTV	31	22	27	21
HTLV-1	24	18	23	35
M-MLV	26	24	21	29

Our goal is to determine the effect of APOBEC3 on the evolution of murine retroviruses, and better understand the effect of APOBEC3 on anti-retroviral drug resistance *in vivo*. Little is known about the activity of antiretroviral drugs against murine retroviruses. These drugs were

designed for HIV-specific use. However, retroviruses share significant functional homology, and thus some anti-HIV therapeutics have previously been shown to inhibit MMTV and M-MLV (Held, 1994; Paprotka, 2010; Sakuma, 2010). Other drugs that are able to inhibit HIV have no activity against these viruses. For example, lamivudine (3TC) effectively impedes HIV reverse transcription, yet has insignificant activity against MLV. This inability to affect MLV reverse transcriptase is thought to be based on a residue within the YMDD sequence motif, which when mutated in HIV-1 RT is known to cause drug resistance (M184V). M-MLV RT naturally contains a valine within this motif (Ndongwe, 2011). Further studies of XMRV, which shares 95% homology with M-MLV, engineered the known AZT resistance mutation of HIV-1 RT, Q151M, into the XMRV RT (Q190M). When expressed as recombinant protein, the mutant protein was five-fold less susceptible to AZT *in vitro*, indicating that homologous drug resistance mutations can occur in multiple retroviral genomes (Ndongwe, 2011). This is helpful in constructing an understanding of potential drug resistance mutations within murine retroviruses, as so little is specifically known. Treatment of these murine retroviruses with antiretrovirals designed for HIV-1 could then serve as a selective evolutionary pressure for viral genomes *in vivo*.

Chapter 3.3: Materials and Methods

Chapter 3.3.1: Cells and Viruses

MMTV virions were isolated from mammary tumors or breast milk of MMTV-infected mice (see Chapter 2.3.1 for protocol details). Viral RNA was isolated using the QIAGEN RNeasy Mini Kit (QIAGEN, Germantown, MD) and reverse transcribed using SuperScript III First-Strand Synthesis SuperMix, based on a modified M-MLV RT (Invitrogen Corp., Carlsbad, CA) (see Chapter 2.3.1). For *in vivo* studies, newborn pups were naturally infected with MMTV RIII virus from their infected mothers.

293 cells stably expressing mouse transferrin receptor 1 were maintained in DMEM supplemented with 10% FBS, 100 U/ml penicillin, 100 ug/ml streptomycin, 2mM L-glutamine and 100 ug/ml geneticin (see Chapter 2.3.1). NIH3T3 murine fibroblast cells were obtained from

ATCC, and cultured in Dulbecco-modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum, 100 U/ml penicillin, 100 ug/ml streptomycin, and 2 mM L-glutamine. 43-D cells are modified NIH3T3 fibroblasts that are stably infected with wild type M-MLV; these cells were cultured in an identical manner. This cell line was used for M-MLV virus production. Cells were grown to 50% confluence in multiple 100 mm culture dishes, at which time the media was changed and reduced to half of the normal volume (5 mL per 100 mm plate). After an additional 24 hours of growth, supernatant from all plates was combined and centrifuged at 500g for 10 minutes at 4°C to remove cells and cellular debris. Further purification of the supernatant was attained by syringe filtration through 0.45 μm filters. The resulting virus-containing media was treated with 20u/mL of recombinant DNase I (Roche, Mannheim, Germany), and incubated for 30 minutes at 37°C. Virus was aliquoted and stored at -80°C.

The virus was titered on NIH3T3 cells, plated at 7 x 10^4 cells/well in 6-well culture dishes one day before infection. Serial ten-fold dilutions were made in 500 μL of culture media (see above), and 8 ug/mL polybrene added before initiating infection at 37°C. After 2 hours, virus-containing media was removed from cells and replaced with 3 mL/well of fresh media. Four days post-infection, media was removed and cells were washed with PBS. MLV Env expressing cells were detected by staining with the mouse 538 env monoclonal antibody for one hour rocking at 4°C. Cells were then washed with PBS with 2% fetal bovine serum, and then incubated with a secondary antibody, Alexa Fluor-conjugated goat anti-mouse IgG (H+L) (Invitrogen, Carlsbad, CA), diluted 1:200 in PBS + 2% FBS, 40 minutes rocking at 4°C. Cells were washed and green fluorescent colonies were counted on a microscope under UV light. Colonies were counted in a well that provided between 5-50 colonies, and viral titer was calculated considering the dilution factor specific to that well (in plaque forming units/mL of virus).

Chapter 3.3.2: *In vitro* infection and drug optimization

293T-TrFR1 cells employed for MMTV infections, using virions isolated from APOBEC3-/- mammary tumors. Cells were pre-treated for two hours with the reverse transcriptase inhibitor AZT (zidovudine), the integrase inhibitor raltegravir, or the reverse transcriptase inhibitor lamivudine, all of which were obtained from the NIH AIDS Research and Reference Reagent Collection. Virus was then added in the presence of drug and 8 ug/ml polybrene, and media with drug was replaced after four hours. Cells were harvested 24 hours after infection, washed with PBS, and total cellular DNA extracted using QIAGEN DNeasy Blood & Tissue Kit (QIAGEN, Germantown, MD).

3T3 cells were used for M-MLV infections. To test for drug susceptibility *in vitro*, cells were plated at 3 x 10^4 cells/well in a 12-well tissue culture dish one day prior to infection. Cell confluency of 60-70% was obtained for optimum infection conditions. Two hours prior to infection, cells were treated with three different antiretroviral drugs, at varying concentrations. All drugs were obtained from the NIH AIDS Research and Reference Reagent Collection, and solubilized in ethanol or water. The protease inhibitor ritonavir was used at 1 μM, 10 μM, and 100 μM, raltegravir was used at 10 nM, 100 nM, and 1 mM, and AZT (zidovudine) was used at 5 μM, 50 μM, and 500 μM. Cells were infected at an MOI of 100, in a volume of 200 μL, with drug added to the appropriate concentration. Two hours after infection, virus was removed and fresh media with drug was applied. Forty-eight hours after infection, cells were washed with PBS, and then removed from the plate in 200 μL of PBS. Isolation of total cellular DNA was completed using the QIAGEN DNeasy Blood & Tissue Kit. Similar analysis was performed using a wider range of concentrations of AZT to determine optimum dose and create a dose curve *in vitro*, with drug concentration ranging from 1 nM to 10 μM.

Chapter 3.3.3: Mice and Sample Collection

APOBEC3-/- and APOBEC3+/+ mice were used in MMTV infection studies. APOBEC3+/+ mice encode the wild-type mouse APOBEC3$^{C57BL/6}$, while APOBEC3-/- contain a

gene trap vector in intron 4 of the APOBEC3 gene, and have been backcrossed to the C57BL/6 background (see Chapter 2.3.2).

APOBEC3-/- and transgenic mice expressing human APOBEC3G were used in MLV infection studies. The transgenic mice were created by injecting a pCAG vector encoding the gene into C57BL/6 blastocysts. The gene is under control of the chicken β-actin promoter, allowing for universal expression, and transcription terminates after addition of the rabbit β-globin poly(A) tail. Resulting pups were screened for the transgene via PCR for human APOBEC3G. Transgenic founders were crossed with APOBEC3-/- mice to obtain heterozygous APOBEC3G mice not expressing mouse APOBEC3-/-. All mice were housed according to the policies of the Institutional Animal Care and Use Committee of the University of Pennsylvania.

Mice infected with MMTV were screened for level of infection at various time points. Blood was drawn by retro-orbital collection into 10% RPMI in PBS with heparin (500 USP units/mL) (Sigma Aldrich Corp., St. Louis, MO). PBMCs were isolated using Gey's solution (150 mM NH_4Cl, 10mM $KHCO_3$, 1mM EDTA, pH 7.3). Whole blood was added to 1 mL of Gey's, shaken for one minute, and centrifuged at 5000g for 5 minutes at 4°C. After removal of the supernatant, this process was repeated once to produce PBMCs in the relative absence of red blood cells. The pellet was resuspended in PBS, and total cellular DNA extracted using the QIAGEN DNeasy Blood & Tissue Kit (QIAGEN, Germantown, MD).

Blood was also collected via retro-orbital bleed from mice used in MLV studies. Here, blood samples were not incubated on ice upon collection. Samples were spun at 16,000g for 10 minutes at room temperature to remove all cellular debris from plasma, which was then removed and stored at -80°C. PBMC DNA was isolated following steps detailed above. When necessary, splenic DNA was also isolated. Spleens were removed after sacrifice and added to PBS on ice. A single cell suspension was created by mechanical disruption through fine mesh to remove fat and the splenic capsule. Red blood cells were lysed once with Gey's solution, with a five-minute incubation on ice prior to addition of equal volume DMEM supplemented with 10%FBS. Cells

were spun at 5000g for 5 minutes at 4°C, and DNA was extracted from the cell pellet using the QIAGEN DNeasy Blood & Tissue Kit.

Chapter 3.3.4: *In vivo* infection and drug treatment

MMTV-infected mice were treated with AZT upon weaning at four weeks of age. Males and females were separated to prevent pregnancy, and mice housed together were either not treated or provided with AZT. Upon weaning, treated mice were injected intraperitoneally with 2 mg AZT reconstituted and diluted in water (approximately 100 mg/kg assuming weight of approximately 12-13 g/mouse). Water bottles containing reconstituted diluted AZT were then provided to these animals until the end of the experiment. Water contained 0.25 mg/mL AZT, providing an oral dose of approximately 40-50 mg/kg (40-50 μg/g) per day assuming water intake of 3 mL/day.

Mice used for MLV studies were infected between 24-48 hours after birth, via intraperitoneal injection with 3.7×10^4 plaque forming units of M-MLV in 100 μL total volume. Upon infection, tissue samples were acquired for genotyping. DNA was isolated using the QIAGEN DNeasy Blood & Tissue Kit, and used for genotyping PCR using the following primers. A3GFor: 5'-GGG ACC CAG ATT ACC AGG AG-3'; A3GRev: 5'-GCA GAT TAT TCC AAG GCT CAA-3'. PCR products were analyzed via agarose gel electrophoresis to distinguish APOBEC3G transgenic mice from APOBEC3-/- mice. Approximately half of all mice were begun on AZT (Zidovudine syrup, 10 mg/mL, Aurobindo Pharma Limited, Dayton, NJ). Dosing experiments were performed in APOBEC3+/+ mice, and AZT doses ranged from 0.01 mg/g to 0.1 mg/g. For viral evolution studies, 0.05 mg/g was administered daily. All mice were weighed daily to ensure proper dosage.

Chapter 3.3.5: Polymerase Chain Reaction

Tumor and milk-derived MMTV-RIII viral cDNA was used for PCR amplification of viral genes, including *env* and the 3' LTR. Primers used included MMTVSagFor: 5'- TCC CGA GAG

TGT CCT ACA CCT AG-3'; MMTVSagRev: 5'-GAA GAT CTT CAA GGG CAA TGC CTT AAT ACT A-3'; MMTVEnvFor: 5'-GAA AGT TAT TGG GCT TAC C-3'; MMTVEnvRev: 5'-AAG TAA CAC AGG CAG ATG TAG G-3'. For *in vitro* infection studies, integrated MMTV was measured by quantitative PCR for strong stop DNA using primers against MMTV 3'LTR: LTRFor: 5'-CGT GAA AGA CTC GCC AGA GCT A-3' and LTRRev: 5'-GAA GAT CTT CAA GGG CAA TGC CTT AAT ACT A-3'. These primers amplify 90bp product from the MMTV RIII U5 region. Equal volumes of cellular DNA was combined with Power SYBR Green PCR Master mix and amplified under standard conditions on an ABI Prism Model 7900HT.

Integrated virus in total PBMC DNA isolated from MMTV-infected mice was assessed via quantitative PCR. The concentration of DNA was measured using a NanoDrop 2000 Spectrophotometer (Thermo Scientific, Waltham, MA), and equal quantities used for a first-round PCR reaction, using primers specific for MMTV RIII, spanning from *pol* to *sag*, MMTVPolFor: 5'-ACC AGA TGA TTC ACA GG-3'; MMTVSagRev: 5'-GAA GAT CTT CAA GGG CAA TGC CTT AAT ACT A-3'. This first-round PCR included 35 cycles at 1 minute 95°C, 1 minute 55°C, and 3 minutes at 72°C using GoTaq DNA Polymerase MasterMix (Promega). An equal volume of the first-round product was transferred to a 386 well plate in triplicate for nested quantitative PCR using the same reverse primer, and a *sag*-specific forward primer LTRFor: 5'-CGT GAA AGA CTC GCC AGA GCT A-3'. Power SYBR Green PCR Master mix (Life Technologies, Carlsbad, CA) was added to the DNA and primers, and samples were amplified under standard conditions on an ABI Prism Model 7900HT.

For both *in vitro* and *in vivo* infection experiments, integrated MLV was also measured by quantitative PCR, using primers specific to MLV *env*: SUMLV2F: 5'-CCT ACT ACG AAG GGG TTG CC -3' and SUMULV2R: 5'-CAC ATG GTA CCT GTA GGG GC -3'. Three μL of PBMC DNA was added to Power SYBR Green PCR Master mix, and amplified under standard conditions as above.

Chapter 3.3.6: Titering of Plasma Viremia

Plasma MLV levels were measured in a similar manner as described in Chapter 3.3.1. Briefly, 3T3 cells were plated in a 6-well cell culture dish 24 hours prior to infection. Plasma was added to fresh DMEM media and two dilutions were produced per sample. The first dilution contained 2.5 μL of plasma, and was used to produce a single 10-fold dilution (0.25 μL total plasma) in a total volume of 500 μL with polybrene. This was added to the plated 3T3 cells and incubated for two hours. Virus was then removed and fresh media added. Four days after infection, cells were analyzed for MLV infection foci as described in Chapter 3.3.1.

Chapter 3.3.7: Sequencing and Statistics

MLV plasma virions were used for sequencing of MLV *pol*. RNA was isolated from 50 μL of plasma using the PicoPure RNA Isolation Kit (Life Technologies, Carlsbad, CA). Plasma was added to 120 μL Extraction Buffer and incubated at 42°C for 30 minutes. The samples were then centrifuged at 3,000g for 2 minutes, and the supernatant was applied to the included column. Samples were then treated as directed by manufacturers protocol. Viral RNA was eluted in 12 μL volume (included Elution Buffer). Total RNA was reverse transcribed using the QIAGEN QuantiTect Reverse Transcription Kit (QIAGEN Corp., Germantown, MD). gDNA Wipeout Buffer was added to remove any cellular or viral DNA contaminants prior to reverse transcription. The reverse transcription reaction was carried out in a total volume of 16 μL, and cDNA was stored at -20°C. MLV *pol* was amplified from viral cDNA using the following primers specific for Moloney virus: MLVRTFor: 5'-TGC CAG TCC CCC CTG GAA CAC G-3' and MLVRTRev: 5'-GGG CAG TTA GAA GAG CTT GC-3'. Phusion High-Fidelity PCR Master Mix (NEB, Ipswich, MA) was used in a 25 μL reaction, with 40 total cycles including 1 minute at 95°C, 1 minute at 55°C, and 1 minute at 72°C.

PCR products were run on a 0.8% agarose gel for product confirmation. DNA bands corresponding to the correct fragment size were cut out of the gel and isolated using QIAGEN

QiaQuick Gel Extraction Kit (QIAGEN Corp., Germantown, MD). The manufacturers protocol was followed, with some modifications for increased DNA purity. The column was incubated with Wash Buffer at 37°C before centrifugation, and this step was repeated for an additional wash. Elution Buffer was heated to 72°C prior to addition to the column, and once applied, was incubated on columns at 37°C for 5 minutes prior to elution.

Purified DNA fragments were cloned into the pCR2.1-TOPO vector, as directed (Invitrogen, Inc.). The resulting plasmids were transformed into OneShot TOP10 chemically competent *E. coli* (Invitrogen, Inc.), and bacteria were plated on LB plates containing 100 µg/mL ampicillin for selection. Following overnight growth at 37°C, colonies were picked and placed into 30 µL ddH_2O for direct *in vitro* plasmid amplification using TempliPhi Amplification Kit according to manufacturers directions (GE Healthcare, Piscataway, NJ). Samples were amplified during an 18-hour incubation to maximize template DNA. Samples were diluted in an additional 40 µL ddH_2O and one strand was sequenced on an ABI 3730XL automated cycle sequencer with ABI BigDye Taq FS Terminator V 3.1 (Applied Biosystems, Foster City, CA) at the University of Pennsylvania DNA Sequencing Facility.

Sequences from each mouse were aligned using ClustalW, and each nucleotide change annotated. Sequences with nucleotide changes were translated into amino acid sequence using ExPASy Translate Tool, and each amino acid change annotated. Mice with multiple viruses harboring the same amino acid sequence change were also recorded.

Statistical analysis was performed using the GraphPad/PRIZM software, with the exception of t-tests, which were performed in Microsoft Excel.

Chapter 3.4: Results

Chapter 3.4.1: Potential long term effects of mAPOBEC3 on MMTV

The primary means by which APOBEC3 inhibits MMTV infection is by blocking reverse transcription (see Chapter 2.4.4). However, G-to-A mutation does occur at a low level in an APOBEC3-dependent manner (<0.5% mutation rate; see Figure 2.2). Thus, we considered the

possibility that over many rounds of replication, this mutation rate may influence the retroviral genome. As described previously, the MMTV genome is similar to HIV in that it has a high prevalence of genomic adenine, suggestive of an evolutionary pressure to maintain this skewed content. Murine leukemia viruses are also inhibited by murine APOBEC3 in a cytidine deaminase-independent manner (Santiago, 2008; Low, 2009). However, studies of recently integrated endogenous MLV sequences suggest that mouse APOBEC3 has caused significant mutation in these viruses, as viruses show an increase in A corresponding to a decrease in G, with the majority of mutations occurring in the mouse APOBEC3$^{C57BL/6}$ preferred 5'-TTC-3' context. Although modern ecotropic MLV is not edited by murine APOBEC3, the editing of these endogenous viruses is hypothesized to have occurred directly prior to integration, and to have played a role in its inactivation (Jern, 2007). This provides further evidence that murine APOBEC3 could have a potential important mutagenic effect.

We thus examined MMTV sequences isolated from APOBEC3+/+ and APOBEC3-/- mice. Both genotypes of mice were infected with identical MMTV RIII and the virus was then passed vertically through milk in multiple generations of mice (approximately 40-50 generations). Viruses were isolated both from tumor tissue and from breast milk, and viral RNA was isolated for sequencing analysis. A 2400 bp segment of genome corresponding to *env* and the 3'LTR (including *sag*) was amplified and sequenced from each sample. We found ten locations at which an A-to-G or T-to-C nucleotide changes occurred in viruses isolated from APOBEC3-/- mice but did not occur in wild type mice (Fig. 3.1). These changes were independent of viral tissue source (tumor or breast milk). This is consistent with the hypothesis that APOBEC3 plays a role in maintaining the high AT content, with lack of this pressure leading to a reversion to higher GC content. However, of the ten changes, two did not fit a model of AT to GC reversion, and could be compensatory mutations to retain infectivity.

Seven of the 10 changes altered the amino acid codon (Fig. 3.1). Several of these nonsynonymous mutations were located near the receptor binding domain, potentially altering viral entry. Most of these amino acid changes cluster in close proximity, further suggestive that

these mutations are not random but instead important for the viral lifecycle. Other nucleotide mutations found in *env* or the 3'LTR were not genotype specific, as they were only seen in single isolates.

A.

Nucleotide (position and consensus)	APOBEC3+/+ Viral Isolates (9 total)	APOBEC3-/- Viral Isolates (6 total)
763; T	T: 4/9 (45%)	C: 5/6 (83%)
842; A	A: 7/9 (78%)	G: 6/6 (100%)
843; A	A: 7/9 (78%)	G: 6/6 (100%)
845; A	A: 6/9 (67%)	G: 6/6 (100%)
859; T	T: 8/9 (89%)	C: 6/6 (100%)
870; A	A: 8/9 (89%)	G: 6/6 (100%)
876; A	A: 9/9 (100%)	G: 6/6 (100%)
991; C*	C: 7/9 (78%)	T: 4/6 (67%)
1018; C*	C: 8/9 (89%)	C: 4/6 (67%)
1020; T	T: 9/9 (100%)	C: 4/6 (67%)

B.

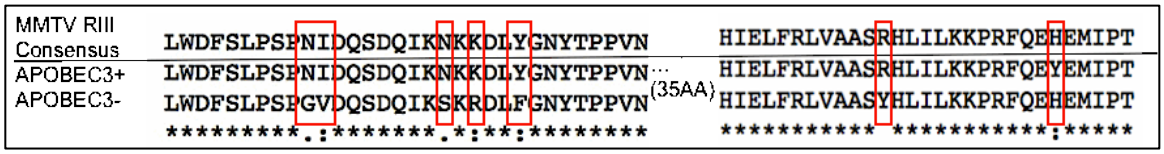


Fig. 3.1. MMTV passaged in APOBEC3-/- mice differs from virus passed in APOBEC3+/+ mice. (A) MMTV RIII was isolated from both mammary tumor tissue and breast milk from APOBEC3-/- and APOBEC+/+ mice, and a 2400 bp region of *env* and 3' LTR was sequenced. Isolate number refers to the number of unique source samples from different mice. Specific nucleotides were found to vary between viruses dependent on the source genotype. The number and fraction of the nucleotide found at these specific *env* locations are listed for viruses from both genotypes of mice. Nucleotide position is indicated based on GenBank accession number DQ767968.1 (Mouse mammary tumor virus envelope glycoprotein Pr73 and superantigen genes). Statistical association of APOBEC3-/- isolated viruses with total A-to-G changes (compared to all other mutations) was confirmed by Fisher's Exact Test ($p<0.01$). Starred nucleotides indicate mutations which do not fit a model of AT to GC reversion, but are notable as they are located within this region and seen in multiple viral isolates. (B) The nucleotide changes found in viruses from APOBEC3-/- mice resulted in amino acid changes clustered together near the receptor binding domain region. These changes are highlighted in red.

Although the viruses isolated from all APOBEC3-/- samples were not identical, most of the nucleotide changes leading to amino acid differences were present simultaneously in individual viruses. This suggests that in addition to the genetic changes that occur in the presence or absence of APOBEC3, there are functional constraints to maintain infectivity that limit the extent of these changes. These data suggest that while mAPOBEC3 does not significantly deaminate MMTV, it does provide pressure for the virus to maintain a high AT content, which may have a detrimental effect on viral fitness. However, verification that these mutations lead to increased fitness would require infection using tissue-isolated MMTV virions, and analyzing replication kinetics over several rounds of replication. MMTV is quite sensitive to small variations in *in vitro* conditions, and in my experience, difficult to maintain in culture for multiple replication cycles (Fine, et al., 1976). Thus, these experiments were not performed.

Chapter 3.4.2: MMTV is restricted by therapeutic antiretrovirals *in vitro*

Sequence analysis of HIV genomes isolated from patients who have failed anti-viral drug therapy suggests that human APOBEC3-mediated G-to-A transitions could play a role in the development of drug-resistant virus. To directly test if such selection occurs *in vivo*, we wanted to test the effect of antiretroviral drugs to MMTV-infected APOBEC3+/+ and APOBEC3-/- mice. We tested several drugs that are used to treat HIV-infected patients to determine their efficacy against MMTV *in vitro*. Virus isolated from APOBEC3-/- mammary tumors was used to infect 293T-mTfR1 cells (expressing the MMTV receptor, mouse transferrin receptor 1) in the presence or absence of several different antiretroviral drugs. Infection levels were measured after 24 hours using quantitative PCR to measure total integrated exogenous MMTV. At the concentrations selected, the thymidine analog AZT was most effective, reducing viral integration approximately 100-fold (Fig. 3.2). Raltegravir was also very effective, while lamivudine was able to reduce infection but with the lowest efficacy of those tested. Our results confirm previous studies that support AZT as active against MMTV infection in a variety of cell types *in vitro* (Indik, 2007;

Okeoma, 2009), and provide evidence for additional antivirals that are efficacious for betaretroviruses. This information led us to select AZT for *in vivo* viral evolution studies, as it is clearly potently antiviral.

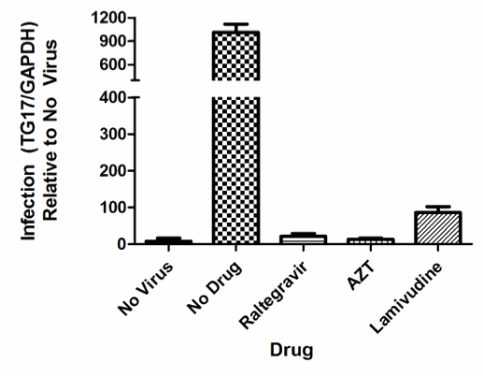


Fig.3.2. Antiretrovirals effectively block MMTV *in vitro*. 293T-mTfR1 cells were infected with MMTV RIII in the presence or absence of antiretroviral drug. Total integrated exogenous MMTV was measured via qPCR one day post-infection. Infection levels are normalized to total cellular DNA using GAPDH as an internal control. Infection was performed in duplicate wells and results averaged; results are representative of two replicates.

Chapter 3.4.3: Selection of antiretroviral drug and dosing for studies of *in vivo* viral evolution

An effective agent for *in vivo* viral evolution studies should have several characteristics. First, it should be highly effective in preventing viral replication. When using an antiretroviral as a selective pressure against the virus, one measurable outcome is viral replication. Without viral evolution and escape, viral replication remains low. However, if virus is able to evolve resistance to the drug, replication will increase even in the presence of drug. It is more difficult to draw clear conclusions using viral replication as a measure if using an ineffective drug that allows continuous viral replication.

Dozens of mutations in HIV are known to cause resistance against the activity of AZT. These mutations are collectively known as “thymidine analog mutations” or TAMs. TAMs are unique sets of mutations within the gene encoding the viral reverse transcriptase, which cause the RT to avoid the effect of the drug, at a potential fitness cost to the virus. TAMs typically occur in sets, with initial mutations that decrease drug susceptibility selected first, and followed by additional compensatory mutations that reduce any deleterious effect of the resistance mutations (Garcia-Lerma, 2005). In measuring APOBEC3-mediated viral evolution, drugs that require a “multiple hit” mechanism, needing several mutations to fully develop resistance, is potentially beneficial over drugs that require a single mutation. APOBEC3 creates only G-to-A mutations, but can cause several G-to-A changes in a single virus. Thus, it is useful to have a variety of targets that could potentially cause drug resistance. One potential downfall is that resistance may take longer and require more rounds of replication to arise. Furthermore, it is possible that some TAMs are comprised solely of amino acids mutations that do not require G-to-A mutations at the nucleotide level. In the presence of drug, these TAMs could arise with or without any additional APOBEC3 mutations, creating a potential source of false positive or false negative results.

Dosing of the chosen drug is also critical for optimizing the possibility of obtaining clear results. As stated previously, too low of a drug dose may allow continued viral replication. This creates an ineffective system, as the virus has little selective pressure to evolve resistance mutations. Further, if resistance mutations do occur, our measure of viral replication to detect

these may fail. Observing an increased rate of viral replication of a drug resistant virus is more difficult to distinguish if viral replication is significant even in the absence of drug. However, suboptimal doses of drug are known to cause more rapid evolution of antiviral drug resistance (Lucas, 2005). HIV-infected patients with non-perfect drug adherence are significantly more likely to acquire drug resistant virus (Harrigan, 2005). Optimal drug doses in patients cause complete cessation of viral replication, and prevents the opportunity for the virus to acquire resistance mutations, through APOBEC3-mediated cytidine deamination, reverse transcriptase errors, or RNA polymerase errors. Thus, in our system, drug should be administered at approximately 90-95% the maximal inhibitory concentration, to establish and maintain a low level of viral replication (to enable mutations to occur) but allow approximately a 10-fold difference between treated and untreated mice (to increase ability to observe rapid changes in viral kinetics). We anticipate that this would decrease time to development of drug resistance.

The IC_{90} value for AZT to treat HIV is approximately 5 μM. However, current recommended dosage of the drug in humans is 600 mg (approximately 8.8 mg/kg assuming 150 lb individual) daily. Further assuming complete bioavailability, serum concentrations with these doses could reach 450 μM, significantly higher than the necessary dose. In MMTV studies, 10 μM was sufficient to completely inhibit viral replication *in vitro* (Indik, 2007). *In vivo*, adult C57BL/6 mice injected with 25 mg/kg of AZT achieved serum concentrations of approximately 50 μM, although levels dropped to 5 μM by 90 minutes after injection (Chow, 1998). In female $B6C3F_1$ mice, doses of 15, 30, and 60 mg/kg given orally led to maximum serum concentrations of 34, 67, and 187 μM respectively (Trang, 1993). The half-life of the drug is also relatively short, ranging from 16-46 minutes (Trang, 1993). However, dosage experiments for inhibiting MMTV have not been performed *in vivo*. Considering the high bioavailability but low half-life, we provided drinking water with 0.25 mg/mL AZT to obtain relatively consistent plasma levels. Assuming average weight of 20 g and consumption of 3 mL of drinking water per mouse per day, the average dose was approximately 40 mg/kg per day.

Chapter 3.4.4: The effects of antiretroviral treatment and APOBEC3 on MMTV replication

MMTV-infected APOBEC3+/+ and APOBEC3-/- mice were treated orally with 40 mg/kg (40 μg/g) of AZT beginning at four weeks after birth, after a single initial dose of 100 mg/kg delivered via intraperitoneal injection. The oral route of drug administration prevented treatment immediately after MMTV infection, which occurs shortly after birth once the pup begins to consume virus-containing breast milk. Thus, drug treatment began once mice were weaned and began consuming solely water. Untreated mice were housed in separate cage and provided drug-free water to serve as controls. However, as infection began considerably earlier than treatment, all mice began with significant initial viral loads. To control for this, we extracted blood and isolated PBMCs upon weaning (at the start of treatment) to establish a baseline level of infection for each mouse. This was measured via level of integrated virus within this target cell population. MMTV does not produce plasma viremia, so cell-associated virus is an appropriate measure of infection (Finke, 2001). Drug treatment continued for 10 weeks, and mice were bled every 3-4 weeks (0, 3, 6, and 10 weeks). After the final bleed, mice were sacrificed and splenocytes harvested for viral integration analysis on this distinct target cell population.

At each time point, PBMCs were isolated from whole blood, and total DNA harvested for PCR analysis. Measuring integrated exogenous virus proved difficult. Single-round PCR or quantitative PCR was not sufficient to amplify measurable quantities of viral DNA from PBMCs. Thus, we utilized a nested PCR-qPCR system to measure comparative levels of integrated virus. Equal concentrations of DNA were first used to amplify a 3 kB region corresponding to the 3' end of the genome, using one MMTV-RIII specific primer to ensure endogenous viruses were not amplified. Equal volumes of the PCR products were then used in a nested quantitative PCR reaction, using the same MMTV-RIII specific primer to again ensure specificity and prevent amplification of endogenous virus from carryover DNA. Multiple dilutions of several test samples were run to ensure amplification was in linear range (not shown).

At the first time point, prior to AZT treatment for any group, average integrated MMTV levels differed based on mouse genotype. APOBEC3+/+ mice displayed a wide range of

infection levels, which may reflect differences in initial infection based on nursing habits and mother-to-mother disparities (Fig. 3.3). APOBEC3-/- mice maintained significantly higher viral loads, approximately 10-fold higher than APOBEC+/+ mice on average (t-test, $p<0.01$). Viral loads in these mice had less variability, perhaps an indicator of viral loads beginning to plateau (Fig. 3.3). Both the difference in viral levels in APOBEC3+/+ vs APOBEC3-/- mice and the potential plateau in APOBEC3-/- mice is consistent with previous results (Okeoma, 2007). However, these studies did not examine PBMCs directly.

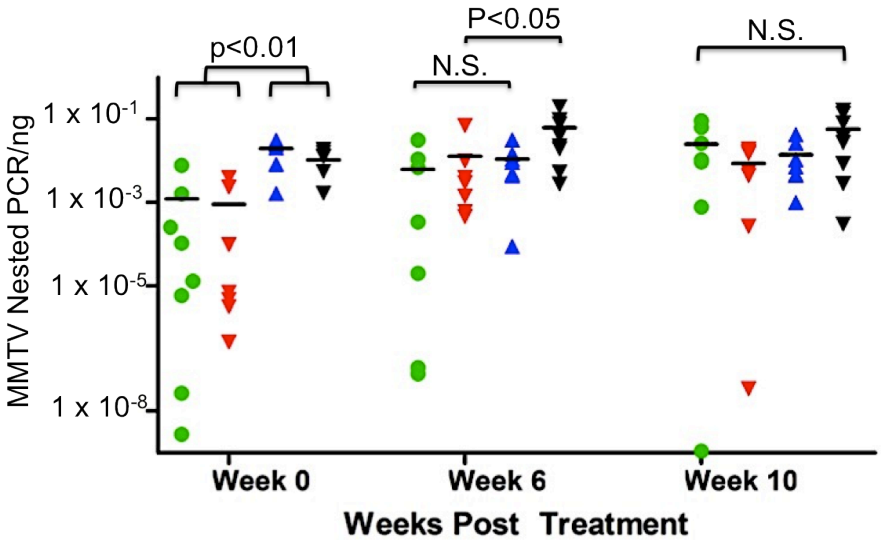


Fig. 3.3. The dual effect of APOBEC3 and AZT on MMTV replication *in vivo*. MMTV-infected APOBEC3+/+ and APOBEC3-/- were treated with the antiretroviral AZT beginning at 4 weeks post-infection. Blood was collected at various time points post-treatment, including a 0 week timepoint to establish baseline MMTV infection. Viral load was measured by nested qPCR of PBMC DNA using primers specific for MMTV RIII. T-tests were performed to determine statistical differences between indicated groups. Each data point represents a single animal. The two outlier data points at Week 10 may have been the result of an inefficient DNA isolation, as low levels of amplification were seen in multiple PCR reactions. Green dots: AZT-treated APOBEC3+/+ mice; Red triangles: untreated APOBEC3+/+ mice; Blue triangles: AZT-treated APOBEC3-/- mice; Black triangles: untreated APOBEC3-/- mice.

After 6 weeks of AZT treatment, APOBEC3+/+ mice maintained a high level of variability in viral loads. Several mice were able to control virus to levels similar to their respective baseline (Fig. 3.3). Untreated APOBEC3+/+ mice showed a significant rise in viral load, although not reaching statistical significance over treated mice. This rise is consistent with the hypothesis that APOBEC3 controls MMTV early in infection, but is unable to completely prevent viral replication. The early kinetics of viral replication were slower than that in APOBEC3-/- mice. However, AZT-treated APOBEC3-/- showed no discernable change in viral load from 3 weeks post-infection to 9 weeks post-infection (Fig. 3.3; comparing Week 0 to Week 6). This could indicate either an efficacious antiretroviral treatment, or a viral load that has plateaued. However, viral loads in untreated APOBEC3-/- mice continued to rise, becoming significantly higher than in treated counterparts ($p<0.05$).

After an additional four weeks of infection and treatment, viral loads in all four groups reached similar levels. Untreated APOBEC3+/+ mice and both groups of APOBEC3-/- mice showed no significant change in viral load from the previous timepoint. Interestingly, the group that differed the most from the experimental midpoint to endpoint was AZT-treated APOBEC3+/+ mice, which saw a marked increase in almost every single member of the group (Fig. 3.3). These results again could be interpreted in two different manners. First, the AZT dose may have been ineffective in controlling MMTV, which provided a second mechanism of slowing replication together with APOBEC3. This control would slow the progression of viral integration events, but eventually viral loads would reach the same putative plateau as APOBEC3-/- and untreated mice. An alternative explanation suggests that AZT in fact was very effective (particularly bearing in mind Week 0 to Week 6 similarities in AZT-treated APOBEC3+/+ mice), but viruses within these mice evolved resistance to the drug. This drug resistance would allow increased replication even in the continued presence of AZT, raising viral loads to similar levels as other experimental groups. This would reflect a potential role for APOBEC3 in promoting retroviral drug resistance *in vivo*.

Chapter 3.4.5: Barriers to secondary analysis of exogenous MMTV variants

If APOBEC3 did cause drug resistant variants to arise in AZT-treated mice, we would expect a higher level of integrated virus with G-to-A mutations within the reverse transcriptase gene. Follow up studies included an attempt to amplify and sequence regions of *pol* that correspond to reverse transcriptase. However, multiple factors prevented effective analysis of integrated exogenous virus. First, as stated previously, amplification of integrated exogenous virus required consecutive rounds of nested PCR. Endogenous and exogenous MMTV differ in very few regions of their genomes, with the most variation occurring in the 3' LTR. No region of *pol* has significant differences in RIII versus all of the endogenous MMTVs found in C57BL/6 mice (*Mtv-8*, *Mtv-9*, and *Mtv-17*, {Peterson, 1985}), although there are a number of single nucleotide differences. This prevents the design of *pol* primers that amplify exogenous RIII over these endogenous viruses. Thus, while we could use an RIII-specific primer for the first round of PCR, second round amplification of the shorter *pol* fragment used primers that could amplify all species of MMTV present. Even with dilution of the first round products, a small amount of second round product was formed even in control samples using uninfected PBMC DNA. This suggests that in our infected experimental samples, contamination with endogenous sequence may exist.

Infection of all mice continued for four weeks prior to beginning AZT treatment. Our quantitative data show that a high baseline infection was established in all groups at this point, although there was a significant difference between APOBEC3-/- and APOBEC3+/+ mice. Unlike in HIV, MMTV infection of target cells within the PBMC population does not cause rapid cell death. A large fraction of cells infected during this initial untreated infection period are thus still present in the samples drawn at later time points. Even if amplification of viral *pol* was completely specific for exogenous MMTV, a significant portion of the amplified virus would be pre-treatment virus, which may or may not carry any APOBEC3-mediated G-to-A mutations. This would make classification of individual viral clone sequences as drug-susceptible or drug-resistant very difficult, particularly without high throughput deep sequencing and computational analysis.

Further complicating factors include the ability for exogenous MMTV to create recombinant viruses with endogenous viruses. Although these endogenous viruses cannot form functional infectious virions, they can be expressed at the RNA level and packaged into virions alongside infectious exogenous viral RNA genomes. Upon infection of a new cell, reverse transcriptase can switch between the two templates to create a recombinant virus that may or may not be infectious (Golovkina, 1994). During initial attempts to sequence exogenous RIII from our PBMC samples, many samples appeared to contain a large percentage of recombinant viruses. This recombination alone could result in drug-resistant viruses, which while noteworthy in its own right, would be an APOBEC3-independent event and potentially confound the search for any novel APOBEC3-mediated drug resistance mutations.

Chapter 3.4.6: MLV is restricted by therapeutic antiretrovirals *in vitro* and *in vivo*

Many of the difficulties we experienced in our MMTV studies could be ameliorated using a different murine retrovirus, Moloney MLV. Mice are naturally infected with MLV through breast milk similar to MMTV, but neonates can also be infected to high titer by intraperitoneal injection with purified virus. This purified MLV is produced *in vitro*, and virus can be titered to ensure mice are infected with the same number of infectious virus particles. MLV causes plasma viremia, which is an easier measure of infectious virus load than total integrated viral DNA, since unlike integrated proviral DNA, it reflects the actively replicating virus pool at any given time. Additionally, plasma-isolated virions can be tested for infectivity *in vitro*. Laboratory and wild mice do contain endogenous MLV viruses similar to M-MLV, and recombinant viruses produced between the two can be infectious. However, there are more regions of dissimilarity between the endogenous viruses and M-MLV than between endogenous and exogenous MMTV. These differences allow specific primer design for easier amplification of exogenous virus without amplifying endogenous viruses from genomic DNA.

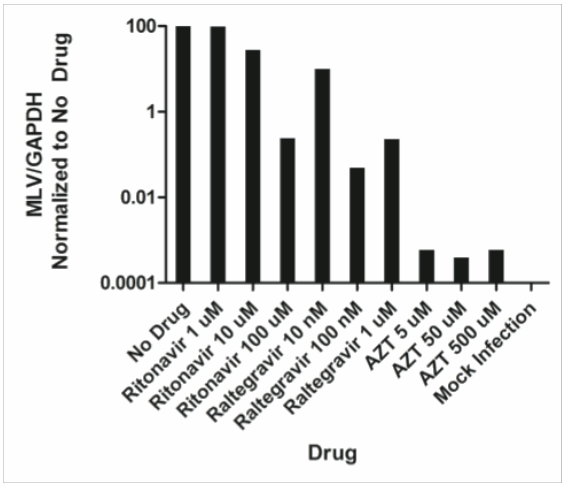
M-MLV reverse transcriptase is a well-studied enzyme, with a known crystal structure. This structure shows significant structural homology to HIV-1 RT, although they share only 25%

sequence identity on the amino acid level (Das and Georgiadis, 2004). Main differences between the structures include regions of the thumb, connection, and RNase domains. Some of this structural difference is attributed to the fact that MLV RT is a 75 kDa monomer and HIV-1 RT is a heterodimer consisting of p66 and p51 subunits. The palm and finger regions, which include the polymerase active site and many catalytically essential residues, have more significant structural and sequence similarities (Das and Georgiadis, 2004). These regions are also important for nucleic acid binding, and in HIV-1 RT, contain the majority of mutations known to induce resistance to reverse transcriptase inhibitors such as AZT (Rhee, 2003). Our analysis focused on this region of MLV *pol*.

Because of the dissimilarities between MLV and HIV-1 reverse transcriptase, we first tested whether AZT inhibits MLV *in vitro*. In order to maintain a similar model to the system we used for MMTV, our goal was to use this drug as a selective agent to induce viral evolution. We simultaneously tested other antiretrovirals for their potential use as well. We tested the integrase inhibitor integrase and the protease inhibitor ritonavir. All three drugs inhibited MLV at concentrations within similar range to those used to treat HIV in humans (ritonavir: 30μM, raltegravir: 3000 nM, AZT: 50 μM [van der Lugt, 2009; Sandkovsky, 2012; Bergshoeff, 2004) (Fig. 3.4).

Because all concentrations of AZT tested were able to prevent infection, we tested susceptibility to a greater range of the drug to determine more exact IC_{50} and IC_{90} concentrations *in vitro*. We found that MLV was susceptible to much lower doses than previously tested, with IC_{50} approximately 0.03-0.04 μM (Fig. 3.4). We believe this sensitivity to AZT, which is within a similar range as the IC_{50} for HIV, makes AZT an appropriate choice for *in vivo* MLV studies, as we are easily able to reach these inhibitory concentrations *in vivo*. We also chose to continue experiments with AZT because large amounts were available at low cost, an important factor in designing long-term experiments involving many animals.

A.



B.

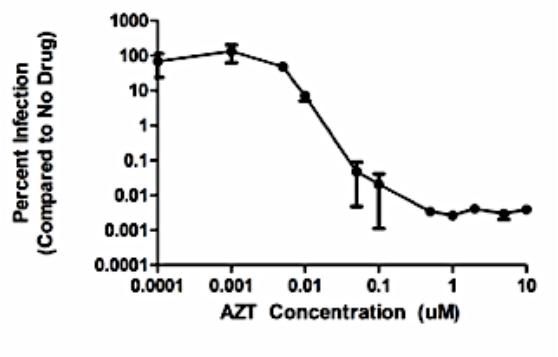


Fig. 3.4. M-MLV is susceptible to antiretroviral drugs *in vitro*. (A) Increasing concentrations of the protease inhibitor ritonavir, the integrase inhibitor raltegravir and the reverse transcriptase inhibitor AZT were added to 3T3 cultures during infection with M-MLV. Infection levels were measured by quantitative PCR of total cellular DNA 24 hours after infection, and normalized to GAPDH. Values were set as proportion of infection in cells not treated with any drug. Each experimental condition was performed in duplicate, and the average value is shown. (B) A wider range of AZT concentrations were used in similar MLV infections to generate an *in vitro* dose curve. Each concentration was applied in triplicate, and levels of infection averaged and normalized to no drug controls.

We next tested viral susceptibility to AZT *in vivo*. As stated previously, drug dosage is critical when used as an evolutionary pressure. To avoid high baseline infection rates seen in MMTV studies, we allowed infection to proceed for only 48 hours before beginning treatment. Mice were also treated using a different mode of delivery, via daily intraperitoneal injection with liquid AZT instead of adding lyophilized AZT to drinking water. Mice were given three different concentrations, which were in approximately the same range (in mg/kg) as that given in the MMTV experiments. Total integrated MLV in PBMCs and splenocytes and infectious plasma viremia were measured to assess viral load. A dose-response effect was seen in this range, confirming that the range chosen was appropriate (Fig. 3.5). Mice treated with 0.05 mg/g AZT had a level of infection that was approximately 1-5% of untreated mice. This suggests a low level of replication, which is necessary for the viral population to obtain and maintain diversity. However, the difference in infection level between this treatment group and untreated mice is sufficient to easily discern, which is important for determining a change in viral load within a single mouse.

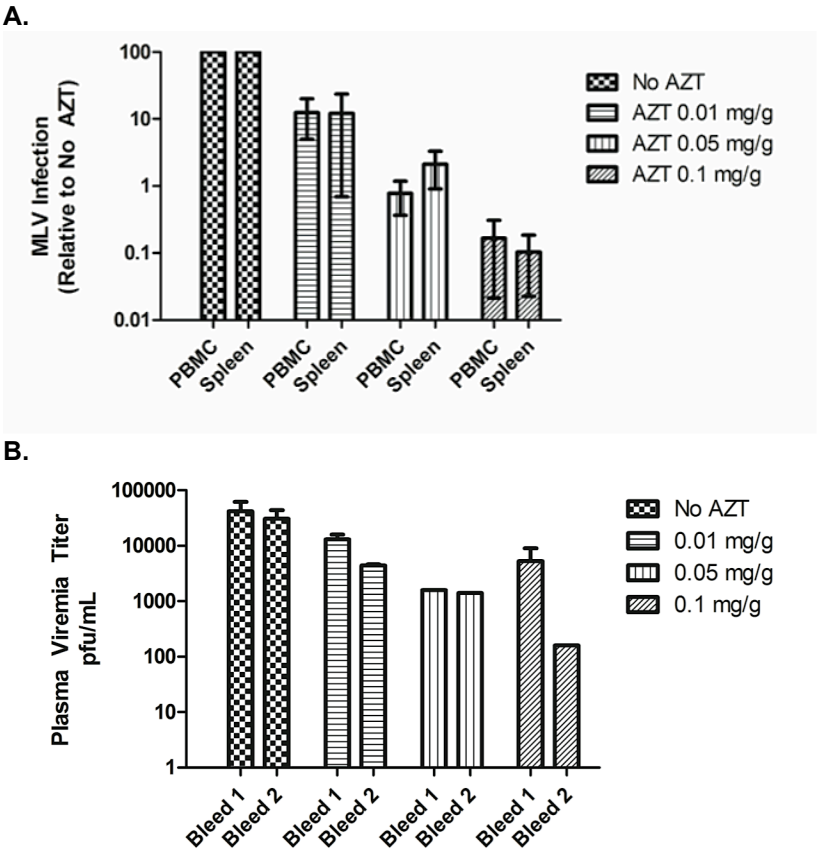


Fig. 3.5. M-MLV is susceptible to AZT *in vivo*. C57BL/6 mice were infected with M-MLV 24-48 hours after birth. Daily intraperitoneal injections of the indicated dose of AZT were given beginning two days post-infection. (A) Two weeks post-infection, mice were bled and total MLV DNA was measured by quantitative PCR from isolated PBMCs and normalized to GAPDH levels. Three weeks post-infection, mice were sacrificed and splenocytes were isolated for similar quantitative analysis. Infection levels are normalized to untreated mice, which are set as 100%. Each group contained three animals, and data are presented as averages with standard error. ANOVA analysis indicates a significant difference between each group ($p<0.01$). (B) Plasma viremia was calculated at both two weeks and three weeks post-infection. Plasma virus was serially diluted and used to infect 3T3 cells. Four days post-infection, cells were stained with an anti-MLV env antibody and fluorescent plaques counted. Average plaque forming units per mL of serum per treatment group are shown. ANOVA analysis indicates a significant difference in viral inhibition between each dose at Bleed 1 and Bleed 2 ($p<0.01$).

Chapter 3.4.7: MLV is restricted and hypermutated in human APOBEC3G-expressing mice

Spiros Stavrou, a postdoctoral fellow in the lab, has recently created transgenic mice that express human APOBEC3G in the absence of murine APOBEC3. These mice express APOBEC3G under a chicken ß-actin promoter, which supports universal expression. APOBEC3G RNA and protein can be detected in M-MLV target tissues (PBMCs, spleen, and thymus), as well as other tissues. APOBEC3G transgene expression by murine PBMCs was about 10-fold lower than that found in human PBMCs. This level was sufficient to affect MLV replication *in vivo*, with transgenic mice carrying approximately 7-8 fold less infectious plasma virus than APOBEC3-/- mice two weeks post-infection (Stavrou, et al., in preparation). The human protein dramatically inhibits infection, and compared to endogenous mAPOBEC3, also causes substantial G-to-A mutations of the *env* region of the M-MLV genome (Stavrou, et al., in preparation). This corroborates previous experiments, which showed that human APOBEC3 is an efficient anti-MLV factor *in vitro* (Mariani, 2003). Because these human APOBEC3G transgenic mice showed higher levels of deamination of the M-MLV genome without completely inhibiting replication, we chose to examine whether human APOBEC3G is capable of generating AZT-resistant M-MLV.

We first tested whether the *pol* gene of MLV was targeted for cytidine deamination by transgenic APOBEC3G. A region of *pol* corresponding to reverse transcriptase was amplified from splenic DNA and from plasma RNA of MLV-infected APOBEC3-/-, APOBEC3+/+, and APOBEC3G-expressing mice, and the DNA was clonally sequenced and assessed for mutations. Virus recovered from transgenic mice showed significantly higher levels of G-to-A mutations than virus isolated from either other genotypes (Table 3.2A). Integrated viral DNA has been shown to harbor more mutations than viral genomes packaged into newly formed virions (Russell, 2009; Sato, 2010). We thus also sought to determine the rate of G-to-A mutations in packaged MLV genomes isolated from cell-free virions. Although we also saw a decrease in mutation rate compared to integrated proviruses, cell-free virions from mice expressing the human APOBEC3G

transgene had significantly more mutations within their genomes than those from APOBEC3+/+ or APOBEC3-/- mice (Table 3.2B). Mutation analysis of regions of *env* and *gag* also show higher rates of G-to-A mutation in APOBEC3G-expressing mice, with an apparent gradient that increases 5' to 3' (Table 3.2C). This is consistent with previous reports that indicate a gradient of mutation that reflects the time the minus-strand DNA remains single stranded (Suspene, 2006; Jern, 2007).

Table 3.2. Human APOBEC3G deaminates M-MLV *in vivo*. APOBEC3-/-, APOBEC3+/+, and transgenic APOBEC3G-expressing mice were infected with MLV 1-2 days after birth. Splenocytes were harvested 16 days post-infection. (A) Total DNA was extracted from a portion of the splenocyte population, and used to amplify a region of *pol* corresponding to integrated M-MLV reverse transcriptase. These amplicons were clonally sequenced and assessed for G-to-A mutations. (B) A separate fraction of splenocytes was cultured for 48 hours *in vitro*, and virus purified from the resulting supernatant. Similar amplification and sequencing analysis was performed on RNA isolated from these viruses. ANOVA tests were performed to assess statistical differences; *= $p<0.05$. (C) Two other regions of the MLV genome were analyzed via clonal sequencing of integrated proviruses within MLV-infected splenocytes two weeks post-infection. G-to-A mutations falling within the APOBEC3G-preferred editing site 5'-GG-3' were calculated for each of three regions of the genome. Figure credit: Spiros Stavrou.

A.

Genotype	G-to-A mutations	Total bp Analyzed	Clones Analyzed	Mutations per kB
APOBEC3-/-	2	16050	21	0.12
APOBEC3+/+	1	24480	31	0.04
APOBEC3G	9	16050	21	0.56*

B.

Genotype	G-to-A mutations	Total bp Analyzed	Clones Analyzed	Mutations per kB
APOBEC3-/-	2	16830	22	0.12
APOBEC3+/+	2	17610	23	0.11
APOBEC3G	5	11670	15	0.42*

C.

Mutation frequency (G to A/total GGs sequenced)			
	gag	***pol***	***env***
APOBEC3G	**.00063**	**.0041**	**.0157**
APOBEC3-/-	**0**	**0**	**.0005**
APOBEC3+/+	**0**	**0**	**0**

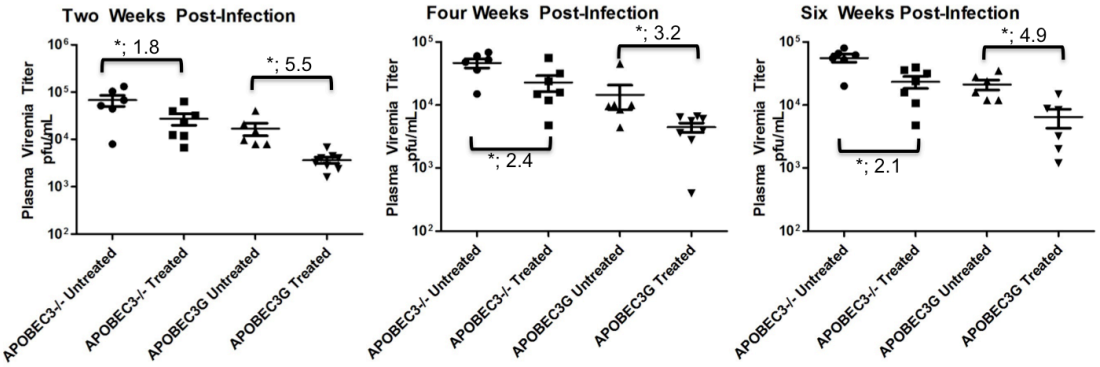
Chapter 3.4.8: Antiretroviral treatment is more effective in APOBEC3G-expressing mice

To study the role of APOBEC3 and viral evolution in the presence of antiretroviral drugs, we infected APOBEC3-/- and APOBEC3G-expressing mice with M-MLV and began treatment with the chosen sub-optimal dose of AZT two days post-infection. During the course of infection, we periodically measured viral loads to determine any large changes in viral load within individual mice. A large increase in infection within an AZT-treated mouse previously able to control infection could indicate the emergence of a drug-resistant virus. At two, four, and six weeks post-infection, blood was collected and both whole plasma and total PBMC DNA was collected. Plasma viral levels were measured by titration on 3T3 cells, and PBMC DNA was analyzed for levels of integrated MLV by quantitative PCR. At two weeks post-infection, significant differences were seen in plasma viremia based on AZT treatment within each genotype (Fig. 3.6A). ANOVA analysis also confirmed statistical differences between each of the four groups at all three time points, confirming that both AZT treatment and genotype both have a significant effect on viral load. Statistical differences were less apparent at the level of integrated provirus, as APOBEC3G mice, but not APOBEC3-/- mice, showed a significant decrease in integration when treated with AZT (Fig. 3.6B). This could be due to the more sensitive nature of the quantitative PCR assay, as it is better able to detect minor variability. However, general trends remain consistent with plasma viremia data, in which mice receiving AZT treatment or expressing APOBEC3G seem to have lower levels of integration. Interestingly, AZT treatment seemed to have a greater effect in mice expressing APOBEC3G, as fold change in plasma viremia between APOBEC3G treated and untreated animals were significantly higher than that seen in APOBEC3-/- treated vs untreated animals. This could indicate a potentially synergistic effect of the restriction factor and the antiretroviral drug.

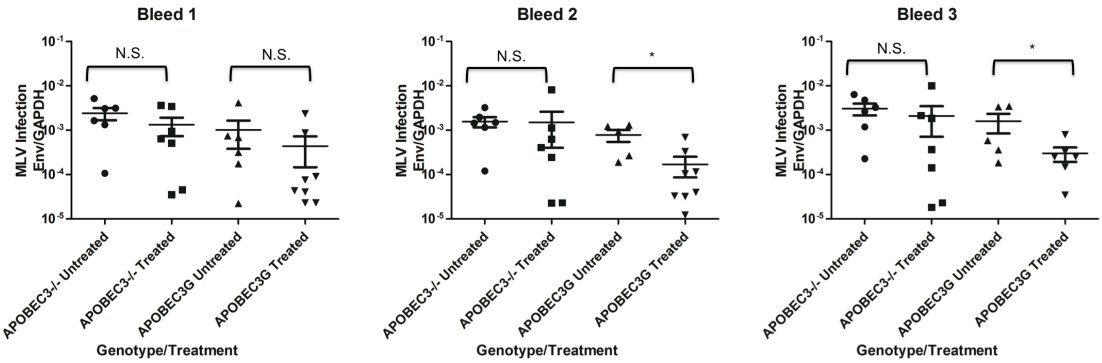
Fig. 3.6. The antiviral effects of APOBEC3G and AZT are distinct early in infection, but become less apparent over time. APOBEC3-/- and APOBEC3G-expressing mice were infected with MLV 48 hours after birth. At specified time points, blood was collected and both

plasma and total PBMC DNA isolated. (A) Plasma viremia was measured by titration of plasma on 3T3 cells. Four days post-infection, cells were stained with anti-MLV env antibodies, and infectious plaques counted. Infectious plaque forming units per mL of plasma were calculated, and plotted per individual mouse. Group averages per time point are shown. T-tests were performed to determine statistical differences between treated and untreated mice within each genotype, at each time point; *= $p<0.01$. The number above each treated vs. untreated group indicates the fold change between the average viremia levels, Untreated/Treated. (B) Total integrated M-MLV was measured by quantitative PCR using primers specific to MLV *env*, and normalized to GAPDH. Statistical analysis was performed as above. (C) Plasma viremia levels for individual mice were plotted chronologically for each experimental group.

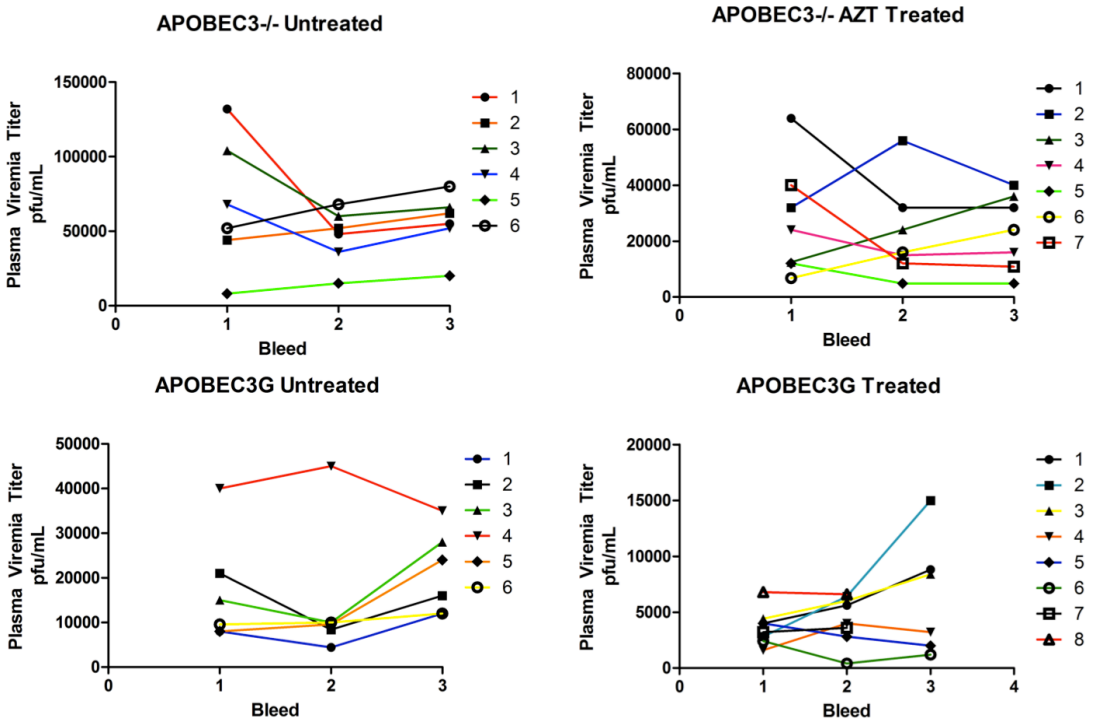
A.



B.



C.



At later time points, infection levels in all groups of mice began to reach comparable levels. This is similar to the effect we saw in MMTV-infected mice (Fig. 3.3). On average, infection levels in untreated mice continued to gradually increase over time as expected, but levels in treated mice also increased on average. As drug was provided at a sub-optimal level, this increase is not unexpected, as the dose may not have completely inhibited viral replication. If drug-resistant virus arose, we might expect a larger rise in viral load over that in other similarly treated mice. Although on average, viral loads increased in all groups, AZT-treated APOBEC3G expressing mice showed the greatest fold increase in viral load between the first and last bleed, which was influenced by virus in three specific mice (Figure 3.6C). Two mice in our APOBEC3-/- treated group also showed over two-fold increase in viral load between the first and the last bleed, while mice in both untreated groups showed less substantial increases. This suggests that the pressure exerted by AZT may be sufficient to induce evolution of drug resistance.

Chapter 3.4.9: Human APOBEC3G affects MLV diversity *in vivo*

We next assessed the effect of both APOBEC3G expression, as well as AZT-treatment induced selection, on viral *pol* nucleotide sequences. Plasma virus cDNA was isolated from a few selected samples at two weeks post-infection. Reverse transcriptase was amplified and clonally sequenced to assess G-to-A mutation levels. In those samples assessed, *pol* sequences from APOBEC3G-expressing mice did not show significantly increased mutagenesis compared to APOBEC3-/- viral sequences at this early time point (Table 3.3A). This may have been a result of our small sample size of mice for this analysis, as our APOBEC3-/- mice showed higher levels of G-to-A mutation than expected, and our APOBEC3G-expressing mice showed lower levels than expected. Viruses circulating in blood are representative of the pool of infectious M-MLV and highly mutated viruses would not be found in this pool. To examine the overall extent of deamination in the AZT-treated mice, we also assessed integrated viral DNA from selected samples at two weeks post-infection. Integrated proviruses isolated from AZT-treated or untreated APOBEC3G-expressing mice contained slightly less G-to-A mutations as those previously sequenced (Table 3.3B; compare to Table 3.2). However, these viruses did maintain significantly higher levels of G-to-A mutations than integrated proviruses sequenced from APOBEC3-/- littermates.

Plasma virions collected at six weeks post-infection were next evaluated for nucleotide changes in *pol*. At this later time point, there were more G-to-A mutations apparent in circulating plasma virus of APOBEC3G-expressing mice (Table 3.3C). Although mutation rate had considerable differences between mice, overall G-to-A mutations on average were significantly higher on in APOBEC3G-expressing mice than APOBEC3-/- mice, irrespective of treatment status (t-test, $p<0.05$). This is indicative of significant mutagenesis, and implies that APOBEC3G is able to influence the viral genome in a relatively short time. G-to-A mutagenesis was not significantly different between AZT-treated and untreated APOBEC3G-expressing mice (t-test, $p=0.15$). Total rates of mutation (including G-to-A) were also not different between AZT-treated and untreated mice, regardless of genotype, suggesting that at least in the region of *pol* assessed, AZT was not sufficient to influence statistically significantly higher levels of total viral

nucleotide diversity within this region. However, in many cases, multiple viruses cloned from an individual mouse carried the same nucleotide change, which could indicate a selective advantage provided by these mutations (see Figure 3.8).

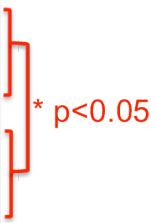
Table 3.3. APOBEC3G affects MLV genomic evolution *in vivo*. APOBEC3G-expressing and APOBEC3-/- mice were infected with MLV and treated with sub-optimal doses of AZT to stimulate development of drug resistant variants. Virus was isolated from either PBMCs (integrated proviruses) or plasma (genomic RNA) at multiple time points as indicated, and *pol* fragments clonally sequenced. (A) Plasma virus was isolated from selected mice at two weeks post-infection, and viral RNA isolated and reverse transcribed for *pol* sequence analysis. No AZT-treated APOBEC3-/- mice were analyzed at this time point. (B) Integrated proviruses were amplified and sequenced from selected mice at two weeks post-infection. The number of all mutations was recorded, and G to A mutations were normalized to total kB sequenced. Mutation levels of mice of the same genotype were combined as shown in red, and a t-test was performed to determine statistical differences. (C) Similar analysis was performed for plasma virions isolated at six weeks post-infection. Multiple mice from each group were assessed at this time point. A t-test was again performed as described above.

A. Circulating plasma virus RNA, 2 weeks post-infection

Genotype/ Treatment	# of Mice	G to A	Other Mutations	# of Clones	Total bp	G to A mutations per kB
APOBEC3-/- No AZT	1	5	18	38	29130	0.17
APOBEC3G +AZT	3	14	35	114	89780	0.16
APOBEC3G No AZT	1	1	10	26	20800	0.05

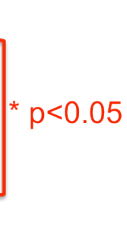
B. Integrated proviral DNA, 2 weeks post-infection

Genotype/ Treatment	# of Mice	# of Clones	Total bp	G to A	Other Mutations	G to A mutations per kB
APOBEC3-/- +AZT	3	52	39000	2	11	0.05
APOBEC3-/- Untreated	1	20	15000	2	2	0.13
APOBEC3G +AZT	2	39	29250	5	10	0.17
APOBEC3G Untreated	2	36	27000	7	7	0.26



C. Circulating plasma virus RNA, 6 weeks post-infection

Genotype/ Treatment	# of Mice	# of Clones	Total bp	G to A	Other Mutations	G to A mutations per kB
APOBEC3-/- +AZT	5	207	159800	36	76	0.23
APOBEC3-/- Untreated	6	267	184300	34	123	0.18
APOBEC3G +AZT	5	218	161800	59	86	0.37
APOBEC3G Untreated	5	221	161470	43	72	0.27



We next considered the effect of G-to-A mutations on MLV reverse transcriptase protein in APOBEC3G-expressing mice. The region sequenced represents about 200 amino acids, comprising about 40% of the total RT structure. This region contains the active site regions of the protein, which is known to be a region targeted for RT inhibitor resistance. It excludes the RNase portion of the enzyme, which has fewer known drug resistance mutations. The portion sequenced also has the highest level of homology with HIV-1 RT (Das and Georgiadis, 2004). Many nucleotide changes do not affect the resulting protein product, as codons maintain redundancy, particularly at the third position. Thus, to fully appreciate the consequence of any nucleotide mutation, amino acid sequence changes were assessed. Although we were particularly interested in G-to-A changes, viral diversity can arise from other factors, including reverse transcriptase errors. To determine the additional effect of APOBEC3G-mediated mutagenesis, we examined amino acid changes resulting from any nucleotide change. Almost all viral clones sequenced were evaluated for amino acid mutations, with the exception of a small fraction that did not translate in frame (likely due to sequencing errors at the 5' end of the amplicon). The number of sequences analyzed at the amino acid level is shown in Table 3.4.

Table 3.4. Viral clones assessed for amino acid mutations.

Genotype/Treatment	# of Viral Clones Assessed for Amino Acid Changes/All Clones Sequenced
APOBEC3-/- AZT-treated	200/207
APOBEC3-/-	260/267

Untreated	
APOBEC3G AZT-treated	208/218
APOBEC3G Untreated	213/221

The goal was to determine the frequency and location of amino acid changes, using our sample of viruses as representative of the viral swarm of each individual mouse. Although it would be expected that mice with more nucleotide changes would result in higher levels of amino acid mutations, we did not see a statistically significant difference in total amino acid mutations between APOBEC3G and APOBEC3-/- isolated virions as we did at the nucleotide level. However, there were significantly more G-to-A mediated amino acid mutations within virions isolated from APOBEC3G mice (Fig. 3.7). The highest level was seen in AZT-treated APOBEC3G mice, suggesting that this selective pressure may lead the virus to maintain increased APOBEC3G-mediated genetic diversity within the targeted gene product.

To determine whether APOBEC3G resulted in an increase in stop codons within the viral population, we assessed the total number of stop codons found per group within the sequenced region. We found few bonafide stop codon-producing mutations, with the untreated APOBEC3-/- group actually showing the largest number (6 stop mutations in 184,300 bp). This is somewhat surprising, considering that a mutation of 5'-TGG-3' to 5'-TGA-3', a potential target for APOBEC3, would result in formation of a stop codon. It is possible that viral genomes containing stop codons are transcribed and packaged into budding virions at a significantly lower rate than intact genomes (Russell, 2009).

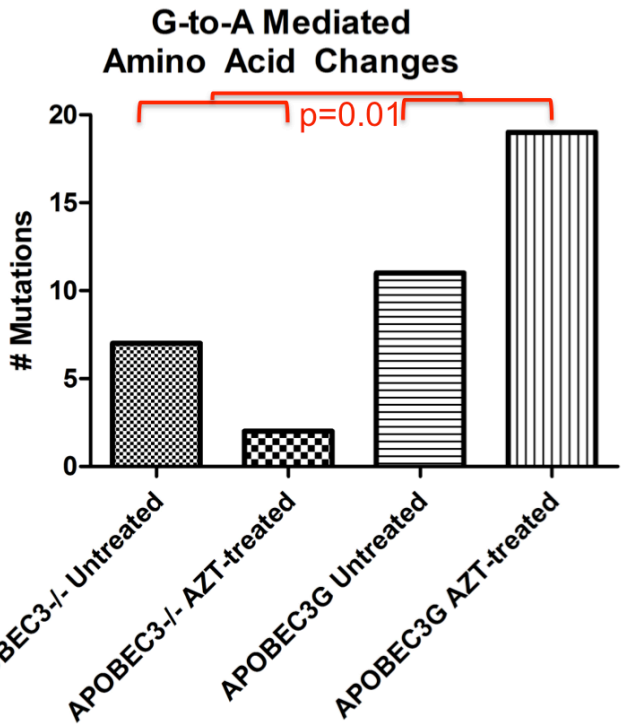


Fig. 3.7. APOBEC3G affects level of viral amino acid diversity *in vivo*. Viruses isolated from APOBEC3G and APOBEC3-/- mice were assessed for G-to-A nucleotide changes compared to the initial viral stock consensus sequence. Viral sequences with G-to-A nucleotide changes were translated to determine the effect on amino acid residues. Total number of G-to-A mediated amino acid changes in all viral clones analyzed per experimental group are shown. A t-test was performed to determine a significant difference between APOBEC3-/- and APOBEC3G-isolated viruses.

We next examined the location of amino acid changes within reverse transcriptase. The region sequenced corresponds to the active site region of the DNA polymerase, including a number of residues with known involvement in nucleic acid binding. Although APOBEC3G-expressing mice did not have an overall increased level of amino acid changes compared to APOBEC3-/- mice, APOBEC3G-induced G-to-A mutations may select for specific mutations in regions more likely to cause drug resistance. All amino acid changes were plotted onto a linear map of the amino acids assessed. Similar regions of the gene were targeted in mice from all four groups (Fig. 3.8).

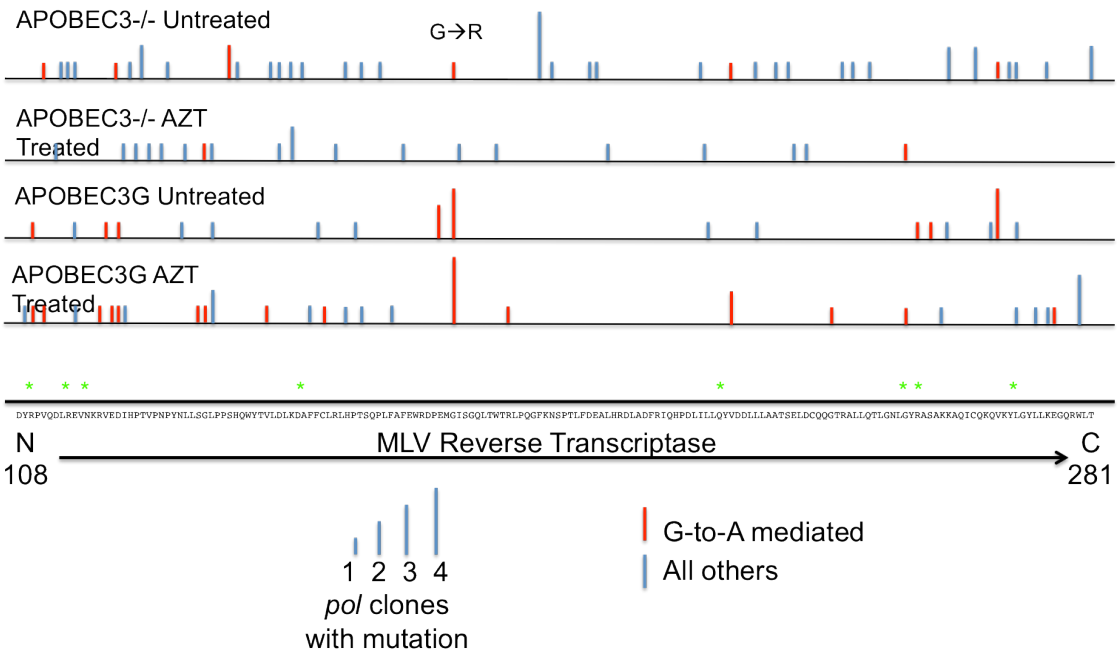


Fig. 3.8. MLV reverse transcriptase amino acid mutations in virions recovered from APOBEC3-/- and APOBEC3G-expressing mice. Plasma virus was isolated from MLV infected mice at 6 weeks post-infection. A portion of *pol* was amplified and clonally sequenced. Nucleotide mutations were annotated, and clones containing mutations were translated. Mutations that led to an amino acid change were recorded, and mapped onto a linear chart of the sequenced region. Each line represents a mutation at the indicated position, and the height of the line indicates how many viruses with this change were found within the experimental group specified. Red lines indicate amino acid changes resulting from G-to-A mutations, and blue lines indicate a change resulting from all other mutations. An APOBEC3G hotspot was targeted in multiple mice, and results in a G→R amino acid change as indicated. * specifies residues known to interact with nucleic acids.

Interestingly, a stretch of amino acids (S180-L217) in viruses isolated from all mice showed very few mutations, potentially indicating a critical region for reverse transcriptase function. This region was also relatively unaffected by mutations at the nucleotide level. Interestingly, this region also contained only a single APOBEC3G residue target (amino acids with 5'-GG-3' within the first two positions of the codon). Although APOBEC3G is able to mutate within other nucleotide contexts other than this preferred target, this perhaps helps explain the lack of mutations in this region in virions isolated from APOBEC3G-expressing mice. Specific residues targeted within AZT-treated and untreated animals were assessed for this APOBEC3G preferred deamination context. One major hotspot was at G178, which contains a G within the following sequence context: 5'-TGGGA-3'. This reflects the preferred editing site of APOBEC3G, and leads to a glycine to arginine mutation. However, this mutation was found at similar levels in both AZT-treated and untreated animals, signifying a lack of selective evolution for this mutation in the presence of drug. This amino acid analysis did not show any mutations found in significantly high levels in any individual mouse or group of mice, nor did it show a region of the gene more heavily targeted by APOBEC3G or in the presence of AZT. However, it does demonstrate the rapid diversification of retroviruses *in vivo*, and provides evidence that this can be augmented by APOBEC3G-mediated deamination.

Chapter 3.5: Discussion

APOBEC3 has been well characterized as an intrinsic antiretroviral factor. Its enzymatic cytidine deaminase activity causes a crippling level of mutagenesis during reverse transcription, preventing establishment of productive infection (Mangeat, 2003; Zhang, 2003; Harris, 2003). In HIV-infected individuals, circulating plasma virus with evident APOBEC3-mediated G-to-A mutations can be found. This presumably sublethal mutagenesis is related to the activity of the HIV accessory protein, Vif. Viruses containing specific Vif mutations that decrease anti-APOBEC3 activity undergo increased APOBEC3 editing (Simon, 2005; De Maio, 2012). However, *in vitro* studies have shown that even with a fully functional Vif, low levels of G-to-A

mutation can be seen in fully infectious virus (Sadler, 2010). Furthermore, in humanized mice infected with wild-type HIV, integrated proviruses with significant G-to-A hypermutation were found (Sato, 2010). In these studies, mutations often led to the formation of errant stop codons; genomes within circulating virions showed significantly fewer mutations, suggesting selection for viral genomes not containing mutations leading to truncated proteins.

Mutagenesis increases viral variability, which helps a viral swarm adapt to new pressures such as antibody mediated immunity or antiviral drugs. APOBEC3-mediated mutagenesis could thus serve the virus as a source of mutagenesis and variability. The presence of G-to-A mutations that do not destroy or significantly diminish infectivity provides principle evidence that APOBEC3 mutations could benefit the virus. Antiretroviral drug resistance is a major problem in HIV treatment. Although drug resistance has significantly declined since the advent of triple therapy, drug non-adherence can still lead to resistance (Clavel and Hance, 2004). Drug resistance develops by viral mutagenesis of specific amino acids necessary for drug efficacy. Some evidence suggests that APOBEC3 could aid the development of drug resistant virus (see Chapter 1.6.6). The absolute contribution of APOBEC3 to viral evolution and consequently drug resistance has not been determined *in vivo*.

We utilized two different models to test the effect of APOBEC3 on viral evolution and drug resistance *in vivo*. Initial studies examined the effect of murine APOBEC3$^{C57BL/6}$-expressing mice on MMTV. Previous work showed that murine APOBEC3 induces very low levels of mutation during single rounds of reverse transcription (see Chapter 2.4.3). However, the MMTV genome has a high A content, which could indicate long term mutagenic pressure from APOBEC3. This mutagenesis may have arisen during the ostensible thousands of years of co-evolution between MMTV and murine APOBEC3. We examined viruses replicating in the long-term absence of APOBEC3, and showed significant sequence reversion to a higher G/C content. Almost all of the viruses with this reversion showed selection for multiple A-to-G mediated amino acid changes within the same region of viral Env, which suggests that this reversion may support

a virus with increased fitness. This also supports present day maintenance of APOBEC3 mutagenic pressure.

To determine the effect of APOBEC3 on viral evolution during a shorter term, we monitored infection levels in MMTV-infected AZT-treated vs. untreated APOBEC3+/+ and APOBEC3-/- mice over a ten-week time course. We showed that AZT treatment was able to block viral replication in mice of both genotypes. This block was maintained in APOBEC3-/- mice over the course of the ten-week treatment. However, in APOBEC3+/+ mice receiving identical doses of drug, viral replication levels significantly increased on average during this time. This suggests that AZT lost efficacy, which may be the result of APOBEC3 mutagenesis and drug resistance. An alternative explanation is that APOBEC3-/- mice reached maximum viral loads by the first time point assessed, and that sub-optimal AZT treatment in APOBEC3+/+ mice allowed for continued viral replication in these mice, allowing overall viral loads to eventually become equal to untreated mice by the last time point assessed. Additional follow up analysis of viral reverse transcriptase changes between endpoint untreated APOBEC3+/+ and AZT-treated APOBEC3+/+ samples, as well as baseline and endpoint viruses from AZT-treated animals would have helped elucidate a role for APOBEC3-mediated mutagenesis on these outcomes.

To overcome some of the experimental challenges of the MMTV system, we next assessed a similar effect of APOBEC3 on retroviral evolution *in vivo* with two important modifications. First, we used Moloney murine leukemia virus. Interestingly, unlike HIV-1 and MMTV, M-MLV does not have a high genomic adenine content, which should make it more susceptible to the editing effects of a cytidine deaminase. Like MMTV, it is restricted by murine APOBEC3 *in vivo* in the absence of viral mutagenesis (Brown and Littman, 2008). However, M-MLV encodes a Gag isoform with a glycosylated N-terminus extension called gGag, which is able to counteract APOBEC3 (see Chapter 1.6.5). The anti-APOBEC3 activity of gGag differs from HIV-1 Vif, which acts to exclude APOBEC3 packaging via proteosomal degradation. Murine APOBEC3 is not excluded from MLV virions, but is still unable to deaminate the virus. Viruses lacking gGag contain viral cores that are less stable compared to cores within viruses expressing

gGag (Stavrou, 2013). It is possible that gGag is able to protect the reverse transcriptase complex from APOBEC3, or that the resulting change in viral stability affects the compartmentalization or release of APOBEC3. Because of this natural resistance to deamination by murine APOBEC3, we used transgenic mice expressing human APOBECG, which is known to cause significant G-to-A hypermutation of MLV *in vitro* (Bishop, 2004).

We showed that the antiretroviral AZT was effective against MLV both *in vitro* and *in vivo*, and performed *in vivo* dosing experiments to determine the optimal dose for viral evolution studies. Using a dose of MLV that allowed approximately 1-5% wild-type viral replication, we showed that AZT and APOBEC3 had a synergistic effect on MLV replication. Both AZT and APOBEC3G restrict virus at reverse transcription, which in turn leads to significantly less integration products. The combination of both is able to substantially prevent new infections, and increasing APOBEC3 activity may prove useful as a therapeutic tool to prevent or inhibit retroviral infections in humans. The most significant increase in viral loads was seen within a few individual mice within the AZT-treated groups, indicating a potential development of mutant escape viruses in these groups. Viral loads in untreated groups also increased on average, but in smaller increments, indicative of continued viral replication and infection of new cellular targets in the absence of the drug.

We sequenced a region of MLV *pol* from *in vivo*-isolated viruses to assess viral diversity. We compared viruses from APOBEC3G-expressing mice in the presence or absence of AZT, a presumed evolutionary pressure, to viruses obtained from APOBEC3-/- mice. We showed that APOBEC3G causes an increase in G-to-A mutations in integrated proviruses early in infection. However, analysis of integrated viruses is not necessarily the most suitable method for analyzing the presence of specific mutations. Potential drug resistance mutations found in integrated proviruses may not necessarily contribute to resistance within circulating virions, if these resistance mutations are found within the only provirus in a cell and the provirus is otherwise lethally mutated.

Of note, the technology utilized also may not provide an accurate display of the actual circulating viruses. Our clonal sequencing efforts may be too insignificant to accurately detect the actual level of mutagenesis. The use of PCR-generated template and transformation of bacterial cells prior to sequencing also provides two additional opportunities for erroneous mutations to arise. Sequencing errors are also a major problem in high-throughput sequencing technology, but samples typically require fewer rounds of PCR prior to sequencing, and samples do not need to be amplified within bacteria. An additional experimental caveat is that the rate of cell turnover after infection with MLV is unknown. MLV infects multiple different cell types, all of which have varying half-lives, and although infection with MLV is not thought to affect cell lifespan, these experiments have not directly been tested within this system. Thus, unlike in an HIV infection in which the infected circulating T-cells represent the most recent rounds of infection, circulating plasma MLV may be produced from the total population of infected cells. If resistance mutations arise later in infection, detecting this resistant population would be more difficult than is the case in HIV-infected patients.

Distinguishing between the mutagenic effects of error-prone reverse transcriptase and APOBEC3 is still critical to determine the relative contribution of both to viral evolution, pathogenesis, and drug resistance. To date, this has been impossible to study *in vivo*, as human samples cannot be experimentally controlled for APOBEC3 expression or initial virus input. At six-weeks post-infection, we found an increased level of G-to-A mutations in plasma virus of APOBEC3G-expressing mice, confirming the ability of APOBEC3 to induce viral diversity during a normal course of infection *in vivo*. Although viruses within APOBEC3-/- mice also showed some viral diversity, the level of nucleotide mutation was significantly higher in APOBEC3G-expressing mice, supporting this antiviral protein as an aid for increasing viral diversity *in vivo*. Interestingly, overall diversity at the amino acid level was not significantly different between APOBEC3G and APOBEC3-/- mice. However, APOBEC3G-expressing mice had significantly higher G-to-A mediated amino acid mutations, demonstrating that APOBEC3G affects the specific population of residues targeted for G-to-A mutagenesis *in vivo*. However, our data shows that APOBEC3G

produces a gradient of deamination, increasing from 5' to 3'. This gradient suggests a lower level of G-to-A mutations within *pol* than in other targets such as *env* or the 3'LTR. Thus, a different selective pressure applied to the virus that targets a gene towards the 3' end of the genome may have a higher chance of being affected by APOBEC3G mutagenesis. Also of note, the rate of viral replication could have an effect on the level of mutations. Increased viral replication allows for additional attack by APOBEC3G, as well as new opportunities for polymerase errors, etc. Our APOBEC3G AZT-treated mice had the lowest level of viral replication, yet still showed a diverse collection of viruses. Thus, APOBEC3 may actually play a larger role in viral diversity than was seen here, if level of nucleotide changes were normalized to number of replication cycles.

In sum, we have demonstrated the effect of APOBEC3G on murine retroviruses *in vivo*. These effects are quite obviously antiviral in nature, in terms of reducing viral load. However, APOBEC3G also increases the rate of viral variability even in a relatively short six-week time period, which could easily aid the virus in escaping immune or drug pressures. Future studies should monitor viral loads and mutagenesis over a longer time period to better understand the long-term effects of APOBEC3. Although we did not see evidence of AZT-resistant viruses in our system, use of different AZT concentrations, other antiretroviral drugs and high throughput next generation sequencing technology may provide better methods for monitoring the influence of APOBEC3G on drug resistance.

Chapter 4: Summary

Chapter 4.1: Discussion

APOBEC3 family members have been well characterized as intrinsic antiretroviral factors. Many previous studies show its potent activity against a large array of viruses *in vitro*, and several murine viruses *in vivo*. However, the large majority of studies have been performed *in vitro*, including investigations into the antiviral mechanisms and the mutagenic effect on viral evolution. Our work has helped elucidate the activity of APOBEC3 *in vivo*.

Chapter 4.1.1: Antiviral mechanisms of endogenous murine APOBEC3 *in vivo*

The mechanism of viral restriction by APOBEC3 is undoubtedly a double-edge sword. As an active cytidine deaminase enzyme, APOBEC3 targets newly synthesized DNA for G-to-A mutations in the coding strand (Mangeat, 2003; Zhang, 2003; Harris, 2003). This mutagenic effect is clearly active *in vivo*, as heavily mutated proviruses can be found in HIV-infected patients, SIV-infected macaques, and HIV-infected humanized mice (Wain-Hobson, 1992; Johnson, 1991; Sato, 2010). Endogenous virion-packaged murine APOBEC3 does not induce cytidine deamination on nascent MMTV reverse transcription products. We have shown that this is not due to an inability for the protein to catalyze the enzymatic reaction. However, significant deamination simply does not occur during reverse transcription even though APOBEC3 is clearly packaged within viral cores and potently affects virus via a cytidine deaminase-independent mode of restriction.

This block to deamination is an interesting phenomenon, as murine retroviruses can be targeted for mutation by APOBEC3 proteins from other species (Bishop, 2004). It is likely that long-term evolution between retroviruses and their host has allowed the more quickly evolving virus to evade the host restriction factors. The evolution of pathogen restriction factors is an important topic in the field, as an understanding of this evolution allows insight into ancestral functions and selective forces, and explains why certain populations are susceptible or resistant to specific pathogens. Analysis of mouse APOBEC3 indicates that it has been in significant genetic conflict during *Mus* evolution, although the specific causes of this conflict are unknown (Sanville, 2010). Interestingly, six of the 10 codons of mAPOBEC3 under the strongest positive selection are within the catalytically active cytidine deaminase domain, five of which encode different amino acids in mouse strains that can or cannot restrict MLV and MMTV. This evolution may have allowed the restrictive form of APOBEC3 to initiate cytidine deamination against ancient retroviruses, but not more recently evolved present-day targets. Alternatively, these changes may be critical for deaminase-independent restriction.

This method of restriction involves inhibiting RT-mediated elongation and accumulation of reverse transcription products, which has been demonstrated for various APOBEC3 proteins *in vitro* by several groups (Bishop, 2006; Iwatani, 2007; Li, 2007; Guo, 2007; Bishop, 2008). Because *in vitro* systems typically use plasmid-based gene expression, it has been argued that cytidine deaminase-independent inhibition is an artifact of APOBEC3 over-expression (Schumacher, 2008; Miyagi, 2007). Our studies indicate that cytidine deaminase-independent inhibition occurs *in vivo*, confirming its role as an important antiviral effect of APOBEC3.

The evolution of two main genetic variants of murine APOBEC3 is noteworthy, as mice encode only a single APOBEC3 gene as opposed to the multiple genes expressed in higher mammals such as cats, horses, and humans (Sanville, 2010). As previously mentioned, this evolution has a significant effect in mice, as mice encoding the APOBEC3$^{C57BL/6}$ variant of the gene are better able to control certain viruses (Santiago, 2008; Okeoma, 2010). The basis for this functional difference has to date been unclear. We have shown that differential APOBEC3 expression levels are critical for this difference. Mice encoding APOBEC3$^{C57BL/6}$ express significantly higher levels at both the RNA and protein level, and are better able to restrict. Heterozygous mice expressing lower levels of their respective APOBEC3 allele also show decreased viral restriction. Thus, higher endogenous expression of APOBEC3 is beneficial for viral restriction, and promotion of increased APOBEC3 expression may be a useful therapeutic tool.

Chapter 4.1.2: APOBEC3-induced viral evolution of murine retroviruses *in vivo*

Although APOBEC3 clearly has antiviral means, viral targets may benefit from the increased mutation rate offered via G-to-A mutagenesis. Murine retroviruses do not suffer the lethal hypermutation seen in other viral-host systems, and the very low-level, sublethal mutation rate that does occur could speed viral evolution. HIV-1 has evolved Vif as an anti-APOBEC3 device, causing rapid degradation and preventing incorporation of APOBEC3 into budding virions. However, Vif is not always completely effective, either due to high APOBEC3 expression or

defective Vif (Sato, 2010; Simon, 2005). Therefore, APOBEC3 is sometimes able to cause sublethal HIV mutations. This is of particular note because a considerable number of antiviral drug resistance mutations are G-to-A based (Berkhout and de Ronde, 2004). These mutations can even arise before drug exposure, allowing for a potential pool of resistant viruses ready for rapid expansion once drug is administered (Mulder, 2008).

Our transgenic APOBEC3G-expressing mice provide an ideal system with which to test the evolutionary effect of APOBEC3 *in vivo*. We present a first look into the timing and magnitude of APOBEC3G mutations in MLV in comparison to mutations found in virus replicating in the complete absence of APOBEC3. Viruses with significant G-to-A mutagenesis were found early in infection at the level of integrated proviruses. After several additional weeks of viral replication, we were able to detect G-to-A mutations within viruses of the circulating viral plasma population. Although we did not see a significant impact on the development of drug resistance under pressure of suboptimal doses of AZT, we did show a significantly higher rate of viral mutagenesis within MLV *pol* within APOBEC3G-expressing mice.

This increase occurred even in the presence of APOBEC3G expression levels considerably lower than those found in human PBMCs. Presumably, higher levels of transgene expression would lead to increased mutation rate, which would have an improved antiviral effect but could induce significantly faster viral evolution. Our lab has also generated additional transgenic mice expressing near-human levels of APOBEC3G. However, when these mice were infected with M-MLV, the levels of infection were several logs lower than in non-transgenic mice and some mice even cleared infection (Stavrou, manuscript in preparation). This is consistent with *in vitro* data using APOBEC3G transfected cells, demonstrating that MLV is highly edited by APOBEC3G. Thus we chose to use the A3G transgenic strain studied here so that we could limit lethal mutations. This complex interplay between the antiviral activity of APOBEC3 versus a potential proviral activity is one to be more closely considered, particularly in note of several attempts to utilize APOBEC3 in therapeutic strategies (Cadima-Couto and Goncalves, 2010). Increasing APOBEC3 levels in addition to antiretroviral drugs was beneficial to decrease viral

replication in our system, as MLV is unable to degrade the protein. However, it is unknown if this effect would be similar in an HIV-1 infection in the presence of an effective Vif that is able to completely degrade the protein. It is possible that increasing the expression of APOBEC3 may have little effect on reducing viral replication, if Vif can handle degradation of an increased level of protein. However, it could increase viral diversity via sublethal mutagenesis, potentially aiding the virus in evading recognition by antiviral cytotoxic T-lymphocytes or antibodies, as well as helping create drug resistance mutations.

Although we did not see a significant increase in total levels of viral diversity at the protein level in viruses isolated from APOBEC3G-expressing mice, we did see a specific trend towards G-to-A induced amino acid changes. We believe that at later time points this increase in G-to-A mediated amino acid changes will result in significantly higher levels of viral diversity, as nucleotide mutations continue to accumulate. Although the higher level of mutation at the integrated provirus level is at least partially obscured from the circulating plasma virus pool through purifying selection, we show that APOBEC3G does make substantial contributions to viral sequence diversity.

Chapter 4.2: Future Directions

This work has focused on the mechanisms and impact of both antiviral and proviral effects of APOBEC3 *in vivo*. Both remain important areas for future investigation. First, the mechanism by which APOBEC3 interferes with reverse transcription and reduces accumulation of viral cDNA is unknown. While MMTV and MLV provide excellent systems for studying viral-host interactions *in vivo*, a better understanding of these mechanisms will likely require substantial *in vitro* biochemical analysis. However, as murine APOBEC3 seems to rely solely on this cytidine deaminase-independent mechanism to restrict murine retroviruses, it may serve as an appropriate system to study *in vitro*.

Furthermore, the means by which murine retroviruses prevent the accumulation of cytidine deaminase-mediated mutations should be further explored. Although gGag serves as a

Vif-like factor for M-MLV, it does not seem to have a specific impact on preventing cytidine deamination by endogenous murine APOBEC3, but instead prevents inhibition of reverse transcription (Stavrou, 2013). Both MMTV and MLV may share a similar method for preventing G-to-A hypermutation. Evolutionary analysis of both the viruses and murine APOBEC3 may provide some insight into regions of the virus or protein that play an important role in this evasion.

The role of APOBEC3 on viral evolution and drug resistance is even less well understood. Our model system of MLV-infected APOBEC3G-expressing mice could be utilized to offer clarity in conjunction with our results. Our experimental model and results provide tools to better design a larger study with better chances of determing the real role of APOBEC3 in viral evolution. First, a long-term study of APOBEC3-mediated mutagenesis in the absence of any selective pressure should be performed. Both integrated proviruses and circulating plasma virus can be easily obtained at various time points without sacrificing animals. Additional data could be obtained from mice expressing higher levels of the transgene to determine a dose effect. This would provide a baseline to understand how viral diversity is affected without a direct evolutionary pressure. Because APOBEC3G causes a 5' to 3' increasing gradient of hypermutation, utilization of selective pressures that target a gene at the 3' end of the genome may be most useful.

Many monoclonal antibodies to MLV exist, and the specific protein targets of these antibodies are known. One potential experimental model would use a neutralizing monoclonal antibody that targets a portion of the viral Env protein that contains specific amino acids residues affected by G-to-A mutations. The monoclonal antibody could be injected into mice similarly to an antiretroviral drug, and the level of viral binding carefully titrated to neutralize the virus to an optimal level. Viral escape would then prevent specific antibody binding and thus neutralization. This could provide an excellent alternative system in which to assess the evolutionary effect of APOBEC3 *in vivo*. Deep sequencing could recover any viral genome containing mutations within the antibody-targeted region, and these viruses could also be grown *ex vivo* in the presence of the monoclonal antibody to assess the ability of the antibody to bind.

A newly created MLV virus engineered to express HIV-1 Vif may also help elucidate the effect of this critical anti-APOBEC3 factor on the level of APOBEC3G-mediated inhibition and mutagenesis (Spiros Stavrou and Ed Browne, data unpublished). Mutations known to affect the activity of Vif could also be engineered into the Vif-containing virus to determine absolute effects, both on viral replication and evolution *in vivo*. For all of these studies, high-throughput next generation sequencing would be extraordinarily beneficial. This technology would allow extensive sampling of the viral population. Furthermore, several mice from experimental and control groups could be combined using barcoded primers for easy identification, to reduce cost and increase sample size (Jabara, 2011).

Our model could also be used for continued studies of the effect of APOBEC3G on drug resistance. Although the reverse transcriptase of MLV shares significant structural homology with that of HIV-1, only a single known AZT-resistance mutation is known for MLV RT. However, this mutation was engineered for experimental purposes, so naturally arising resistance mutations, similar to those known for HIV, have not been identified. This mutation is not mediated by a G-to-A mutation (Ndongwe, 2011). Regardless, it was not seen in virus isolated from any of our AZT-treated mice, irrespective of genotype. Although we initially reasoned that AZT would work well in our system based on the requirement for multiple resistance mutations (TAMs), other drugs have been used with better results in similar evolution studies *in vitro*. For example, lamivudine (3TC) resistance arises via a single nucleotide substitution in HIV-1 reverse transcriptase, which is mediated by an APOBEC3G targeted G-to-A mutation (at M184) (Kim, 2010). While this drug has little activity against MLV, other drugs with a similarly distinct resistance mutation pattern may work better within our murine model. In conclusion, our studies offer a basic foundation for an almost unlimited collection of important data, which together could potentially affect therapeutic prospects for HIV and other human retroviral pathogens.

REFERENCES

1. Abudu A, Takaori-Kondo A, Izumi T, et al. Murine retrovirus escapes from murine APOBEC3 via two distinct novel mechanisms. *Curr Biol*. 2006;16(15):1565--1570.

2. Aguiar RS, Lovsin N, Tanuri A, Peterlin BM. Vpr.A3A chimera inhibits HIV replication. *J Biol Chem*. 2008;283(5):2518--2525.

3. Albritton LM, Tseng L, Scadden D, Cunningham JM. A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection. *Cell*. 1989;57(4):659--666.

4. Armitage AE, Deforche K, Chang CH, et al. APOBEC3G-induced hypermutation of human immunodeficiency virus type-1 is typically a discrete "all or nothing" phenomenon. *PLoS Genet*. 2012;8(3):e1002550.

5. Baltimore D. RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature*. 1970;226(5252):1209--1211.

6. Balvay L, Lastra ML, Sargueil B, Darlix JL, Ohlmann T. Translational control of retroviruses. *Nat Rev Microbiol*. 2007;5:128--140.

7. Barre-Sinoussi F, Chermann JC, Rey F, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). 1983. *Rev Invest Clin*. 2004;56(2):126-129.

8. Baumann JG. Intracellular restriction factors in mammalian cells-- an ancient defense system finds a modern foe. *Curr HIV Res*. 2006;4:141--168.

9. Beale RC, Petersen-Mahrt SK, Watt IN, Harris RS, Rada C, Neuberger MS. Comparison of the differential context-dependence of DNA deamination of APOBEC enzymes: Correlation with mutation spectra in vivo. *J Mol Biol*. 2004;337(3):585--596.

10. Bentz J, Mittal A. Deployment of membrane fusion protein domains during fusion.. *Cell Biol Int*. 2000;24(11):819--838.

11. Berger G, Durand S, Fargier G, et al. APOBEC3A is a specific inhibitor of the early phases of HIV-1 infection in myeloid cells. *PLoS Pathog*. 2011;7(9):e1002221.

12. Bergshoeff AS, Fraaij PLA, Verweij C, et al. Plasma levels of zidovudine twice daily compared with three times daily in six HIV-1-infectec children. *J Antimicrob Chemother*. 2004;54:1152--1154.

13. Berkhout B, Das AT, Beerens N. HIV-1 RNA editing, hypermutation, and error-prone reverse transcription. *Science*. 2001;292:7.

14. Berkhout B, de Ronde A. APOBEC3G versus reverse transcriptase in the generation of HIV-1 drug resistance mutations. *AIDS*. 2004;18(13):186--1863.

15. Bernhard W. Electron microscopy of tumor cells and tumor viruses. *Cancer Res*. 1958;18:491--509.

16. Bieniasz P. Restriction factors: A defense against retroviral infection. *Trends Microbiol*. 2003;11(6):286--291.

17. Bishop KN, Holmes RK, Malim MH. Antiviral potency of APOBEC proteins does not correlate with cytidine deamination. *J Virol*. 2006;80(17):8450--8458.

18. Bishop KN, Holmes RK, Sheehy AM, Davidson NO, Cho S, Malim MH. Cytidine deamination of retroviral DNA by diverse APOBEC proteins. *Curr Biol*. 2004;14(15):1392--1396.

19. Bishop KN, Verma M, Kim EY, Wolinsky SM, Malim MH. APOBEC3G inhibits elongation of HIV-1 reverse transcripts. *PLoS Pathog*. 2008;4(12):e1000231.

20. Bittner JJ. Some possible effects of nursing on the mammary gland tumor incidence in mice. *Science*. 1936;84(2172):162.

21. Bonvin M, Achermann F, Greeve I, et al. Interferon-inducible expression of APOBEC3 editing enzymes in human hepatocytes and inhibition of hepatatis B virus resplication. *Hepatology*. 2006;43(6):1364--1374.

22. Boone LR, Skalka AM. Viral DNA synthesized in vitro by avian retrovirus particles permeabilized with melittin. II. evidence for a strand displacement mechanism in plus-strand synthesis. *J Virol*. 1981;37(1):293--304.

23. Boucher CA, Cammack N, Schipper P, et al. High-level resistance to (-) enantiomeric 2'-deoxy-3'-thiacytidine in vitro is due to one amino acid substitution in the catalytic site of human immunodeficiency virus type 1 reverse transcriptase. *Antimicrob Agents Chemother*. 1993;37(10):2231--2234.

24. Bray M, Prasad S, Dubay JW, et al. A small element from the mason-pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication rev-independent. *PNAS*. 1994;91:1256--1260.

25. Brekelmans P, van Soest P, Voerman J, Platenburg PP, Leenen PJ, van Ewijk W. Transferrin receptor expression as a marker of immature cycling thymocytes in the mouse. *Cell Immunol*. 1994;159:331--339.

26. Browne EP, Littman DR. Species-specific restriction of Apobec3-mediated hypermutation. *J Virol*. 2008;87(16):1305--1313.

27. Bukrinsky MI, Sharova N, Dempsey MP, et al. Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. *PNAS*. 1992;89:6580--6584.

28. Burns DP, and Desrosiers RC. Envelope sequence variation, neutralizing antibodies, and primate lentivirus persistance. *Curr Top Microbiol Immuno*. 1994;188:185--219.

28B. Cadima-Couto I, and Goncalves J. Toward Inhibition of Vif-APOBEC3 Interaction: Which Protein to Target? *Advances in Virology*. 2010; 2010: 1-10.

29. Callahan R, Smith GH. Common integration sites for MMTV in viral induced mouse mammary tumors. *J Mammary Gland Biol Neoplasia*. 2008;13:309--321.

30. Campbell S, Vogt VM. In vitro assembly of virus-like particles with rous sarcoma virus gag deletion mutants: Identification of the p10 domain as a morphological determinant in the formation of spherical particles.. *J Virol*. 1997;71(6):4425--4435.

31. Carlson L, de Marco A, Oberwinkler H, et al. Cryo electron tomography of native HIV-1 budding sites. *PLoS Pathog*. 2010;6(11):e1001173.

32. Chesebro B, Wehrly K. Identification of a non-H-2 gene (rfv-3) influencing recovery from viremia and leukemia induced by friend virus complex. *PNAS*. 1979;76(1):425--429.

33. Chesebro B, Wehrly K. Identification of a non-H-2 gene (rfv-3) influencing recovery from viremia and leukemia induced by friend virus complex. *PNAS*. 1979;76(1):425--429.

34. Chiu Y, Greene WC. The APOBEC3 cytidine deaminases: An innate defensive network opposing exogenous retroviruses and endogenous retrolements. *Annu Rev Immunol*. 2008;26:317--353.

35. Chow H, Brookshier G, Li P. Tissue deposition of zidovudine and its phosphorylated metabolits in zidovudine-treated healthy and retrovirus infected mice. *Pharm Res*. 1998;15(1):139.

36. Clavel F, Hance AJ. HIV drug resistance. *N Engl J Med*. 2004;350:1023--1035.

37. Coffin JM. Retroviridae and their replication. In: Fields BN, ed. *Virology.* New York: Raven Press; 1996:1767--1848.

38. Coffin JM, Hughes SH, Varmus HE, eds. *Retroviruses.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2002.

39. Cohen JC, Varmus HE. Endogenous mammary tumour virus DNA varies among wild mice and segregates during inbreeding. *Nature*. 1979;278(5703):418--423.

40. Conticello SG, Thomas CJF, Petersen-Mahrt SK, Neuberger MS. Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol Biol Evol*. 2005;22(2):367--377.

41. Corbin A, Prats AC, Darlix JL, Sitbon M. A nonstructural gag-encoded glycoprotein precursor is necessary for efficient spreading and pathogenesis of murine leukemia viruses. *J Virol*. 1994;68(6):3857--3867.

42. Crick F. Central dogma of molecular biology. *Nature*. 1970;227(5258):561--563.

43. Das D, Georgiadis MM. The crystal structure of the monomeric reverse transcriptase from moloney murine leukemia virus. *Structure*. 2004;12(5):819--829.

44. Davis NL, Rueckert RR. Properties of a ribonucleoprotein particle isolated from nonidet P-40-treated rous sarcoma virus. *J Virol*. 1972;10(5):1010--1020.

45. De Maio FA, Rocco CA, Aulicino PC, Bologna R, Mangano A, Sen L. Unusual substitutions in HIV-1 vif from children infected perinatally without progression to AIDS for more than 8 years without therapy. *J Med Virol*. 2012;84(12):1844--1852.

46. Delelis O, Carayon K, Saib A, Deprez E, Mouscadet J. Integrase and integration: Biochemical activities of HIV-1 integrase. *Retrovirology*. 2008;5:114.

47. Delelis O, Zamborlini A, Thierry S, Saib A. Chromosomal tethering and proviral integration. *Biochim Biophys Acta*. 2010;1799(3-4):207-216.

48. Dickson C, Peters G. Proteins encoded by mouse mammary tumour virus. *Curr Top Microbiol Immuno*. 1983;106:1--34.

49. Doehle BP, Schafer A, Wiegand HL, Bogerd HP, Cullen BR. Differential sensitivity of murine leukemia virus to APOBEC3-mediated inhibition is governed by virion exclusion. *J Virol*. 2005;79(13):8201--8207.

50. Dougherty WG, Semler BL. Expression of virus-encoded proteinases: Functional and structural similarities with cellular enzymes. *Microbiol Rev*. 1993;57(4):781--822.

51. Duggan J, Okonta H, Chakraborty J. Transmission of moloney murine leukemia virus (ts-1) by breast milk. *J Gen Virol*. 2006;87(Pt 9):2679--2684.

52. Dunn TB, Moloney JB, Green AW, Arnold B. Pathogenesis of a virus-induced leukemia in mice. *J Natl Cancer Inst*. 1961;26:189--221.

53. Einfeld D. Maturation and assembly of retroviral glycoproteins. In: Krausslich HG, ed. *Morphogenesis and maturation of retroviruses.* ; 1996:133--176.

54. Elder JH, Lerner DL, Hasselkus-Light CS, et al. Distinct subsets of retroviruses encode dUTPase. *J Virol*. 1992;66:1791--1794.

55. Elder JH, McGee JS, Alexander S. Carbohydrate side chains of rauscher leukemia virus envelope glycoproteins are not required to elicit a neutralizing antibody response. *J Virol*. 1986;57(1):340--342.

56. Elis E, Ehrlich M, Prizan-Ravid A, Laham-Karam N, Bacharach E. P12 tethers the murine leukemia virus pre-integration complex to mitotic chromosomes. *PLoS Pathog*. 2012;8(12):e1003103.

57. Ellerman V, Bang O. Experimentelle leukamie bei huhnern. *Zentralbl Bakteriol Parasitenkd Infectionskr Hyg Abt*. 1908;46:595--609.

58. Esnault C, Priet S, Ribet D, Heidmann O, Heidmann T. Restriction by APOBEC3 proteins of endogenous retroviruses with an extracellular life cycle: Ex vivo effects and in vivo "traces" on the murine IAPE and human HERV-K elements. *Retrovirology*. 2008;5:75.

59. Fassati A, Goff SP. Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. *J Virol*. 2001;75:3626--3635.

60. Fine DL, Arthur LO, Young LJ. Cell culture factors influencing in vitro expression of mouse mammary tumor virus. *In Vitro*. 1976; 12(10): 693-701.

61. Finke D, Baribaud F, Diggelmann H, Acha-Orbea H. Extrafollicular plasmablast B cells play a key role in carrying retroviral infection to peripheral organs. *J Immunol*. 2001;166(10):6266--6275.

62. Finston WI, Champoux JJ. RNA-primed initiation of moloney murine leukemia virus plus strands by reverse transcriptase in vitro.. *J Virol*. 1984;51(1):26--33.

63. Fischer U, Huber J, Boelens WC, Mattajt LW, Luhrmann R. The HIV-1 rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell*. 1995;82(3):475--483.

64. Fisher AG, Ensoli B, Ivanoff L, et al. The sor gene of HIV-1 is required for efficient virus transmission in vitro. *Science*. 1987;237(4817):888--893.

65. Fourati S, Malet I, Binka M, et al. Partially active HIV-1 vif alleles facilitate viral escape from specific antiretrovirals. *AIDS*. 2010;24(15):2313--2321.

66. Freed EO. HIV-1 gag proteins: Diverse functions in the virus life cycle. *Virology*. 1998;251(1):1--15.

67. Freed EO, Risser R. The role of envelope glycoprotein processing in murine leukemia virus infection. *J Virol*. 1987;61(9):2852--2856.

68. Gabuzda DH, Lawrence K, Langhoff E, et al. Role of vif in replication of human immunodeficiency virus type 1 in CD4+ T lymphocytes. *J Virol*. 1992;66(11):6489--6495.

69. Gandhi SK, SIliciano JD, Bailey JR, Siliciano RF, Blankson JN. Role of APOBEC3G/F-mediate hypermutation in the control of human immunodeficiency virus type 1 in elite suppressors. *J Virol*. 2008;82(6):3125--3130.

70. Garcia-Lerma JG. Diversity of thymidine analogue resistance genotypes among newly diagnosed HIV-1-infected persons. *J Antimicrob Chemother*. 2005;56(2):265--269.

71. Gheysen D, Jacobs E, de Foresta F, et al. Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells.. *Cell*. 1989;59(1):103--112.

72. Gillick K, Pollpeter D, Phalora P, Kim EY, Wolinsky SM, Malim MH. Suppression of HIV-1 infection by APOBEC3 proteins in primary human CD4(+) T cells is associated with inhibition of processive reverse transcription as well as excessive cytidine deamination. *J Virol*. 2013;87(3):1508--1517.

73. Goff S. Retrovirus restriction factors. *Mol Cell*. 2004;16:849--859.

74. Goff S. Retroviral reverse transcriptase: Synthesis, structure, and function. *J Acquir Immune Defic Syndr*. 1990;3(8):817--831.

75. Goila-Gur R, Strebel K. HIV-1 vif, APOBEC, and intrinsic immunity. *Retrovirology*. 2008;5:51.

76. Golovkina TV, Dudley JP, Ross SR. Superantigen activity is needed for mouse mammary tumor virus spread within the mammary gland. *J Immunol*. 1998;161:2375--2382.

77. Golovkina TV, Jaffe AB, Ross SR. Coexpression of exogenous and endogenous mouse mammary tumor virus RNA in vivo results in viral recombination and broadens the virus host range. *J Virol*. 1994;68(8):5019--5026.

78. Golovkina TV, Jaffe AB, Ross SR. Coexpression of exogenous and endogenous mouse mammary tumor virus RNA in vivo results in viral recombination and broadens the virus host range. *J Virol*. 1994;68(8):5019--5026.

79. Golovkina TV, Prescott JA, Ross SR. Mouse mammary tumor virus-induced tumorigenesis in sag transgenic mice: A laboratory model of natural selection. *J Virol*. 1993;67:7690--7694.

80. Grand RJA. Acylation of viral and eukaryotic proteins. *J Biochem*. 1989;258:625--638.

81. Gross L. Development and serial cellfree passage of a highly potent strain of mouse leukemia virus. *Proc Soc Exp Biol Med*. 1957;94(4):767--771.

82. Guntaka RV. Transcription termination and polyadenylation in retroviruses. *Microbiol Mol Biol Rev*. 1993;57(3):511--521.

83. Hakata Y, Landau NR. Reversed functional organization of mouse and human APOBEC3 cytidine deaminase domains. *J Biol Chem*. 2006;281(48):36624--36631.

84. Hamard-Peron E, Muriaux D. Retroviral matrix and lipids, the intimate interaction. *Retrovirology*. 2011;7(8):15.

85. Harrigan PR, Hogg RS, Dong WWY, et al. Predictors of HIV drug-resistance mutations in large antiretroviral-naive cohort initiating triple antiretroviral therapy. *J Infect Dis*. 2005;191(3):339--347.

86. Harris ME, Hope TJ. RNA export: Insights from viral models. *Essays Biochem*. 2000;36:115--127.

87. Harris RS, Bishop KN, Sheehy AM, et al. DNA deamination mediates innate immunity to retroviral infection. *Cell*. 2003;113(6):803--809.

88. Harris RS, Hultquist JF, Evans DT. The restriction factors of human immunodeficiency virus. *J Biol Chem*. 2012;287:40875--40883.

89. Harris RS, Petersen-Mahrt SK, Neuberger MS. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell*. 2002;10(5):1247--1253.

90. Hartley JW, Rowe WP. Naturally occurring murine leukemia viruses in wild mice: Characterization of a new "amphotropic" class. *J Virol*. 1976;19(1):19--25.

91. Haseltine WA, Kleid DG, Panet A, Rothenberg E, Baltimore D. Ordered transcription of RNA tumor virus genomes.. *J Mol Biol*. 1976;106(1):109--131.

92. Held W, Shakhov AN, Izui S, et al. Superantigen-reactive CD4+ T cells are required to stimulate B cells after infection with mouse mammary tumor virus. *J Exp Med*. 1993;177(2):359--366.

93. Held W, Waanders GA, Acha-Orbea H, MacDonald HR. Reverse transcriptase-dependent and -indepedent phases of infection with mouse mammary tumor virus: Implications for superantigen function. *J Exp Med*. 1994;180(6):2347--2351.

94. Holmes RK, Malim MH, Bishop KN. APOBEC-mediated viral restriction: Not simply editing? *Trends Biochem Sci*. 2007;32(3):118--128.

95. Holtz CM, Sadler HA, Mansky LM. APOBEC3G cytosine deamination hotspots are defined by both sequence context and single-stranded DNA secondary structure. *Nucleic Acids Res*. 2013;41(12):6139--6148.

96. Hulme AE, Perez O, Hope TJ. Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. *PNAS*. 2011;108(24):9975--9980.

97. Ignatowicz L, Kappler J, Marrack P. The effects of chronic infection with a superantigen producing virus. *J Exp Med*. 1992;175:917--923.

98. Indik S, Gunzburg WH, Kulich P, Salmons B, Rouault F. Rapid spread of mouse mammary tumor virus in cultured human breast cells. *Retrovirology*. 2007;4:73.

99. Iwatani Y, Chan DS, Wang F, et al. Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. *Nucleic Acids Res*. 2007;35(21):7096--7108.

100. Jabara CB, Jones CD, Roach J, Anderson JA, Swanstrom R. Accurate sampling and deep sequencing of the HIV-1 protease gene using a primer ID. *PNAS*. 2011;108(50):20166--20171.

101. Janeway C, Travers P, Walport M, Shlomchik M, eds. *Immunobiology.* Fifth Edition ed. New York, NY: Garland Science; 2001.

102. Janini M, Rogers M, Birx DR, McCutchan FE. Human immunodeficiency virus type 1 DNA sequences genetically damaged by hypermutation are often abundant in patient peripheral blood mononuclear cells and may be generated during near-simultaneous infection and activation of CD4(+) T cells. *J Virol*. 2001;75(17):7973--7986.

103. Jarmuz A, Chester A, Bayliss J, et al. An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics*. 2002;79(3):285--296.

104. Jern P, Russell RA, Pathak VK, Coffin JM. Likely ROle of APOBEC3G-mediated G-to-A mutations in HIV-1 evolution and drug resistance. *PLoS Pathog*. 2009;5(4):e1000367.

105. Jern P, Stoye JP, Coffin JM. Role of APOBEC3 in genetic diversity among endogenous murine leukemia viruses. *PLoS Genet*. 2007;3(10):2014--2022.

106. Johnson PR, Hamm TE, Goldstein S, Kitov S, Hirsch VM. The genetic fate of molecularly cloned simian immunodeficiency virus in experimentally infected macaques. *Virology*. 1991;185:217--228.

107. Johnson SF, Telesnitsky A. Retroviral RNA dimerization and packaging: The what, how, when, where, and why. *PLoS Pathog*. 2010;6(10):e1001007.

108. Kaiser SM, Emerman M. Uracil DNA glycosylase is dispensable for human immunodeficiency virus type 1 replication and does not contribute to the antiviral effects of the cytidine deaminase Apobec3G. *J Virol*. 2006;80:875--882.

109. Kappler JW, Staerz U, White J, Marrack PC. Self-tolerance eliminates T cells specific for mls-modified products of the major histocompatibility complex. *Nature*. 1988;332(6159):35--40.

110. Kelly NJ, Palmer MT, Morrow CD. Selection of retroviral reverse transcription primer is coordinated with tRNA biogenesis. *J Virol*. 2003;77(16):8695--8701.

111. Kieffer TL, Kwon P, Nettles RE, Han Y, Ray SC, Siliciano RF. G-->A hypermutation in protease and reverse transcriptase regions of human immunodeficiency virus type 1 residing in resting CD4+ T cells in vivo. *J Virol*. 2005;79(3):1975--1980.

112. Kim EY, Bhattacharya T, Kunstman K, et al. Human APOBEC3G-mediated editing can promote HIV-1 sequence diversification and accelerate adaptation to selective pressure. *J Virol*. 2010;84(19):10402--10405.

113. Kim JW, Closs EI, Albritton LM, Cunningham JM. Transport of cationic amino acids by the mouse ecotropic retrovirus receptor. *Nature*. 1991;352:725--728.

114. Klein KC, Reed JC, Lingappa JR. Intracellular destinies: Degradation, targeting, assembly, and endocytosis of HIV gag. *AIDS Rev*. 2007;9:150--161.

115. Knipe DM, Howley PM, eds. *Fields virology.* Fifth ed. Lippincott Williams & Wilkins; 2007.

116. Kohli RM, Abrams SR, Gajula KS, Maul RW, Gearhart PJ, Stivers JT. A portable hot spot recognition loop transfers sequence preferences from APOBEC family members to activation-induced cytidine deaminase. *J Biol Chem*. 2009;284(34):22898--22904.

117. Kohli RM, Maul RW, Guminski AF, et al. Local sequence targeting in the AID/APOBEC family differentially impacts retroviral restriction and antibody diversification. *J Biol Chem*. 2010;285(52):40956--40964.

118. Kolokithas A, Rosenke K, Malik F, et al. The glycosylated gag protein of a murine leukemia virus inhibits the antiretroviral function of APOBEC3. *J Virol*. 2010;84(20):10933--10936.

119. Koning FA, Goujon C, Bauby H, Malim MH. Target cell-mediated editing of HIV-1 cDNA by APOBEC3 proteins in human macrophages. *J Virol*. 2011;85(24):13448--13452.

120. Koning FA, Newman ENC, Kim EY, Kunstman K, Wolinsky SM, Malim MH. Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. *J Virol*. 2009;83(18):9474--9485.

121. Korman AJ, Bourgarel P, Meo T, Rieckhof GE. The mouse mammary tumour virus long terminal repeat encodes a type II transmembrane glycoprotein. *EMBO J*. 1992;11(5):1901--1905.

122. Kozak CA. The mouse "xenotropic" gammaretroviruses and their XPR1 receptor. *Retrovirology*. 2010;7:101.

123. Krebs FC, Hogan TH, Quiterio S, Gartner S, Wigdahl B. Lentiviral LTR-directed expression, sequence variation, and disease pathogenesis. In: Kuiken C, Foley B, Hahn B, et al, eds. *HIV sequence compendium 2001.* Los Alamos, NM: Theoretical Biology and Biophysics Group; 2001:29--70.

124. Kulkosky J, Skalka AM. HIV DNA integration: Observations and interferences. *J Acquir Immune Defic Syndr*. 1990;3(9):839--851.

125. Laimins LA, Gruss P, Pozzatti R, Khoury G. Characterization of enhancer elements in the long terminal repeat of moloney murine sarcoma virus. *J Virol*. 1984;49(1):183--189.

126. Land AM, Bail TB, Luo M, et al. Human immunodeficiency virus (HIV) type 1 proviral hypermutation correlates with CD4 count in HIV-infected women from kenya. *J Virol*. 2008;82(16):8172--8182.

127. Langlois MA, Kemmerich K, Rada C, Neuberger MS. The AKV murine leukemia virus is restricted and hypermutated by mouse APOBEC3. *J Virol*. 2009;83(22):11550--11559.

128. LaRue RS, Andresdottir V, Blanchard Y, et al. Guidelines for naming nonprimate APOBEC3 genes and proteins. *J Virol*. 2009;83(2):494-497.

129. Layne SP, Merges MJ, Dembo M, Spouge JL, Nara PL. HIV requires multiple gp120 molecules for CD-4 mediated infection. *Nature*. 1990;346(6281):277--279.

130. Lee S, Potempa M, Swanstrom R. The choreography of HIV-1 proteolytic processing and virion assembly. *J Biol Chem*. 2012;287(49):40867--40874.

131. Leider JM, Palese P, Smith FI. Determination of the mutation rate of a retrovirus. *J Virol*. 1988;62(9):3084--3091.

132. Li J, Hakata Y, Takeda E, et al. Two genetic determinants acquired late in mus evolution regulate the inclusion of exon 5, which alters mouse APOBEC3 translation efficiency. *PLoS Pathog*. 2012;8(1):e1002478.

133. Li Y, Golemis E, Hartley JW, Hopkins N. Disease specificity of nondefective friend and moloney murine leukemia viruses is controlled by a small number of nucleotides. *J Virol*. 1987;61(3):693--700.

134. Low A, Okeoma CM, Lovsin N, et al. Enhanced replication and pathogenesis of moloney murine leukemia virus in mice defective in the murine APOBEC3 gene. *Virology*. 2009;385(2):455--463.

135. Lucas GM. Antiretroviral adherence, drug resistance, viral fitness and HIV disease progression: A tangled web is woven.. *J Antimicrob Chemother*. 2005;55(4):413--416.

136. Madani N, Kabat D. An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral vif protein. *J Virol*. 1998;72(12):10251--10255.

137. Mak J, Kleiman L. Primer tRNAs for reverse transcription. *J Virol*. 1997;71(11):8087--8095.

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

139. Mansky LM. Forward mutation rate of human immunodeficiency virus type 1 in a T lymphoid cell line. *AIDS Res Hum Retroviruses*. 1996;12(4):307--314.

140. Mariani R, Chen D, Schrofelbauer B, et al. Species-specific exclusion of APOBEC3G from HIV-1 virions by vif. *Cell*. 2003;114(1):21--31.

141. McClure MO, Sommerfelt MA, Marsh M, Weiss RA. The pH independence of mammalian retrovirus infection. *J Gen Virol*. 1990;71:767--773.

142. McNally MT. RNA processing control in avian retroviruses. *Front Biosci*. 2008;13:3869--3883.

143. Mehta A, Kinter MT, Sherman NE, Driscoll DM. Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA. *Mol Cell Biol*. 2000;20(5):1846--1854.

144. Meric C, Darlix JL, Spahr PF. It is rous sarcoma virus protein P12 and not P19 that binds tightly to rous sarcoma virus RNA. *J Mol Biol*. 1984;173(4):531--538.

145. Michalides R, Schlom J. Relationship in the nucleic acid sequences between mouse mammary tumor virus variants. *PNAS*. 1975;72(11):4635--4639.

146. Michalides R, van Ooyen A, Nusse R. Mouse mammary tumor virus expression and mammary tumor development. *Curr Top Microbiol Immuno*. 1983;106:57--78.

147. Miller DG, Edwards RH, Miller AD. Cloning of the cellular receptor for amphotropic murine retroviruses reveals homology to that for gibbon ape leukemia virus. *PNAS*. 1994;91(1):78--82.

148. Mitchell RS, Beitzel BF, Schroder ARW, et al. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol*. 2004;2(8):e234.

149. Miyagi E, Opi S, Takeuchi H, et al. Enzymatically active APOBEC3G is required for effiicient inhibition of human immunodefiency virus type 1. *J Virol*. 2007;81(24):13346--13353.

150. Moelling K, Bolognesi DP, Bauer H. Polypeptides of avian RNA tumor viruses. 3. purification and identification of a DNA synthesizing enzyme. *Virology*. 1971;45(1):298--302.

151. Mulder LCF, Harari A, Simon V. Cytidine deamination induced HIV-1 drug resistance. *PNAS*. 2008;105(14):5501--5506.

152. Mullner M, Salmons B, Gunzburg WH, Indik S. Identification of the rem-responsive element of mouse mammary tumor virus. *Nucleic Acids Res*. 2008;36:6284--6294.

153. Naghavi MH, Goff SP. Retroviral proteins that interact with the host cell cytoskeleton. *Curr Opin Immunol*. 2007;19:402--407.

154. Navia MA, Fitzgerald PM, McKeever BM, et al. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature*. 1989;337(6208):615--620.

155. Neil S, Bieniasz P. Human immunodeficiency virus, restriction factors, and interferon. *J Interferon Cytokine Res*. 2009;29(9):569--580.

156. Newman ENC, Holmes RK, Craig HM, et al. Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity. *Curr Biol*. 2005;15(2):166--170.

157. Nisole S, Saib A. Early steps of retrovirus replicative cycle. *Retrovirology*. 2004;1:9.

158. Niwa O, Yokota Y, Ishida H, Sugahara T. Independent mechanisms involved in suppression of the moloney leukemia virus genome during differentiation of murine teratocarcinoma cells.. *Cell*. 1983;32(4):1105--1113.

159. Okeoma CM, Huegel AL, Lingappa JR, Feldman MD, Ross SR. APOBEC3 proteins expressed in mammary epithelial cells are packaged into retroviruses and can restrict transmission of milk-borne virions. *Cell Host Microbe*. 2010;8(6):534--543.

160. Okeoma CM, Lovsin N, Peterlin BM, Ross SR. APOBEC3 inhibits mouse mammary tumour virus replication in vivo. *Nature*. 2007;445(7130):927--930.

161. Okeoma CM, Low A, Bailis W, Fan HY, Peterlin BM, Ross SR. Induction of APOBEC3 in vivo causes increased restriction of retrovirus infection. *J Virol*. 2009;83(8):3486--3495.

162. Okeoma CM, Petersen J, Ross SR. Expression of murine APOBEC3 alleles in different mouse strains and their effect on mouse mammary tumor virus infection. *J Virol*. 2009;83(7):3029--3038.

163. Ooms M, Letko M, Binka M, Simon V. The resistance of human APOBEC3H to HIV-1 NL4-3 molecular clone is determined by a single amino acid in vif. *PLoS One*. 2013;8(2):e57744.

164. Pace C, Keller J, Nolan D, et al. Population level analysis of human immunodeficiency virus type 1 hypermutation and its relationship with APOBEC3G and vif genetic variation. *J Virol*. 2006;80(18):9259--9269.

165. Paillart JC, Shehu-Xhilaga M, Marquet R, Mak J. Dimerization of retroviral RNA genomes: An inseparable pair. *Nat Rev Microbiol*. 2004;2:461--472.

166. Paprotka T, Venkatachari NJ, Chaipan C, et al. Inhibition of xenotropic murine leukemia virus-related virus by APOBEC3 proteins and antiviral drugs. *J Virol*. 2010;84(11):5719--5729.

167. Payvar F, DeFranco D, Firestone GL, et al. Sequence-specific binding of glucocorticoid receptor to MTV DNA at sites within and upstream of the transcribed region. *Cell*. 1983;35(2 Pt 1):381-392.

168. Peng G, Greenwell-Wild T, Nares S, et al. Myeloid differentiation and susceptibility to HIV-1 are linked to APOBEC3 expression. *Blood*. 2007;110(1):393--400.

169. Peterson DO, Kriz KG, Marich JE, Toohey MG. Sequence organization and molecular cloning of mouse mammary tumor virus DNA endogenous to C57BL/6 mice. *J Virol*. 1985;54(2):525--531.

170. Petit V, Guetard D, Renard M, et al. Murine APOBEC1 is a powerful mutator of retroviral and cellular RNA in vitro and in vivo. *J Mol Biol*. 2009;385:65--78.

171. Pincus T, Rowe WP, Lilly F. A major genetic locus affecting resistance to infection with murine leukemia viruses. II. apparent identity to a major locus described for resistance to friend murine leukemia virus. *J Exp Med*. 1971;133(6):1234--1241.

172. Pinter A, Honnen WJ. O-linked glycosylation of retroviral envelope gene products. *J Virol*. 1988;62(3):1016--1021.

173. Poiesz BJ, Ruscetti FW, Gazdar AF, Bunn PA, Minna JD, Gallo RC. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *PNAS*. 1980;77(12):7415--7419.

174. Prats AC, De Billy G, Wang P, Darlix JL. CUG initiation codon used for the synthesis of a cell surface antigen coded by the murine leukemia virus. *J Mol Biol*. 1989;205(2):363--372.

175. Rambaut A, Posada D, Crandall KA, Holmes EC. The causes and consequences of HIV evolution. *Nat Rev Genet*. 2004;5(1):52--61.

176. Refsland EW, Harris RS. The APOBEC3 family of retroelement restriction factors. *Curr Top Microbiol Immunol*. 2013;371:1-27.

177. Refsland EW, Stenglein MD, Shindo K, Albin JS, Brown WL, Harris RS. Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: Implications for HIV-1 restriction. *Nucleic Acids Res*. 2010;38(13):4274--4284.

178. Rein A. Murine leukemia viruses: Objects and organisms. *Adv Virol*. 2011;2011:403419.

179. Renard M, Henry M, Guetard D, Vartanian JP, Wain-Hobson S. APOBEC1 and APOBEC3 cytidine deaminases as restriction factors for hepadnaviral genomes in non-humans in vivo. *J Mol Biol*. 2010;400(3):323--334.

180. Reuss FU, Coffin JM. The mouse mammary tumor virus transcription enhancers for hematopoietic progenitor and mammary gland cells share functional elements. *J Virol*. 2000;74:8183--8187.

181. Rhee S, Gonzales MJ, Kantor R, Betts BJ, Ravela J, Shafer RW. Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucleic Acids Res*. 2003;31(1):298--303.

182. Roa A, Hayashi F, Yang Y, et al. RING domain mutations uncouple TRIM5a restriction of HIV-1 from inhibition of reverse transcription and acceleration of uncoating. *J Virol*. 2012;86(3):1717--1727.

183. Robinson HL. Retroviruses and cancer. *Rev Infect Dis*. 1982;4(5):1015--1025.

184. Roos WH, Ivanovska IL, Evilevitch A, Wuite GJL. Viral capsids: Mechanical characteristics, genome packaging and delivery mechanisms. *Cell Mol Life Sci*. 2007;64(12):1484--1497.

185. Ross SR. Mouse mammary tumor virus molecular biology and oncogenesis. *Viruses*. 2010;2(9):2000--2012.

186. Ross SR. Are viruses inhibited by APOBEC3 molecules from their host species? *PLoS Pathog*. 2009;5:e1000347.

187. Ross SR. MMTV infectious cycle and the contribution of virus-encoded proteins to transformation of mammary tissue. *J Mammary Gland Biol Neoplasia*. 2008;13(3):299--307.

188. Ross SR, Schofield JJ, Farr CJ, Bucan M. Mouse transferrin receptor 1 is the cell entry receptor for mouse mammary tumor virus. *PNAS*. 2002;99(19):12386--12390.

189. Rous P. A transmissable avian neoplasm (sarcoma of the common fowl). *J Exp Med*. 1910;201:320.

190. Rulli SJ, Mirro J, Hill SA, et al. Interactions of murine APOBEC3 and human APOBEC3G with murine leukemia viruses. *J Virol*. 2008;82(13):6566--6575.

191. Russell RA, Moore MD, Hu WS, Pathak VK. APOBEC3G induces a hypermutation graduent: Purifying selection at multiple steps during HIV-1 replication results in levels of G-to-A mutations that are high in DNA, intermediate in cellular viral RNA, and low in virion RNA. *Retrovirology*. 2009;6:16.

192. Sadler HA, Stenglein MD, Harris RS, Mansky LM. APOBEC3G contributes to HIV-1 variation through sublethal mutagenesis. *J Virol*. 2010;84(14):7396--7404.

193. Sakuma R, Sakuma T, Ohmine S, Silverman RH, Ikeda Y. Xenotropic murine leukemia virus-reltaed virus is susceptible to AZT. *Virology*. 2010;397(1):1--6.

194. Sandkovsky U, Swindells S, Robbins BL, Nelson SR, Acosta EP, Fletcher CV. Measurement of plasma and intracelular concentrations of raltegravir in patients with HIV infection. *AIDS*. 2012;26(17):2257--2259.

195. Sandrin V, Cosset F. Intracellular versus cell surface assembly of retroviral pseudotypes is determined by the cellular localization of the viral glycoprotein, its capacity to interact with gag, and the expression of the nef protein. *J Biol Chem*. 2007;281:528--542.

196. Santiago ML, Benitez RL, Montano M, Hasenkrug KJ, Greene WC. Innate retroviral restriction by Apobec3 promotes antibody affinity maturation in vivo. *J Immunol*. 2010;185(2):1114--1123.

197. Santiago ML, Montano M, Benitez R, et al. Apobec3 encodes Rfv3, a gene influencing neutralizing antibody control of retrovirus infection. *Science*. 2008;321:1343--1346.

198. Sanville B, Dolan M, Wollenburg K, et al. Adaptive evolution of mus Apobec3 includes retroviral insertion and positive selection at two cluster os residues flanking the substrate groove. *PLoS Pathog*. 2010;6(7):e1000974.

199. Sarafianos SG, Das K, Clark AD, et al. Lamivudine (3TC) resistance in HIV-1 reverse transcriptase involves steric hindrance with B-branched amino acids. *PNAS*. 1999;96:10027--10032.

200. Sato K, Izumi T, Misawa N, et al. Remarkable lethal G-to-A mutations in vif-proficient HIV-1 provirus by individual APOBEC3 proteins in humanized mice. *J Virol*. 2010;84(18):9546--9556.

201. Sawyer SL, Emerman M, Malik HS. Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. *PLoS Biol*. 2004;2(9):e275.

202. Schafer A, Bogerd HP, Cullen BR. Specific packaging of APOBEC3G into HIV-1 virions is mediated by the nucleocapsid domain of the gag polyprotein precursor. *Virology*. 2004;328:163--168.

203. Schiff RD, Oliff A. The pathophysiology of murine retrovirus-induced leukemias. *Crit Rev Onc/Hem*. 1986;5(3):257--323.

204. Schulman HM, Ponka P, Wilczynska A, Gauthier Y, Shyamala G. Transferrin receptor and ferritin levels during murine mammary gland development. *Biochim Biophys Acta*. 1989;1010:1--6.

205. Sethi AK, Celentano DD, Gange SJ, Moore RD, Gallant JE. Association between adherence to antiretroviral therapy and human immunodeficiency virus drug resistance. *Clin Infect Dis*. 2003;37(8):1112--1118.

206. Sharma S, Miyanohara A, Friedmann T. Separable mechanisms of attachment and cell uptake during retrovirus infection. *J Virol*. 2000;74:10790--10795.

207. Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral vif protein. *Nature*. 2002;418:646--650.

208. Simon JH, Sheehy AM, Carpenter EA, Fouchier RA, Malim MH. Mutational analysis of the human immunodeficiency virus type 1 vif protein. *J Virol*. 1999;73(4):2675--2681.

209. Simon V, Zennou V, Murray D, Huang Y, Ho DD, Bieniasz P. Natural variation in vif: Differential impact on APOBEC3G/3F and a potential role in HIV-1 diversification. *PLoS Pathog*. 2005;1(1):e6.

210. Slosberg BN, Montelaro RC. A comparison of the mobilities and thermal transitions of retrovirus lipid envelopes and host cell plasma membranes by electron spin resonance spectroscopy. *Biochim Biophys Acta*. 1982;689(2):393--402.

211. Smith AJ, Srinivasakumar N, Hammarskjold ML, Rekosh D. Requirements for incorporation of Pr160gag-pol from human immunodeficiency virus type 1 into virus-like particles. *J Virol*. 1993;67:2266--2275.

212. Soros VB, Yonemoto W, Greene WC. Newly synthesized APOBEC3G is incorporated into HIV virions, inhibited by HIV RNA, and subsequently activated by RNase H. *PLoS Pathog*. 2007;3(2):e15.

213. South TL, Summers MF. Zinc- and sequence-dependent binding to nucleic acids by the N-terminal zinc finger of the HIV-1 nucleocapsid protein: NMR structure of the complex with the psi-site analog, dACGCC. *Protein Sci*. 1993;2(1):3--19.

214. Sova P, Volsky DJ. Efficiency of viral DNA synthesis during infection of permissive and nonpermissive cells with vif-negative human immunodeficiency virus type 1. *J Virol*. 1993;67(10):6322--6326.

215. Stavrou S, Nitta T, Kotla S, et al. Murine leukemia virus glycosylated gag blocks apolipoprotein B editing complex 3 and cytosolic sensor access to the reverse transcription complex. *PNAS*. 2013;110(22):9078--9083.

216. Suspene R, Rusniok C, Vartanian JP, Wain-Hobson S. Twin gradients in APOBEC3 edited HIV-1 DNA reflect the dynamics of lentiviral replication. *Nucleic Acids Res*. 2006;34(4677):-4684.

217. Suspene R, Sommer P, Henry M, et al. APOBEC3G is a single-stranded DNA cytidine deaminase and functions independently of HIV reverse transcriptase. *Nucleic Acids Res*. 2004;32(8):2421--2429.

218. Takeda E, Tsuji-Kawahara S, Sakamoto M, et al. Mouse APOBEC3 restricts friend leukemia virus infection and pathogenesis in vivo. *J Virol*. 2008;82(22):10998--11008.

219. Tanese N, Telenitsky A, Goff SP. Kinetics of synthesis, structure and purification of avian sarcoma virus-specific DNA made in the cytoplasm of acutely infected cells. *J Virol*. 1991;65(8):4387--4397.

220. Temin HM, Mizutani S. RNA-dependent DNA polymerase in virions of rous sarcoma virus. *Nature*. 1970;226(5252):1211--1213.

221. Teng B, Burand CF, Davidson NO. Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science*. 1993;260(5115):1816--1819.

222. Trang JM, Prejean JD, James RH, Irwin RD, Goehl TJ, Page JG. Zidovudine bioavailability and linear phamacokinetics in female B6C3F1 mice. *Drug Metab Dispos*. 1993;21:189--193.

223. van der Kuyl AC, Berkhout B. The biased nucleotide composition of the HIV genome: A constant factor in a highly variable virus. *Retrovirology*. 2012;9:92.

224. van der Lugt J, Lange J, Avihingsanon A, et al. Plasma concentrations of generic lopinavir/ritonavir in HIV type-1 infected individuals. *Antivir Ther*. 2009;14(7):1001--1004.

225. Van Valen L. A new evolutionary law. *Evol Theory*. 1973;1:1--30.

226. Varela-Echavarria A, Garvey N, Preston BC, Dougherty JP. Comparison of moloney murine leukemia virus mutation rate with the fidelity of its reverse transcriptase in vitro. *J Biol Chem*. 1992;267:24681--24688.

227. Varmus HE, Heasley S, Kung HJ, et al. Kinetics of synthesis, structure and purification of avian sarcoma virus-specific DNA made in the cytoplasm of acutely infected cells. *J Mol Biol*. 1978;120(1):55--82.

228. Vasudevan AA, Smits SH, Hoppner A, Haussinger D, Koenig BW, Munk C. Structural features of antiviral DNA cytidine deaminases. *Biol Chem*. 2013.

229. Vetter ML, D'Aquila RT. Cytoplasmic APOBEC3G restricts incoming vif-positive human immunodeficiency virus type 1 and increases two-long terminal repeat circle formation in activated T-helper-subtype cells. *J Virol*. 2009;83(17):8646--8654.

230. Vetter ML, D'Aquila RT. Cytoplasmic APOBEC3G restricts incoming vif-positive human immunodeficiency virus type 1 and increases two-long terminal repeat circle formation in activated T-helper-subtype cells. *J Virol*. 2009;83(17):8646--8654.

231. Vogt VM, Eisenman RN. Identification of a large polypeptide precursor of avian orcornavirus proteins. *PNAS*. 1973;70(6):1734--1738.

232. von Schwedler U, Song J, Aiken C, Trono D. Vif is crucial for human immunodeficiency virus type 1 proviral DNA synthesis in infected cells. *J Virol*. 1993;67(8):4945--4955.

233. Wain-Hobson S. Human immunodeficiency virus type 1 quasispecies in vivo and ex vivo. *Curr Top Microbiol Immuno*. 1992;176:181--193.

234. Wang X, Ao Z, Chen L, Kobinger G, Peng J, Yao X. The cellular antiviral protein APOBEC3G interacts with HIV-1 reverse transcriptase and inhibits its function during viral replication. *J Virol*. 2012;86(7):3777--3786.

235. Wight DJ, Boucherit VC, Nader M, Allen DJ, Taylor IA, Bishop KN. The gammaretroviral p12 protein has multiple domains that function during the early stages of replication. *Retrovirology*. 2012;9:83-4690-9-83.

236. Yilmaz A, Bolinger C, Boris-Lawrie K. Retrovirus translation initiation: Issues and hypotheses derived from study of HIV-1. *Curr HIV Res*. 2006;4(2):131-139.

237. Yu Q, Konig R, Pillai S, et al. Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat Struct Mol Bio*. 2004;11(5):435--442.

238. Yu X, Yu Y, Liu B, Luo K, Kong, W., Mao, P., Yu XF. Induction of APOBEC3G ubiquitination and degradation by an HIV-1 vif-Cul5-SCF complex. *Science*. 2003;302(5647):1056--1060.

239. Zachary JF, Baszler TV, French RA, Kelley KW. Mouse moloney leukemia virus infects microglia but not neurons even though it induces motor neuron disease. *Mol Psychiatry*. 1997;2(2):104--106.

240. Zhang H, Dornadula G, Orenstein J, Pomerantz RJ. Morphologic changes in human immunodeficiency virus type 1 virions secondary to intravirion reverse transcription: Evidence indicating that reverse transcription may not take place within the intact viral core. *J Hum Virol*. 2000;3:165--172.

241. Zhang H, Yang B, Pomerantz RJ, Zhang C, Arunachalam SC, Gao L. The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature*. 2003;424(6944):94--98.

242. Zhang Y, Rassa JC, deObaldia ME, Albritton LM, Ross SR. Identification of the receptor binding domain of the mouse mammary tumor virus envelope protein. *J Virol*. 2003;77(19):10468--10478.

243. Zheng YH, Irwin D, Kurosu T, Tokunaga K, Sata T, Peterlin BM. Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. *Virology*. 2004;78(11):6073--6076.