Mouse strain-specific splicing of Apobec3

A Thesis
Submitted to the Faculty
Of the
WORCESTER POLYTECHNIC INSTITUTE

In Partial Fulfillment of the Requirements for the
Degree of Master of Science
In Biochemistry

By

Ryan E. Casey

APPROVED:

_____	_____	_____
Kristen K. Wobbe, Ph.D.	Victor L. Boyartchuk, Ph.D.	
WPI	UMass Medical School	
Project Advisor	Primary Investigator	

ABSTRACT

Host resolution of viral infection is dependent upon components of the innate and acquired immune system. The mammalian protein Apobec3 plays an important role as part of the immune system's innate defenses through its modification of reverse transcribed viral DNA. Recently, Apobec3 was found to directly inhibit HIV-1 and HBV replication through deaminating newly transcribed deoxycytidine residues to deoxyuridine. The ability of mouse and simian Apobec3 variants to inhibit human retroviruses and vice versa highlights the utility of analyzing cross-species homologues. To better understand this editing enzyme, differentially pathogen-susceptible inbred mice were used as an experimental model.

The purpose of this project is to examine the effects of murine Apobec3 (muA3) alternative splicing on its DNA-editing characteristics. Three distinct Apobec3 isoforms were isolated from pathogen-susceptible BALB/cByJ ("C") inbred mice, and two Apobec3 isoforms came from pathogen-resistant C57BL/6ByJ ("Y") mice. The five muA3 isoforms were cloned, sequenced, and expressed from a constitutive promoter in a haploid *Saccharomyces cerevisia* strain. MuA3 DNA-editing activity was measured via the CAN1 forward mutation assay.

The five isoforms studied in this project were discovered to be strain-specific. One isoform from each mouse strain mutated the yeast CAN1 locus significantly. Additionally, both muA3 isoform mRNAs derived from the pathogen-resistant Y mice were found to persist at a higher level (2.7 -12.4 fold) than any of the C mouse isoforms. This suggests that the absence of exon 5 or some other signal in the Y mice may influence transcript stability. Evidence also suggests that the murine Apobec3 start

codon is actually 33bp upstream of its reference start, with implications for previous research performed using muA3. Sequencing analysis of genomic DNA revealed the presence of a 4bp insertion in a region of BALB/cByJ muA3 which may have disrupted an intronic splicing enhancer signal. Furthermore, a novel BALB/cByJ Apobec3 isoform was characterized. This is the first report of strain-specific processing with regard to muA3.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I extend my warmest gratitude to **Dr. Victor Boyartchuk**, my primary advisor, mentor and friend. The knowledge gained during my two years of research in his laboratory far exceeds that which is detailed in this thesis.

I also thank my WPI advisor, **Dr. Kristin Wobbe**, for her five years of patient mentoring and guidance.

The following University of Massachusetts Medical School researchers were instrumental to this project: **Paul Kaufman, Ph.D.** for generously sharing his yeasts, reagents, laboratory equipment and invaluable expertise. His post-doctoral students, **Corey Smith** and **Judy Erkmann** deserve equal praise for their positive encouragement.

Thanks to **Sujatha Patnala, Oleg Garifulin, Ph.D., ZanMei Qi, Ph.D.,** and **Christian Kraft, Ph.D.**, of the Boyartchuk laboratory, for putting my laboratory hurdles in perspective. I have only the greatest respect for their tireless and passionate pursuit of scientific excellence.

Additionally, I thank **Ms. Angela Chester**, **Mr. Pablo Cortez**, and **Mr. Osvaldo Mercado** for regularly sharing their good humor and wisdom.

My beautiful girlfriend and soon-to-be graduate student **Haley Menard** carried me in more ways than one through the past year. You are truly a wonderful person.

Finally, I thank my family – especially **Dad** and **Mom** – for your lifetime of positive encouragement.

I don't think I would have made it without you guys.

LIST OF FIGURES

LIST OF TABLES

BACKGROUND

Apobec3 and Innate Immunity

Apobec3, named for its similarity to the apolioprotein B mRNA editing protein Apobec1, is a DNA-editing enzyme found in all mammals. Members of the human Apobec3 family, including Apobec3G (A3G), Apobec3F (A3F) and Apobec3B (A3B), were recently linked to host resistance of mutant forms of HIV-1. Research thereafter expanded the range of human Apobec3 targets to include exogenous reverse-transcriptase coding viruses such as Hepatitis-B Virus (HBV) and Human T-cell leukemia virus type I (HTLV-I), as well as a number of endogenous retroelements.

Most research to date has focused on human A3G, though A3A, A3B, A3C, and A3F have been found to possess similar antiretroviral properties *in vitro* (Langlois *et al.*, 2005; Zheng *et al*., 2004). However, the continued presence of these viral infections *in vivo* has led to the investigation of these infections in other mammals. As part of these efforts, researchers have found that the transfection and expression of mouse or primate Apobec3 is capable of inhibiting retroviral replication in human cells (Cullen 2006, Esnault *et al*., 2006, Xu *et al*., 2004). This inhibition is due to species-specific differences in protein structure which is unrecognizable by retroviral defense proteins. For instance, the HIV-1 virus encodes a *Vif* protein, which recognizes Apobec3 and targets it for degradation. *Vif* can not bind the mouse and primate Apobec3 homologues, which are then able to bind and hypermutate the nascent retroviral DNA. These findings highlight murine Apobec3's potential for better understanding host innate immunity.

Apobec3: Retroviral Defense Protein

Apobec3's role in regulating HIV-1 was first described in 2002, when a team identified Apobec3G (A3G) expression as conferring "non-permissive" status to cells which were formerly "permissive" to HIV-1Δ*Vif* infection (Sheehy *et al.*, 2002). Shortly thereafter, a number of independent groups showed that A3G exerted its antiviral effect through binding the HIV-1Δ*Vif Gag* protein and "hitch-hiking" with the newly synthesized HIV-1 *ΔVif* virion (MacDuff and Harris, 2006). Upon entering a target cell, the HIV-1Δ*Vif* virion unpackages to reverse transcribe its RNA genome, allowing Apobec3G to interact with the RNA/cDNA/reverse transcriptase complex and edit the nascent cDNA minus strand. The uracil-containing HIV-1 is then thought to undergo degradation or integrate into the host genome in an replication-incompetent form (Mangeat *et al.*, 2003; Zhang *et al.*, 2003). Apobec3F was later found to exhibit the same anti-HIV-1 *ΔVif* effect (Zheng *et al.*, 2004). Confirming this *in vitro* data with *in vivo* observations, sequencing of viral DNA from long-term nonprogressor AIDS patients showed hypermutation patterns matching the sequence specificities of A3G and A3F (Cullen, 2006).

However, the majority of humans still succumb to retroviruses such as HIV-1 regardless of Apobec3 expression. This is due to HIV-1's *Vif* protein, which catalyzes the polyubquitination of Apobec3 attempting to gain access to the assembling HIV-1 virion. The *Vif* protein recruits Apobec3 to an E3 uibiquitin ligase, which in turn marks Apobec3 for proteasome-dependent degradation (Yu *et al*., 2003). *Vif* therefore essentially disarms Apobec3 in cells infected with HIV-1. Of note, a 2004 study discovered that a single amino-acid substitution to A3G (<0.0027% total protein) was enough to confer cellular

resistance to infection from HIV-1 producing *Vif* (Xu *et al*., 2004).

In 2005, Apobec3G was found to restrict HIV-1 infection in "resting", non-permissive human $CD4^+$ T cells. This contrasted with "activated" $CD4^+$ T cells, which were permissive to HIV-1 infection. A3G was found to reside in an enzymatically inert high-molecular-mass (HMM) complex in permissive $CD4^+$ T cells while existing as a low-molecular-mass (LMM) complex in non-permissive cells. The authors incorporated recent findings describing viral RNA as critical to the formation of an A3G-HIV-1 complex (Khan *et al*., 2005). By treating HIV-1 activated, permissive cells with RNAase, the HMM Apobec3G changed to a LMM complex. This transformation of the A3G from an RNA-bound HMM state to an active, LMM state rendered the cell non-permissive to HIV-1 infection. Interestingly, sequenced HIV-1 cDNA from non-permissive cells revealed a low level of G→A hypermutation, suggesting A3G acts via some other antiretroviral mechanism (Chiu *et al.*, 2005). This not only provided a possible HIV-1 treatment avenue, but lent to previous reports of Apobec3 proteins acting in a DNA-editing independent manner (Turelli *et* al., 2004; Newman *et al*., 2005). Whether Apobec3 members act entirely through cytidine deamination or whether it utilizes a secondary mechanism is still contentious.

Apobec3 Proteins: Potential Regulator of Genomic Stability

The role of Apobec3 extends to the regulation of endogenous retroelements. This discovery was not surprising, as the 100 million year old Apobec3 gene likely served a purpose prior to the introduction of HIV-1 into the human population around 100 years

ago (MacDuff and Harris, 2006). Retrotransposons are a similarly ancient feature in mammalian genomes, comprising nearly 50% of total genome content (Kazazian, 2004). Of the two characterized classes of retrotransposons, long terminal repeat (LTR) retrotransposons share a high degree of functional similarity with exogenous retroviruses. LTR retrotransposons differ from retroviruses such as HIV-1 due to their inability to move from its cell of origin (hence their categorization as an endogenous rather than exogenous retroelement). Rather, the LTR retrotransposon periodically moves from one genomic location to another via a process involving reverse transcription. Though beneficial for introducing genetic variation necessary for evolution, unchecked retrotransposition activity could lead to gross instability (Kazazian, 2004; Schumacher *et al*., 2005). Thus, researchers were curious whether Apobec3 members are engaged in limiting the movement of such retrotransposons.

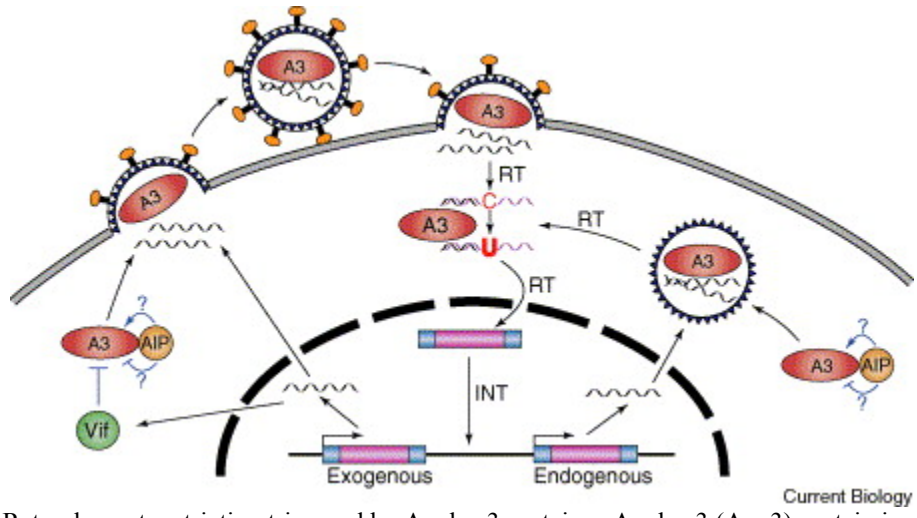

Figure 1: Retroelement restriction triggered by Apobec3 proteins. Apobec3 (Apo3) protein incorporates into assembling retroelement particles. Upon reverse transcription of viral RNA, Apo3 deaminates cytosines in the first DNA strand which then re-integrates in a mutated form (MacDuff and Harris, 2006).

Members of the Apobec3 family limit the movement of yeast LTR retrotransposons primarily via cytidine deamination (**Figure 1**). A3C, A3G, A3F and murine A3 limit LTR

retrotransposition, with follow up research confirming that A3G exerts its antiretroviral activity in a similar manner to previously described work with HIV-1 (Dutko *et al*., 2005). Specifically, the Apobec3 members were associating with the retrotransposon virus-like particles and inducing cytidine deamination of the cDNA transcript. A3G's DNA-editing and subsequent inhibition of LTR retrotransposition was echoed by a 2005 study performed by Harris and colleagues. Interestingly, that study described the cytoplasmically-localizedA3G editing the yeast genome in a similar manner. It is likely that A3G is regulated in mammalian cells as to avoid massive nuclear DNA hypermutation and subsequent cell death.

Human Apobec3 exerts similar antiretroviral activity against murine LTR retrotransposons. A3A, A3B and A3G regulate murine LTR retrotransposition, with A3A acting in non-DNA editing manner (Bogerd et al, 2006). These findings were expanded in human cells, with human, mouse, & monkey Apobec3 proteins, through cytidine deamination, inhibiting murine LTR retrotransposition events (Esnault et al, 2006). The authors concluded that the equal effectiveness of human, mouse and monkey Apobec3G against the LTR retrotransposons demonstrated the strong overall functional conservation of these innate defense proteins during evolution.

Apobec3 thus appears to play a role in regulating genomic stability through control of retroelements. However, preliminary research into this question utilizing Apobec3 knock-out mice found that the gene was apparently inessential to mouse survival, development, and fertility (Mikl *et al*., 2005).

Murine Apobec3

Murine Apobec3 (muA3) is DNA-editing protein belonging to the cytidine and deoxycytidylate deaminase superfamily of enzymes. Like members of the human Apobec3 family, muA3 is thought to play a role as a host retroviral restriction factor (Russell *et al*., 2005; Kobayashi *et al*., 2004). The murine Apobec3 protein also shares the same dual cytidine deamination catalytic sites found in members of the human Apobec3 family as well as a 32% amino acid similarity to its closest homolog. These two enzymatic active sites are comprised of the zinc-coordination ligands C/HXE and PCXXC (**Figure 2**), in which the histidine and the cysteines coordinate zinc while the glutamate participates in deamination (MacDuff and Harris, 2006).

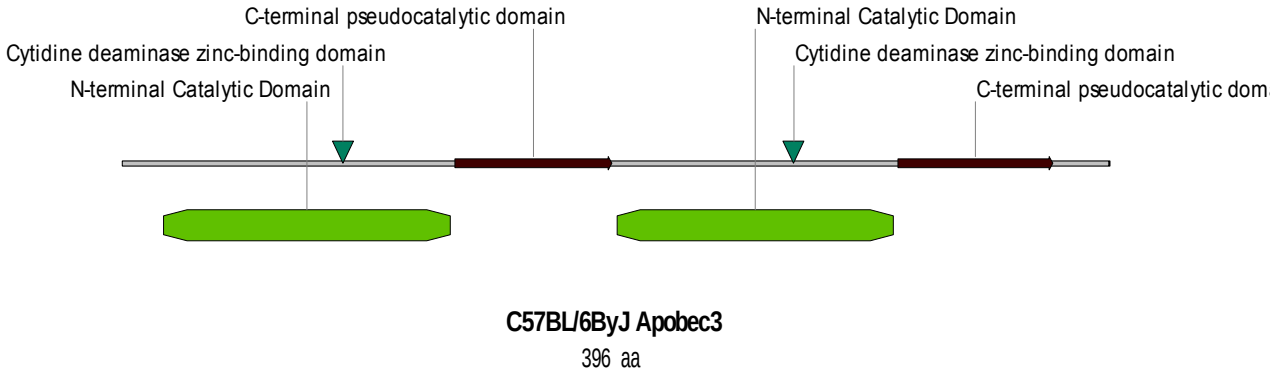

Figure 2: Murine Apobec3 reference protein derived from C57BL/6ByJ mouse. Constructed using Vector NTI software, feature data compiled using the Expert Protein Analysis System proteomics server of the Swiss Institute of Bioinformatics (http://www.expasy.org/).

However, muA3 prefers to edit the deoxycytidine, marked by an asterisk in the viral DNA target sequence TTC*, while human A3G prefers to edit CCC* (Doehle *et al*., 2005). MuA3 expression occurs primarily in the spleen and bone marrow, with a lesser level measured in the thymus.

BALB/c and C57BL/6 Inbred Mouse Strain differential susceptibility to Murine Leukemia Virus: a role for Apobec3?

Differential pathogen sensitivities of inbred mouse strains allows for the identification of host genes critical to disease progression (Risser et al., 1985). Two well-characterized mouse strains (and closely related substrains), BALB/cJ and C57BL/6J, have been identified as possessing differential sensitivity to a number of murine leukemia viruses (Panoutsakopoulou *et al*. 1998, Risser *et al*. 1985, Super *et al*. 1999). Three murine leukemia viruses - E-55+, Friend, and Moloney – are described along with the genomic region thought to be responsible for their restriction. Apobec3 is notably located within these proposed viral control regions.

E-55+ murine leukemia virus (E-55+ MLV), cited as a mouse model for HIV infection, is a chronic retrovirus that affects BALB/c and C57BL/6 mice differently (Ezzel, C. 1993, Avidan *et al*. 1995). Of the three main stages of viral infection – acute, latent and leukemic – the mice differ only during the final period of pathogenesis. During the acute phase, characterized as the first two weeks of E-55+ MLV infection, both mouse strains suffer from large numbers of virus-infected cells in the spleen and bone marrow. The infection then enters a four to seven month latent period, where the virus retreats to the

lymphoid tissue due to a host T-cell mediated immune response (Avidan *et al*, 1995). During this stage of pathogenesis, viral DNA is undetectable in both mouse strains (McEacheron *et al*., 1994). After the latent period, resistant C57BL/6 mice (termed long-term nonprogressors, LTNP), remain chronically infected but do not progress to the final infection stages during their lifetime (Panoutsakopoulou *et al*., 1999). In contrast, susceptible BALB/c mice (termed progressors), advance to the leukemic and eventually terminal phase within seven months of the initial infection.

A 1995 study by Avidan *et al*. demonstrated that the difference in E-55+MLV pathogenesis exhibited by the BALB/c-H-2^k and C57BL/10-H-2^k strains, close relatives of BALB/c and C57BL/6, was controlled by a dominant gene/genes not linked to the major histocompatability complex (MHC). The two strains shared the same MHC haplotype, H-2^k, ruling out differences in antigen presentation as a cause of the differential susceptibility. A F_1 cross (BALB/c-H-2^k X C57BL/10-H-2^k) resulted in progeny with the LTNP phenotype, indicating that resistance to E-55+ MLV is a dominant trait. As the MHC plays a crucial role in the acquired immune response, its irrelevance to E-55+ MLV viral clearance suggests that a component of the innate immune system – of which muA3 is a part – is instead responsible for host resistance.

The applicability of these findings to BALB/c and C57/BL6 mice was later verified by the same laboratory in 1998, where BALB/c and C57BL/6 mice were found to exhibit the same immune response phenotype as BALB/c-H-2^k and C57BL/10-H-2^k mice during the acute phase of E-55+ MLV infection, despite having different MHC haplotypes

(Panoutsakopoulou *et al*., 1998). The BALB/c-H-2^k strain is congenic to BALB/c (which expresses the H-2^d haplotype) and C57BL/10-H-2^k is congenic to C57BL/10 (expresses the H-2^b haplotype). C57BL/10 is closely related to C57BL/6, and both strains share the same H-2^b MHC haplotype.

Panoutsakopoulou *et al*.(1999) proposed a gene, *Rev1*, as responsible for the LTNP phenotype observed in C57BL mice. *Rev1*'s location was investigated using microsatellite-based mapping of LTNP (BALB/c-H-2^k X C57BL/10-H-2^k) F_1 mice with progressor BALB/c-H-2^k mice. Though the ratio of P to LTNP backcross progeny was 1:3, suggesting that two or more loci controlled the LTNP trait, only one region on chromosome 15 was found to be significantly ($p < 0.0001$) linked. The authors noted that the 20-cM linked region of chromosome 15 happened to contain another proposed gene labeled *Rfv3*, involved in host recovery from Friend virus complex (FV). The possibility that *Rev1* and *Rfv3* were the same gene was dismissed as BALB/c and C57BL mice produce the same quantities of E-55+ antibodies while differing in antibody production to FV. Of note is *Apobec3*'s presence within the D15Mit68 and D15Mit32 markers of the *Rev1* candidate region, raising the possibility that it plays a role in viral clearance.

Friend virus complex (FV), composed of a replication-competent Friend murine leukemia virus and a defective spleen focus-forming virus (SFFV), infection results in acute erythroleukemia in susceptible murine strains (Hasenkrug *et al.,* 1995). As with E-55+ MLV, the BALB/c strain is susceptible to the FV infection while C57BL/6 is resistant. The susceptible BALB/c mice fail to produce FV-neutralizing antibodies after infection,

exhibiting viremia before succumbing to FV-induced erythroleukemic splenomegaly. Resistant C57BL/6 mice produce FV-specific antibodies within two weeks of infection, which clear all FV plasma viremia within an additional two weeks (Super et al., 1999).

Preliminary experiments to determine the genetic basis of the differential FV sensitivity were performed by Chesebro and Wehrly (1978). They discovered a single, dominant gene in the resistant C57BL background which was not linked to the MHC. This proposed gene was assigned the name *Rfv3*, recovery from Friend virus number 3. C57BL/6 mice were identified as a $Rfv3^{r}/Rfv3^{r}$ genotype, while the BALB/c mice were termed $Rfv3^{s}/Rfv3^{s}$. The presence of at least one *Rfv3* resistant allele in recombinant mice was enough to induce recovery from FV leukemia in mice possessing the Friend virus complex susceptible MHC haplotype (Super *et al*., 1999).

Subsequent mapping experiments by Hasenkrug et al. (1995) and Super et al. (1999) narrowed the *Rfv3* gene to a region of chromosome 15. The 1995 study mapped *Rfv3* to a 20-centimorgan interval of chromosome 15. Interestingly, the data showed that *Rfv3* was not linked to any of the most obvious candidate genes, such as those controlling immunoglobin light and heavy chain characteristics, the MHC, or T-cell receptors. This region, does, however, contain the *Rev1* region and Apobec3. Thus, *Apobec3* lies within regions mapped to control two viral infections hypothesized as under the control of the innate, rather than acquired, immune system.

Moloney murine leukemia virus (Mo-MLV), a chronic retrovirus, induces T-

lymphoblastic lymphoma or myeloid leukemia in susceptible Babl/c mice while C57Bl/6 mice are relatively resistant (Pak and Faller 1996, Risser 1985). Mo-MLV only encodes viral structural genes and primarily targets T-cells for tumor development. Both mouse strains are resistant to Mo-MLV challenge from early adulthood onwards (Risser *et al*., 1985). Much like E-55+ murine leukemia virus and Friend virus complex, Mo-MLV is characterized by a long latent period after the initial immune response. This latent period lasts for approximately 15 weeks in susceptible BALB/c mice before tumor induction is detectable (Pak and Faller, 1996).

In 1985, Risser *et al.* examined the genetics of susceptibility between BALB/c and C57BL/6 mice using the CXB recombinant inbred mouse strain set. They concluded that more than one gene is responsible for Mo-MLV resistance in BALB/c mice, as nonparental patterns of Mo-MLV resistance emerged in the CXB recombinant inbreds. Risser et al., also found that though BALB/c and C57BL/6 exhibited similar susceptibilities to Abelson murine leukemia virus (A-MLV) as they did to Mo-MLV, the genes controlling this viral sensitivity were separate and unlinked. The authors concluded that, "…genetic control of Mo-MLV disease susceptibility may relate to host responses to viral proteins…" (Risser et al., 1985).

Thus for E-55+ murine leukemia virus, friend virus complex, and Moloney murine leukemia virus, BALB/c is susceptible while C57BL/6 is resistant. *Apobec3* is located within or close to the proposed host susceptibility regions for these viruses, advancing the hypothesis that *Apobec3* may play a critical role during these and other viral infections.

Apobec3 is differentially expressed in BALB/cByJ and C57BL/6ByJ mice

To identify primary differences in gene expression between C57BL/6ByJ (Y) and BALB/cByJ (C) mice, an Affymetrix GeneChip analysis of differential gene expression in spleen tissue was performed previously by the Boyartchuk laboratory. The resulting expression data was then normalized and sorted to identify genes with an absolute average expression difference of at least 2-fold. Among the 377 genes identified by this method was Apobec3.

As depicted in **Table 1**, pathogen-resistant Y mice exhibit 6.64-fold greater muA3 expression level compared to BALB/cByJ mice.

C57BL/6ByJ	C57BL/6ByJ	BALB/cByJ	BALB/cByJ	Student's T-test	Y/C fold difference
1840.4	1902.8	342.6	231.6	0.06387	6.64339

Table 1: muApobec3 expression in spleen tissue. Two mice were processed per strain and their labeled RNA hybridized to a MOE430 Affimetrix GeneChip array. Signal significance was evaluated using the Student's T-test statistical analysis.

L. monocytogenes pathogenesis has been widely utilized as a tool to better understand the innate immune system (Boyartchuk *et al.,* 2004). C57BL/6ByJ mice Apobec3 transcripts are nearly 5 times as prevalent compared to BALB/cByJ mice 24 hours post-infection with the bacteria *Listeria monocytogenes* (**Table 2**).

C57BL/6ByJ	C57BL/6ByJ	C57BL/6ByJ	BALB/cByJ	BALB/cByJ	BALB/cByJ	Student's T-test	Y/C fold difference
1242.5	1693.8	1562.8	335.7	304.9	267.6	0.00028	4.93351

Table 2: muApobec3 expression in spleen tissue 24hrs post-infection with the bacteria *L monocytogenes*.

The observation that recombinant inbred mice displaying differential susceptibility to *L. monocytogenes* also display different levels of muA3 suggests the two events may be

related.

The higher level of muA3 expression in Y mice compared to C mice suggests that the Y mice possess a greater quantity of muA3 protein. This increased presence of muA3 protein and the gene's location in or near a host disease susceptibility locus may thus be related to the inability [v1]of C mice to resist infection relative to Y mice.

The motivation for further investigation into murine Apobec3 was therefore threefold: 1) muA3's structural and functional similarity to human Apobec3, 2) muA3's proximity to a region of the genome hypothesized to regulate murine viral susceptibility, and 3) differential muA3 expression between mouse strains differentially susceptible to viral and bacterial pathogenesis. Current research into muA3 is limited to its antiretroviral properties in the context of HIV-1 or HBV infection, without regard for the existence of different Apobec3 isoforms. Examining Apobec3 strain-specific splicing in the context of DNA-editing activity would ideally lend further understanding of disease susceptibility, pathogen recognition and gene regulation.

The Yeast CAN1 forward mutation assay

In *Saccharomyces cerevisiae* yeasts a 1773bp CAN1 gene encodes a high-affinity arginine permease protein. If this arginine permease is inactivated by mutation, yeast are unable to transport extracellular arginine into the cell (**Figure 3**) and must synthesize the amino acid. L-canavanine is a toxic arginine analog which is also transported by arginine permease. By culturing yeast cells on media containing L-canavanine, L-canavanine is

incorporated into growing peptides with lethal consequences for the cell. Thus, yeast cells which are mutated at the CAN1 gene are capable of growth on L-canavanine containing media. Therefore canavanine-resistance is widely utilized as a means of caculating mutation rates in yeast (Drake, 1991). The CAN1 forward mutation assay detects mutations at the CAN1 locus. This assay only applies when yeasts are grown under high-nitrogen conditions, such as those provided by SD or SC media, since low-nitrogen conditions activate other arginine transport pathways.

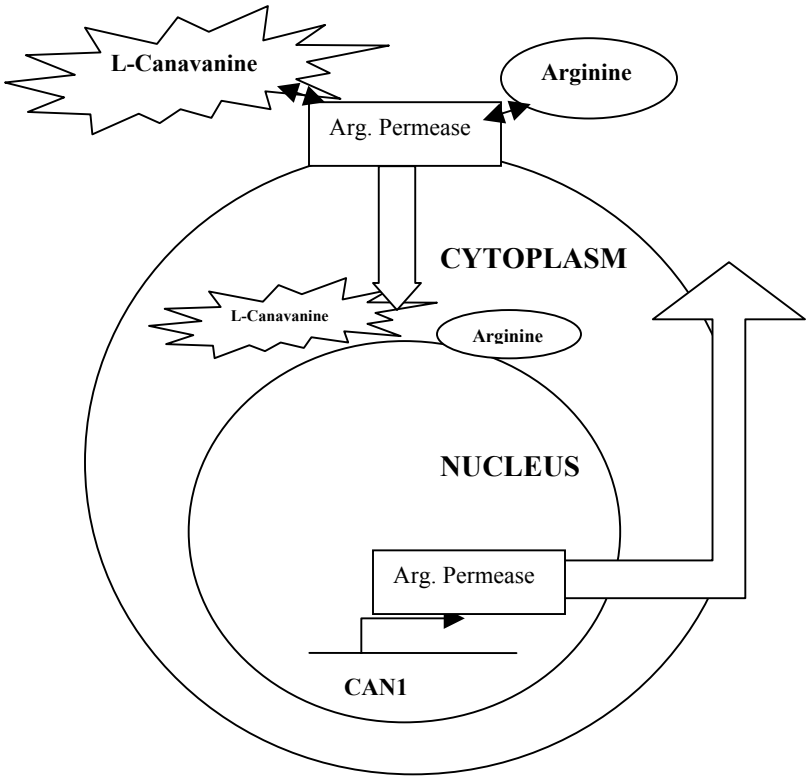

Figure 3: Yeast Arginine Permease pathway. CAN1 gene encodes arginine permease, which exits yeast cell to transport extracellular arginine/L-canavanine back into cytoplasm. Yeast cells which are mutated at CAN1 are unable to transport arginine/L-canavanine and thus survive growth on L-canavanine containing media.

The CAN^R forward mutation assay has been used to examine the regulation of genome stability by both yeast and mammalian genes, including human Apobec3 (Craven *et al.*,

2002; Schumacher *et al*., 2005). In the case of Schumacher *et al*., human Apobec3G was found to hypermutate the CAN1 locus ~18-fold more frequently than the vector control. Human Apobec3G was postulated to act as a regulator of endogenous and exogenous retroelements *in vivo*.

PROJECT PURPOSE

The aim of this investigation was to characterize the DNA-editing activity and strain-specificity of five murine Apobec3 isoforms. The yeast *Saccharomyces cerevisiae* served as a eukaryotic DNA template with which to measure the rate of muA3-induced mutations at the CAN1 yeast gene, as previously reported (Schumacher *et al*., 2005). Elucidating the relation of strain-specific alternative splicing with enzymatic activity and transcript stability will lead to a better understanding of host viral resolution.

METHODS AND MATERIALS

Murine Spleen and Liver RNA isolation

BALB/cByJ and C57BL/6ByJ spleen tissue RNA was isolated using Invitrogen TRIzol® Reagent. All centrifugation steps were performed at 12,000 x g unless otherwise stated. Mice were previously sacrificed and their organs stored in RNAlater solution (Ambion). Tissue samples (50-100mg) were transferred to 1mL TRIzol® and homogenized using a Fisher Scientific Tissuemiser®. The solution was spun for 10 minutes at 4˚C and the supernatant transferred to a fresh 2.0mL tube. After incubating the mixture at room temperature for 5 minutes, 200µL chloroform was added, the tube was shaken by hand for 15 seconds and then incubated at room temperature for 3 minutes. Centrifuged tube for 15 minutes at 4˚C and transferred top aqueous phase to fresh microcentrifuge tube. Incubated tube with 500µL isopropanol at room temperature for 10 minutes followed by a spin down at 4˚C for 10 minutes. The pellet was washed with 75% ethanol, vortexed and then centrifuged at 7,500 x g for 5 minutes at 4˚C. After decanting the supernatant, the pellet was air dried for ~15 minutes, redissolved in 50µL deionised water and stored at -20˚C.

RNA samples were quantitated via optical density (OD) using a Beckman Coulter DU 640 Spectrophotometer.

cDNA synthesis

Oligo $d(T)_{15}$ primer was used to prime the mRNA transcripts. In a microcentrifuge tube, 2µg spleen RNA from either a BALB/cByJ or C57BL/6ByJ mouse was combined with

10uM oligo $d(T)_{15}$ to obtain 0.5µg primer/1.0µg RNA final concentration. The tubes were incubated for 5 minutes at 70˚C and placed on ice. Afterwards, the tubes were briefly spun down. To allow for reverse transcription of the primed mouse mRNA, 75µL of the following solution was created and added to each RNA sample: Promega AMV Reverse Transcriptase (RT) 5X buffer, 1mM complete dNTP mix, 40 units of Promega RNasin Ribonuclease inhibitor, 30 units Promega AMV, and Promega Nuclease-Free H_20. The mixtures were then incubated at 42˚C for 1 hour and stored at -20˚C.

Apobec3 Amplification

Gene specific primers were designed to amplify Apobec3 and Apobec3-X (which utilizes a start codon 33bp upstream of the reference sequence designated start codon) from the BALB/cByJ and C57BL/6ByJ spleen cDNA libraries (**Table 3**). The Apobec3 and Apobec3-X gene primers added an *EcoRI* restriction enzyme site to both the 5' and 3' termini of the PCR products. An endogenous Apobec3 isoform lacking exon 2 ("Apobec3 Δexon2") was inadvertently amplified from the BALB/cByJ mouse using the Apobec3 primers and cloned in subsequent steps.

Target	Primer	Sequence
Apobec3	A3-F-**EcoRI**	AACC**GAATTC**ATGGGACCATTCTGTCTGGGATG
	A3-R-**EcoRI**	AACC**GAATTC**CTTGAATCTCTTCTTGCCTCTCA
Apobec3X	A3X-F-**EcoRI**	AACC**GAATTC**TGCTCACAGAAAATGCAACC
	A3-R-**EcoRI**	AACC**GAATTC**CTTGAATCTCTTCTTGCCTCTCA

Table 3: Primers used to amplify gene of interest and add terminal *EcoRI* restriction sites for future digestion.

A 100µL master mix was created for the Apobec3 and Apobec3-X genes: 1X PCR Buffer, filtered H_20, 0.2mM dNTP mix, 0.4µM forward and reverse primer, 0.05

Units/µL Sigma Jumpstart DNA Polymerase, and 6.0µL cDNA. PCR reactions were performed using the DNA Engine DYAD – Peltier Thermal Cycler and software version 1.08. Protocol: 1 minute denaturing at 94˚C, then 30 cycles of 30 seconds at 94˚C, 30 seconds at 54˚C, 1 minute and 30 seconds at 68˚C. This was followed by a 10 minute final extension and indefinite hold at 4˚C.

Analysis of Apobec3 amplification by Gel Electrophoresis

The amplified gene products were resolved on a 1.2% agarose gel run at 110V in 1X TAE buffer. 5µL of amplification reaction was mixed with 6X loading buffer (60% glycerol, 10mM Tris, 60mM EDTA, bromophenol blue)and loaded into each well of an agarose gel containing 0.00004% ethidium bromide. Approximately 0.7µg Invitrogen 1kb plus DNA ladder was used as a size standard. Apobec3 DNA presence and size was verified using a UVP BioImaging Systems Epichemi^3 Darkroom at 360nm and analyzed using LabWorks^{TM} Image Acquisition and Analysis Software (version 4.5). Apobec3 was quantified[v2] by measuring absorbance(OD) at 260nm using a Beckman Coulter DU 640 Spectrophotometer.

Apobec3 Purification

PCR product was purified using the PureLink^{TM} PCR Purification Kit (Cat # K3100-01) (Invitrogen) according to manufacturer's protocol. 380µL Binding Buffer HC was added to 95µL PCR product and inverted multiple times. The total volume was transferred to a PureLink^{TM} spin column in a collection tube. The sample was centrifuged at 10,000 x g for 1 minute to bind the DNA to the column and the flow-

through discarded. The column was then washed with 650µL Wash Buffer, centrifuged at 10,000 x g for 1 minute, the flow-through discarded, and the column spun for 2 minutes at 18,000 x g. The spin column was transferred to a labeled 1.7mL PureLink™ elution tube, 50µL 1X New England Biolabs EcoRI digestion buffer was added, and the column allowed to incubate at room temperature for 1 minute. The column was then spun at 18,000 x g for 2 minutes to elute the purified DNA and the final product stored at -20˚C.

Apobec3 and Yeast Expression Vector 5'/3' Enzymatic Modification

Apobec3 and Apobec3-X were enzymatically modified prior to insertion into a correspondingly digested pJR1138, a shuttle vector that can be propagated both in bacteria and yeasts, (gift of Paul Kaufman, UMass Medical School). Apobec3 and Apobec3-X, previously eluted in 50µL 1X NE EcoRI Unique buffer, were digested in the following mix: 1X NE Bovine Serum Albumin and 10 Units EcoRI RE for two to twelve hours at 37˚C. The digested Apobec3 insert was purified again using the PureLink™ PCR Purification Kit.

The pJR1138 vector was digested to accept the Apobec3 and Apobec3-X DNA inserts. A 40µL master mix was created to digest two *EcoRI* restriction sites from 7ug pJR1138: 1X NE *EcoRI* Unique buffer, 1X NE Bovine Serum Albumin and 10 Units *EcoRI* RE. The digestion mix was incubated at 37˚C for a minimum of 12 hours and stored at -20˚C.

To reduce the frequency of vector background during cloning, an alkaline phosphatase

was used ("Antartic Phosphatase", New England Biolabs) to remove the 5' overhang phosphate group. The following components were incubated for 15 minutes at 37˚C in a total volume of 10µL: 1µg (*EcoRI*-digested) pJR1138 vector, 1X Antartic Phosphatase reaction buffer, 1µL Antartic Phosphatase, milli-Q H_20. The modified pJR1138 vector was then purified using the PureLink™ PCR Purification Kit.

Ligation of Apobec3 products into Yeast Expression Vector

Ligation of the Apobec3 and Apobec3-X into yeast expression vector pJR1138 was performed using T4 DNA Ligase enzyme (New England Biolabs). A 10µL mix was created with the following components: 1X T4 Ligase buffer, 400 Units T4 DNA Ligase, 1.0µL digested, purified pJR1138 (0.1ug/µL) and 16µL digested, purified gene insert. The solution was incubated for 3 to 12 hours at ~20˚C and then stored at -20˚C.

Cloning Apobec3 into pJR1138 vector

Selection of individual clones and amplification of the newly created plasmids was performed using Invitrogen™ Subcloning Efficiency™ DH5-α™ chemically competent *E. coli* cells (Cat # 18265-017). 1.0µL plasmid (~5ng DNA) and 25µL thawed DH5-α cells were mixed. The solution was incubated on ice for 30 minutes, heat shocked in a heating block at 42˚C for 30 seconds and incubated on ice for 2 minutes. 180µL of warmed LB broth was added to the cells and the mixture incubated in a water bath at 37˚C for 1 hour. After incubation, the cells were plated at 50µL and 150µL volumes onto LB-Kanamycin (50µg/mL) plates for overnight incubation at 37˚C.

Kanamycin resistant colonies were selected after 20 hours incubation for further cellular proliferation and plasmid purification. The Eppendorf FastPlasmid® Mini kit protocol was followed to obtain pure pJR1138-Apobec3 plasmids. Briefly, a kanamycin resistant colony containing a plasmid was inoculated into 2mL LB-Kanamycin (50ug/mL) media for 12-16 hours. 1.5mL of overnight culture was transferred to a 2mL culture tube, spun down (12,000 x g for 1 minute), and resuspended in 400µL ice-cold lysis solution. The solution was vortexed at maximum speed for 30 seconds, incubated at room temperature for ~3 minutes, transferred to a column assembly and spun down for 1 minute at maximum speed. 400µL wash buffer was added and the assembly spun down again. The filtrate was discarded and the assembly spun down for 1 minute. The column was transferred to a fresh collection tube, 50µL of water was added to the center of the column and spun down for 1 minute. The eluted DNA was quantitated and then stored at -20˚C.

MuApobec3 Sequence Verification

Sequencing of muApobec3 contained within the pJR1138 vector was performed according to UMass Medical Nucleic Acid Facility Guidelines. For each muA3 candidate, the following master mix was created: 4µL H20, 2µL plasmid DNA (0.1-0.2ug/µL), 2µL of 1uM primer, and 2µL of 5X Big Dye. The reaction was run in a PCR cycler with the following program: 95˚C initial denaturation for 20sec, then 28 cycles of 95˚C – 1min, 48˚C – 1min, 60˚C – 4min30sec followed by 4˚C indefinitely. 1µL of 2% SDS was added to each tube which were then incubated at 95˚C for 5 minutes before cooling to room temperature.

The sequencing reactions were processed according to the DyeEx 2.0 Spin Protocol for Dye-Terminator Removal. Spin columns were vortexed briefly, loosened a quarter turn and the bottom closure snapped off. The columns were then placed into a 2mL collection tube and spun for 3 minutes at 750 x g. Columns were transferred to a clean centrifuge tube and the sequencing reaction loaded onto the center of the gel bed. The assembly was spun down for 3 minutes at 750 x g, the spin column was removed and the eluted DNA dried at 90˚C. Samples were sequenced at the UMass Medical School Sequencing Facility. Sequences were compared against the geome.ucsc.edu reference sequences ("NM030255" and "BC003314") using VectorNTI software (Invitrogen) to verify absence of mutations.

Isolation of α-mating type Haploid Yeast

The BY4743 (MATa/α his3Δ1/his3Δ1 leu2Δ0 /leu2Δ0 lys2Δ0/LYS2 MET15/met15Δ0 ura3Δ0 /ura3Δ0) yeast was utilized during this project. The yeast was a gift of Prof. Paul Kaufman, Ph.D., of the University of Massachusetts Medical School. All yeast incubations were performed at 30˚C unless otherwise stated. Media formulations can be found in **Appendix B**.

Frozen yeast stock was streaked onto a YPD plate and grown overnight. An individual colony was streaked from this plate onto a synthetic complete plate containing all essential amino acids except arginine SC/-Arg) and incubated overnight. A colony was also streaked onto a SC plate which lacked leucine supplement (SC/-Leu) to verify the

yeasts' leucine auxotrophy. The yeast was then streaked onto a presporulation plate and left overnight. This was repeated again on a fresh presporulation plate. One colony was innoculated into a 15 x 100mm test tube containing 2mL sporulation media and incubated in a roller drum for 5 days at 25˚C, followed by up to 3 days at 30˚C. Yeast were periodically checked for sporulation, visible as a tetrad formation, via light microscope.

Upon observation of tetrads, the yeast culture was spun down (1,000 x g for 5 minutes) and a small quantity of cells were inoculated into a 1.5ml tube containing 40µL 1M sorbitol by sterile tootpick. To this mixture, 1µL of zymolyase solution (20mg/mL zymolyase in 50% glycerol) was added. After centrifuging for 3-4 seconds, the tube was incubated at 30˚C for ~5 minutes. 300µL of water was added and an aliquot spread down the center of a YPD plate.

Tetrads were dissected using an Olympus BHTU-BH-2 Binocular Microscope. Tetrads were identified and micromanipulated to disperse into their four sub-components. These four spores were placed at 1cm intervals along the plate. This was repeated with at least 4 tetrads. The plate was incubated overnight. After colony growth became visible, the spores were replica plated onto a YPD plate and two minimal media plates containing an "α" and "**a**[v4]" haploid mating strain, respectively. The plates were again allowed to incubate overnight. Mating type of the four spores were determined by inspection for colony growth (α spores grew on the **a** mating strain and vice-versa). Plates were wrapped in Parafilm and stored at 4˚C.

Transformation of Haploid Yeast with muApobec3 plasmids

A colony of an α mating type was grown overnight in 10mL YPD media at 30˚C. The stationary growth phase yeast (OD_{600} >2.0) were added 90mL YPD and incubated until the media reached an OD_{600} between 0.5 and 0.8. The cells were spun down (1,000 x g for 5 minutes), resuspended in H_20, spun down again, and resuspended in 1mL 0.1M Lithium Acetate (pH 7.5). The yeast solution was incubated on a roller drum for ≥2 hours at 30˚C. A master mix was prepared for each transformant and a negative control: 5µL pJR1138-muA3 plasmid (0.1 to 5ug), 40µL salmon single stranded-DNA, and 100µL yeast cells. The yeast was incubated at 30˚C for 30 minutes, 700µL Li-PEG solution (40% w/v PEG 3550 in 0.1M LiAc) was added, and the mixture incubated at 30˚C for ≥ 1 hour. The yeast was incubated in a 42˚C water bath for 15 minutes before 250µL was plated onto SC/-Leucine plates. The plates were allowed to dry and then were stored at 30˚C until colonies appeared.

CAN^R Forward Mutation Assay

Yeast transformant colonies were inoculated into SC/-Leu broth and incubated at 30˚C with agitation for 3-4 days. The vector contains a 2µm constitutive promoter which continuously created muA3 transcripts. The yeast cultures were plated onto SC/-Arg/+L-Canavanine (60mg/L) to obtain CAN^R mutants. Viable cell counts were obtained by plating a dilution to YPD rich medium. Viable colonies were counted after 2 days, and CAN^R colonies after 3 days. Accurate values for mutation frequency were obtained using multiple independent cultures (n= 5) and repeating each experiment three times.

Isolation of yeast total RNA

RNA from all six yeast transformants (C, CΔ2, CX, Y, YX and pJR1138 vector only) [v5]plus untransformed yeast was isolated using a hot phenol extraction method. One colony per yeast strain was inoculated into 3mL YPD media and grown at 30˚C to stationary phase. The stationary yeast was inoculated into 17mL YPD and grown at 30˚C for 2-3 hours until OD_{600} between 0.5 and 0.1. The yeast was then spun down (5 min at 1,000 x g) and resuspended in 1mL AE buffer (50mM NaOAC, pH 5.2 + 10mM EDTA). 65µL 25% SDS and 1mL acidic (pH 4.2) phenol were added to the mixture. The yeast solution was incubated at 60˚C for 10 minutes, with vortexing every minute, followed by a 5 minute incubation on ice. The solution was spun down in a microcentrifuge at >12,000 x g for 15 minutes. The aqueous phase was collected in a new tube and 1mL chloroform added. The mixture was gently shaken and spun down for 10 minutes at 3,000 RPM. The aqueous phase was again collected, spun down in 1mL chloroform to remove any remaining phenol and transferred to a new collection tube. To the RNA solution, 1/10 volume 3M NaOAc (pH 5.2) and 1 volume of isopropanol were added. This solution was spun down at 4˚C for 45 minutes at > 12,000 x g. The pellet was washed with 70% EtOH and spun again at 4˚C for 5 minutes. The supernatant was poured off, the pellet allowed to dry in a ~80˚C heat block, and 50µL milli-Q H_20 added. RNA was quantitated and stored at -20˚C.

Residual DNA was removed from the RNA preparations using the RNase-Free DNase Set (Qiagen,) and RNA was subsequently purified using the RNeasy® Mini Kit [v6](Qiagen,)100µL of RNA sample was mixed with 350µL buffer RLT before adding

250µL EtOH. The sample was loaded onto an RNeasy mini column, centrifuged for 15 seconds at >8,000 x g before adding 350µL buffer RW1 and centrifuging for 15 seconds at >8,000 x g again. 10µL DNase I stock solution was combined with 70µL buffer RDD on the RNeasy silica-gel membrane and incubated at 20-30˚C for 15 minutes. 350µL buffer RW1 was added to the column and spun down for 15 seconds at >8,000 x g. Transferred column to fresh collection tube, added 500µL buffer RPE and spun down for 2 minutes at >8,000 x g. Transferred column to fresh collection tube, added 50µL milli-Q H_20, spun down for 1 minute at 8,000 x g and stored at -20˚C.

Real-time PCR analysis of wt-α strains for presence of muApobec3 mRNA

For real-time PCR verification of muA3 mRNA expression in the yeast transformants, the following 10µL master mix was created for each sample: 5µL RNA (30ng/µL), 3.95µL H_20, 10µL 2X Sybr Green mixture, 0.45µL enzyme mix (4:1 taq polymerase to reverse transcriptase) and 0.6µL 1uM forward + reverse primer. Primers specific for the actin gene were used as an experimental control (**Table 4**. Additionally, a no-reverse transcriptase control was used for each sample. RT-PCR cycles were as follows: 50˚C for 25 minutes, followed by 35 cycles of 94˚C for 20 seconds, 56˚C for 30 seconds, 72˚C for 30 seconds, with a final melting curve analysis from 65˚C to 95˚C, reading every 1˚C.

Primer	Sequence
A3-F-RT-PCR	GCTACTACCACCGCATGAA
A3-R-RT-PCR	TGGACCGAATCTTATCAAGGAA
Actin-F-RT-PCR	ATCACCGCTTTGGCTCCAT
Actin-R-RT-PCR	CCAATCCAGACGGAG[v7]

Table 4: Primers for muA3 and Actin genes used in rt-pcr.

CAN1 Sequencing

Yeast colonies which survived on media containing L-canavanine were isolated and their DNA extracted for PCR and sequencing of the mutant CAN1 gene. One colony from C, YX and pJR1138 yeast transformants was isolated from six independent plates containing L-canavanine. The colony was suspended in 30µL 1M sorbitol, vortexed for 15 seconds, and then heated at 90˚C for 4 minutes. The tube was spun down at ≥12,000 x g for 1 minute and the supernatant transferred to a fresh tube on ice. The following PCR mix was created for each 80µL PCR reaction: 60µL H_20, 8µL 10X Promega PCR buffer, 4.8µL 25mM $MgCl_2$, 0.64µL 25mM dNTP mix, 3µL 25% Triton X-100, 0.7µL Promega Taq Polymerase and 1.7µL 10uM F/R CAN1 Primer mix. The PCR cycles were as follows: 94˚C for 4, with 35 cycles of 94˚C for 30 seconds, 56˚C for 1 minute, 72˚C for 2 minutes, followed by 72˚ for 6 minutes and 4˚C indefinitely. 5µL of the PCR products were run on a gel to verify CAN1 amplification and the DNA purified as previously described. To sequence the CAN1 gene, three internal primers were used (**Table X**). Sequencing was performed as previously described. Sequencing results were analyzed against the reference CAN1 gene using VectorNTI software.

Primer	Sequence
Can1-F	GTTGGATCCAGTTTTTAATCTGTCGTC
Can1-R	TTCGGTGTATGACTTATGAGGGTG
Can1-Seq-R	CAGTGGAACTTTGTACGTCC
Can1-Seq-F1	TTCTGTCACGCAGTCCTTGG
Can1-Seq-F2	AACTAGTTGGTATCACTGCT

Table 5: Primers used to amplify and sequence CAN1 gene from yeast.

RESULTS

5' Upstream start codon identified in BALB/cByJ and C57BL/6ByJ muA3

During the process of designing primers for muApobec3, an ATG codon was identified 33bp upstream of the reference start. Though located in the putative untranslated region, this codon maintained the reading frame of the transcript (**Figure 4**). This added an additional 11 amino acids -QPQRLGPRAGM- to the start of the muA3 peptide.

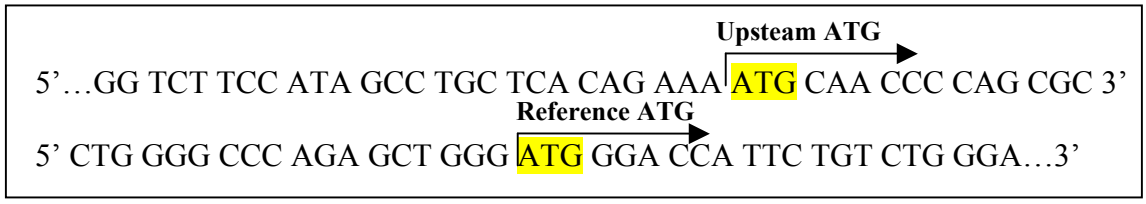

Figure 4: muA3 mRNA sequence at transcription start point (www.genome.ucsc.edu).

Primers were designed to amplify both from the upstream and reference ATG for both mouse strains. Thus, a "C" and "CX" for BALB/cByJ and BALB/cByJ 5' extended start and a "Y and "YX" for C57BL/6ByJ and C57BL/6ByJ 5' extended start isoforms were isolated from spleen tissue. Together, these constituted four of the five isoforms studied.

Apobec3 Alternative Splicing is Strain-Specific

During PCR amplification of mouse spleen cDNA, it was found that Y and C mice express muA3 isoforms in a strain-specific manner (**Figure 5**). C mice expressed a larger form of the gene whereas Y mice expressed an isoform approximately 100bp shorter. Comparison of these isoforms with a mouse genome database suggested these were identical to the two known muA3 sequences: 1287bp (containing exon 5) and 1188bp (missing exon 5).

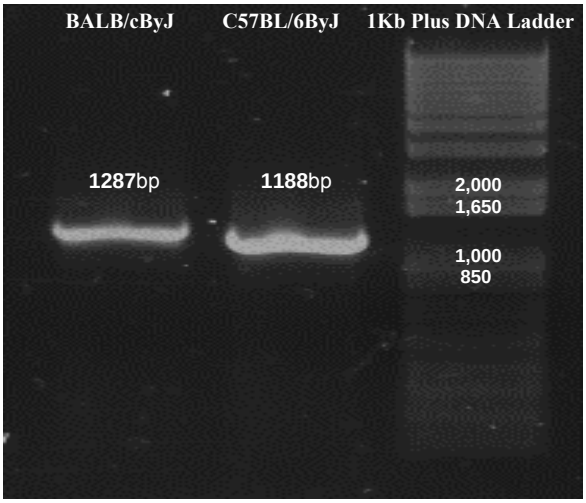

Figure 5: Murine Apobec3 amplified from BALB/cByJ and C57BL/6ByJ spleen cDNA. The BALB/cByJ Apobec3 isoform is larger than the C57BL/6ByJ Apobec3 isoform by ~99bp.

To verify that the two isoforms of muA3 differed with respect to exon 5, primers were created to selectively amplify from exon 4 to exon 5 and from exon 4 to exon 6. PCR and subsequent gel analysis revealed that BALB/cByJ A3 contains exon5, evidenced by a band in the "exon 4→5" column, while C57BL/6ByJ muA3 does not (**Figure 5**). Analysis of C and Y liver cDNA for muA3 exons 5 resulted in the same gene expression profile, though with lower band intensity (data not shown). This is the first report of strain-specific alternative splicing with regard to murine Apobec3.

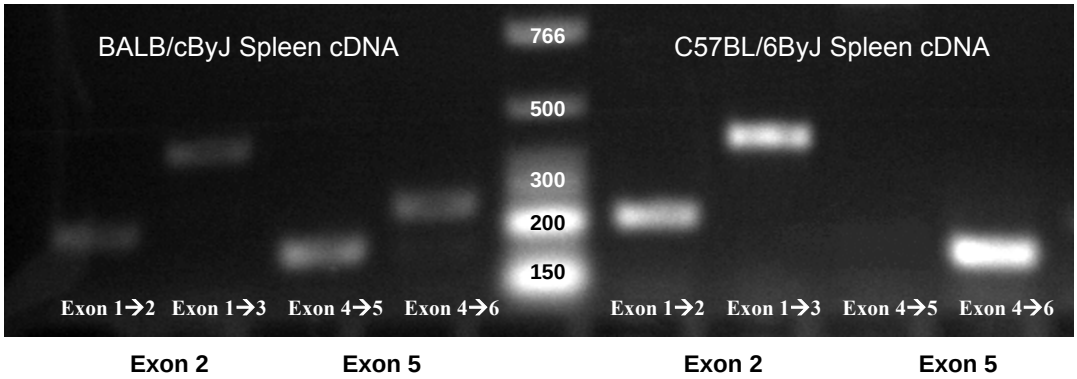

Figure 6: Visualization of Y and C mouse spleen PCR products investigating presence/absence of exon 2 and exon 5 in muA3. C and Y muA3 appear to contain exon 2. C muA3 contains exon 5 while Y muA3 does not.

During the sequencing process, a novel isoform of muA3 was discovered. This isoform is unique to the BALB/cByJ strain and differs from the known BALB/cByJ muA3 in that it lacks exon 2. To confirm that the muA3 Δexon2 transcript exists endogenously, new primers were designed to amplify exons 1 through 3 of muA3. PCR analysis suggests that muA3 Δexon2 transcripts exist in BALB/cByJ spleenic cDNA and is absent from C57BL/6ByJ (**Figure 7**). Though the faint band indicates that an exon 2-deficient muA3 transcript exists in C mice and not in Y, poor visualization conditions and different primers above in **Figure 6** did not independently confirm this observation.

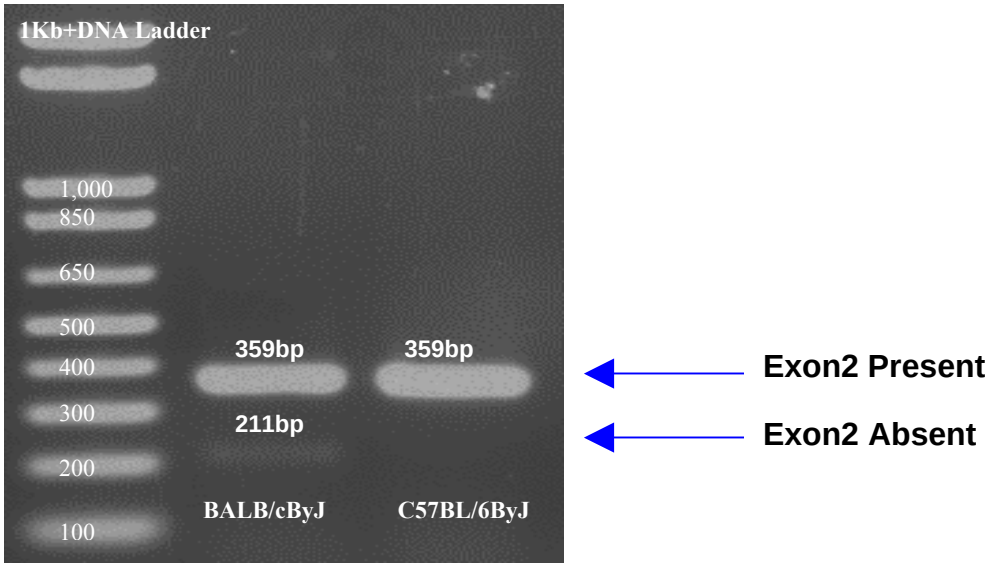

Figure 7: PCR for presence of exon2 in Apobec3 spleen cDNA. Note the faint band present in BALB/cByJ mouse matching predicted 211bp muA3 product missing exon2.

A comparison of the two major strain-specific isoforms with the CΔexon 2 sequence using the blat function of genome.ucsc.edu is depicted below in **Figure 8**.

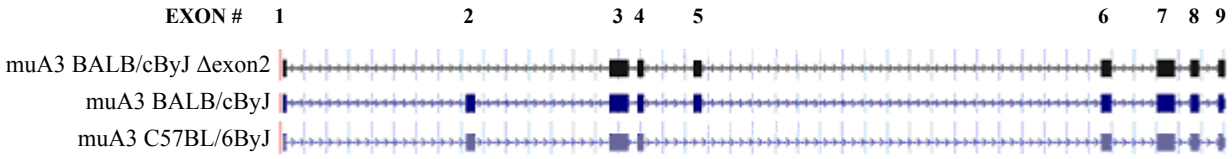

Figure 8: muA3 Δexon2 sequence compared with reference sequences for BALB/cByJ and C57BL/6ByJ mice (genome.ucsc.edu).

BALB/cByJ contains deletion in potential splice enhancer signal 5' to exon 5

To investigate the reason(s) that C57BL/6ByJ and BALB/cByJ mice alternatively splice exon5, 1000bp 5' and 1000bp 3' from exon 5 were sequenced from genomic BALB/cByJ DNA (**Figure 9**). The two intron sequences were then compared against the reference C57BL6/ByJ genomic sequence (www.genome.ucsc.edu).

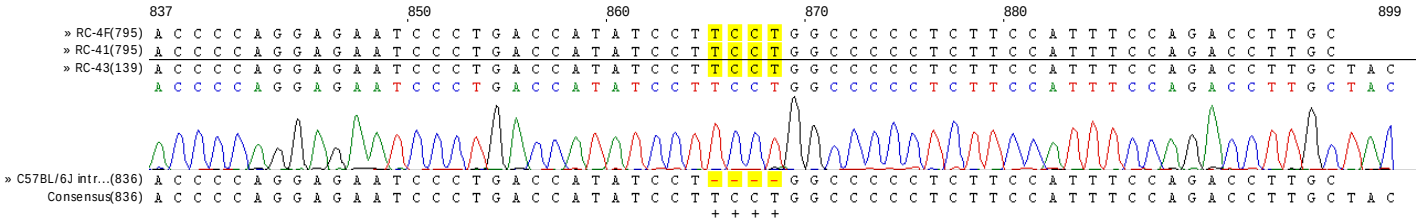

Figure 9: Comparison of C57BL/6ByJ and BALB/cByJ intron 4. The reference sequence for C57BL/6ByJ differs by the absence of a 4bp insert present in the BALB/cByJ DNA. This may influence splicing of exon 5.

The most striking difference between the C57BL/6ByJ and BALB/cByJ mice was the presence of a 4bp insertion in the BALB/cByJ mouse intron 4. This insertion occurs just 28 bp upstream of the start of exon5 and may influence the inclusion of this exon during transcript processing.

muA3 Isoforms Mutate Yeast Genomic DNA at Different Rates

The yeast CAN1 forward mutation assay was utilized to investigate the five muA3 isoforms' editing activity. Haploid yeast were transformed with the five muApobec3 isoforms, plus a vector-only control, to investigate their DNA-editing activity. Transformed yeast containing constitutively-expressed muA3 were grown for 3 days and plated on L-canavanine containing media to select for yeast mutated at the CAN1 gene.

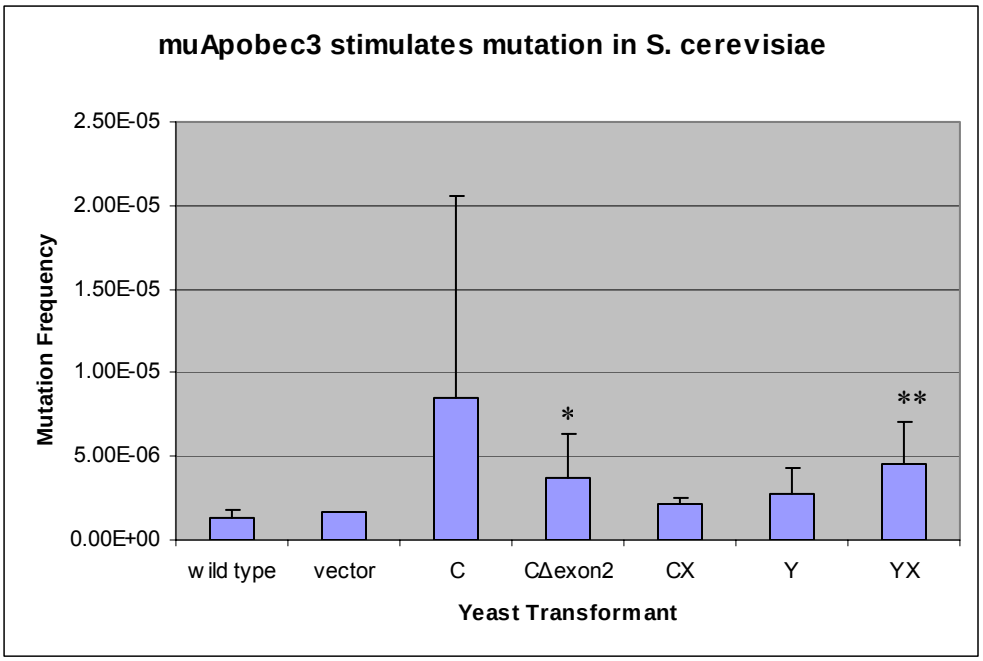

Figure 10: CAN1 mutation rate in haploid yeast. The muApobec3 missing exon2 from the BALB/cByJ mouse and the muApobec3 with the extended start codon from the C57BL/6ByJ mouse both significantly increased the CAN1 mutation rate compared to the vector only (results of 3 independent experiments each analyzing 5 indepenent isolates, * = $p < 0.05$, ** = $p < 0.01$; Error bars = standard deviation).

The results indicate that two of the five isoforms mutated the CAN1 locus with statistical significance. The BALB/cByJ derived CΔexon2 and C57BL/6ByJ derived YX forms of muA3 mutated CAN1 2.12 and 2.60-fold more than the vector control, respectively. This was a surprising finding, as the Y form of muA3 shares 99% sequence homology with YX yet failed to mutate CAN1 significantly. Additionally, the absence of exon2 appears to allow the BALB/cByJ muA3 isoform to mutate CAN1 more than the C or CX forms. Though the C isoform had the highest recorded mutation rate relative to the vector control (4.89-fold), significant variation within the recorded mutation rates rendered the data borderline statistically significant ($p = 0.05$).

C57BL/6ByJ muA3 mRNA stability is greater than BALB/cByJ

Real-time PCR analysis of the transformed yeast was performed to verify the presence of

muA3 mRNAs. RNA was extracted from each of the yeast transformants – C, CΔexon2, CX, Y and YX – plus the vector control and untransformed yeast via a hot phenol extraction method. Yeast mRNAs were then analyzed using rt-pcr for a short fragment spanning muA3 exons 8 to 9. The yeast actin gene served as an internal control, along with a dilution curve of the template RNAs.

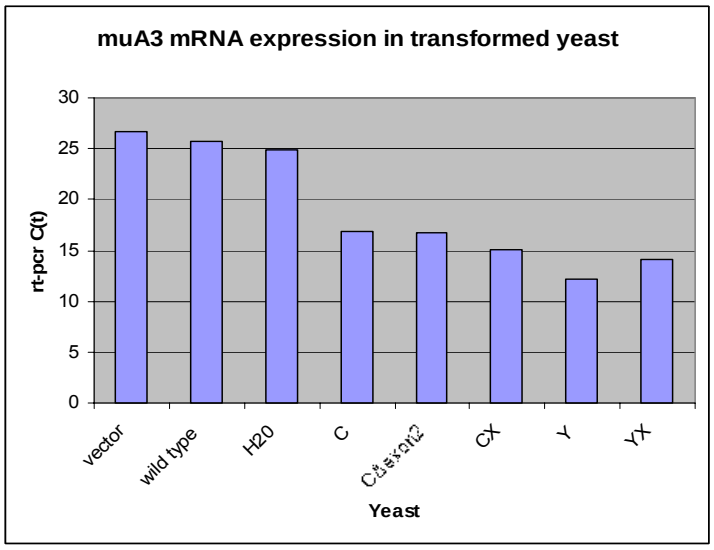

Figure 11: Number of cycles required to detect exponential increase in ds-DNA.

Both the Y and YX muA3 reactions indicated higher mRNA levels relative to C, CΔexon2, and CX (**Figure 11**). This is despite nearly equal actin expression from all 5 samples (data not shown). When the cycle numbers are normalized to an internal mRNA quantitative control, large differences between the C and Y isoforms are apparent (**Figure 12**[v8]).

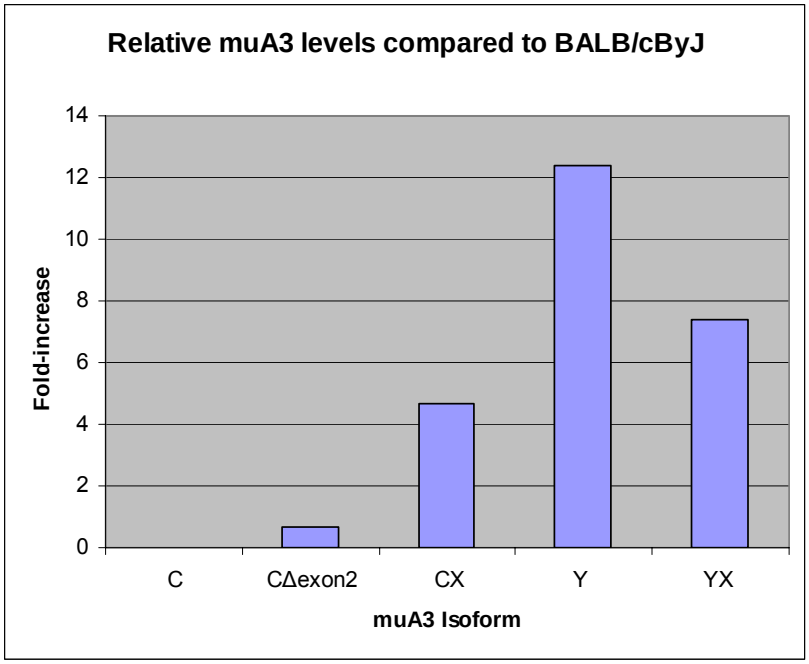

Figure 12: Fold-difference in murine Apobec3 mRNA isolated from transformed yeast. The Y and YX isoforms are the most stable.

The largest difference is the 12.4 fold-difference between the C57BL/6ByJ (Y) and BALB/cByJ (C) transcript levels. Similarly, the YX isoform is 7.4 fold more common than the C isoform. This finding was unexpected, as all five isoforms were constitutively expressed from the same GPD promoter. The observed differences in C57BL/6ByJ (Y and YX) muA3 mRNAs levels compared to the BALB/cByJ (C, CΔexon2, and CX) isoforms in yeast suggest the presence of different secondary structure or other mRNA stability factors.

Additionally, it is notable that the C57BL/6ByJ – YX form of muA3 was expressed at a lower level than Y (**Figure 12**), yet possessed a higher level of catalytic activity **(Figure 10).** Specifically, YX isoform mRNA was 5-fold less prevalent than Y isoform mRNA yet was observed to significantly mutate the CAN1 locus 22% more frequently. This

suggests that the extended start isoform of C57BL/6ByJ may be the true, endogenous form of muA3. However, a protein Blast search using ncbi.nlm.nih.gov for the 11 additional amino acids included in the upstream start peptide did not result in any Apobec3 matches.

DISCUSSION

Alternative splicing effects muA3 DNA-editing activity

The two active DNA-editing muA3 isoforms, BALB/cByJ – CΔexon2 and C57BL/6ByJ - YX, did not mutate the CAN1 locus to the same degree as previously reported for human Apobec3 (Schumacher *et al*., 2005). Instead of the ~18-fold CAN^R increase observed in another haploid yeast strain, the largest CAN^R increase was for C57BL/6ByJ muA3 with a 2.6-fold difference compared to the vector control. The observation of similar mutation rates for the two mouse strains was unexpected. Murine Apobec3, like its human homologue, is capable of inhibiting retroviral infection *in vitro* through DNA-editing. Since the strains differed with regard to viral pathogenesis, it was hypothesized that the Apobec3 isoforms' activity would mirror their progenitor's phenotype. The absence of exon 5, approximately 8% of the Apobec3 transcript, in the C57BL/6ByJ strain was expected to alter the tertiary structure of the protein such that DNA-editing activity would be significantly increased.

Since the CAN^R mutation rates for the Y and C apobec isoforms are not significantly different given the expected mutation rate, it is possible that the difference between the two strains' editing capacity is quantitative instead of qualitative. The observed high concentration of Y and YX muA3 transcripts compared to C, CΔexon2 and CX both in mouse spleen and yeast RNA isolates suggests that the transcripts are differentially stable. The observation that the C57BL/6ByJ – YX form of Apobec3 was enzymatically more active than the Y isoform despite being less prevalent suggests that the YX form may be the true form of muA3 *in vivo*. This has implications for previous research

performed with muA3, as the experimental muA3 would have been a truncated form lacking optimal enzymatic efficiency.

Additionally, the absence of Apobec3 outside of mammals combined with the observation of muA3's editing capacity shows that murine cofactors are unnecessary for muA3 activity.

C57BL/6ByJ vs. BALB/cByJ: Differential muA3 mRNA stability?

Sequence elements within an mRNA transcript determine its half-life. Signals promoting decay or stability are found in the 5' intranslated region (5' UTR), the coding region and the 3' UTR (Knapinska *et al*., 2005). Given that the C57BL/6ByJ and BALB/cByJ mouse muA3 transcripts were derived from cDNA and thus do not possess 5' or 3' UTR regions, it is likely that they contain different coding region instability determinants (CRDs). Further research is necessary to elucidate the presence of instability determinants in the BALB/cByJ muA3 mRNA.

Role of Alternative Splicing in generation of muA3 functional diversity

Recent analysis of microarray data suggests that the number of protein-coding genes in organism does not correlate with its cellular complexity. For example, humans and mice possess approximately 20,000-25,000 protein-coding genes, which is about the same as *Arabidopsis thaliana*. Alternative splicing, the process by which pre-mRNAs are spliced into different arrangements, is believed to play a major role in generating the complexity noted in higher eukaryotic organisms (Blencowe, 2006).

Currently, the most common type of alternative splicing event - approximately 1/3 known events - involves the cassette-type alternative exon (**Figure 13**).

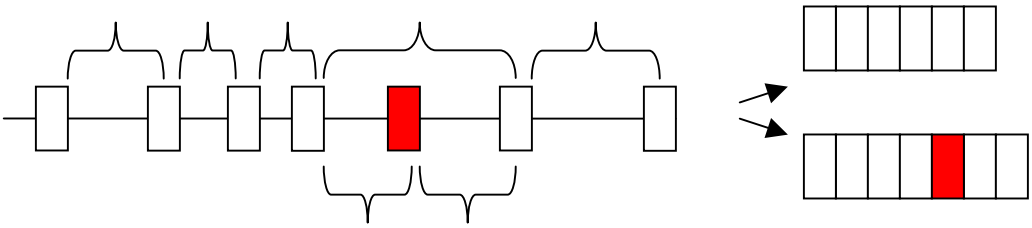

Figure 13: Cassette alternative exon. The fifth exon is alternatively spliced in this example. Adapted from Blencowe 2006.

The intronic/exonic signals responsible for alternative splicing are not yet clear, though data suggests that they lack the information necessary for splicing machinery to distinguish correct from cryptic splice sites (Sun and Chasin, 2000). These signals are generally from 5 to 10bp in length, are degenerate and are recognized by the RNA binding domains of a diverse range of factors during spliceosome formation (**Figure 14**).

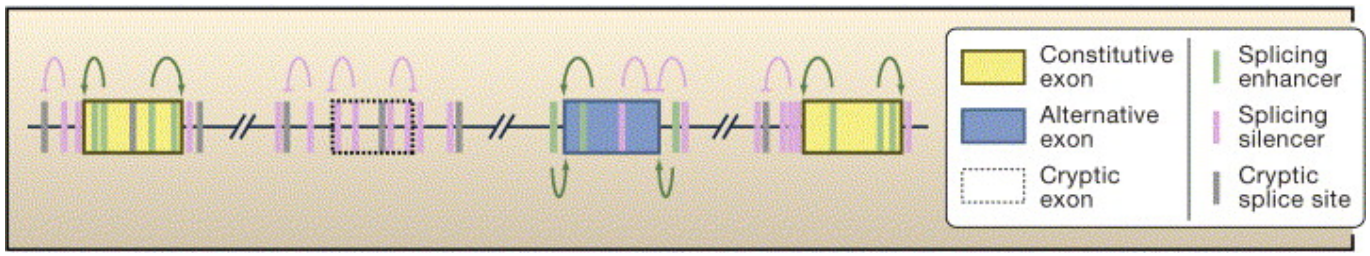

Figure 14: Elements of a putative splicing code. The diagram illustrates approximate relative distributions of splicing enhancers and silencer sequences in pre-mRNA in relation to constitutive, alternative, and pseudoexons, as well as cryptic splice sites (Blencowe, 2006).

Recently, a preliminary code of enhancer and silencer elements was developed to predict splice-site selection (Wang *et al*., 2004). This code, modeled in **Figure 14**, shows how splice enhancer sequences work in tandem to select for alternative exons. The observation of a 4bp deletion in C57BL/6ByJ muA3 3' distal of exon5 raises the possibility that a critical splicing enhancer signal was disrupted. Unlike constitutively spliced exons which contain silencer motifs in the intronic region proximal to the exon, alternatively spliced exons possess enhancer sequences (Wang *et al*., 2004). This disruption effectively may have prevented the splicing of exon5 into the final muA3 transcript. Alternatively, the deletion may have interfered with the ability of an alternating arginine/serine domain containing cofactor (U2AF) to recruit additional splicing elements (such as small nuclear riboncleoprotein particles – snRNPs) during muA3 transcript modification (**Figure 15**). Such a lack of U2AF binding, seen in part B, would lead to the skipping of the alternative muA3 exon in C57BL6/ByJ mice.

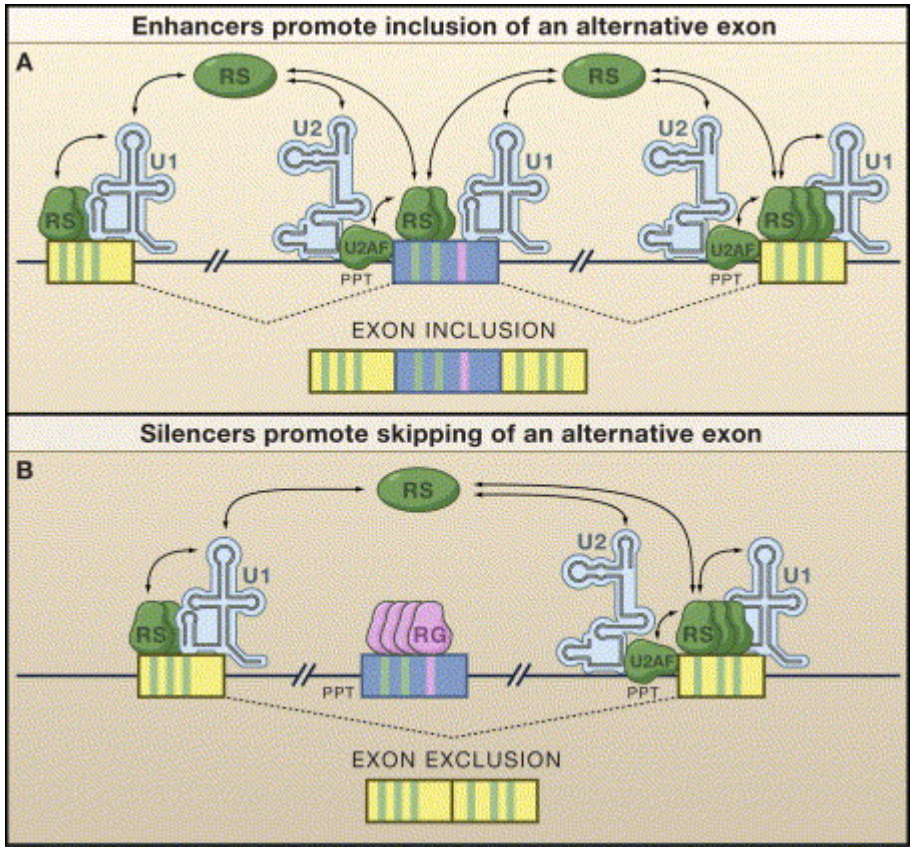

Figure 15: Regulation of Crossintron and Crossexon Interactions by SR and hnRNP Proteins. U1 and U2 small nuclear ribonucleoprotein particles (snRNPs) associate with the 5′ splice site and branch site, respectively. Binding of U2 snRNP to the branch site is promoted by the binding of an alternating arginine/serine (RS) domain-containing factor (U2AF) to the polypyrimidine tract (PPT) (Blencowe, 2006).

Sun and Chasin suggested that intronic suppressor signals play a larger role in alternative splicing than exonic enhancers, strengthening the hypothesis that the 4bp intronic difference between C and Y mice influences splicing (2004).

The observation of a strain-specific low level alternative splicing variant, CΔexon2, was not entirely surprising. Such minor splice variants are proposed to allow for the rapid evolution of modified or novel protein functions under strong evolutionary selection conditions (Xing and Lee, 2005). These variants may not often have an impact upon cellular physiology, while the more abundant variants maintain critical gene activities.

Over time, these minor variants are either lost or positively selected for in the population (Blencowe, 2006). It is feasible that pathogen-susceptible BALB/cByJ mice may be under greater selective pressure to improve its innate defenses compared to C57BL/6ByJ mice and thus has more minor splice variants.

Appendix A: Future directions.

A number of exciting new therapies for the treatment of retroviral infection have been proposed. Among them is the upregulation of Apobec3 members in a cell-specific manner. For instance, A3B inhibits HIV-1 regardless of *Vif*'s presence though it is not normally expressed in HIV-1 target cells (Chiu and Greene, 2006). Likewise, A3G is potentially regulated by IFN-α, which allows cells infected with HIV-1 to overcome infection (Peng *et al*., 2006). Stimulating A3B or IFN-α expression may be a viable treatment option in the future to render retrovirial target cells non-permissive to infection.

Another promising avenue of research is the inhibition of the HIV-1 *Vif* protein. Assuming that upregulation of individual Apboec3's is potentially dangerous to genomic integrity (as evidenced by yeast A3G assays), and IFN-α upregulation similarly chaotic given its connection to a range of signaling pathways, disabling the *Vif* protein has been proposed. One method described in 2004 describes the highly specific recognition pattern that *Vif* requires to identify target Apobec3. Changing even a single amino acid on A3G and other Apobec3's could render the *Vif* and thus entire HIV-1 infection harmless (Xu *et al*., 2004). Other proposed methods include identifying molecules which disrupt the *Vif* – Apobec3G/F interaction and/or blocking the recruitment of E3 ligases by *Vif* (Chiu and Greene, 2006).

Finally, a growing body of work suggests A3G in a catalytically inactive state is able to inhibit retroelements as efficiently as catalytically active A3G (Newman *et al.*, 2005; Schumacher *et al*., 2005; Turelli *et al*., 2004). Defining this activity may provide future

retroviral treatment strategies.

Appendix B: Boyartchuk Laboratory Specialty Yeast Media

Plates (directions for 2L media)

→ SC/-Arg pH 5.9

Materials: Two 2L flasks, Bacteriological Agar, YNB w/out amino acids, SC/(-)Arg AA supplement powder, One stir-bar, 10M NaOH

1) Perform the following for a 2L "media" flask:
    a. Add 13.6g YNB w/out amino acids
    b. Add 3.5g –Arg AA supplement powder
    c. Adjust pH to 5.9 using 10M NaOH
    d. Add large stir-bar
    e. Add 950mL milli-Q H_20
2) Add the following to a 2L "agar" flask
    a. 35g Bacteriological agar
    b. Add 950mL milli-Q H_20
3) Autoclave the flasks together on a **40 minute sterilization cycle**
4) Under fume hood, slowly pour the "agar" flask into the "media" flask on stirrer
5) Add 100mL 40% glucose solution to flask
6) Mix total solution > 5 minutes on stirrer
7) Aseptically pour plates (approx 80 plates/2 L)

→ SC/-Leu pH 5.9

SAME AS **SC/+**, EXCEPT THAT **(-) LEUCINE AA SUPPLEMENT** IS USED IN PLACE OF **(-) Arg AA SUPPLEMENT**!

→ SC/-Arg/Canavanine 60mg/L pH 5.9

SAME AS **SC/-Arg**, EXCEPT THAT 3mL OF L-CANAVANINE SULFATE (20mg/mL) IS ADDED AFTER COOLING FOR **EACH LITER OF MEDIA** (i.e. 6mL canavanine solution is added for a 2L preparation). THE CANAVANINE IS STORED IN THE VB REFRIGERATOR.

→ Presporulation (directions for 500mL)

This recipe is adopted from the Saccharomyces Genome Deletion Project website (http://www-sequence.stanford.edu).
Combine the following in a 1L flask:

5% D-glucose - 25g
3% Difco TSB broth - 15g

1% Difco yeast extract - 5g
2% Bacto agar - 10g
milli-Q H20 to the final volume of 500mL

Autoclave on 15 minute cycle and pour. Plates should be stored at 4C and last up to 2 weeks.

Liquid Media

→ SC/-Arg pH 5.9

SAME AS SC/-Arg PLATE EXCEPT THAT 35g BACTERIOLOGICAL AGAR IS OMITTED

→ SC/-Leu pH 5.9

SAME AS SC/-Leu PLATE EXCEPT THAT 35g BACTERIOLOGICAL AGAR IS OMITTED

AA Supplement Powder Media

→ SC/(-)Arg AA Supplement

Combine 1g of the following amino acids in a 50mL test tube: Ala, Asp, Asc, Cys, Glu, Gln, Gly, Iso, Lys, Met, Phe, Pro, Ser, Thr, Tyr, Val, His, Ura, Trp. Add 1g Adenine hemisulfate. Add 3g Leu. Wrap in aluminum foil.

→ SC/(-)Leu AA Supplement

Combine 1g of the following amino acids in a 50mL test tube: Ala, Arg, Asp, Asc, Cys, Glu, Gln, Gly, Iso, Lys, Met, Phe, Pro, Ser, Thr, Tyr, Val, His, Ura, Trp. Add 1g Adenine hemisulfate. Wrap in aluminum foil.

LITERATURE REVIEW

Avidan, N., Tumas-Brundae, K., Sieck, T.G., Blank, K.J. (1995) Effect of non-H-2-linked genes on anti-virus immune responses and long-term survival in mice persistently infected with E-55+ murine leukemia virus. *Virology*. Jun; 211:507-515.

Blencowe, B. (2006) Alternative Splicing: New Insights from Global Analyses. *Cell*. Jul; 126: 37-47.

Bogerd, H., Wiegand, H., Doehe, B., Lueders, K., Cullen, B.R. (2006) Apobec3A and Apobec3B are potent inhibitors of LTR-retrotransposition function in human cells. *Nucleic Acids Research*. Jan; 34(1): 89-95.

Boyartchuk, V., Rojas, M., Yan, B., Jobe, O., Hurt, N., Dorfman, D., Higgins, D., Dietrich, W., Kramnik, I. (2004) The Host Resistance Locus *sst1* Controls Innate Immunity to *Listeria monocytogenes* Infection in Immunodeficient Mice. *J. Immunology.* 173: 5112–5120.

Chesebro, B. and Wehrly, K. (1978) Identification of a non-H-2 gene (Rfv-3) influencing recovery from viremia and leukemia induced by Friend virus complex. *Proc. Natl. Acad. Sci USA*. Jan; 76(1): 425-429.

Chiu, Y., Greene, W.C. (2006) APOBEC3 Cytidine Deaminases: Distinct Antiviral Actions along the Retroviral Life Cycle. *J. Biol. Chem.* Mar; 281(13): 8309-8312.

Craven, R., Greenwell, P., Dominska, M., Petes, T. (2002) Regulation of Genome Stability by *TEL1* and *MEC1*, Yeast Homologs of the Mammalian ATM and ATR Genes. *Genetics*. June; 161: 493-507.

Cullen, B.R. (2006) Minireview: Role and Mechanism of Action of the Apobec3 Family of Antiretroviral Resistance Factors. *J. Virology*. Feb; 80(3): 1067-1076.

Doehle, B., Schäfer, A., Wiegand, H., Bogerd, H., Cullen, B. (2005) Differential Sensitivity of Murine Leukemia Virus to APOBEC3-Mediated Inhibition Is Governed by Virion Exclusion. *J. Virology*. July; 79(13): 8201-8207.

Drake, J. (1991) A constant rate of spontaneous mutation in DNA-based microbes. *PNAS*. Aug; 88: 7160-7164.

Dutko, J., Schäfer, Kenny, A., Cullen, B., Curcio, M.J. (2005) Inhibition of a Yeast LTR Retrotransposon by Human APOBEC3 Cytidine Deaminases. *Curr. Biol.* Apr; 15: 661-666.

Esnault, C., Heidmann, O., Delebecque, F., Dewannieux, M., Ribet, D., Hance, A., Heidmann, T., Schwartz, O. (2005) Apobec3G cytidine deaminase inhibits

retrotransposition of endogenous retroviruses. *Nature*. Jan; 433: 430-433.

Esnault. C., Millet, J., Schwartz, O., Heidmann, T. (2006) Dual inhibitory effects of Apobec family proteins on retrotransposition of mammalian endogenous retroviruses. *Nucleic Acids Research*. Feb; 34(5): 1522-1531.

ExPASy Proteomics Server. Retrived June 25, 2006 from http://www.expasy.org.

Ezzel, C. (1993) On Borrowed Time; Long-term survivors of HIV-1 Infection. *Journal of NIH Research*. 5(7): 77-82.

Goff. S.P. (2004) Review: Retrovirus Restriction Factors. *Molecular Cell*. Dec; 16: 849-859.

Kasenkrug, K., Valenzuela, A., Letts, V., Nishio, J., Chesebro, B., Frankel, W. (1995) Chromosome Mapping of *Rfv3*, a Host Resistance Gene to Friend Murine Retrovirus. *J. Virology.* Apr; 69(4): 2617-2620.

Kazazian, H.H. (2004) Review: Mobile Elements: Drivers of Genome Evolution. *Science*. Mar; 303: 1626-1632.

Khan, M., Kao, S., Miyagi, E., Takeuchi, H., Goila-Gaur, R., Opi, S., Gipson, C., Parslow, T., Ly, H., Strebel, K. (2005) Viral RNA Is Required for the Association of APOBEC3G with Human Immunodeficiency Virus Type 1 Nucleoprotein Complexes. *J. Virology.* May; 79(9): 5870-5874.

Klinken, S., Hartley, J., Fredrickson, T., Rapp, U., Morse, H.C. (1989) Susceptibility to *raf* and *raf/myc* Retroviruses is Governed by Different Genetic Loci. *J. Virology.* May; 63(5): 2411-2414.

Knapinska, A., Irizarry-Barreto, P., Adusumalli, S., Androulakis, I., Brewer, G. (2006) Molecular Mechanisms Regulating mRNA stability: Physiological and Pathological Significance. *Curr. Genomics*. 6: 1-16.

Kobayashi, M., Takaori-Kondo, A., Shindo, K., Abudu, A., Fukunaga, K., Uchiyama, T. (2004) APOBEC3G Targets Specific Virus Species. *J. Virology.* Aug; 78(15): 8238-8244.

Langlois, M., Beale, R., Conticello, S., Neuberger, M.S. (2005) Mutational comparison of the single-domained Apobec3c and double-domained Apobec3F/G anti-retroviral cytidine deaminases provides insight into their DNA target site specificities. *Nucleic Acids Res.*. April; 33(6): 1913-1923.

MacDuff, D., Harris, R.S. (2006) Directed DNA deamination by AID/APOBEC3 in immunity. *Curr Biol.* Mar; 16(6): R186-R189.

Malkinson, A., Radcliffe, R., Bauer, A. (2002) Quantitative trait locus mapping of susceptibilities to butylated hydroxytoluene-induced lung tumor promotion and pulmonary inflammation in CXB mice. *Carcinogenesis.* Mar; 23(3): 411-417.

Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., Trono, D. (2003) Broad antiretroviral defence by human Apobec3G through lethal editing of nascent reverse transcripts. *Nature.* Jul; 424(6944): 99-103.

Markowitz, D., Goff, S., Bank, A. (1988) A Safe Packaging Line for Gene Transfer: Separating Viral Genes on Two Different Plasmids. *J. Virology*. 62(4): 1120-1124.

McEachron, D., Tumas, K., Blank, K., Prystowsky, M.B. (1994) Environmental Lighting Alters the Infection Process in an Animal Model of AIDS. *Pharmacology Biochemistry and Behavior*. 51(4): 947-952.

Mikl, M., Watt, I., Lu, M., Reik, W., Davies, S., Neuberger, M., Rada, C. (2005) Mice Deficient in APOBEC2 and APOBEC3. *Molecular and Cellular Biology*. Aug; 25(16): 7270-7277.

Newmann, E., Holmes, R., Craigm H., Klein, K., Lingappa, J., Malim, M., Sheehy, A. M. (2005) *Curr Biol.* Jan; 15: 166-170.

Pak, J. and Faller, D.V. (1996) Moloney Murine Leukemia Virus Activates NF-κB. *J. Virology*. Jun; 70(6): 4167-4172.

Panoutsakopoulou, V., Hunter, K., Sieck, T., Blankenhorn, E., Blank, K.J. (1999) Genetic Regulation of Long-Term Nonprogression in E-55+ Murine Leukemia Virus Infection in Mice. *J. Virology*. Nov; 73(11): 9232-9236.

Panoutsakopoulou, V., Little, C., Sieck, T., Blankenhorn, E., Blank, K.J. (1998) Differences in the Immune Response During the Acute Phase of E-55+ Murine Leukemia Virus Infection in Progressor BALB and Long Term Nonprogressor C57BL Mice. *Immunology*. Feb; 161: 17-26.

Peng, G., Lei, K., Jin, W., Greenwell-Wild, T., Wahl, S.M. (2006) Induction of Apobec3 family proteins, a defensive maneuver underlying interferon-induced anti-HIV-1 activity. *J. Expt. Med.* Jan; 203(1): 41-46.

Risser, R., Kaehler, D., Lamph, W.W. (1985) Different Genes Control the Susceptibility of Mice to Moloney or Abelson Murine Leukemia Virus. *J. Virology*. May; 55(3): 547-553.

Russell, R., Wiegand, H., Moore, M., Schäfer, A., McClure, M., Cullen, B. (2005) Foamy Virus Bet Proteins Function as Novel Inhibitors of the APOBEC3 Family of Innate Antiretroviral Defense Factors. *J. Virology*. July; 79(14): 8724-8731.

Schumacher, A., NIssley, D., Harris, R.S. (2005) Apobec3G hypermutates genomic DNA

and inhibits Ty1 retrotransposition in yeast. *Proc. Natl. Acad. Sci.USA*. July; 102(28): 9854-9859.

Sheehy, A., Gaddis, N., Choi, J., Malim, M.H. (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature*. Aug; 418(6898): 646-650.

Stanford SOURCE. Retrieved June 30, 2006 from http://source.stanford.edu.

Sun, H. and Chasin, L. (2000) Multiple Splicing Defects in an Intronic False Exon. *Mol. and Cell Biol.* Sept; 20(17): 6414-6425.

Super, H., Hasenkrug, K., Simmons, S., Brooks, D., Konzek, R., Sarge, K., Morimoto, R., Jenkins, N., Gilbert, D., Copeland, N., Frankel, W., Chesebro, B. (1999) Fine Mapping of the Friend Retrovirus Resistance Gene, *Rfv3*, on Mouse Chromosome 15. *J. Virology.* Sept; 73(9): 7848-7852.

Turelli, P., Mangeat, B., Jost, S., Vianin, S., Trono, D. (2004) Inhibition of Hepatitis B Virus Replication by APOBEC3G. *Science* Mar; 303: 1829.

U.S. National Institutes of Health, National Library of Medicine, National Center for Biotechnology Information. Entrez Gene. [Online] Available at <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed> Accessed April 26, 2006.

Wang, Z., Rolish, M., Yeo, G., Tung, V., Mawson, M., and Burge, C. (2004) Systematic identification and analysis of exonic splicing silencers. *Cell*. 119: 831-845.

Xing, Y., and Lee, C. (2005) Evidence of functional selection pressure for alternative splicing events that accelerate evolution of protein subsequences. *PNAS*. 102(38): 13526-13531.

Xu, H., Svarovskaia, E., Barr, R., Zhang, Y., Khan, M., Strebel, K., Pathak, V. K. (2004) A single amino acid substitution in human APOBEC3G antiretroviral enzyme confers resistance to HIV-1 virion infectivity factor-induced depletion. *PNAS.* April: 101(15); 5652-5657.

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

Zheng, Y., Irwin, D., Kurosu, T., Tokunaga, K., Sata, T., Peterlin, B.M. (2004) Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. *J. Virol.* 78: 6073-6076.