Developing a Single-Cycle Infectious System to Study an ERV-K Retroviral Envelope

Author: Rana Elias Akleh

Boston College Electronic Thesis or Dissertation, 2017

Developing a Single-Cycle Infectious System to Study an ERV-K Retroviral Envelope

Rana Elias Akleh

A thesis submitted to the faculty of

the department of Biology

in partial fulfillment

of the requirements for the degree of

Master of Science

Boston College
Morrissey College of Arts and Sciences
Graduate School

September 2017

Developing a Single-Cycle Infectious System to Study an ERV-K Retroviral Envelope

Rana Elias Akleh

Advisor: Welkin Johnson, Ph.D.

Endogenous Retroviruses (ERVs) are “fossilized” retroviruses of a once exogenous retrovirus located in the genome of extant vertebrates. Retroviral infection results in a provirus integration into the host genome. An infection of a germline cell could lead to the provirus potentially being inherited by the offspring of the infected individual. Once in the genome, the provirus becomes subject to evolutionary processes and can become either lost or fixed in a population, remaining as “fossils” long after the exogenous retrovirus has gone extinct[23].

Notably, 8% of the human genome consists of ERVs[30]. Human Endogenous Retrovirus Type K (HERV-K)(HML-2) family is of particular interest. HERV-K integrations are as old as 30-35 million years, endogenizing before the separation of humans and Old World Monkeys. However, there are human specific insertions, some as young as 150,000 – 250,000 years, making them the youngest insertion in the human genome. There are over 90 insertions in the human genome; the bulk is shared by all humans[44,47]. Transcripts of HERV-K genes are upregulated in multiple cancer and tumor cell lines [14,39,46], as well as in HIV-1 infected patients [7,11,29].

Just as there are human specific insertions of ERV-K, there are also Old World Monkey specific insertions[44]. I have identified an intact endogenous retroviral envelope open reading frame on chromosome 12 of the rhesus macaque genome. This viral envelope-encoding sequence, which I refer to as rhERV-K *env*, retains all the canonical features of a retroviral Env protein. An alignment between rhERV-K *env* and a consensus sequence of HERV-K, HERV-Kcon *env,* shows a 70% amino acid sequence identity.

For experimental purposes, reconstructed HERV-K envelopes have been incorporated into virions of Human Immunodeficiency virus (HIV-1)[19,26,49], Murine Leukemia Virus (MLV)[12], and Vesicular stomatitis Virus (VSV)[26,41,49]. While these approaches have illuminated some aspects of HERV-K Env-mediated entry, to date a cell-surface receptor has not been identified for any ERV-K Env. This could be due to its low infectivity levels[12,26,49], its seemingly broad cell tropism limiting identification of null cell lines[26,49], or possibly the HERV-K consensus reconstructions are not an accurate representation of the progenitor HERV-K virus.

I am interested in understanding how the ERV-K retrovirus accessed the human germline (some 150,000 – 250,000 years ago). To do this, I focused specifically on the envelope proteins of HERV-K and rhERV-K, with the goal of analyzing the ERV-K entry process. The identification and inclusion of rhERV-K Env in this study is meant to circumvent the possibility that the previously described consensus reconstructions of human HERV-K Env are not representative, and may also

provide a means to compare the endogenization process in the human/ape and old-world monkey lineages. I focused on developing two systems for single-cycle infection, one based on Mason-Pfizer Monkey Virus (MPMV) (which has not been done before), and a second based on MLV, which has previously been reported on. MPMV, like HERV-K, is a betaretrovirus, and I reasoned that possibly using a betaretrovirus would overcome some of the low-infectivity issues associated with prior attempts using HIV and MLV.

To develop a system for examining function of the ERV-K Env proteins, I addressed 3 issues:

1. Are the HERV-K Env and rhERV-K Env proteins expressed and properly processed?
2. Can they be incorporated into virions of a heterologous virus?
3. Are ERV-K pseudotyped virions infectious?

I have answered these questions in the following thesis.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis advisor, Professor and Chair of Biology at Boston College, Dr. Welkin Johnson, for his continuous support of my Master of Science study and related research. His patience, motivation, and knowledge were a major part in finishing my thesis. Furthermore, his guidance helped me to the proper career trajectory, and for that I am extremely grateful. Thanks, Jefe!

I would also like to thank everyone in the Johnson lab for their continuous moral support and their troubleshooting advice. Specifically, I would like to thank Andrea Kirmaier and Kate Halm for their intellectual contributions to my research. I would also like to thank my lab mates for making the lab environment an enjoyable one, notably Ismael Ben Fofana, Kate Halm, Laura Hall, Sergio Ita, Andrea Kirmaier, and Anindita Sinha. Also, I would like to thank Elise Gray for her friendship and constant support throughout my time at Boston College.

Furthermore, I would like to thank the staff of the Biology Department at Boston College, significantly Colette McLaughlin, Peter Marino, and Diane Butera for helping me in my transition to Boston College and their constant moral support during my time there. They were always helpful, available, and supportive and I am honored to know such wonderful people.

Lastly and most importantly, I would like to thank my family: my parents, Juheina and Elias Akleh, my brother, Amir Akleh, and my second mother, Noha Akleh, for their unconditional love and support. I also would like to thank Samer Bou Jawdeh for his encouragement and strength during the last part of my research. Knowing I had such a strong support system enabled me to persevere. I love you all.

CHAPTER 1: INTRODUCTION

1.1 Retroviruses

Retroviruses are enveloped single-stranded positive sense RNA viruses which replicate via a double-stranded DNA intermediate that integrates into the host genome becoming simply a new set of cellular genes. The virus can then be transmitted indefinitely by the cell during DNA replication and cell division[50]. Retroviruses have been isolated from numerous animal species, including fish, mice, chicken, cats, cattle, horses, monkeys, and humans[24]. Most notorious of the retrovirus family are the *Lentiviridae*, such as the Human Immunodeficiency Virus (HIV), a sexually transmitted virus discovered in the mid-1980s[50]. This particular retrovirus caused a worldwide epidemic of acquired immune deficiency syndrome (AIDS) which has claimed the lives of millions of people and still remains rampant in many parts of the world.

Retrovirus virions are spherical enveloped particles about 100 nm in diameter and are characterized by the presence of reverse transcriptase (**Figure 1**)[9,50]. The virion contains two copies of the viral positive-stranded RNA genome, complexed with the nucleocapsid protein (NC). The genome is capped at the 5' end and polyadenylated at the 3' end. After infection and reverse-transcription, the DNA provirus is flanked by 5' and 3' Long Terminal Repeats (LTRs), which contain a 150- to 200-nt repeated sequence (R) with unique regions adjacent at either end designated U5 and U3. Downstream of the U5 is the primer binding sequence (PBS) in which a specific cellular transfer RNA (tRNA) binds, serving as the primer for the initiation of synthesis of DNA complementary to the genome RNA.

The retroviral genome consists of three major genes, each translated as a polyprotein: *gag, pol* and *env.* The *gag* gene encodes the nucleocapsid protein (NC), the capsid protein (CA), which forms an icosahedral or conical core, and the matrix protein (MA), which coats the inner surface of the membrane. The *pol* gene encodes the viral enzymes that are associated with the viral core: a protease (PR), an integrase (IN), and a reverse transcriptase (RT). The *env* gene encodes the envelope protein (Env) containing an external surface protein (SU), bound by noncovalent interactions to the transmembrane protein (TM). All these viral proteins are synthesized as polyproteins, which are cleaved by viral and cellular proteases into the individual viral proteins that then assemble to form an infectious virion [9,50].

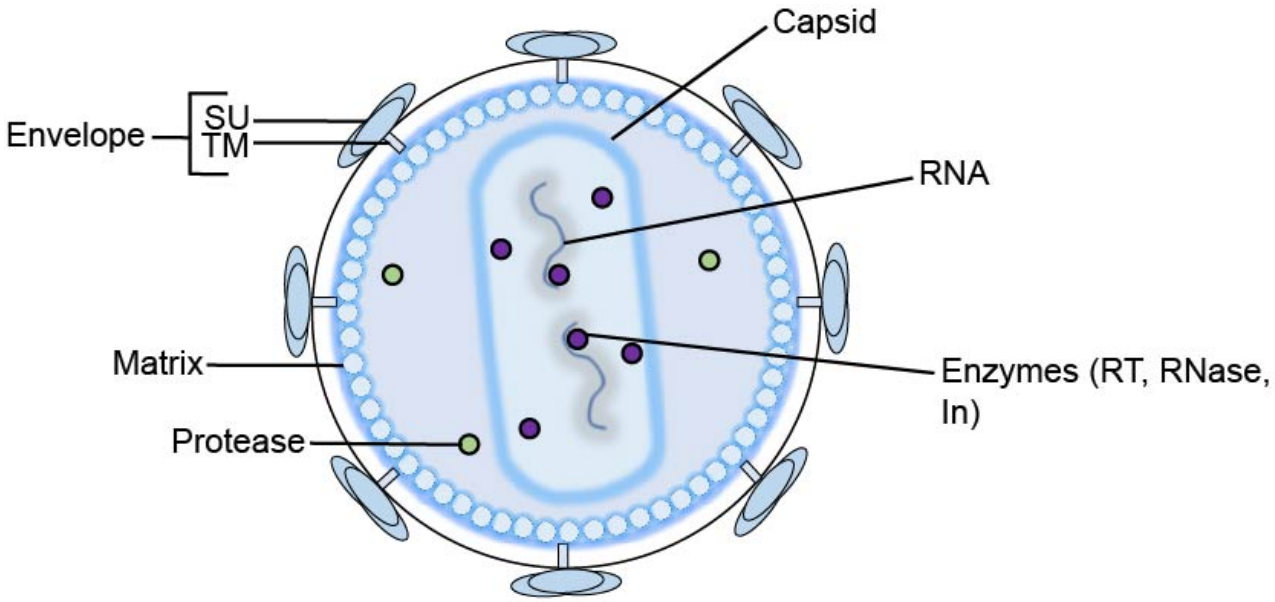

Figure 1. Retrovirus*. A typical retrovirus structure.*

The retrovirus is defined by its unique replication cycle which is based on reverse transcription and integration of its genome (**Figure 2**)[9,50]. A retrovirus recognizes and binds to a receptor on the host cell, which then mediates fusion and entry into the cell. The capsid is released into the cell cytoplasm and the RNA genome is reverse transcribed into double-stranded DNA (dsDNA). The dsDNA is imported into the nucleus, where it is integrated into the host cell genome and is now termed the provirus. The proviral genes become transcribed and translated. Env proteins are translated at the endoplasmic reticulum (ER) and transported through the Golgi apparatus and the endosomal compartment before arriving at the plasma membrane. In contrast, Gag and Gag-Pol proteins are translated in the cytosol and interact with each other to initiate assembly of the viral core. The viral proteins then assemble at the cellular membrane where the virion buds from the host cell; this activates the viral protease, which cleaves Gag and Pol into their individual structural and enzymatic proteins rearranging into a mature virion. At this stage the virion is infectious and can begin its replication cycle anew [9].

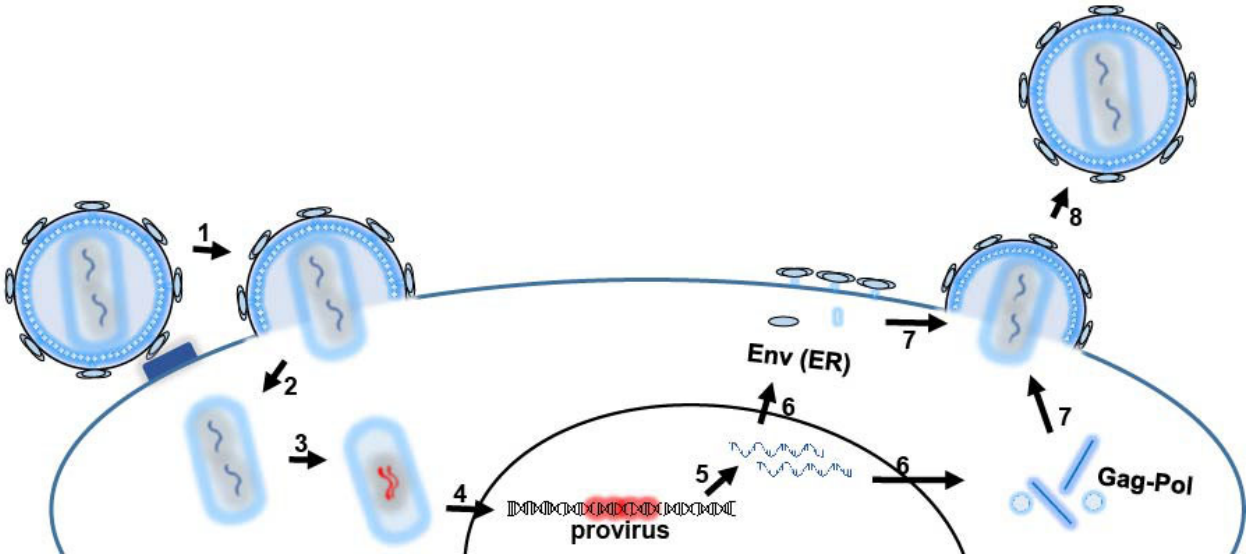

***Figure 2. Retroviral Life Cycle**. **1.** A retrovirus recognizes and binds to the receptor on the host cell mediating fusion and entry into the cell. **2.** The CA is released into the cytoplasm. **3.** The RNA is reverse transcribed into dsDNA. **4.** The dsDNA integrates into the host genome, and is now termed the provirus. **5.** The proviral genes are transcribed and translated. **6.** The Env proteins are transported through the Golgi apparatus arriving at the plasma membrane. The Gag and Pol proteins are translated in the cytosol assembling the viral core. **7.** The viral proteins assemble at the cell membrane. **8.** The virion buds and matures.*

1.2 Endogenous Retroviruses

Endogenous Retroviruses (ERVs) are remnants of a once exogenous retrovirus located in the genome of extant vertebrates [24]. Retroviral infection results in a provirus integration into the host genome (**Figure 3**). If present in the germ line cells, the provirus could potentially be inherited and can persist by vertical transmission. Once in the genome, the provirus becomes subject to evolutionary processes and can become either lost or fixed in a population, remaining as “fossils” long after the exogenous retrovirus has gone extinct. The majority of these “fossilized” retroviruses, however, are unable to produce infectious virions because of mutations that disrupt reading frames or gene expression.

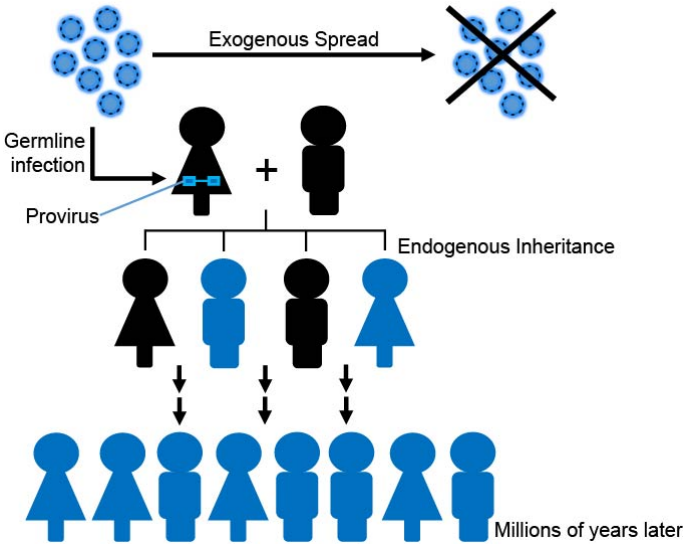

Figure 3. Process of Endogenization*. A retroviral infection of a germline cell results in a provirus integration into the host genome. Once in the host genome, it becomes subject to evolutionary processes and can become lost or fixed in a population remaining as "fossils" long after the exogenous retrovirus has gone extinct.*

ERVs have been found in many vertebrate genomes, including fish [20], mice [2,33], lemurs [17], rabbits [25], colugos [18], koalas [45], armadillos [35], cats[16,22], dogs[16,36], sheep[5], pandas[37], multiple simian species[6,43,48], humans[12,47,49], and many more. Sequencing information has revealed a significant portion of vertebrate genomes consisting of ERVs, with about 8% in the human genome[30], for example. Interestingly, some ERV genes have been adapted by the host organism for new and specialized functions. A well-known example is the *env* gene of human endogenous retrovirus W, *syncytin-1*[15,31]. Syncytin-1 is a cell-cell fusion protein expressing during fetal development which acts in the formation of the fused trophoblast layer of the cells surrounding the fetus. ERVs have also been coopted

for viral resistance. For example, Refrex-1 is a truncated Env encoded by ERV-DC7 and ERV-DC16 that acts as an antiretroviral factor in domestic cats inhibiting infection by feline leukemia virus subgroup D[22,34]. Friend virus susceptibility protein 4 in mice, derived from an endogenous ecotropic Murine Leukemia Virus (MLV), protects from exogenous MLV infection[33,34]. Expression of endogenous Jaagsiekte sheep retrovirus (enJSRV) Env in sheep restricts exogenous JSRV infection[3,4]. These data suggests that these remnants of a once exogenous retrovirus not only serve as “fossils” within the genome but have contributed to genomic evolution in multiple species.

1.3 Human Endogenous Retrovirus Type-K Family

Human Endogenous Retrovirus Type-K (HERV-K)(HML-2) family includes the most recently formed proviral integrations[44,47] in humans found to date. While the presence of HERV-K(HML-2) sequences in humans and Old World Monkeys indicates that the group is 30 – 35 million years old, there are human-only insertions thought to have integrated as recent as 150,000 – 250,000 years ago[44,47]. Thus, the human genome has proviruses that contain nearly intact reading frames, with few or no mutations and proteins that remain functional *in vitro*. However, there are no replication-competent HERV-K viruses found exogenously. There are over 90 proviral insertions and over 1,000 solo long terminal repeats (LTRs) present in the human genome, the bulk of which is shared by all humans[44,47]. Transcripts of HERV-K genes are upregulated in multiple

cancer and tumor cell lines[10,14,39,40], as well as in HIV-1 infected patients[7,11,29], indicating its potential for contributing to human genomic evolution.

Previous work has been conducted with reconstructed ancestral HERV-K consensus sequences, HERV-Kcon and *Phoenix*[12,19,49]. While there is no replication-competent HERV-K virus, expression of the HERV-Kcon and *Phoenix* proviruses produces virus-like particles that weakly infect human and nonhuman cell lines[12,49]. Prior studies on individual HERV-K proteins showed Gag was functionally expressed and processed into MA, NC, and CA[10,13,27]. Gag and Pro proteins mediated efficient assembly and processing of virus-like particles. Furthermore, Env was functionally expressed and processed into its two subunits[12,49]. HERV-K Env from both HERV-Kcon and *Phoenix* reconstructions displayed a broad cell tropism, infecting a range of human and nonhuman cell lines[13,26,42,49]. Env was also shown to inhibit tetherin[32], a restriction factor blocking the release of viral particles from the host cell by keeping them tethered to the cell surface. HERV-Kcon was also shown to resist inhibition from the human restriction factor apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like (APOBEC) 3G, but it is inhibited by APOBEC 3F[32]. These data elucidate possible characteristics for the now-extinct HERV-K virus. Because of the very recent integration of HERV-K, researchers have a rare opportunity to understand the process of endogenization as it is occurring within the human genome through extensive characterization and functional studies of the proviral sequences.

1.4 Retrovirus Envelope

The process of endogenization is still relatively poorly understood. Endogenization begins upon retroviral entry into the host cell, mediated by Env proteins. Studying Env of the most recently endogenized retrovirus family, HERV-K, can elucidate the early stages of endogenization through functional characterization. The HERV-K family resembles retroviruses in the *Betaretrovirus* subfamily[44,47]. Viruses that belong to this group include JSRV, mouse mammary tumor virus (MMTV), enzootic nasal tumor virus (ENTV), and Mason-Pfizer monkey virus (MPMV). The Envs of viruses within this subfamily strongly dictate viral tropism through receptor recognition.

Envelope mediates retroviral fusion and entry into a host cell[9,21]. Unprocessed envelope contains a signal peptide domain (SP), a surface subunit (SU), and a transmembrane subunit (TM) (**Figure 4**). The SP initiates transmembrane synthesis in the ER and ultimately is cleaved off by a host signal peptidase. During Env processing, the Env becomes N-glycosylated in the endoplasmic reticulum and the Golgi apparatus[9]. The host-protease furin then cleaves the SU-TM precursor at the furin cleavage site, generating the two subunits, which are noncovalently associated as hetero-trimers on the cell surface. The SU subunit is responsible for receptor recognition, while the TM subunit mediates fusion and cell entry. Only fully processed envelope can mediate retroviral fusion and entry in the cell. Through functional characterization of HERV-K Env, we can understand how the retrovirus first accessed the human germline 150,000 – 250,000 years ago.

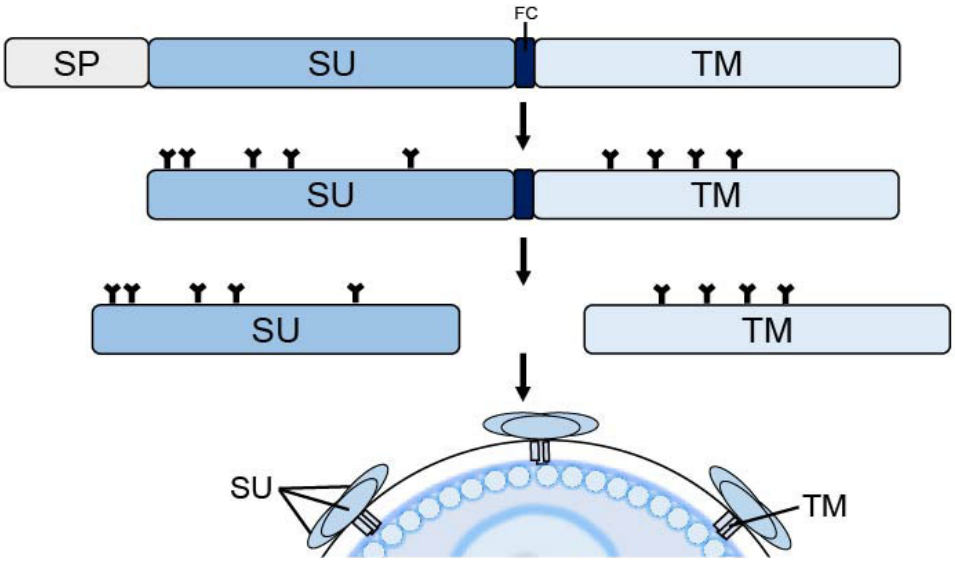

Figure 4. Envelope processing. *Full-length envelope contains a signal peptide (SP) domain which initiates transmembrane synthesis and is cleaved by a host peptidase. The Env becomes N-glycosylated and is cleaved at the furin cleavage site (FC) by endogenous proteases. The SU and TM subunits are noncovalently associated as trimers on the cell surface. SU - surface subunit. TM - transmembrane domain.* ***Y*** *– glycosylation sites.*

Just as there are human-specific insertions of ERV-K, there are also Old World Monkey specific insertions. Focusing on the rhesus macaque genome, I identified an intact endogenous retroviral envelope open reading frame on chromosome 12. Soon thereafter, a group published the same findings of ERV-K proviral insertions on chromosome 12, while also finding two other insertion sites on chromosomes 5 and 11[48]. This viral envelope-encoding sequence, referred to as rhERV-K Env, retains all the canonical features of a retroviral Env protein, and spliced mRNA product encoding rhERV-K Env was isolated from macaque cells[19]. An alignment between rhERV-K *env* and HERV-K *env* shows an 82% amino acid sequence

identity (APPENDIX B). These data suggest that rhERV-K *env* potentially retains its original function.

1.5 Open Questions

To date, there are no exogenous replication competent HERV-K viruses. In order to functionally characterize HERV-K Env, an infectious system using Env must be made. The unique retroviral lifecycle can be exploited in a process called pseudotyping (**Figure 5**). Env and Gag-Pol proteins are translated in different parts of the cell and united at the cell surface to form a virion. We can exploit this by expressing ERV-K Env *in trans* with Gag-Pol from a heterologous retrovirus to create a pseudotyped virions, with which functional studies of ERV-K Env can be conducted.

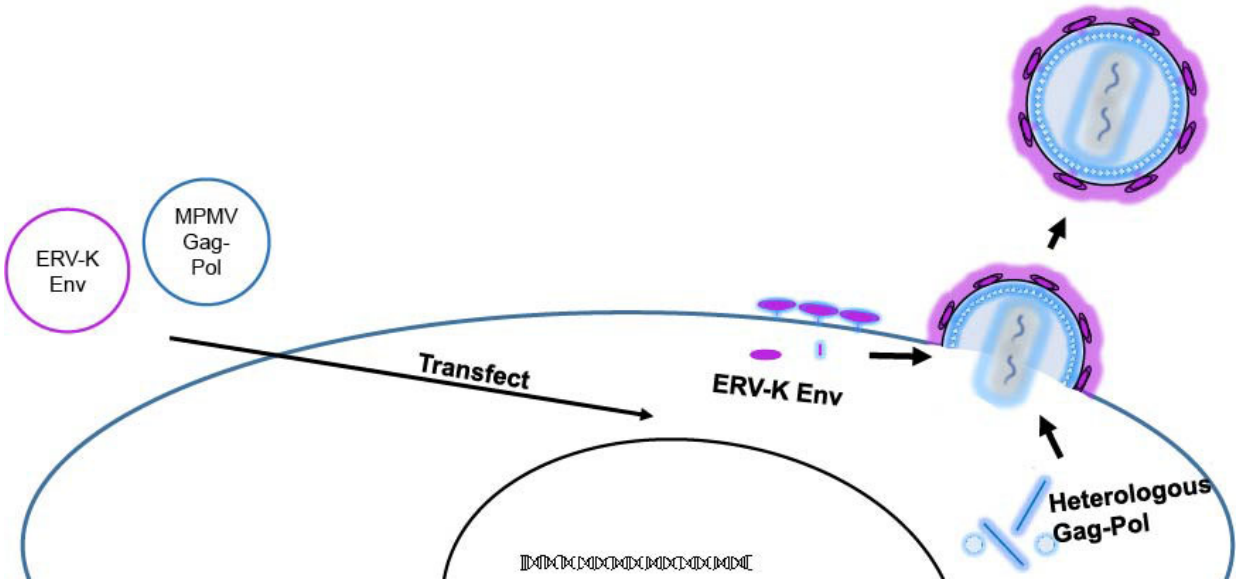

Figure 5. Process of pseudotyping virions*. To make functional ERV-K virions, ERV-K Env are expressed in trans with Gag-Pol of a heterologous retrovirus. The proteins assemble at the cell surface forming an ERV-K pseudotyped virion.*

Infectious systems using HERV-K Env have been established. For experimental purposes, Gag-Pol cores of Human Immunodeficiency Virus (HIV-1)[19,26,49], Murine Leukemia Virus (MLV)[12], and Vesicular stomatitis Virus (VSV)[26,42,49] were pseudotyped with reconstructed HERV-K envelopes. While these approaches have illuminated some aspects of HERV-K Env-mediated entry, to date a cell-surface receptor has not been identified for any ERV-K Env. This could be due to its low infectivity levels[12,26,49], its seemingly broad tropism, limited identification of ERV-K null cell lines[13,26,49], or possibly the HERV-K consensus reconstructions are not an accurate representation of the progenitor HERV-K virus.

To understand how the ERV-K retrovirus accessed the germline, I focused specifically on the envelope proteins of HERV-K and rhERV-K, with the goal of analyzing the ERV-K entry process. The identification and inclusion of rhERV-K Env in this study is meant to circumvent the possibility that previously described consensus reconstructions of human HERV-K Env are not representative, and it may also provide a means to compare the endogenization process in the human/ape and Old-World monkey lineages. I focused on developing a system based on MPMV, which has not been attempted before, and MLV, which has been previously done. MPMV, like HERV-K, is a betaretrovirus, and I reasoned that using a betaretrovirus core could overcome some of the low-infectivity issues associated with prior attempts using HIV and MLV cores.

To develop a system for examining function of the ERV-K Env proteins, I had to address three issues:

1. Are the HERV-K Env and rhERV-K Env proteins expressed and properly processed?
2. Can they be incorporated into virions of a heterologous virus?
3. Are ERV-K pseudotyped virions infectious?

Here, I describe the development of a single-cycle ERV-K infectious system based on MPMV.

CHAPTER 2: MATERIALS AND METHODS

2.1 ERV-K sequences

The *env* gene of the HERV-Kcon reconstructed ancestral HERV-K(HML-2) was previously described[49]. To find the rhERV-K *env*, HERV-Kcon Env was used as a probe in a *tblastn* search in the genome of *Macaca mulatta*. A fully intact rhERV-K *env* was found in the *Macaca mulatta* BAC clone CH250-211O116 from chromosome 12 (AC200900.3). The rhERV-K *env* sequence is described in APPENDIX A.

2.2 Cell lines

Human embryonic kidney 293T cells (HEK293T) (ATCC CRL-11268; American Type Culture Collection, Manassas, VA), human lung carcinoma A549 cells (ATCC CCL-185), and feline kidney CRFK cells (ATCC CCL-94) were maintained at 37°C and 5% CO_2 in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 1% penicillin-streptomycin, and 25mM HEPES.

2.3 Plasmids

The *env* gene of the HERV-Kcon reconstructed ancestral HERV-K(HML-2) and a codon-optimized version of the *env* gene of rhERV-K with an N-terminal AviTag were ordered from GeneArt (ThermoFisher Scientific). Both HERV-Kcon *env* and rhERV-K *env* were cloned into pcDNA3.1(+) and pCRV1 expression vectors (gift of Paul Bieniasz, Rockefeller University) using EcoRV and NotI sites, producing the plasmids pcDNA3.1-HERV-Kcon *env*, pcDNA3.1-rhERV-K *env*, pCRV1-HERV-Kcon *env,* and pCRV1-rhERV-K *env*. Envelope constructs were cloned without the AviTag, pCRV1-HERV-Kcon *env* ΔAviTag and pCRV1-rhERV-K *env* ΔAviTag. pVSV-G (ClonTech) was used as a positive envelope control for

pseudotyping virions. Production of pseudotyped Mason Pfizer Monkey Virus (MPMV) virions was carried out using pSARM-EGFP (produced by WE Diehl). Production of pseudotyped Murine Leukemia Virus (MLV) virions was carried out using pCIGB and pLXIN-EGFP.

2.4 Viruses

To make ERV-K pseudotyped virions, pCRV1 envelope constructs without the AviTag were used. Pseudotyped virions were produced by cotransfection of HEK293T/17 cells with appropriate plasmids using GenJet Transfection Reagent (SignaGen Laboratories, Ijamsville, MD). Briefly, 600,000 – 750,000 HEK293T/17 cells were seeded per well in a 6-well plate the day before transfection. For transfection in a T75 flask, 7.5 x 10^6 cells were seeded. At 48 hours after transfection, supernatant was collected and centrifuged at 1,500 rpm for 5 min. The resulting supernatant was used for infection. Pseudotyped MPMV virions were produced by cotransfection of a 1:1 ratio of envelope (pCRV1-HERV-Kcon *env* ΔAviTag, pCRV1-rhERV-K *env* ΔAviTag, or pVSV-G) and pSARM4-EGFP. Pseudotyped MLV virions were produced by cotransfection of envelope, pLXIN-EGFP, and pCIGB.

2.5 Transfection

600,000 HEK293T/17 cells were seeded in 6-well plates the day before transfection. For protein expression, 1 ug of DNA was transfected using 6 uL GenJet reagent. At 48 hours after transfection, cells were lysed with 200 ul of cold IP Lysis Buffer for 5 – 10 min on ice and added in a 1:1 ratio of 2X Laemmli buffer. To harvest viral supernatant for western blot, the supernatant was centrifuged at

3,000 rpm for 5 min. The supernatant was then added to 6.5 ml of D10 and filtered through a 0.45 µm membrane. The cleared supernatant was layered on 3 ml of 20% sucrose and centrifuged at 35,000 rpm for 1.5 hours at 4℃ (Rotor SW41TI). The media was aspirated and the virion pellet was resuspended in 100 ul of 2X Laemmli buffer. The samples were then heated at 99℃ for 10 min and used for western blot.

2.6 Western Blot

Samples were run on a 10% polyacrylamide gel: 4.1 ml of H_2O, 3.3 ml of 30% Acrylamide, 2.5 ml of 1.5 M Tris-HCl pH 8.8, 100 ul of 10% SDS, 100 ul of 10% Ammonium persulfate (APS), 5 ul of TEMED. The proteins were transferred to a polyvinylidene difluoride (PVDF) membrane in a wet transfer apparatus at 100V / gel overnight at 4℃. Membranes were then blocked with 5% milk in PBST (0.05% Tween-20 in PBS) for at least 1 hour and primary antibody was added overnight at 4℃. The membranes were washed with PBST in 10 minute intervals for 50 minutes and secondary antibody was added for one hour. The membranes were washed as previously described and exposed. The following primary antibodies were used: 1:1000 Mouse Anti-HERV-K Env (AUSTRAL Biologicals HERM-1811-5), 1:500 Rabbit Anti-MPMV CA. The secondary antibodies were at a dilution of 1:5000.

2.7 Infectivity Assay

600 ul of the virion supernatant was added to 50,000 cells per well in a 12-well plate seeded the day before infection. At 4 - 5 hours after infection, 500 ul of warm D10 was added to each well. At 48 hours after infection, cells were imaged for

GFP^+ cells. For FACS analysis, cells were washed with 150 ul of warm trypsin. 150 ul of warm trypsin was added to the cells until no longer adherent. 1 ml of D10 was added to the wells. The cells were centrifuged at 1,500 rpm for 5 minutes and resuspended in 150 ul of 2% paraformaldehyde in PBS.

2.8 Concentrating Virions

7.5×10^6 HEK293T/17 cells were seeded in a T75 flask a day before transfection. 8 µg of DNA was transfected in the cells using the GenJet Transfection Reagant. At 48 hours after transfection, the supernatant was removed and centrifuged at 1,500 rpm for 5 minutes. The resulting supernatant was then added to YM-50 Centriprep Centrifugal Filter Unit with Ultracel-50 membrane columns (Millipore Sigma 4311). The supernatant was centrifuged twice, once at 1,500 g for 10 minutes, and a second time at 1,500 g for 5 minutes. The concentrated virions were added to cells as previously described.

CHAPTER 3: RESULTS

3.1 rhERV-K *env*

To find the rhERV-K *env*, HERV-Kcon Env was used as a probe in a *tblastn* search in the genome of *Macaca mulatta*. A fully intact rhERV-K *env* was found in the *Macaca mulatta* BAC clone CH250-211O116 from chromosome 12 (AC200900.3). The codon optimized rhERV-K *env* was created via GeneArt's algorithm and the gene was ordered via GeneArt (APPENDIX A).

3.2 Expession of ERV-K Envs

To test expression of rhERV-K Env and confirm expression of HERV-Kcon Env, expression plasmids were transfected in triplicate into HEK293T cells and the cell lysates were harvested 48 hours after transfection. Expression was not detectable by western blot using the pcDNA3.1(+) vector. However, when *env*s were cloned into the pCRV1 vector, expression was detected via western blot. A western blot using anti-HERV-K Env antibody showed expression and processing of full-length HERV-Kcon and rhERV-K Envs, with protein products showing up around 103 kDa and 38 kDa, for full-length and processed Env respectively (**Figure 6**). Notably, the antibody against HERV-K Env cross-reacts and can detect rhERV-K Env. Furthermore, the western blots demonstrate that processing of the Envs is incomplete since there were high levels of full-length envelope expression.

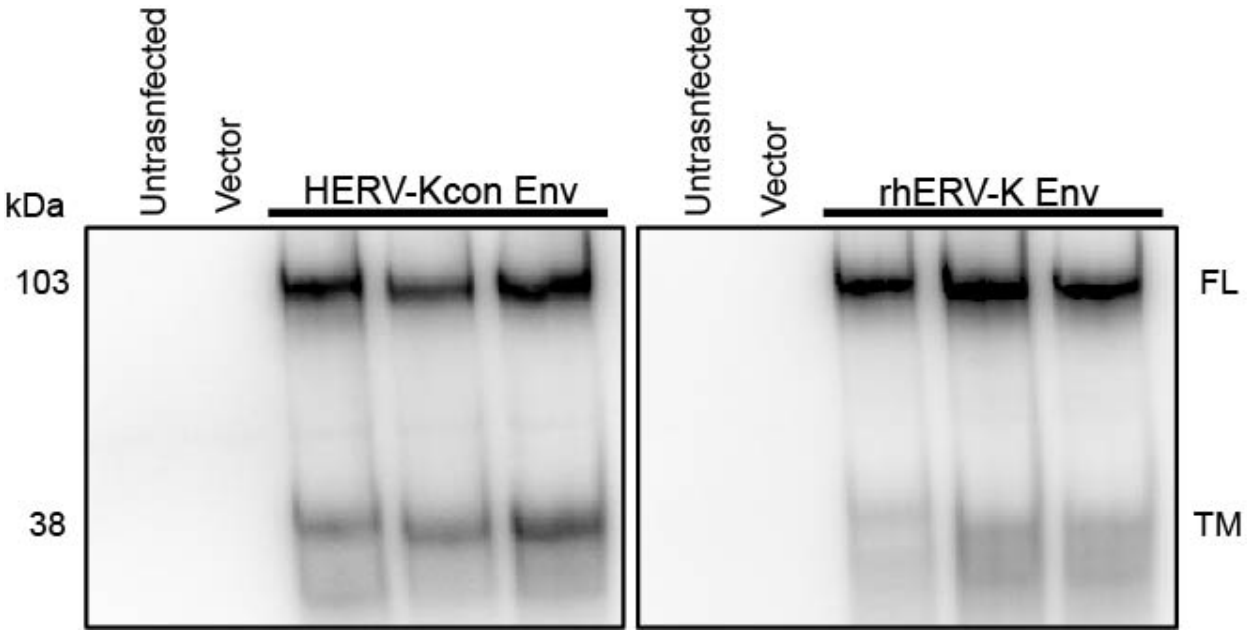

Figure 6. Expression of ERV-K Envelopes*. 1 µg pCRV1 HERV-Kcon env and pCRV1 rhERV-K env were transfected into HEK293T cells and cell lysates were harvested. A western blot against the transmembrane domain (TM) of HERV-Kcon Env revealed full-length (FL) unprocessed and TM subunits; the presence of both a FL and TM band indicates partial processing of the Env polyproteins.*

Expression of the envelopes occurred using the pCRV1 vector, but did not occur in the pCDNA3.1(+) vector because transcripts were unable to export out of the nucleus. pCRV1 encodes HIV-1 Rev and has a Rev Response Element (RRE). RNA transcripts are spliced to produce Rev, which is imported into the nucleus after translation. Rev binds the RRE of the transcribed RNAs before splicing and exports the unspliced and partially spliced transcripts into the cytoplasm[38]. Translation can then occur. The envelopes in the pCDNA3.1(+) vector were transcribing mRNA (RT-PCR data not shown), but translation was not occurring, indicating that mRNA was getting stuck in the nucleus. This problem was

overcome by cloning the envelopes in the pCRV1 backbone, which contained *rev* and an RRE.

3.3 Pseudotyping MPMV particles

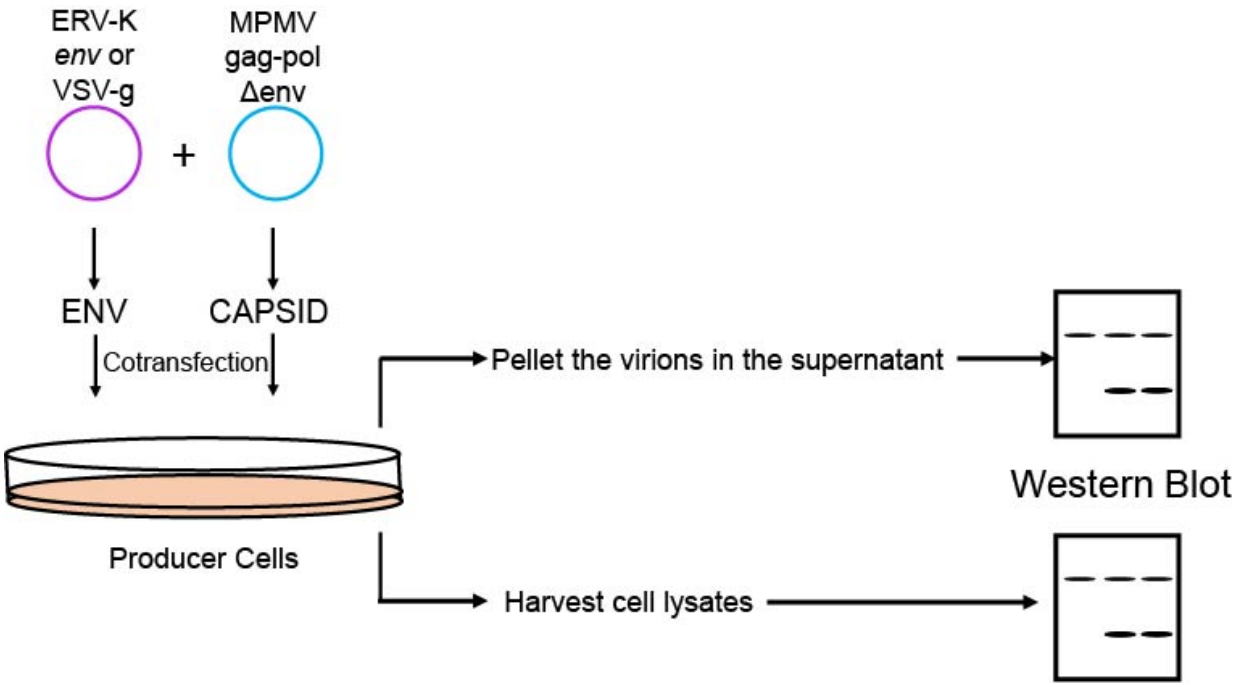

Figure 7. Pseudotying MPMV with ERV-K Envs. *ERV-K envs were cotransfected with MPMV gag-pol. The producers cell lysates were harvested to check for protein expression via western blot. The supernatants were harvested to check for virion release via western blot.*

ERV-K envelope constructs, as well as a VSV-G-encoding construct as a positive control, were cotransfected with MPMV *gag-pol Δenv* plasmids. The producer cell lysates were harvested and visualized by a western blot to look for ERV-K Env and MPMV CA expression (**Figure 7**). A western blot of the producer cell lysates stained for HERV-K Env showed expression and processing of full-length HERV-Kcon and rhERV-K Envs (**Figure 8**). A western blot against MPMV CA showed CA expression

in the producer cells. This data illustrates proper expression of viral proteins in the producer cells.

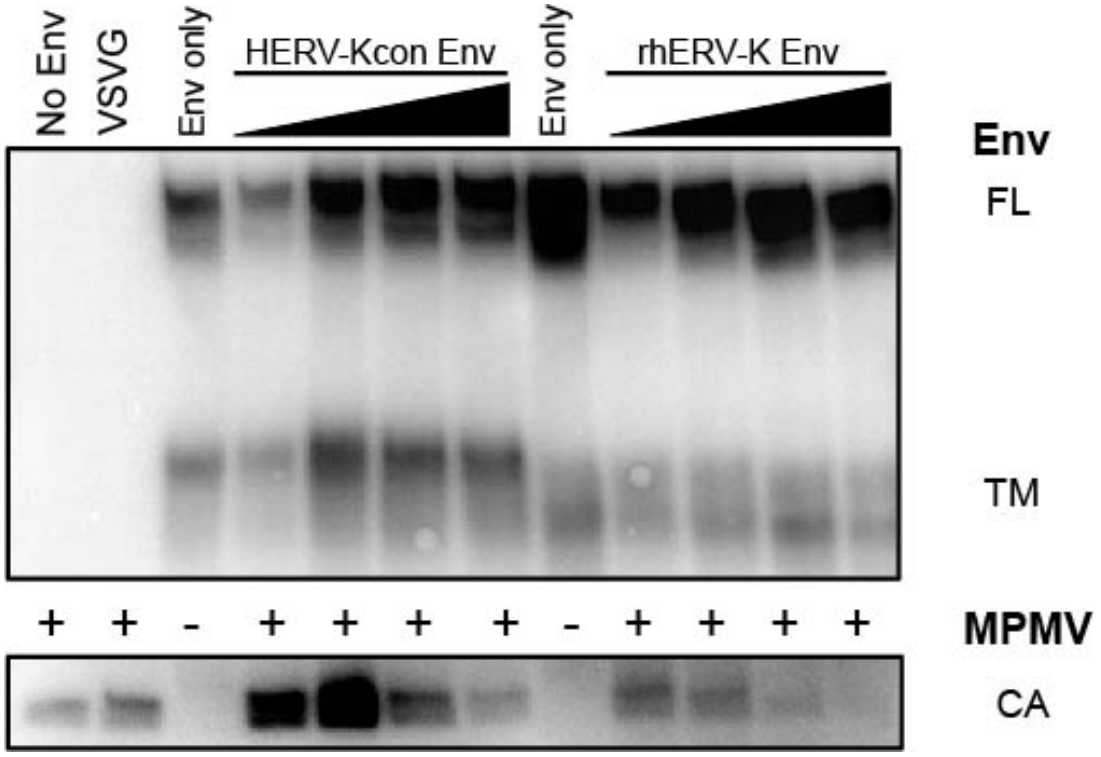

Figure 8. Western blot of Producer Cell Lysates. *Producer cells were cotransfected with 0.5, 1, 1.5, and 2 µg of ERV-K envelope constructs and 1 µg of pSARM4-EGFP. The producer cell lysates were harvested 48 hours after transfection and checked for Env and MPMV Capsid (CA) expression. There is Env expression and processing of both HERV-Kcon Env and rhERV-K Env. There is also MPMV CA expression in the producer cell lysates indicating that the producer cells were making all proteins required for the production of ERV-K pseudotyped virions.*

If the Envs were incorporated into virion particles, then Env-pseudotyped virions would be released from the cell into the supernatant. To test for incorporation, the virions in the supernatant were pelleted and analyzed by western blot to check for ERV-K Env and MPMV CA protein release (**Figure 1Figure 9**). An Env-only control was used to determine the amount of background release of Env into the

supernatant without an MPMV core. A western blot probed for expression of HERV-K Env showed full-length and processed HERV-Kcon and rhERV-K Envs in the supernatant indicating Env release. A western blot probed for MPMV CA revealed CA release into the supernatant, indicating that both ERV-K Envs were incorporated into MPMV virions that were released into the supernatant.

It's important to note high levels of full-length Env in the supernatant, which indicates full-length Env incorporation into MPMV virions. It is possible incorporation of full-length uncleaved Env into virions may interfere with infectivity (**Figure 9**). Furthermore, a 1:1 ratio of *env* to MPMV *gag-pol* is optimal for incorporation to occur, as evidenced by the highest levels of both Env and MPMV CA in the viral supernatant (lane indicated by the [*] asterisk).

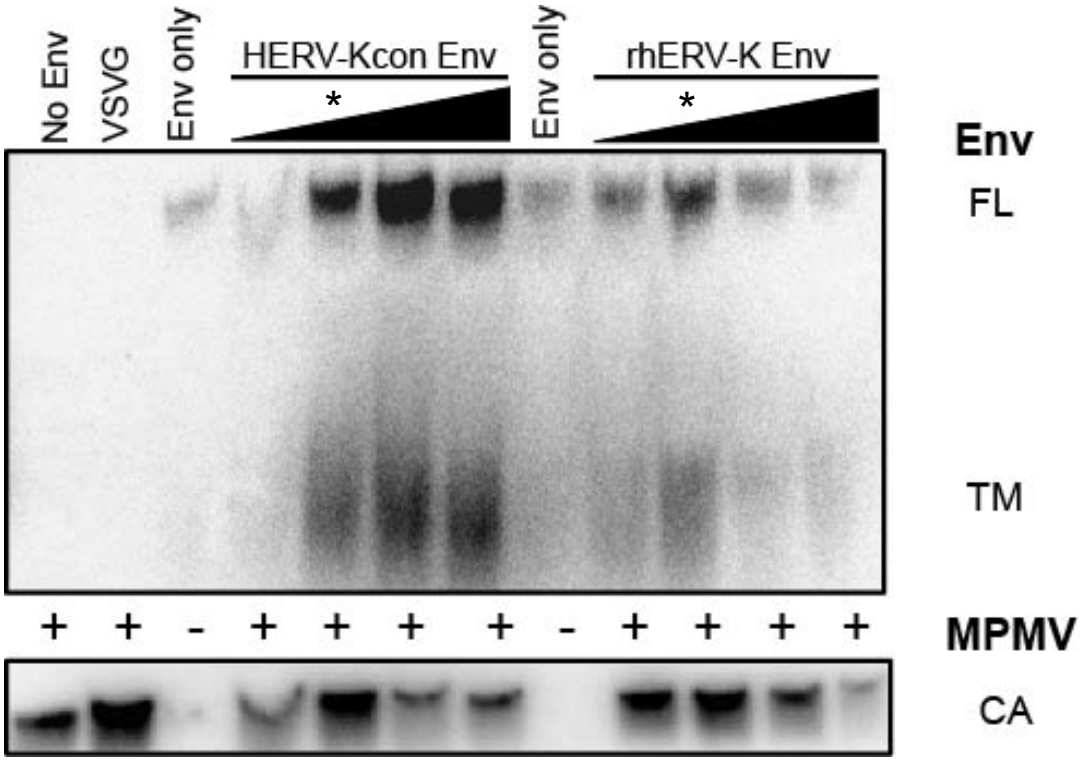

Figure 9. Western blot of Virions in the Supernatant. *Producer cells were cotransfected with 0.5, 1, 1.5, and 2 µg of ERV-K envelope constructs and 1 µg of pSARM4-EGFP. The supernatant was harvested and checked for Env and MPMV Capsid (CA) release. There is full-length and processed Env release for both HERV-Kcon Env and rhERV-k Env, as well as CA release in the supernatant indicating ERV-K Envs were incorporated into MPMV and were released from the cell. The Env only control serves as the background for Env release without MPMV CA.*

3.4 Infectivity of ERV-K Pseudotyped Virions

2 mls of the supernatants from the producer cells containing the pseudotyped virions were added to HEK293T/17, A549, and CRFK cells (**Figure 10**). At 48 hours after infection, the cells were harvested for FACS analysis.

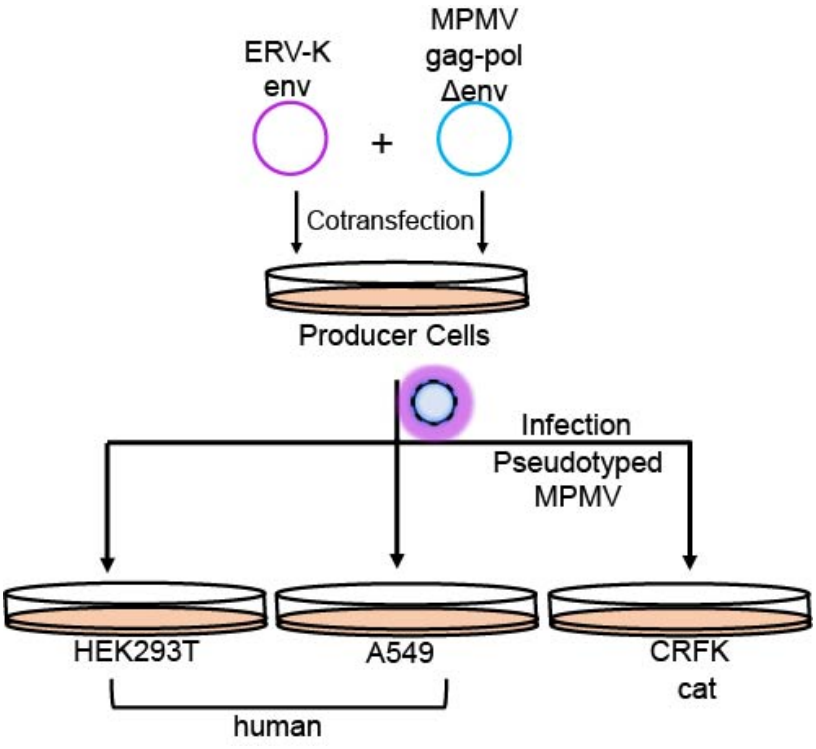

Figure 10. Infectivity Experimental Design. *ERV-K env plasmids were co-transfected with MPMV gag-pol at 1:1 ratio (1µg each). The supernatant was harvested 48 hours after transfection and added to two human cell lines, HEK293T and A549, and one feline cell line, CRFK. At 5 hrs after infection, 500 ul of warm D10 was added. At 48 hrs after infection, the cells were harvested for FACS analysis.*

ERV-K pseudotyped virions were able to mediate entry into CRFK cells, but not HEK293T/17 cells and A549 cells (**Figure 11**). The positive control VSVG-MPMV mediated entry into 80% of CRFK cells. FACS analysis revealed that the highest levels of infectivity occurred with the 1:1 ratio of transfected *env* to MPMV *gag-pol.* At this optimum ratio, HERV-Kcon MPMV infected 0.22% of CRFK cells, whilst rhERV-K Env MPMV infected 0.032% of CRFK cells. However, rhERV-K MPMV infectivity, while suggesting fusion and entry into the cell, is inconclusive. There was no entry of the ERV-K pseudotyped virions in the human cell lines,

HEK293T/17 and A549. The VSVG positive control infected HEK293T/17 and A549 cells at a low rate of 0.42% and 33.6%, respectively. Therefore, no conclusions can be drawn from those cell types.

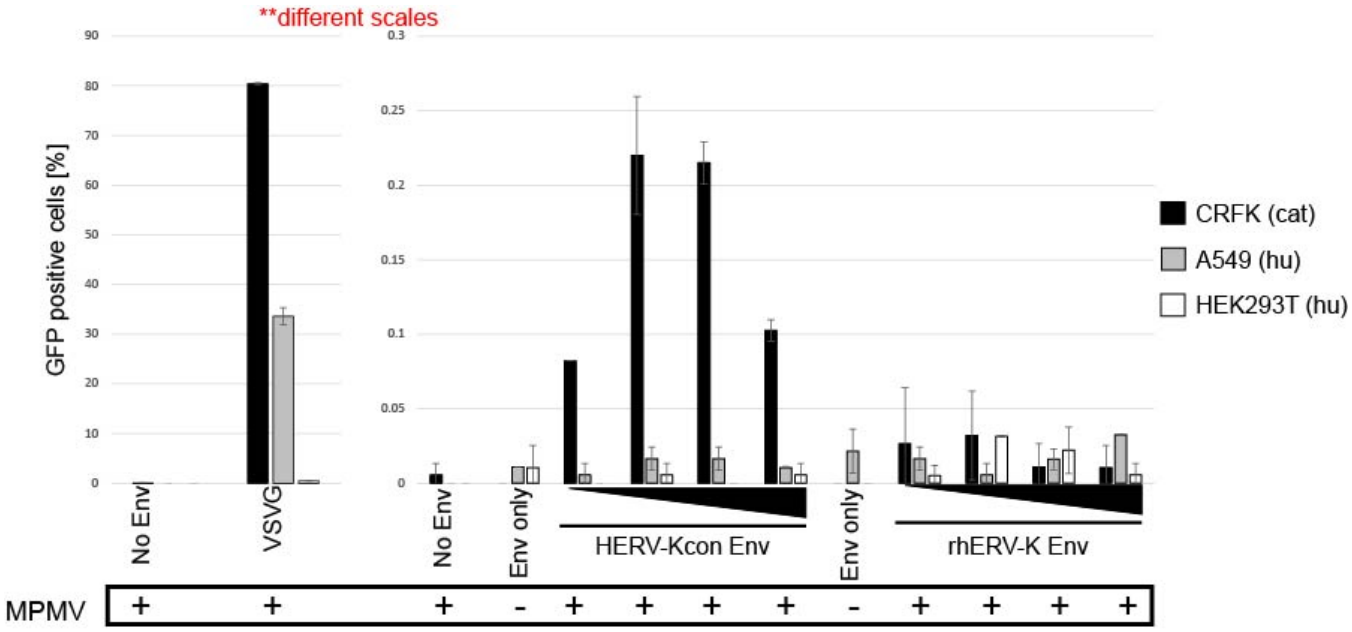

Figure 11. FACS Analysis of ERV-K Pseudotyped MPMV Virions on different cell types. *Both HERV-Kcon and rhERV-K Env pseudotyped MPMV virions were able to infect CRFK cells, but were unable to infect A549 and HEK293T cells. The positive control was VSVG pseudotyped MPMV, which was able to infect CRFK cells at a high level, A549 at a mediocre level, and unable to infect HEK293T indicating a technical issue. hu – human cell line.*

3.5 Increasing Infectivity of ERV-K Pseudotyped MPMV virions

In order to develop an infectious system, I sought to increase infectivity in CRFK cells. First, the optimum time to harvest virions after transfection of the plasmids was determined. 2 ml of supernatants were harvested 24, 36, and 48 hours after transfection and added to CRFK cells. At 48 hours after infection, the cells were harvested for FACS analysis. The data revealed that harvesting supernatant at

later time points increased ERV-K pseudotyped MPMV virion infectivity in CRFK cells (**Figure 12**). Harvesting virions at 48 hours after transfection was optimal and it increased the infectivity two-fold from previous experiments. HERV-Kcon MPMV infected up to 0.47% of CRFK cells, while rhERV-K MPMV infected up to 0.13% of CRFK cells.

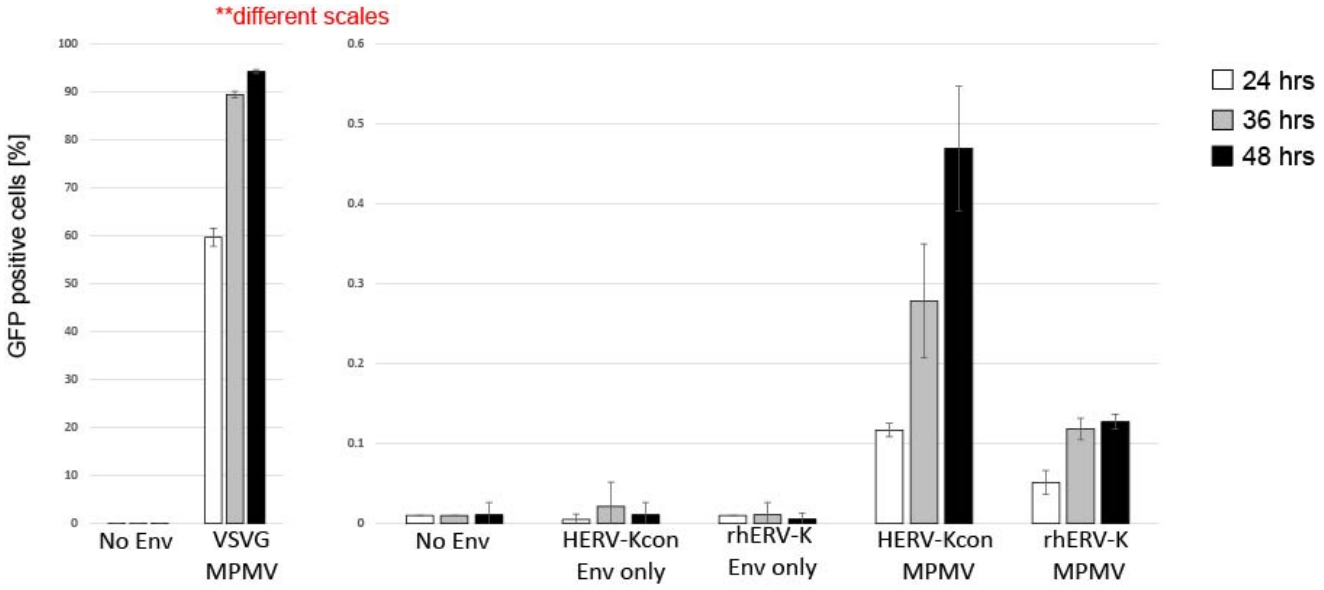

Figure 12. FACS Analysis of Infectivity of ERV-K Pseudotyped MPMV virions harvested at 24, 36, and 48 hrs after transfection. *Pseudotyped virions harvested 48 hours after transfection were able to infect CRFK cells at the highest level. No Env was used as the Gag-Pol-only negative control. The Envs only were also used as a background control.*

To further increase infectivity of pseudotyped virions, concentrating columns were used to concentrate virions. 7.5 x 10^6 HEK293T/17 cells were seeded and transfected with the plasmids. 14 ml of virion supernatant were collected and concentrated to 4 ml of virion supernatant using YM-50 concentrating columns. Both unconcentrated and concentrated virions were used to infect CRFK cells and

48 hours after infection, the cells were harvested for FACS analysis (**Figure 13**). Since VSVG-MPMV virions previously demonstrated a 94% infectivity level, VSVG-MPMV virions were diluted 15-fold and then concentrated. Unconcentrated VSVG-MPMV virions infected 13.3% of CRFK cells, while concentrated VSVG-MPMV virions infected 27.3% of CRFK cells. Concentrating ERV-K pseudotyped MPMV virions slightly increased infectivity. Unconcentrated HERV-Kcon MPMV infected 2.8% of CRFK cells, while concentrated virions infected 3.8% of CRFK cells. Unconcentrated rhERV-K MPMV infected 1.4% of CRFK cells, while concentrated virions infected 1.7% of CRFK cells. The data reveals that transfecting a higher amount of cells increases infectivity as shown by unconcentrated virions, and that concentrating the virions using the concentration columns increased infectivity slightly.

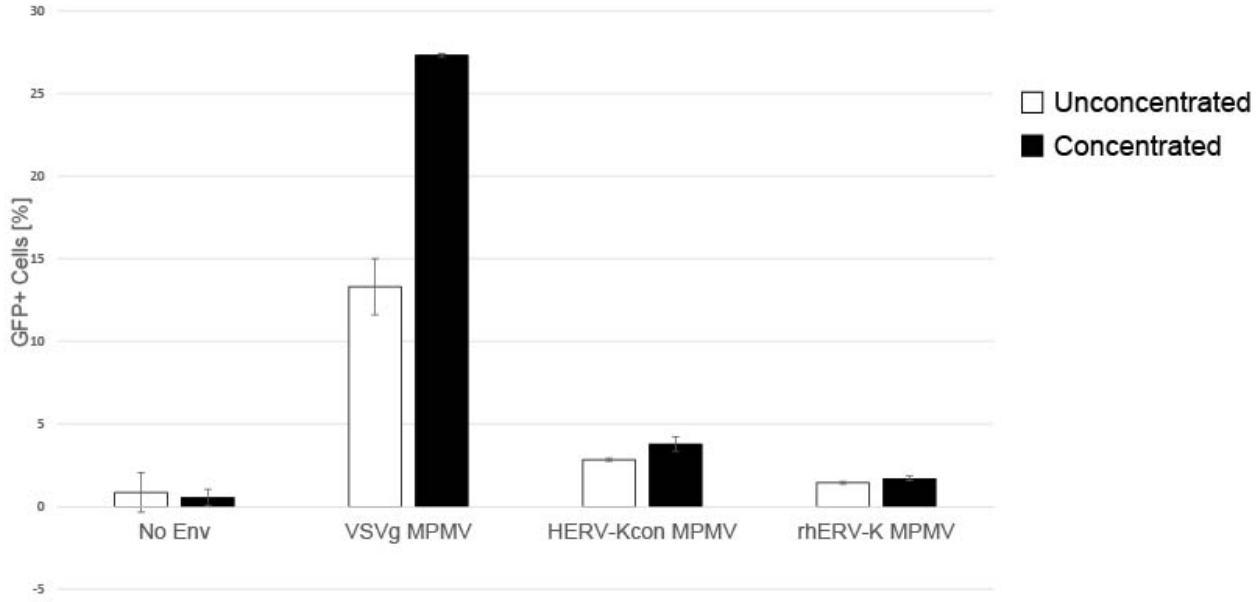

Figure 13. Concentrating ERV-K Pseudotyped MPMV Virions slightly increased infectivity. *No Env was the Gag-Pol-only negative control. VSVG MPMV was used as the positive control and was diluted 1:15 before concentrating.*

3.6 Addition of exogenous furin

According to **Figure 9**, there is a significant amount of full-length envelope in released virions. Full-length envelope could potentially be incorporated into MPMV virions interfering with infectivity. Only processed envelope can mediate fusion and entry into the cell. To increase processing of envelope, a furin construct was cotransfected with the envelope constructs. At 48 hours after transfection, cells were lysed. A western blot against HERV-K Env revealed that addition of exogenous furin did not increase processing of full-length envelope (**Figure 14**). Interestingly it did change the banding pattern of full-length envelope revealing proteins at the 103 kDa and 79 kDa mark as compared to the 103 kDa protein in the full-length envelope ERV-K Env without addition of furin, most likely indicating different glycosylation patterns.

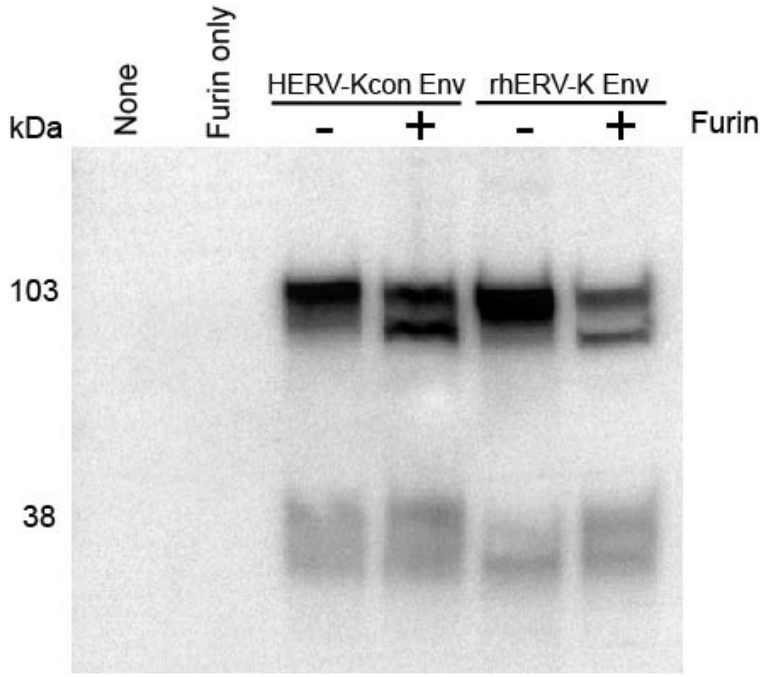

Figure 14. Addition of Exogenous Furin to the ERV-K Envs. *Co-transfection of ERV-K env plasmids and furin expression plasmids did not yield significantly higher envelope processing.*

3.7 Pseudotyping MLV Virions

Previous reports have demonstrated HERV-Kcon Env incorporation with MLV cores. I sought to pseudotype MLV virions to develop an ERV-K infectious system based on MLV. HEK293T/17 cells were co-transfected with varying amounts of envelope constructs ranging from 0.3 µg to 2 µg, varying amounts of MLV *gag-pol Δenv* ranging from 0.3 µg to 2 µg, and varying amounts of packagable GFP construct ranging from 0.3 µg to 2 µg. The plasmids were transfected at varying ratios of *env: gag-pol: gfp.* At 48 hours after transfection, 2 mL of virion supernatant were harvested and placed onto HEK293T/17 and CRFK cells. At 48 hours after infection, cells were imaged for GFP+ cells. While VSVG-MLV virions were able to

infect both HEK293T/17 cells and CRFK cells, ERV-K pseudotyped MLV virions were unable to infect either cell type.

CHAPTER 4: DISCUSSION

I have developed a single-cycle ERV-K infectious system based on MPMV. While HERV-K-based single-cycle systems based on lentiviruses and gammaretroviruses exist, this is the first report of an HERV-K infectious system based on MPMV, a betaretrovirus. I also discovered an ERV-K Env homolog in the rhesus macaque genome, referred to as rhERV-K Env. rhERV-K Env was expressed, processed and found capable of incorporating into MPMV virions. This is the first report of functional characterization of a rhERV-K Env.

In order to create an ERV-K infectious system, ERV-K Envs need to be functional. I have demonstrated that both ERV-K Envs are expressed and processed (**Figure 6**). This is the first report showing rhERV-K Env expression *in vitro*. To detect the expression of the envelopes, an antibody against HERV-K Env TM was used. This antibody cross reacted with the rhERV-K Env. This is of interest because the antibody recognizes and binds to the TM of HERV-K Env and rhERV-K Env. However, the primary amino acid sequences of the TM regions of the two Envs differ quite a bit (APPENDIX B) so the recognition of the antibody must be at the conserved sites of the HERV-Kcon TM and rhERV-K Env TM.

The sizes of the full-length and processed envelopes were as predicted, 103 kDa and 38 kDa respectively. The smear of the processed envelope products in the western blot in **Figure 6** is likely due to post-translational glycosylation altering the sizes. Notably, the envelopes were not fully processed as indicated by the presence of the full-length envelope in the western blot. Co-transfection of the

envelope constructs with exogenous furin constructs did not increase processing of full-length envelope, but instead slightly changed the banding pattern of the western blot showing a full length protein at size 79 kDa, which is the size of the full length protein with no glycosylation. This indicates that co-expression with exogenous furin possibly altered the glycosylation patterns of the full-length envelope (**Figure 14**).

Expression of the envelope could only occur from a pCRV1 backbone and did not occur from a pcDNA3.1(+) backbone. pCRV1 encodes for HIV-1 Rev and contains a RRE to which Rev can bind, mediating transport of unspliced RNA out of the nucleus. RT-PCR performed on RNA extracted from HEK293T cells transfected with pcDNA3.1(+) ERV-K showed that ERV-K transcripts were being made (data not shown), but the lack of expression detected via western blot indicates that the problem was probably a failure of ERV-K ENV mRNA to export out of the nucleus.

Incorporation of ERV-K Env into MPMV virions was successful and pseudotyped virions mediated fusion and entry into CRFK cells, indicating that CRFK cells express the receptor required for HERV-Kcon and rhERV-K Env recognition (**Figure 8**, **Figure 9**, **Figure 11**). This is the first report demonstrating that rhERV-K Env is functional, can pseudotype MPMV virions, and can mediate fusion and entry into a cell. However, infectivity of the ERV-K pseudotyped MPMV virions in CRFK cells is low. This could be due to full-length envelope incorporation into virions. As shown in **Figure 9**, there was a significant amount of full-length envelope being

incorporated into MPMV virions. Full-length envelope is unable to mediate fusion and entry into the cell. Only processed envelope is able to mediate fusion and entry into the cell. Unfortunately, co-transfecting furin expression plasmids with the ERV-K *envs* did not increase envelope processing (**Figure 14**).

In order to increase infectivity of the ERV-K pseudotyped MPMV virions, the virions were concentrated using YM-50 concentrating columns. The samples were centrifuged in the concentrating columns twice, concentrating the virions from a 14 ml volume to a 4 ml volume. This yielded slightly higher infectivity, up to ~4% for HERV-Kcon pseudotyped MPMV virions, and up to ~2% for rhERV-K pseudotyped MPMV virions (**Figure 13**). It is also notable that scaling up the amount of virus-producing cells from a 6-well to a T75 flask increased infectivity, as shown in **Figure 13**. Scaling up the amount of virus-producing cells increased the amount of virions in the supernatant, which in turn would increase infectivity on CRFK cells.

Future experiments should be aimed to further increase the infectivity levels of the ERV-K pseudotyped MPMV virions to characterize ERV-K entry and to identify the receptor for the ERV-K Envs. One of the simplest approaches would be to increase the amount of virions in the supernatant to concentrate. This study transfected 7.5 x 10^6 HEK293T/17 cells to make pseudotyped MPMV virions. The virions in the supernatant from the transfection were concentrated to a smaller volume. If double the amount of HEK293T/17 cells were transfected, more virions would be created. By concentrating a larger number of virions into a smaller volume, there would be

more virions to recognize the receptor and mediate fusion into the cell, potentially increasing infectivity levels.

Another simple method to improve infectivity would be to try infecting other cell types and lines. In the interest of time, this study only focused on CRFK cells. However, there are various cell lines that could potentially express the receptor on the surface of the cell in higher quantities. Because of increased receptor expression, the ERV-K pseudotyped virions would be able to recognize the receptor and mediate fusion and entry more easily than in a cell line with sparse cell receptor expression. CRFK cells are feline cells. While my data indicates that CRFK cells express the ERV-K receptor on the cell surface, it is highly likely that rhERV-K Env is better adapted to use its orthologous receptor on rhesus macaque cells. This study did not use rhesus macaque cell lines for infectivity. It is probable that rhesus macaque cell lines express the receptor for rhERV-K Env and it is important to test the infectivity of rhERV-K pseudotyped virions on rhesus-origin cells. However, it does remain possible that rhERV-K *env* have been coopted for viral resistance. Even though rhesus macaque cells lines express the receptor for rhERV-K Env, endogenous rhERV-K *env* could be downregulating expression of the receptor inhibiting exogenous infection.

Unfortunately, pseudotyping MLV with the ERV-K Envs was unsuccessful. In parallel to creating an ERV-K based system in betaretroviruses, I sought to create an ERV-K based system in gammaretroviruses since it was previously reported.

While the control, pseudotyping with VSVg was successful, there was no infectivity seen with ERV-K pseudotyped MLV. A possible explanation is that Serinc5, a cellular protein with antiviral effects, is reducing the infectivity of the pseudotyped MLV virions. A recent paper identified that co-transfection of virus-producing cells with a Serinc5 expression plasmid reduced infectivity and entry of MLV and this was determined by envelope meaning that different envelopes resulted in differential susceptibility to Serinc5[1]. MLV encodes a glycosylated form of gag, glycogag, which counteracts the effects of Serinc5. It was demonstrated that a small amount of Glycogag was able to counteract endogenous Serinc5 in HEK293T cells[1]. Thus, for this study, Serinc5 in the virus-producing HEK293T cells could have potentially been reducing the infectivity of pseudotyped MLV virions. This effect was not seen with VSVg pseudotyped MLV virions, which could explain why the control worked well, but the ERV-K pseudotyped virions did not. To test if this is the case, virus-producing plasmids need to be co-transfected with a MLV glycogag expression plasmid. Since it was previously demonstrated that even a small amount of Glycogag can counteract endogenous Serinc5 in HEK293T cells, it is predicted that this experiment would result in increased infectivity of ERV-K pseudotyped MLV virions.

Once the ERV-K infectious system is optimized to high infectivity levels, ideally around 80-100% (but lower infectivity levels will work), it can be utilized to find a receptor. In order to do that, the cell tropism of both envelopes must be determined. It stands to reason that some cell lines will be refractory to infection by ERV-K-

MPMV, indicating a lack of receptor expression. Once such a null cell line and at least one permissive cell line are identified, two main strategies can be pursued to find the receptor. The first is through a cDNA library screen, which has been previously successfully used to identify receptors for other viral envelopes[8,43]. Once null cell lines and permissive cell lines are established, a packagable cDNA library from the permissive cell line will be created. The permissive cell line will contain the transcript for the receptor. The permissive cell line cDNA library will be transduced into the null cell line and ERV-K pseudotyped GFP-expressing virions will be used to infect the transduced null cell line. The cells will be screened for permissivity by means of GFP expression. GFP+ cells will be isolated by flow cytometry and expanded. Genomic DNA will be extracted from those cells any transduced genes will be identified by PCR. Once potential target genes are identified, gain-of-function studies and loss-of-function studies will identify the receptor of ERV-K.

The second method to identify the cell surface receptor is a proteomics approach similar to the one used to find the SARS Coronavirus receptor[28]. Briefly, soluble ERV-K envelope proteins fused to a tag will be incubated with the cell extracts of permissive cell lines. The mixture will be put through a column containing beads designed to bind to the tag. After multiple washes, elution of the soluble envelope and the bound proteins will occur. This eluate will then be run on a SDS-Page gel and the bands will be cut out and analyzed via mass spectrometry.

This study is the initial step in figuring out how ERV-K retroviruses accessed the human germline some 150,000 – 250,000 years ago. Since there is no exogenous replication competent HERV-K, an infectious system is required to characterize the envelope proteins and identify the receptor. This was the first study demonstrating that ERV-K envelope proteins can be incorporated into MPMV virions. Furthermore, this is the first study showing that rhERV-K Env is functional and can mediate fusion and entry into CRFK cells. This infectious system, once optimized, can be used to find a receptor for this once exogenous retrovirus. Using this infectious system, we can determine how this virus accessed the germline cell in the first place and learn more about the early stages of endogenization.

APPENDIX A

Leader Sequence
Furin Cleavage Site
Fusion domain
Immunosuppressive Domain
Conserved Cysteine Residues
Transmembrane Anchor Domain

rhERV-K *env* sequence (AC200900.3).

ATG AAC CCA TCG GAG ATG CAA AGA AAA GCG CCT CCA CGG AGA CAG AAA
M N P S E M Q R K A P P R R Q K
CAC CGC AAT CGA GCA CCA TTG ACT CGC ATG ATG AAC CAA GTG ATG ATA
H R N R A P L T R M M N Q V M I
TCA GAA GAA CAG ATG AAG TCA CCA CGC ACC AAG AAG GCG GAG CTG CCG
S E E Q M K S P R T K K A E L P
ACC TGG GCA CAG TTA AAG AAG CTG ACA CCG TTA GCT GGA AAA AGC CTA
T W A Q L K K L T P L A G K S L
GCT AGC ACA AAG GTG ACA CAA ACC CCA GAA AAA ATG CTG CTT ACA GCT
A S T K V T Q T P E K M L L T A
TTA ATG ATT GTA TCA ACG GTG GTA AGT CTC CCC ATG CCT GCA GGA GCA
L M I V S T V V S L P M P A G A
GCT GCA GCT AAT TAT ACC TAC TGG GCC TAT GTG CCT TTC CCG CCC TTA
A A A N Y T Y W A Y V P F P P L
ATT CGG GCA GTT ACA TGG ATG GAT AAT CCT ATT GAA GTA TAT GTT AAT AAT
I R A V T W M D N P I E V Y V N N
AGT GTG TGG GTA CCT GGT CCC ACA GAT GAT CGT TGC CCT GCC AAA CCG
S V W V P G P T D D R C P A K P
GAG GAA GAA GGA ATG ATG ATA AAT ATT TCC ATT GGG TAT CGT TAT CCT
E E E G M M I N I S I G Y R Y P
CCT ATT TGC CTA GGG AGA GCA CCA GGA TGT TTA ATG CCT GCT ATT CAA
P I C L G R A P G C L M P A I Q
AAT TGG TTG GTA GAA GTA CCT ACT GTC AGT CCC ACC AGT AGA TTT ACT
N W L V E V P T V S P T S R F T
TAT CAC ATG GTA AGC GGA ATG TCA CTC AAA CCA CAG GTA AAC TAT TTA
Y H M V S G M S L K P Q V N Y L
CAA GAC TTT TCT TAT CAA AGA TCA TTA AAA TTT AGG CCA AAA GGG AAA CCT

Q D F S Y Q R S L K F R P K G K P

TGC CCC AAA GAG ATT TCC AGA GAA TCG AAA GAT TTA GTT TGG GAA GAA

C P K E I S R E S K D L V W E E

TGT GTG GCC GAT AGT GCA GTG ATA TTA CAA AAC AAT ACA TTC GGA ACA

C V A D S A V I L Q N N T F G T

GTT ATA GAT TGG GCA CCT AGA GGT CAA TTC TAC CAC AAT TGC ACA GGA

V I D W A P R G Q F Y H N C T G

CAA ACT CAA TTC TGT CCC AGT GCA CTA GTG AGT CCA ACT GTT GAC AGT

Q T Q F C P S A L V S P T V D S

GAT TTA ACG GAA AAT TTA GAT AAA CAT AAG CAC AAA AAA TTA CAG TCT TTC

D L T E N L D K H K H K K L Q S F

TAC CCT TGG ATA TGG GGA GAA AAG GGA ATG TCT ACT CCA AGA CCA AAA

Y P W I W G E K G M S T P R P K

ATG ATA AGT CCT GTT TTT GGT CCT GAA CAT CCA GAA TTA TGG AGA CTT

M I S P V F G P E H P E L W R L

ACT GTG GCT TCA TAC CGC CTT AGA ATT TGG TCT GGA AAT CAA ACT ATA

T V A S Y R L R I W S G N Q T I

GAA ACA AGA GAT TAT AAG CCA TTT TAC TCT ATC AAC CTA AAT TCC AGT CTA

E T R D Y K P F Y S I N L N S S L

ACA GTT CCT TTA CAA AGT TGT GTA AAG CCC CCT TAT ATG TTA GTC ATA

T V P L Q S C V K P P Y M L V I

GGA AAT ATA GTT ATT AAA CCA GAC TCC CAA ACT ATA ACT TGT GAA AAT TGC

G N I V I K P D S Q T I T C E N C

AGA TTG TTT ACT TGC ATT GAT TCA ACT TTT GAT TGG CAG CAC CGT ATT CTA

R L F T C I D S T F D W Q H R I L

CTG GTG AGA GCA AGA GAA GGC GTG TGG ATC CCT GTG TCC ATG GAC CGA

L V R A R E G V W I P V S M D R

CCA TGG GAG GCC TCA CCA TCC ATC CAT ATT TTG ACT GAA GTA TTA AAA

P W E A S P S I H I L T E V L K

GGC GTT TTA AGT AGA TCC AAA AGA TTC ATT TTT ACT TTA ATT GCA GTG ATT

G V L S R S K R F I F T L I A V I

ATG GGA TTA ATT GCA GTC ACA GCT ACA GCC TCT GTG GCA GGA GTT GCA

M G L I A V T A T A S V A G V A

TTG CAC TCT TCT GTT CAG ACA GTA AGC TTT GTT GAC AAT TGG CAA AAG

L H S S V Q T V S F V D N W Q K

AAT TCC ACA AGG TTG TGG AAT TCA CAA TCT GGT ATC GAT CAA AAA TTG

N S T R L W N S Q S G I D Q K L

GCA AAT CAA ATT AAT GAT CTT AGA CAA ACC GTC ATT TGG ATG GGA GAT

A N Q I N D L R Q T V I W M G D

AGA CTC ATG AGC TTG GAA CAT CGT TTC CAG TTA CAG TGT GAC TGG AAT

R L M S L E H R F Q L Q C D W N

ACG TCA GAT TTT TGT ATT ACA CCC CAA GTT TAT AAT GAG TCT AAA CAT CAC

T S D F C I T P Q V Y N E S K H H

TGG GAC ATG GTT AGA CGC CAT CTA CAG GGA AGA GAA GAT AAT CTC ACT

W D M V R R H L Q G R E D N L T

TTA GAC ATT TCT AAA TTA AAA GAA CAA ATT TTT GAA GCC TCT CAA AGC CAC

L D I S K L K E Q I F E A S Q S H

TTA AAT ATT GTG CCT GGA GCT GAG GCA TTA GAT CAA GTG GCA AAA AAT

L N I V P G A E A L D Q V A K N

CTT TAT GAA TTA AAC CCC ACG ACT TGG ATT AAG TCT ATT GGA AAC **TCT ACT**

L Y E L N P T T W I K S I G N S T

GCA ATA AAT TTT GGA ATT ATG TGT CTC TGT TTA ATC AGC TTG TTT TTA GTG

A I N F G I M C L C L I S L F L V

TGC TGG ACC AGT CGA AGA ATC CTG CGT CAA AAT CGA GAG AAC GAA CAA

C W T S R R I L R Q N R E N E Q

GCC TTC ATC GCC ATG GCA CAT TTA TAT AGA GGA AAA GGG AGG GAG AAC

A F I A M A H L Y R G K G R E N

GTT GCG GGA AGT CAG GGA CCT TGA

V A G S Q G P *

rhERV-K codon-optimized *env* sequence.

ATGAATCCCAGCGAAATGCAACGGAAAGCTCCCCCTAGGCGACAGAAGCAC
AGGAACCGCGCCCCTCTGACCAGAATGATGAACCAGGTGATGATCAGCGA
GGAACAGATGAAGTCACCACGCACAAAAAAAGCGGAACTGCCCACATGGG
CCCAGCTGAAGAAGTTGACTCCACTTGCAGGTAAATCCCTGGCATCCACCA
AGGTGACTCAGACTCCCGAAAAGATGCTTCTGACTGCTCTTATGATCGTGTC
CACCGTGGTGTCCCTGCCTATGCCTGCCGGCGCCGCCGCCGCAAACTATA
CCTACTGGGCCTACGTCCCCTTTCCCCCTCTCATCAGGGCAGTGACTTGGA
TGGACAATCCCATCGAGGTGTATGTCAATAATAGCGTGTGGGTACCCGGAC
CTACGGACGACAGGTGTCCAGCAAAGCCTGAGGAGGAAGGCATGATGATT
AATATAAGTATCGGATACCGGTACCCTCCAATTTGTCTCGGCCGAGCGCCC
GGCTGCCTTATGCCTGCAATACAGAATTGGCTGGTGGAAGTCCCTACCGTG
TCACCCACATCCCGCTTCACTTACCACATGGTGTCAGGAATGAGCCTGAAA
CCGCAGGTGAACTACCTGCAAGATTTCTCCTACCAAAGGTCTTTGAAATTTC
GGCCTAAAGGCAAGCCATGTCCCAAAGAGATCTCCAGAGAGTCCAAGGATC
TGGTATGGGAAGAGTGTGTCGCAGACAGCGCTGTTATCCTGCAGAACAATA
CCTTCGGAACCGTGATCGACTGGGCTCCTAGAGGACAGTTTTACCATAATT
GCACCGGTCAAACCCAGTTCTGTCCTAGCGCTCTTGTGTCCCCGACAGTAG
ATAGTGACCTTACCGAAAACTTGGACAAGCACAAACACAAGAAGCTGCAGT
CATTCTACCCGTGGATCTGGGGCGAAAAGGGTATGTCCACTCCCCGCCCGA
AGATGATTTCTCCTGTCTTCGGTCCAGAGCACCCAGAACTGTGGAGACTGA
CTGTGGCCAGTTATAGGCTGAGAATCTGGTCTGGAAACCAGACTATCGAGA
CCCGAGATTACAAACCTTTTTACAGCATCAACCTCAACTCATCCCTTACAGT
CCCACTTCAGTCATGTGTGAAGCCCCCATATATGCTCGTAATTGGCAATATC
GTCATTAAACCCGATTCTCAAACAATAACCTGCGAGAACTGCCGCTTGTTTA
CCTGCATTGACAGTACCTTCGATTGGCAGCATCGAATACTGCTCGTGAGGG
CCCGGGAGGGTGTGTGGATCCCAGTCAGCATGGACAGGCCCTGGGAAGCC
TCTCCATCTATACATATTCTCACAGAGGTGCTCAAGGGAGTCCTGAGCAGAT
CTAAGCGCTTCATCTTCACTCTGATTGCTGTGATCATGGGATTGATCGCAGT
GACAGCCACTGCTAGTGTGGCCGGGGTTGCGCTCCACAGTAGCGTCCAAA

CAGTGTCCTTCGTGGACAATTGGCAGAAGAACTCCACACGGCTGTGGAACT
CTCAGAGTGGAATCGACCAAAAACTGGCGAATCAAATAAACGATCTGAGAC
AGACAGTTATCTGGATGGGTGACCGGCTGATGTCCCTCGAACATAGATTCC
AACTGCAGTGCGACTGGAACACAAGCGACTTCTGCATCACTCCTCAGGTTT
ATAACGAGAGCAAACACCATTGGGATATGGTGCGCCGCCATCTGCAGGGG
CGAGAAGATAACCTGACCCTGGACATTAGCAAGCTGAAGGAACAAATCTTC
GAGGCCTCCCAGAGTCACCTTAATATTGTGCCTGGAGCCGAGGCCCTCGAT
CAGGTAGCAAAAAATCTTTACGAGCTCAATCCAACCACTTGGATTAAGTCTA
TAGGGAACTCTACGGCCATCAATTTTGGAATAATGTGTCTTTGCCTCATCTC
ACTCTTCCTGGTCTGTTGGACATCCAGAAGAATTCTTCGGCAGAACCGCGA
GAACGAGCAGGCATTCATCGCCATGGCACACCTCTATCGCGGTAAAGGCAG
AGAGAACGTGGCTGGGTCCCAGGGCGCCGGCGCCGGCCTGAACGACATC
TTCGAGGCCCAGAAGATCGAGTGGCACGAGTAG

APPENDIX B

```
>Protein alignment 6 Alignment of 2 sequences: HERV-Kcon Env , rhERV-K Env

Score = 3092.0, Identities = 575/700 (82%),
Positives = 631/700 (90%), Gaps = 13/700 (1%)

HERV-Kcon Env      1 MNPSEMQRKAPPRRRRHRNRAPLTHKMNKMVTSEEQMKLPSTKKAEPPTWAQLKKLTQLA  60
                     MNPSEMQRKAPPRR++HRNRAPLT  MN+++ SEEQMK P TKKAE PTWAQLKKLT LA
rhERV-K Env        1 MNPSEMQRKAPPRRQKHRNRAPLTRMMNQVMISEEQMKSPRTKKAELPTWAQLKKLTPLA  60

HERV-Kcon Env     61 TKYLENTKVTQTPESMLLAALMIVSMVVSLPMPAGAAAANYTYWAYVPFPPLIRAVTWMD 120
                      K L +TKVTQTPE MLL ALMIVS VVSLPMPAGAAAANYTYWAYVPFPPLIRAVTWMD
rhERV-K Env       61 GKSLASTKVTQTPEKMLLTALMIVSTVVSLPMPAGAAAANYTYWAYVPFPPLIRAVTWMD 120

HERV-Kcon Env    121 NPIEVYVNDSVWVPGPIDDRCPAKPEEEGMMINISIGYRYPPICLGRAPGCLMPAVQNWL 180
                     NPIEVYVN+SVWVPGP DDRCPAKPEEEGMMINISIGYRYPPICLGRAPGCLMPA+QNWL
rhERV-K Env      121 NPIEVYVNNSVWVPGPTDDRCPAKPEEEGMMINISIGYRYPPICLGRAPGCLMPAIQNWL 180

HERV-Kcon Env    181 VEVPTVSPISRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKPCPKEIPKESKNTEV 240
                     VEVPTVSP SRFTYHMVSGMSL+P+VNYLQDFSYQRSLKFRPKGKPCPKEI +ESK+
rhERV-K Env      181 VEVPTVSPTSRFTYHMVSGMSLKPQVNYLQDFSYQRSLKFRPKGKPCPKEISRESKD--- 237

HERV-Kcon Env    241 LVWEECVANSAVILQNNEFGTIIDWAPRGQFYHNCSGQTQSCPSAQVSPAVDSDLTESLD 300
                     LVWEECVA+SAVILQNN FGT+IDWAPRGQFYHNC+GQTQ CPSA VSP VDSDLTE+LD
rhERV-K Env      238 LVWEECVADSAVILQNNTFGTVIDWAPRGQFYHNCTGQTQFCPSALVSPTVDSDLTENLD 297

HERV-Kcon Env    301 KHKHKKLQSFYPWEWGEKGISTPRPKIVSPVSGPEHPELWRLTVASHHIRIWSGNQTLET 360
                     KHKHKKLQSFYPW WGEKG+STPRPK++SPV GPEHPELWRLTVAS+ +RIWSGNQT+ET
rhERV-K Env      298 KHKHKKLQSFYPWIWGEKGMSTPRPKMISPVFGPEHPELWRLTVASYRLRIWSGNQTIET 357

HERV-Kcon Env    361 RDRKPFYTVDLNSSLTVPLQSCVKPPYMLVVGNIVIKPDSQTITCENCRLLTCIDSTFNW 420
                     RD KPFY+++LNSSLTVPLQSCVKPPYMLV+GNIVIKPDSQTITCENCRL TCIDSTF+W
rhERV-K Env      358 RDYKPFYSINLNSSLTVPLQSCVKPPYMLVIGNIVIKPDSQTITCENCRLFTCIDSTFDW 417

HERV-Kcon Env    421 QHRILLVRAREGVWIPVSMDRPWEASPSVHILTEVLKGVLNRSKRFIFTLIAVIMGLIAV 480
                     QHRILLVRAREGVWIPVSMDRPWEASPS+HILTEVLKGVL+RSKRFIFTLIAVIMGLIAV
rhERV-K Env      418 QHRILLVRAREGVWIPVSMDRPWEASPSIHILTEVLKGVLSRSKRFIFTLIAVIMGLIAV 477

HERV-Kcon Env    481 TATAAVAGVALHSSVQSVNFVNDWQKNSTRLWNSQSSIDQKLANQINDLRQTVIWMGDRL 540
                     TATA+VAGVALHSSVQ+V+FV++WQKNSTRLWNSQS IDQKLANQINDLRQTVIWMGDRL
rhERV-K Env      478 TATASVAGVALHSSVQTVSFVDNWQKNSTRLWNSQSGIDQKLANQINDLRQTVIWMGDRL 537

HERV-Kcon Env    541 MSLEHRFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQIFEA 600
                     MSLEHRFQLQCDWNTSDFCITPQ+YNES+HHWDMVRRHLQGREDNLTLDISKLKEQIFEA
rhERV-K Env      538 MSLEHRFQLQCDWNTSDFCITPQVYNESKHHWDMVRRHLQGREDNLTLDISKLKEQIFEA 597

HERV-Kcon Env    601 SKAHLNLVPGTEAIAGVADGLANLNPVTWVKTIGSTTIINLILILVCLFCLLLVCRCTQQ 660
                     S++HLN+VPG EA+  VA  L  LNP TW+K+IG++T IN  ++ +CL  L LVC  +++
rhERV-K Env      598 SQSHLNIVPGAEALDQVAKNLYELNPTTWIKSIGNSTAINFGIMCLCLISLFLVCWTSRR 657

HERV-Kcon Env    661 LRRDSDHRERAMMTMAVLSKRKGG-NVGKSKRDQIVTVSV 699
                     + R +   E+A + MA L + KG  NV  S+
rhERV-K Env      658 ILRQNRENEQAFIAMAHLYRGKGRENVAGSQ--------- 688
```

REFERENCES

1. Ahi, Y. S., S. Zhang, Y. Thappeta, A. Denman, A. Feizpour, S. Gummuluru, B. Reinhard, D. Muriaux, M. J. Fivash, and A. Rein. Functional interplay between murine leukemia virus glycogag, serinc5, and surface glycoprotein governs virus entry, with opposite effects on gammaretroviral and ebolavirus glycoproteins. *MBio* 7:, 2016.

2. Aiewsakun, P., and A. Katzourakis. Endogenous viruses : Connecting recent and ancient viral evolution. *Virology* 479–480:26–37, 2015.

3. Armezzani, A., F. Arnaud, M. Caporale, G. di Meo, L. Iannuzzi, C. Murgia, and M. Palmarini. The signal peptide of a recently integrated endogenous sheep betaretrovirus envelope plays a major role in eluding gag-mediated late restriction. *J. Virol.* 85:7118–7128, 2011.

4. Armezzani, A., L. Murphy, T. E. Spencer, M. Palmarini, and F. Arnaud. The evolutionary interplay between exogenous and endogenous sheep betaretroviruses. In: Viruses: Essential Agents of Life. 2012, pp. 293–307.

5. Arnaud, F., M. Varela, T. E. Spencer, and M. Palmarini. Endogenous retroviruses: Coevolution of endogenous betaretroviruses of sheep and their host. , 2008.

6. Barbulescu, M., G. Turner, M. Su, R. Kim, M. I. Jensen-Seaman, A. S. Deinard, K. K. Kidd, and J. Lenz. A HERV-K provirus in chimpanzees, bonobos and gorillas, but not humans. *Curr. Biol.* 11:779–783, 2001.

7. Bhardwaj, N., F. Maldarelli, J. Mellors, and J. M. Coffin. HIV-1 infection leads to increased transcription of HERV-K (HML-2) proviruses in vivo but not to increased virion production. *J. Virol.* 88:11108–11120, 2014.

8. Blanco-Melo, D., R. J. Gifford, and P. D. Bieniasz. Co-option of an endogenous retrovirus envelope for host defense in hominid ancestors. *Elife* 6:, 2017.

9. Coffin, J. M., S. H. Hughes, and H. E. Varmus. Retroviruses. 1997.at <http://www.ncbi.nlm.nih.gov/pubmed/21433340>

10. Contreras-Galindo, R., M. H. Kaplan, P. Leissner, T. Verjat, I. Ferlenghi, F. Bagnoli, F. Giusti, M. H. Dosik, D. F. Hayes, S. D. Gitlin, and D. M. Markovitz. Human endogenous retrovirus K (HML-2) elements in the plasma of people with lymphoma and breast cancer. *J. Virol.* 82:9329–36, 2008.

11. Contreras-Galindo, R., P. López, R. Vélez, and Y. Yamamura. HIV-1

infection increases the expression of human endogenous retroviruses type K (HERV-K) in vitro. *AIDS Res. Hum. Retroviruses* 23:116–122, 2007.

12. Dewannieux, M., S. Blaise, and T. Heidmann. Identification of a functional envelope protein from the HERV-K family of human endogenous retroviruses. *J. Virol.* 79:15573–15577, 2005.

13. Dewannieux, M., F. Harper, A. Richaud, C. Letzelter, D. Ribet, G. Pierron, and T. Heidmann. Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. *Genome Res.* 16:1548–1556, 2006.

14. Dudley, J. P., J. A. Mertz, S. Bhadra, M. Palmarini, and C. A. Kozak. Endogenous retroviruses and cancer. In: Retroviruses and Insights into Cancer. 2011, pp. 119–162.doi:10.1007/978-0-387-09581-3_5

15. Dupressoir, A., G. Marceau, C. Vernochet, L. Benit, C. Kanellopoulos, V. Sapin, and T. Heidmann. Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. *Proc. Natl. Acad. Sci.* 102:725–730, 2005.

16. Garcia-Etxebarria, K., M. Sistiaga-Poveda, and B. M. Jugo. Endogenous retroviruses in domestic animals. *Curr. Genomics* 15:256–65, 2014.

17. Gifford, R. J., A. Katzourakis, M. Tristem, O. G. Pybus, M. Winters, and R. W. Shafer. A transitional endogenous lentivirus from the genome of a basal primate and implications for lentivirus evolution. *Proc Natl Acad Sci U S A* 105:20362–20367, 2008.

18. Han, G. Z., and M. Worobey. Endogenous lentiviral elements in the weasel family (mustelidae). *Mol. Biol. Evol.* 29:2905–2908, 2012.

19. Hanke, K., P. Kramer, S. Seeher, N. Beimforde, R. Kurth, and N. Bannert. Reconstitution of the Ancestral Glycoprotein of Human Endogenous Retrovirus K and Modulation of Its Functional Activity by Truncation of the Cytoplasmic Domain. *J. Virol.* 83:12790–12800, 2009.

20. Henzy, J. E., R. J. Gifford, C. P. Kenaley, and W. E. Johnson. An Intact Retroviral Gene Conserved in Spiny-Rayed Fishes for over 100 My. *Mol. Biol. Evol.* 34:msw262, 2016.

21. Henzy, J. E., W. E. Johnson, and J. E. Henzy. Pushing the endogenous envelope. , 2013.

22. Ito, J., T. Baba, J. Kawasaki, and K. Nishigaki. Ancestral Mutations Acquired in Refrex-1, a Restriction Factor against Feline Retroviruses, during its Cooption and Domestication. *J. Virol.* 90:1470–1485, 2016.

23. Johnson, W. E. Endogenous Retroviruses in the Genomics Era. *Annu. Rev. Virol.* 2:135–159, 2015.

24. Johnson, W. E. Endogenous Retroviruses in the Genomics Era. , 2015.doi:10.1146/annurev-virology-100114-054945

25. Keckesova, Z., L. M. J. Ylinen, G. J. Towers, R. J. Gifford, and A. Katzourakis. Identification of a RELIK orthologue in the European hare (Lepus europaeus) reveals a minimum age of 12 million years for the lagomorph lentiviruses. *Virology* 384:7–11, 2009.

26. Kramer, P., V. Lausch, A. Volkwein, K. Hanke, O. Hohn, and N. Bannert. The human endogenous retrovirus K(HML-2) has a broad envelope-mediated cellular tropism and is prone to inhibition at a post-entry, pre-integration step. *Virology* 487:121–128, 2016.

27. Kraus, B., K. Boller, A. Reuter, and B. S. Schnierle. Characterization of the human endogenous retrovirus K Gag protein: identification of protease cleavage sites. *Retrovirology* 8:21, 2011.

28. Kuhn, J. H., W. Li, H. Choe, and M. Farzan. Angiotensin-converting enzyme 2: A functional receptor for SARS coronavirus. , 2004.

29. Kuyl, A. C. Van Der. HIV infection and HERV expression : a review. *Retrovirology* 9:6, 2012.

30. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* 409:860–921, 2001.

31. Lavialle, C., G. Cornelis, A. Dupressoir, O. Heidmann, E. Supe, and I. G. Roussy. Paleovirology of “ syncytins ”, retroviral env genes exapted for a role in placentation. , 2013.

32. Lemaitre, C., F. Harper, G. Pierron, T. Heidmann, and M. Dewannieux. The HERV-K Human Endogenous Retrovirus Envelope Protein Antagonizes Tetherin Antiviral Activity. *J. Virol.* 88:13626–13637, 2014.

33. Mager, D. L., and J. P. Stoye. Mammalian Endogenous Retroviruses. 1–20.doi:10.1128/microbiolspec.MDNA3-0009-2014.Correspondence

34. Malfavon-Borja, R., and C. Feschotte. Fighting fire with fire: endogenous retrovirus envelopes as restriction factors. *J. Virol.* 89:4047–50, 2015.

35. Malicorne, S., C. Vernochet, G. Cornelis, B. Mulot, F. Delsuc, O. Heidmann, T. Heidmann, and A. Dupressoir. Genome-Wide Screening of Retroviral Envelope Genes in the Nine-Banded Armadillo (Dasypus novemcinctus, Xenarthra) Reveals an Unfixed Chimeric Endogenous Betaretrovirus Using the ASCT2 Receptor. *J. Virol.* 90:8132–8149, 2016.

36. Martínez Barrio, Á., M. Ekerljung, P. Jern, F. Benachenhou, G. O. Sperber, E. Bongcam-Rudloff, J. Blomberg, and G. Andersson. The first sequenced carnivore genome shows complex host-endogenous retrovirus relationships. *PLoS One* 6:, 2011.

37. Mayer, J., K. Tsangaras, F. Heeger, M. ??vila-Arcos, M. D. Stenglein, W. Chen, W. Sun, C. J. Mazzoni, N. Osterrieder, and A. D. Greenwood. A novel endogenous betaretrovirus group characterized from polar bears (Ursus maritimus) and giant pandas (Ailuropoda melanoleuca). *Virology* 443:1–10, 2013.

38. Pollard, V. W., and M. H. Malim. THE HIV-1 REV PROTEIN. *Annu. Rev. Microbiol.* 52:491–532, 1998.

39. Rakoff-Nahoum, S., P. J. Kuebler, J. J. Heymann, M. E Sheehy, G. M Ortiz, G. S Ogg, J. D. Barbour, J. Lenz, A. D. Steinfeld, and D. F. Nixon. Detection of T lymphocytes specific for human endogenous retrovirus K (HERV-K) in patients with seminoma. *AIDS Res. Hum. Retroviruses* 22:52–56, 2006.

40. Rhyu, D.-W., Y.-J. Kang, M.-S. Ock, J.-W. Eo, Y.-H. Choi, W.-J. Kim, S.-H. Leem, J.-M. Yi, H.-S. Kim, and H.-J. Cha. Expression of Human Endogenous Retrovirus env Genes in the Blood of Breast Cancer Patients. *Int. J. Mol. Sci.* 15:9173–83, 2014.

41. Robinson, L. R., and S. P. J. Whelan. Infectious entry pathway mediated by the human endogenous retrovirus K envelope protein. *J. Virol.* 90:JVI.03136-15, 2016.

42. Robinson, L. R., and S. P. J. Whelan. Infectious entry pathway mediated by the human endogenous retrovirus K envelope protein. *J. Virol.* JVI.03136-15, 2016.doi:10.1128/JVI.03136-15

43. Soll, S. J., S. J. D. Neil, and P. D. Bieniasz. Identification of a receptor for an extinct virus. *Proc. Natl. Acad. Sci. U. S. A.* 107:19496–19501, 2010.

44. Subramanian, R. P., J. H. Wildschutte, C. Russo, and J. M. Coffin. Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. *Retrovirology* 8:90, 2011.

45. Tarlinton, R. E., J. Meers, and P. R. Young. Retroviral invasion of the koala genome. *Nature* 442:79–81, 2006.

46. Wang-Johanning, F., a R. Frost, G. L. Johanning, M. B. Khazaeli, a F. LoBuglio, D. R. Shaw, and T. V Strong. Expression of human endogenous retrovirus k envelope transcripts in human breast cancer. *Clin. Cancer Res.*

7:1553–1560, 2001.

47. Wildschutte, J. H., Z. H. Williams, M. Montesion, R. P. Subramanian, J. M. Kidd, and J. M. Coffin. Discovery of unfixed endogenous retrovirus insertions in diverse human populations. *Proc. Natl. Acad. Sci.* 201602336, 2016.doi:10.1073/pnas.1602336113

48. Wu, H. L., E. J. Léon, L. T. Wallace, F. A. Nimiyongskul, M. B. Buechler, L. P. Newman, P. A. Castrovinci, R. Paul Johnson, R. J. Gifford, R. Brad Jones, and J. B. Sacha. Identification and spontaneous immune targeting of an endogenous retrovirus K envelope protein in the Indian rhesus macaque model of human disease. *Retrovirology* 13:6, 2016.

49. Young, N. L., and P. D. Bieniasz. Reconstitution of an infectious human endogenous retrovirus. *PLoS Pathog.* 3:0119–0130, 2007.

50. Zündorf, I. Fundamentals of Molecular Virology. Von Nicholas H. Acheson. *Pharm. Unserer Zeit* 36:479–479, 2007.